A computational approach to visual ecology with deep reinforcement learning

Sacha Sokoloski[1,2], and Philipp Berens[1,2]

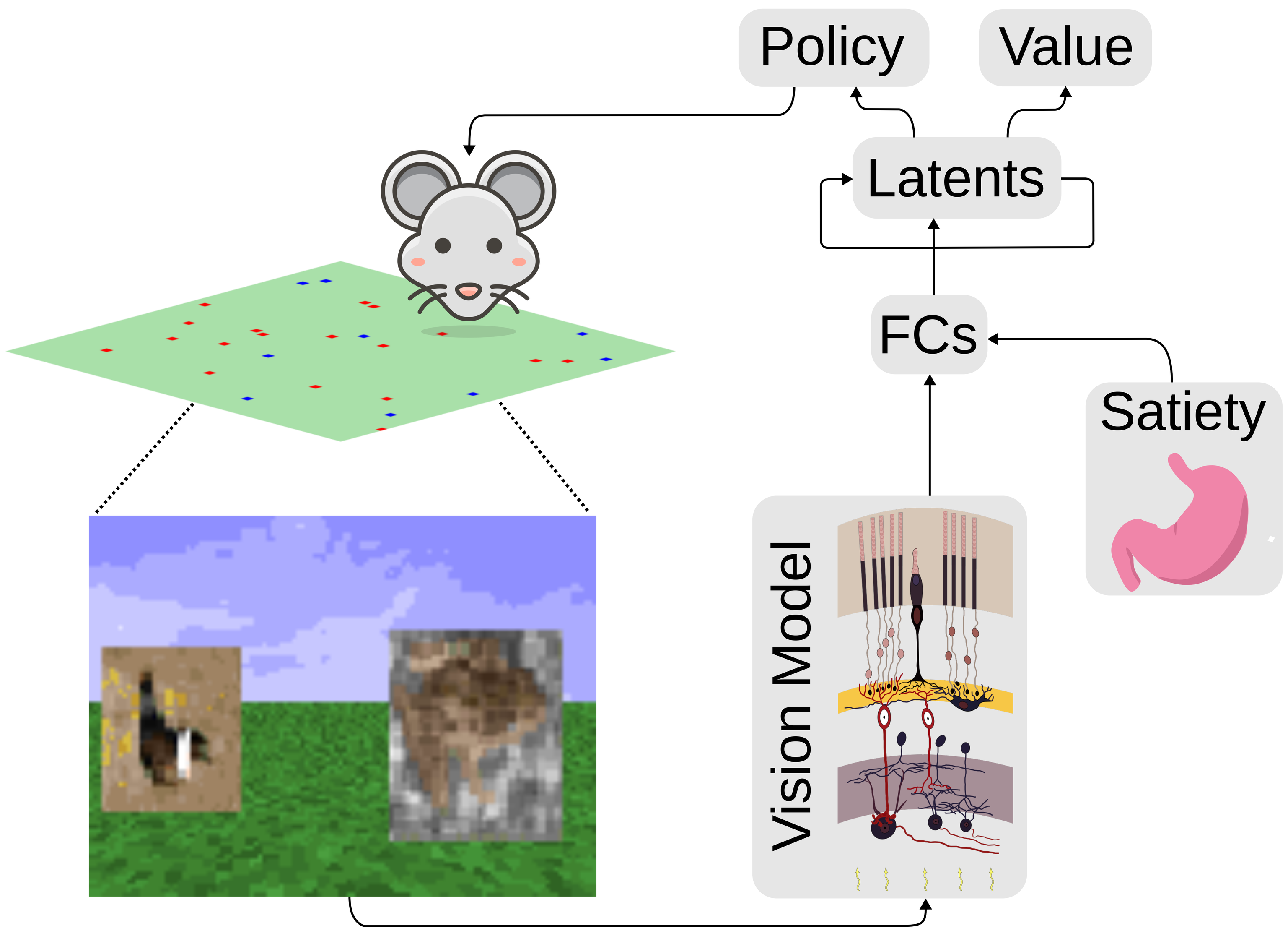
1. Hertie Institute for AI in Brain Health, Tübingen; 2. University of Tübingen

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Embedding vision models in RL environments

To study how ecology constrains visual processing, we present a reinforcement learning (RL) framework in which an agent aims to survive in a 3-d environment that it perceives through a vision model. Our code is built on

- **PyTorch**, the DNN library we use to implement vision models,
- **Sample Factory** [2], a deep RL library that we use to train agents, and
- **ViZDoom** [3], an RL environment engine based on *Doom*.



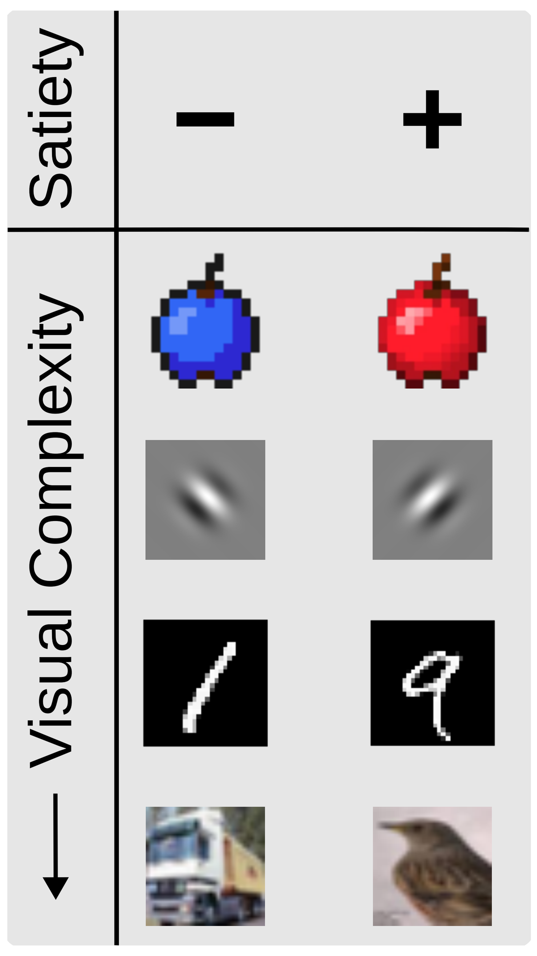
Varying the visual complexity of a task

The agent has "satiety" from 0 to 100, and continuously loses satiety. To survive and thrive, it must solve a visually-guided foraging task.

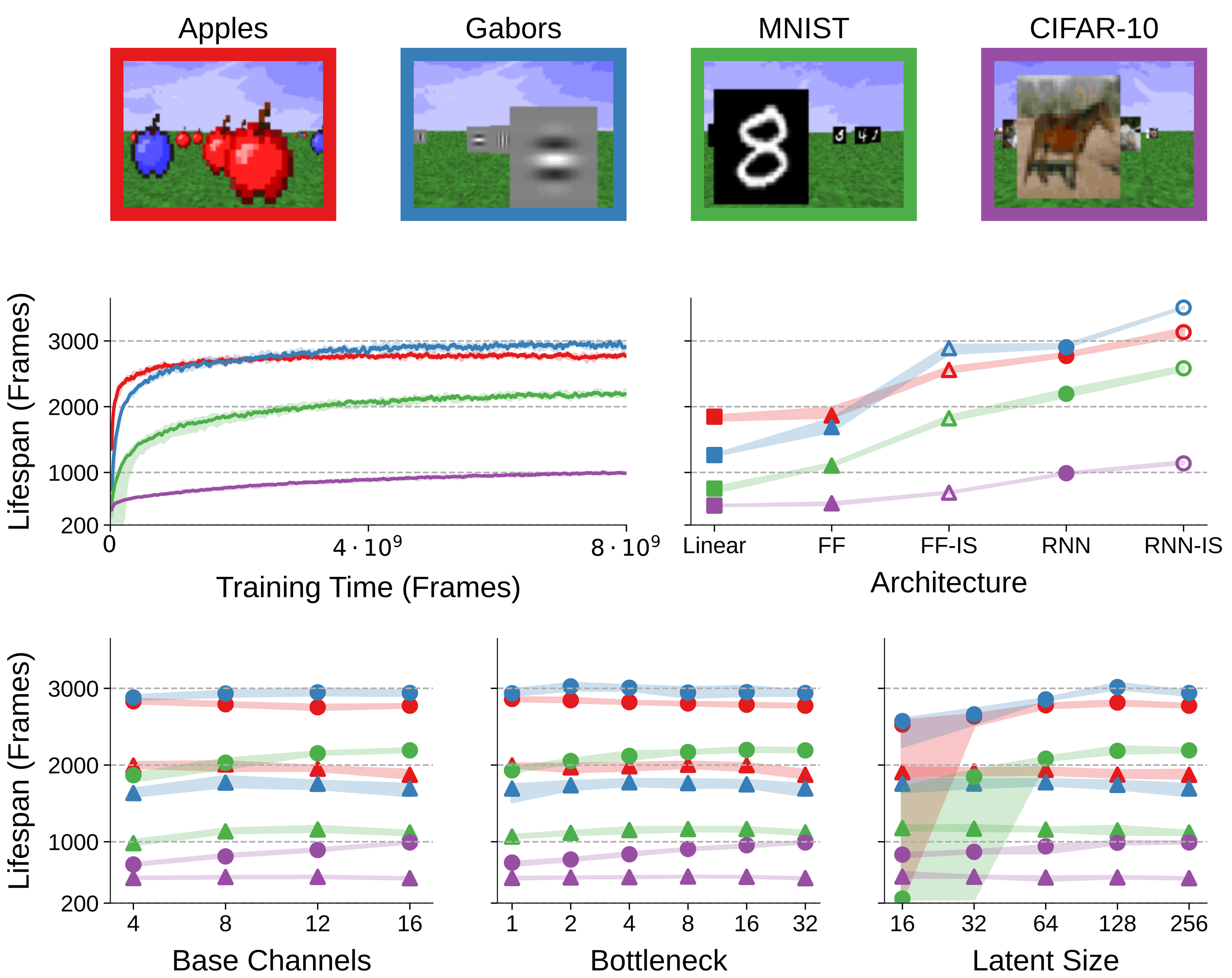
- Nourishment and poison are randomly distributed in the environment.
- We modulate task difficulty by varying the diversity of the nourishment/poison.

RL Primer: We describe the world in terms of states s, actions a, and rewards r.

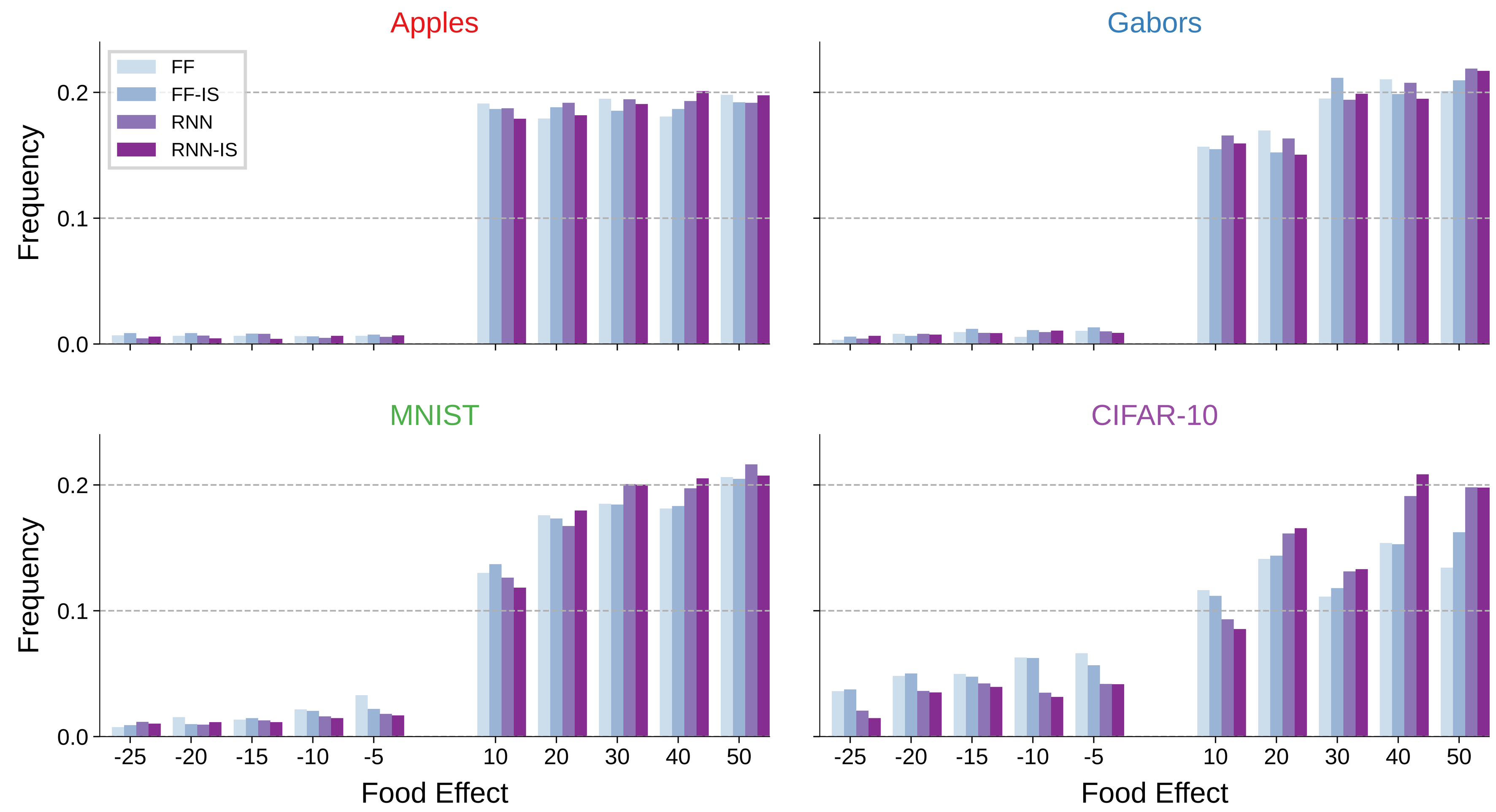
- **transition distribution:** $p(s' \mid s, a)$
- **policy:** $\pi(s) = p(a \mid s)$
- **action-value function:** $Q^\pi(s, a) = \mathbb{E}^\pi[\sum_{t=0}^{\infty} \gamma^t R_t \mid S_0 = s, A_0 = a]$
- **policy gradient theorem:** $\frac{\partial}{\partial\theta}\mathbb{E}^\pi[\sum_{t=0}^{\infty} \gamma^t R_t] = \int_S \mu^\pi(s) \int_A \frac{\partial \pi(s|a)}{\partial\theta} Q^\pi(s, a) da\, ds$



Brain complexity scales with visual complexity

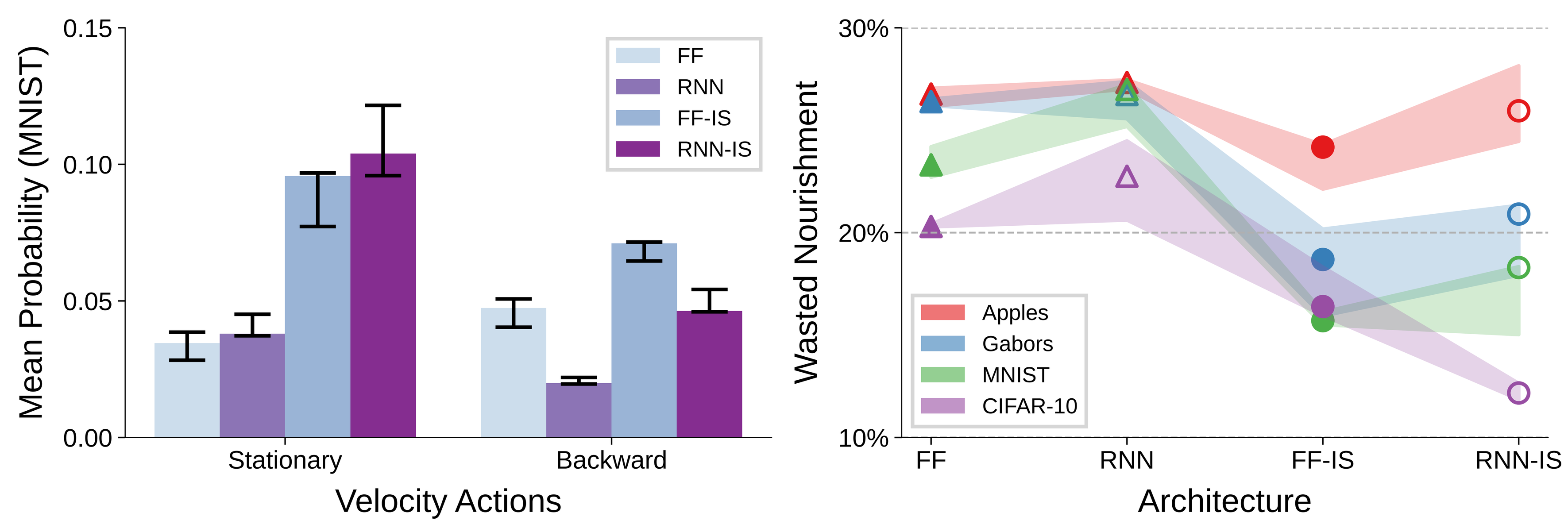


Complex recognition recruits recurrent connectivity



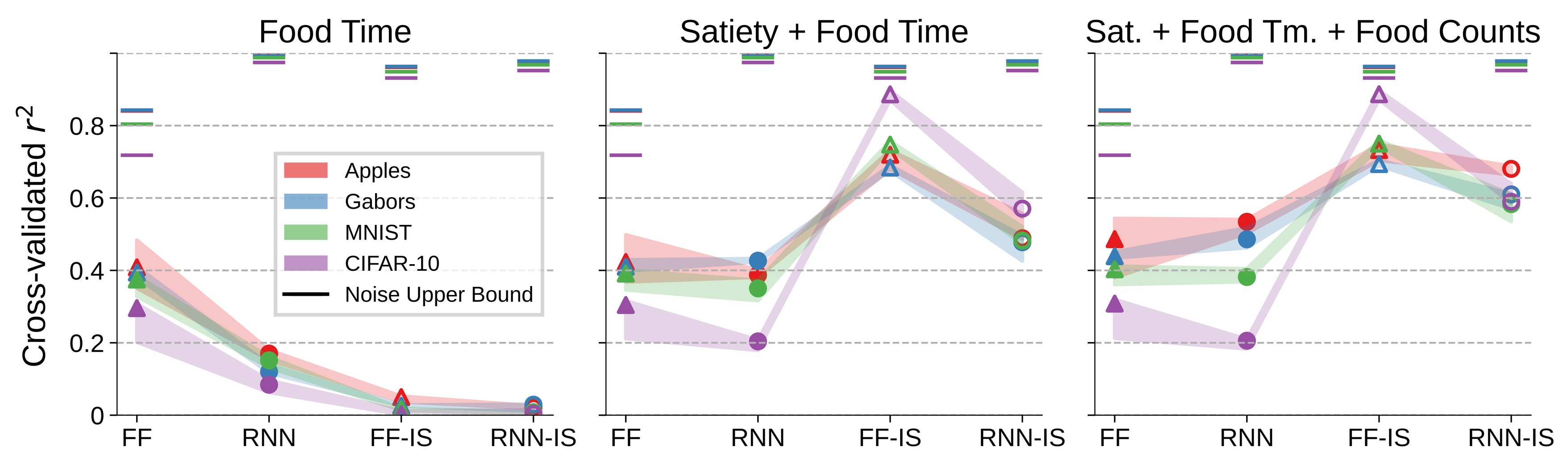
- Histogram of food pickup frequency shows how good the agent is at distinguishing foods.

Satiety signals enable unique behaviours



- Satiety signals allow agents to strategize around nourishment magnitude.
- Wasted nourishment occurs when nourishments yields 100 > 0.

Brain architecture drives different representations



- Linear regression on the estimated value function $\hat{V}$ shows the importance of environmental variables to its representation of the world.
- $V^\pi(s) = \mathbb{E}^\pi[\sum_{t=0}^{\infty} \gamma^t R_t \mid S_0 = s]$.

Conclusion and Outlook

We have lain foundations for modelling ecological constraints on vision and shown:

1. how to bridge established results from DNN vision models [1] into a dynamic RL setting,
2. that rich behaviours and representations emerge even on simple tasks, and
3. that brain architecture and visual complexity interact in nontrivial ways.

Building on this, our next goals are to explore:

1. more complex environmental objects such as predators,
2. adding metabolic costs for the magnitude of network activity to induce "implicit" efficient coding, and
3. modelling tradeoffs between minimizing image reconstruction error and task-relevant efficient codes.

References

[1] Jack Lindsey et al. "A Unified Theory of Early Visual Representations from Retina to Cortex through Anatomically Constrained Deep CNNs". International Conference on Learning Representations. 2018.

[2] Aleksei Petrenko et al. "Sample Factory: Egocentric 3D Control from Pixels at 100000 FPS with Asynchronous Reinforcement Learning". International Conference on Machine Learning. 2020.

[3] Christopher Schulze et al. "ViZDoom: DRQN with Prioritized Experience Replay, Double-Q Learning and Snapshot Ensembling". Advances in Intelligent Systems and Computing. Cham, 2019.

HERTIE INSTITUTE FOR AI IN BRAIN HEALTH

Funding

This work has been funded by the DFG CRC 1233 "Robust Vision". We thank the Hertie Stiftung.

Contact

philipp.berens@uni-tuebingen.de
www.hertie.ai/data-science

GEMEINNÜTZIGE Hertie Stiftung