

Dueling Network Architectures for Deep Reinforcement Learning

ECE 239AS Poster Presentation

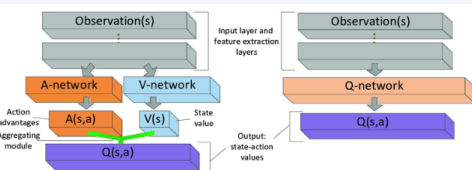
Sarat Bhargava Eden Haney Radhika Nayar

BACKGROUND

The dueling architecture for Deep Q-learning (DQN) leverages a unique neural network architecture customized for model-free RL [1] to improve performance of the standard deep Q-learning algorithms.

This architecture advances an existing neural network by improving already established algorithms such as the robust Double Deep Q-learning (Double DQN) algorithm [2] and Deep Q-Networks (DQN) [3].

ARCHITECTURE



Dueling Q-network (left) and a standard single-stream Q-network (right). Image from [3].

The dueling network architecture adopts two sequences (or streams) of fully connected layers.

The two streams provide separate estimates of:

1. state-value function
2. state-dependent action advantage function

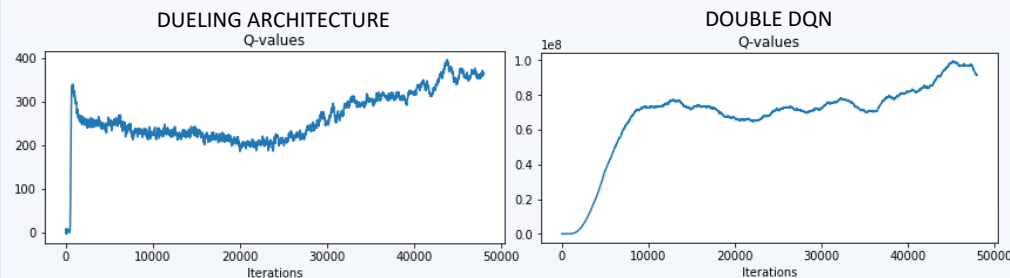
The two streams are then combined to output a set of Q-values, one for each action.

Key motivation: by separating two estimators, the dueling architecture can learn which states are valuable, without learning the effect of each action for each state.

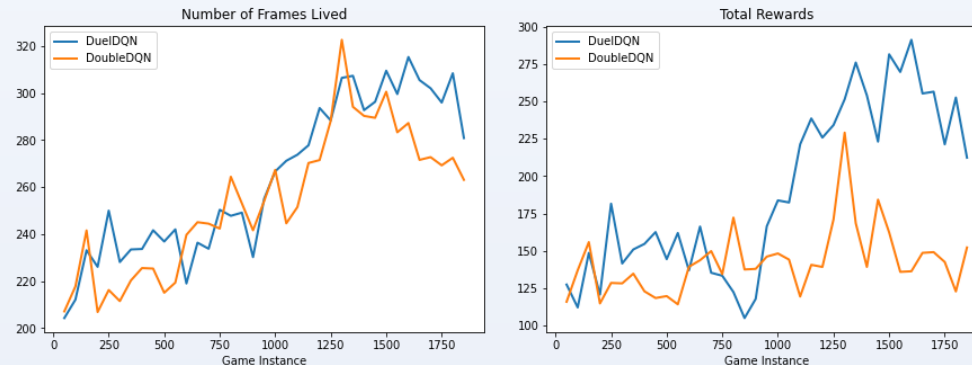
IMPLEMENTATION

Implemented the **Dueling Architecture** on top of **Double DQN** in the **Space Invaders** games on the Atari environment, given only **raw pixel observations** and **game rewards**

RESULTS



In both algorithms the Q-values plateau initially before increasing around 200k iterations. We also see the Double DQN is grossly overvaluing the Q-values.



The Dueling Architecture and Double DQN learn to live for approximately the same number of frames.

We see an increase in score around the same time as the increase in number of frames lived and the Q-values

Final Score Statistics

	mean	Std. Dev.
Baseline (No Training)	174.7	82.0
Dueling	212.2	116.6
Double DQN	32.2	70.5

Scores were obtained by running each algorithm for 500k iterations (~1800 game instances).

The Dueling Architecture average score increases significantly from the baseline. Double DQN has much lower score.

CONCLUSION

Based on our results, we conclude that the Dueling architecture advances Double DQN in Space Invader Atari game domain.

Advantages of the Dueling Architecture over Double DQN:

1. Better approximation of the state values due to the dueling architecture which decouples value and advantage in deep Q-networks, while sharing a common feature learning module.
2. Unlike Double DQN, the dueling architecture is robust to noise caused by difference in scales of average action gap and average state value gap during updates.

REFERENCES

1. Z. Wang, T. Schaul, M. Hessel, H. van Hasselt, M. Lanctot, and N. de Freitas, "Dueling network architectures for deep reinforcement learning," arXiv:1511.06581 [cs], Apr 2016.
2. H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," arXiv:1509.06461 [cs], Dec 2015. arXiv: 1509.06461.
3. V. Mine, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, and et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, p. 529-533, Feb 2015.