Modeling electrocorticography signals on the macaque inferior temporal cortex in space, time and frequency domains using hierarchical visual features of convolutional neural networks

*Hiroto Date[1,2], Keisuke Kawasaki[2], Mete Ozay[1], Takumi Hongo[2], Isao Hasegawa[2], Takayuki Okatani[1]

1: Graduate School of Information Sciences, Tohoku University, Japan
2: Graduate School of Medicine and Dental Sciences, Niigata University, Japan

hdate@vision.is.tohoku.ac.jp

Introduction

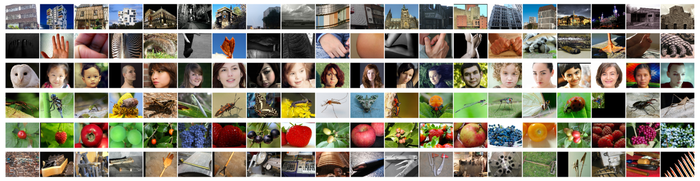
Recent studies on neural frequency-specific activities in primate visual cortex suggest that 1) different frequency bands are related to band-specific directions of cortical information processing [1], and that 2) different frequency bands contain band-specific information. However, it has been unclear what kind of visual features are related to each frequency band. Utilizing deep convolutional neural networks (CNNs), which are currently the most biologically plausible models of the visual system in terms of performance and representation, we addressed this issue by investigating the following questions:

1. Is the relationship between CNNs and electrocorticography (ECoG) activities frequency-specific?
2. Do different frequency bands contain band-specific visual features?
3. How are these band-specific neural representations distributed in space and time?

Materials and Methods

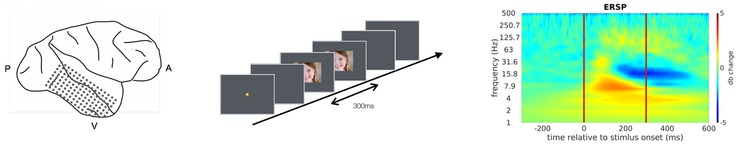
Image set

- 6 object categories (building, body part, face, insect, fruit, tool)
- 600 natural images for each category



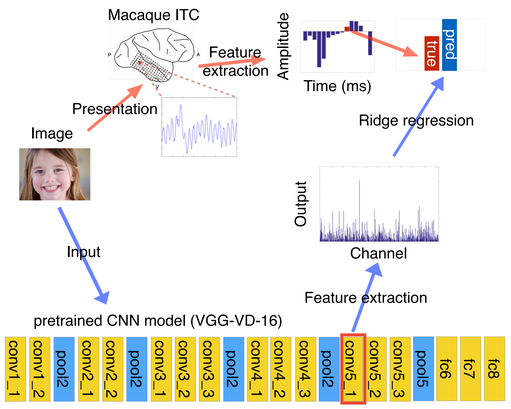
Measuring ECoG activities on macaque inferior temporal cortex (ITC)

- 2 female macaque monkeys
- 128-channel ECoG covering from posterior ITC to anterior ITC
- We computed the amplitude of each frequency (1-250 Hz) by complex Morlet wavelet convolution.



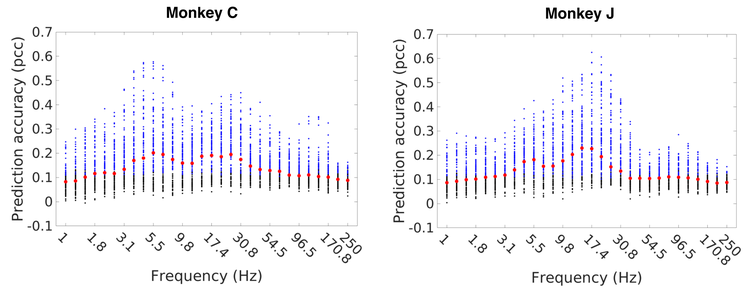
Encoding ECoG activities from hierarchical representations of CNNs

- We extracted internal representations of conv/fc layers in a pretrained CNN model (VGG-VD-16 [2]).
- An encoding model is specified by one ECoG electrode, time window, frequency, and CNN layer.
- We first trained each encoding model by ridge regression, and then evaluated each model's prediction accuracy with test set.
- Prediction accuracy was calculated as the Pearson's correlation coefficient (PCC) between observed and predicted ECoG activities.



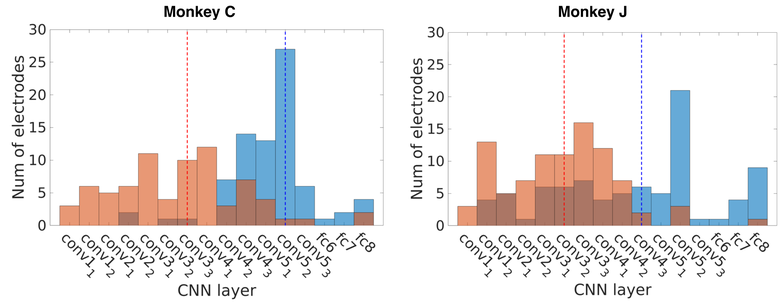
Results

1. Theta and beta-gamma band activities were better predicted from CNN representations



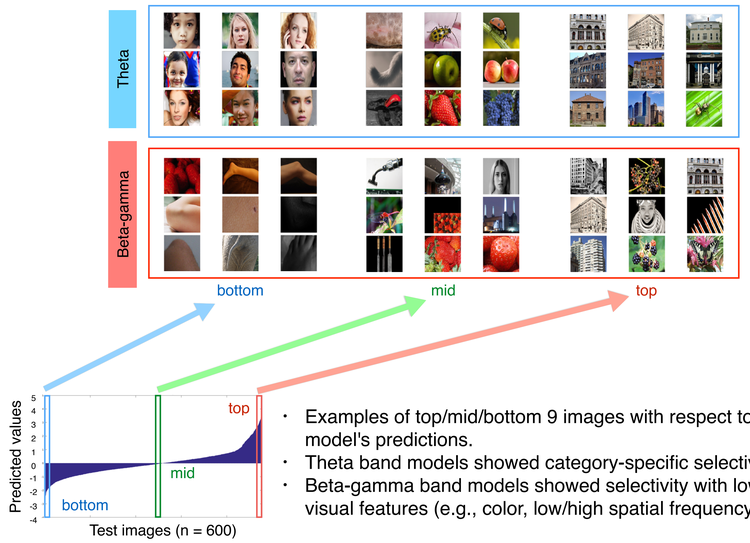
- Prediction accuracy of each frequency (with the best layer and time window of each electrode)
- Blue points: electrodes greater than the chance level ($p < 0.001$, permutation test), black points: electrodes less than or equal to the chance level, red points: median over electrodes

2. Theta and beta-gamma band activities were better predicted from higher and lower CNN layers, respectively



- Visualizations of assigned CNN layers for theta (blue) and beta-gamma (red) activities (PCC at the best time window of all electrodes)
- Vertical lines: median layer of each band's results

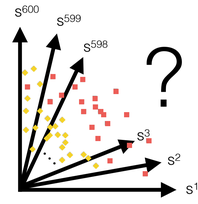
3. Category-specific selectivity of theta band models

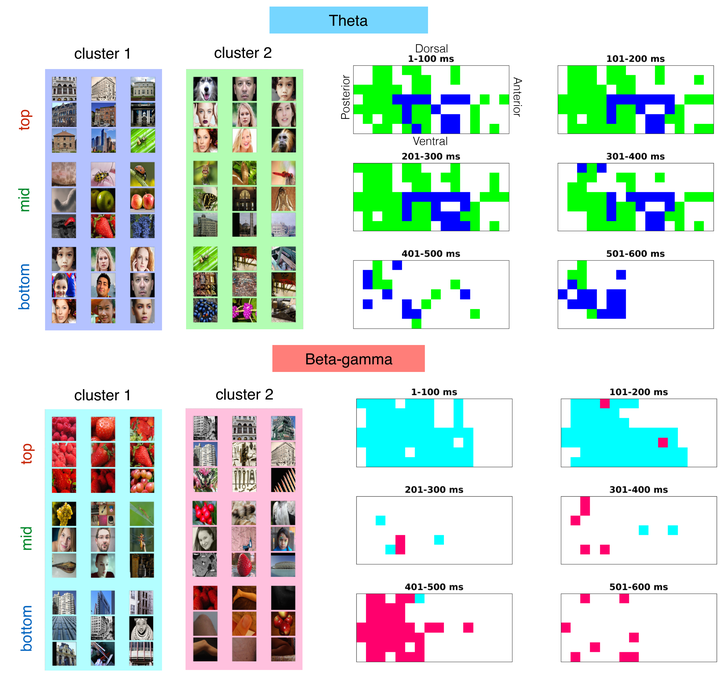


- Examples of top/mid/bottom 9 images with respect to each model's predictions.
- Theta band models showed category-specific selectivity.
- Beta-gamma band models showed selectivity with low-level visual features (e.g., color, low/high spatial frequency).

4. Theta and beta-gamma models were distributed in several spatiotemporal clusters

- We clustered encoding models (electrode, time window) of theta and beta-gamma bands by k-means clustering.
- Strong/weak amplitude of each cluster have different selectivity (e.g., building, face, red, fruit, high/low spatial frequency)
- Spatially and temporally overlapping clusters of theta and beta-gamma bands have different selectivity (cluster 2 of theta versus cluster 1 of beta-gamma).





Conclusion

- Theta and beta-gamma band activities are more related to visual features than other bands.
- Theta band activities are related to higher- and category-level visual features (object category, shape), while beta-gamma band activities to lower-level visual features (color, low/high spatial frequency)
- Theta and beta-gamma band activities have several spatiotemporal clusters that show different selectivity.

References
[1] Michalareas, G., Vezoli, J., Van Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. Neuron, 89(2), 384-397.
[2] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. ICLR 2015, 1–14.

Acknowledgements (grant): KAKENHI (15H05919) and CREST, JST