Reinforcement Learning Over Complex Naturalistic Scenes

Daniel B. Ehrlich (dbehrlich@berkeley.edu)
Department of Psychology, 2121 Berkeley Way
Berkeley, CA 94704 United States

Anne G.E. Collins (annecollins@berkeley.edu)
Department of Psychology, 2121 Berkeley Way, Helen Wills Neuroscience Institute
Berkeley, CA 94704 United States

Abstract:

Significant strides have been made in understanding reward-based learning in neuroscience and psychology. However, many experimental paradigms employ overly simplistic stimuli, diverging from real-life experiences. Here, we investigate how humans learn over well-controlled naturalistic stimuli from the ImageNet dataset. We used a pre-trained neural network to partition scenes into categories of visual and semantic features, drawn from differing levels of abstraction. Recent research highlights the intricate interaction between learning and cognitive processes, like working memory, attention, and perception. We hypothesized that the level of feature abstraction might impact recruitment of cognitive processes during learning. Indeed, we found that far from learning parameters being constant across these varied types of scene constructions, the types of features used impacted overall performance and the learning parameters drawn from best fit models.

Keywords: reinforcement learning; reward; neural network

Introduction

The complex cacophony of natural scenes sits a far distance from the simple stimuli used in most laboratory reinforcement learning (RL) experiments. Limited literature exists to help bridge that gap, partially due to the difficulty of working with complex and variable natural scenes. Prior research on naturalistic stimuli in RL hints at the importance of considering how stimuli impact learning but has tended to use highly restricted stimuli sets (Farashahi et al., 2020). In this study we develop a method to sort arbitrary visual image sets into well-controlled, structured categories for RL tasks. We then analyze human RL behavior on a set of categories drawn from a diverse natural scene collection. RL and working memory (WM) are impacted by natural images (Yoo, Keglovitis & Collins, 2023; Brady, Störmer & Alvarez, 2016). We hypothesized that these cognitive processes might differentially impact reward learning over differing stimuli features. We test this theory by manipulating the level of semantic abstraction of the stimulus categories in the reward learning task.

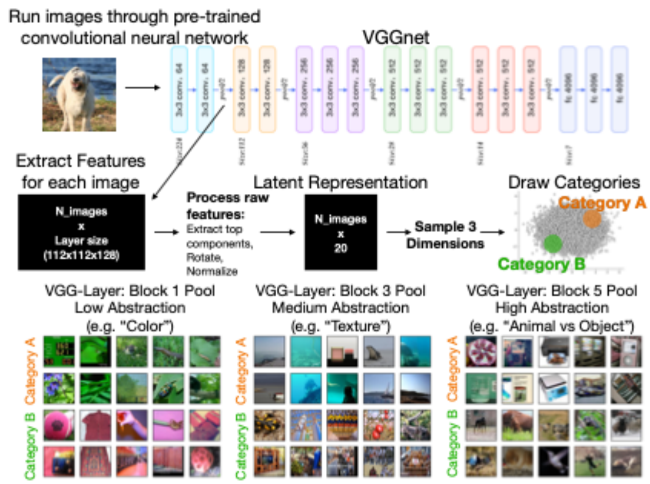

Figure 1. Summary of Automated Category Generation Algorithm. (Top row) Run images through pre-trained ConvNet; (Middle row) Extract features from unit activity, pre-process, draw two categories; (Bottom row) Example image categories by network depth.

Methods

Category Generation

We developed a protocol for automated stimuli set construction that controls for overall distance between two categories over different types of image features. In brief, we use the activations from intermediate layers in a pre-trained deep convolutional network to create a latent space on which we define categories. For this paper we use the ImageNet dataset (Deng et al., 2009) and VGG-net (Simonyan & Zisserman, 2014) but methods broadly generalize to alternative networks and datasets. Layer depth of latent representations correlates to the level of semantic abstraction (Yamins & DiCarlo, 2016). We use the block 1, 3 and 5 pooling layers of VGG-net as our latent representations for “low”, “medium” and “high” abstraction, respectively. Our method allows us to generate pairs of image categories that are well controlled in their within category similarity and across category distance, with variance coming from features at specific abstraction levels.

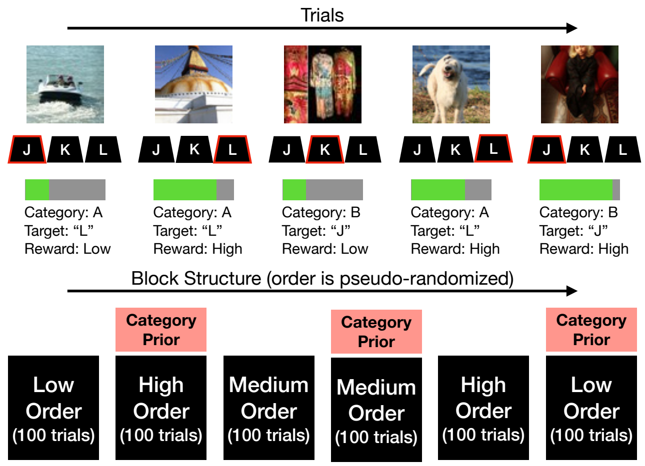

Figure 2: RL Task Structure. (Top) On each trial participants saw a single image from our dataset, respond by pressing one of three keys, then received a High or Low reward; (Bottom) Block structure for task. Each level of abstraction was run twice, once with a "category prior" at the block start. Blocks were independent, each with new categories and images.

Task Structure

Participants learned category-key associations from reward. Each image was an exemplar of one of two categories in that block and was presented only once in the experiment. Each category had one high value key. To isolate the effect of category understanding on reward learning, we introduced a "category prior" in 3 blocks: participants were first shown 8 labeled images from each of that block's categories to allow them to form category prototypes prior to reward learning onset. Example images were not used in the learning phase.

Modeling Approach

Working Memory (WM) Model

$$WM_{a,r} = stim_t^{a,r}$$

$$WM_{t+1} = (1-\lambda)WM_t + \lambda WM_0$$

$$Vwm^{a,stim} = dist(WM_{a,r=1}, stim_{t+1}) - dist(WM_{a,r=0}, stim_{t+1})$$

$$p(a|stim) \propto exp(\beta Vwm^{a,stim})$$

Reinforcement Learning (RL) Model

$$Vrl_t^{a,stim} = \theta_t^a \cdot stim_t$$

$$p(a|stim) \propto exp(\beta Vrl^{a,stim})$$

$$\theta_{t+1}^a = \theta_t^a + \alpha \nabla L(\theta_t^a)$$

For feature RL:

$$Vrl_t^{a,stim} = \sum_l^{level} w_l * \theta_t^{a,l} \cdot stim_t^l$$

RLWM Model

$$p(a|stim) = \rho^l p_{RL}(a|stim) + (1-\rho^l)(p_{WM}(a|stim)$$

Eqs 1: WM: WM buffer, λ: decay, V: value, β: softmax temp, θ: RL approximation params, α: learning rate, w: RL feature weight, ρ: strategy mixing param

We fit reinforcement learning (RL), working memory (WM) and hybrid (RLWM) models adapted to naturalistic scenes (Collins et al., 2014). The RL model uses a value function approximation approach. The WM model stores recent stimuli representations as response exemplars. The RLWM model combines these policies with a fit mixing parameter. Models were fit with standard best practices (Wilson & Collins, 2019).

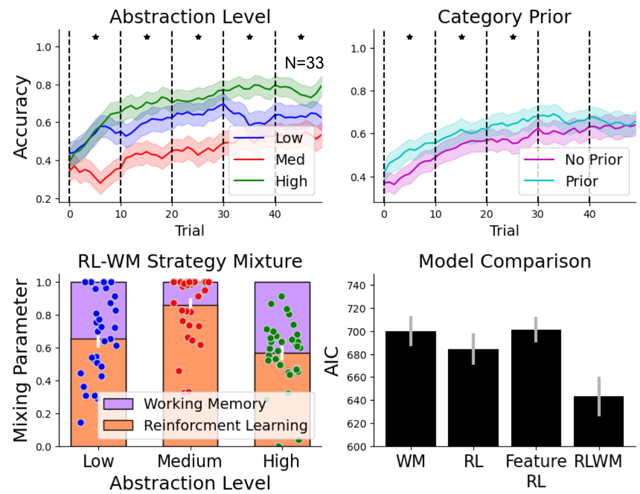

Figure 3: Behavioral and Model Results. (Top Left) Moving average of accuracy (chose high reward key) by level of category abstraction. * indicates p<0.05 within the bin (dashed lines) (Top Right) As left but for blocks with and without a category prior. (Bottom Left) Best fit mixing parameters for RL-WM by abstraction (Bottom Right) Model comparison

Results

There was a strong effect of abstraction level on performance (One-way ANOVA, F=14.4, p=2.6e-5). Category prior availability impacted early learning (bins: 0-30) but not asymptotic performance (Rel. T-Test, T=-2.8,-2.8,-2.9, p=0.008,0.009,0.007). Model fitting results showed that the RLWM model outperformed the WM, RL and a modified RL model with features weighted by abstraction level. (T-Test, T=-6.7,5.3,4.3, p=2e-5,2e-4,9e-4). In RLWM, participants recruited significantly more WM in the high and low abstraction levels than medium (T-Test, T=-3.9,-7.9,p=5e-4,5e-9).

Conclusion

Studying reinforcement learning in naturalistic stimuli will help extend prior insights into more realistic environments. Here we show that learning performance and cognitive mechanisms depend on level of abstraction, even with category differences controlled. Intriguingly our data indicates a U-shaped pattern in WM perhaps indicating preferential access to early and late visual information. In future work, we will explore how category priors impact cognitive strategy.

References

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). *Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli*. *Proceedings of the National Academy of Sciences*, *113*(27), 7459-7464.

Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). *Working memory contributions to reinforcement learning impairments in schizophrenia*. *Journal of Neuroscience*, *34*(41), 13747-13756.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: *A large-scale hierarchical image database*. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

Farashahi, S., Xu, J., Wu, S. W., & Soltani, A. (2020). *Learning arbitrary stimulus-reward associations for naturalistic stimuli involves transition from learning about features to learning about objects*. *Cognition*, *205*, 104425.

Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. *arXiv preprint arXiv:1409.1556*.

Wilson, R. C., & Collins, A. G. (2019). *Ten simple rules for the computational modeling of behavioral data*. *Elife*, *8*, e49547.

Yamins, D. L., & DiCarlo, J. J. (2016). *Using goal-driven deep learning models to understand sensory cortex*. *Nature neuroscience*, *19*(3), 356-365.

Yoo, A. H., Keglovits, H., & Collins, A. G. (2023). *Lowered inter-stimulus discriminability hurts incremental contributions to learning*. *Cognitive, Affective, & Behavioral Neuroscience*, *23*(5), 1346-1364.