

Attention Guided Imitation Learning and Reinforcement Learning

Ruohan Zhang

Department of Computer Science, The University of Texas at Austin

Introduction

In end-to-end learning of visuomotor behaviors, algorithms such as imitation learning, reinforcement learning (RL), or a combination of both, have achieved remarkable successes. However millions of training samples are often required to achieve the desired performance. Hence in practice human demonstration can be used to speedup learning and improve sample efficiency (Silver et al. 2016; Hester et al. 2018).

When an agent learns from human demonstration, it may benefit from knowing where the human is allocating his or her visual attention, which can be inferred from the gaze. Humans have high acuity foveal vision only in the central 1-2 visual degrees of the visual field, covering the width of a finger at arm's length. Consequently humans learn to move their foveae to the correct place at the right time in order to process important task-relevant visual features (Hayhoe and Ballard 2005). A wealth of information is encoded in human gaze behaviors—for example, the priority of one object over another in obtaining longterm reward. Therefore human attention mechanism can be viewed as a powerful visual feature extractor for reward-seeking decision tasks.

We propose a framework that uses learned human visual attention model to guide the learning process of an imitation learning or reinforcement learning agent. We have collected high-quality human action and eye-tracking data while playing Atari games in a carefully controlled experimental setting. We have shown that incorporating a learned human gaze model into deep imitation learning yields promising results. We plan to further improve the existing algorithm for imitation learning, and incorporating visual attention into deep reinforcement learning. Meanwhile we are studying how visual attention is formed in a biological neural network and hope to transfer that knowledge to artificial neural networks.

Proposal

Related and Accomplished Work The relation between attention and RL has been extensively studied by neuroscience community (Roelfsema and van Ooyen 2005; Leong et al. 2017). In machine learning, (Mnih et al. 2014) pioneered efforts to combine deep RL with visual attention,

where attention is treated as a sequential decision problem and is jointly trained with the control policy.

In my previous work we proposed the AGIL (Attention Guided Imitation Learning) framework (Zhang et al. 2017; 2018). With collected human gaze data in Atari games, we first train a deep network that can predict human visual attention 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 action prediction accuracy and task performance, see Fig. 1a. Here human attention information helps the learning agent infer the correct decision state of the human by highlighting the relevant visual features for that decision in a high-dimensional state space.

Prior to deep RL, my research focused on a modularized solution to the RL problems. Modular RL assumes that humans have limited attention resources hence cannot pay attention to all state features at the same time. Therefore a complex task should be decomposed into modules, where each module is a simpler Markov decision process with reduced state space that is factorized from the original state space (Rothkopf and Ballard 2013). My work used modular RL as a cognitive model for human behaviors, and developed a corresponding sample-efficient modular inverse reinforcement learning (IRL) algorithm to retrieve the intrinsic reward of human. Using virtual-reality and motion capture, we collected human navigation decisions in a virtual room with multiple subtasks. With the proposed modular IRL algorithm, we were able to create an avatar which can infer relative importance (reward) of different objects and goals from observed human behavior, and generate walking trajectories that are human-like in novel environment. The modular IRL significantly outperformed a standard Bayesian IRL in modeling human behaviors.

Improving AGIL The key challenge for visual attention guided learning is to fuse the attention information into the decision (policy) network. In previous work a naive pipeline architecture was used (predicting gaze first then feed the result into the policy network via element-wise masking). Fusing multiple channels of information is well studied in the computer vision community (Feichtenhofer, Pinz, and Zisserman 2016). We find that the design choices of a fusion network, such as where (which layer) and how (which arith-

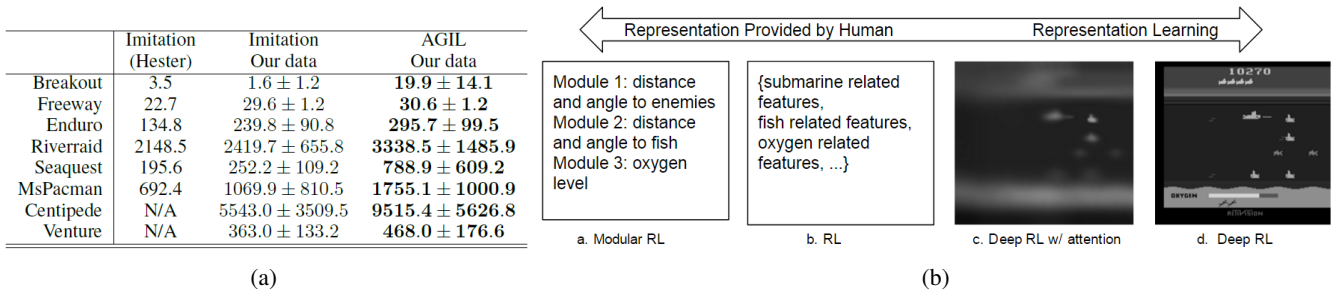


Figure 1: (a) A comparison of game scores (mean \pm standard deviation) between plain imitation learning from a previous work (Hester et al. 2018), plain imitation learning using our dataset, and AGIL. (b) The state space provided to different reinforcement learning algorithms for Atari Seaquest game.

metic operator) to fuse, significantly affect the task performance. We plan to continue experimenting multiple network architectures to fuse attention information.

Attention guided reinforcement learning Given our initial progress in modular RL and AGIL, it is natural to incorporate visual attention into deep RL. The assumption made here is that visual features that capture human attention are likely to be informative for deep RL. Deep Q-network has demonstrated the effectiveness of end-to-end learning of visuomotor tasks (Mnih et al. 2015). However, for games such as Seaquest and MsPacman—which typically involve multiple tasks—the performance is below human level. In addition, DQN takes millions of samples to train. The above issues could be potentially alleviated by combining AGIL and deep RL where attention model can help extract features to speedup learning and to indicate task priority. Fig. 1b illustrates the relation between the modular RL, the standard RL, the attention-guided deep RL proposed, and the standard deep RL in terms of how they define their state space.

In addition, although the human attention is likely to be useful for guiding a learning agent, it might be more important to study whether that an agent can learn its own attention mechanism from experience. At current stage we are working on understanding how attention is realized in animal nervous system from a neural coding standpoint (Ballard and Zhang 2018). Our model suggests that the attention gain effect may reflect the use of additional neurons by a computational process. It is possible that a similar attention mechanism can be implemented using the current artificial neural networks, therefore a learning agent can learn attention and decision jointly and efficiently.

Below is a brief timeline for research plan.

1. Sep.-Dec.2018: Collecting another 22 Atari games’ human eye-tracking and decision data. The data is made available to encourage similar research in Atari domain.
2. Sep.-Nov.2018: Experimenting different fusion architectures for AGIL.
3. Nov.2018-Feb.2019: Experimenting how to incorporate human attention into deep RL.
4. Feb.2019-: Investigating the possibility to implement an neural attention mechanism such that a deep RL agent

can jointly learn attention and policy.

References

- Ballard, D., and Zhang, R. 2018. Cortical spike multiplexing using gamma frequency latencies. *bioRxiv* 313320.
- Feichtenhofer, C.; Pinz, A.; and Zisserman, A. 2016. Convolutional two-stream network fusion for video action recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1933–1941.
- Hayhoe, M., and Ballard, D. 2005. Eye movements in natural behavior. *Trends in cognitive sciences* 9(4):188–194.
- Hester, T.; Vecerik, M.; Pietquin, O.; Lanctot, M.; Schaul, T.; Piot, B.; Sendonaris, A.; Dulac-Arnold, G.; Osband, I.; Agapiou, J.; et al. 2018. Deep q-learning from demonstrations. *Association for the Advancement of Artificial Intelligence (AAAI)*.
- Leong, Y. C.; Radulescu, A.; Daniel, R.; DeWoskin, V.; and Niv, Y. 2017. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* 93(2):451–463.
- Mnih, V.; Heess, N.; Graves, A.; and Kavukcuoglu, K. 2014. Recurrent models of visual attention. In *Advances in Neural Information Processing Systems*, 2204–2212.
- Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A. A.; Veness, J.; Bellemare, M. G.; Graves, A.; Riedmiller, M.; Fidjeland, A. K.; Ostrovski, G.; et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533.
- Roelfsema, P. R., and van Ooyen, A. 2005. Attention-gated reinforcement learning of internal representations for classification. *Neural computation* 17(10):2176–2214.
- Rothkopf, C. A., and Ballard, D. H. 2013. Modular inverse reinforcement learning for visuomotor behavior. *Biological cybernetics* 107(4):477–490.
- Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; Van Den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; et al. 2016. Mastering the game of go with deep neural networks and tree search. *Nature* 529(7587):484–489.
- Zhang, R.; Liu, Z.; Hayhoe, M. M.; and Ballard, D. H. 2017. Attention guided deep imitation learning. *Cognitive Computational Neuroscience*.
- 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)*.