A Deep Learning-based system for Collaborative Robotics: real-time monitoring and gesture recognition at the edge

Cristina Nuzzi {c.nuzzi@unibs.it}

This project focuses on **Collaborative Robotics**, one of the of the Industry 4.0 paradigm. The issues this research aims to solve are:

1. to define a natural and easy-to-use gesture language as the communication method between human operators and robots;
2. to set up a real-time monitoring system that can understand the gestures of the operator and, at the same time, can take care of the safety strategies;
3. to port the whole system intelligence on an embedded platform to remove or at least to reduce the usage of a PC.

The novelty of the project is in the proposed approach, which intensely uses **Deep Learning algorithms** to create the intelligence of the system, along with **Machine Vision** techniques and a suitable number of vision sensors.

The Problem

In industrial environments robotic systems and manipulators are usually kept inside safety cages to protect the operators. In the new Industry 4.0 paradigm, however, the safety cage is removed, and the operators share the same workspace of the machines. This is why the **safety** of the human workers is the first of the three central issues of Collaborative environments.

The second issue is the **communication method** between humans and robots: how can machines and humans truly collaborate if they cannot interact naturally with each other? As shown in [1], an efficient human-human team communicate through vision, voice and by constantly knowing the position of the other members of the team. The aim of this research is to port this concept on a human-robot team, using vision sensors such as cameras to monitor the scene in real-time. This allows the development of a real-time people detection and tracking, which means that the system can constantly know the relative position between the humans in the scene and the robot. It is then straightforward to apply a suitable number of safety strategies on the system to ensure the safety of the operators [2]. While voice commands are widely used for a large number of consumer-end applications, due to the disturbances of industrial environments in this research only visual data are considered for the communication method, in the form of gesture commands. Natural gestures, though, are not suitable for industrial environments, where the safety of the operator is placed first and it is essential to avoid wrong or not specifically requested commands. Although a set of gestures for human-robot collaboration has already been proposed [3], in this research one of the aims is to design a complete, intuitive and user-friendly language. A first step has already been made in our first publication [4]. The novelty of the project is characterized by the intense use of Deep Learning strategies, which have been refined through the years and only recently have been intensely researched also for industrial applications.

A third issue is characterized by the need of **reducing the hardware** in terms of both dimension and costs. Now, the trend is to port the applications on embedded platforms, which are less expensive and require less space. These platforms can be easily mounted on-site, for example near the robot, and can be easily accessed by using a common laptop or a remote connection.

The focus of the proposed research project is the development of a system for Collaborative Robotics. The system must perform the following main tasks at the same time and in real-time:

1. to monitor the position of human operators entering the scene (**detection and tracking**);
2. to intervene with safety procedures according to the mutual position of the robot and the operators (**safety**);
3. to understand the gesture commands the operator is performing, and to translate them into operative actions for the robot, according to the predetermined list of allowed commands (**collaboration**).

This system will be the intelligence-on-site of the robotic system, and the final goal is to **deploy it on an embedded platform to be mounted on-site**.

References

[1] H. Liu e L. Wang, «Gesture recognition for human-robot collaboration: A review,» *International Journal of Industrial Ergonomics,* 2017.

[2] S. De Salvia, S. Pasinetti e G. Sansoni, *Sviluppo di una barriera di sicurezza virtuale per applicazioni robotiche: acquisizioni multiple e simultanee mediante telecamere a tempo di volo,* Brescia: Università degli Studi di Brescia, 2017.

[3] P. Barattini, C. Morand e N. M. Robertson, «A proposed gesture set for the control of industrial collaborative robots,» in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012, pp. 132 - 137.

[4] C. Nuzzi, S. Pasinetti, M. Lancini, F. Docchio e G. Sansoni, «Deep Learning based Machine Vision: first steps towards a hand gesture recognition set up for Collaborative Robots,» in *Proceedings of the 2018 IEEE International Workshop on Metrology for Industry 4.0 and IoT*, Brescia, Italy, 2018.