Object Detection for Wheeled Mobile Robot Based Using Deep Learning: Systematic Review

Hapsari Peni Agustin Tjahyaningtijas [1*], Parama Diptya Widayaka [2], Raissa Alfatikarani [3], Sayyidul Aulia Alamsyah [4], Pradini Puspitaningayu [5], L. Endah Cahya Ningrum [6], Jessie R. Paragas [7]
[1, 2, 3, 4, 5, 6] Department of Electrical Engineering, Universitas Negeri Surabaya, Indonesia
[7] Information Technology Department, Eastern Visayas State University, Tacloban, Philippines
Email: [1] hapsaripeni@unesa.ac.id, [2] paramawidayaka@unesa.ac.id, [3] raissa.20003@mhs.unesa.ac.id, [4] sayyidulalamsyah@unesa.ac.id, [5] pradinip@unesa.ac.id, [6] endahningrum@unesa.ac.id, [7] jessie.paragas@evsu.edu.ph
*Corresponding Author

***Abstract*—The integration of deep learning technologies in object detection has significantly enhanced the capabilities of wheeled mobile robots, making them more efficient and intelligent in navigating complex environments. These technologies enable more accurate pattern recognition, adaptability to diverse environmental conditions, and improved autonomous decision-making capabilities. This study aims to explore the evolution and current trends in object detection for wheeled mobile robots, with a specific focus on the application of deep learning technologies as a foundational driver for system advancements. The objectives include analyzing the contribution of deep learning to improving object detection accuracy, system efficiency, and the robots' adaptability to dynamic environments. The methodology employed in this study is a Systematic Literature Review (SLR), comprising several key steps: formulating research questions, identifying relevant research sources, utilizing specific keywords for data collection, disseminating the gathered data, and analyzing the findings to address the research questions. Data is sourced exclusively from the Scopus digital database, focusing on publications from 2019 to 2024. The collected data, formatted in RIS, is subsequently analyzed and visualized using VOSviewer. The outcomes of this research include insights into the growth of research publications in recent years, the identification of key trends in object detection methodologies for wheeled mobile robots, the exploration of interconnections between critical concepts in the field, and the mapping of the knowledge network based on relevant keywords. Special emphasis is placed on the pivotal role of deep learning technologies in driving object detection advancements, including accuracy and system efficiency enhancements.**

***Keywords*—Mobile Robots; Object Detection; VOSViewer; Review Paper.**

I. Introduction

The rapid advancements in artificial intelligence and robotics have significantly expanded the capabilities of autonomous systems, particularly in object detection for wheeled mobile robots. Object detection, a fundamental aspect of computer vision, enables machines to identify and recognize objects within images or videos [1]**.** Wheeled mobile robots are increasingly utilized across various sectors, including logistics, manufacturing, and public services, where their ability to accurately detect surrounding objects is crucial for operational efficiency and safety [2]. Deep learning has emerged as a powerful tool in this domain, offering high-accuracy solutions for object detection through models such as Convolutional Neural Networks (CNN) and their derivatives [3].

A recent study [4] proposed an enhancement to the YOLOv4 algorithm by replacing CSPDarknet53 with a pruned GhostNet and applying depthwise separable convolution for improved efficiency. The dataset used was The DJI Robomaster Objects in Context (DJI ROCO), which features complex scenarios. The algorithm achieved a precision of 88.89%, recall of 87.12%, F1-score of 88.00%, and an mAP (0.5) of 86.84%, with a model size of 42.5 MB, making it lighter than the original YOLOv4. Despite its efficiency, the approach requires further optimization for real-time detection and seamless integration with robotic systems. Additionally, a study [5] presents a CNN-based method for detecting and characterizing cabbage and red cabbage for robotic fertilization in strip cropping. Implemented in ROS, the system integrates image acquisition, processing, and robotic trajectory planning. The dataset comprises 1,638 images (1280x960 pixels) collected from experimental fields, augmented to improve model robustness under varying lighting conditions. The trained CNN achieved an average accuracy of 90.5% and less than 3% validation error, outperforming conventional methods with an efficiency of 90.5% versus 50.6%. Limitations include potential constraints in detecting other plant types and applicability to low-power hardware, suggesting further research for broader integration. Recent studies [6] on object detection have shown its critical role in enabling autonomous systems to perceive and interact with their environment. This study leverages monocular vision and deep learning through Convolutional Neural Networks (CNNs) for autonomous navigation in mobile robots. The proposed model was trained on a custom dataset captured using a low-cost RGB-D sensor from a Hand-Controlled Mobile Robot (HCMR) in diverse indoor settings. The model achieved a mean accuracy of 77%, comparable to more expensive systems like Microsoft Kinect. While demonstrating effective low-cost navigation, limitations include offline training not suited for real-time use and discrete decision-making that may not fully address continuous state space requirements. Another study [7] combined Brain–Computer Interface (BCI) with object detection for robotic grasping, utilizing Overlapping Object GraspNet (OOGNet) to achieve a mAPg of 80.4%. Despite its promising results, the system requires extensive user training and struggles with performance variability in unstructured environments.

Another study [8] presents an innovative brain-actuated robotic arm system designed for autonomous 3D object grasping using a Brain–Computer Interface (BCI) based on EEG signals, specifically motor imagery and P300 patterns. The system integrates a 6-DOF robotic arm with an RGB-D camera and employs the Overlapping Object GraspNet (OOGNet) for accurate object detection and grasp prediction in cluttered environments. A custom Multi-Object Grasping Dataset, consisting of 784 annotated images, was developed to train the model. The system achieved high performance, with EEG classification accuracies of 95.58% (MI) and 96.3% (P300), and an mAPg of 80.4% for grasp detection, outperforming traditional methods. Advantages include reduced cognitive load and efficient real-time operation, although limitations persist, such as performance challenges in diverse environments and the need for user training.

The next study [9] classification and state recognition of agricultural robots using multi-source vibration data. The methodology integrates denoised and non-linearly enhanced vibration signals, processed via the Mallat wavelet algorithm and transformed into images using Gramian Angular Summation Fields (GASF). These images serve as input to an Attention-fused Residual Convolutional Neural Network (ANR-CNN), which utilizes channel and spatial attention mechanisms to enhance feature extraction. The model achieves a high classification accuracy of 92.35%, outperforming existing networks like GoogLeNet and ResNet50. While offering efficient real-time performance and robust accuracy, the approach's complexity and limited testing on diverse terrains highlight areas for further research and potential enhancement. These studies highlight the rapid progress in object detection for robotic applications while also revealing key limitations such as real-time processing challenges, adaptability to dynamic environments, and computational efficiency.

However, this systematic review aims to provide a comprehensive overview of recent advancements in object detection for wheeled mobile robots utilizing deep learning techniques. To address these challenges, this systematic review aims to provide a comprehensive overview of recent advancements in object detection for wheeled mobile robots utilizing deep learning techniques. The primary objective is to analyze existing research, highlight emerging trends, and evaluate the latest methodologies in object detection for robotic applications. The research contribution is threefold:

1. identifying key trends and challenges in recent object detection methods for wheeled mobile robots
2. systematically reviewing state-of-the-art deep learning-based approaches and their effectiveness,
3. providing insights and recommendations for future improvements in real-time object detection and navigation strategies for autonomous robots.

II. METHODS

The methodology employed in this research is a systematic literature review (SLR) as shown in Fig. 1. SLR is a structured approach for locating, assessing, and analyzing all existing research pertinent to a specific research question, subject matter, or area of interest [10]. The second stage involves a search process to identify data that will be used for analysis as shown in Fig. 2 which is illustrates the systematic process used to search and collect relevant research articles from the Scopus digital database. The flowchart begins with the selection of a digital database, followed by the definition of search keywords to ensure that the retrieved literature aligns with the study's objectives. An initial search is then conducted, after which the effectiveness of the chosen keywords is evaluated. If the keywords are found to be ineffective, the process loops back to a refinement step where keywords are adjusted to improve search accuracy. If the keywords are effective, the relevant search results are collected, and the selected articles are extracted and downloaded in RIS format for further analysis. The flowchart visually represents the structured approach taken to ensure a systematic and reproducible literature review process. This structured methodology ensures that only high-quality and relevant research publications are included in the study, providing a solid foundation for analyzing trends in deep learning-based object detection for wheeled mobile robots.

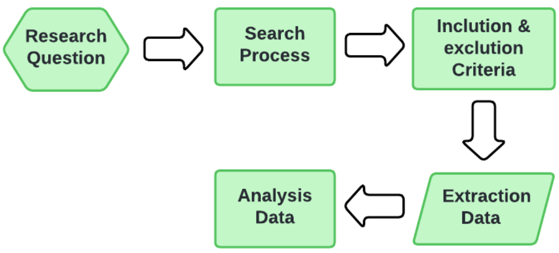

Fig. 1. SLR method

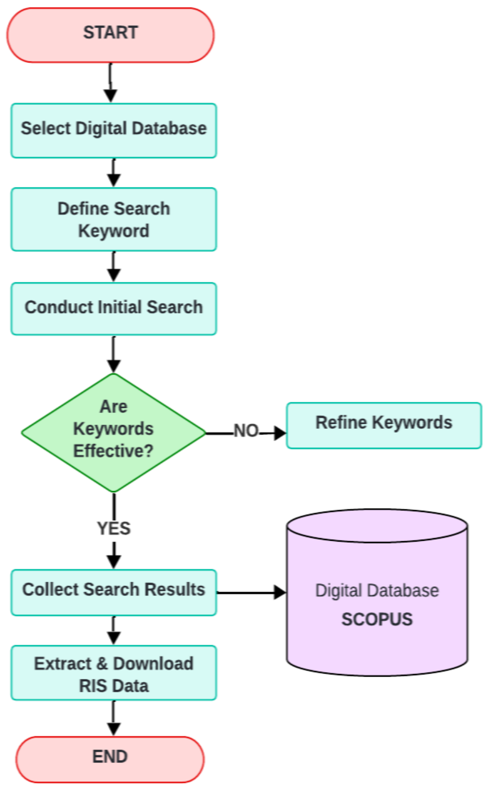

Fig. 2. Search data flowchart

A. *Research Question*

- RQ1: What research has been conducted on the topic of object detection for wheeled mobile robots using deep learning?
- RQ2: How has the research on object detection for wheeled mobile robots using deep learning evolved over time?
- RQ3: How many studies on object detection for wheeled mobile robots using deep learning were identified during the search process?

B. *Inclution & Exclution Criteria*

This stage aims to determine the data needed for this research, the following criteria were applied:

- Publication year 2019 – 2024.
- Focus on keyword object detection, mobile robot, deep learning.
- Scopus Digital Database.
- Only journal articles and conference papers were considered, while reviews, book chapters, and non-peer-reviewed materials were excluded.

C. *Extraction Data*

In this stage, the data required to address the Research Question is identified. The data used in this study is in RIS format, which has been downloaded from the digital database Scopus, comprising a total of 327 data.

D. *Analysis Data*

At this stage, analyzing the collected data is essential to address the research questions. In this study, the researchers utilized VOSViewer software for data analysis. VOSViewer was employed for:

- Keyword co-occurrence analysis to identify research trends.
- Bibliographic coupling to explore relationships between studies.
- Cluster visualization to categorize related topics.

To ensure transparency and reproducibility, the VOSViewer parameters used in this analysis were explicitly defined. The threshold for keyword co-occurrence was set at a minimum of five occurrences, while the **Association Strength** method was applied for normalization. Additionally, the **LinLog/modularity clustering algorithm** was employed to effectively group related research topics, facilitating a clearer understanding of the field's structure.

By leveraging VOSViewer, this study systematically maps the research landscape, identifying existing trends, research gaps, and potential future directions. These findings contribute to a deeper understanding of the progression of object detection methodologies in mobile robotics. Future studies may consider integrating qualitative synthesis methods or meta-analysis techniques to further enhance the robustness and comprehensiveness of bibliometric insights.

III. RESULTS AND DISCUSSION

This section will present and discuss the findings related to object detection for wheeled mobile robots using deep learning, as outlined in the research questions for this systematic review.

A. *RQ1: Research Conducted to the Topic Object Detection for Mobile Robot Using Deep Learning*

Based on Fig. 3, Fig. 4, and Fig. 5, illustrate the research landscape on object detection for wheeled mobile robots using deep learning. The keyword analysis reveals a strong correlation between *object detection*, *object recognition*, *mobile robots*, and *deep learning*, as indicated by the larger circles in the visualizations. Additionally, related topics such as *CNN*, *semantics*, *image processing*, *machine learning*, *intelligent robots*, *navigation*, and *autonomous mobile robots* are identified as emerging subfields.

These findings highlight the interdisciplinary nature of object detection in mobile robotics, demonstrating its relevance in various applications, including robotic vision, intelligent navigation, and autonomous decision-making. However, while deep learning has significantly improved object detection performance, challenges such as real-time processing constraints, dataset limitations, and model generalizability in dynamic environments remain critical concerns for researchers.

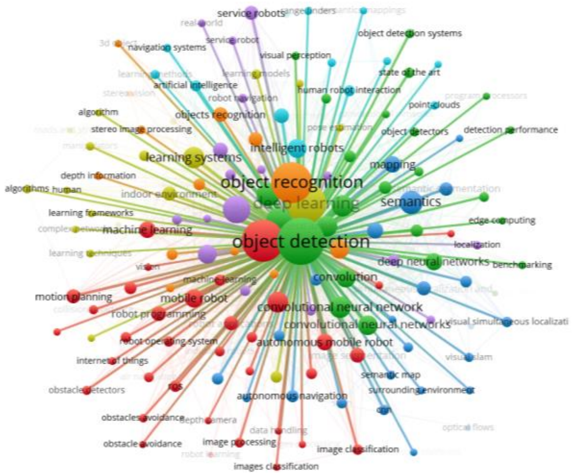

Fig. 3. VOSViewer visualization of keyword "Object Detection"

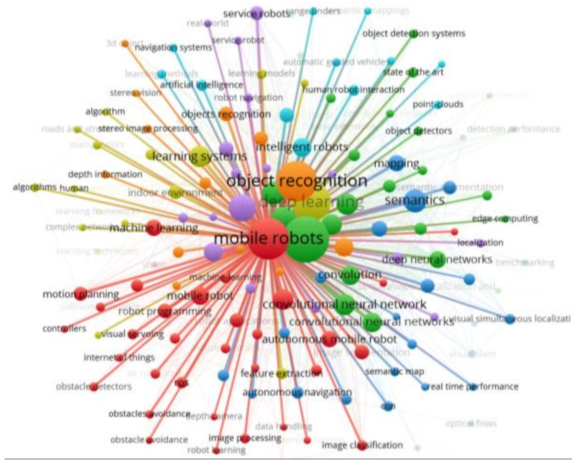

Fig. 4. VOSViewer visualization of keyword "Mobile Robots"

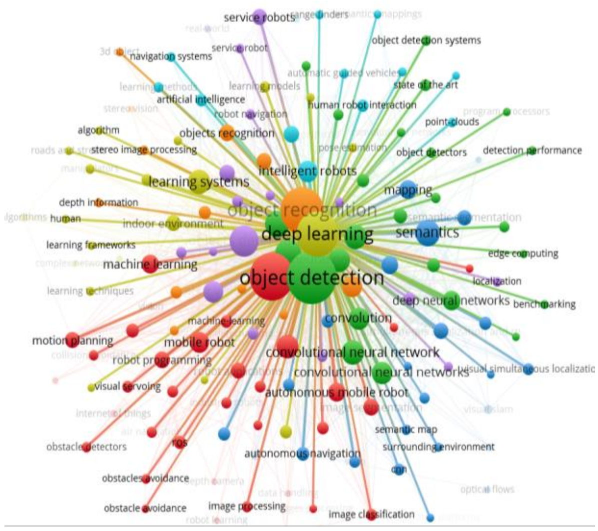

Fig. 5. VOSViewer visualization of keyword "Deep Learning"

B. *RQ2: Research on Object Detection for Wheeled Mobile Robots Using Deep Learning has Significantly Evolved Over Time*

Based on Fig. 6, Fig. 7, and Fig. 8, there is a yellow circle (yellow color gradient) indicating that the topic of Object Detection for Mobile Robots Using Deep Learning in 2022 is expected to continue evolving. The focus of this development will be on several key aspects, including visual simultaneous localization (SLAM) for real-time environment mapping, robot vision for enhancing scene understanding and perception, obstacle avoidance to improve autonomous navigation in dynamic environments, robot operating systems or seamless integration of deep learning models, semantic segmentation for object classification and contextual awareness, optical radar for depth perception and enhanced spatial understanding, object recognition and detection models to improve classification accuracy, and real-time processing for efficient deployment in edge devices. In other words, these topics will receive more attention in the context of changes and advancements in these fields from 2022 onward.

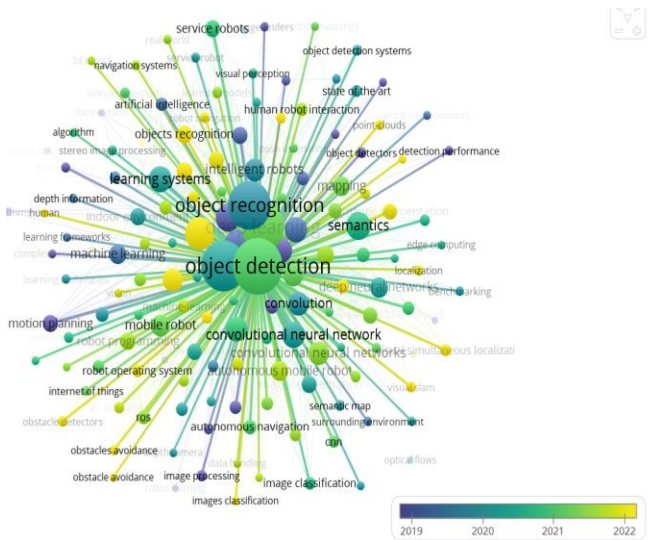

Fig. 6. Overlay visualization of keyword "Object Detection"

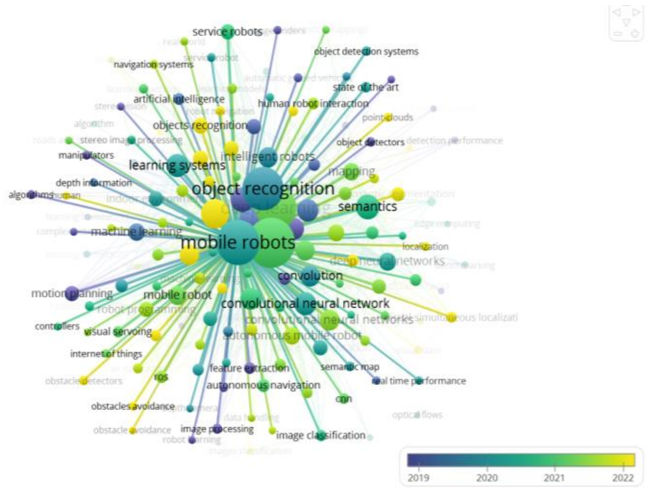

Fig. 7. Overlay visualization of keyword "Mobile Robots"

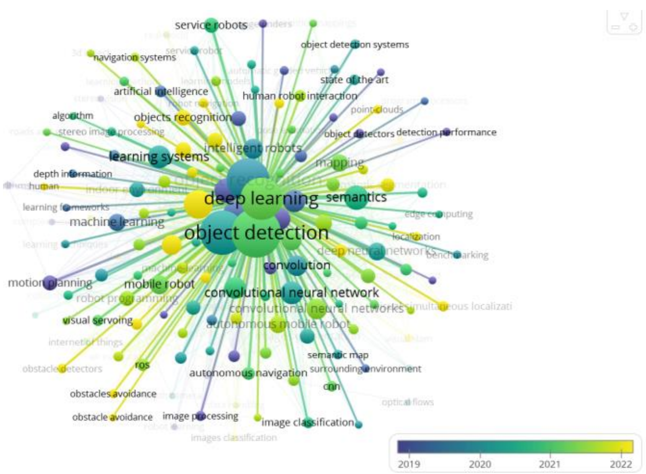

Fig. 8. Overlay visualization of keyword "Deep Learning"

As outlined in Table I, various deep learning-based methods have been employed to enhance object detection performance in wheeled mobile robots. The application of algorithms such as YOLO, CNN, and EfficientNet has been observed across different contexts, including indoor service robots, agricultural automation, and security surveillance. YOLO has been widely adopted for real-time object detection due to its high inference speed, with some studies reporting accuracy rates of up to 99%. However, its performance is highly dependent on dataset availability and may degrade in complex environments with varying illumination conditions. Meanwhile, lightweight models such as MobileNet and EfficientNet offer a balance between accuracy and computational efficiency, making them suitable for mobile robotic platforms with limited processing power. Recent research has also explored hybrid approaches that integrate multiple deep learning techniques, including the combination of CNN with transformer-based architectures, to improve robustness against occlusions and environmental variations. Despite these improvements, ongoing challenges remain in ensuring model generalization, handling occlusions, and adapting to diverse environmental conditions. Addressing these issues is essential for further advancing the practical deployment of object detection models in mobile robotics.

C. RQ3: A Total of Studies on Object Detection for Wheeled Mobile Robots Using Deep Learning were Identified During the Search Process

The density visualization shown in Fig. 9 presents a density visualization of research trends in object detection for wheeled mobile robots, highlighting the most frequently occurring keywords and their relationships. The bright yellow areas represent central research themes, including object detection, object recognition, deep learning, and mobile robots, which continue to be the primary focus of studies in this field. The green areas indicate moderate keyword density, where terms such as robot vision, learning systems, and intelligent robots remain relevant but receive comparatively less attention. In contrast, the dark blue regions signify less-explored topics with weaker connections, suggesting potential opportunities for future research.

This analysis provides valuable insights into the evolution of research trends and identifies areas that warrant further exploration. One promising direction is the integration of transformer-based architectures, which have demonstrated superior performance in feature extraction and contextual understanding. Additionally, multimodal learning approaches that combine visual data with other sensory inputs, such as LiDAR and infrared, could enhance object detection capabilities in complex environments. Another critical area of interest is the development of adaptive learning models that can adjust to new environments with minimal retraining, thereby improving the robustness of object detection systems. Furthermore, optimizing deep learning models for real-time processing on embedded systems remains a significant challenge, requiring novel techniques to reduce computational overhead while maintaining high detection accuracy. By analyzing term co-occurrence patterns and emerging research trends, this study highlights key advancements in the field while providing a foundation for future studies to explore new methodologies and technological innovations in object detection for wheeled mobile robots.

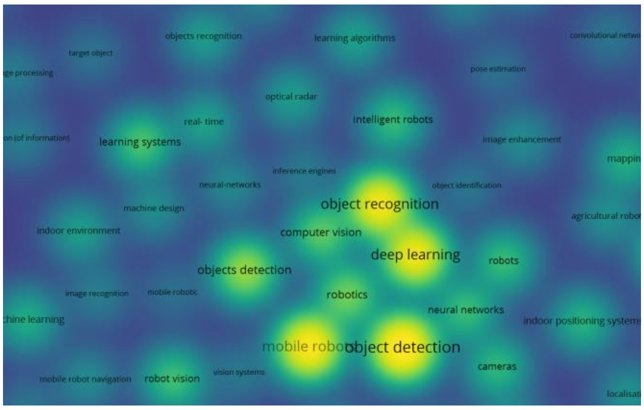

Fig. 9. VOSViewer density visualization

TABLE I. RESEARCH COMPARISON

Author	Methods	Performance
Ammar, et al [11]	Object detection and localization mobile robots. Kernel Density Estimation (KDE) Method	achieve a detection rate of 40.7 FPS
Min-Fan, et al[12]	Security robot using Three CNN models, (AlexNet, VGGNet, and GoogleNet)	The experiment showed accuracy of the three CNN models is 0.95, 0.95, and 0.6
Medrano, et al[13]	Robot navigation Five classification methods: GB, XGB, LGBM, KNN, SVM	GB = 0.7297 XGB = 0.7253 LGBM = 0.7208 KNN = 0.6754 SVM =0.7297
Zijing Song, et al[14]	Mobile robot using YF-SLAM Method	RMSE YF-SLAM= 0.01 RMSE DYNA-SLAM= 0.02 RMSE DS-SLAM= 0.03 RMSE ORB-SLAM2= 0.7
Caixia He, et al[15]	Mobile robot using SVM algorithm, SSD, improve algorithm	SVM algorithm= 96.65% SSD= 96.78% Improve algorithm= 98.8%
Li Y, et al[16]	Mobile robot using PID algorithm	Acquisition rate= 0.81 Tracked target position= 0.46 Angle smoothness= 1.02
Yuichi Sasaki, et al[17]	Robot DANIEL using YOLO	α= No display of object recognition β= display only one object recognition result γ= display 80 kinds of object recognition
Li Z., et al[18]	Mobile robot using YOLOV8	Mean Absolute Percentage Error of 0.77
Dimitrij, et al[19]	Autonomous forklifts using YOLOV7	MAPE SYNTH Model = 0.096 MAPE LOCO SYNTH Model = 0.740 MAPE TOMIE SYNTH Model = 0.005
Marta, et al[20]	Indoor service robot using Efficient Net	Accuracy = 0.97
Kaur, et al[21]	Robot Operating System-based mobile robot using CNN	Accuracy87.73%
Alajami, et al[22]	Distributed multi-robot localization strategy using CNN	The results show that this proposed method is robust and accurate while maintaining a degree of simplicity and efficiency in costs
Homayun, et al[23]	Multi Robotic System Simultaneous Localization and Mapping (SLAM)	Accuracy 93.4%
Adel, et al[24]	Robot Vision using YOLOv5	Accuracy 0.98
Feiyu, et al[25]	Mobile robot using YOLOv7	Accuracy 95%
Stefan, et al[26]	Mobile Robots in Logistic Environments using CNN	Accuracy of 99%
Kanta, et al[27]	Domestic service robots (DSRs) using MRR	Accuracy 80%

Petar, et al[28]	Autonomous robots for vineyard maintenance using YOLOV5	Accuracy 81%
Lesia, et al[29]	Mobile robot using YOLOV5	Accuracy 95%
Dinda Pramanta[30]	EV3-Robot, Mobile robot YOLO	Latency below 1,000 ms.
Duy, et al[31]	Robot navigation using CNN	Accuracy 97.17%
Ali, et al[32]	Mobile robot using YOLOv3	Accuracy 99%
Pel An, et al[33]	Mobile robot using PSPL-3D methods	Accuracy 71,63%
Divya, et al[34]	Mobile robot using MobileNet with Vision Transformer	Accuracy 94,3%
Yangqing, et al[35]	Home Service Robots using YOLO	Accuracy 90.72%
Yin Jia, et al[36]	Autonomous mobile robot using RCNN Resnet 50	Accuracy 82,33%
Zhang, et al[37]	Service robot using CNN	Accuracy 98,1%
Alif, et al[38]	Robot navigation MobileNetv3	Accuracy 83,7%
Salman, et al[39]	Robot motion planning Siamese using Neural Network	Accuracy 96,87%
Yadav, et al[40]	Mobile robot PicoDet model	Mean Average Precision (mAP) = 92.38%
Siddhant, et al[41]	Mobile robotic platforms DNN	Mean Average Precision (mAP)= 90.51%
Zhang, et al[42]	Mobile robot navigation GC YOLOv3	IoU = 6,35%
Peng Ding, et al[43]	Mobile robots in security scenes YOLOv4	mAP = 34,51%
Weifeng, et al[44]	Mobile robots YOLOv5	Accuracy 99.9%
Jesse, et al[45]	Mobile robots Deep Object Pose Estimation	Accuracy 64,8%
Myo, et al[46]	Mobile robots YOLOv4	Accuracy 97%
Jeonghoon, et al[47]	Mobile robots Siamese Network	Accuracy 93,95%
Sneha, et al[48]	Mobile robots YOLOv4	Accuracy 90%
Lim Kim, et al[49]	Mobile robots AFAM–EfficientDet network	Accuracy 90%
Sudeep, et al[50]	Robot System With autonomy Semantic segmentation method	Accuracy 98,2%
Donghun, et al[51]	Autonomous Mobile robots YOLOv5	Accuracy 97,5%
Raihan, et al[52]	Mobile robots MobileNet	Accuracy 98%
Zhengxue, et al[53]	Mobile robots PV-RCNN	Accuracy 92%
Rodrigo, et al[54]	Mobile robots using U-Net	Accuracy 90%
Hwang, et al[55]	Mobile robots CNN	Accuracy 99%
Wu, et al[56]	Mobile robots YOLO-SLAM	Accuracy 98,13%
Zhaohui, et al[57]	Mobile robots YOLOv5-AH	Accuracy 92%
Mouna, et al[58]	Indoor robots EfficientDet	Accuracy 89%
Sonay, et al[59]	Mobile robots Regression models	Accuracy 98%
Shabnam, et al[60]	Mobile robots using CNN	Accuracy 88,4%
Griffin, et al[61]	Task-Focused Few-Shot Object Detection (TFOD) & Detection-Based Manipulation	Speed pick-and-place: 124,6 pick/hour; Gain visual-servo control +16,7%; Gain depth estimation +25,0%
Lei, X, et al[62]	YOLOv8-R (pruned + TensorRT)	mAP 0.5 = 97,3%; mAP 0.5:0.95 = 82,1%; jetson nano 639,8 FPS
Xu, Z, et al[63]	Ensemble Lightweight DODT (RGB-D)	Position error = 0,11 m; Velocity error = 0,23 m/s (real-time onboard)
Xu, et al[64]	Multi-branch Faster R-CNN + Latency SLA scheduler	52 % Latency Reduced & 11.1 % Accuracy gain then YOLOv3 in Benchmark Video.
Wu, et al[65]	SSD300 & YOLOv3 model-parallelism	Increase in FPS up to 1.8× on heterogeneous edge devices without significant loss of accuracy.
Cabrera, et al[66]	CNN global descriptor for omnidirectional images	Outperform traditional methods in recall accuracy and runtime < 100 ms per image.
Patruno, et al[67]	CNN for feature detection & visual odometry	Avg. localization error 0.02 m at 1 Hz frame rate
Zhang, et al[68]	YOLOv3 + SVM classifier	mAP = 0.92; 45 FPS in indoor scenarios
Protasov, et al[69]	CNN + YOLOv3 on six 360° cameras	30 FPS; detection accuracy = 0.87
Li, et al[70]	HSV + background subtraction + CNN	95 % catch success (3–6 m); 120 ms latency
Hu, et al[71]	YOLOv3 + ROS navigation	Detection accuracy = 89.2 %; 25 FPS on Jetson Nano
He, et al[72]	LiDAR + USB camera + Faster R-CNN	Accuracy = 0.91%; 150 ms inference
Wang, et al[73]	LiDAR + YOLOv3-tiny	Detection accuracy = 0.87; mapping error 0.15 m
Ruchanovsky et al[74]	NAV-YOLO (YOLO variant) + ROS	100 % obstacle avoidance; 28 FPS on TurtleBot3
Yamamoto et al[75]	Weight-quantized SqueezeNet	Accuracy = 93 %; model footprint = 0.8 MB
Tahara, et al[76]	Data augmentation + adaptive feature fusion (DETR)	+12 % small-object detection accuracy on OmniCam
Patruno, et al[77]	Visual odometry + CNN	Avg. position error 0.03 m
Liu, et al[78]	DETR + 3D convolution	mAP3D = 85.2 % (KITTI); 22 FPS
Manglani, et al[79]	YOLOv4-tiny + A* path planning	100 % collision-free navigation; 30 FPS
Zeng, et al[80]	IBN-YOLOv5s (YOLOv5s enhanced with Spatial Pyramid Pooling and Instance-Batch Norm) running on ROS	Intersection over Union (IoU): 0.96

IV. CONCLUSION

Based on the results and discussion of the research findings related to object detection for wheeled mobile robots using deep learning, as analyzed through VOSviewer, it can be concluded that research trends in this area are actively evolving and expanding. This growth is closely associated with key topics such as robot vision, obstacle avoidance, visual simultaneous localization, and advancements in deep learning models. The integration of these themes highlights a growing focus on enhancing the capabilities of mobile robots for autonomous navigation and real-time decision-making in dynamic environments.

Despite these advancements, several challenges remain, including the need for more efficient deep learning models that balance accuracy and computational efficiency, as well as the integration of multimodal sensor data for improved robustness. Future research should explore the application of emerging techniques such as transformer-based architectures, edge computing, and self-supervised learning to further optimize object detection performance. Addressing these challenges will be crucial in positioning wheeled mobile robots as intelligent, adaptable systems capable of operating effectively across diverse and complex scenarios

Future research should explore the application of emerging techniques such as transformer-based architectures, edge computing, and self-supervised learning to further optimize object detection performance. In addition, addressing key research gaps—such as developing lightweight deep learning models with lower computational overhead, integrating multi-sensor fusion techniques (e.g., LiDAR, thermal imaging, and depth sensors), and improving adaptive learning mechanisms—will be essential for enhancing the robustness of object detection systems. These advancements will be crucial in positioning wheeled mobile robots as intelligent, adaptable systems capable of operating effectively across diverse and complex scenarios.

ACKNOWLEDGMENT

This research has been supported through the Institute for Research and Community Service (LPPM) and the Directorate of Innovation, Ranking and Scientific Publication, Universitas Negeri Surabaya.

REFERENCES

[1] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," *Proc. IEEE*, vol. 111, no. 3, pp. 257–276, 2023, doi: 10.1109/JPROC.2023.3238524.

[2] H. Zhang, Y. Zhang, and T. Yang, "A survey of energy-efficient motion planning for wheeled mobile robots," *Ind. Rob.*, vol. 47, no. 4, pp. 607–621, 2020, doi: 10.1108/IR-03-2020-0063.

[3] N. Manakitsa, G. S. Maraslidis, L. Moysis, and G. F. Fragulis, "A Review of Machine Learning and Deep Learning for Object Detection, Semantic Segmentation, and Human Action Recognition in Machine and Robotic Vision," *Technologies*, vol. 12, no. 2, 2024, doi: 10.3390/technologies12020015.

[4] Y. Hu, G. Liu, Z. Chen, and J. Guo, "Object Detection Algorithm for Wheeled Mobile Robot Based on an Improved YOLOv4," *Appl. Sci.*, vol. 12, no. 9, 2022, doi: 10.3390/app12094769.

[5] C. Cruz Ulloa, A. Krus, A. Barrientos, J. del Cerro, and C. Valero, "Robotic Fertilization in Strip Cropping using a CNN Vegetables Detection-Characterization Method," *Comput. Electron. Agric.*, vol. 193, 2022, doi: 10.1016/j.compag.2022.106684.

[6] W. K. Sleaman, A. A. Hameed, and A. Jamil, "Monocular vision with deep neural networks for autonomous mobile robots navigation," *Optik (Stuttg).*, vol. 272, 2023, doi: 10.1016/j.ijleo.2022.170162.

[7] T. H. Lin, C. T. Chang, and A. Putranto, "Tiny machine learning empowers climbing inspection robots for real-time multiobject bolt-defect detection," *Eng. Appl. Artif. Intell.*, vol. 133, p. 108618, 2024, doi: 10.1016/j.engappai.2024.108618.

[8] A. Rakshit, S. Pramanick, A. Bagchi, and S. Bhattacharyya, "Autonomous grasping of 3-D objects by a vision-actuated robot arm using Brain–Computer Interface," *Biomed. Signal Process. Control*, vol. 84, 2023, doi: 10.1016/j.bspc.2023.104765.

[9] J. Guo *et al.*, "Supervised learning study on ground classification and state recognition of agricultural robots based on multi-source vibration data fusion," *Comput. Electron. Agric.*, vol. 219, 2024, doi: 10.1016/j.compag.2024.108791.

[10] H. Zhang and M. Ali Babar, "Systematic reviews in software engineering: An empirical investigation," *Inf. Softw. Technol.*, vol. 55, no. 7, pp. 1341–1354, 2013, doi: 10.1016/j.infsof.2012.09.008.

[11] A. Y. Naich and J. R. Carrión, "LiDAR-Based Intensity-Aware Outdoor 3D Object Detection," *Sensors*, vol. 24, no. 9, pp. 1–17, 2024, doi: 10.3390/s24092942.

[12] M. F. R. Lee and Z. S. Shih, "Autonomous Surveillance for an Indoor Security Robot," *Processes*, vol. 10, no. 11, pp. 1–32, 2022, doi: 10.3390/pr10112175.

[13] M. A. Kevin, J. Crespo, J. Gomez, and C. Alfaro, "Advanced System for Enhancing Location Identification through Human Pose and Object Detection," *Machines*, vol. 11, no. 8, 2023, doi: 10.3390/machines11080843.

[14] Z. Song, W. Su, H. Chen, M. Feng, J. Peng, and A. Zhang, "VSLAM Optimization Method in Dynamic Scenes Based on YOLO-Fastest," *Electron.*, vol. 12, no. 17, 2023, doi: 10.3390/electronics12173538.

[15] C. He and L. He, "Deep Learning-based Mobile Robot Target Object Localization and Pose Estimation Research," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, pp. 1325–1333, 2023, doi: 10.14569/IJACSA.2023.01406140.

[16] Y. Li, G. Yang, L. He, L. Zhao, T. Zhao, and H. Sun, "Deep Vision Guided Mobile Robot Dual Axis PID Dynamic Real Time Adjustment Following," in *2023 WRC Symposium on Advanced Robotics and Automation, WRC SARA 2023*, 2023. doi: 10.1109/WRCSARA60131.2023.10261835.

[17] Y. Sasaki, T. Kamegawa, and A. Gofuku, "Effect of display of YOLO's object recognition results to HMD for an operator controlling a mobile robot," *Artif. Life Robot.*, vol. 28, no. 2, pp. 323–331, 2023, doi: 10.1007/s10015-023-00856-0.

[18] Z. Li, C. Li, and P. Munoz, "Blueberry Yield Estimation Through Multi-View Imagery with YOLOv8 Object Detection," in *2023 ASABE Annual International Meeting*, 2023, doi: 10.13031/aim.202300883.

[19] D. M. Holm, P. Junge, J. Rutinowski, and J. Fottner, "Investigation of Deep Learning Datasets for Warehousing Logistics," *Proc. Conf. Prod. Syst. Logist.*, pp. 119–128, 2023, doi: 10.15488/15311.

[20] M. Rostkowska and P. Skrzypczyński, "Optimizing Appearance-Based Localization with Catadioptric Cameras: Small-Footprint Models for Real-Time Inference on Edge Devices," *Sensors*, vol. 23, no. 14, 2023, doi: 10.3390/s23146485.

[21] P. Kaur, S. Kansal, and V. P. Singh, "Investigation of the prediction of wildlife animals and its deployment using the robot," *J. F. Robot.*, 2023, doi: 10.1002/rob.22243.

[22] A. A. Alajami, N. Palau, S. Lopez-Soriano, and R. Pous, "A ROS-based distributed multi-robot localization and orientation strategy for heterogeneous robots," *Intell. Serv. Robot.*, vol. 16, no. 2, 2023, doi: 10.1007/s11370-023-00457-7.

[23] M.-L. Tham, "Internet of robotic things for mobile robots: Concepts, technologies, challenges, applications, and future directions," *Digit. Commun. Networks*, vol. 9, May 2023, doi: 10.1016/j.dcan.2023.05.006.

[24] A. Merabet, A. V Latha, F. A. Kuzhippallil, M. Rahimipour, J. Rhinelander, and R. Venkat, "Robot Vision and Deep Learning for Automated Planogram Compliance in Retail BT - Robotics, Computer Vision and Intelligent Systems," *J. Filipe and J. Röning*, pp. 21–30, 2024.

[25] F. Jia, M. Afaq, B. Ripka, Q. Huda, and R. Ahmad, "Vision- and Lidar-Based Autonomous Docking and Recharging of a Mobile Robot for Machine Tending in Autonomous Manufacturing Environments," *Appl. Sci.*, vol. 13, no. 19, 2023, doi: 10.3390/app131910675.

[26] S. D. Achirei, R. Mocanu, A. T. Popovici, and C. C. Dosoftei, "Model-Predictive Control for Omnidirectional Mobile Robots in Logistic Environments Based on Object Detection Using CNNs," *Sensors*, vol. 23, no. 11, 2023, doi: 10.3390/s23114992.

[27] K. Kaneda, S. Nagashima, R. Korekata, M. Kambara, and K. Sugiura, "Learning-To-Rank Approach for Identifying Everyday Objects Using a Physical-World Search Engine," *IEEE Robot. Autom. Lett.*, vol. 9, no. 3, 2024, doi: 10.1109/LRA.2024.3352363.

[28] P. Slaviček, I. Hrabar, and Z. Kovačić, "Generating a Dataset for Semantic Segmentation of Vine Trunks in Vineyards Using Semi-Supervised Learning and Object Detection," *Robotics*, vol. 13, no. 2, 2024, doi: 10.3390/robotics13020020.

[29] L. Mochurad, Y. Hladun, Y. Zasoba, and M. Gregus, "An Approach

for Opening Doors with a Mobile Robot Using Machine Learning Methods," *Big Data Cogn. Comput.*, vol. 7, no. 2, 2023, doi: 10.3390/bdcc7020069.

[30] D. Pramanta, N. Fuengfusin, A. R. Syulistyo, and H. Tamukoh, "YOLO real-time object detection on EV3-Robot using FPGA hardware Accelerator," *Proc. Int. Conf. Artif. Life Robot.*, pp. 416–419, 2024, doi: 10.5954/icarob.2024.os15-1.

[31] D. A. Nguyen, K. N. Hoang, N. T. Nguyen, and H. N. Tran, "Enhancing Indoor Robot Pedestrian Detection Using Improved PIXOR Backbone and Gaussian Heatmap Regression in 3D LiDAR Point Clouds," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3351868.

[32] A. A. Abed, A. Al-Ibadi, and I. A. Abed, "Vision-Based Soft Mobile Robot Inspired by Silkworm Body and Movement Behavior," *J. Robot. Control*, vol. 4, no. 3, 2023, doi: 10.18196/jrc.v4i3.16622.

[33] P. An *et al.*, "Leveraging Self-Paced Semi-Supervised Learning with Prior Knowledge for 3D Object Detection on a LiDAR-Camera System," *Remote Sens.*, vol. 15, no. 3, 2023, doi: 10.3390/rs15030627.

[34] D. Sirala and K. S. Nagla, "Accuracy in Object Detection by Using Deep Learning Method," in *Lecture Notes in Mechanical Engineering*, 2024, doi: 10.1007/978-981-99-4594-8_5.

[35] Y. Ye, X. Ma, X. Zhou, G. Bao, W. Wan, and S. Cai, "Dynamic and Real-Time Object Detection Based on Deep Learning for Home Service Robots," *Sensors*, vol. 23, no. 23, 2023, doi: 10.3390/s23239482.

[36] Y. Jia, B. Ramalingam, R. E. Mohan, Z. Yang, Z. Zeng, and P. Veerajagadheswar, "Deep-Learning-Based Context-Aware Multi-Level Information Fusion Systems for Indoor Mobile Robots Safe Navigation," *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23042337.

[37] Y. Zhang, M. Yin, H. Wang, and C. Hua, "Cross-Level Multi-Modal Features Learning With Transformer for RGB-D Object Recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 12, 2023, doi: 10.1109/TCSVT.2023.3275814.

[38] A. W. Ramadhan, B. S. B. Dewantara, Setiawardhana, and F. Akzada, "Lidar-based Human Detection for a Mobile Robot in Social Environment using Deep Learning," in *IES 2023 - International Electronics Symposium: Unlocking the Potential of Immersive Technology to Live a Better Life, Proceeding*, 2023, doi: 10.1109/IES59143.2023.10242550.

[39] A. E. Salman and M. R. Roman, "Augmented reality-assisted gesture-based teleoperated system for robot motion planning," *Ind. Rob.*, vol. 50, no. 5, 2023, doi: 10.1108/IR-11-2022-0289.

[40] R. Yadav, R. Halder, A. Thakur, and G. Banda, "A Lightweight Deep Learning-based Weapon Detection Model for Mobile Robots," in *ACM International Conference Proceeding Series*, 2023, doi: 10.1145/3610419.3610489.

[41] S. Panigrahi, P. Maski, and A. Thondiyath, "Real-time biodiversity analysis using deep-learning algorithms on mobile robotic platforms," *PeerJ Comput. Sci.*, vol. 9, 2023, doi: 10.7717/peerj-cs.1502.

[42] L. Zhang, L. Ji, and Z. Ma, "Target positioning method based on B-spline level set and GC Yolo-v3," *Cluster Comput.*, vol. 27, no. 3, 2024, doi: 10.1007/s10586-023-04164-x.

[43] P. Ding, H. Qian, Y. Zhou, and S. Chu, "Object detection method based on lightweight YOLOv4 and attention mechanism in security scenes," *J. Real-Time Image Process.*, vol. 20, no. 2, 2023, doi: 10.1007/s11554-023-01263-1.

[44] W. Kong *et al.*, "Object Detection Algorithm of Stairs Based on Multi-information Fusion," in *ACM International Conference Proceeding Series*, 2023, doi: 10.1145/3598151.3598187.

[45] J. Roch, J. Fayyad, and H. Najjaran, "DOPESLAM: High-Precision ROS-Based Semantic 3D SLAM in a Dynamic Environment," *Sensors*, vol. 23, no. 9, 2023, doi: 10.3390/s23094364.

[46] M. M. Aung, D. Maneetham, P. N. Crisnapati, and Y. Thwe, "Enhancing Object Recognition for Visually Impaired Individuals using Computer Vision," *Int. J. Eng. Trends Technol.*, vol. 72, no. 4, pp. 297–305, 2024, doi: 10.14445/22315381/IJETT-V72I4P130.

[47] J. Kwak, K. M. Yang, Y. J. Lee, M. G. Kim, and K. H. Seo, "Variance Optimization Based on Guided Anchor Siamese Network for Target-of-interest Object Recognition in Autonomous Mobile Robots," *Int. J. Control. Autom. Syst.*, vol. 21, no. 11, 2023, doi: 10.1007/s12555-022-0542-5.

[48] S. K. Hegde, R. Dharmalingam, and S. Kannan, "Intelligent German traffic sign and road barrier assist for autonomous driving in smart cities," *Multimed. Tools Appl.*, vol. 83, no. 22, 2024, doi: 10.1007/s11042-023-16435-1.

[49] T. L. Kim, S. Arshad, and T. H. Park, "Adaptive Feature Attention Module for Robust Visual–LiDAR Fusion-Based Object Detection in Adverse Weather Conditions," *Remote Sens.*, vol. 15, no. 16, 2023, doi: 10.3390/rs15163992.

[50] S. Sharan, P. Nauth, and J.-J. Domínguez-Jiménez, "Accurate semantic segmentation of RGB-D images for indoor navigation," *J. Electron. Imaging*, vol. 31, no. 06, 2022, doi: 10.1117/1.jei.31.6.061818.

[51] D. Shin, J. Cho, and J. Kim, "Environment-Adaptive Object Detection Framework for Autonomous Mobile Robots," *Sensors*, vol. 22, no. 19, 2022, doi: 10.3390/s22197647.

[52] R. Kabir, Y. Watanobe, M. R. Islam, K. Naruse, and M. M. Rahman, "Unknown Object Detection Using a One-Class Support Vector Machine for a Cloud–Robot System," *Sensors*, vol. 22, no. 4, 2022, doi: 10.3390/s22041352.

[53] Z. Zhou, L. Li, A. Fürsterling, H. J. Durocher, J. Mouridsen, and X. Zhang, "Learning-based object detection and localization for a mobile robot manipulator in SME production," *Robot. Comput. Integr. Manuf.*, vol. 73, 2022, doi: 10.1016/j.rcim.2021.102229.

[54] R. E. A. Ancona, L. G. C. Ramírez, and O. O. G. Frías, "Indoor Localization and Navigation based on Deep Learning using a Monocular Visual System," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, 2021, doi: 10.14569/IJACSA.2021.0120611.

[55] H. C. Wang, W. Z. Chen, Y. L. Huang, and J. J. Zhuang, "Adaptive System for Object Detection and Picking Based on Efficient Convolutional Neural Networks," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 94, 2022, doi: 10.1007/978-3-030-89776-5_1.

[56] W. Wu, L. Guo, H. Gao, Z. You, Y. Liu, and Z. Chen, "YOLO-SLAM: A semantic SLAM system towards dynamic environment with geometric constraint," *Neural Comput. Appl.*, vol. 34, no. 8, 2022, doi: 10.1007/s00521-021-06764-3.

[57] Z. Dong *et al.*, "Automatic Pavement Crack Detection Based on YOLOv5-AH," in *2022 12th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, CYBER 2022*, 2022, doi: 10.1109/CYBER55403.2022.9907394.

[58] M. Afif, R. Ayachi, Y. Said, and M. Atri, "An evaluation of EfficientDet for object detection used for indoor robots assistance navigation," *J. Real-Time Image Process.*, vol. 19, no. 3, 2022, doi: 10.1007/s11554-022-01212-4.

[59] S. Duman, A. Elewi, and Z. Yetgin, "Distance Estimation from a Monocular Camera Using Face and Body Features," *Arab. J. Sci. Eng.*, vol. 47, no. 2, 2022, doi: 10.1007/s13369-021-06003-w.

[60] S. Sadeghi Esfahlani, A. Sanaei, M. Ghorabian, and H. Shirvani, "The Deep Convolutional Neural Network Role in the Autonomous Navigation of Mobile Robots (SROBO)," *Remote Sens.*, vol. 14, no. 14, 2022, doi: 10.3390/rs14143324.

[61] B. Griffin, "Mobile Robot Manipulation using Pure Object Detection," in *Proceedings - 2023 IEEE Winter Conference on Applications of Computer Vision, WACV 2023*, 2023, doi: 10.1109/WACV56688.2023.00063.

[62] X. Lei, Y. Chen, and L. Zhang, "Real-Time SLAM and Faster Object Detection on a Wheeled Lifting Robot with Mobile-ROS Interaction," *Appl. Sci.*, vol. 14, no. 14, 2024, doi: 10.3390/app14145982.

[63] Z. Xu, X. Zhan, Y. Xiu, C. Suzuki, and K. Shimada, "Onboard Dynamic-Object Detection and Tracking for Autonomous Robot Navigation With RGB-D Camera," *IEEE Robot. Autom. Lett.*, vol. 9, no. 1, 2024, doi: 10.1109/LRA.2023.3334683.

[64] R. Xu *et al.*, "ApproxDet: Content and contention-aware approximate object detection for mobiles," in *SenSys 2020 - Proceedings of the 2020 18th ACM Conference on Embedded Networked Sensor Systems*, 2020, doi: 10.1145/3384419.3431159.

[65] Y. Wu, L. Liu, and R. Kompella, "Parallel Detection for Efficient Video Analytics at the Edge," in *Proceedings - 2021 IEEE 3rd International Conference on Cognitive Machine Intelligence, CogMI 2021*, 2021, doi: 10.1109/CogMI52975.2021.00035.

[66] J. J. Cabrera, V. Román, A. Gil, O. Reinoso, and L. Payá, "An experimental evaluation of Siamese Neural Networks for robot localization using omnidirectional imaging in indoor environments," *Artif. Intell. Rev.*, vol. 57, no. 8, 2024, doi: 10.1007/s10462-024-10840-

0.

[67] C. Patruno, V. Renò, M. Nitti, N. Mosca, M. di Summa, and E. Stella, “Vision-based omnidirectional indoor robots for autonomous navigation and localization in manufacturing industry,” *Heliyon*, vol. 10, no. 4, 2024, doi: 10.1016/j.heliyon.2024.e26042.

[68] K. J. Singh, D. S. Kapoor, K. Thakur, A. Sharma, and X. Z. Gao, “Computer-Vision Based Object Detection and Recognition for Service Robot in Indoor Environment,” *Comput. Mater. Contin.*, vol. 72, no. 1, 2022, doi: 10.32604/cmc.2022.022989.

[69] S. Protasov *et al.*, “CNN-based Omnidirectional Object Detection for HermesBot Autonomous Delivery Robot with Preliminary Frame Classification,” *2021 20th Int. Conf. Adv. Robot. ICAR 2021*, pp. 517–522, 2021, doi: 10.1109/ICAR53236.2021.9659386.

[70] S. T. Kao and M. T. Ho, “Ball-catching system using image processing and an omni-directional wheeled mobile robot,” *Sensors*, vol. 21, no. 9, 2021, doi: 10.3390/s21093208.

[71] M. Liu, M. Chen, Z. Wu, B. Zhong, and W. Deng, “Implementation of Intelligent Indoor Service Robot Based on ROS and Deep Learning,” *Machines*, vol. 12, no. 4, 2024, doi: 10.3390/machines12040256.

[72] F. He and L. Zhang, “Design of Indoor Security Robot based on Robot Operating System,” *J. Comput. Commun.*, vol. 11, no. 05, 2023, doi: 10.4236/jcc.2023.115008.

[73] T. Y. Lin and J. G. Juang, “Application of 3D point cloud map and image identification to mobile robot navigation,” *Meas. Control (United Kingdom)*, vol. 56, no. 5–6, 2023, doi: 10.1177/00202940221136242.

[74] N. Adiuku, N. P. Avdelidis, G. Tang, A. Plastropoulos, and Y. Diallo, “Mobile Robot Obstacle Detection and Avoidance with NAV-YOLO,” *Int. J. Mech. Eng. Robot. Res.*, vol. 13, no. 2, pp. 219–226, 2024, doi: 10.18178/ijmerr.13.2.219-226.

[75] Q. Huang, “Weight-Quantized SqueezeNet for Resource-Constrained Robot Vacuums for Indoor Obstacle Classification,” *AI*, vol. 3, no. 1, 2022, doi: 10.3390/ai3010011.

[76] K. Tahara and N. Hirose, “Ex-DoF: Expansion of Action Degree-of-Freedom with Virtual Camera Rotation for Omnidirectional Image,” in *Proceedings - IEEE International Conference on Robotics and Automation*, 2022, doi: 10.1109/ICRA46639.2022.9812301.

[77] C. Patruno, R. Colella, M. Nitti, V. Renò, N. Mosca, and E. Stella, “A vision-based odometer for localization of omnidirectional indoor robots,” *Sensors (Switzerland)*, vol. 20, no. 3, 2020, doi: 10.3390/s20030875.

[78] D. Rapado-Rincon, H. Nap, K. Smolenova, E. J. van Henten, and G. Kootstra, “MOT-DETR: 3D single shot detection and tracking with transformers to build 3D representations for agro-food robots,” *Comput. Electron. Agric.*, vol. 225, no. February, p. 109275, 2024, doi: 10.1016/j.compag.2024.109275.

[79] S. Manglani, “Real-time Vision-based Navigation for a Robot in an Indoor Environment,” *arXiv preprint arXiv:2307.00666*., 2023.

[80] J. Zeng and J. Fu, “Basketball robot object detection and distance measurement based on ROS and IBN-YOLOv5s algorithms,” *PLoS One*, vol. 19, pp. 1–18, 2024, doi: 10.1371/journal.pone.0310494.