DEEP LEARNING THE EEG MANIFOLD

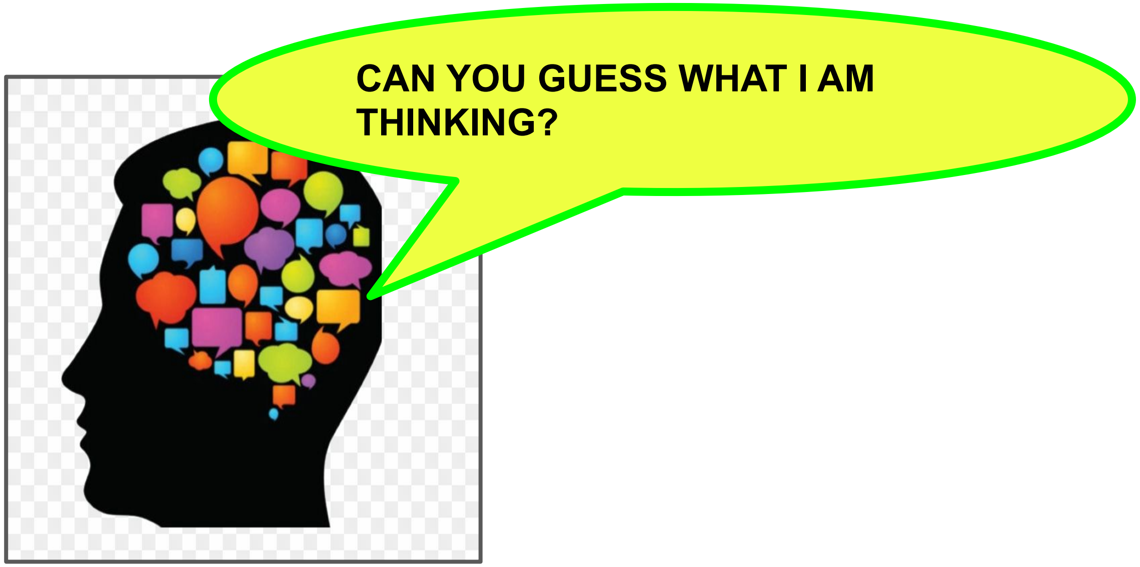
FOR PHONOLOGICAL CATEGORIZATION FROM ACTIVE THOUGHTS

Pramit Saha Sidney Fels Muhammad Abdul-Mageed

Human Communication Technologies Lab, Natural Language Processing Lab

University of British Columbia

Inferring Imagined Speech from EEG



- **Automatic Recognition of imagined speech** from **brain** (covert speech - thinking in form of sound, without muscle or articulatory movements).

- Provides vocal communication strategies for controlling devices through **speech commands interpreted from brain signals**, **Neuro-prosthetic help**, Equips people with '**voice to thoughts**'.

Subject-independent classification of phonological categories

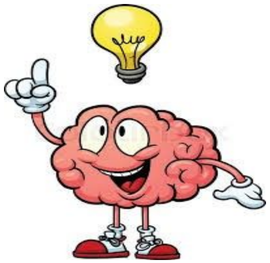
We specifically perform five binary classification tasks:

1. presence/absence of bilabial
2. presence/absence of nasal
3. presence/absence of consonant (closed vocal tract)
4. presence/absence of /uw/
5. presence/absence of /iy/

DATASET: Imagined speech EEG data corresponding to 7 **phonemic/syllabic** (***/iy/, /piy/, /tiy/, /diy/, /uw/, /m/, /n/***) as well as 4 words (***pat, pot, knew, gnaw***) of 14 participants, with each prompt presented 11 times to each individual.

(from KARA ONE Dataset published by University of Toronto)

Motivation for proposed mixed Neural Network



Attempt to follow neural pathway behind speech.

Higher cognitive processes underlying speech involve :

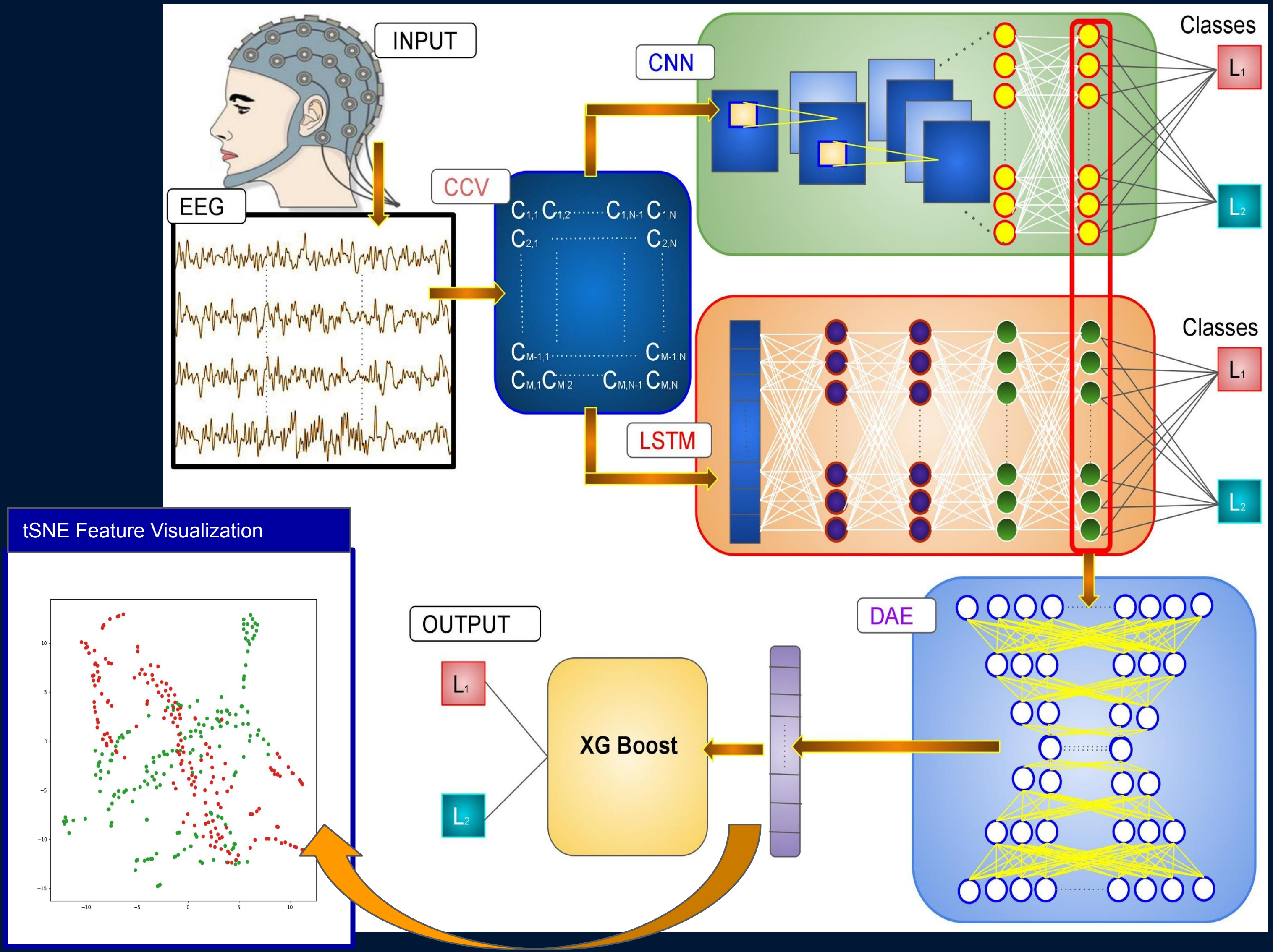
(1) frequent exchange between different parts of brain [encoded here in terms of joint variability inside Channel Cross-covariance matrix (CCV)]

(2) Incorporation of intermediate feedback loops [modeled here through backprops in intermediate CNN-RNN]

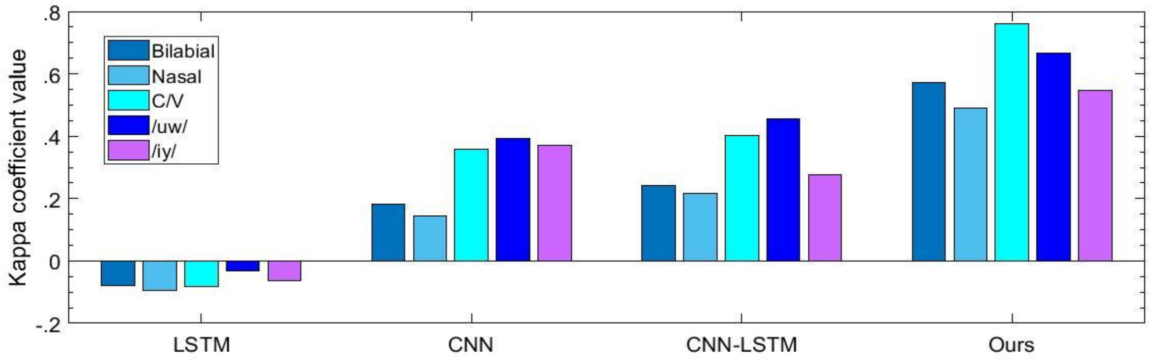
(3) utilization of past information stored [here, through LSTM]

(4) hierarchical combination of several feature extractors [implemented through hierarchical mixture of different networks]

PROPOSED MIXED NEURAL NETWORK



Performance Analysis



Method	±Bilab	±Nasal	C/V	±/uw/	±/iy/
LSTM	46.07	45.31	45.83	48.44	46.88
CNN	59.16	57.20	67.88	69.56	68.60
CNN+LSTM	62.03	60.89	70.04	72.76	63.75
Our Mixed	78.65	74.57	87.96	83.25	77.30

(Train-Dev-Test:80-10-10 split)

Performance Comparison

	±Bilabial	± Nasal	C/V	±/uw/	±/iy/
[1]	56.64	63.5	18.08	79.16	59.6
[2]	53	47	25	74	53
Ours	**75.55**	**73.45**	**85.23**	**81.99**	**73.30**

(Leave-one-subject out cross-validation)

[1] Zhao, Shunan, and Frank Rudzicz. "Classifying phonological categories in imagined and articulated speech." *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015.

[2] Sun, Pengfei, and Jun Qin. "Neural networks based eeg-speech models." *arXiv preprint arXiv:1612.05369* (2016).

Conclusion

- We report a novel hierarchical deep neural network architecture composed of parallel CNN-LSTM and a deep autoencoder for phonological prediction from imagined speech EEG data.
- Previous approaches mostly deal with subject dependent classification of EEG into discrete vowel or word labels.
- This work investigates a subject-invariant mapping of EEG data with different phonological categories, varying widely in terms of underlying articulator motions (eg: involvement or non-involvement of lips and velum, variation of tongue movements etc)
- Not only does our model outperform the existing method by 22.51%, but also shows sufficient reliability and robustness based on our methodic analyses.
- Our ability to successfully exploit phonological information suggests existence of underlying articulatory footprints associated with active thoughts.

THE UNIVERSITY OF BRITISH COLUMBIA

THANK YOU