ECSE Final Year Project, Semester2, 2021:
Learning to play hide and seek, An active visual guidance system

Department of Electrical and Computer Systems Engineering

Author : Daming Wang
Supervisor : Dr. Michael Burke

MONASH University

Overview

Exploring unknown environments has become a common robotic task in various modern industries. In this case, **learning from human demonstration** is an effective way for robot agents to imitate human actions.

This research proposes an approach to teach the robot agents to **recognize the latent information of interest to a demonstrator.** The ranking models can automatically **evaluate new frames** in the video sequences.

Figure 1. Potentially Applicable Agents

The traditional operation mode for **the robot exploration task** is still based on **remote control**. Robot agents are expected to **autonomously** learn and identify the target and conduct navigation and planning.

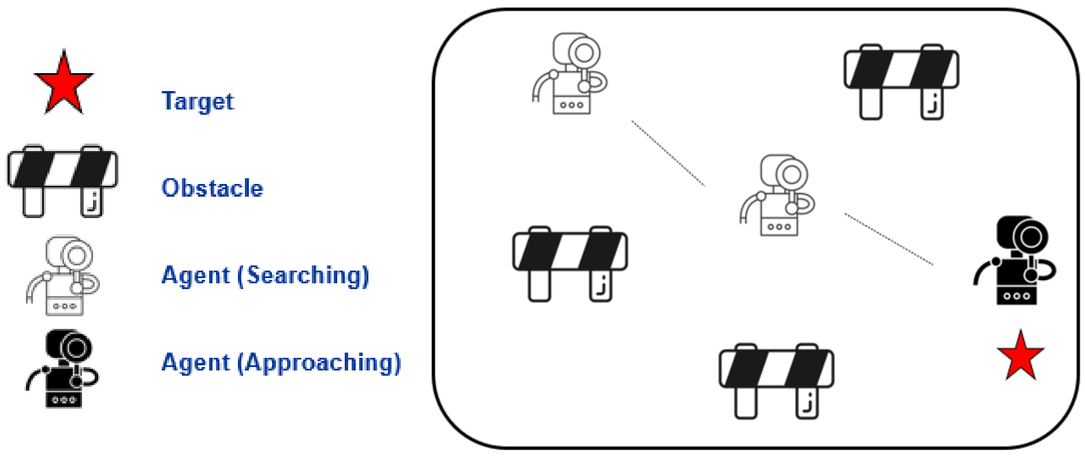

Figure 2. Autonomous Searching Example

Figure 3. Demonstration Sequences Example

- **Our task is to train a model that can predict how important an image is using only a few sequences of demonstrated images.**

Image Pairwise Comparison & Siamese Network

The project adopts **Unsupervised Learning** and **pairwise compares** the image pairs by **Learning from Time**. Computing the **Siamese network [1]** presents the image pairs' differences, further training the image temporal ranking model.

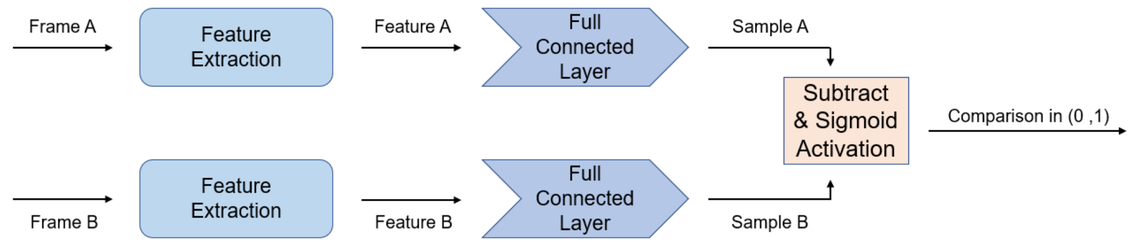

Figure 4. Schematic Diagram for The Siamese Network

TrueSkill Algorithm & Probabilistic Temporal Ranking

The **deterministic ranking models** predict the rewards based on the image features directly. However, to improve performance further, the **probabilistic ranking models[2]** adopt the ***TrueSkill***[3] idea and predict rewards with **uncertainty**.

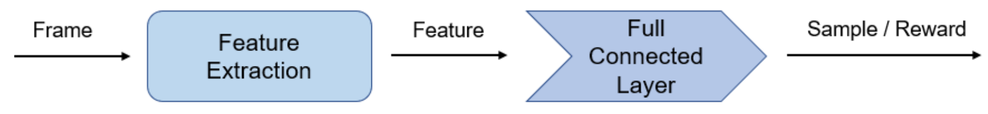

Figure 5. Schematic Diagram for the Deterministic Temporal Ranking Model

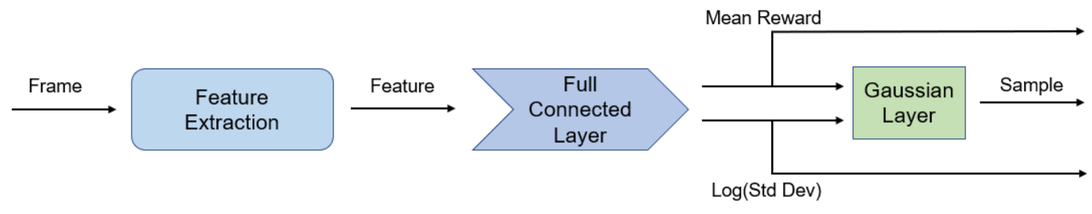

Figure 6. Schematic Diagram for the Probabilistic Temporal Ranking Model

Outcome & Case Study

The Proposed Ranking Models can localize the target in the frames. Compared with DTR model, PTR model can concentrate on the target better with **carrying rewards uncertainty**.

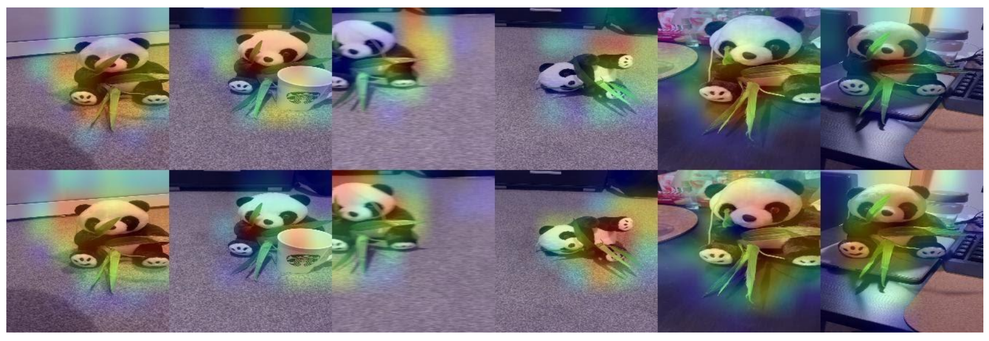

Figure 7. Target localization Saliency (Upper: DTR; Lower: PTR)

Both **DTR** and **PTR** approaches are experimented with different scenarios to imitate seeking targets or discover new environments in the real world.

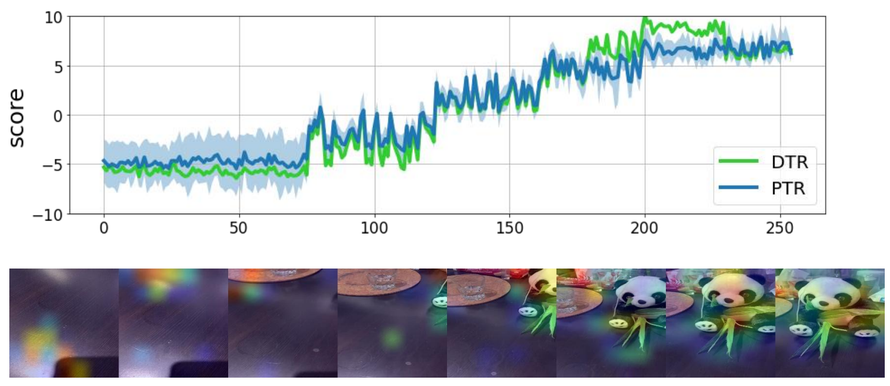

Figure 8. Ranking For a Monotonic Process

Further Development

The **Data Augmentation approach** might be adopted to expand the datasets diversity. The **alter labelling approaches** might improve ranking performance. **Furthermore, the image ranking system could be applied in real world navigation and discovery tasks.**

References

[1] Koch, Gregory. Siamese neural networks for one-shot image recognition. ICML Deep Learning Workshop, 2015.
[2] Michael Burke, Katie Lu, Daniel Angelov, Artras Straiys, Craig Innes, Kartic Subr, Subramanian Ramamoorthy "Learning rewards for robotic ultrasound scanning using probabilistic temporal ranking" Institute of Perception, Action and Behaviour, University of Edinburgh, Edinburgh, UK
[3] Ralf Herbrich, Tom Minka, and Thore Graepel. TrueSkill: a Bayesian skill rating system. In Advances in neural information processing systemsm, 2007.