深度增强学习方向论文整理

Alex-zhai · 18 天前

一. 开山鼻祖DQN

1. Playing Atari with Deep Reinforcement Learning，V. Mnih et al., NIPS Workshop, 2013.

2. Human-level control through deep reinforcement learning, V. Mnih et al., Nature, 2015.

二. DQN的各种改进版本（侧重于算法上的改进）

1. Dueling Network Architectures for Deep Reinforcement Learning. Z. Wang et al., arXiv, 2015.

2. Prioritized Experience Replay, T. Schaul et al., ICLR, 2016.

3. Deep Reinforcement Learning with Double Q-learning, H. van Hasselt et al., arXiv, 2015.

4. Increasing the Action Gap: New Operators for Reinforcement Learning, M. G. Bellemare et al., AAAI, 2016.

5. Dynamic Frame skip Deep Q Network, A. S. Lakshminarayanan et al., IJCAI Deep RL Workshop, 2016.

6. Deep Exploration via Bootstrapped DQN, I. Osband et al., arXiv, 2016.

7. How to Discount Deep Reinforcement Learning: Towards New Dynamic Strategies, V. François-Lavet et al., NIPS Workshop, 2015.

8. Learning functions across many orders of magnitudes，H Van Hasselt，A Guez，M Hessel，D Silver

9. Massively Parallel Methods for Deep Reinforcement Learning, A. Nair et al., ICML Workshop, 2015.

10. State of the Art Control of Atari Games using shallow reinforcement learning

11. Learning to Play in a Day: Faster Deep Reinforcement Learning by Optimality Tightening（11.13更新）

12. Deep Reinforcement Learning with Averaged Target DQN（11.14更新）

三. DQN的各种改进版本（侧重于模型的改进）

1. Deep Recurrent Q-Learning for Partially Observable MDPs, M. Hausknecht and P. Stone, arXiv, 2015.

2. Deep Attention Recurrent Q-Network

3. Control of Memory, Active Perception, and Action in Minecraft, J. Oh et al., ICML, 2016.

4. Progressive Neural Networks

5. Language Understanding for Text-based Games Using Deep Reinforcement Learning

6. Learning to Communicate to Solve Riddles with Deep Distributed Recurrent Q-Networks

7. Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation

8. Recurrent Reinforcement Learning: A Hybrid Approach

四. 基于策略梯度的深度强化学习

深度策略梯度：

1. End-to-End Training of Deep Visuomotor Policies

2. Learning Deep Control Policies for Autonomous Aerial Vehicles with MPC-Guided Policy S earch

3. Trust Region Policy Optimization

深度行动者评论家算法：

1. Deterministic Policy Gradient Algorithms

2. Continuous control with deep reinforcement learning

3. High-Dimensional Continuous Control Using Using Generalized Advantage Estimation

4. Compatible Value Gradients for Reinforcement Learning of Continuous Deep Policies

5. Deep Reinforcement Learning in Parameterized Action Space

6. Memory-based control with recurrent neural networks

7. Terrain-adaptive locomotion skills using deep reinforcement learning

8. Compatible Value Gradients for Reinforcement Learning of Continuous Deep Policies

9. SAMPLE EFFICIENT ACTOR-CRITIC WITH EXPERIENCE REPLAY（11.13更新）

搜索与监督：

1. End-to-End Training of Deep Visuomotor Policies

2. Interactive Control of Diverse Complex Characters with Neural Networks

连续动作空间下探索改进：

1. Curiosity-driven Exploration in DRL via Bayesian Neuarl Networks

结合策略梯度和Q学习：

1. Q-PROP: SAMPLE-EFFICIENT POLICY GRADIENT WITH AN OFF-POLICY

CRITIC（11.13更新）

2. PGQ: COMBINING POLICY GRADIENT AND Q-LEARNING（11.13更新）

其它策略梯度文章：

1. Gradient Estimation Using Stochastic Computation Graphs

2. Continuous Deep Q-Learning with Model-based Acceleration

3. Benchmarking Deep Reinforcement Learning for Continuous Control

4. Learning Continuous Control Policies by Stochastic Value Gradients

五. 分层DRL

1. Deep Successor Reinforcement Learning

2. Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation

3. Hierarchical Reinforcement Learning using Spatio-Temporal Abstractions and Deep Neural Networks

4. Stochastic Neural Networks for Hierarchical Reinforcement Learning – Authors: Carlos Florensa, Yan Duan, Pieter Abbeel （11.14更新）

六. DRL中的多任务和迁移学习

1. ADAAPT: A Deep Architecture for Adaptive Policy Transfer from Multiple Sources

2. A Deep Hierarchical Approach to Lifelong Learning in Minecraft

3. Actor-Mimic: Deep Multitask and Transfer Reinforcement Learning

4. Policy Distillation

5. Progressive Neural Networks

6. Universal Value Function Approximators

7. Multi-task learning with deep model based reinforcement learning（11.14更新）

8. Modular Multitask Reinforcement Learning with Policy Sketches （11.14更新）

七. 基于外部记忆模块的DRL模型

1. Control of Memory, Active Perception, and Action in Minecraft

2. Model-Free Episodic Control

八. DRL中探索与利用问题

1. Action-Conditional Video Prediction using Deep Networks in Atari Games

2. Curiosity-driven Exploration in Deep Reinforcement Learning via Bayesian Neural Networks

3. Deep Exploration via Bootstrapped DQN

4. Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation

5. Incentivizing Exploration In Reinforcement Learning With Deep Predictive Models

6. Unifying Count-Based Exploration and Intrinsic Motivation

7. #Exploration: A Study of Count-Based Exploration for Deep Reinforcemen Learning（11.14更新）

8. Surprise-Based Intrinsic Motivation for Deep Reinfo rcement Learning（11.14更新）

九. 多Agent的DRL

1. Learning to Communicate to Solve Riddles with Deep Distributed Recurrent Q-Networks

2. Multiagent Cooperation and Competition with Deep Reinforcement Learning

十. 逆向DRL

1. Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization

2. Maximum Entropy Deep Inverse Reinforcement Learning

3. Generalizing Skills with Semi-Supervised Reinforcement Learning（11.14更新）

十一. 探索+监督学习

1. Deep learning for real-time Atari game play using offline Monte-Carlo tree search planning

2. Better Computer Go Player with Neural Network and Long-term Prediction

3. Mastering the game of Go with deep neural networks and tree search, D. Silver et al., Nature, 2016.

十二. 异步DRL

1. Asynchronous Methods for Deep Reinforcement Learning

2. Reinforcement Learning through Asynchronous Advantage Actor-Critic on a GPU（11.14更新）

十三：适用于难度较大的游戏场景

1. Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation, T. D. Kulkarni et al., arXiv, 2016.

2. Strategic Attentive Writer for Learning Macro-Actions

3. Unifying Count-Based Exploration and Intrinsic Motivation

十四：单个网络玩多个游戏

1. Policy Distillation

2. Universal Value Function Approximators

3. Learning values across many orders of magnitude

十五：德州poker

1. Deep Reinforcement Learning from Self-Play in Imperfect-Information Games

2. Fictitious Self-Play in Extensive-Form Games

3. Smooth UCT search in computer poker

十六：Doom游戏

1. ViZDoom: A Doom-based AI Research Platform for Visual Reinforcement Learning

2. Training Agent for First-Person Shooter Game with Actor-Critic Curriculum Learning

3. Playing FPS Games with Deep Reinforcement Learning

4. LEARNING TO ACT BY PREDICTING THE FUTURE（11.13更新）

5. Deep Reinforcement Learning From Raw Pixels in Doom（11.14更新）

十七：大规模动作空间

1. Deep Reinforcement Learning in Large Discrete Action Spaces

十八：参数化连续动作空间

1. Deep Reinforcement Learning in Parameterized Action Space

十九：Deep Model

1. Learning Visual Predictive Models of Physics for Playing Billiards

2. J. Schmidhuber, On Learning to Think: Algorithmic Information Theory for Novel Combinations of Reinforcement Learning Controllers and Recurrent Neural World Models, arXiv, 2015. arXiv

3. Learning Continuous Control Policies by Stochastic Value Gradients

4.Data-Efficient Learning of Feedback Policies from Image Pixels using Deep Dynamical Models

5. Action-Conditional Video Prediction using Deep Networks in Atari Games

6. Incentivizing Exploration In Reinforcement Learning With Deep Predictive Models

二十：DRL应用

机器人领域：

1. Trust Region Policy Optimization

2. Towards Vision-Based Deep Reinforcement Learning for Robotic Motion Control

3. Path Integral Guided Policy Search

4. Memory-based control with recurrent neural networks

5. Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection

6. Learning Deep Neural Network Policies with Continuous Memory States

7. High-Dimensional Continuous Control Using Generalized Advantage Estimation

8. Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization

9. End-to-End Training of Deep Visuomotor Policies

10. DeepMPC: Learning Deep Latent Features for Model Predictive Control

11. Deep Visual Foresight for Planning Robot Motion

12. Deep Reinforcement Learning for Robotic Manipulation

13. Continuous Deep Q-Learning with Model-based Acceleration

14. Collective Robot Reinforcement Learning with Distributed Asynchronous Guided Policy Search

15. Asynchronous Methods for Deep Reinforcement Learning

16. Learning Continuous Control Policies by Stochastic Value Gradients

机器翻译：

1. Simultaneous Machine Translation using Deep Reinforcement Learning

目标定位：

1. Active Object Localization with Deep Reinforcement Learning

目标驱动的视觉导航：

1. Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning

自动调控参数：

1. Using Deep Q-Learning to Control Optimization Hyperparameters

人机对话：

1. Deep Reinforcement Learning for Dialogue Generation

2. SimpleDS: A Simple Deep Reinforcement Learning Dialogue System

3. Strategic Dialogue Management via Deep Reinforcement Learning

4. Towards End-to-End Learning for Dialog State Tracking and Management using Deep Reinforcement Learning

视频预测：

1. Action-Conditional Video Prediction using Deep Networks in Atari Games

文本到语音：

1. WaveNet: A Generative Model for Raw Audio

文本生成：

1. Generating Text with Deep Reinforcement Learning

文本游戏：

1. Language Understanding for Text-based Games Using Deep Reinforcement Learning

无线电操控和信号监控：

1. Deep Reinforcement Learning Radio Control and Signal Detection with KeRLym, a Gym RL Agent

DRL来学习做物理实验：

1. LEARNING TO PERFORM PHYSICS EXPERIMENTS VIA DEEP REINFORCEMENT LEARNING（11.13更新）

DRL加速收敛：

1. Deep Reinforcement Learning for Accelerating the Convergence Rate（11.14更新）

利用DRL来设计神经网络：

1. Designing Neural Network Architectures using Reinforcement Learning（11.14更新）

2. Tuning Recurrent Neural Networks with Reinforcement Learning（11.14更新）

3. Neural Architecture Search with Reinforcement Learning（11.14更新）

控制信号灯：

1. Using a Deep Reinforcement Learning Agent for Traffic Signal Control（11.14更新）

二十一：其它方向

避免危险状态：

1. Combating Deep Reinforcement Learning's Sisyphean Curse with Intrinsic Fear （11.14更新）

DRL中On-Policy vs. Off-Policy 比较：

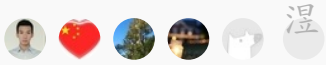
1. On-Policy vs. Off-Policy Updates for Deep Reinforcement Learning（11.14更新）

注1：小伙伴们如果觉得论文一个个下载太麻烦，可以私信我，我打包发给你。

注2：欢迎大家及时补充新的或者我疏漏的文献。

深度学习（Deep Learning） 深度增强学习（Deep Reinforcement Learning） 机器学习

414 分享 举报



16 条评论

写下你的评论

韩劼群

最近放出来许多2017ICLR的投稿，有不少是关于DRL的，我目前读过里面比较有意思的有：1. Learning to Play in a Day: Faster Deep Reinforcement Learning by Optimality Tightening， 2. PGQ: Combining policy gradient and Q-learning， 3. Q-Prop: Sample-Efficient Policy Gradient with An Off-Policy Critic， 4. Sample Efficient Actor-Critic with Experience Replay，5. Learning to Act by Predicting the Future。1,2,4都应用在了Atari Games上，3,4 应用在Robotics continuous control上，5 在 Doom Full Deathmatch track 中赢得了第一名。

17 天前 6 赞

lin feng

赞

17 天前 1 赞

Alex-zhai（作者）回复 **韩劼群** 查看对话

最近比较忙，还没来得及关注，我加一下。谢谢啦。

17 天前 2 赞

朱军

可以在github上建个repo

17 天前

吴江

计算机论文的学习不要脱离代码，在各位同学看论文时请尽量在github上撸点代码，有时候他们的程序实现跟论文写得完全不是一回事。

17 天前 4 赞

张夏天

赞

17 天前

Alex-zhai（作者）回复 **吴江** 查看对话

谢谢老师的建议。

17 天前 1 赞

天清

感谢

16 天前

韩劼群 回复 **朱军** 查看对话

可以看一下这个 repo，muupan/deep-reinforcement-learning-papers ，不过作者最近好像没啥更新

16 天前

机械键盘侠

mark

16 天前 1 赞