CESI LINEACT

Title : Emergence of Contextual Awareness in Robotics through Social Learning for the Detection of Industrial Risks and Energy Inefficiencies in Workplace Settings

Domaines scientifiques : Artificial Intelligence, Robotics, Social Interactions

Mots clés : Social Robotics, Interactive Learning, Contextual Awareness

Supervisors

Thesis Supervisor :

Fabrice DUVAL, Associate professor

Co-supervisors :

Hakim GUEDJOU, Associate professor

Beatrice BIANCARDI, Associate professor

Research Work

Abstract

This thesis explores how a robot can develop contextual vigilance capabilities through social learning in interaction with humans. By observing and reproducing demonstrations, the robot learns to detect critical situations in its environment. The goal is to design a system capable of adapting to various contexts, such as energy efficiency or industrial safety. The approach relies on intuitive interactions to enable non-experts to transfer their knowledge to the robot. This work lies at the intersection of cognitive robotics, artificial intelligence, and human sciences.

The thesis

Scientific context

In a world where work environments are becoming increasingly complex and dynamic, robotic systems must go beyond basic perception capabilities and develop a more refined contextual understanding (Balažević *et al.*, 2023; Ni *et al.*, 2023). *The concept of contextual vigilance refers to an agent's ability to proactively detect risky or inefficient situations by considering the relationships between objects, human actions, and the environment. This capability is crucial for applications in industrial safety (e.g., detecting anomalies or hazardous configurations) as well as in energy-efficiency efforts (e.g., identifying waste due to active heating in a ventilated room).*

Social learning, and more specifically learning from demonstration (Argall *et al.*, 2009; Correia and Alexandre, 2024)*, emerges as a promising strategy for equipping robots with such skills. This approach enables non-expert users to intuitively transfer contextual knowledge to robots (knowledge that is often difficult to model explicitly) by directly demonstrating relevant situations* (Engelbracht *et al.*, 2024; Luo *et al.*, 2024). *Combined with advances in multimodal perception and adaptive learning, this method supports the development of autonomous robots capable of generalizing learned behaviors to diverse contexts, while remaining understandable and acceptable to humans.*

This line of research lies at the intersection of cognitive robotics, human-centered artificial intelligence, and behavioral sciences.

Subject

The primary objective of this thesis is to design a learning framework that enables a robot to develop contextual vigilance that is, the ability to autonomously and appropriately detect problematic situations in its environment. The originality of this approach lies in its use of social learning: the robot does not learn from pre-labeled datasets, but directly from human users through demonstrations and situated interactions (Engelbracht *et al.*, 2024; Luo *et al.*, 2024)*. These situations may span a wide range of domains, from industrial safety to energy efficiency, including issues of compliance or abnormal functioning.*

The scientific core of the thesis focuses on developing models capable of interpreting complex scenes by leveraging relevant contextual signals—not only the objects present, but also their spatial and functional relationships, associated usages, and the underlying intentions of humans (Arashpour, Ngo and Li, 2021; Balažević *et al.*, 2023; Dolatyabi, Regan and Khodayar, 2025)*. Learning from demonstration should enable the robot to encode flexible and generalizable representations, so that it can ultimately identify new occurrences of these situations without direct supervision. Special attention will be given to designing interaction modalities that are simple and accessible to non-expert users, allowing this learning capability to be integrated into real-world work settings.*

In a second phase, the robot will also learn—again through demonstration—corrective actions to respond to the situations it has identified. For this phase, existing state-of-the-art models in behavior learning by demonstration and gesture recognition will be employed (Billard *et al.*, 2008; Argall *et al.*, 2009; Correia and Alexandre, 2024)*, and adapted to the constraints of the given context. This second component will close the loop between detection, decision-making, and action, ensuring that the robot not only serves as an alert system but actively contributes to resolving anomalies or malfunctions.*

Finally, attention will be paid to the effect of this collaboration on the behaviors and motivations of the users themselves. Teaching the robot about situations and corrective actions may have a feedback effect on humans (Koh, Lee and Lim, 2018)*: enhancing attentiveness, improving adherence to best practices, or even shifting behaviors and motivations. These dynamics will be explored through the lens of cognitive science theories, such as self-perception theory* (Bem, 1972) *or cognitive dissonance* (Festinger, 1957)*, to understand how interaction with a learning robot could become a lever for individual—and eventually organizational—change.*

Prior works in the laboratory

This PhD project naturally builds on the ongoing work of the supervising research team, which has established expertise in the key domains involved. Previous research has notably focused on the generation of synthetic data for visual learning in robotics, particularly in complex industrial environments, providing a solid methodological foundation for tackling contextual recognition of problematic situations (Laignel *et al.*, 2024)*.*

The team has also explored mechanisms of adaptation in human-agent interaction, studying how to dynamically adjust the behavior of a socially interactive agent based on user reactions and preferences : an essential aspect for structuring effective pedagogical interactions during learning-from-demonstration phases (Biancardi, Dermouche and Pelachaud, 2021)*.*

Finally, special attention has already been given to the influence of individual user characteristics on learning dynamics in social robotics, which informs the current reflection on the mutual impact of human-robot interaction (Guedjou *et al.*, 2024)*.*

Taken together, this body of work provides a coherent and complementary scientific foundation that fully supports the ambition of the project: to foster contextual vigilance in robots through social learning, while also analyzing the effects of this learning relationship on the human counterpart.

Work program

The 36-month work plan is structured into two main phases, each organized around theoretical, technical, and experimental objectives. The first and longer phase (months 1 to 24) will focus on learning to detect problematic situations through demonstration. It will include an in-depth literature review on contextual perception models and learning from demonstration, a comparative analysis of models using public datasets, and the definition of experimental scenarios centered on contextual vigilance (energy performance and industrial safety). This phase will also involve the creation of a dedicated dataset, the development of a contextual perception model, the design of human-robot interaction interfaces, and an initial real-world experiment. The second phase (months 25 to 36) will focus on learning corrective gestures, drawing on and adapting state-of-the-art models to the identified use cases. A second experiment, focused on industrial safety, will also be conducted. Both phases will be accompanied by ongoing evaluation of the impact of the interaction on user behavior. The entire project will rely on the TIAGo robot as the experimental platform.

Production scientifique/technique attendue

The thesis aims to produce several scientific and technical contributions. On the theoretical level, it will propose an original framework for social learning applied to the contextual detection of problematic situations in robotics, including a formalization of generalization mechanisms based on human demonstrations. On the technical side, it will involve the development of a robotic system integrating perception, situation recognition, and gesture execution modules, all validated in real-world environments. Publications in international conferences and journals in robotics, AI, and human-machine interaction are expected, along with a functional demonstrator showcasing the robot's capabilities for contextual adaptation and interactive learning. An experimental analysis of the behavioral impact of interaction will further enrich the thesis's multidisciplinary results and may be submitted to international interdisciplinary journals. Finally, as part of the Pays de la Loire regional call for projects, the thesis will also contribute to a broader effort to promote open science within society. The PhD candidate will thus engage in science communication and public outreach activities aimed at a general audience.

Context

Lab presentation

CESI LINEACT (UR 7527), *Laboratory* for Digital Innovation for Businesses and Learning to Support the Competitiveness of Territories, anticipates and accompanies the technological mutations of sectors and services related to industry and construction. The historical proximity of CESI with companies is a determining element for our research activities. It has led us to focus our efforts on applied research close to companies and in partnership with them. A human-centered approach coupled with the use of technologies, as well as territorial networking and links with training, have enabled the construction of cross-cutting research; it puts humans, their needs and their uses, at the center of its issues and addresses the technological angle through these contributions.

Its research is *organized* according to two interdisciplinary scientific teams and several application areas.

- Team 1 "Learning and Innovating" mainly concerns Cognitive Sciences, Social Sciences and Management Sciences, Training Techniques and those of Innovation. The main scientific

objectives are the understanding of the effects of the environment, and more particularly of situations instrumented by technical objects (platforms, prototyping workshops, immersive systems...) on learning, creativity and innovation processes.

- Team 2 "Engineering and Digital Tools" mainly concerns Digital Sciences and Engineering. The main scientific objectives focus on modeling, simulation, optimization and data analysis of cyber physical systems. Research work also focuses on decision support tools and on the study of human-system interactions in particular through digital twins coupled with virtual or augmented environments.

These two teams develop and cross their research in application areas such as Industry 5.0, Construction 4.0 and Sustainable City, Digital Services. Areas supported by research platforms, mainly that in Rouen dedicated to Factory 5.0 and those in Nanterre dedicated to Factory 5.0 and Construction 4.0.

Links to the research axes of the research team involved

This PhD project is fully aligned with the strategic priorities of the CESI Lineact research laboratory. By developing robots capable of learning to detect problematic situations and respond to them through interaction with human operators, the thesis aims to contribute to the transformation of work environments in industry, smart buildings, and urban services. It proposes an innovative approach to embedded and interactive learning, offering concrete opportunities to enhance performance, safety, and organizational sustainability. Moreover, by promoting active collaboration between humans and machines, the project also addresses the challenges of future education by placing users in an active role as knowledge transmitters, while simultaneously fostering their own skill development. This cross-disciplinary positioning—at the intersection of robotics, cognitive and human sciences, and digital applications—makes the project a relevant lever to support the digital transformation of regions, driving both technological innovation and the evolution of professional practices.

Organisation

Funding : *CESI, Pays de la Loire Region*

Location : *Saint Nazaire*

Starting date : *October 2025*

Durée : *3 years*

CESI LINEACT

Your Hiring Process

Application Procedure : based on application file and interview.

Please send your application to fduval@cesi.fr, hguedjou@cesi.fr, bbiancardi@cesi.fr with the subject line: :

« [Application] Emergence of Contextual Awareness in Robotics through Social Learning for the Detection of Industrial Risks and Energy Inefficiencies in Workplace Settings »

Your application must include :

- **A detailed Curriculum Vitae**. If there are any gaps in your academic background, please provide an explanation ;
- **A cover letter** explaining your motivation for pursuing a doctoral thesis ;
- **Academic transcripts** for Master 2, including corresponding grade reports ;
- Any other document you consider relevant.

*Please submit all documents in a single **.zip** file named: **LASTNAME_Firstname.zip**.*

Skills :

Scientific and technical skills :

- *Background in machine learning (experience in learning from demonstration would be an asset)*
- *Good understanding of supervised learning principles, with initial experience developing models using Python (PyTorch or TensorFlow)*
- *Basic understanding of computer vision and robotic perception (e.g., image processing, object detection, or segmentation)*
- *Interest in human-robot interaction*
- *Awareness of challenges in human-machine collaboration, with the ability to design or evaluate simple interactive behaviors or interfaces*
- *Curiosity about cognitive or behavioral sciences as applied to robotics*
- *Motivation for applied projects involving the development, implementation, and evaluation of user-interactive systems, with attention to detail in data collection and analysis*

Soft skills :

- Ability to work independently, with a proactive and curious mindset
- Strong teamwork and interpersonal skills
- Attention to detail and a rigorous approach to work

Bibliography :

Arashpour, M., Ngo, T. and Li, H. (2021) 'Scene understanding in construction and buildings using image processing methods: A comprehensive review and a case study', *Journal of Building Engineering*, 33, p. 101672. Available at: https://doi.org/10.1016/j.jobe.2020.101672.

Argall, B.D. *et al.* (2009) 'A survey of robot learning from demonstration', *Robotics and Autonomous Systems*, 57(5), pp. 469–483. Available at: https://doi.org/10.1016/j.robot.2008.10.024.

Balažević, I. *et al.* (2023) 'Towards In-context Scene Understanding'. arXiv. Available at: https://doi.org/10.48550/arXiv.2306.01667.

Bem, D.J. (1972) 'Self-Perception Theory1', in L. Berkowitz (ed.) *Advances in Experimental Social Psychology*. Academic Press, pp. 1–62. Available at: https://doi.org/10.1016/S0065-2601(08)60024-6.

Biancardi, B., Dermouche, S. and Pelachaud, C. (2021) 'Adaptation Mechanisms in Human–Agent Interaction: Effects on User's Impressions and Engagement', *Frontiers in Computer Science*, 3. Available at: https://doi.org/10.3389/fcomp.2021.696682.

Billard, A. *et al.* (2008) 'Robot Programming by Demonstration', in B. Siciliano and O. Khatib (eds) *Springer Handbook of Robotics*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1371–1394. Available at: https://doi.org/10.1007/978-3-540-30301-5_60.

Correia, A. and Alexandre, L.A. (2024) 'A survey of demonstration learning', *Robotics and Autonomous Systems*, 182, p. 104812. Available at: https://doi.org/10.1016/j.robot.2024.104812.

Dolatyabi, P., Regan, J. and Khodayar, M. (2025) 'Deep Learning for Traffic Scene Understanding: A Review', *IEEE Access*, 13, pp. 13187–13237. Available at: https://doi.org/10.1109/ACCESS.2025.3529289.

Engelbracht, T. *et al.* (2024) 'SpotLight: Robotic Scene Understanding through Interaction and Affordance Detection'. arXiv. Available at: https://doi.org/10.48550/arXiv.2409.11870.

Festinger, L. (1957) *A theory of cognitive dissonance*. Stanford University Press (A theory of cognitive dissonance), pp. xi, 291.

Guedjou, H. *et al.* (2024) 'The Influence of Extraversion on a Robot Developmental Learning in a Human Robot Interaction', in *2024 IEEE International Conference on Development and Learning (ICDL). 2024 IEEE International Conference on Development and Learning (ICDL)*, pp. 1–6. Available at: https://doi.org/10.1109/ICDL61372.2024.10644954.

Koh, A.W.L., Lee, S.C. and Lim, S.W.H. (2018) 'The learning benefits of teaching: A retrieval practice hypothesis', *Applied Cognitive Psychology*, 32(3), pp. 401–410. Available at: https://doi.org/10.1002/acp.3410.

Laignel, A. *et al.* (2024) 'Synthetic datasets for 6D Pose Estimation of Industrial Objects: Framework, Benchmark and Guidelines', in *2024 The 11th International Conference on Industrial Engineering and Applications - proceedings of 2024 The 5th International Conference on Industrial Engineering and Industrial Management (IEIM 2024)*. Nice, France. Available at: https://hal.science/hal-04389164 (Accessed: 11 March 2025).

Luo, H. *et al.* (2024) 'Learning Visual Affordance Grounding From Demonstration Videos', *IEEE Transactions on Neural Networks and Learning Systems*, 35(11), pp. 16857–16871. Available at: https://doi.org/10.1109/TNNLS.2023.3298638.

Ni, J. *et al.* (2023) 'Deep learning-based scene understanding for autonomous robots: a survey', *Intelligence & Robotics*, 3, pp. 374–401. Available at: https://doi.org/10.20517/ir.2023.22.