Ana I. Pereira · Armando Mendes ·
Florbela P. Fernandes · Maria F. Pacheco ·
João P. Coelho · José Lima (Eds.)

Communications in Computer and Information Science 1982

Optimization, Learning Algorithms and Applications

Third International Conference, OL2A 2023
Ponta Delgada, Portugal, September 27–29, 2023
Revised Selected Papers, Part II

Part 2

Springer

Ana I. Pereira · Armando Mendes ·
Florbela P. Fernandes · Maria F. Pacheco ·
João P. Coelho · José Lima
Editors

Optimization, Learning Algorithms and Applications

Third International Conference, OL2A 2023
Ponta Delgada, Portugal, September 27–29, 2023
Revised Selected Papers, Part II

Springer

Editors
Ana I. Pereira
Instituto Politécnico de Bragança
Bragança, Portugal

Florbela P. Fernandes
Instituto Politécnico de Bragança
Bragança, Portugal

João P. Coelho
Instituto Politécnico de Bragança
Bragança, Portugal

Armando Mendes
University of Azores
Ponta Delgada, Portugal

Maria F. Pacheco
Instituto Politécnico de Bragança
Bragança, Portugal

José Lima
Instituto Politécnico de Bragança
Bragança, Portugal

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-53035-7 ISBN 978-3-031-53036-4 (eBook)
https://doi.org/10.1007/978-3-031-53036-4

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2024
Chapters 13 and 23 are licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). For further details see license information in the chapters.

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

The volumes CCIS 1981 and 1982 contains the refereed proceedings of the III International Conference on Optimization, Learning Algorithms and Applications (OL2A 2023), a hybrid event held on September 27–29.

OL2A provided a space for the research community in optimization and learning to get together and share the latest developments, trends and techniques as well as develop new paths and collaborations. OL2A had the participation of more than four hundred participants in an online and face-to-face environment throughout three days, discussing topics associated with areas such as optimization and learning and state-of-the-art applications related to multi-objective optimization, optimization for machine learning, robotics, health informatics, data analysis, optimization and learning under uncertainty and 4th industrial revolution.

Six special sessions were organized under the topics Learning Algorithms in Engineering Education, Optimization in the SDG context, Optimization in Control Systems Design, Computer Vision Based on Learning Algorithms, Machine Learning and AI in Robotics and Machine Learning and Data Analysis in Internet of Things. The event had 66 accepted papers. All papers were carefully reviewed and selected from 172 submissions. All the reviews were carefully carried out by a scientific committee of 115 PhD researchers from 23 countries.

The OL2A 2023 volume editors,

September 2023

Ana I. Pereira
Armando Mendes
Florbela P. Fernandes
Maria F. Pacheco
João P. Coelho
José Lima

Organization

General Chairs

Ana I. Pereira	Polytechnic Institute of Bragança, Portugal
Armando Mendes	University of the Azores, Portugal

Program Committee Chairs

Florbela P. Fernandes	Polytechnic Institute of Bragança, Portugal
M. Fátima Pacheco	Polytechnic Institute of Bragança, Portugal
João P. Coelho	Polytechnic Institute of Bragança, Portugal
José Lima	Polytechnic Institute of Bragança, Portugal

Special Session Chairs

João P. Teixeira	Polytechnic Institute of Bragança, Portugal
José Cascalho	University of the Azores, Portugal

Technology Chairs

Paulo Medeiros	University of the Azores, Portugal
Rui Pedro Lopes	Polytechnic Institute of Bragança, Portugal

Program Committee

Ana Isabel Pereira	Polytechnic Institute of Bragança, Portugal
Abeer Alsadoon	Charles Sturt University, Australia
Ala’ Khalifeh	German Jordanian University, Jordan
Alberto Nakano	Federal University of Technology – Paraná, Brazil
Alexandre Douplik	Ryerson University, Canada
Ana Maria A. C. Rocha	University of Minho, Portugal
Ana Paula Teixeira	University of Trás-os-Montes and Alto Douro, Portugal
André Pinz Borges	Federal University of Technology – Paraná, Brazil

André Rodrigues da Cruz	Federal Center for Technological Education of Minas Gerais, Brazil
Andrej Košir	University of Ljubljana, Slovenia
António José Sánchez-Salmerón	Universitat Politècnica de València, Spain
António Valente	University of Trás-os-Montes and Alto Douro, Portugal
Armando Mendes	University of the Azores, Portugal
Arnaldo Cândido Júnior	Federal Technological University – Paraná, Brazil
B. Rajesh Kanna	Vellore Institute of Technology, India
Bilal Ahmad	University of Warwick, UK
Bruno Bispo	Federal University of Santa Catarina, Brazil
C. Sweetlin Hemalatha	Vellore Institute of Technology, India
Carlos Henrique Alves	CEFET - Rio de Janeiro, Brazil
Carmen Galé	University of Zaragoza, Spain
Carolina Gil Marcelino	Federal University of Rio de Janeiro, Brazil
Christopher Expósito Izquierdo	University of Laguna, Spain
Clara Vaz	Polytechnic Institute of Bragança, Portugal
Damir Vrančić	Jožef Stefan Institute, Slovenia
Dhiah Abou-Tair	German Jordanian University, Jordan
Diamantino Silva Freitas	University of Porto, Portugal
Diego Brandão	CEFET - Rio de Janeiro, Brazil
Dimitris Glotsos	University of West Attica, Greece
Eduardo Vinicius Kuhn	Federal Technological University – Paraná, Brazil
Elaine Mosconi	Université de Sherbrooke, Canada
Eligius M. T. Hendrix	Malaga University, Spain
Elizabeth Fialho Wanner	Federal Center for Technological Education of Minas Gerais, Brazil
Felipe Nascimento Martins	Hanze University of Applied Sciences, The Netherlands
Florbela P. Fernandes	Polytechnic Institute of Bragança, Portugal
Florentino Fernández Riverola	University of Vigo, Spain
Francisco Sedano	University of León, Spain
Fredrik Danielsson	University West, Sweden
Gaukhar Muratova	Dulaty University, Kazakhstan
Gediminas Daukšys	Kauno Technikos Kolegija, Lithuania
Gianluigi Ferrari	University of Parma, Italy
Glaucia Maria Bressan	Federal University of Technology – Paraná, Brazil
Glotsos Dimitris	University of West Attica, Greece
Humberto Rocha	University of Coimbra, Portugal
João Paulo Carmo	University of São Paulo, Brazil
João Paulo Coelho	Polytechnic Institute of Bragança, Portugal
João Paulo Teixeira	Polytechnic Institute of Bragança, Portugal

Jorge Igual	Universitat Politécnica de Valencia, Spain
Jorge Ribeiro	Polytechnic Institute of Viana do Castelo, Portugal
José Boaventura-Cunha	University of Trás-os-Montes and Alto Douro, Portugal
José Cascalho	University of the Azores, Portugal
José Lima	Polytechnic Institute of Bragança, Portugal
José Ramos	Nova University Lisbon, Portugal
Joseane Pontes	Federal University of Technology – Ponta Grossa, Brazil
Josip Musić	University of Split, Croatia
Juan A. Méndez Pérez	University of Laguna, Spain
Juan Alberto García Esteban	University de Salamanca, Spain
Júlio Cesar Nievola	Pontifícia Universidade Católica do Paraná, Brazil
Kristina Sutiene	Kaunas University of Technology, Lithuania
Laura Belli	University of Parma, Italy
Lidia Sánchez	University of León, Spain
Lino Costa	University of Minho, Portugal
Luca Davoli	University of Parma, Italy
Luca Oneto	University of Genoa, Italy
Luca Spalazzi	Marche Polytechnical University, Italy
Luis Antonio De Santa-Eulalia	Université de Sherbrooke, Canada
Luís Coelho	Polytechnic Institute of Porto, Portugal
M. Fátima Pacheco	Polytechnic Institute of Bragança, Portugal
Mahmood Reza Khabbazi	University West, Sweden
Manuel Castejón Limas	University of León, Spain
Marc Jungers	Université de Lorraine, France
Marco Aurélio Wehrmeister	Federal University of Technology – Paraná, Brazil
Marek Nowakowski	Military Institute of Armoured and Automotive Technology in Sulejowek, Poland
Maria do Rosário de Pinho	University of Porto, Portugal
Martin Hering-Bertram	Hochschule Bremen, Germany
Matthias Funk	University of the Azores, Portugal
Mattias Bennulf	University West, Sweden
Michał Podpora	Opole University of Technology, Poland
Miguel Ángel Prada	University of León, Spain
Mikulas Huba	Slovak University of Technology in Bratislava, Slovakia
Milena Pinto	Federal Center of Technological Education Celso Suckow da Fonseca, Brazil
Miroslav Kulich	Czech Technical University Prague, Czech Republic
Nicolae Cleju	Technical University of Iasi, Romania

Paulo Alves	Polytechnic Institute of Bragança, Portugal
Paulo Leitão	Polytechnic Institute of Bragança, Portugal
Paulo Lopes dos Santos	University of Porto, Portugal
Paulo Medeiros	University of the Azores, Portugal
Paulo Moura Oliveira	University of Trás-os-Montes and Alto Douro, Portugal
Pavel Pakshin	Nizhny Novgorod State Tech University, Russia
Pedro Luiz de Paula Filho	Federal Technological University – Paraná, Brazil
Pedro Miguel Rodrigues	Catholic University of Portugal, Portugal
Pedro Morais	Polytechnic Institute of Cávado e Ave, Portugal
Pedro Pinto	Polytechnic Institute of Viana do Castelo, Portugal
Roberto Molina de Souza	Federal University of Technology – Paraná, Brazil
Rui Pedro Lopes	Polytechnic Institute of Bragança, Portugal
Sabrina Šuman	Polytechnic of Rijeka, Croatia
Sancho Salcedo Sanz	Alcalá University, Spain
Sandro Dias	Federal Center for Technological Education of Minas Gerais, Brazil
Sani Rutz da Silva	Federal Technological University – Paraná, Brazil
Santiago Torres Álvarez	University of Laguna, Spain
Sara Paiva	Polytechnic Institute of Viana do Castelo, Portugal
Shridhar Devamane	Global Academy of Technology, India
Sławomir Stępień	Poznań University of Technology, Poland
Sofia Rodrigues	Polytechnic Institute of Viana do Castelo, Portugal
Sudha Ramasamy	University West, Sweden
Teresa Paula Perdicoulis	University of Trás-os-Montes and Alto Douro, Portugal
Toma Rancevic	University of Split, Croatia
Uta Bohnebeck	Hochschule Bremen, Germany
Virginia Castillo	University of León, Spain
Vítor Duarte dos Santos	Nova University Lisbon, Portugal
Vitor Pinto	University of Porto, Portugal
Vivian Cremer Kalempa	State University of Santa Catarina, Brazil
Wojciech Giernacki	Poznań University of Technology, Poland
Wojciech Paszke	University of Zielona Gora, Poland
Wynand Alkema	Hanze University of Applied Sciences, The Netherlands
Zahia Guessoum	University of Reims Champagne-Ardenne, France

Using LiDAR Data as Image for AI to Recognize Objects in the Mobile Robot Operational Environment

Marek Nowakowski[1](✉), Jakub Kurylo[6], João Braun[2,3,4], Guido S. Berger[2,4,5], João Mendes[2,4], and José Lima[2,3,4]

[1] Military Institute of Armoured and Automotive Technology, Okuniewska 1, 05-070 Sulejówek, Poland
marek.nowakowski@witpis.eu

[2] Research Centre in Digitalization and Intelligent Robotics (CeDRI) - Instituto Politécnico de Braganca, Campus de Santa Apolónia, Bragança, Portugal
{jbneto,guido.berger,jllima}@ipb.pt

[3] INESC Technology and Science, Porto, Portugal

[4] Laboratory for Sustainability and Technology in Mountain Regions (SusTEC) - Instituto Politécnico de Bragança, Bragança, Portugal

[5] Engineering Department, School of Sciences and Technology, Universidade de Trás-os-Montes e Alto Douro (UTAD), 5000-801 Vila Real, Portugal

[6] Bialystok University of Technology, 45A, Wiejska Street, 15-351 Bialystok, Poland

Abstract. Nowadays, there has been a growing interest in the use of mobile robots for various applications, where the analysis of the operational environment is a crucial component to conduct our special tasks or missions. The main aim of this work was to implement artificial intelligence (AI) for object detection and distance estimation navigating the developed unmanned platform in unknown environments. Conventional approaches are based on vision systems analysis using neural networks for object detection, classification, and distance estimation. Unfortunately, in the case of precise operation, the used algorithms do not provide accurate data required by platforms operators as well as autonomy subsystems. To overcome this limitation, the authors propose a novel approach using the spatial data from laser scanners supplementing the acquisition of precise information about the detected object distance in the operational environment.

In this article, we introduced the application of pretrained neural network models, typically used for vision systems, in analysing flat distributions of LiDAR point cloud surfaces. To achieve our goal, we have developed software that fuses detection algorithm (based on YOLO network) to detect objects and estimate their distances using the MiDaS depth model. Initially, the accuracy of distance estimation was evaluated through video stream testing in various scenarios. Furthermore, we have incorporated data from a laser scanner into the software, enabling precise distance measurements of the detected objects.

The paper provides discussion on conducted experiments, obtained results, and implementation to improve performance of the described modular mobile platform.

Keywords: Convolutional Neural Network · Depth Estimation · Point Clouds

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
A. I. Pereira et al. (Eds.): OL2A 2023, CCIS 1982, pp. 118–131, 2024.
https://doi.org/10.1007/978-3-031-53036-4_9

1 Introduction

Mobile robots have emerged as versatile tools that can be utilized for a wide range of applications, including internal logistics, security operations, reconnaissance, special-purpose field operations, visual inspections as well as demining tasks [1–3]. In conflict-affected regions like Ukraine, the presence of landmines and unexploded ordnance poses a significant threat to civilian populations and hinders post-conflict recovery efforts. The clearance of these hazardous remnants of war is a critical undertaking that requires careful planning, specialized equipment, and skilled personnel. Mobile robots equipped with advanced perception systems have emerged as valuable assets in demining operations, enhancing safety and efficiency while reducing the risk to human lives.

In addition to demining requirements in open outdoor areas, the detection of explosive materials remains a crucial challenge, especially in crowded buildings such as airports, where the persistent threat of terrorism is a concern. Ensuring public safety requires the integration of advanced technologies based on mobile platforms, equipped with innovative navigation systems, to effectively operate in complex indoor environments as well as using some additional devices [4]. Advanced perception systems, such as vision-based object recognition and depth estimation, empower mobile robots to precisely identify and locate suspicious objects.

The versatility of mobile robots in various environments underscores the imperative of their adoption in security operations. Different configurations and constructions have been developed including wheeled, tracked, or legged platforms as well as aerial and underwater unmanned vehicles [5, 6]. Each type of platform offers unique advantages for specific applications and environments. The choice of platform structure depends on factors such as terrain, mobility requirements, payload capacity, and the tasks the robot needs to perform. Researchers and engineers continue to explore and develop innovative platform designs to meet the diverse needs of mobile robotic systems. Moreover, the selection of an appropriate perception system, as well as efficient object analysis and classification algorithms, plays a significant role in enabling accurate and timely decision-making during missions. Certain tasks and missions may demand autonomous modes of operation for mobile robots, which highlights the need for the development of robust navigation algorithms. In the existing literature, numerous methods have been extensively studied and documented, focusing on the integration of vision systems and active sensors such as 3D LiDARs [7, 8]. While vision systems remain a principal component for object recognition and depth estimation, they are susceptible to adverse weather conditions commonly encountered in outdoor environments, such as dust, heavy rain, and fog [9]. These conditions can degrade the accuracy and reliability of vision-based distance estimation. LiDARs offer high-resolution scans and precise distance measurements, making them an attractive alternative.

In this article, the importance of advanced perception systems for mobile robots in special-purpose applications is highlighted. The selection of a suitable perception system, considering environmental conditions and mission objectives is described. Object analysis and classification techniques are introduced. The paper investigates the usage of vision systems for object recognition and depth estimation, discussing their advantages and limitations in different operational environments. Additionally, developed software was examined using pretrained neural network models for analysing flat distributions of LiDAR point cloud surfaces to precise measuring distance from detected object. The conducted experiments, presenting the obtained results and exploring potential implementations for scenarios requiring accurate spatial in special-purpose applications are discussed.

2 Described of Developed Mobile Robot Platform

The developed wheeled mobile robot represents a versatile solution that can be effectively deployed for a wide range of tasks in both indoor and outdoor environments as shown in Fig. 1. Its modular design allows the integration of different modules and components according to operational requirements. Developed architecture enables users to control the robot from a safe distance while also benefiting from its ability to perform tasks autonomously according to advanced defined paths.

a) b)

Fig. 1. View of developed mobile robot platform (a) and control station (b).

The mobile platform is equipped with key modules including electric drives, batteries, and control panels, as well as electric energy distribution circuits. These modules provide the necessary power and control systems for the robot's movement and overall operation. Their robust design ensures reliable performance in challenging outdoor environments.

The remote operation system is equipped with radio communication modules allowing to transfer control data and vision stream from installed cameras. The onboard computer processes the data received from the robot's sensors, performs advanced algorithms for decision-making, and enables real-time operation (Fig. 2). The radio communication transceivers facilitate reliable and efficient command transmission between the robot and the control station.

The developed platform is suitable for optional modules that can be installed exclusively on the specific mission requirements. One of the proposed integrations (Fig. 3) is a neutron-based explosives detector module that can facilitate demining operations, allowing the robot to identify explosive materials in the environment [10]. Alternatively,

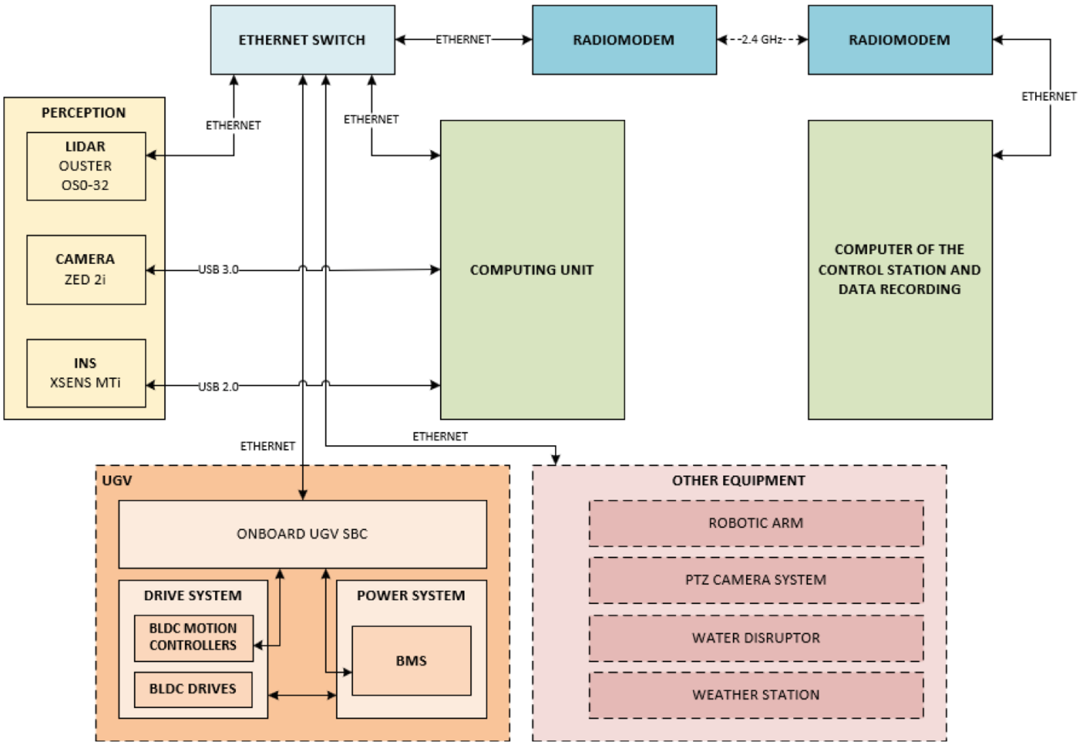


Fig. 2. The modular architecture of the developed mobile robot system.

an observation head module can be employed for reconnaissance purposes, providing enhanced visual capabilities and intelligence gathering capabilities.

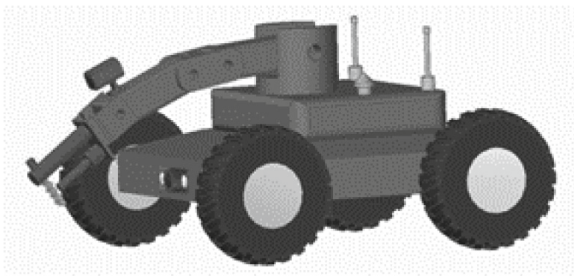


Fig. 3. Integration example of the robot equipped with the neutron-based detector [10].

The control station, equipped with a user application as the interface between the operator and the mobile robot. It enables operators to monitor and control the robot's actions, access real-time data, and interact with the robot's functionalities. The user application provides a user-friendly interface that facilitates efficient command and control operations, allowing operators to make informed decisions and adjust the robot's behaviour as needed.

Critical to the robot's operational capabilities is the perception system, which plays a crucial role in environment analysis. The basic system is based on vision and incorporates distinct types of cameras such as stereo, daylight, or thermal, along with optional active sensors like LiDARs or radars. Based on captured data from sensors, the robot can detect objects, estimate distances, and create a detailed understanding of its environment. Perception system allows the robot to navigate through complex layouts, avoid obstacles, and make informed decisions, ensuring safe and efficient operation in various operational environments, both indoors and outdoors. Despite hardware it is required to integrate software algorithms and tools to gather information about objects, obstacles, vehicles, pedestrians, buildings, and more.

a) b)

Fig. 4. View of ZED 2 stereo camera (a) and IP cameras (b).

The selection of a vision system for a mobile robot depends on its specific needs and operating conditions, with options ranging from monocular cameras for basic obstacle detection to stereo and RGB-D cameras for enhanced depth perception and environmental analysis. In our mobile platform, we commonly use the ZED 2 stereo camera (Fig. 4) from Stereolabs which has two 4K resolution sensors that capture images from different perspectives with a maximum resolution of 3840×2160 pixels for colour and 2560×720 pixels for depth [11]. In the case of advanced missions like reconnaissance, the integration of the Mobile Surveillance Head can enhance the mobile robot's capabilities. This solution incorporates high-resolution cameras, advanced video analytics, and real-time monitoring, enabling detailed imagery, intelligent object detection, and tracking for enhanced situational awareness and reconnaissance tasks.

Navigation in outdoor environment requires localization technologies that combines GNSS signal and Aided Inertial Navigation System (INS) in conjunction with fitted external wheel speed sensors to provide precise position estimation and motion tracking capabilities. The GNSS signal, received from satellite constellations such as GPS, GLONASS, or Galileo, allows the mobile platform to determine its global position with high accuracy. Processing signals from multiple satellites, the robot can calculate its coordinates and align itself with a global reference system to follow predefined paths or reach specific waypoints. In situations where GNSS signal reception may be disturbed, such as in urban canyons or dense foliage, the Aided Inertial Navigation System (INS) are used to estimate the robot's position, velocity, and orientation using data from accelerometers and gyroscopes.

Unknown environments pose significant challenges for mobile platforms navigation, requiring systems that can accurately perceive and map the surroundings. In some conditions vision-based systems have limitations due to adverse weather conditions like rain, snow, or dust significantly reducing reliable perception and obstacle detection. To overcome these limitations, it is necessary to integrate 3D sensors like LiDARs that emit pulses and utilize time-of-flight measurements [12]. Laser scanners can accurately capture detailed information about the environment, including object distances and positions.

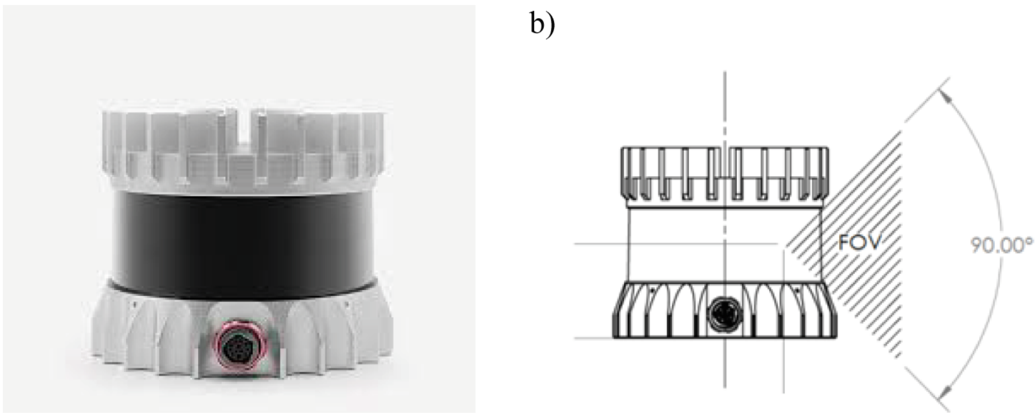
a)
b)


Fig. 5. View of Ouster LiDAR (a) and indicated field of view (b) [17].

The optical signal from the laser is modulated to carry information about the surrounding environment. Analysing the reflected signals, the LiDAR system can gather information not only about distance but also reflectivity and other properties of objects and surfaces in the environment. Combining the distance information from multiple laser pulses, the system can build a real-time 3D map of the environment. This map provides detailed spatial information, allowing the robot to operate effectively in various lighting conditions and weather. LiDARs are not affected by factors such as darkness or precipitation, enabling reliable object detection at a long range. This capability provides early warnings to the path-planning system, enabling the robot to adapt its trajectory accordingly.

In the architecture of developed mobile platform, different type of Ouster LiDARs (Fig. 5) are used due to long-range capabilities and high-resolution imaging resulting in a precise and dense point cloud representation of the surrounding environment. The sensor data is seamlessly integrated into the mobile platform's perception and decision-making systems in real-time through Ethernet communication, using UDP packets (Table 1).

Table 1. Main parameters of used Ouster LiDARs.

Parameter	Ouster OS0-32	Ouster OS1-128
Channels	32	128
Range accuracy	±1.5–5 cm	±0.7–5 cm
Field of view (vertical)	90°	45°
Angular resolution (vertical)	2.125°	0.35°
Points per second	655,360	2,621,440

This integration of the described perception system enables the mobile platform to operate in both indoor and outdoor environments. Remote control and optional autonomous functions can be used for a wide range of tasks, including security operations, reconnaissance, demining, and more. Its modular structure and adaptable sensor and module integration make it a powerful and flexible tool for various mission requirements and operational scenarios.

3 Vision Sensors and LiDARs Comparison

Accurate environmental perception is essential for enabling mobile robots to navigate their surroundings effectively. In the context of autonomous operation, it is important to understand the differences between vision systems, which capture 2D images, and lidars, which generate a 3D point cloud offering complementary information about detected objects.

Data collected from cameras in the form of images or videos can be processed using computer vision techniques to extract valuable information. Vision systems can provide detailed information about the appearance and characteristics of objects in the environment. The following methods are commonly employed for analysing image data [13]:

- Image Segmentation:
  Image segmentation divides an image into multiple segments, where each segment corresponds to a different object in the scene. Techniques such as thresholding, edge detection, or region growing can be used to perform image segmentation. Segments can then be classified as obstacles or other objects of interest. Authors in [18] do an extensive explanation on the subject.
- Object Detection:
  Object detection involves detecting instances of objects within an image, such as vehicles, pedestrians, or buildings. Deep learning-based object detection algorithms, such as Faster R-CNN or YOLO, are commonly used for this task. These algorithms can learn to detect objects of interest by training on large datasets. The innovation in this area is recent and long. For a better understanding, the reader is referred to [19].
- Deep Learning-Based Semantic Segmentation:
  Semantic segmentation assigns a label to each pixel in an image, such as "car," "building," or "obstacle." Deep learning-based semantic segmentation algorithms, such as U-Net or SegNet, can learn to segment an image into different classes of objects. This provides a more detailed understanding of the scene. A comprehensive review can be seen in [20].
- Stereo Vision:
  Stereo vision utilizes two cameras to estimate the depth of objects in a scene, as shown in Fig. 6. Computing disparities between the two images, the depth of each pixel can be estimated. Stereo vision can be used to detect obstacles by identifying regions in the depth image that correspond to objects in front of the camera.

  LiDAR-based obstacle recognition presents a different approach, utilizing point cloud data to perceive the environment. LiDAR systems emit laser pulses and measure the time it takes for the pulses to bounce back, allowing for the creation of detailed

Fig. 6. Captured Image from the ZED2i stereo camera.

3D maps. The following methods are commonly used for analysing LiDAR data [14, 21].

- Point Cloud Segmentation:
  Point cloud segmentation involves separating LiDAR data into distinct groups of points that correspond to different objects in the environment. Clustering algorithms, such as the k-means algorithm or the Euclidean clustering algorithm, can be used to group points that are close to each other in space. The resulting clusters can be classified as obstacles or other objects of interest.
- Ground Segmentation:
  Ground segmentation aims to remove ground points from the LiDAR data and extract non-ground objects. The RANSAC (Random Sample Consensus) method is commonly employed, which fits a plane to the ground points and separates them from the rest of the data. The remaining points can then be classified as obstacles.
- Simultaneous Localization and Mapping (SLAM):
  SLAM combines LiDAR data with other sensory information, such as wheel odometry or IMU data, to build a real-time map of the environment. SLAM algorithms can detect obstacles by analysing differences between the current and previous maps and identifying changes corresponding to the presence of obstacles.

Vision systems for mobile robot navigation have limitations that can affect precision, such as sensitivity to lighting, adverse weather conditions and limited field of view. However, these limitations can be overcome by incorporating LiDAR sensors as a separate data source, complementing the capabilities of vision sensors. It should underline that vision sensors provide high-resolution imagery, but LiDAR sensors offer additional parameters such as distance to objects and signal strength (indicated in the scale at the lower right corner), which provide valuable information about the environment (Fig. 7).

A typical environment with dense vegetation poses significant challenges regarding the classification of objects against the complex background. However, the integration of LiDAR technology proves to be highly advantageous in this scenario. LiDAR's ability to capture high-resolution point clouds allows the robot to overcome these challenges and achieve more accurate data.

This information can be particularly useful in scenarios where visual data from cameras may be limited, such as low-light conditions or when objects are partially a)

b)

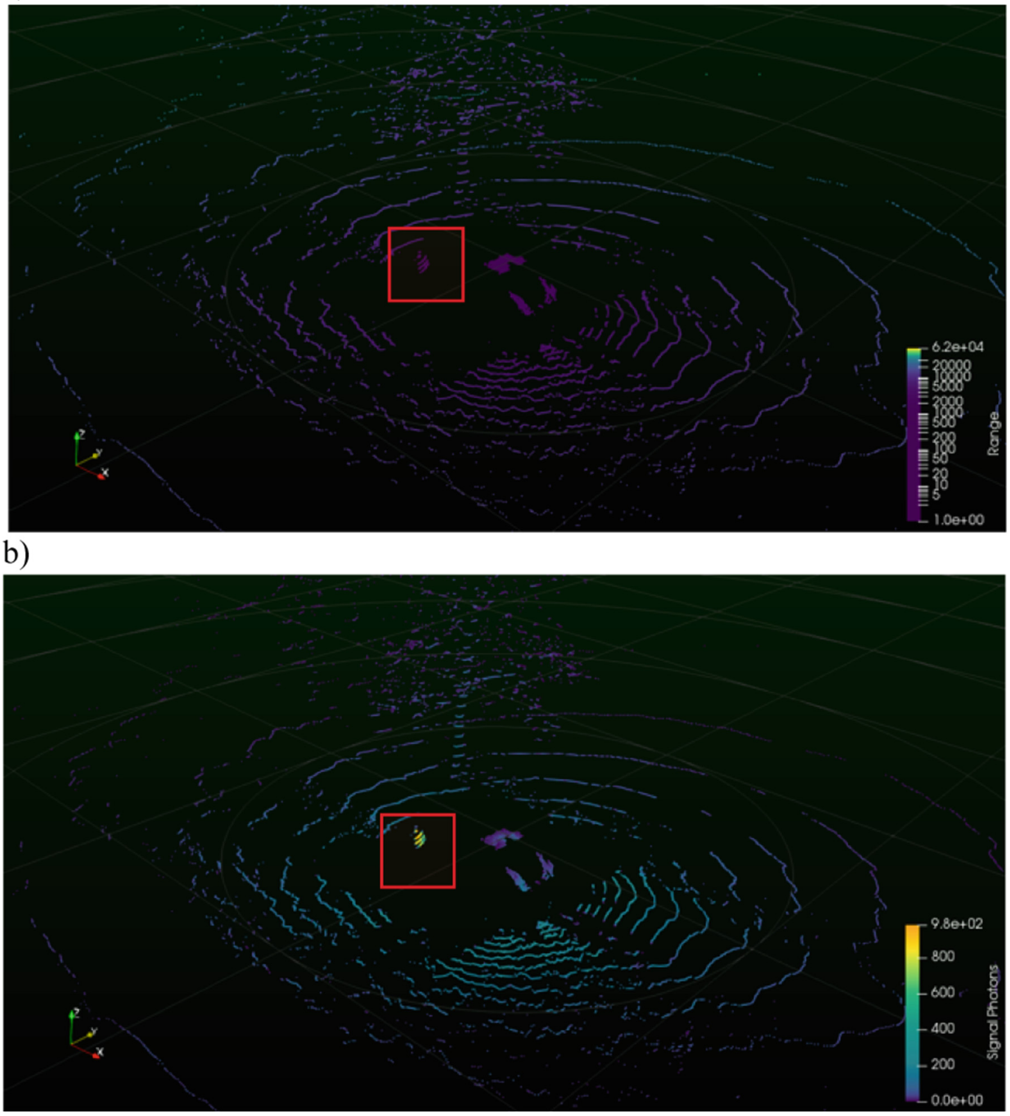


Fig. 7. 3D environment representation considering the distance from the sensor (a) and reflected energy as signal photons (b).

obscured. The proposed approach in the paper uses LiDAR data as source for already trained neutral network for vision system to utilize additional parameters, enabling more accurate object detection, classification, and scene understanding for precise navigation and decision-making in various mobile robot applications.

4 Implementation of AI for Object Recognition in Operational Environment

Precisely measuring the distance to objects is not only important for mobile robot navigation but also critical in certain scenarios like operation in dangerous zones to mitigate any potential risks. This capability to perceive the spatial layout of surroundings in real-time enhances the overall safety and reliability of the unmanned platforms. Accurate distance estimation is particularly critical in operational environments where the robot may encounter complex terrains, dynamic obstacles, or narrow passageways.

Vision systems are valuable sources of object detection for mobile robots, utilizing neural network models (Fig. 8) such as AdaBins for absolute depth maps and MiDaS for inverse depth estimations [15]. These advanced models enhance distance estimation and object recognition, enabling mobile robots to navigate and interact with their environment more effectively, improving their overall capabilities.

For instance, authors in [16] proposed an overall system for monocular depth estimation, shown in Fig. 8. It consists of a front-end CNN with an encoder and decoder (blue blocks) that extracts features and generates an initial depth map from the input image. The researchers propose a Structured Attention guided CRF model (grey box) to refine the depth estimation further. This CRF model incorporates attention maps (green boxes) to highlight essential regions in the image and jointly inferred features (light blue boxes) through a message-passing algorithm. Integrating attention into the CRF framework aims to enhance depth estimation accuracy by leveraging contextual information.

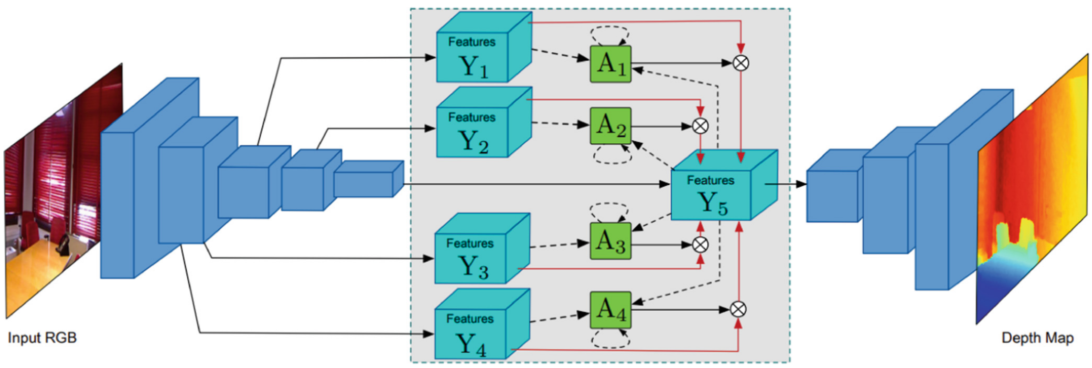


Fig. 8. Architecture for depth estimation using convolutional neural network [16].

In our study, we have evaluated precision of developed model fusing the object detection algorithm You Only Look Once (YOLO) and depth estimation using machine learning model from an arbitrary input image (MiDaS). The model employs a convolutional neural network architecture, specifically a fully convolutional neural network (FCN). The FCN architecture allows the model to take an entire image as input and produce a corresponding depth map. Typically, the depth regression represents the distance of each pixel from the camera in a grayscale, where brighter pixels indicate objects that are closer to the camera and darker pixels represent objects that are quite far.

This software was extensively assessed in operational environments to assess its performance and reliability. Tests were conducted based on typical traffic scenarios to evaluate the accuracy of our distance estimation under good weather, lighting, and environmental conditions.

The results depicted in Fig. 9 rely solely on vision-based distance estimations. Throughout the testing phase, the operational environment division into two zones. The first one is the Near Zone, encompassing distances up to 4.5 m from the mobile robot, requiring precise operation. This zone demands high accuracy and responsiveness to navigate around nearby obstacles or potential hazards. The second region is the Far

Fig. 9. Examples of detected objects using fused model in operational environment with assigned estimated distance.

Zone, where objects are positioned beyond the 4.5-m, allowing the mobile robot to move freely without hindrance.

Unfortunately, all measurements did not provide the expected level of precision required for effective operation in various conditions. Our investigation highlighted the crucial role of precise distance measurements for efficient operation in dynamic environments, leading us to investigate the integration of LiDAR sensors (that provide precise measured distance as a parameter in their output data).

5 Flat View Projection from LiDAR as Source for AI Models

In our work, we have proposed an innovative approach that combines data from lidars represented as flat images as source (Fig. 10) for already trained neutral network models for object detection. Integrating laser scanner data into the analysis, we achieve precise distance estimation for the detected objects utilizing data from the high-quality OS1-128 LiDAR sensor, based on an existing dataset provided by the manufacturer [17].

Fig. 10. Example of flat view from Ouster OS1–128 LiDAR point cloud surfaces.

Using the same neutral network model YOLO, like in-depth analysis of vision system, we have classified the trained objects in operational environment. After object detection it is possible to accurately determine the distance and gain additional information provided by the Ouster sensor.

a)

b)

Fig. 11. Examples of detected objects using flat view with assigned precise distance.

The integration of convolutional neural network model (YOLO) for real-time object detection system with laser scanner data in our approach allows achieve precise object detection and accurate distance estimation, as shown in Fig. 11.

Each point has also encoded distance and due to high resolution dense point cloud can be treated like image form camera as source for already trained CNN. Calculating the average distance from a classified set of points (associated with the recognized object) our approach enables accurate localization and distance estimation even in complex spatial environments.

Laser scanners provide precise measurement from dense point cloud that enhance the perception system of mobile robots, enabling comprehensive environmental information. This innovative method opens new possibilities for advanced object detection, accurate length estimation, and leveraging the rich data from laser scanners to improve decision-making in diverse mobile robot applications.

6 Results Discussion

The article presents a detailed overview of a wheeled robot with a modular construction designed to conduct special missions in challenging terrains. These tasks often require precise object detection and accurate distance estimation capabilities. Considering operational capabilities dedicated software was developed that utilizes the YOLO algorithm and the MiDaS model for image analysis. We utilised data collected from unstructured and unexplored environments to validate the estimation model detailed in Sect. 4. Distance estimation was obtained through camera data captured across two distinct settings: one imitating real-world scenarios in densely vegetated surroundings and the other in a characteristic urban structure.

The limitations of vision systems concerning distance estimation are underlined. To overcome this challenge, the authors proposed a solution that uses a developed neural network model to analyse the projection of LiDAR point cloud distributions. This approach provides more accurate and reliable distance information, enhancing the robot's capabilities in terms of situation awareness and control.

The spatial dataset used in this study is based on Ouster 128 output derived from a typical urban landscape. This dataset was obtained by downloading high-resolution point cloud representation with precise distance measurements and flat-view imaging from the manufacturer's website [17].

The proposed approach combining vision systems and LiDAR technology enables the wheeled robot to overcome the limitations of traditional vision-based distance estimation. Through this integration, the robot gains the ability to operate with enhanced precision and efficiency, effectively navigating and perceiving its surroundings. The results of this approach are illustrated in Fig. 11.

Acknowledgments. This work was supported under research work no. 55.23615.PR and 55.2022489.PL at the Military Institute of Armoured and Automotive Technology.

References

1. Petrişor, S.M., Simion, M., Bârsan, G., Hancu, O.: Humanitarian demining serial-tracked robot: design and dynamic modeling. Machines **11**, 548 (2023). https://doi.org/10.3390/machines11050548
2. Rubio, F., Valero, F., Llopis-Albert, C.: A review of mobile robots: concepts, methods, theoretical framework, and applications. Int. J. Adv. Robot. Syst. **16**(2) (2019). https://doi.org/10.1177/1729881419839596
3. Jung, Y.H., et al.: Development of multi-sensor module mounted mobile robot for disaster field investigation. Gottingen Copernicus GmbH (2022). https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-1103-2022
4. Janczak, D., Walendziuk, W., Sadowski, M., Zankiewicz, A., Konopko, K., Idzkowski, A.: Accuracy analysis of the indoor location system based on bluetooth low-energy RSSI measurements. Energies **15**, 8832 (2022). https://doi.org/10.3390/en15238832
5. Janos, R., Sukop, M., Semjon, J., et al.: Conceptual design of a leg-wheel chassis for rescue operations. Int. J. Adv. Robot. Syst. **14**(6) (2017). https://doi.org/10.1177/1729881417743556

6. Russo, M., Ceccarelli, M.: A survey on mechanical solutions for hybrid mobile robots. Robotics **9**, 32 (2020). https://doi.org/10.3390/robotics9020032
7. Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., Bennamoun, M.: Deep learning for 3D point clouds: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **43**(12), 4338–4364 (2021). https://doi.org/10.1109/TPAMI.2020.3005434
8. Khan, D., Cheng, Z., Uchiyama, H., Ali, S., Asshad, M., Kiyokawa, K.: Recent advances in vision-based indoor navigation: a systematic literature review. Comput. Graph. **104**, 24–45 (2022). https://doi.org/10.1016/j.cag.2022.03.005. ISSN 0097-8493
9. Zhang, Y., Carballo, A., Yang, H., Takeda, K.: Perception and sensing for autonomous vehicles under adverse weather conditions: a survey. ISPRS J. Photogram. Remote Sens. **196**, 146–177 (2023). https://doi.org/10.1016/j.isprsjprs.2022.12.021. ISSN 0924-2716
10. Silarski, M., Nowakowski, M.: Performance of the SABAT neutron-based explosives detector integrated with an unmanned ground vehicle: a simulation study. Sensors **22**, 9996 (2022). https://doi.org/10.3390/s22249996
11. https://www.stereolabs.com/zed-2/. Accessed 02 June 2023
12. Yang, T., et al.: 3D ToF LiDAR in mobile robotics: a review. arXiv preprint arXiv:2202.11025 (2022)
13. Sivaraman, S., Trivedi, M.M.: Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis. IEEE Trans. Intell. Transp. Syst. **14**(4), 1773–1795 (2013)
14. Alaba, S., Gurbuz, A., Ball, J.: A Comprehensive Survey of Deep Learning Multisensor Fusion-based 3D Object Detection for Autonomous Driving: Methods, Challenges, Open Issues, and Future Directions. TechRxiv (2022)
15. Ranftl, R., Bochkovskiy, A., Koltun, V.: Vision transformers for dense prediction. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 12159–12168 (2021). https://doi.org/10.1109/ICCV48922.2021.01196
16. Xu, D., Wang, W., Tang, H., Liu, H., Sebe, N., Ricci, E.: Structured attention guided convolutional neural fields for monocular depth estimation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3917–3925 (2018). https://doi.org/10.1109/CVPR.2018.00412
17. https://ouster.com/resources/lidar-sample-data/autonomous-vehicle-sample-data/. Accessed 02 June 2023
18. Wu, Q., Castleman, K.R.: Image segmentation. In: Microscope Image Processing, pp. 119–152. Academic Press (2023)
19. Zou, Z., Chen, K., Shi, Z., Guo, Y., Ye, J.: Object detection in 20 years: a survey. Proc. IEEE **111**(3), 257–276 (2023). https://doi.org/10.1109/JPROC.2023.3238524
20. Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **44**(7), 3523–3542 (2022). https://doi.org/10.1109/TPAMI.2021.3059968
21. Gomes, T., Matias, D., Campos, A., Cunha, L., Roriz, R.: A survey on ground segmentation methods for automotive LiDAR sensors. Sensors **23**(2), 601 (2023)