Individual differences in prefrontal coding of visual features

Qi Lin & Hakwan Lau
Center for Brain Science, RIKEN
Wako, Saitama, Japan

Abstract: The lateral prefrontal cortex (LPFC) is commonly associated with high-level cognition, such as attention, language and cognitive control. However, recent work has demonstrated that it is also critical for basic perceptual functions including object recognition. Here we characterize the role of LPFC in visual processing with computational models. Using a dataset of human fMRI data at 7T, we built encoding models relating visual features extracted from a deep neural network (the image encoder of a CLIP [Contrastive Language–Image Pre-training] network) to brain responses to thousands of natural images. Within each of the eight subjects, we were able to robustly predict responses in patches of LPFC, most notably in FEF (frontal eye field) and vlPFC (ventrolateral PFC) regions. Leveraging these robust encoding models, we then explored the representational structures and screened for images with high predicted responses in LPFC . We found striking individual differences in the coding schemes of LPFC. In contrast, the coding scheme of the ventral visual stream remains more consistent across individuals. Overall, our study demonstrates the under-appreciated role of LPFC in visual processing and suggests that LPFC may underlie the idiosyncrasies in how different individuals experience the visual world. Methodologically, these findings may also explain why previous group studies have often failed to observe robust visual functions in LPFC, as subjects' responses may need to be calibrated individually.

Results and Discussion

Results presented here are based on data from the Natural Scenes Dataset (NSD; [1]). Each of the eight subjects viewed 9209~10,000 unique scene images while being scanned in a 7T MRI scanner and performing a continuous recognition memory task.

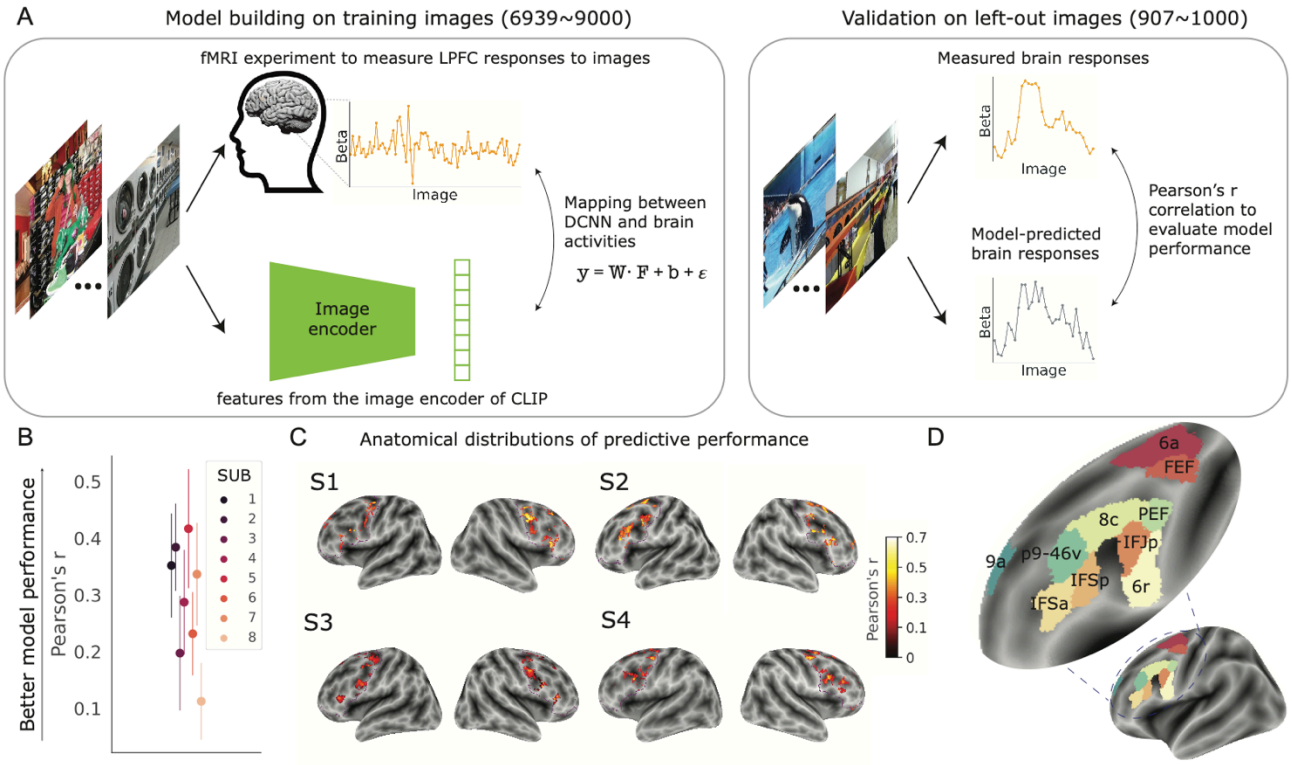

Figure 1. Encoding models robustly predict visual responses in patches of LPFC. (A) Schematic of the cross-validated prediction pipeline. (B) Average Pearson's r between predicted and observed responses in the top 10% LPFC vertices in the held-out data. Each dot represents a subject. Error bars

represent +/- 1 SD. (C) Anatomical distributions of the predictive performance in four example subjects. The contour of the LPFC mask is marked in purple. (D) The 10 parcels with the largest average proportion of visually-sensitive vertices. Parcel labels are defined by [3].

Encoding models robustly predict LPFC activities. To build encoding models of LPFC within each subject (see Figure 1A for a schematic of the prediction pipeline), for the presented images, we first extracted activations (vector length: 512) from the image encoder of a CLIP network (ViT-B/32 backbone; [2]). We divided the images into a training set and a test set. Separately for each vertex in a liberal LPFC mask defined based on [3] (see the purple contours in Figure 1C), we built a ridge regression model to map the CLIP-img features to the brain responses (i.e., single-trial beta estimates) evoked by the training images. We reasoned that only a small proportion of the liberally defined LPFC mask would show sensitivity to visual information and therefore only retained the vertices where the explained variance from the regression models *within the training set* was in the top 10% among all the LPFC vertices. For each of these selected vertices, we then fixed the regression weights and tested the performance of the encoding model in the held-out images. Model performance was assessed as the Pearson's r between predicted vs. observed brain responses to the held-out images. Across all eight subjects, we observed robust prediction of LPFC in the held-out data (Figure 1B; all ps $< .001$ based on 10000 bootstrapping iterations within each subject). Figure 1C shows the anatomical distribution of prediction performance from four example subjects.

To further investigate the anatomical location of visually sensitive vertices, within each parcel defined by [3], we calculated the proportion of bilateral vertices that showed cross-validated Pearson's $r > 0.1$. This was done separately within each subject. Figure 1D shows the 10 parcels with the largest average proportion of visually-sensitive vertices. Consistent with findings from non-human primates (e.g., [4], [5], also reviewed in [6]), we also identified visually-sensitive regions in FEF (frontal eye field) and vlPFC (ventrolateral PFC).

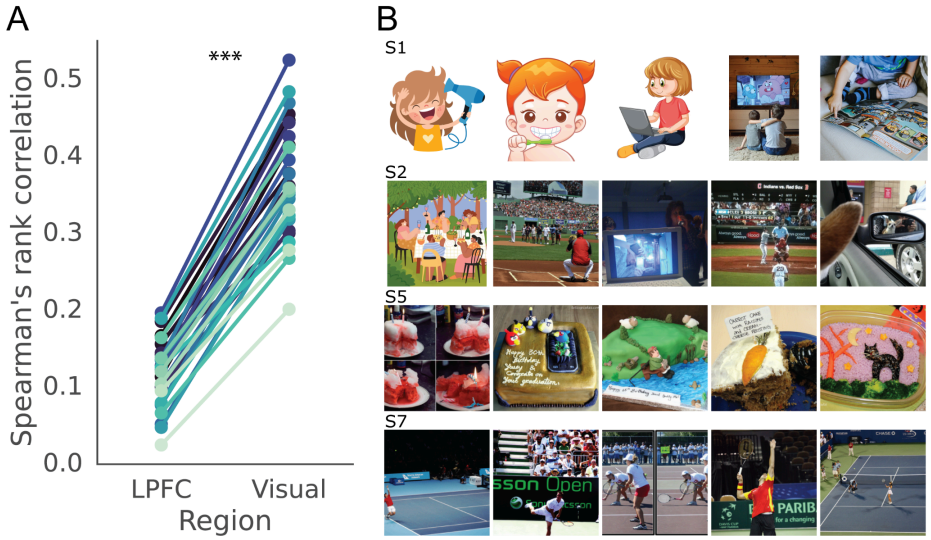

Figure 2. Individual differences in prefrontal coding of visual features. (A) Across-subject similarity of the representational structures in LPFC vs. visual regions. Each line represents a pair of subjects. (B) The 5 images with the largest predicted LPFC responses in 4 representative subjects. Images with identifiable human faces have been replaced by drawings or unidentifiable images.

The coding scheme of PFC is more variable across individuals. Although recent work using the same dataset and whole-brain encoding models also presented results showing predictable vertices in frontal regions (e.g., [7]), these works tend to focus on the visual regions. To fill in this gap, we next asked: what are the visual features being encoded by these visually-sensitive regions in LPFC? And how do they differ from the coding schemes of the ventral visual regions?

To compare the coding scheme of PFC to that of ventral visual regions, we focused on the set of images that were seen three times by all 8 subjects (412 images in total). Within each ROI of each

subject, we calculated the pairwise euclidean distances between the averaged patterns of activities for these images, using only the vertices with cross-validated Pearson's $r > 0.1$. We then compared the similarity (measured with Spearman's rank correlation) of these representational structures (i.e. sets of pairwise activity pattern similarities for all images) between pairs of subjects within each ROI. As shown in Figure 2A, we found that the representational structures of these images showed lower across-subject consistency in LPFC than in the ventral visual regions (Wilcoxon's rank test: $p < .001$). To make sure that this is not simply a result of the higher predictability and larger size of the ventral visual regions than LPFC, we subsampled vertices in ventral visual regions to match the number of included vertices as well as the overall range and mean of cross-validated Pearson's r values in LPFC (i.e., mean difference in cross-validated Pearson's r between the subsampled visual vertices and LPFC < 0.009 within each subject). The observed pattern of lower across-subject consistency in LPFC than in the ventral visual regions still holds after the subsampling ($p < .001$ based on 1000 subsampling iterations).

In addition to the individual differences observed in multivariate responses, inspired by recent work [8], we screened all 73000 images used as stimuli in the NSD dataset through our robust encoding models. Shown in Figure 2B are the 5 images predicted to evoke the largest overall averaged responses in the LPFC of the four representative subjects. Corroborating the representational structure results, inspection of these LPFC-activating images reveals a striking variance across individuals: each subject's LPFC seems to overall prefer a different kind of stimuli.

Conclusion

Capitalizing on the recent developments of computer vision models and massive sampling of neural responses to natural images, we present evidence for robust prediction within regions of LPFC based on visual features. Furthermore, we demonstrate that the coding profiles of these LPFC regions are more variable across individuals compared to higher-level visual regions. Overall, our study suggests that far from being only a content-free region for high-level cognition, LPFC represents rich visual information, may play an important role in visual processing and underlie the idiosyncrasies in how individuals experience the visual world. These results may also partially explain why some previous studies have reported null results for content-specific visual responses in the LPFC; the substantive individual differences reported here might have undermined the sensitivity of studies using mass-univariate approaches with an identical set of stimuli over large groups of subjects. We are currently investigating what are the cognitive implications of these individual differences in LPFC coding.

References:

[1] Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., ... & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience*, *25*(1), 116-126.

[2] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning* (pp. 8748-8763). PMLR.

[3] Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171-178.

[4] Haile, T. M., Bohon, K. S., Romero, M. C., & Conway, B. R. (2019). Visual stimulus-driven functional organization of macaque prefrontal cortex. *Neuroimage*, *188*, 427-444.

[5] Kar, K., & DiCarlo, J. J. (2021). Fast recurrent processing via ventrolateral prefrontal cortex is needed by the primate ventral stream for robust core visual object recognition. *Neuron*, *109*(1), 164-176.

[6] Kornblith, S., & Tsao, D. Y. (2017). How thoughts arise from sights: inferotemporal and prefrontal contributions to vision. *Current Opinion in Neurobiology*, *46*, 208-218.
[7] Wang, A. Y., Kay, K., Naselaris, T., Tarr, M. J., & Wehbe, L. (2023). Better models of human high-level visual cortex emerge from natural language supervision with a large and diverse dataset. *Nature Machine Intelligence*, *5*(12), 1415-1426.
[8] Ratan Murty, N. A., Bashivan, P., Abate, A., DiCarlo, J. J., & Kanwisher, N. (2021). Computational models of category-selective brain regions enable high-throughput tests of selectivity. *Nature Communications*, *12*(1), 5540.
[9] Mehrer, J., Spoerer, C. J., Jones, E. C., Kriegeskorte, N., & Kietzmann, T. C. (2021). An ecologically motivated image dataset for deep learning yields better models of human vision. *Proceedings of the National Academy of Sciences*, *118*(8), e2011417118.