Safe Robot Navigation in Indoor Healthcare Spaces

E. Toulkeridou, A. Kourris, E. G. Christoforou, R. Julia Ros, M. Bosch, R. Lopez, A. Perrot, A. Godart, N. Ramdani, C. S. Pattichis, *Fellow, IEEE,* A. S. Panayides, *Senior Member, IEEE*

Abstract— **This study proposes a safe robot navigation in healthcare spaces methodology and aids towards the wider application of studies on Human-Robot Interaction in less structured environments. Using state-of-the-art computer vision methods, the robots should be able to detect the presence and identify the type of dynamic obstacles inside their visual field and adapt their navigation accordingly.**

I. Introduction

Human-Robot Interaction (HRI) is the field of study that explores the use of robotic systems by humans; moreover, it investigates, understands, and evaluates the communication between robots and humans [1]. As new robotic applications emerge, HRI research interest is growing, driven by an ever-increasing social impact. A key application which has seen widespread utilisation over the past decades is robot adaptation to large industrial and warehouse spaces [1]. Yet, this is not the case for the use of robots in hospital and commercial spaces, whose less structured environments require robots that are able to adapt to crowds, and that are aware of how to avoid deadlocks [2]. Moreover, the robots should be able to conduct human-aware navigation and interact with other robots while overcoming the challenge of untrained healthcare personnel.

A key objective for such applications is the training of robots for the detection of dynamic obstacles using computer vision methods. The latter is a non-trivial process, due to the wide selection of established and emerging deep learning architectures; the most suitable method needs to adapt to the specific task, producing a robust human-aware algorithm. Thus, the robots should be able to recognise other robots and humans alike in their vicinity, predict their motions in real time, and adjust their movement accordingly.

II. Methodology

A. Computer Vision for Human and Object Detection

Human and object recognition (including robots) research has gained significant attention from the community, due to its widespread usage across diverse industry setups and applications, ranging from autonomous self-driving cars to drone monitoring and hospital spaces navigation. The goal is to train the robots to recognise humans and other robots in their visual field using embedded or 3rd party camera sensors. In this study, we opted for YOLO (You Only Look Once) deep learning model, since it can perform fast and accurate real-time object detection, using a single forward propagation though the neural network. YOLO scans the video feed as a series of sequential images, and for each one, it provides object identification, localisation, and categorization. More specifically, YOLO splits the image into an S×S grid, and for each cell in the grid it predicts multiple bounding boxes per object of interest. Additionally, for each box it predicts confidence scores and a class probability. As it is based on regression, the pipeline is quite simple and hence computationally efficient for real-time performance.

B. Autonomous Navigation

In the proposed approach and system setup, autonomous mobile robots are controlled with the open-source Robot Operating System (ROS). A newly developed Command Manager executable allows the user to send on the ROS network complex commands using simple string-encoded messages and read the feedbacks. ROS topics are used for node communication and for publishing and reading command messages and feedback. With a fleet of robots endowed with a centralised management system, for inter-robot communication we use the MQTT protocol, which is the standard messaging protocol used for Internet of Things (IoT) setups. A dedicated ROS node provides functionality to bidirectional bridge between the MQTT and ROS messages.

III. Ongoing and Future Work

YOLOv5 PyTorch model using the MS COCO dataset was employed to pre-train our network and for hyper parameters fine-tuning [3]. Then, a custom annotated dataset, extending the MS COCO dataset categories to identify mobile robots was used for training. Ongoing work includes comparing our YOLOv5 approach with the new YOLOv7 model [4]. Future works entails investigating how video properties such as video resolution and frame rate affect prediction accuracy as well as energy consumption for edge processing applications.

References

[1] Goodrich M.A. and Schultz A.C., "Human-Robot Interaction: A Survey," *Foundations and Trends in Human-Computer Interaction*, 2017, vol. 1 no.3 , pp. 203-275.

[2] N. Ramdani *et al.*, "A Safe, Efficient and Integrated Indoor Robotic Fleet for Logistic Applications in Healthcare and Commercial Spaces: The ENDORSE Concept," *2019 20th IEEE MDM*, 2019, pp. 425-430.

[3] Sharma A., “Training the YOLOv5 Object Detector on a Custom Dataset,” PyImageSearch, D. Chakraborty, et al., eds., 2022.

[4] Wang C-Y. et. al., “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” *arXiv*:2207.02696, 2022.

* This work was funded from the European Union’s Horizon 2020 MSCA RISE Action under grant agreement No. 823887.

E. Toulkeridou, E. G. Christoforou and C. S. Pattichis are with the University of Cyprus, Nicosia, Cyprus (e-mail: toulkeridou.evropi@ucy.ac.cy).

R. Julia Ros, M. Bosch, and R. Lopez are with the R&D Department, Robotnik Automation, Spain (e-mail: rjuliaa@robotnik.es).

A. Perrot, A. Godart, and N. Ramdani are with the PRISME Laboratory, Université d'Orléans, Bourges, France (e-mail: nacim.ramdani@univ-orleans.fr).

A. S. Panayides and A. Kourris are with the VIDEOMICS group, CYENS Centre of Excellence, 1016 Nicosia, Cyprus (phone: +357 22 747575; e-mail: a.panayides@cyens.org.cy).