Opening of a Postdoc on Disentangled Latent Manipulation Learning for Dexterous Robotic Manipulation

@Imagine Team of LIRIS lab, Ecole Centrale de Lyon

Advisors: Liming Chen (Professor, LIRIS CNRS UMR 5205, ECL), Emmanuel Dellandréa (Associate Professor, ECL)

SUBJECT

1. Context and main aims

Dexterous manipulation of objects is a core task in robotics. Because of the design complexity needed for robot controllers even for simple manipulation tasks for humans, *e.g.*, pouring a cup of tea, robots currently in use are **mostly limited to specific tasks within a known environment**.

Disentangled Representation Learning is to learn a task-independent representation which captures an underlying low-dimensional representation of the observed world factorized into a number of variation factors, *e.g.*, object properties, geometry, lighting, *etc*. Its theoretical advantages for supervised learning and reinforcement learning, *e.g.,* data efficiency for subsequent machine learning, the generalization skills, transferability and interpretability of learnt models, have been highlighted in [Bengio et al. 2013] and there exists a large body of literature, *e.g.*, face editing and recognition under various factors including face identity, lighting, pose facial expression [Tran et al.@CVPR2017][Marriot et al.@FG2020] in computer vision, contextualized word embeddings and language models [Vaswani et al.@NIPS2017][Brown et al.2020] in NLP, and more recently [Higgins et al.@ICML2017][Whitney et al.@ICRL2020] in RL for robotics. Within the project, We are building upon our expertise on disentangled latent feature space learning [Marriott et al.@CVPR2021] and manipulation subspace learning [Katyara et al.@IEEE TCDS2021], and extending the previous research on representation learning to robot manipulation control and performing simulation enabled self-supervised representation learning to learn embeddings of states and action sequences and capture the structure of the environment's dynamics, and thereby improve data efficiency in our end-to-end RL for dexterous bimanual robotic manipulation.

Despite uncountable number of potential applications enabled by such a robotic manipulation system, we will focus our attention on three use cases that we have been implementing through 4 research projects, namely 1) bin-picking widely required in logistics and industrial assembly line, 2) waste sorting for better waste recycling and environment protection, and 3) assistance systems for "stick-to-bed" patients or elders with limited physical ability in their daily life object manipulation tasks, *e.g.*, fetching a bottle of water and pouring it into a glass. They are closely related to three ongoing research projects within the group, namely the 3-year LEARN-REAL project, the PSPC FAIR WASTES project and the CHIRON project.

2. Research strategy and main milestones

This research project aims to extend the large body of research works on disentangled representation learning, currently mostly focused on the observation or state space, to cover the action space and the dynamics of a robot agent, *i.e.*, various task level robot controllers, for the purpose of dexterous manipulation. Such a disentangled latent manipulation feature space is to capture the structure of the underlying low-dimensional manipulation manifold with the following three components: T1 - Disentangled latent representation learning of the observed environment; T2 - Latent action space; and T3 - Dynamics-aware embeddings. The last task T4 will rely upon the previous results to perform transfer learning of the robot bimanual manipulation skills based on the latent manipulation representation. The overall goal of this thesis is to make it possible for efficient policy learning despite sparse rewards and long-horizons in manipulation tasks and enable control from disentangled low-dimensional state space despite much increased degree of freedom due to low-level image input, tactile sensing and bimanual fingered soft hands, using hierarchical representations of temporally abstract actions.

3. Skills

The candidate must:

- be fluent in french or english
- be fluent or even expert with linux, C ++ and Python languages
- be familiar with Cmake, ROS and Gitlab

If necessary, skills development will be offered for the mastery of ROS (Robot Operating System, to control robots and sensors), Gazebo (to simulate robots), Docker (installations and dependencies), OpenCV, machine learning frameworks and Deep Learning for Vision (Cuda, Theano, Torch, Tensor Flow).

He (She) should like teamwork and will be called to help in writing novel research proposals, collaborate with other PhD students and supervise interns.

4. Duration and Preferred starting date

24 months; July 1st 2022

5. Application

Applications should include a detailed curriculum vitae, brief statements of interests and two reference letters.

Applications and letters should be sent via electronic mail to:

- Pr. Liming Chen (liming.chen@ec-lyon.fr)