DIPARTIMENTO DI INGEGNERIA
CORSO DI DOTTORATO IN INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE -
PHD COURSE IN INDUSTRIAL AND INFORMATION ENGINEERING - 34TH CYCLE

Title of the research activity:	Machine learning tools for collaborative robotics
State of the Art:	The robotic revolution is calling for more and more advanced algorithms and solutions, to cope with the huge number of problems foreseen in all the potential application fields. Just to mention a few scenarios: driverless cars, drones, autonomous vehicles for custom applications, humanoid robots for edutainment and impaired people support, down to advanced AGV and cooperative robots for industries aligned with the 4.0 paradigm. The development of smart, robust and efficient robotic systems is strongly dependent, among other, on machine learning tools. The Isarlab group has developed a string expertise on the field, and is now addressing the problem of smart human robot collaboration and the problem of smart localization and navigation. The common framework is the use of computer vision solutions and ideas. As for the smart human robot collaboration, one of the open problem if that of allowing the robot to "understand" and "describe" in natural language the surroundings. This will allow the robot to nicely interact with the human operators, both in industrial and daily-life scenarios. Some results along this line have been presented in [Rohrbach et al., 2015], [Barbu et al., 2012], [Venugopalan et al., 2014], [Baraldi et al., 2017]. Initial results from the proponent are given in [Cascianelli et al., 2018]. As for smart localization and navigation, the general framework is that of Visual Odometry (VO) based on deep learning solutions. While geometry based VO is well accepted, it still is suffers from a number of robustness issues [D. Scaramuzza, F. Fraundorfer, 2011], which can be overcame by a learned approach, i.e., by algorithms where the input-output relationship between an acquired image flow and the associated motion parameters is computed by means of a deep learning model. [T. A. Ciarfuglia at alii, 2014]. At the same time, machine leaning techniques are today extremely important for the analysis of the large amount of data and signals that are produced continuously by modern industrial production plants, especially in the area of predictive diagnostics and of quality control. In this context the study of "time series mining techniques" [Fu, 2011] is a natural complement of the above mentioned machine learning techniques based on machine vision that can be

	used together for the intelligent conduction of industry 4.0 plants.
Short description and objectives of the research activity:	The PhD project is aimed at developing innovative solutions for the above mentioned problems. As for the collaborative issue, the objective is to explore approaches based on synthetic description of the scene automatically generated by deep-learning based algorithms. The problem will be addressed both based on data set currently under development at Isarlab, and on additional data to be produced. As for the localization and navigation issue, the focus will be on end-to-end policies, aimed at computing the control signals for the robot directly from the measured sensor streaming, e.g., directly from the video images. The interaction between the two problems will also addressed, and extensive experimental campaigns carried out with available and foreseen robotic equipment.
Bibliography:	[Pillonetto et al., 2014] Pillonetto, Gianluigi, et al. "Kernel methods in system identification, machine learning and function estimation: A survey." Automatica 50.3 (2014): 657-682. [Hassani et al., 2014] Hassani, Vahid, Tegoeh Tjahjowidodo, and Thanh Nho Do. "A survey on hysteresis modelling, identification and control." Mechanical systems and signal processing 49.1-2 (2014): 209-233. [Rohrbach et al., 2015] Rohrbach, M. Rohrbach, and B. Schiele, "The long-short story of movie description," in Proc. German Conf. Pattern Recognition , 2015, pp. 209–221. [Barbu et al., 2012] A. Barbu et al. , "Video in sentences out," in Proc. 28th Conf. Uncertainty Artif. Intell. , 2012, pp. 102–112. [Venugopalan et al., 2014] S. Venugopalan, H. Xu, J. Donahue, M. Rohrbach, R. Mooney, and K. Saenko, "Translating videos to natural language using deep recurrent neural networks," in Proc. 2015 Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol. , 2014, pp. 1495–1504. [Baraldi et al., 2017] Baraldi, C. Grana, and R. Cucchiara, "Hierarchical boundary-aware neural encoder for video captioning," in Proc. 2017 IEEE Conf. Comput. Vis. Pattern Recognit. , 2017, pp. 3185–3194. [Cascianelli et al., 2018] Cascianelli, Silvia, et al. "Full-GRU Natural Language Video Description for Service Robotics Applications." to appear in IEEE Robotics and Automation Letters. [D. Scaramuzza, F. Fraundorfer, 2011], Visual odometry [tutorial], IEEE Robot. Autom. Mag. 18 (4) (2011) 80–92. [T. A. Ciarfuglia at alii, 2014], Evaluation of non-geometric methods for visual odometry. Robotics and Autonomous Systems (2014), Vol. 62, No. 12, pp. 1717-1730. [Fu, 2011] Fu, T. C. (2011). A review on time series data mining. Engineering Applications of Artificial Intelligence, 24(1), 164-181.
Scientific coordinator (s)	Prof. Paolo Valigi
Contact (s)	Paolo.valigi@unipg.it Https://isar.unipg.it