Doctoral Dissertation
Doctoral Program in Management and Production Engineering (35th Cycle)

A Deep Learning and Augmented Reality Framework for 3D human-machine interaction

By

Leonardo Tanzi

Supervisor(s):
Prof.ssa Federica Marcolin, Supervisor

Doctoral Examination Committee:
Prof. Motyl Barbara, Università degli Studi di Udine
Prof. Speranza Domenico, Università degli Studi di Cassino
Prof. Baronio Gabriele, Università di Brescia
Prof. Gloria Antonio, Università degli Studi di Napoli Federico II
Prof. Montrucchio Bartolomeo, Politecnico di Torino

Politecnico di Torino
2022

Abstract

In the context of Industry 4.0, human-machine interaction (HMI) is a multidisciplinary field of study focusing on the design of computer technology and the interaction between humans (the users) and computers. In recent years, Artificial Intelligence (AI) and particularly Deep Learning (DL) has introduced huge improvements in the field. The same applies to Augmented Reality (AR), a technology that offers the possibility to build intuitive and immersive interfaces between the user and the machine, creating tools and methods for the approach for augmented humans. In different complex situations, a user, both expert and beginner, could improve his performance with the help of an assistance application.

Following these needs, this study focuses on building an adaptable and general-purpose framework for assistance that could be applied in different areas, such as the biomedical and the industrial field, to underline the scalability of this approach. The proposed framework is composed of two layers: the first one focuses on the DL analysis of the data, while the second on the AR assistance.

After a brief overview, we focused on a literature review of augmented human approaches to critically analyze the state-of-the-art for identifying the baseline of DL and AR frameworks, discussed in Chapter 1.

In Chapter 2, we described the first layer, which comprises two projects related to classification. Concerning the biomedical area, we worked on classification DL algorithms to analyze images related to anatomical human regions, i.e., bone X-Rays. This research was carried out in collaboration with the A.O.U Città della Salute e della Scienza of Turin. In the industrial setting, we worked on DL algorithms to exploit 2D/3D data to support the classification of human emotions in collaboration with the Polytechnic University of Milan.

In Chapter 3, we presented the results of these three years of research, i.e., the second layer of the framework and the consequent two fully working pipelines. The first project focuses on computer-assisted prostatectomy, while the second on computer-assisted troubleshooting. The resulting final framework composes the two layers to build a computer-assistance tool that analyzes the video stream and returns 3D information to augment it.

Finally, Chapter 4 presents a discussion and concludes this research.