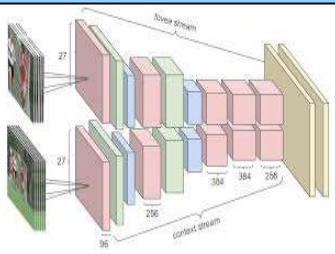
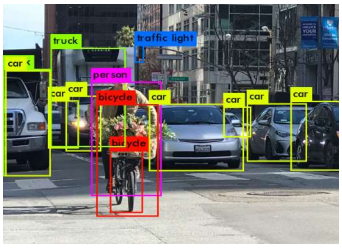
TELECOMMUNICATIONS, COMPUTER ENGINEERING, AND PHOTONICS INSTITUTE
Sant'Anna School of Advanced Studies – Pisa

Retis Real-Time Systems Laboratory

Reinforcement Learning

The first lecture is scheduled on March 16, 2027 at 9:00
Visit the course web page for registration and connecting to the channel.
http://retis.sssup.it/~giorgio/courses/neural/nn.html

Giorgio Buttazzo

Objectives: This course focuses on the Reinforcement Learning paradigm. It introduces the objective, and the main algorithms proposed in the literature, including tabular methods, deep Q-learning models, policy gradient methods, and model-based approaches. Several examples are illustrated and critical implementation issues are also discussed.
Hours: 20. **CFU**: 2.

Course program

1. **Reinforcement Learning**. Basic concepts and intuitions. The problem of temporal credit assignment. Early models: The box model and its neural implementation. A taxonomy of existing RL methods.
2. **State value estimation methods**: Temporal difference (TD) methods. Monte Carlo, n-step TD, and TD-lambda estimation methods.
3. **Control methods**: Q-learning, SARSA algorithms, Q-lambda, and SARSA-lambda.
4. **Implementation issues**: Implementing the Q-learning algorithm in C. Structuring the program. Data structures and main functions. Examples of applications: The grid world, Inverted pendulum, Amoeba, Shooting targets, Visual tracking, Lunar lander, Crawling robot, Self-driving car.
5. **Deep Reinforcement Learning**. Deep Q-learning models: DQL, Double-DQL, Dueling-DQL, Rainbow DQL, Recurrent DQL.
6. **Policy gradient methods**. The REINFORCE algorithm. Actor-critic methods: DDPG, TD3, A3D, SAC, and PPO.
7. **Stochastic RL**: A statistical approach to reinforcement learning for stochastic environments.
8. **Curiosity-based learning**: Variational Information Maximizing Exploration learning. Surprise-based Intrinsic Motivation. Intrinsic Curiosity Module. Random Network Distillation.
9. **Imitation learning**: Behavioral cloning. Iterative policy learning. Inverse Reinforcement Learning (IRL): Feature expectation matching, Maximum Entropy IRL, Guided Cost Learning. Generative Adversarial Imitation Learning. Imitation by diffusion.
10. **Model-based RL**: How to learn a model. The Dyna approach. Dyna extensions: Deep Dyna-Q, D3Q, Dyna-PINN, MBPO. Tree-search methods: Minimax and Monte Carlo Tree Search. AlphaGo.