

Roberto Bigazzi

Ph.D in Computer Vision and Deep Learning

+39 334 8756245 | ✉ roberto.bigazzi.1995@gmail.com | 🏠 robertobigazzi.it | 🌐 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 Professor Rita Cucchiara. The main topics of my research are Computer Vision and Deep Learning, particularly on Embodied AI tasks and challenges. 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 Professor Marco Gribaudo. During my Ph.D. program, I spent a period as a visiting student researcher at Stanford University in the Autonomous Systems Lab (ASL) directed by Professor Marco Pavone.

Knowledge and Technical Skills

Topics	Visual Navigation, Multimodal Learning, Natural Language Processing, Reinforcement Learning, Deep Learning
Programming	Python, Java, Javascript, C++, C#, C, MATLAB, SQL, Microcontroller Programming (Arduino)
Frameworks and Other Languages	PyTorch, OpenCV, Numpy, Pandas, \LaTeX , Scikit-learn, Tensorflow, Keras, ROS, Git, Android Programming, MySQL
Languages	Italian (Mother tongue), English (Proficient <i>TOEIC (C1)</i> , <i>FCE (B2)</i>)

Education

University of Modena and Reggio Emilia

Ph.D. in Information and Communication Technologies: Computer Vision and Deep Learning

Modena, Italy

Nov. 2019 - Mar. 2023

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

Stanford University

Visiting Student Researcher

Stanford, California, United States

May 2022 - Aug. 2022

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

Polytechnic University of Milan

Master of Science in Computer Science and Engineering

Milan, Italy

Sep. 2017 - Oct. 2019

- Thesis with Prof. Marco Gribaudo: "Enhancing Spatial Navigation in Robot-Assisted Surgery: a case study"

Technische Universität Wien

Erasmus+ Exchange Semester

Vienna, Austria

Oct. 2018 - Feb. 2019

Polytechnic University of Milan

Bachelor of Science in Software Engineering

Milan, Italy

Sep. 2014 - Sep. 2017

Istituto Istruzione Superiore Roncalli

Liceo Scientifico Tecnologico

Poggibonsi (SI), Italy

Sep. 2009 - Jul. 2014

Experience

AlmageLab - University of Modena and Reggio Emilia

Postdoctoral Research Fellow

Modena, Italy

Mar. 2023 - present

Research on Visual Navigation and Multimodal Learning

AlmageLab - University of Modena and Reggio Emilia

Research Fellow

Modena, Italy

Nov. 2019 - Mar. 2023

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

Nuova Didactica

Lecturer

Modena, Italy

Mar. 2024 - Apr. 2024

Deep Learning Application course

Prometeia

Lecturer

Bologna, Italy

Nov. 2021

Python and Machine Learning course for Banco BPM

IFOA Bologna

Lecturer

Data Analysis and Data Visualization course

Bologna, Italy

Mar. 2021 - Apr. 2021

IFOA Modena

Lecturer

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

Modena, Italy

Jan. 2021 - Feb. 2021

Polytechnic University of Milan

Master Thesis

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.

Milan, Italy

Mar. 2019 - Sep. 2019

Cyber-Physical Systems Group - Technische Universität Wien

Project Developer

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.

Vienna, Austria

Oct. 2018 - Jan. 2019

Publications

2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*

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

Niyati Rawal, Roberto Bigazzi, Lorenzo Baraldi, Rita Cucchiara

Seattle, USA

In the last few years, the research interest in Vision-and-Language Navigation (VLN) has grown significantly. VLN is a challenging task that involves an agent following human instructions and navigating in a previously unknown environment to reach a specified goal. Recent work in literature focuses on different ways to augment the available datasets of instructions to improve navigation performance by exploiting synthetic training data. In this work, we propose AIGeN, a novel architecture inspired by Generative Adversarial Networks (GANs) that produces meaningful and well-formed synthetic instructions to improve navigation agents' performance. The model is composed of a Transformer decoder (GPT-2) and a Transformer encoder (BERT). During the training phase, the decoder generates sentences for a sequence of images describing the agent's path to a particular point while the encoder discriminates between real and fake instructions. Moreover, we devise an additional trajectory-matching loss term to improve instruction generation. Experimentally, we evaluate the quality of the generated instructions and perform extensive ablation studies. Additionally, we generate synthetic instructions for 217K trajectories using AIGeN on Habitat-Matterport 3D Dataset (HM3D) and show an improvement in the performance of an off-the-shelf VLN method. The validation analysis of our proposal is conducted on REVERIE and R2R and highlights the promising aspects of our proposal, achieving state-of-the-art performance.

2024 *International Conference on Robotics and Automation (ICRA)*, (Collaboration with Stanford University)

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

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

Yokohama, Japan

Robots require a semantic understanding of their surroundings to operate in an efficient and explainable way in human environments. In the literature, there has been an extensive focus on object labeling and exhaustive scene graph generation; less effort has been focused on the task of purely identifying and mapping large semantic regions. The present work proposes a method for semantic region mapping via embodied navigation in indoor environments, generating a high-level representation of the knowledge of the agent. To enable region identification, the method uses a vision-to-language model to provide scene information for mapping. By projecting egocentric scene understanding into the global frame, the proposed method generates a semantic map as a distribution over possible region labels at each location. This mapping procedure is paired with a trained navigation policy to enable autonomous map generation. The proposed method significantly outperforms a variety of baselines, including an object-based system and a pretrained scene classifier, in experiments in a photorealistic simulator.

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

London, UK

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.

2022 Robotics and Automation Letters (RA-L) + International Conference on Robotics and Automation (ICRA)

Focus on Impact: Indoor Exploration with Intrinsic Motivation

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

Philadelphia, USA

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

Montreal, Canada

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

Milan, Italy

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

2023 22st International Conference on Image Analysis and Processing (ICIAP), (Honorable Mention for Best Paper Award)

Towards Explainable Embodied Navigation and Recounting

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

Udine, Italy

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

Nicosia, Cyprus

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

Lecce, Italy

Under Review

Where is my Teddy Bear? Embodied Object-Goal Navigation Towards Personalized Targets

Roberto Bigazzi, Francesco Olivato, Niyati Rawal, Lorenzo Baraldi, Rita Cucchiara

Under Review

UNMuTe: Unifying Navigation and Multimodal Dialogue-like Text Generation

Niyati Rawal, Roberto Bigazzi, Lorenzo Baraldi, Rita Cucchiara

Program Committees

2023	Evaluator , ELLIS Ph.D. Program	
2021-now	Peer Reviewer , IEEE Robotics and Automation Letters (RA-L)	
2022-now	Peer Reviewer , IEEE Geoscience and Remote Sensing Letters (GRSL)	
2022-now	Peer Reviewer , IEEE Pattern Recognition Letters (PRL)	
2020	Peer Reviewer , Transactions on Multimedia Computing Communications and Applications (TOMM)	
2024	Peer Reviewer , European Conference on Computer Vision (ECCV)	<i>Milan, Italy</i>
2024	Peer Reviewer , IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)	<i>Seattle, USA</i>
2024	Peer Reviewer , ACM International Conference on Multimedia	<i>Melbourne, Aus</i>
2024	Peer Reviewer , IEEE International Conference on Robotics and Automation (ICRA)	<i>Yokohama, Japan</i>
2024	Peer Reviewer , IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	<i>Waikoloa, USA</i>
2023	Peer Reviewer , ACM International Conference on Multimedia	<i>Ottawa, Canada</i>
2023	Peer Reviewer , 22nd International Conference on Image Analysis and Processing (ICIAP)	<i>Udine, Italy</i>
2023	Peer Reviewer , IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)	<i>Detroit, USA</i>
2023	Peer Reviewer , IEEE International Conference on Robotics and Automation (ICRA)	<i>London, UK</i>
2022	Peer Reviewer , IAPR International Conference on Pattern Recognition (ICPR)	<i>Montreal, Canada</i>
2022	Peer Reviewer , 21st International Conference on Image Analysis and Processing (ICIAP)	<i>Lecce, Italy</i>
2021	Peer Reviewer , ACM International Conference on Multimedia	<i>Chengdu, China</i>
2020	Peer Reviewer , ACM International Conference on Multimedia	<i>Seattle, US (virtual)</i>

Teaching Activities

2020-23	Teaching Assistant , Computer Architectures
2021-22	Project Tutor , AI for Automotive
2021-22	Lecturer , AI for Automotive: SLAM Lessons
2020-21	Project Tutor , Neural Network Computing, AI and Machine Learning for Automotive
2019-20	Project Tutor , Computer Vision and Cognitive Systems

Presentations

2019-23	Presenter , "Into The Future" University Orientation Program
2022	Presenter , "Humanities & Intelligence" Laboratory
2021	Presenter , "In the mind of robots, how AI teaches robots to navigate" Laboratory

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	<i>Siena, Italy</i>
2017	Interactive Internet of Things e Smart Object Design Course , Polytechnic University of Milan	<i>Milan, Italy</i>
2013	European Computer Driving Licence Certificate , ECDL	

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.