Last updated August 15, 2022.

Alireza Kazemipour

Website: alirezakazemipour.github.io
Email: kazemipour.alireza@gmail.com
LinkedIn: alireza-kazemipour
GitHub: alirezakazemipour

Education

University of Alberta — Edmonton, Canada
M.Sc. in Computing Science — Jan 2023–TBD

K. N. Toosi University of Technology — Tehran, Iran
B.Sc. in Electrical-Control Engineering — Sep 2015–Sep 2020

- Thesis: "Continuous Control With Deep Reinforcement Learning"
- Demonstration of the superiority of different state-of-the-art Deep RL methods in Continuous Space settings compared to conventional non-adaptive Control approaches such as PDI Controller.

Research Interests

- (Deep) (Multi-Agent) (Self-Play) Reinforcement Learning
- Deep (Unsupervised) (Semi-Supervised) (Representation) Learning
- Adaptive Autonomous Agents Learning
- Robotics

Academic Projects

Deep Deterministic Policy Gradient and Hindsight Experience Replay (DDPG and HER)
Using the DDPG to control continuously and the HER to mitigate the problem of sparsity of rewards. [Project page]

- Implemented the DDPG + HER to train a 7DOF manipulator to fetch, pick and place a box in a sparse-reward but multi-goal environment.

Diversity is All You Need (DIAYN) [Project page]
Learning Skills without a Reward Function.

- The DIAYN (with the SAC as the backend) was used to train a Hopper robot, a Bipedal walker, and the notoriously challenging Mountain Car agent to learn useful and Diverse skills without supervision.

Soft Actor-Critic (SAC) [Continuous SAC page] [Discrete SAC page]
Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor.

- A Humanoid was trained to walk using the SAC (as the SOTA method in Continuous settings in RL), as well as an agent with the discrete version of the SAC who was trained to play Ms Pacman.

Proximal Policy Optimization Algorithms (PPO) [Mario page] [Ant & Walker2d page]
Policy Gradient methods that alternate between sampling data and optimizing a ”surrogate” objective.

- The PPO was applied to solving 29/32 levels on Super Mario Bros and to training simulated Ant and Walker2d robots to locomote.

Twin Delayed Deep Deterministic Policy Gradient (TD3) [Project page]
Addressing Function Approximation Error in Actor-Critic Methods.

- The TD3 algorithm was implemented to train Ant and Hopper robots in proper forward movement.

Sample Efficient Actor-Critic with Experience Replay (ACER) [Project page]

An Actor-Critic Deep RL method with experience replay that is stable and sample efficient.

- The ACER method was implemented in conjunction with the A3C (Asynchronous Methods for Deep RL) method on Space Invaders and Pong, showing improvements in the sample efficiency metric.

Self-Imitation Learning (SIL) [Project page]

A simple off-policy actor-critic algorithm that learns to reproduce the agent's past good decisions.

- The SIL method was implemented in conjunction with the A2C (synchronous version of A3C) on the game of Freeway, showing exploitation of good past experiences drives to deep exploration.

Trust Region Policy Optimization (TRPO) [Project page]

An iterative procedure for optimizing policies, with guaranteed monotonic improvement.

- The TRPO was applied to the games of Beam Rider and Ms Pacman and the promised sample efficiency of this 2nd order method was observed given the run-time of other 1st order optimization based methods.

Actor Critic using Kronecker-Factored Trust Region (ACKTR) [Project page]

Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation.

- Applied the ACKTR on the games of Seaquest and Ms Pacman and successfully implemented the K-FAC optimizer to apply Natural Gradients on the weights of the Neural Net needed in Atari benchmark.

Exploration by Random Network Distillation (RND) [Project page]

Exploration based on intrinsic rewards (Novelty Seeking).

- Implemented the RND (using the PPO as the backend with Convolutional and Recurrent policies) to solve the game of Montezuma's Revenge and the first level of the game of Super Mario Bros that both demand intelligent exploration and are sparse-reward environments.

Rainbow [Project page]

Combining Improvements in Deep Reinforcement Learning.

- Atari games Pong and Boxing were taught to an agent by the full improvements (Prioritized Experience, Noisy Nets, Dueling Architecture, etc.) to the DQN paper.

Distributional RL [Project page]

Comprehensive implementation of Deep Distributional Reinforcement Learning Algorithms.

- Implemented C51 (Categorical DQN), IQN (Implicit Quantile Networks), QRDQN (Quantile Regression DQN) and FQF (Fully parameterized Quantile Function) to train an agent playing Space Invaders.

Tabular Reinforcement Learning

Fundamental Reinforcement Learning algorithms in tabular format to solve Taxi gym environment.

- Off-Policy Temporal Difference Learning (Q-Learning). [Project Page]
- State–action–reward–state–action (SARSA). [Project Page]
- Backward View of TD(λ) both by Q-Learning and SARSA. [Q(λ) Project Page] [SARSA(λ) Project Page]
- Combination of Q-learning and Q-planning (Dyna-Q). [Project Page]

Neural Networks Without Frameworks [Project page]

Implementation of different Neural Network configurations without using scientific frameworks.

- Implemented different configurations and ideas involved in Deep Learning without using scientific frameworks like TensorFlow and PyTorch in: Python (with and without taking advantage of NumPy), C++ and Java.

Cycle GAN

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. [Project page]

- Implemented the Cycle GAN to produce fake horse images from real zebra images and vice versa.

Auxiliary and Deep Convolutional GANs

Using DCGAN to generate realistic-looking images and Auxiliary GAN to improve quality. [Project page]

- Implemented the DCGAN and the AUXGAN to produce images similar to the MNIST dataset.

Deep Dream and Style Transfer

Using the Deep Dream to visualize a CNN's layer output and the Style Transfer to transfer an image's style. [Project Page]

- Implemented the Deep Dream to visualize a mixture of five-layer outputs of the Inception-V3 model with frozen weights that had been obtained by being trained previously on the ImageNet dataset.
- Implemented the Style Transfer to transfer the style of the Last Supper image to an image of the Limmat River in Zurich by using the VGG19 model to extract and deploy the style.

Face Detection and Facial Expression Classification

The final project of Fundamentals of Computer Vision course. [Project Description]

- Face detection was accomplished by Cascade Detectors with Local Binary Pattern features, then the expression was categorized by a CNN.

Persian Digits Classification

Designing and preparing instructions of 13th lab of the Fundamentals of Computer Vision course.

- Utilized Linear and RBF Support-Vector Machines, K-Nearest Neighbors, and Random Forest methods to classify images of handwritten Persian digits.

Academic Experience

26th International Computer Conference, Computer Society of Iran [Homepage]

Lecturer — Mar 2021

- Presented some recent Deep RL advances and challenges associated with Deep RL algorithms' implementations. [slides]

Fundamentals of Computer Vision Course [Homepage]

Teaching Assistant — Feb 2015–Jul 2020

- Designed and prepared instructions of 13th lab (Machine Learning session) of the course and moreover, was responsible to evaluate and grade 20% of students based on their performance in each lab of the course.
- Instructor: Dr. Behrooz Nasihatkon

Signals and Systems Course

Head Teaching Assistant — Sep 2019–Feb 2020

- Was Responsible to design and arrange assignments and weekly reporting of students' outcomes.
- Instructor: Dr. Maryam mohebbi

KN2C Robotics Team [Homepage]

Research Assistant — Sep 2017–Sep 2019

- Computer Vision and A.I researcher at the Micro Aerial Vehicle section.
- Supervisor: Dr. Hamid D. Taghirad

Awards and Honors

- Got Admitted to the MSc program in Computing Science at the University of Alberta — 2022
- **3rd place** in The RoboCup Iran Open Competitions in the Unmanned Aerial Vehicle League. [link] — 2018
- **6th place** in The RoboCup Asia-Pacific Competitions in the Unmanned Aerial Vehicle League. [link] — 2018
- Ranked within the **top 0.8%** in Iran's National Bachelor University Exam among 252,000 participants. — 2015

Online Courses

- **Deep Reinforcement Learning**
  CS 285 at UC Berkeley
- **Designing, Visualizing and Understanding Deep Neural Networks**
  CS W182 / 282A at UC Berkeley
- **Reinforcement Learning Specialization** [Certificate]
  University of Alberta on Coursera
- **Improving Deep NNs, Hyperparameter Tuning, Regularization, and Optimization** [Certificate]
  DeepLearning.AI on Coursera
- **Neural Networks and Deep Learning** [Certificate]
  DeepLearning.AI on Coursera
- **Structuring Machine Learning Projects** [Certificate]
  DeepLearning.AI on Coursera
- **Machine Learning** [Certificate]
  Stanford | Online on Coursera
- **Algorithms Specialization**[1] [Certificate]
  Stanford | Online on Coursera

Technical Skills

- **Programming Languages:** Python, C/C++, Java, Bash, VHDL
- **Libraries:** PyTorch, TensorFlow, Keras, Gym, NumPy, Scikit-learn, OpenCV
- **Frameworks:** Qt, ROS, MATLAB, Simulink and Octave
- **Version Control Systems:** Git
- **Linux Distros:** Ubuntu, Manjaro

Languages

- **Farsi:** Native
- **English:** Professional Proficiency

❗References, further information, and proofs are available upon request.

[1] **Projects**: *Karatsuba's algorithm, Counting number of Inversions in an Array, Quick Sort, Randomized Contraction algorithm to compute the Min Cut, TwoSum Problem, Dijkstra, Kosaraju's algorithm, Median Maintenance, Greedy Clustering, Greedy Job Scheduling, Huffman Coding, Knapsack, Prim's algorithm for the MST problem, Maximum-Weight Independent Set, 2-SAT, BellmanFord, Travelling Salesman with Heuristic approach*