Active Vision with Predictive Coding and Uncertainty Minimization

Abdelrahman Sharafeldin
Machine Learning Center
Georgia Institute of Technology
abdo.sharaf@gatech.edu

Nabil Imam
School of Computational Science and Engineering
Georgia Institute of Technology
nimam6@gatech.edu

Hannah Choi
School of Mathematics
Georgia Institute of Technology
hannahch@gatech.edu

Abstract

We present an end-to-end procedure for embodied visual exploration based on two biologically inspired computations: predictive coding and uncertainty minimization. The procedure can be applied in a task-independent and intrinsically driven manner. We evaluate our approach on an active vision task, where an agent actively samples its visual environment to gather information. We show that our model builds unsupervised representations through exploration that allow it to efficiently categorize visual scenes. We further show that using these representations for downstream classification leads to superior data efficiency and learning speed compared to other baselines while maintaining lower parameter complexity. Finally, the modularity of our model allows us to probe its internal mechanisms and analyze the interaction between perception and action during exploratory behavior.

1 Introduction

Biological organisms interact with the world in cycles of perception and action. These two process are intertwined and interact with one another to guide animal behavior Guillery and Sherman (2011); Guillery (2005); Linson et al. (2018); Friston (2010b). Visual perception, for example, is not passive. Rather, we actively sample our visual field in search for information, a process referred to as active vision in neuroscience and psychology Yarbus (1967); Hayhoe and Ballard (2005); Krajbich et al. (2010); Land and Tatler (2009); Friston et al. (2012). In contrast, most models of artificial intelligence (AI), e.g. convolutional neural networks (CNN), treat perception and action as separate processes and aim to optimize performance with respect to task-specific objectives, e.g. classification accuracy. In this paper, we integrate two theories from systems neuroscience to develop a combined perception-action model for intrinsically-driven active sensing. We base the perception component of our model on the theory of predictive coding Rao and Ballard (1999). According to predictive coding, the brain maintains a generative model of the world which it uses to predict its sensory input. The goal of perception, therefore, is to infer the latent states of this generative model such as to minimize prediction error. The action component of our model is based on the proposition that the brain minimizes uncertainty of inferred latent states during exploratory behavior Butko and Movellan (2010, 2008); Little and Sommer (2013); Friston (2010b). Due to intractability of the uncertainty reduction objective (or equivalently, information gain objective), most models that optimize it rely either on sample-inefficient reinforcement learning methods or on restrictive assumptions that make it easier to evaluate. In our approach, we use a deep generative model based on predictive coding

Information-Theoretic Principles in Cognitive Systems Workshop at the 37th Conference on Neural Information Processing Systems (NeurIPS 2023).

that allows us to optimize a Monte Carlo (MC) approximation to the information gain objective in a fully differentiable manner without assuming explicit knowledge of the true generative model of the environment. We show that this approximation, even when done in a greedy fashion, leads to a highly efficient exploration strategy.

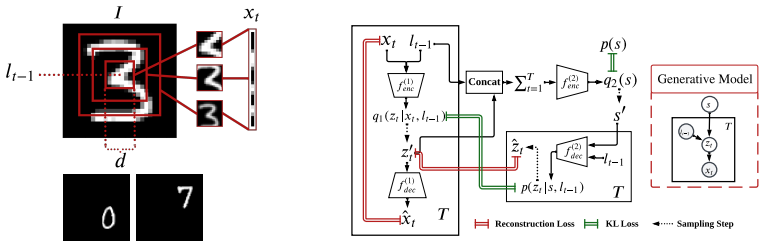
Our model integrates perception and action within an end-to-end differentiable procedure and can be applied in a task-independent manner without the need for extrinsic reward signals. To illustrate, we evaluate our model on an active vision task, where the model has a band-limited sensor used to perceive small patches of a hidden image through a limited number of fixations. We show that, in this setting, the model is able to learn the spatial relationships between pixels of a given image, as demonstrated by its ability to generate full meaningful images by combining smaller generated patches at different locations. Furthermore, we show that although these representations are learned unsupervised, they enable a downstream classifier to quickly reach high test performance with fewer training data and lower parameter complexity. We compare these results to a feedforward network receiving full images as well as to other popular baselines from RL literature including the Recurrent Attention Model (RAM) (Mnih et al., 2014), VIME (Houthooft et al., 2016), and Plan2Explore (Sekar et al., 2020). Finally, because of the modularity of our framework, we are able to probe the relationship between perception and action throughout learning and exploration. For example, we show that during active vision, the model learns representations that reflect the properties of the data and the structure of the task. Our approach demonstrates the promise of integrating neuroscientific theories of perception and action into embodied AI agents, and we hope that it will motivate more research in this area.

A survey of related work is provided in Appendix D.

2 Model Description

2.1 Task

We apply our model to an active vision task, where the state and action spaces are continuous and the state space is high-dimensional. In this task, the model explores a hidden image through a sequence of fixations. Each fixation is a sample of the image at a given location. Furthermore, 'foveated' samples can be extracted by the process illustrated in Figure 1a. Specifically, let l_{t-1} denote the location of sample x_t from the input image I. We use normalized coordinates so $l_{t-1} \in [-1, 1] \times [-1, 1]$, with $(-1, -1)$ corresponding to the top left corner. We first extract N_{fov} patches (the number of red squares in Fig.1a) of increasing size, all centered at l_{t-1}. We then downsample all patches so they all have the same size $d \times d$. The patches are then flattened and concatenated to generate x_t, which is the input to the model. Note that this is the same setup used in Mnih et al. (2014).



(a) (b)

Figure 1: (a) Foveation setup for the bandlimited sensor in the active vision task (top) and translated MNIST examples (bottom). (b) Generative model and architecture for the active vision model. Shaded and unshaded circles represent observed and latent variables, respectively.

2.2 Perception

The active vision perception model is based on a simple hierarchical two-level generative model shown in Figure 1b. Let I denote the image presented on a given trial. The higher level of the perception model encodes a single abstract representation s which may reflect high-level properties of the class to which I belongs. The lower level contains individual units whose activations z_t at time t are entirely driven by the sensory input x_t observed at location l_{t-1} in the image.

For notational simplicity, we will often omit the conditioning in the variational posteriors, e.g. use $q(z_{1:T}, s)$ to refer to $q(z_{1:T}, s|x_{1:T}, l_{0:T-1})$. The ELBO for our generative model, derived in Appendix B, is found to be

$$\mathcal{L}_{ELBO} = \sum_{t=1}^{T} \mathbb{E}_q[\log p(x_t|z_t)] - \mathbb{E}_q\Big[\frac{p(s)}{q_2(s|z_{1:T}, l_{0:T-1})}\Big] - \sum_{t=1}^{T} \mathbb{E}_q\Big[\frac{p(z_t|s, l_{t-1})}{q_1(z_t|x_t, l_{t-1})}\Big] \quad (1)$$

Throughout our experiments, we assume the prior over s to be a standard Gaussian. We also assume the likelihood distributions, $p(x_t|z_t)$ and $p(z_t|s, l_{t-1})$, as well as the variational posteriors, q_1 and q_2, to be Gaussian with means and variances parameterized by feed-forward neural networks. The full architecture of the model with these networks is shown in Figure 1b and is described below.

The perception architecture consists of two variational autoencoders, one for the posterior over $z_{1:T}$ and one for the posterior over s. The encoders and decoders for both VAEs are simple feedforward networks. A more detailed description of each model component is included in Appendix C.

2.3 Action

The goal of action selection in our model is to maximize the expected reduction uncertainty. This goal leads to the following value function for a given action (or fixation location) l_t

$$V(l_t|x_{1:t}, l_{0:t-1}) := H(s|x_{1:t}, l_{0:t-1}) - \mathbb{E}_{p(x_{t+1}|s,l_t)}\Big[H(s|x_{1:t+1}, l_{0:t})\Big], \quad (2)$$

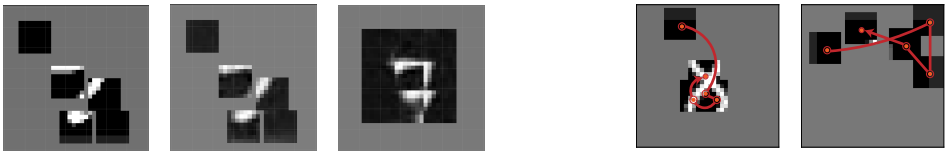
where $H(\cdot)$ denotes the Shannon entropy. Intuitively, $V(l_t)$ quantifies how much information the agent expects to gain as a result of observing the input image at location l_t. Computing the expectation in 2 is intractable. So, instead, we compute an approximation of it using a Monte Carlo (MC) sampling approach. To make our model end-to-end differentiable, we use a neural network to select fixation locations that maximize the MC-approximated objective. This action network can be trained with gradient descent because all the terms in 2 are computed from the neural networks in the perception model, which receive fixation locations as part of their input. Therefore, it's possible to compute gradients with respect to the output of the action network. The pseudocode describing our differentiable approach for Bayesian Action Selection is included in Algorithm 1 in Appendix E.

3 Results

We test our active vision model on multiple image data datasets including MNIST, fashion MNIST, and grayscale CIFAR-10. First, we test the model's ability to produce meaningful images by generating and combining small patches at different locations. This ability reflects an implicit understanding of the spatial relationships between pixels on a given image. Figure 2 shows examples of trials in which a random sequence of patches is given to the perception model. At the end of the sequence, the model infers the abstract state s which is then used by the decoder network to generate patches at the (unobserved) nine central locations. As seen in Figures 2a and S2a, the generated patches show a meaningful image of a digit consistent with the random patches observed by the model. This demonstrates that the model successfully learns the spatial relationships between patches corresponding to individual digits in a completely unsupervised manner, which explains its superior performance during classification later on.

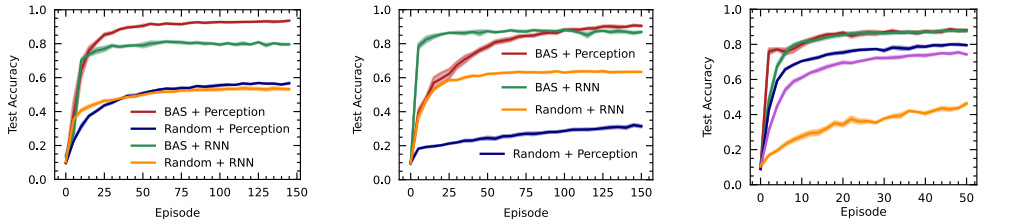
We also examined how the representations learned by the perception model affect what actions are selected. In the centered MNIST dataset, the most informative location about the class of an image is the center. Therefore, a strategy that minimizes uncertainty would ideally choose to fixate at the center most of the time. Figures 2b and S2b show that this is exactly the case; the BAS strategy almost always chooses the center as its second fixation location after the initial random fixation. This shows that the statistical regularities in the environment are reflected in the behavior of the action model. We also studied the effect of action selection on the representations learned by the perception model. By training two perception models on randomly collected and BAS-collected data, we find that the BAS-trained model is able to learn much better representations that are well clustered (according to class) in the principal component space. These results are shown in Figure S4.

Despite the model being trained with unsupervised objectives, we tested its representations on a downstream image classification task, where only a separate decision network is trained with the supervised classification loss. We also tested the model's translation invariance using the translated MNIST dataset which consists of 60×60 pixel images with a handwritten digit placed at a random



(a) (b)

Figure 2: (a) Original patches of input images (left) and their reconstructions (middle). After the model infers an abstract representation, it is able to generate an imagined digit at the unobserved locations (right). (b) Fixation sequences generated using BAS (left) and random strategies (right).



(a) (b) (c)

Figure 3: Performance on (a) the centered MNIST dataset (N = 5) and (b) the translated MNIST dataset (N = 5). Error bars indicate SEM. (c) Training speed comparisons.

location in the image (Figure 1a). To investigate the role of each model component, we evaluate four variants of our model corresponding to the combinations of two options for action selection, **BAS** versus **Random**, and two options for classifier input: **Perception** representations versus states from an **RNN** used to integrate collected observations. Figures 3a and 3b show performance on the centered and translated MNIST datasets, respectively. In general, our BAS strategy yields a better performance on both tasks. Furthermore, when the perception representations are used as input to the classifier, the accuracy is higher compared to using a separate RNN to integrate past observations, indicating that these representations are more informative about the data.

Finally, since our model is trained unsupervised to capture informative parts of an image, we ask if it leads to improved the computational efficiency and training speed for downstream classification. To test this, we look at the learning speed of a downstream classifier trained with full images (Full Images + FF) versus one trained with BAS-collected patches on the translated MNIST data. We also included three popular baselines from the RL literature in this evaluation: the Recurrent Attention Model (RAM) Mnih et al. (2014), VIME Houthooft et al. (2016), and Plan2Explore Sekar et al. (2020). Note that since BAS selects a few locations to observe, the total number of supervised trainable parameters is approximately 50% less for our model than for the model trained with full images. Figure 3c shows that, in addition to having lower parameter complexity, our method learns much faster than all other baselines and achieves higher asymptotic performance than VIME, Plan2Explore, and Full Images + FF. In terms of data efficiency (Figure S3), our model is able to reach higher test accuracy after observing significantly fewer training examples for the first time, indicating its superior ability in generalization and few-shot learning. A description of these experiments and hyperparameters used is included in Appendix E.

4 Conclusion

We developed a novel, biologically-inspired model of active sensing by combining two theories from neuroscience: predictive coding for perception and uncertainty minimization for action. While these two theories have been utilized previously, our model incorporates them in a unique, scalable, and end-to-end framework, enabling flexible intrinsically-driven exploration for embodied AI. We show that this model, despite being driven in an purely intrinsic manner, discovers information action policies and generalizable sensory representations, giving us insight into possible computational strategies employed by the brain.

References

Amin, S., Gomrokchi, M., Aboutalebi, H., Satija, H., and Precup, D. (2020). Locally persistent exploration in continuous control tasks with sparse rewards. *CoRR*, abs/2012.13658.

Barto, A. G. (2013). Intrinsic motivation and reinforcement learning. *Intrinsically motivated learning in natural and artificial systems*, pages 17–47.

Butko, N. J. and Movellan, J. R. (2008). I-pomdp: An infomax model of eye movement. In *2008 7th IEEE International Conference on Development and Learning*, pages 139–144.

Butko, N. J. and Movellan, J. R. (2010). Infomax control of eye movements. 2(2):91–107.

Friston, K. (2010a). The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138.

Friston, K., Adams, R. A., Perrinet, L., and Breakspear, M. (2012). Perceptions as hypotheses: saccades as experiments. *Frontiers in psychology*, 3:151.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2017). Active inference: a process theory. *Neural computation*, 29(1):1–49.

Friston, K. J. (2010b). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11:127–138.

Gregor, K., Danihelka, I., Graves, A., Rezende, D., and Wierstra, D. (2015). Draw: A recurrent neural network for image generation. In Bach, F. and Blei, D., editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1462–1471, Lille, France. PMLR.

Guillery, R. (2005). Anatomical pathways that link perception and action. In *Cortical Function: a View from the Thalamus*, volume 149 of *Progress in Brain Research*, pages 235–256. Elsevier.

Guillery, R. and Sherman, S. M. (2011). Branched thalamic afferents: What are the messages that they relay to the cortex? *Brain Research Reviews*, 66(1):205–219. Camillo Golgi and Modern Neuroscience.

Han, K., Wen, H., Zhang, Y., Fu, D., Culurciello, E., and Liu, Z. (2018). Deep predictive coding network with local recurrent processing for object recognition. *CoRR*, abs/1805.07526.

Hayhoe, M. and Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4):188–194.

Hazan, E., Kakade, S. M., Singh, K., and Soest, A. V. (2018). Provably efficient maximum entropy exploration. *CoRR*, abs/1812.02690.

Higgins, I., Matthey, L., Pal, A., Burgess, C. P., Glorot, X., Botvinick, M. M., Mohamed, S., and Lerchner, A. (2016). beta-vae: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*.

Houthooft, R., Chen, X., Duan, Y., Schulman, J., Turck, F. D., and Abbeel, P. (2016). Curiosity-driven exploration in deep reinforcement learning via bayesian neural networks. *CoRR*, abs/1605.09674.

Jiang, L. P. and Rao, R. P. N. (2021). Predictive coding theories of cortical function.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., and Saul, L. K. (1999). *An Introduction to Variational Methods for Graphical Models*, page 105–161. MIT Press, Cambridge, MA, USA.

Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13(10):1292—1298.

Land, M. F. and Tatler, B. W. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford University Press.

Linson, A., Clark, A., Ramamoorthy, S., and Friston, K. (2018). The active inference approach to ecological perception: General information dynamics for natural and artificial embodied cognition. *Frontiers in Robotics and AI*, 5.

Little, D. Y. and Sommer, F. T. (2013). Learning and exploration in action-perception loops. *Frontiers in neural circuits*, 7:37.

Lotter, W., Kreiman, G., and Cox, D. D. (2016). Deep predictive coding networks for video prediction and unsupervised learning. *CoRR*, abs/1605.08104.

Marino, J. (2022). Predictive Coding, Variational Autoencoders, and Biological Connections. *Neural Computation*, 34(1):1–44.

Ménard, P., Domingues, O. D., Jonsson, A., Kaufmann, E., Leurent, E., and Valko, M. (2020). Fast active learning for pure exploration in reinforcement learning. *CoRR*, abs/2007.13442.

Millidge, B., Song, Y., Salvatori, T., Lukasiewicz, T., and Bogacz, R. (2022). Backpropagation at the infinitesimal inference limit of energy-based models: Unifying predictive coding, equilibrium propagation, and contrastive hebbian learning.

Mnih, V., Heess, N., Graves, A., and kavukcuoglu, k. (2014). Recurrent models of visual attention. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc.

Mohamed, S. and Rezende, D. J. (2015). Variational information maximisation for intrinsically motivated reinforcement learning. *Advances in neural information processing systems*, 28.

Ororbia, A. and Mali, A. (2023). Convolutional neural generative coding: Scaling predictive coding to natural images.

Ororbia, A., Mali, A., Giles, C. L., and Kifer, D. (2020). Continual learning of recurrent neural networks by locally aligning distributed representations. *IEEE Transactions on Neural Networks and Learning Systems*, 31(10):4267–4278.

Ororbia, A. and Mali, A. A. (2021). Backprop-free reinforcement learning with active neural generative coding. *CoRR*, abs/2107.07046.

Ororbia, A., Mali, A. A., Kifer, D., and Giles, C. L. (2019). Lifelong neural predictive coding: Sparsity yields less forgetting when learning cumulatively. *CoRR*, abs/1905.10696.

Pathak, D., Agrawal, P., Efros, A. A., and Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. *CoRR*, abs/1705.05363.

Rao, R. P. N. and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2:79–87.

Salvatori, T., Song, Y., Lukasiewicz, T., Bogacz, R., and Xu, Z. (2021). Predictive coding can do exact backpropagation on convolutional and recurrent neural networks. *CoRR*, abs/2103.03725.

Schmidhuber, J. (1991). A possibility for implementing curiosity and boredom in model-building neural controllers. In *Proc. of the international conference on simulation of adaptive behavior: From animals to animats*, pages 222–227.

Sekar, R., Rybkin, O., Daniilidis, K., Abbeel, P., Hafner, D., and Pathak, D. (2020). Planning to explore via self-supervised world models. In *ICML*.

Still, S. and Precup, D. (2012). An information-theoretic approach to curiosity-driven reinforcement learning. *Theory in Biosciences*, 131:139–148.

Storck, J., Hochreiter, S., Schmidhuber, J., et al. (1995). Reinforcement driven information acquisition in non-deterministic environments. In *Proceedings of the international conference on artificial neural networks, Paris*, volume 2, pages 159–164.

Sun, Y., Gomez, F. J., and Schmidhuber, J. (2011). Planning to be surprised: Optimal bayesian exploration in dynamic environments. *CoRR*, abs/1103.5708.

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).

Yarbus, A. L. (1967). *Eye movements and Vision*. Plenum Press.

A Relationship between Predictive Coding and Variational Autoencoders

According to the predictive coding framework, the brain maintains a generative model of the world which approximates a mapping between observed sensory input and hidden states of the environment. This is illustrated in Figure S1. Perception, therefore, corresponds to inverting this model to infer hidden states, while learning corresponds to updating the parameters of this model based on prediction errors. Here, we outline the relationship between hierarchical predictive coding as presented in Rao and Ballard (1999) and the framework of Variational Auto-encoders in machine learning [Kingma and Welling (2013)]. A similar outline of this relationship is given in Jiang and Rao (2021) and Marino (2022) (with more details on the connections between theory and biology).

To simplify the discussion, we assume the generative model consists of two hierarchical layers (an input layer and a sensory layer) as shown in Figure S1. In reality, the sensory areas in the brain contain many more hierarchical levels, and this discussion can be easily extended to multi-layer hierarchical generative models.

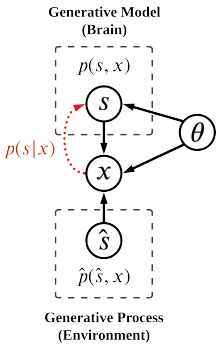


Figure S1: Simple two-layer hierarchical generative model, parameterized by θ, which approximates the true distribution of the generative process giving rise to observations x. The predictive coding framework postulates that neural activities encode hidden state estimates, e.g. s. Thus, perception corresponds to inverting the generative model and inferring those neural activities (red arrow).

The goal of inference is to find the best estimate $s^\star$ of the true hidden state $\hat{s}$ given an observation x. In the predictive coding framework, this is done by maximizing the posterior distribution $p(s|x)$, which is known as maximum a posteriori (MAP) estimation. Equivalently, we can maximize $\log p(s|x)$ since the log is a monotonic function in p. This problem can be formulated as follows

$$s^\star = \arg\max_s \log p(s|x) \tag{3}$$

$$= \arg\max_s \log \frac{p(x|s)p(s)}{p(x)} \tag{4}$$

$$= \arg\max_s [\log p(x|s) + \log p(s)] \tag{5}$$

To perform this optimization, we adopt some assumptions about the likelihood distribution $p(x|s)$ and the prior on the hidden state $p(s)$. In their original implementation, Rao and Ballard assume the following parameterizations

$$p(x|s) = \mathcal{N}(f(\boldsymbol{W}s); \sigma_x^2 \boldsymbol{I}) \tag{6}$$

$$p(s) = \mathcal{N}(\mu_s; \sigma_s^2 \boldsymbol{I}) \tag{7}$$

where $\boldsymbol{W}$ is a weight matrix, $f(\cdot)$ is a non-linearity, and $\mathcal{N}(\mu, \sigma \boldsymbol{I})$ is an isotropic gassuain with mean μ and covariance $\sigma \boldsymbol{I}$. Without loss of generality, we can simplify this further by assuming the prior $p(s)$ is a standard gaussian, i.e. $\mu_s = \mathbf{0}$ and $\sigma_s^2 = 1$. Substituting this into equation 5, we get

$$s^\star = \arg\max_s \ \log \mathcal{N}(f(\boldsymbol{W}s); \sigma_x^2 \boldsymbol{I}) + \log \mathcal{N}(0; \boldsymbol{I}) \tag{8}$$

$$= \arg\min_s \ \frac{1}{\sigma_x^2} \|x - f(\boldsymbol{W}s)\|_2^2 + \|s\|_2^2 \tag{9}$$

Equation 9 is the predictive coding objective for the simple generative model in Figure S1. The first term in the objective is a reconstruction loss (or prediction error) and the second term is a regularization term which ensures that the inferred state $s^\star$ is consistent with the prior over s. To learn the parameters $\boldsymbol{W}$, the same objective is minimized with respect to $\boldsymbol{W}$ while fixing the inferred state s.

The predictive coding formulation described above attempts to find a point estimate $s^\star$ which maximizes $p(s|x)$. An alternative is to find the full posterior distribution

$$p(s|x) = \frac{p(x|s)p(s)}{p(x)} = \frac{p(x|s)p(s)}{\int_s p(x, s) ds} \tag{10}$$

This is intractable to do exactly since it requires evaluating an integral over the continuous space of hidden states. Variational inference Jordan et al. (1999) allows us to approximate the posterior $p(s|x)$

with some *variational* posterior $q(s)$. Given a family of distributions $\mathcal{Q}$ defined over the space of hidden states s, we aim to find the distribution $q(s) \in \mathcal{Q}$ which minimizes the objective

$$D_{KL}\Big(q(s)||p(s|x)\Big) = \mathbb{E}_{s\sim q(s)}\Big[\log \frac{q(s)}{p(s|x)}\Big] \tag{11}$$

$$= \mathbb{E}_{s\sim q(s)}\Big[\log \frac{q(s)p(x)}{p(x,s)}\Big] \tag{12}$$

$$= \log p(x) - \mathbb{E}_{s\sim q(s)}\Big[\log \frac{p(x,s)}{q(s)}\Big] \tag{13}$$

The quantity $\mathbb{E}_{s\sim q(s)}\Big[\log \frac{p(x,s)}{q(s)}\Big]$ is known as the evidence lower bound (ELBO) because, due to the non-negativity of the KL divergence, it forms a lower bound on the log evidence

$$\log p(x) \geq \mathbb{E}_{s\sim q(s)}\Big[\log \frac{p(x,s)}{q(s)}\Big] = \mathbb{E}_{s\sim q(s)}[\log p(x|s)] - D_{KL}\Big(q(s)||p(s)\Big) \tag{14}$$

Therefore, minimizing the KL term $D_{KL}\Big(q(s)||p(s|x)\Big)$ amounts to maximizing the ELBO since $\log p(x)$ does not depend on either s or the parameters of the model. When we factor in the assumption in Equation 6, the ELBO reduces to the objective

$$-\mathbb{E}_{s\sim q(s)}\Big[\|x - f(\boldsymbol{W}s)\|_2^2\Big] - D_{KL}\Big(q(s)||p(s)\Big) \tag{15}$$

To see the correspondence between the ELBO in Equation 15 and the predictive coding objective in Equation 9, note that the first term in Equation 15 leads to the minimization of the reconstruction error, while the second term constrains the deviation of the posterior $q(s)$ from the prior $p(s)$. Finally, We can train neural networks to optimize the ELBO in equation 15 using the framework of Variational Autoencoders Kingma and Welling (2013).

B Derivation of the ELBO for the perception model

The generative graphical model in Figure 1b admits the following factorization of the joint likelihood

$$p(x_{1:T}, z_{1:T}, s|l_{0:T-1}) = p(x_{1:T}|z_{1:T})p(z_{1:T}|l_{0:T-1}, s)p(s) \tag{16}$$

$$= p(s)\prod_{t=1}^{T} p(x_t|z_t)\prod_{t=1}^{T} p(z_t|l_{t-1}, s) \tag{17}$$

Similarly, the joint posterior factorizes as follows

$$q(z_{1:T}, s|x_{1:T}, l_{0:T-1}) = q_1(z_{1:T}|x_{1:T}, l_{0:T-1})q_2(s|z_{1:T}, x_{1:T}, l_{0:T-1}) \tag{18}$$

$$= q_1(z_{1:T}|x_{1:T}, l_{0:T-1})q_2(s|z_{1:T}, l_{0:T-1}) \tag{19}$$

$$= q_2(s|z_{1:T}, l_{0:T-1})\prod_{t=1}^{T} q_1(z_t|x_t, l_{t-1}), \tag{20}$$

We can, therefore, express the log likelihood and posterior probabilities as

$$\log p(x_{1:T}, z_{1:T}, s|l_{0:T-1}) = \sum_{t=1}^{T}\log p(x_t|z_t) + \sum_{t=1}^{T}\log p(z_t|l_{t-1}, s) + \log p(s) \tag{21}$$

$$\log q(z_{1:T}, s|x_{1:T}, l_{0:T-1}) = \log q_2(s|z_{1:T}, l_{0:T-1}) + \sum_{t=1}^{T}\log q_1(z_t|x_t, l_{t-1}) \tag{22}$$

Using the log joint likelihood in Equation 21 and the log posterior in Equation 22, we can obtain the ELBO on the log marginal likelihood as follows

$$\log p(x_{1:T}|l_{0:T-1}) = \log \mathbb{E}_q\Big[\frac{p(x_{1:T}, z_{1:T}, s|l_{0:T-1})}{q(s, z_{1:T}|x_{1:T}, l_{0:T-1})}\Big] \tag{23}$$

$$\geq \mathbb{E}_q\Big[\log p(x_{1:T}, z_{1:T}, s|l_{0:T-1}) - \log q(s, z_{1:T}|x_{1:T}, l_{0:T-1})\Big] \tag{24}$$

$$\begin{aligned} &= \mathbb{E}_q\Big[\sum_{t=1}^{T}\log p(x_t|z_t) + \sum_{t=1}^{T}\log p(z_t|l_{t-1}, s) + \log p(s) \\ &\qquad - \sum_{t=1}^{T}\log q_1(z_t|x_t, l_{t-1}) - \log q_2(s|z_{1:T}, l_{0:T-1})\Big] \end{aligned} \tag{25}$$

$$\begin{aligned} &= \sum_{t=1}^{T}\mathbb{E}_q\Big[\log p(x_t|z_t)\Big] + \sum_{t=1}^{T}\mathbb{E}_q\Big[\log\frac{p(z_t|l_{t-1}, s)}{q_1(z_t|x_t, l_{t-1})}\Big] \\ &\qquad + \mathbb{E}_q\Big[\log\frac{p(s)}{q_2(s|z_{1:T}, l_{0:T-1})}\Big], \end{aligned} \tag{26}$$

where Equation 24 follows from Jensen's inequality.

C Network Architecture for the Active Vision Perception Model

Lower-level VAE: At each time step t, the model receives an observation x_t which, together with the corresponding location l_{t-1}, is passed through an encoder network to infer the posterior over sensory representations $q_1(z_t)$. Let $f^{(1)}_{enc,\,\mu}(x_t, l_{t-1})$ and $f^{(1)}_{enc,\,\sigma}(x_t, l_{t-1})$ denote the outputs of the lower-level encoder network at time t. In our experiments, we assume $q_1(z_t)$ is an isotropic Gaussian $\mathcal{N}(z_t|\mu^z_t, \sigma^z_t I)$, where $\mu^z_t = f^{(1)}_{enc,\,\mu}(x_t, l_{t-1})$, and $\sigma^z_t = \exp\left(\frac{1}{2} f^{(1)}_{enc,\,\sigma}(x_t, l_{t-1})\right)$.

The lower-level VAE decoder network takes sensory representations z_t and outputs the likelihood distribution $p(x_t|z_t)$. We assume this distribution is Gaussian $\mathcal{N}(x_t|\hat{x}_t, I)$, where $\hat{x}_t$ is the output of the decoder.

Higher-level VAE: At the end of a fixation sequence, the higher-level encoder network receives the sum of past sensory representations and uses it to infer the posterior over abstract representations $q_2(s)$. Similar to $q_1(z_t)$, we assume $q_2(s)$ is an isotropic Gaussian $\mathcal{N}(s|\mu^s, \sigma^s I)$ parameterized by the output of the higher-level encoder $f^{(2)}_{enc}(h_T)$. The decoder network at this level receives an abstract representation s and a query location l_{t-1}, and predicts a distribution over the corresponding lower-level representations $p(z_t|s, l_{t-1})$.

Generative Mechanism We can generate new data from the model as follows. First, we sample an abstract representation s from a standard Gaussian distribution. Then, we pick a query location l' from the interval $[-1, 1] \times [-1, 1]$. Then, we pass s and l' through the higher-level decoder $f^{(2)}_{dec}$ which outputs a distribution $p(z')$. We sample z' from this distribution and pass it through the lower-level decoder $f^{(1)}_{dec}$ which outputs an observation x' that is the same size as the model's retina.

D Survey of related work

Intrinsic motivation Our approach can be regarded as an intrinsically-motivated exploration strategy Barto (2013). In intrinsically-motivated exploration, an agent learns exploratory behavior in the absence of any extrinsic reward signals. Instead of extrinsic reward, exploration is guided by intrinsic value, which in our case is based on the expected uncertainty reduction associated with an action. The uncertainty is measured with respect to the agent's perception model, which is learned in a completely unsupervised manner. Other types of intrinsic signals have been used for autonomous exploration, such as prediction error Schmidhuber (1991); Pathak et al. (2017), space coverage Hazan et al. (2018); Amin et al. (2020), and visitation count Ménard et al. (2020). Intrinsic motivation strategies offer the advantage of representations that generalize to different tasks in the same environment since there is no dependence on a specific reward function. The closest family of intrinsic motivation approaches to ours are information-theoretic approaches, discussed below.

Information-theoretic exploration in reinforcement learning Information gain has been used to promote autonomous exploration in multiple approaches such as Storck et al. (1995); Sun et al. (2011); Still and Precup (2012). However, these approaches rely on state-action enumeration to compute information gain, which limits their applicability to settings with discrete state and action spaces. In contrast, our framework is general and can be applied to both discrete and continuous settings. In the discrete setting, the most similar work to ours is Little and Sommer (2013); we test our model in a maze navigation task similar to the one used there. The main difference between their approach and ours is that we do not assume explicit knowledge about the true generative model of the environment. Instead, the perception component of our architecture learns a generative model through collected experiences in an end-to-end manner. In the continuous setting, our work is most similar to Houthooft et al. (2016) and Mohamed and Rezende (2015) in deep reinforcement learning (RL). Our approach is different from those two approaches in that it can be applied in model-based settings since the perception component of our model explicitly learns the transition dynamics of the environment, enabling the generation of *imagined* trajectories that can be used for model-based planning and training. In contrast, those two approaches rely on model-free methods by modifying the reward function to include an information gain component.

Active vision and visual attention in machine learning We apply our model to the task of active vision. Here, our work is related to the Recurrent Attention Model (RAM) by Mnih et al. (2014) and the DRAW model by Gregor et al. (2015), but differs from those models in four key aspects. First, the perception and action components of our model are trained in a completely unsupervised, task-independent manner. During the classification task, only one feedforward decision network (separate from the main model) is trained with the classification loss. The learned representations can then be used for arbitrary tasks: to illustrate, we use a simple feedforward decision network (separate from the main model) to achieve high performance on a downstream image classification task. Second, despite the sequential nature of this task, our model solves it using end-to-end feedforward networks, greatly reducing the amount of computation compared to the recurrent architectures used in Mnih et al. (2014) and Gregor et al. (2015). Third, in contrast to Gregor et al. (2015), our model does not assume access to the full image in the training loss function, which is consistent with the assumption of bandlimited sensing. Finally, our model makes explicit links to ideas in neuroscience enabling the testing of functional hypotheses in a modern machine learning setting.

Active inference and the free energy principle In general, the theoretical formulation of our approach is most similar to the active inference formulation in neuroscience Friston (2010a); Friston et al. (2017). However, there are two differences. First, action selection in active inference relies on minimizing a generalized Expected Free Energy (EFE), whereas we use a more specific uncertainty reduction objective geared towards exploratory behavior. Second, current implementations of active inference use enumerated trajectories to minimize EFE, which limits their applicability to discrete state and action spaces. In contrast, our approach combines perception and action into an integrated and scalable neural network model, readily applicable to diverse tasks.

Predictive coding in machine learning There is a large body of work adapting the theory of predictive coding to machine learning problems, ranging from computer vision Lotter et al. (2016); Han et al. (2018); Ororbia and Mali (2023), gradient-based optimization Salvatori et al. (2021); Millidge et al. (2022), lifelong learning Ororbia et al. (2019), and temporal learning Ororbia et al. (2020). However, these models apply the theory in the context of passive perception. Although some

recent work combines predictive coding models with action Ororbia and Mali (2021), they do not focus on autonomous exploration.

E Experiment settings and hyperparameters for active vision

E.1 MNIST classification

We trained our model on the regular MNIST dataset, the translated MNIST dataset (described in the main text), and the fashion MNIST dataset. First, the perception model was pre-trained in a completely unsupervised manner with randomly selected fixation locations. Afterwards, we continued to train the perception model with the unsupervised loss while actions were selected using our BAS strategy. At the same time, a separate decision network was trained to take as input the inferred state s at the end of trial output a class label at the end of the trial. The gradients from the classification loss were only used to update the parameters of the decision network. For all experiments, the encoder and decoder networks of both the the lower-level and the higher-level VAEs were feedforward networks with two layers, each with 256 hidden units followed by rectified linear unit (ReLU) activation functions. The action network was a two-layer feedforward network with 64 and 32 hidden units each. When perception states were used for decision making, the decision network was a two-layer feedforward network with 256 hidden units each. When an RNN was used to integrate past observations for decision making, the hidden size of the RNN decision network was chosen to be the same as the dimensionality of the abstract state s. Table 1 lists the hyperparameters used for each type of experiment. Hyperparameters were adjust ad hoc based on the resulting accuracy obtained on a separate validation set. In these experiments, we also use a regularization hyperparameter β as a scalar multiplying the KL term in the ELBO objective for the perception model Higgins et al. (2016). Algorithm 1 describes the pseudocode for Bayesian Action Selection in active vision.

Algorithm 1 Bayesian Action Selection in Active Vision

Input: observations $x_{1:t}$, locations $l_{0:t-1}$, perception model F, action network ψ_t, number of MC samples K
$q_t(s) = F.Encode(x_{1:t}, l_{0:t-1})$
$l_t = \psi_t(\mathbb{E}[q_t(s)])$
$p(x_{t+1}) = F.Decode(q_t(s), l_t)$
Draw K samples from $p(x_{t+1})$
for $k = 1$ **to** K **do**
  $q^{(k)}_{t+1}(s) = F.Encode(x_{1:t}, x^{(k)}_{t+1}, l_{0:t-1})$
end for
$\tilde{V}(l_t) = H(q_t(s)) - \frac{1}{K}\sum_{k=1}^{K} H\left(q^{(k)}_{t+1}(s)\right)$
Update action network parameters using gradient descent on $\tilde{V}$: $\psi_{t+1} = \psi_t + \mu\nabla_{\psi_t}\tilde{V}(l_t)$
return: selected action l_t and updated action network ψ_{t+1}

Table 1: Settings for Centered, Translated, and Fashion MNIST Experiments

Hyper-parameter	Centered MNIST	Translated MNIST	Fashion MNIST
# pre-training episodes	0	10	10
# fixations (n)	3	4	5
Patch dim (d)	8	12	6
# foveated patches (N_{fov})	1	3	1
Foveation scale	—	2	—
z dim	32	64	64
s dim	64	128	128
σ_{action}	0.15	0.15	0.05
Action network lr	0.001	0.001	0.001
Perception model lr	0.001	0.001	0.001
Decision network lr	0.001	0.001	0.001
β	0.1	0.1	1.1
Batch size	64	64	64

E.2 Grayscale CIFAR-10

We tested our perception model on grayscale CIFAR-10 images to see if it can capture the overall structure and statistics of natural images. Table 2 lists the hyperparameters used for these experiments.

Table 2: Settings for Grayscale CIFAR-10 Experiments

Hyper-parameter	Setting
# fixations (n)	6
Patch dim (d)	12
# foveated patches (N_{fov})	1
z dim	32
s dim	64
Perception model lr	0.001
β	0.01
Batch size	64

E.3 Learning speed and data efficiency comparisons

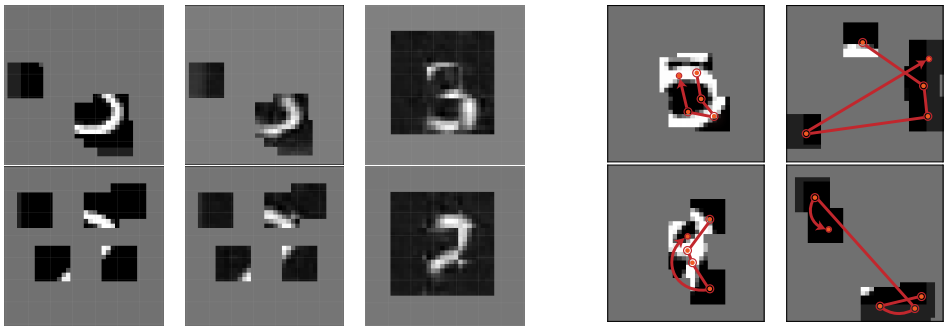
All tests reported in Section **??** were performed on the translated MNIST dataset. Our approach (BAS + FF), described in the main text, was compared to two baselines: the Recurrent Attention Model (RAM) Mnih et al. (2014), and a feedforward (FF) neural network receiving full images as input (Full Images + FF). In all three cases, the decision network (the network that outputs class labels) consisted of two hidden layers each with 128 hidden units followed by ReLU activation functions.

For BAS + FF, the perception model was pretrained unsupervised for 10 epochs with a random action selection strategy. The architectures of the perception model and the action network were the same as those described in Section E.1. For RAM, the dimensionality of h_g and h_l in the glimpse network was 64 and the dimensionlaity of g was 128. The hidden size of the RNN was chosen to match the dimensionality of the abstract representation s in our perception model, which was 128. The location network had one hidden layer with 64 hidden units. Similar to our model, the location was drawn from a two-component Gaussian (with a pre-determined fixed variance) parameterized by the output of the location network. Hyperparameters for both RAM and BAS were adjusted ad hoc to optimize performance on a validation set (separate from the MNIST test set). Those hyperparameters are listed in Table 3. In the Full Images + FF case, the only hyperparameters adjusted were the batch size and the learning rate which were fixed at 64 and 0.001, respectively.

Table 3: Hyperparameters for learning speed and data efficiency tests

Hyper-parameter	RAM	BAS + FF
# fixations (n)	3	3
Patch dim (d)	12	12
# foveated patches (N_{fov})	3	3
Foveation scale	2	2
z dim	—	64
s dim	—	128
h_g	64	—
h_l	64	—
RNN hidden size	128	—
σ_{action}	0.05	0.15
Action network lr	0.001	0.001
Perception model lr	—	0.001
Decision network lr	0.001	0.001
Core and glimpse networks lr	0.001	—
β	—	0.1
Batch size	64	64

F Supplementary figures



(a) (b)

Figure S2: (a) More examples showing the generative ability of the perception model, similar to Figure 2. (b) More examples of BAS versus random fixation sequences.

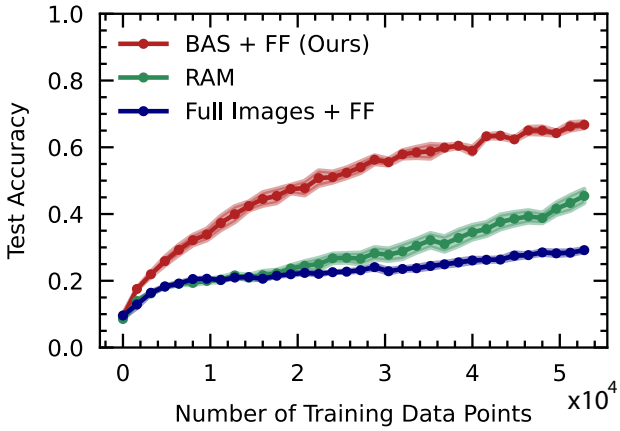
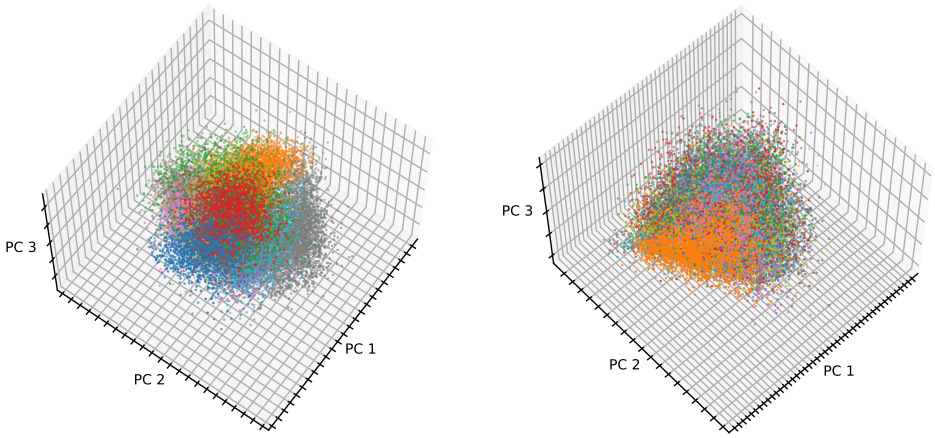
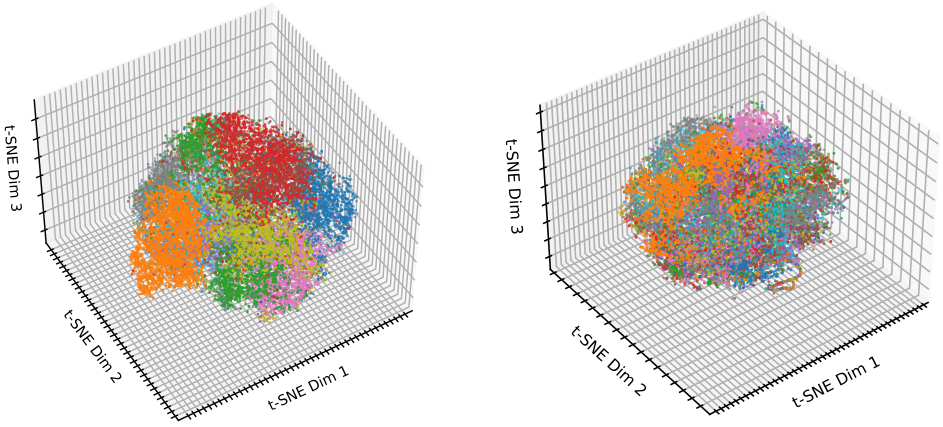


Figure S3: Data efficiency and generalization. We compare our method versus RAM and Full Images + FF in terms of their test performance during the first episode of supervised training, when the classification networks see the data for the first time. Each point on the plot represents the test score after observing only x training data points for the first time. Shaded error bars represent the SEM (n = 5 random seeds).



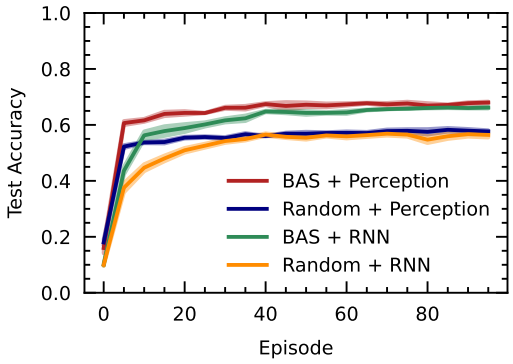
(a) (b)

Figure S4: PCA projections of the latent representations learned by (a) BAS-trained perception model, and (b) the randomly-trained perception model. Each point in the PC space correspond the projection of the inferred state s for a given input image. Points are colored based on the class of their corresponding input images.

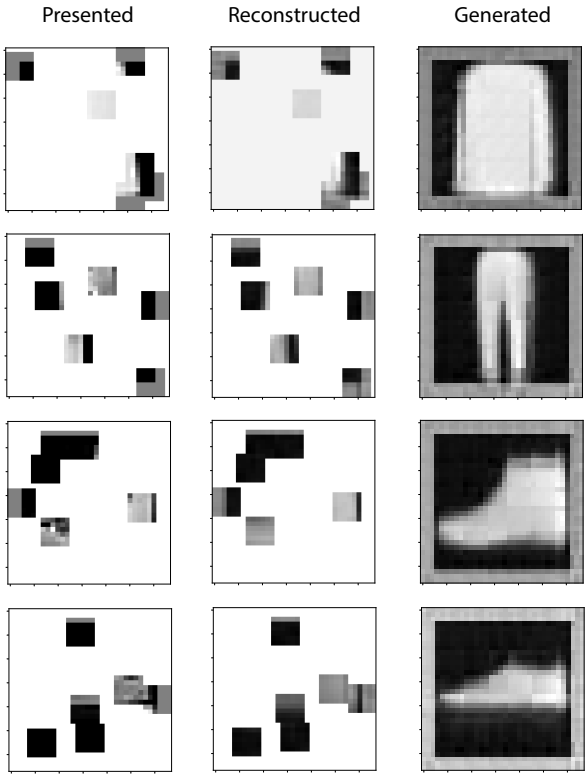


(a) (b)

Figure S5: Same as in Figure S4 but using t-SNE Van der Maaten and Hinton (2008) to visualize projections. (a) t-SNE projections of the BAS-trained perception model. (b) t-SNE projections for the ranomly-trained perception model.



(a)



(b)

Figure S6: Results on Fashion MNIST. (a) Classification performance on the fashion MNIST dataset during the active vision task. (b) Examples demonstrating the generative ability of the perception model. The original patches presented are shown on the left and their reconstructions are shown in the middle. Right shows images generated by first generating small patches at various locations (not seen during presentation) and combining them to form the final image. These results show the perception model is able to capture the spatial relationships associated with elements in the dataset.

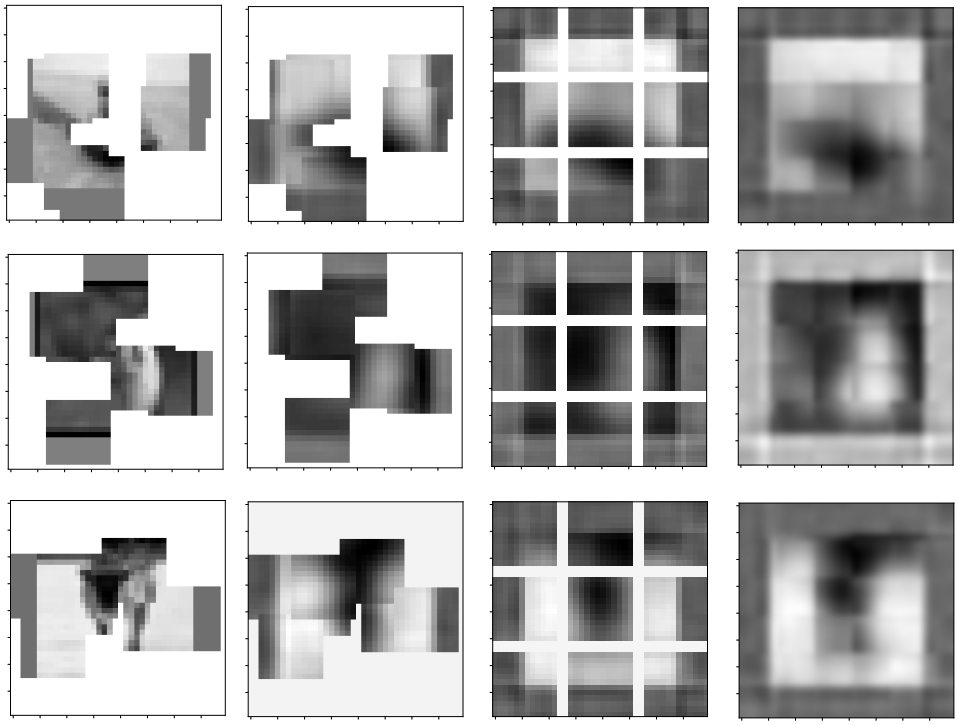


Figure S7: Results on grayscale CIFAR-10. Original presented patches are shown on the left and their reconstructions are shown on the second column. Third and fourth column show images generated by the perception model combining smaller patches at different locations. The last column has more patches at continguous locations followed by bicubic smoothing for illustration. These results show that, despite the simplicity of our model's architecture, it is still able to capture the overall structure and statistics in natural images.

References

Amin, S., Gomrokchi, M., Aboutalebi, H., Satija, H., and Precup, D. (2020). Locally persistent exploration in continuous control tasks with sparse rewards. *CoRR*, abs/2012.13658.

Barto, A. G. (2013). Intrinsic motivation and reinforcement learning. *Intrinsically motivated learning in natural and artificial systems*, pages 17–47.

Butko, N. J. and Movellan, J. R. (2008). I-pomdp: An infomax model of eye movement. In *2008 7th IEEE International Conference on Development and Learning*, pages 139–144.

Butko, N. J. and Movellan, J. R. (2010). Infomax control of eye movements. 2(2):91–107.

Friston, K. (2010a). The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138.

Friston, K., Adams, R. A., Perrinet, L., and Breakspear, M. (2012). Perceptions as hypotheses: saccades as experiments. *Frontiers in psychology*, 3:151.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2017). Active inference: a process theory. *Neural computation*, 29(1):1–49.

Friston, K. J. (2010b). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11:127–138.

Gregor, K., Danihelka, I., Graves, A., Rezende, D., and Wierstra, D. (2015). Draw: A recurrent neural network for image generation. In Bach, F. and Blei, D., editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1462–1471, Lille, France. PMLR.

Guillery, R. (2005). Anatomical pathways that link perception and action. In *Cortical Function: a View from the Thalamus*, volume 149 of *Progress in Brain Research*, pages 235–256. Elsevier.

Guillery, R. and Sherman, S. M. (2011). Branched thalamic afferents: What are the messages that they relay to the cortex? *Brain Research Reviews*, 66(1):205–219. Camillo Golgi and Modern Neuroscience.

Han, K., Wen, H., Zhang, Y., Fu, D., Culurciello, E., and Liu, Z. (2018). Deep predictive coding network with local recurrent processing for object recognition. *CoRR*, abs/1805.07526.

Hayhoe, M. and Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4):188–194.

Hazan, E., Kakade, S. M., Singh, K., and Soest, A. V. (2018). Provably efficient maximum entropy exploration. *CoRR*, abs/1812.02690.

Higgins, I., Matthey, L., Pal, A., Burgess, C. P., Glorot, X., Botvinick, M. M., Mohamed, S., and Lerchner, A. (2016). beta-vae: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*.

Houthooft, R., Chen, X., Duan, Y., Schulman, J., Turck, F. D., and Abbeel, P. (2016). Curiosity-driven exploration in deep reinforcement learning via bayesian neural networks. *CoRR*, abs/1605.09674.

Jiang, L. P. and Rao, R. P. N. (2021). Predictive coding theories of cortical function.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., and Saul, L. K. (1999). *An Introduction to Variational Methods for Graphical Models*, page 105–161. MIT Press, Cambridge, MA, USA.

Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13(10):1292—1298.

Land, M. F. and Tatler, B. W. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford University Press.

Linson, A., Clark, A., Ramamoorthy, S., and Friston, K. (2018). The active inference approach to ecological perception: General information dynamics for natural and artificial embodied cognition. *Frontiers in Robotics and AI*, 5.

Little, D. Y. and Sommer, F. T. (2013). Learning and exploration in action-perception loops. *Frontiers in neural circuits*, 7:37.

Lotter, W., Kreiman, G., and Cox, D. D. (2016). Deep predictive coding networks for video prediction and unsupervised learning. *CoRR*, abs/1605.08104.

Marino, J. (2022). Predictive Coding, Variational Autoencoders, and Biological Connections. *Neural Computation*, 34(1):1–44.

Ménard, P., Domingues, O. D., Jonsson, A., Kaufmann, E., Leurent, E., and Valko, M. (2020). Fast active learning for pure exploration in reinforcement learning. *CoRR*, abs/2007.13442.

Millidge, B., Song, Y., Salvatori, T., Lukasiewicz, T., and Bogacz, R. (2022). Backpropagation at the infinitesimal inference limit of energy-based models: Unifying predictive coding, equilibrium propagation, and contrastive hebbian learning.

Mnih, V., Heess, N., Graves, A., and kavukcuoglu, k. (2014). Recurrent models of visual attention. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc.

Mohamed, S. and Rezende, D. J. (2015). Variational information maximisation for intrinsically motivated reinforcement learning. *Advances in neural information processing systems*, 28.

Ororbia, A. and Mali, A. (2023). Convolutional neural generative coding: Scaling predictive coding to natural images.

Ororbia, A., Mali, A., Giles, C. L., and Kifer, D. (2020). Continual learning of recurrent neural networks by locally aligning distributed representations. *IEEE Transactions on Neural Networks and Learning Systems*, 31(10):4267–4278.

Ororbia, A. and Mali, A. A. (2021). Backprop-free reinforcement learning with active neural generative coding. *CoRR*, abs/2107.07046.

Ororbia, A., Mali, A. A., Kifer, D., and Giles, C. L. (2019). Lifelong neural predictive coding: Sparsity yields less forgetting when learning cumulatively. *CoRR*, abs/1905.10696.

Pathak, D., Agrawal, P., Efros, A. A., and Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. *CoRR*, abs/1705.05363.

Rao, R. P. N. and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2:79–87.

Salvatori, T., Song, Y., Lukasiewicz, T., Bogacz, R., and Xu, Z. (2021). Predictive coding can do exact backpropagation on convolutional and recurrent neural networks. *CoRR*, abs/2103.03725.

Schmidhuber, J. (1991). A possibility for implementing curiosity and boredom in model-building neural controllers. In *Proc. of the international conference on simulation of adaptive behavior: From animals to animats*, pages 222–227.

Sekar, R., Rybkin, O., Daniilidis, K., Abbeel, P., Hafner, D., and Pathak, D. (2020). Planning to explore via self-supervised world models. In *ICML*.

Still, S. and Precup, D. (2012). An information-theoretic approach to curiosity-driven reinforcement learning. *Theory in Biosciences*, 131:139–148.

Storck, J., Hochreiter, S., Schmidhuber, J., et al. (1995). Reinforcement driven information acquisition in non-deterministic environments. In *Proceedings of the international conference on artificial neural networks, Paris*, volume 2, pages 159–164.

Sun, Y., Gomez, F. J., and Schmidhuber, J. (2011). Planning to be surprised: Optimal bayesian exploration in dynamic environments. *CoRR*, abs/1103.5708.

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).

Yarbus, A. L. (1967). *Eye movements and Vision*. Plenum Press.