Relational Neural Control: a Method for Investigating Functional Relationships Across the Brain

Alessandro T. Gifford (alessandro.gifford@gmail.com)
Department of Education and Psychology, Freie Universität Berlin
Kaiserswerther Str. 16-18, 14195 Berlin, Germany

Maya A. Jastrzebowska (maya.jastrzebowska@fu-berlin.de)
Department of Education and Psychology, Freie Universität Berlin
Kaiserswerther Str. 16-18, 14195 Berlin, Germany

Johannes J.D. Singer (johannes.singer@arcor.de)
Department of Education and Psychology, Freie Universität Berlin
Kaiserswerther Str. 16-18, 14195 Berlin, Germany

Radoslaw M. Cichy (rmcichy@zedat.fu-berlin.de)
Department of Education and Psychology, Freie Universität Berlin
Kaiserswerther Str. 16-18, 14195 Berlin, Germany

Abstract

The functional relationships between early- and mid-level retinotopic regions of interest (ROIs) of the human visual cortex are not entirely understood. We address this gap by introducing Relational Neural Control (RNC), a neural-control-based method that jointly controls the activity in multiple ROIs by selecting images that align or disentangle their responses. We applied RNC on retinotopic visual cortex using the Neural Encoding Dataset (NED), a massive dataset of synthetic fMRI responses to naturalistic images. RNC found stimulus images that significantly aligned or disentangled both univariate and multivariate responses between retinotopic areas, and these controlling images contained interpretable visual patterns. Our contributions are threefold. First, we provide new quantitative and qualitative findings on functional similarities and differences across retinotopic areas. Second, we introduce RNC as a generalist method for controlling neural responses and uncovering functional relationships across the brain. Third, we release NED with tutorials, hoping it will boost research in cognitive computational visual neuroscience.

Keywords: vision; neural control; neural function; encoding models; fMRI; retinotopic visual cortex; method/data release

Introduction

Early- and mid-level retinotopic ROIs of the human ventral visual stream (i.e., V1, V2, V3, V4) implement key stages of visual information processing. However, the functional relationships of these regions remain incompletely known: how are these ROIs quantitatively encoding visual information, what visual information is being encoded, and how are these properties changing between ROIs? To address these questions, we introduce Relational Neural Control (RNC), a neural-control-based (Bashivan, Kar, & DiCarlo, 2019; Lehky, Sejnowski, & Desimone, 1992; Ponce et al., 2019; Walker et al., 2019) method that jointly controls multiple ROIs by selecting visual stimuli that either align or disentangle their responses.

Methods

Dataset

We trained fMRI encoding models (St-Yves & Naselaris, 2018) of retinotopic ROIs of all 8 Natural Scenes Dataset (NSD) subjects (Allen et al., 2022), using different random seeds and up to 9,000 naturalistic stimulus images per subject, and used the trained models to predict fMRI responses for all 73,000 NSD images. This large battery of synthetic fMRI responses allowed us to investigate, through RNC, the functional properties of retinotopic ROIs in an exploratory, data-driven fashion.

Experiments

To uncover functional relationships between early- and mid-level retinotopic areas, we implemented RNC independently to the synthetic fMRI responses of each pairwise ROI combination (i.e., V1 vs. V2, V1 vs. V3, V1 vs. V4, V2 vs. V3, V2 vs. V4, V3 vs. V4). Furthermore, since the visual information contained in fMRI responses can be investigated at both the global mean and the population code level, we implemented two RNC algorithmic variants: univariate and multivariate control.

Univariate Control Univariate control assumes that visual information is encoded at the level of univariate responses (i.e., the average responses across all voxels within a ROI). This variant used a ranking procedure to find stimulus images that best aligned or disentangled the univariate responses between ROI pairs. This resulted in four neural control conditions (25 controlling images per condition): two control conditions in which both ROIs have aligned univariate responses (i.e., both ROIs have either high or low responses), and two control conditions in which both ROIs have disentangled univariate responses (i.e., one ROI has high responses while the other ROI has low responses, or vice versa).

Multivariate Control Multivariate control assumes that visual information is encoded at the level of multivariate responses (i.e., the population response patterns of all voxels within a ROI). This variant used genetic optimization (Ponce et al., 2019) to find stimulus images that best aligned or disentangled the multivariate responses between ROI pairs, as measured by representational similarity analysis (RSA) (Kriegeskorte, Mur, & Bandettini, 2008). This resulted in two neural control conditions (50 controlling images per condition): one control condition in which both ROIs have aligned multivariate responses (i.e., high RSA correlation score), and one control condition in which both ROIs have disentangled multivariate responses (i.e., low RSA correlation score).

Subject Cross Validation To assess whether the RNC results generalize across subjects, we trained and tested both RNC variants in a leave-one-subject-out fashion: we found the controlling images on N-1 subjects, and tested them on the left out subject, for all 8 subjects.

Results and Discussion

Both RNC variants found stimulus images that successfully aligned and disentangled the fMRI responses between retinotopic ROIs (**Figure 1**).

For univariate control, images that aligned both ROIs led to higher absolute univariate responses compared to images that disentangled them, and disentangling images led to higher absolute univariate responses in pairs of non-adjacent (e.g., V1 vs. V4) compared to adjacent (e.g., V2 vs. V3) ROIs (**Figure 1A**, upper triangular matrix). These observations are in line with the geometry of manifolds of univariate responses for all NSD images (**Figure 1A**, lower triangular matrix). These manifolds indicate a positive relationship between the univariate responses of both ROIs (i.e., high alignment), and that this relationship is strongest for adjacent ROIs. Similarly for multivariate control, images that aligned both ROIs resulted in RSA correlation scores very close to ceiling (i.e., r=1), the disentangling images significantly decorrelated ROI

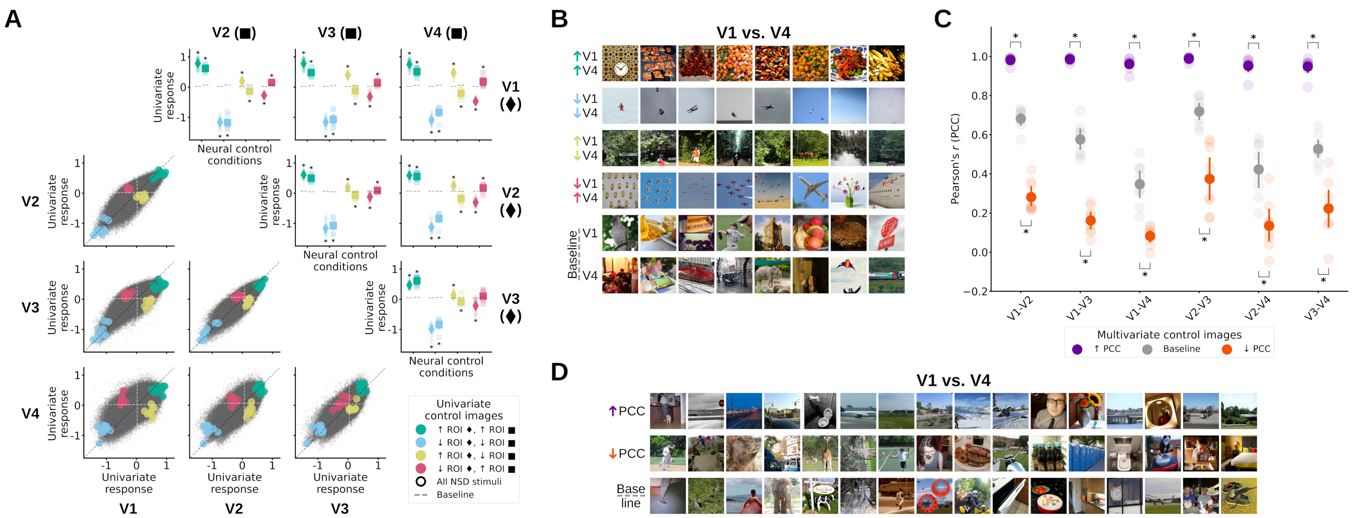

Figure 1: Results of RNC applied to retinotopic ROIs of the 8 NSD subjects. **A.** Neural control scores for the univariate RNC variant. **B.** Examples of controlling images for the univariate RNC variant (V1 vs. V4 comparison). **C.** Neural control scores for the multivariate RNC variant. **D.** Examples of controlling images for the multivariate RNC variant (V1 vs. V4 comparison).

pairs responses, and pairs of non-adjacent ROIs could be best decorrelated (**Figure 1C**). These results suggest that retinotopic ROIs might share both functional similarities and differences in visual information processing, that the similarities might be more pronounced, and that the differences might increase the further away two ROIs are from each other.

Next, we visually inspected the controlling images (here we describe the images from the V1 vs. V4 comparison). For univariate control (**Figure 1B**), the disentangling images suggest that V1 is more responsive (than V4) to high spatial frequencies (e.g., vegetation), and V4 is more responsive (than V1) to objects on low spatial frequency backgrounds (e.g., planes on a sky background). High spatial frequencies and objects are both present in aligning images leading to high univariate responses in both ROIs (e.g., objects on cluttered backgrounds), whereas they are both lacking in aligning images leading to low responses in both ROIs (e.g., small or no objects on uniform backgrounds). This suggests that two ROIs might result in aligned responses not only due to functional similarities, but also in cases where these ROIs preferentially respond to different visual information, which is either co-existing or co-non-existing in controlling images. For multivariate control (**Figure 1D**), the aligning images often contained empty regions (e.g., the sky) in the upper half, whereas the disentangling images did not. Thus, since V1 and V4 are both retinotopic areas, alignment might be largely driven by both ROIs similarly encoding topological properties of images (e.g., the spatial location of objects).

Together, we propose RNC, a new neural-control-based method for discovery of neural functional relationships, and showcase its applicability and potential on early- and mid-level retinotopic visual ROIs, whose visual information encoding properties are not well understood (we are currently collecting fMRI responses for the univariate and multivariate controlling images to validate our findings on real neural data). RNC has three key features. First, given that the brain is an interconnected system, functional properties are best understood if taking into account cross-regional relationships: RNC uncovers both functional similarities and differences between several brain regions. Second for a multifaceted understanding of functional properties, RNC provides both quantitative (i.e., neural control scores) and qualitative (i.e., controlling stimuli) solutions. Third, RNC is a flexible method which can be adapted to different data modalities (e.g., fMRI, ECoG, behavior, AI models); cognitive modalities (e.g., vision, language, audition), stimuli (e.g., hand-picked parameterized images, text from large language corpora, AI-generated audio), data information levels (e.g., univariate responses, multivariate responses, frequency oscillations), spatial scales (e.g., single neurons, populations of neurons, entire brain areas), relationship complexities (e.g., comparing functional relationships between two, three or more ROIs), and research approaches (e.g., hypothesis-based, exploratory).

Tutorials and NED Release

To facilitate RNC adoption and the discovery of new functional relationships between visual areas, we created online tutorials[1] where users can interactively implement univariate and multivariate control on the Neural Encoding Dataset (NED): synthetic fMRI responses for 150,000 naturalistic images, over all 8 NSD subjects, and a choice of 23 ROIs spanning the whole visual hierarchy. We also release NED[2], hoping it will contribute to the cognitive computational visual neuroscience field both practically and paradigmatically: practically, by being immediately usable for hypothesis testing, data-driven exploration, methods development, and model building; paradigmatically, by introducing a new research paradigm centered on synthetic large-scale neural datasets.

[1] https://www.alegifford.com/projects/rnc
[2] https://www.alegifford.com/projects/ned

Acknowledgments

This research was supported by German Research Council (DFG) grants (CI 241/1-1, CI 241/1-3, CI 241/1-7, INST 272/297-1) and the European Research Council (ERC) starting grant (ERC-StG-2018-803370). The authors would like to thank the HPC Service of FUB-IT, Freie Universität Berlin, for computing time (DOI: 10.17169/refubium-26754).

References

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.

Bashivan, P., Kar, K., & DiCarlo, J. J. (2019). Neural population control via deep image synthesis. *Science*, *364(6439)*, eaav9436.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 4.

Lehky, S. R., Sejnowski, T. J., & Desimone, R. (1992). Predicting responses of nonlinear neurons in monkey striate cortex to complex patterns. *Journal of Neuroscience*, *12(9)*, 3568–3581.

Ponce, C. R., Xiao, W., Schade, P. F., Hartmann, T. S., Kreiman, G., & Livingstone, M. S. (2019). Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences. *Cell*, *177(4)*, 999–1009.

St-Yves, G., & Naselaris, T. (2018). The feature-weighted receptive field: an interpretable encoding model for complex feature spaces. *NeuroImage*, *180*, 188–202.

Walker, E. Y., Sinz, F. H., Cobos, E., Muhammad, T., Froudarakis, E., Fahey, P. G., ... Tolias, A. S. (2019). Inception loops discover what excites neurons most using deep predictive models. *Nature neuroscience*, *22(12)*, 2060–2065.