RObot Manipulation Learning by Demonstration using **G**enerative **A**dversarial **N**etworks (ROGAN)

University of Adelaide, Australia, AIML
IMT Atlantique, France, LabSTICC/HAAL

1 PhD thesis proposal summary

Following its introduction in industrial manufacturing, robotics is nowadays entering into new application domains through its introduction in urban environments. This provides opportunities but also poses challenges in redefining conventional practices and methods, as is the case in healthcare and well-being that can become more user-oriented, by offering services remotely and with higher degree of autonomy.

Our research interest in this thesis proposal concerns the development of such services through the use of humanoid service robots, via machine learning methods capable of transferring human demonstrated skills to the robot. This topic is broadly referred as learning/programming from demonstration (LfD) and deals with task-based coupling of robot action with its sensory/perceptual data. This thesis provides the opportunity to build upon recent work in deep learning and probabilistic graphical models research [12, 6], mainly applied in computer vision applications, to vision-guided robot manipulation skills in use cases related to personal assistance of human users by a companion robot [4]. The thesis is the product of synergy between the teams HAAL of CNRS Lab-STICC in IMT-A and PGM group of University of Adelaide, specialized in smart spaces and cognitive robotics on the one hand and high-end artificial intelligence based on deep neural networks and probabilistic graphical models on the other.

The thesis will address challenges at a technological and methodological level. In the latter case, it will primarily pursue a task-based mapping of human kinematics to robot arm kinematics, namely, addressing the correspondence problem, via Generative Adversarial Networks (GANs) [10, 8]. This will be the case of supervised human demonstrations given to the robot within the LivingLab facility of IMT-A which emulates the real operating conditions of an assisting robot along human users and experts. Complementary to such demonstrations, the thesis will also explore weakly supervised learning from unannotated video demonstration, which would allow training on significantly richer data sources and reduce the dependence for accurate human to robot kinematics. At a technical level, the doctoral candidate will evaluate the applicability of the developed algorithms on robot arms of 6 and 4 degrees of freedom and of interchangeable grippers [2, 3] as well as varying operational contexts, e.g. on heterogeneous robotic platform (cf. Turtlebot2i [1]). The newly developed robot functionalities will be then integrated into the smart house integration protocol of the LivingLab [13], allowing the composition of novel assistive services driven by use case scenarios that involve other interconnected sensors and actuators.

2 Envisioned methodological and technical approach

One particular area that we envision to be of particular interest for this project is imitation learning [7] in which the robot learns to mimic human behaviour through small number of demonstrations. This learning is done utilizing deep learning. However, since the number of human provided demonstrations is limited compared to the potentially large number of states that the robot can be in, imitation learning is particularly challenging. As such, we use generative models, in particular GANs, to learn the underlying structure of the demonstrations. GANs are recently proposed generative models that are shown to be capable of capturing the intrinsic structure of complex data distributions such as images. In GANs two deep neural networks called the generator and the discriminator compete. The generator seeks to generate samples that look like the ones from the true distribution. For instance in our imitation learning of the robot, it generates samples of the robot behaviour that look like the human's. On the other hand, the discriminator is tasked with distinguishing the actions generated from the true distribution (human actions) vs. the ones from the robot's. The discriminator learns a "distance" between the two distributions that gives us a measure of the robots ability to mimic human demonstration. For instance, in [5] the algorithm learns to map the human hand movements to the pose estimates using GANs. All in all, these generative models allow the robot to experience the situations that it might come across without a need for a large training dataset.

These imitation learning models are embedded into reinforcement learning [14] so that the robot experiments with various strategies to interact with the environment and the generative model provides the intrinsic rewards. Within the reinforcement learning, the GAN's generator is the policy that learns to map the states the robot is into the actions. Examples of such approaches are recently investigated in theoretical level such as [11, 9].

3 Thesis perimeter

The doctoral student is expected to have a federating role between the two teams of IMT-A and UoA along with an increased inter-site mobility. The research activities of the thesis are in line with a collaboration initiated via a Memorandum of Understanding (MoU) between IMT-A and UoA, further sponsored by Brittany region, the French state and South Australia.

In detail, the hosting facilities for the thesis are:

- **IMT Atlantique Bretagne/Pays de la Loire (Brest campus)**, a public institute of superior education (postgraduate) and research, that accredits Master diplomas to engineers up to Doctoral degree. The Lab-STICC/HAAL[1] research team specialises in information and communication technologies for assistance to people, combining skills from informatics and robotics: domotics protocols, embedded systems, service robotics, robotic learning et interactive TV. It collaborates with medical personnel, ergonomists, sociologists. IHSEV disposes state-of-the-art robotic equipment (mobile robots (TurtleBot2 and variants), humanoids (Pepper, RB-1, Poppy), robotic arms (Kinova), etc), offering an attractive working environment for hosting research activities of international visibility.

- **The University of Adelaide** is one of the top 1% universities worldwide and a member of the group of eight in Australia. It is recognised globally as a leading research university in particular in computer vision where it is ranked 3rd in the world. From 2018, the *Australian Institute for Machine Learning*[2] (AIML) was officially established at the University of Adelaide with a mission to carrying out research to provide benefits to society and preparing the next generation of experts. AIML has now more than 120 members with a range of expertise in various theoretical and practical areas of artificial intelligence, computer vision and deep learning. Collaboration with world-leading institutions to develop high-tech solutions is at the heart of AIML.

4 Candidate profile

Holder of a postgraduate diploma, Master of research or engineer diploma in the domains of Robotics, Mechatronics, Computer Science or associated field. Fluency in English is required, a spirit of collaboration and of initiative in the face of technological challenges.

Theoretical skills: Solid background in one or more of the following domains: *deep learning, computer vision, robotic system engineering.*

Technical skills: Experience in one more or more of the following technologies/tools: *DL frameworks (pytorch, tensorflow, etc), Robot Operating System*

[1] `https://bit.ly/2kKXh1I`
[2] `https://www.adelaide.edu.au/aiml/`

(ROS), scientific computing tools (Scikit, numpy, OpenCV, PCL), 3D simulation (Gazebo)

The interested candidates should provide the following:

- CV
- Motivation letter
- Academic notes transcript

and optionally, letters of recommendation. Interested candidates should send by e-mail (with subject "[ROGAN]:application") their applications to **Panagiotis Papadakis** and **Ehsan Abbasnejad**:

panagiotis.papadakis@imt-atlantique.fr;ehsan.abbasnejad@adelaide.edu.au

References

[1] `https://github.com/Interbotix/turtlebot2i`. Turtlebot2i platform.

[2] `https://www.kinovarobotics.com/en/assistive-technologies/jaco-assistive-robotic-arm`. JACO Kinova arm.

[3] `https://www.kinovarobotics.com/en/products/robotic-arm-series/mico-robotic-arm`. MICO Kinova arm.

[4] `https://www.robotnik.eu/manipulators-2/rb-one/`. RB1 Robotic manipulator.

[5] M. Abdi, E. Abbasnejad, C. P. Lim, and S. Nahavandi. 3d hand pose estimation using simulation and partial-supervision with a shared latent space. *CoRR*, abs/1807.05380, 2018.

[6] P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. F. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, Ç. Gülçehre, H. F. Song, A. J. Ballard, J. Gilmer, G. E. Dahl, A. Vaswani, K. R. Allen, C. Nash, V. Langston, C. Dyer, N. Heess, D. Wierstra, P. Kohli, M. Botvinick, O. Vinyals, Y. Li, and R. Pascanu. Relational inductive biases, deep learning, and graph networks. *CoRR*, abs/1806.01261, 2018.

[7] A. Billard, S. Calinon, R. Dillmann, and S. Schaal. *Robot Programming by Demonstration*, pages 1371–1394. 2008.

[8] C. Chou, J. Chien, and H. Chen. Self adversarial training for human pose estimation. *CoRR*, abs/1707.02439, 2017.

[9] J. Fu, K. Luo, and S. Levine. Learning robust rewards with adversarial inverse reinforcement learning. *CoRR*, abs/1710.11248, 2017.

[10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27*, pages 2672–2680. 2014.

[11] J. Ho and S. Ermon. Generative adversarial imitation learning. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 4565–4573. 2016.

[12] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.

[13] C. Lohr, P. Tanguy, and J. Kerdreux. xaal: A distributed infrastructure for heterogeneous ambient devices. *J. Intelligent Systems*, 24(3):321–331, 2015.

[14] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.