

# Roberto Bigazzi

Ph.D in Computer Vision and Deep Learning

Modena, 41124, Italia

☎ +39 334 8756245 | ✉ roberto.bigazzi.1995@gmail.com | 🏠 bigazzon.github.io | 📄 roberto-bigazzi | 🎓 R. Bigazzi

## Summary

I'm a Post-Doc Researcher at the University of Modena and Reggio Emilia, in the AlmageLab research laboratory supervised by the Professor Rita Cucchiara. The main topics of my research are Computer Vision and Deep Learning, in particular on robotic visual navigation. I pursued my Ph.D. at AlmageLab research laboratory after completing my Master's Degree in Computer Science and Engineering at the Polytechnic University of Milan, with the Professor Marco Gribaudo. During my Ph.D. program, I spent a period as visiting student researcher at Stanford University in the Autonomous Systems Lab (ASL) directed by the Professor Marco Pavone.

## Education

### University of Modena and Reggio Emilia

Modena, Italy

DOCTOR OF PHILOSOPHY IN INFORMATION AND COMMUNICATION TECHNOLOGIES: COMPUTER VISION AND DEEP LEARNING

Nov. 2019 - Mar. 2023

- Multimedia Data Learning at AlmageLab under the supervision of Prof. Rita Cucchiara

### Stanford University

Stanford, California, United States

VISITING STUDENT RESEARCHER

May 2022 - Aug. 2022

- Research on Visual Navigation at Autonomous Systems Lab (ASL) under the supervision of Prof. Marco Pavone

### Polytechnic University of Milan

Milan, Italy

MASTER OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

Sep. 2017 - Oct. 2019

- Final Grade: 108/110. Thesis with Prof. Marco Gribaudo: "Enhancing Spatial Navigation in Robot-Assisted Surgery: a case study"

### Technische Universität Wien

Vienna, Austria

ERASMUS+ EXCHANGE SEMESTER

Oct. 2018 - Feb. 2019

### Polytechnic University of Milan

Milan, Italy

BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING

Sep. 2014 - Sep. 2017

- Final Grade: 99/110

### Istituto Istruzione Superiore Roncalli

Poggibonsi, Siena, Italy

LICEO SCIENTIFICO TECNOLOGICO

Sep. 2009 - Jul. 2014

## Publications

### CONFERENCE PROCEEDINGS

2023 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)

#### Embodied Agents for Efficient Exploration and Smart Scene Description

Roberto Bigazzi, Silvia Cascianelli, Lorenzo Baraldi, Marcella Cornia, Rita Cucchiara

The development of embodied agents that can communicate with humans in natural language has gained increasing interest over the last years, as it facilitates the diffusion of robotic platforms in human-populated environments. As a step towards this objective, in this work, we tackle a setting for visual navigation in which an autonomous agent needs to explore and map an unseen indoor environment while portraying interesting scenes with natural language descriptions. To this end, we propose and evaluate an approach that combines recent advances in visual robotic exploration and image captioning on images generated through agent-environment interaction. Our approach can generate smart scene descriptions that maximize semantic knowledge of the environment and avoid repetitions. Further, such descriptions offer user-understandable insights into the robot's representation of the environment by highlighting the prominent objects and the correlation between them as encountered during the exploration. To quantitatively assess the performance of the proposed approach, we also devise a specific score that takes into account both exploration and description skills. The experiments carried out on both photorealistic simulated environments and real-world ones demonstrate that our approach can effectively describe the robot's point of view during exploration, improving the human-friendly interpretability of its observations.

**Focus on Impact: Indoor Exploration with Intrinsic Motivation**

Roberto Bigazzi, Federico Landi, Silvia Cascianelli, Lorenzo Baraldi, Marcella Cornia, Rita Cucchiara

Exploration of indoor environments has recently experienced a significant interest, also thanks to the introduction of deep neural agents built in a hierarchical fashion and trained with Deep Reinforcement Learning (DRL) on simulated environments. Current state-of-the-art methods employ a dense extrinsic reward that requires the complete a priori knowledge of the layout of the training environment to learn an effective exploration policy. However, such information is expensive to gather in terms of time and resources. In this work, we propose to train the model with a purely intrinsic reward signal to guide exploration, which is based on the impact of the robot's actions on the environment. So far, impact-based rewards have been employed for simple tasks and in procedurally generated synthetic environments with countable states. Since the number of states observable by the agent in realistic indoor environments is non-countable, we include a neural-based density model and replace the traditional count-based regularization with an estimated pseudo-count of previously visited states. The proposed exploration approach outperforms DRL-based competitors relying on intrinsic rewards and surpasses the agents trained with a dense extrinsic reward computed with the environment layouts. We also show that a robot equipped with the proposed approach seamlessly adapts to point-goal navigation and real-world deployment.

2022 26TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR)

**Spot the Difference: A Novel Task for Embodied Agents in Changing Environments**

Federico Landi, Roberto Bigazzi, Marcella Cornia, Silvia Cascianelli, Lorenzo Baraldi, Rita Cucchiara

Embodied AI is a recent research area that aims at creating intelligent agents that can move and operate inside an environment. Existing approaches in this field demand the agents to act in completely new and unexplored scenes. However, this setting is far from realistic use cases that instead require executing multiple tasks in the same environment. Even if the environment changes over time, the agent could still count on its global knowledge about the scene while trying to adapt its internal representation to the current state of the environment. To make a step towards this setting, we propose Spot the Difference: a novel task for Embodied AI where the agent has access to an outdated map of the environment and needs to recover the correct layout in a fixed time budget. To this end, we collect a new dataset of occupancy maps starting from existing datasets of 3D spaces and generating a number of possible layouts for a single environment. This dataset can be employed in the popular Habitat simulator and is fully compliant with existing methods that employ reconstructed occupancy maps during navigation. Furthermore, we propose an exploration policy that can take advantage of previous knowledge of the environment and identify changes in the scene faster and more effectively than existing agents. Experimental results show that the proposed architecture outperforms existing state-of-the-art models for exploration on this new setting.

2020 25TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR) (ORAL PRESENTATION)

**Explore and Explain: Self-supervised Navigation and Recounting**

Roberto Bigazzi, Federico Landi, Marcella Cornia, Silvia Cascianelli, Lorenzo Baraldi, Rita Cucchiara

Embodied AI has been recently gaining attention as it aims to foster the development of autonomous and intelligent agents. In this paper, we devise a novel embodied setting in which an agent needs to explore a previously unknown environment while recounting what it sees during the path. In this context, the agent needs to navigate the environment driven by an exploration goal, select proper moments for description, and output natural language descriptions of relevant objects and scenes. Our model integrates a novel self-supervised exploration module with penalty, and a fully-attentive captioning model for explanation. Also, we investigate different policies for selecting proper moments for explanation, driven by information coming from both the environment and the navigation. Experiments are conducted on photorealistic environments from the Matterport3D dataset and investigate the navigation and explanation capabilities of the agent as well as the role of their interactions.

**OTHER RESEARCH WORK**

2021 19TH INTERNATIONAL CONFERENCE ON COMPUTER ANALYSIS OF IMAGES AND PATTERNS (CAIP)

**Out of the Box: Embodied Navigation in the Real World**

Roberto Bigazzi, Federico Landi, Marcella Cornia, Silvia Cascianelli, Lorenzo Baraldi, Rita Cucchiara

2021 21ST INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING (ICIAP)

**Embodied Navigation at the Art Gallery**

Roberto Bigazzi, Federico Landi, Silvia Cascianelli, Marcella Cornia, Lorenzo Baraldi, Rita Cucchiara

UNDER REVIEW

**Indoor Semantic Region Mapping: Language-grounded Visual Recognition for Embodied Navigation**

Roberto Bigazzi, Shreyas Kousik, Lorenzo Baraldi, Rita Cucchiara, Marco Pavone

UNDER REVIEW

**AIGeN: An Adversarial Approach for Instruction Generation in Vision-and-Language Navigation**

Niyati Rawal, Roberto Bigazzi, Lorenzo Baraldi, Rita Cucchiara

## Towards Explainable Embodied Navigation and Recounting

Samuele Poppi, Roberto Bigazzi, Niyati Rawal, Marcella Cornia, Silvia Cascianelli, Lorenzo Baraldi, Rita Cucchiara

## Program Committees

---

2021-now **Peer Reviewer**, IEEE Robotics and Automation Letters (RA-L)

2022-now **Peer Reviewer**, IEEE Transactions on Geoscience and Remote Sensing (TGRS)

2022-now **Peer Reviewer**, IEEE Pattern Recognition Letters (PRL)

2023 **Peer Reviewer**, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)

*Detroit, USA*

2023 **Peer Reviewer**, IEEE International Conference on Robotics and Automation (ICRA)

*London, UK*

2022 **Peer Reviewer**, IAPR International Conference on Pattern Recognition (ICPR)

*Montreal, Canada*

2021 **Peer Reviewer**, ACM International Conference on Multimedia

*Chengdu, China*

2020 **Peer Reviewer**, Transactions on Multimedia Computing Communications and Applications (TOMM)

2020 **Peer Reviewer**, ACM International Conference on Multimedia

*Seattle, US (virtual)*

## Other Activities

---

2020-23 **Teaching Assistant**, Computer Architectures

2019-23 **Presenter**, “Into The Future” University Orientation Program

2022 **Presenter**, “Humanities & Intelligence” Laboratory

2021-22 **Project Tutor**, AI for Automotive

2021-22 **Lecturer**, AI for Automotive: SLAM Lessons

2021 **Presenter**, “In the mind of robots, how AI teaches robots to navigate” Laboratory

2020-21 **Project Tutor**, Neural Network Computing, AI and Machine Learning for Automotive

2019-20 **Project Tutor**, Computer Vision and Cognitive Systems

## Experience

---

### AlmageLab - University of Modena and Reggio Emilia

*Modena, Italy*

RESEARCH FELLOW

*Nov. 2019 - present*

Research during the Doctorate at AlmageLab covering these topics in Computer Vision: Deep Learning, Reinforcement Learning, Visual Navigation, Robotics, Machine Learning, Domain Adaptation.

### Prometeia

*Bologna, Italy*

LECTURER

*Nov. 2021*

Python and Machine Learning course for Banco BPM

### IFOA Bologna

*Bologna, Italy*

LECTURER

*Mar. 2021 - Apr. 2021*

Data Analysis and Data Visualization course

### IFOA Modena

*Modena, Italy*

LECTURER

*Jan. 2021 - Feb. 2021*

Deep Learning, Artificial Intelligence e Neurolinguistic Processing (IBM, SAS) course

### Polytechnic University of Milan

*Milan, Italy*

MASTER THESIS

*Mar. 2019 - Sep. 2019*

Master Thesis with the title “Enhancing Spatial Navigation in Robot-Assisted Surgery: a case study” supervised by Professor Marco Gribaudo.

Abstract: Use of augmented reality in order to assist surgeons during surgery for prostatectomies, trying to solve the problem of identifying unambiguous landmarks from the video recorded by the camera of the “da Vinci Surgical System” and speeding up the steps needed to apply the 3D model superimposed to the prostate.

### Cyber-Physical Systems Group - Technische Universität Wien

*Vienna, Austria*

PROJECT DEVELOPER

*Oct. 2018 - Jan. 2019*

Project performed on the creation of a vertical farming system during the course of Internet of Things of the Professor Ezio Bartocci at TU Wien.

Main topics of the project: BLE Communication, Microcontroller Programming, Cloud Computing, Arrowhead Framework.

# Languages

---

## Italian

Mother tongue

## English

**Writing:** Fluent, **Reading:** Fluent, **Listening:** Fluent

- TOEIC (C1): Grade 960/990
- FCE (B2): Grade C

# Certificates

---

## SCHOOLS AND WORKSHOPS

2021 **4th Advanced Course on Data Science and Machine Learning**, ACDL

*Siena, Italy*

2017 **Interactive Internet of Things e Smart Object Design Course**, Polytechnic University of Milan

*Milan, Italy*

2013 **European Computer Driving Licence Certificate**, ECDL

# Skills

---

**Topics** Deep Learning, Machine Learning, Neural Networks, Multimodal Learning, Parallel Programming

**Programming** Python, Java, Javascript, C++, C#, C, MATLAB, SQL, Microcontroller Programming (Arduino)

**Other** PyTorch, Numpy, Pandas,  $\text{\LaTeX}$ , Scikit-learn, Tensorflow, Keras, Git, Android Programming, MySQL

# Interests

---

**Professional** Computer Vision, Artificial Intelligence, Robotics, Internet of Things, UI Design

**Personal** Technology, Gaming, Sport, Cinema

According to law 679/2016 of the Regulation of the European Parliament of 27th April 2016, I hereby express my consent to process and use my data provided in this document and application for recruiting purposes.