Attention Guided Imitation Learning and Reinforcement Learning

The University of Texas at Austin

Ruohan Zhang (zharu@utexas.edu)

Doctoral Consortium-127

Abstract

When intelligent agents learn visuomotor behaviors from human demonstrations, they may benefit from knowing where the human is allocating visual attention, which can be inferred from their gaze. A wealth of information regarding intelligent decision making is conveyed by human gaze allocation; hence, exploiting such information has the potential to improve the agents' performance. With this motivation, we propose the AGIL (Attention Guided Imitation Learning) framework. We collect high-quality human action and gaze data while playing Atari games, driving in a virtual city, and walking outdoor. Using these data, we first train a deep neural network that can predict human gaze positions with high accuracy (the gaze network) and then train another network to predict human actions (the policy network). Incorporating the learned attention model from the gaze network into the policy network significantly improves the task performance. Current work involves using attention learned from human to facilitate the training process of deep reinforcement learning algorithms, as well as understanding attention mechanism in the biological nervous systems.

Keywords: Attention; Gaze; Imitation Learning; Reinforcement Learning; Neural Coding

Introduction

- Imitation learning: extracting initial biases as well as strategies how to approach a visuomotor task from demonstrations of humans [Schaal, 1999]
- Deep imitation learning: using a deep neural network to extract such knowledge
- One concern: The sensory system of a human demonstrator is different from a machine's
  - Humans have foveal vision with high acuity for only 1-2 visual degrees

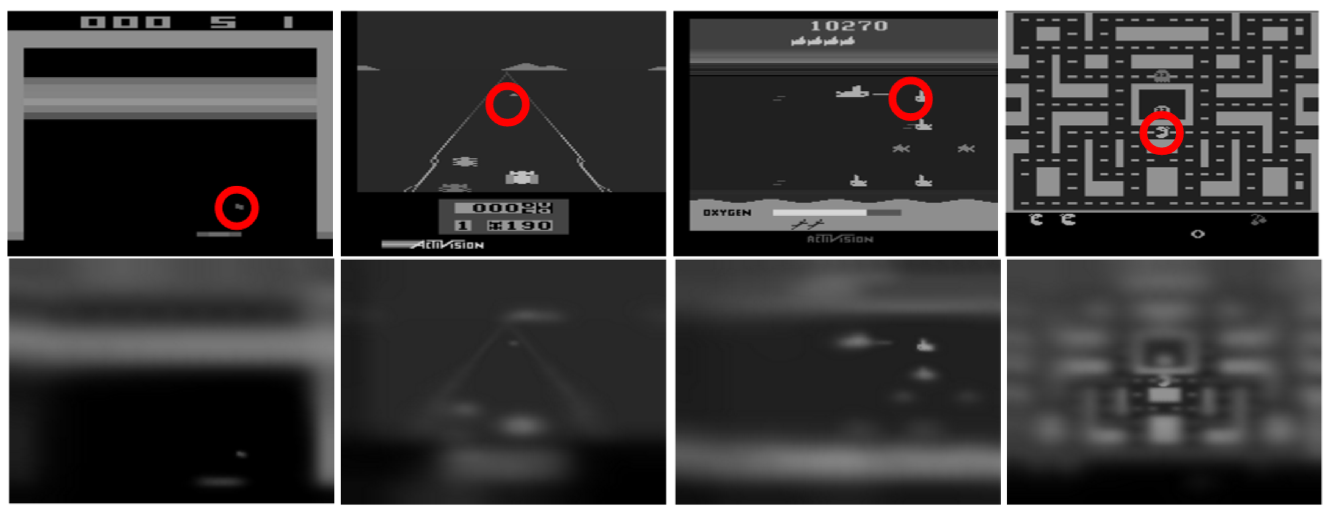
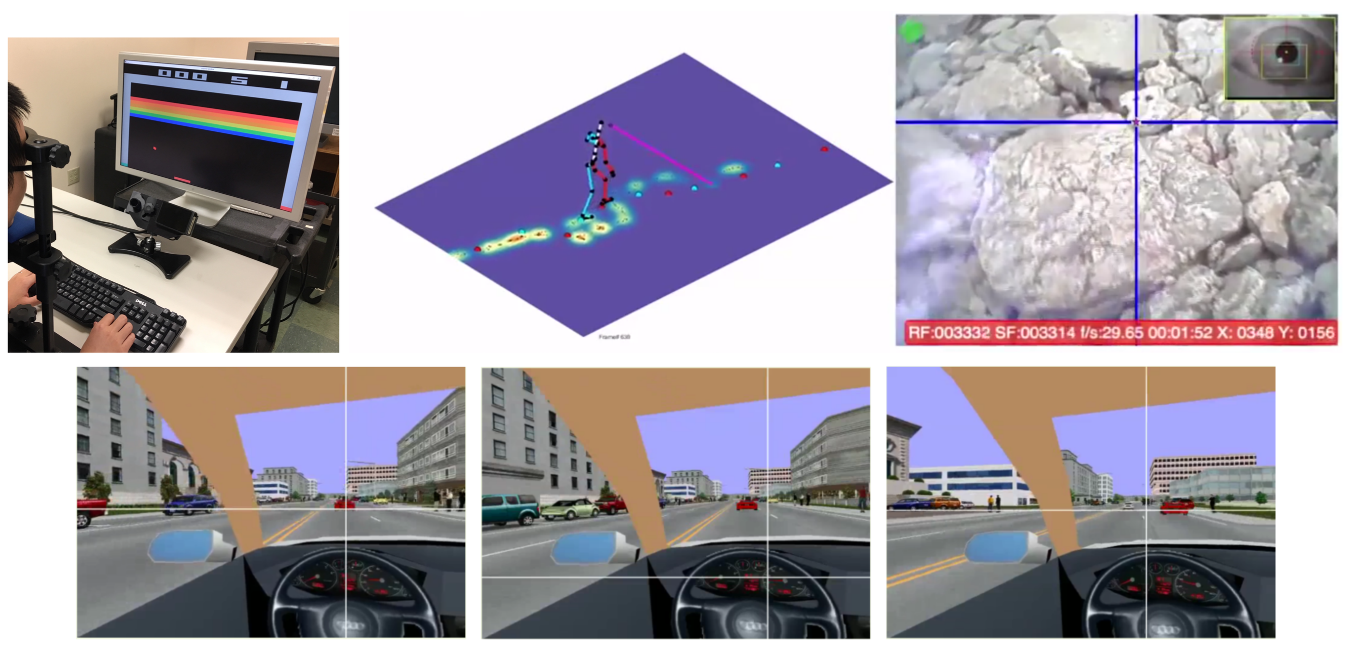


Figure 1: Foveal vision. Red circles indicate gaze positions. The images are generated using the foveated rendering algorithm [Perry and Geisler, 2002]

- Foveal vision leads to eye movements/gaze behaviors to solve the POMDP problem
- For visuomotor tasks, gaze indicates visual features that matter for the current decision – a very strong clue why that decision was made
- Approach: learn a visual attention model from human gaze data, then use the learned model to guide the process of learning the actions

Data Collection: Gaze + Action

- 20 Atari games using Arcade Learning Environment [Bellemare et al., 2012, Zhang et al., 2018]
- Virtual urban driving in a virtual environment with Logitech driving platform [Johnson et al., 2014]
- Outdoor walking on a rough terrain with full-body motion capture [Matthis et al., 2018]



(a) Leading car (b) Speedometer (c) Lane

Learning to Predict Human Gaze

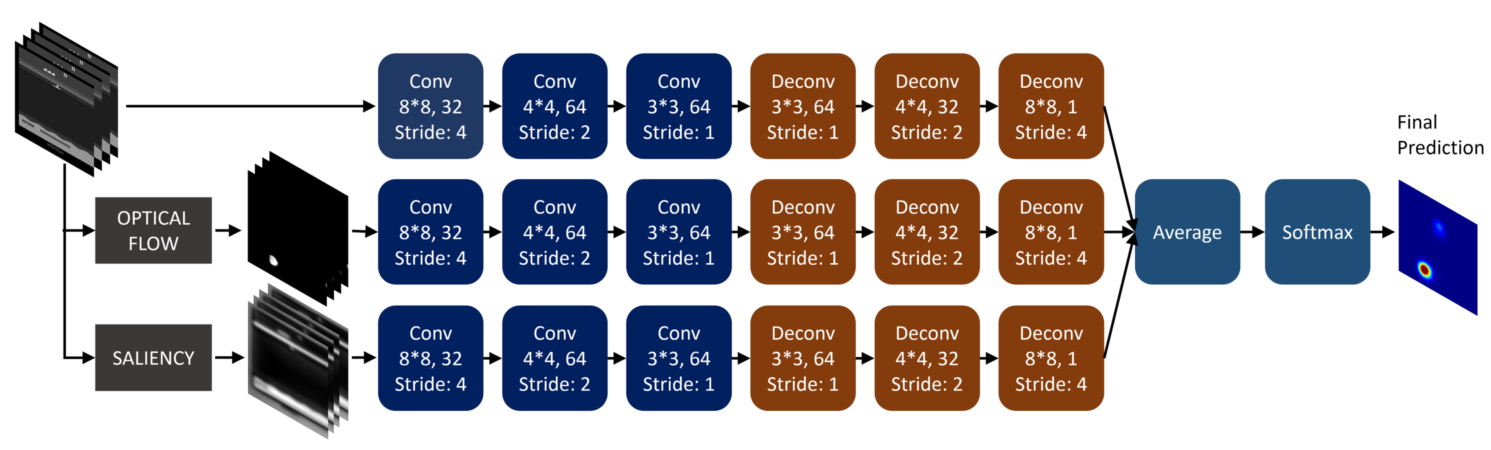


Figure 2: The three-channel gaze network [Zhang et al., 2018]. The top channel takes in images, the mid channel takes in the corresponding optical flow , and the bottom channel takes in the bottom-up image saliency. The final output is a gaze saliency map that indicates the predicted probability distribution of the gaze trained using KL divergence.

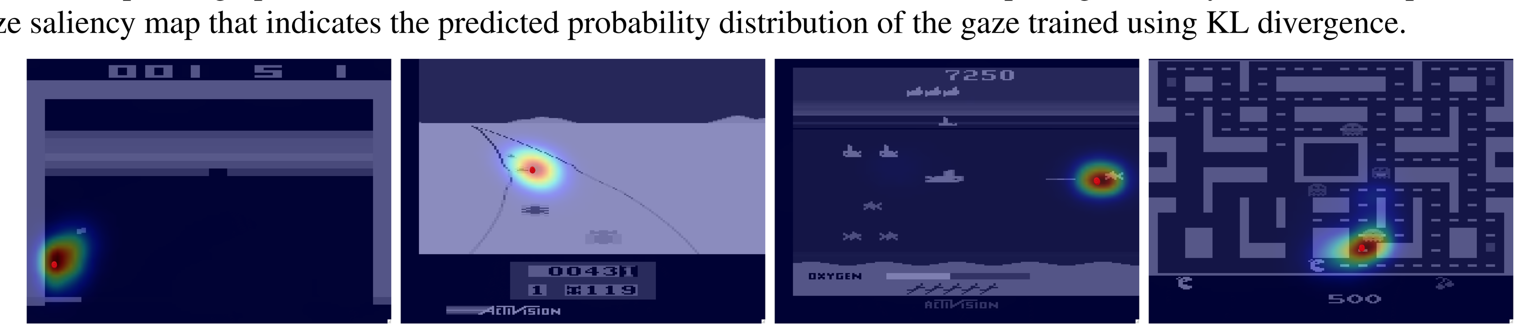


Figure 3: Visualization of gaze prediction results for eight games. The solid red dot indicates the ground truth human gaze position. The heatmap shows the model's prediction as a saliency map, computed using the gaze network. Average area under the curve (AUC) score across 8 games on testing dataset: **0.965**.

Learning to Predict Human Actions

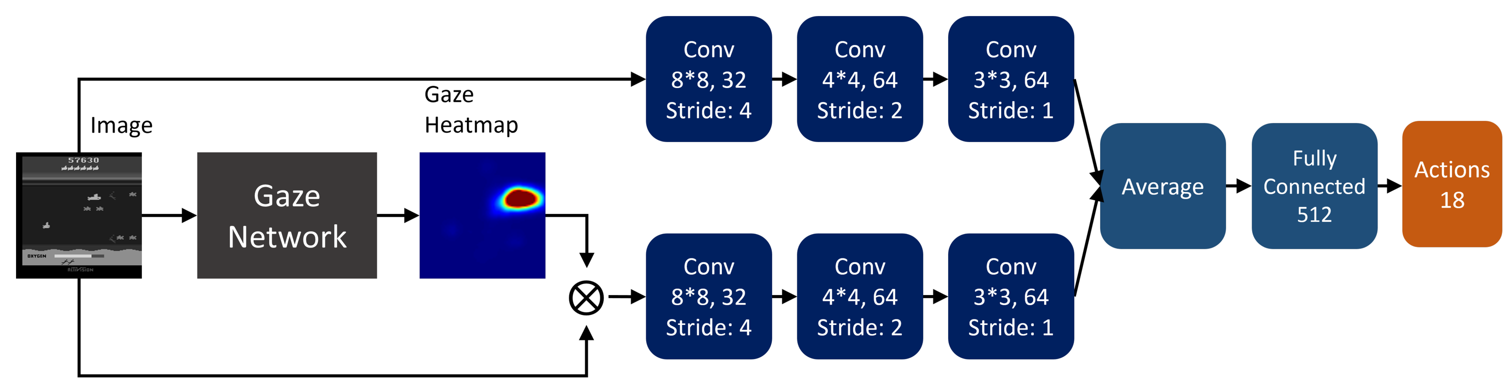


Figure 4: The policy network architecture for imitating human actions [Zhang et al., 2018]. The top channel takes in the current image frame and the bottom channel takes in the masked image which is an element-wise product of the original image and predicted gaze saliency map by the gaze network.

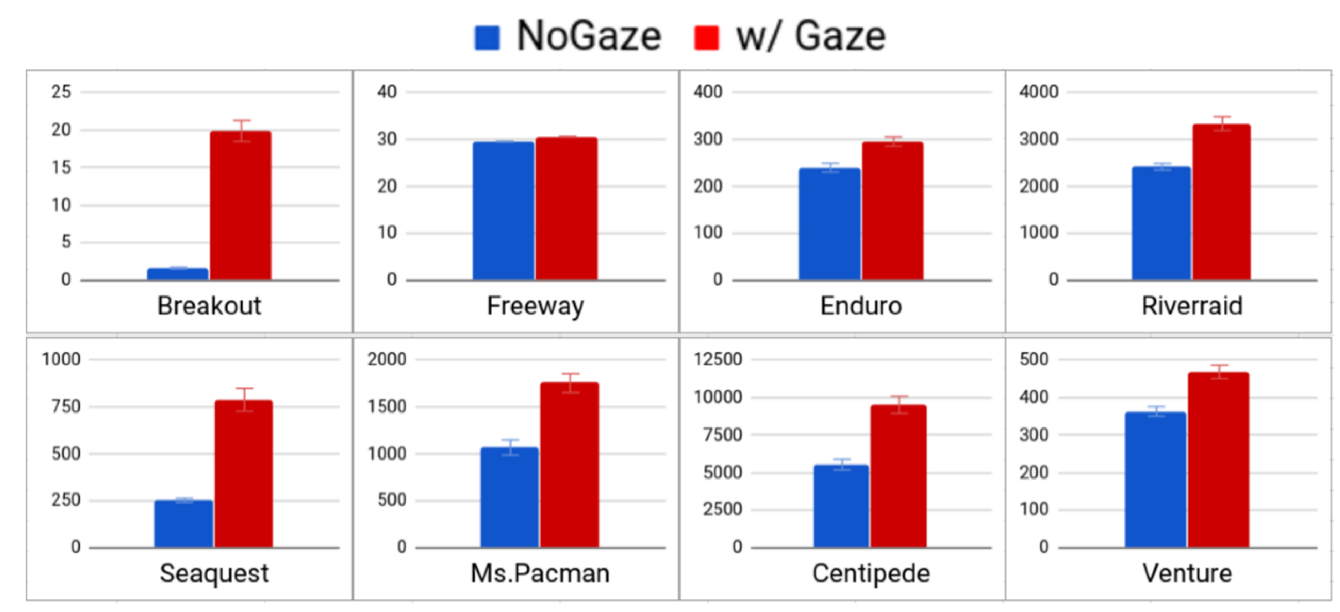


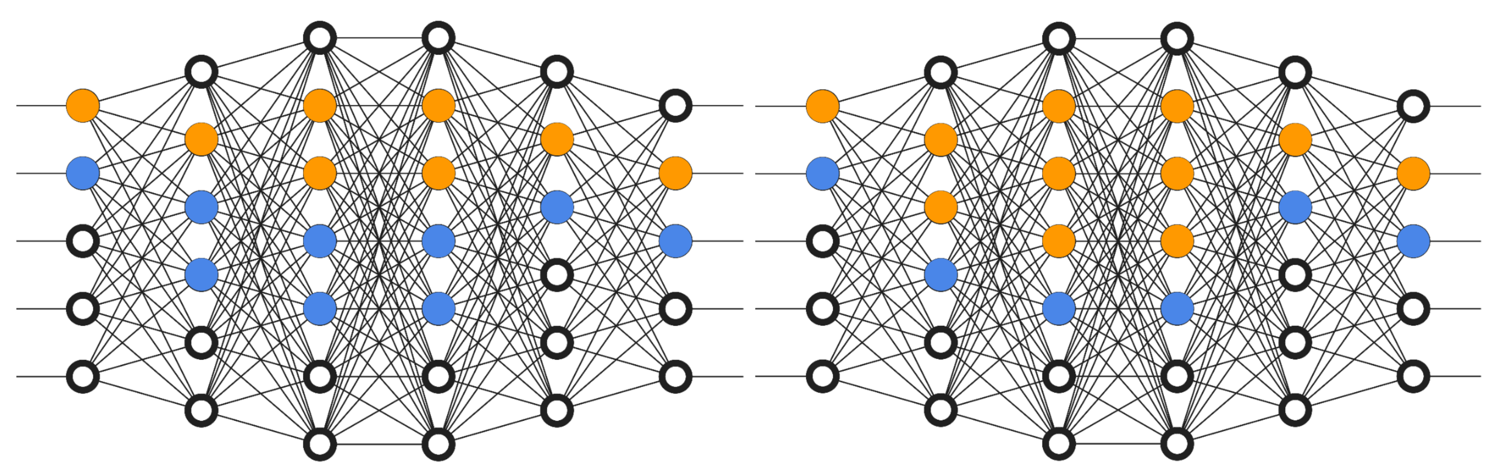
Figure 5: Game scores across 8 games. Visual attention can help the imitator learn better from human demonstration.

- Hypothetically, visual attention simplifies the learning problem by indicating the object of interest for current decision

- There is still a large performance gap between the human player and the learning agent, what else is missing from an imitation learning perspective?
- Attention guided deep reinforcement learning – applying learned attention model to speedup RL

The Neural Mechanism of Attention

- Motivation: ultimately, machines should be able to learn its own attention mechanism from experience, which requires a deeper understanding of attention mechanism itself
- A neural multiplexing model: neurons communicate like radio stations [Ballard and Zhang, 2018]
- Neurons dedicated to a computational process tuned to a particular frequency in gamma range
- Attention reflects the use of additional neurons by a computational process by switching these neurons to the frequency band associated with that process
- Initial analysis of awake mouse V1 neural recording data tends to support this theory. It could constitute an advance in understanding attention and designing artificial neural networks



Acknowledgement

The work is supported by NIH grant EY05729, NSF grant CNS-1624378 and Google AR/VR Research Award.

References

[Ballard and Zhang, 2018] Ballard, D. and Zhang, R. (2018). Cortical spike multiplexing using gamma frequency latencies. *bioRxiv*, page 313320.

[Bellemare et al., 2012] Bellemare, M. G., Naddaf, Y., Veness, J., and Bowling, M. (2012). The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research*.

[Johnson et al., 2014] Johnson, L., Sullivan, B., Hayhoe, M., and Ballard, D. (2014). Predicting human visuomotor behaviour in a driving task. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1636):20130044.

[Matthis et al., 2018] Matthis, J. S., Yates, J. L., and Hayhoe, M. M. (2018). Gaze and the control of foot placement when walking in natural terrain. *Current Biology*, 28(8):1224–1233.

[Perry and Geisler, 2002] Perry, J. S. and Geisler, W. S. (2002). Gaze-contingent real-time simulation of arbitrary visual fields. In *Electronic Imaging 2002*, pages 57–69. International Society for Optics and Photonics.

[Schaal, 1999] Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in cognitive sciences*, 3(6):233–242.

[Zhang et al., 2018] Zhang, R., Liu, Z., Zhang, L., Whritner, J. A., Muller, K. S., Hayhoe, M. M., and Ballard, D. H. (2018). Agil: Learning attention from human for visuomotor tasks. *European Conference on Computer Vision (ECCV)*.