International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 6, December 2025, pp. 5894~5903
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i6.pp5894-5903

Robotic product-based manipulation in simulated environment

Juan Camilo Guacheta-Alba[1], Anny Astrid Espitia-Cubillos[2], Robinson Jimenez-Moreno[3]
[1]Mechanical Engineering Program, Engineering Faculty, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[2]Industrial Engineering Program, Engineering Faculty, Universidad Militar Nueva Granada, Bogotá, Colombia
[3]Mechatronic Engineering Program, Engineering Faculty, Universidad Militar Nueva Granada, Bogotá, Colombia

Article Info

Article history:

Received Nov 18, 2024
Revised Jul 14, 2025
Accepted Sep 14, 2025

Keywords:

Artificial intelligence
Computer vision
Deep learning
Industry 5.0
Machine learning

ABSTRACT

Before deploying algorithms in industrial settings, it is essential to validate them in virtual environments to anticipate real-world performance, identify potential limitations, and guide necessary optimizations. This study presents the development and integration of artificial intelligence algorithms for detecting labels and container formats of cleaning products using computer vision, enabling robotic manipulation via a UR5 arm. Label identification is performed using the speeded-up robust features (SURF) algorithm, ensuring robustness to scale and orientation changes. For container recognition, multiple methods were explored: edge detection using Sobel and Canny filters, Hopfield networks trained on filtered images, 2D cross-correlation, and finally, a you only look once (YOLO) deep learning model. Among these, the custom-trained YOLO detector provided the highest accuracy. For robotic control, smooth joint trajectories were computed using polynomial interpolation, allowing the UR5 robot to execute pick-and-place operations. The entire process was validated in the CoppeliaSim simulation environment, where the robot successfully identified, classified, and manipulated products, demonstrating the feasibility of the proposed pipeline for future applications in semi-structured industrial contexts.

Corresponding Author:

Anny Astrid Espitia Cubillos
Industrial Engineering, Faculty of Engineering, Universidad Militar Nueva Granada
Carrera 11, 101-80, Bogotá, Colombia
Email: anny.espitia@unimilitar.edu.co

1. INTRODUCTION

Virtual environments stand out as a means of research development that allows the validation of algorithms in multiple environments and purposes as varied as autonomous navigation [1], object recognition and localization [2], security systems in transportation [3], collaborative production environments [4], drone operation [5], among others. Within these virtual environments it is feasible to implement machine learning and computer vision algorithms [6], operating with cameras, deep learning algorithms and image processing for product oriented applications [7], where this type of application underpin Industry 5.0 [8]. Advances in Industry 4.0 [9] and 5.0 [10] allow orienting efforts in the improvement of the production chain, thus involving the use of robots in manufacturing operations [11], collaborative work with robots [12] and object manipulation [13], [14] and their integration with different interaction algorithms with aforementioned virtual environment.

Despite the growing adoption of digital twins and industrial simulation, there is a lack of integrated virtual frameworks that combine product detection, classification, and robotic manipulation in semi-structured environments [15]. This is an important aspect in the task of promoting technological progress and incursion to improve the supply chain in different areas. In order to guide this task, this study proposes a methodology to apply industry 5.0 in small and medium production companies in virtual environments with a set of algorithms related to identification and manipulation of product through machine learning and computer vision by robotic manipulator. For its development, products are replicated in a virtual environment, where edge-based algorithms [16], [17] are used to discriminate objects of interest, identify features using the speeded-up robust features (SURF) algorithm [18], [19] and deep learning [20], [21] for subsequent manipulation by a UR5 robotic arm [22], [23].

The main objective of this work is to develop and validate a hybrid computer vision pipeline for identifying and manipulating visually similar products using an industrial robotic platform in a fully simulated environment. The key contributions include integration of heterogeneous AI techniques for object detection and robotic control; implementation of a real-time classification and manipulation cycle via the CoppeliaSim–MATLAB interface; and evaluation of the method under realistic conditions aligned with Industry 5.0 principles. In the absence of standardized datasets, a custom simulation-based dataset was created for validation, the difficulty of distinguishing products with similar geometry and branding highlights the applicability of the proposed approach to future industrial automation.

2. METHOD

Methodology proposed is illustrated in Figure 1 and consists of five phases. The first phase involves selecting the application environment. The second phase focuses on defining the artificial intelligence algorithm for product label identification. In third phase, the algorithm for package recognition is established. Fourth phase involves programming the movement of the robotic arm for product handling. In the final phase, the algorithms are integrated, and complete simulation is executed. CoppeliaSim serves as simulation environment, while MATLAB is used for programming; two software are connected through the remote API.

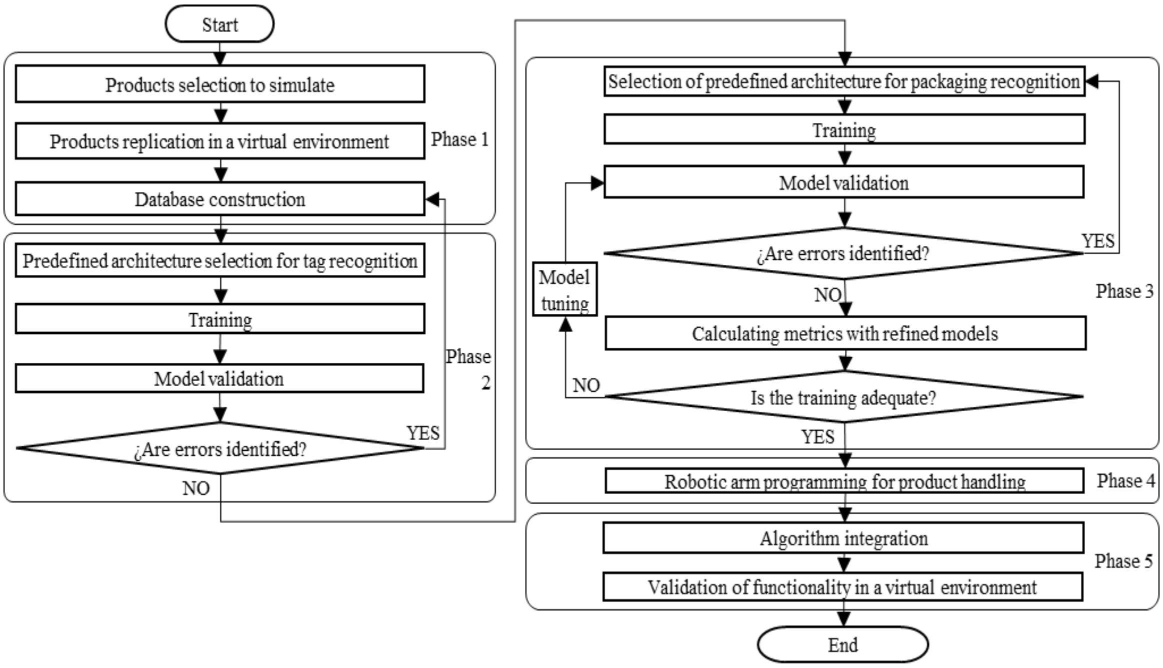


Figure 1. Methodology flowchart

To establish an industrial application that frames concepts of industry 5.0 in small and medium-sized companies, a liquid cleaning product are chosen, oriented towards most commonly used in both domestic and organizational environments. For this purpose, website of the micro company Industrias Novaquim [24], which offers a wide variety of formats, labels and designs of different cleaning products, was consulted, among others. This review allows understanding product preferences and ergonomics of packaging.

With this information, the selected products are filtered into their CAD formats. Five packaging formats are chosen corresponding to volumes of 500 cc, 800 cc, ½ gallon, 1 gallon and 2 gallons. For simulation specifications, six product types are selected: dishwasher, neutral cleaner, multipurpose cleaner, polymeric wax, bleach and ultra cleaner. Figure 2 exposes the comparison between real products and their representation.

Figure 2(a) shows the different presentations of these products in real photographs. Since 5 types of packaging, 6 types of labels, 7 product texture colors and 3 cap colors are used, 55 different products have been created for simulation, selecting the most representative and realistic combinations. In Figure 2(b), these configurations can be observed in the Coppelia simulation environment, showing their labels and colors.

(a) (b)

Figure 2. Comparison between (a) real products [24] and (b) their representation in Coppelia

To generate the CoppeliaSim database, objects are rotated in 10-degree increments, capturing images at each step to produce 1,980 images, divided into 70% for training, 15% for validation, and 15% for testing. To ensure robustness against orientation, scale, and illumination variations, the SURF algorithm is used. It begins by calculating points of interest from the determinant of the scale-normalized Hessian matrix, as shown in (1), where L_{xx}, L_{yy}, and L_{xy} are second-order image derivatives at scale σ. Descriptors are extracted and matched against a label database using a nearest-neighbor ratio threshold of 0.8, a uniqueness constraint, and a region of interest (ROI) is defined to improve accuracy. The label with more matches is selected, and a bounding box is assigned to localize the object.

$$\mathcal{H}(x, y, \sigma) = \sigma^2\left(L_{xx}(x, y, \sigma) \cdot L_{yy}(x, y, \sigma) - L_{xy}(x, y, \sigma)^2\right) \quad (1)$$

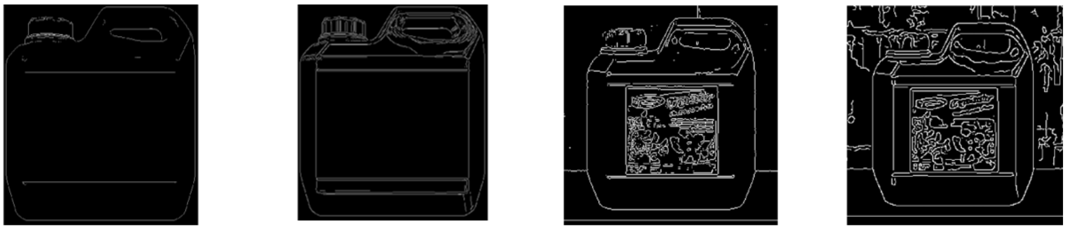
Figure 3 shows the edges of patterns to be used, Figure 3(a) corresponds to a 500 cc product, Figure 3(b) to an 800 cc product, Figure 3(c) to a ½ gallon product, Figure 3(d) to 1 gallon, and Figure 3(e) to 2 gallons, these patterns serve as a reference in classification strategy, allowing to evaluate effectiveness of the algorithms in identification and recognition of the different objects. Comparison of these edges with processed images is essential to ensure the accuracy and robustness of the methods to be evaluated.

(a) (b) (c) (d) (e)

Figure 3. Border patterns for cleaning products: (a) 500 cc, (b) 800 cc, (c) ½ gallon, (d) 1 gallon, and (e) 2 gallons

MATLAB connects to CoppeliaSim via its remote API, using a vision sensor to capture images from the simulation. Each image is converted to grayscale, and edge detection is performed using the Sobel and Canny methods [25]. Figure 4 shows the results for 2-gallon package: Sobel in Figure 4(a) and Canny in Figure 4(b) using the 3D viewer, and Sobel in Figure 4(c) and Canny in Figure 4(d) using CoppeliaSim. The simulated images introduce more noise, with Canny detecting more edges but also capturing excessive detail, which may hinder classification. Notably, both filters also detect the label edges, which is undesirable since only the outer contour is relevant. This confirms that both filters provide more detail than necessary for the task.

You only look once (YOLO) network is a deep learning model known for its speed and accuracy in object detection, used in real-time applications [26]. In this study the YOLO v2 architecture was configured to locate and classify five types of objects, other pretrained deep learning models considered such as Faster R-CNN and EfficientDet, which typically offer higher detection accuracy, especially for small or overlapping objects, at the cost of increased computational demand. YOLO v2 was ultimately chosen for its balance between speed and sufficient accuracy within the constraints of the targeted industrial scenario.



(a) (b) (c) (d)

Figure 4. Results of the application of filters on the 2-gallon package Sobel in (a) and Canny in (b) using the 3D viewer, and Sobel (c) and Canny in (d) using CoppeliaSim

The motion of the UR5 robot from it is current to target position and orientation is programmed using polynomial interpolation (tpoly) to generate smooth joint trajectories. Inverse kinematics is applied to compute the joint values corresponding to the target pose. At each step, the computed joint positions are sent to CoppeliaSim to simulate the robot's movement, while the end-effector position is recorded in an output matrix. The robot is programmed to classify an object by capturing images through a camera, identifying its label and packaging type using SURF and YOLO algorithms, respectively, and transporting the object to a designated shelf through a predefined sequence of poses. Once the task is completed, the robot returns to its initial position.

3. RESULTS AND DISCUSSION

After defining and implementing the detection and classification algorithms, this section presents the results of their application, as well as a comparative analysis of their performance. The code implemented in MATLAB for feature detection in images using SURF algorithm is executed with the Fast Hessian algorithm. Figure 5 shows the results of the SURF algorithm applied to different cleaning product formats, using the six selected labels: dishwasher, neutral cleaner, multipurpose cleaner, polymeric wax, bleach and ultra cleaner. These results confirm the correct performance of the algorithm in the task of label detection and classification, ensuring that the system can properly identify each product according to its corresponding label.

Figure 5. Results of SURF algorithm applied to different products formats with their respective labels

The Hopfield network operates on binary images represented as bit matrices, where each pixel is either 0 or 1. During the update process, it iteratively adjusts the bit values in the image until they converge to one of the stored patterns, thus minimizing system energy. This behavior makes it suitable for tasks like image retrieval and denoising. Therefore, a Hopfield network is trained using Hebb's rule [27], with five reference images processed with the Sobel filter and then this network is used to correct a distorted image, the objective is that the network converges to the reference image like the test image.

The algorithm was integrated with Coppelia, enabling direct image capture from the simulation environment. In Figure 6, the updates for each iteration of the Hopfield network are shown for a real-time image captured with the Sobel filter applied to enhance interaction with the environment. Figure 6(a) presents iteration 1, Figure 6(b) the iteration 2, Figure 6(c) the iteration 4 and Figure 6(d) the iteration 8. This iterative visualization demonstrates how the algorithm functions, converging to the trained pattern. This integration is essential for accurate product detection and improvement of system performance within the simulation.

The designed Hopfield network was then applied to the 1980 images, initially processing all images with an edge detection filter. This was done to evaluate the network’s performance and validate its accuracy. Therefore, the confusion matrix for the 5 objects is presented in Table 1. This table reveals that the Hopfield network is not an ideal strategy, as a single 2D pattern cannot represent the model in all possible orientations.

Additionally, the network struggles to discriminate noise and is only effective for binary images with few pixels. Although the network is fast, simple to program, and execute, it is not functional for this application.

(a) (b) (c) (d)

Figure 6. Iterative updates of the Hopfield network for image: Iteration (a) 1, (b) 2, (c) 4, and (d) 8

Table 1. Confusion matrix for classification of 5 objects using the Hopfield algorithm

	Predicted 500 cc	Predicted 800 cc	Predicted ½ gal	Predicted 1 gal	Predicted 2 gal
Actual 500 cc	150 (37.88%)	64	73	52	57
Actual 800 cc	50	180 (45.45%)	63	66	37
Actual ½ gal	74	46	100 (25.25%)	79	61
Actual 1 gal	66	49	71	140 (35.35%)	50
Actual 2 gal	63	44	57	56	150 (37.88%)

2D cross-correlation is a powerful technique for comparing binary images, allowing the detection of similarities and patterns through displacements [28]. To determine the cross-Correlation for two binary images, first, it is verified that they have compatible dimensions so as not to generate error; then, it is calculated in MATLAB and the maximum value of the resulting correlation matrix is extracted, indicating the highest degree of coincidence between the two images in question. To compare the cross-correlation, three examples of edges calculated in the Coppelia simulation software were selected, the results are presented at Table 2, which will be used to quantify the correlation metric and evaluate the effectiveness of the classification strategy.

Table 2. Cross-correlation metric results for pattern edges

Cross-correlation	500cc	800cc	½ gallon	1 gallon	2 gallons
½ gallon	0.120322	0.121047	**0.122017**	0.106261	0.094171
800cc	0.122909	**0.164446**	0.096968	0.109216	0.109216
1 gallon	0.122399	0.124067	0.094241	**0.188348**	0.095863

This metric was then applied to all 1980 images, once again starting with the Sobel filter and calculating correlation between patterns and image taken directly from CoppeliaSim. The confusion matrix for classification of the 5 objects is presented in Table 3. As seen in the table, the results improved significantly with this method, showing strong performance for the 500cc and 1-gallon objects. This improvement occurred because the shape of the solids does not change significantly when rotated. However, the classification for the other objects, especially the ½ gallon and 2-gallon objects, was not as accurate due to their more complex geometry. Although this strategy is simple and fast, it is only effective for cylindrical objects. For objects with rectangular prism shapes, this method proves inefficient for classification.

Table 3. Confusion matrix for classification of 5 objects using 2D cross-correlation

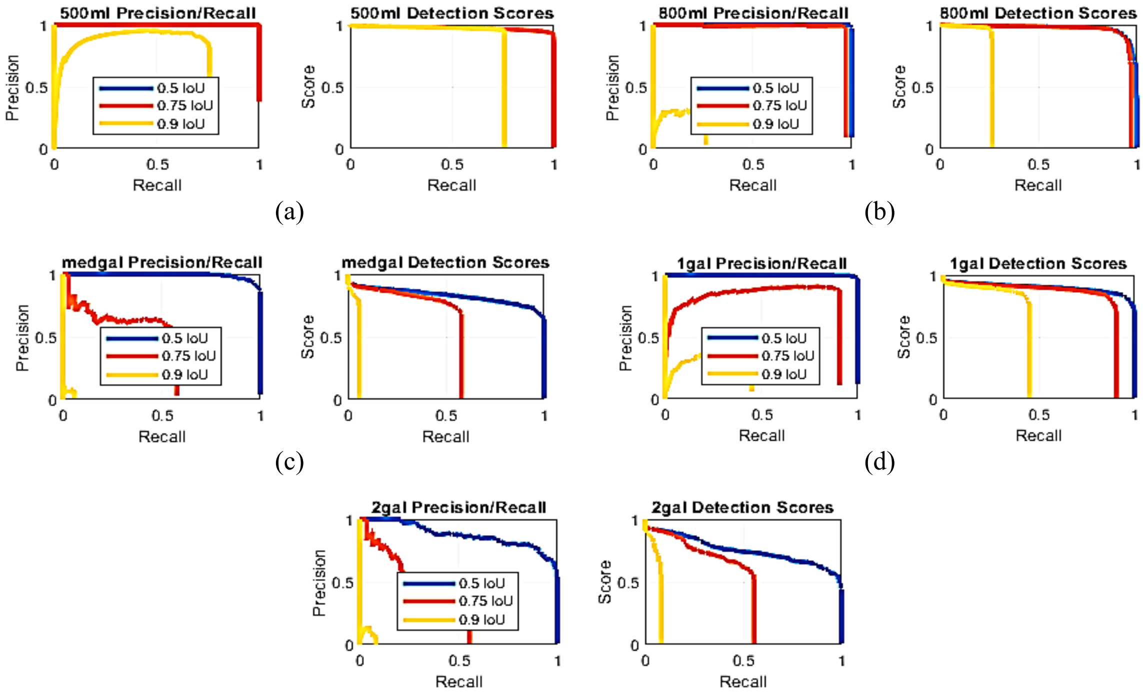
	Predicted 500cc	Predicted 800cc	Predicted ½ gal	Predicted 1 gal	Predicted 2 gal
Actual 500cc	360 (90.91%)	12	15	3	6
Actual 800cc	28	246 (62.12%)	47	35	40
Actual ½ gal	50	60	162 (40.91%)	75	49
Actual 1 gal	16	14	25	321 (81.06%)	20
Actual 2 gal	55	44	66	53	178 (44.95%)

Both the cross-correlation and the SURF algorithm were integrated to precisely define the type of object in camera's field of view and thus proceed to its classification. Figure 7 shows the metrics obtained and detection of objects in two cases, using images taken directly from the Coppelia simulation environment. Figure 7(a) shows the 800cc container and Figure 7(b) show a product in ½ gallon container. This fusion of techniques allows to improve the effectiveness in the identification and classification of the analyzed products.

The YOLO configuration aimed to balance training time and detection accuracy for five distinct product containers. The process began with loading a pretrained model and defining the network architecture, including anchor box estimation and an input image size of 720×720 pixels with three color channels. During training, feature extraction layer was set to leaky ReLU 5, with five anchor boxes and the RMSProp optimizer. The key training parameters included an initial learning rate of 0.001, mini-batch size of 8, a maximum of 10 epochs, a piecewise learning rate schedule, and a drop period of 5 epochs. The training process, converges near iteration 10 where the loss error is reduced to zero, shows the evolution of the error loss across epochs, highlighting the model's convergence behavior and learning stability. Figure 8 shows the precision vs. recall plots and the score vs. recall plots for each of the five objects, using the test images. Subplots correspond to: Figure 8(a) 500 cc, Figure 8(b) 800 cc, Figure 8(c) ½ gallon, Figure 8(d) 1 gallon, and Figure 8(e) 2 gallons. When analyzing these plots, it is evident that the model performed adequately in detecting all five objects during testing, with the best results observed for the 500 ml, 800 ml, and 1-gallon containers.

(a) (b)

Figure 7. Object detection and classification results using cross correlation and SURF algorithm for cleaning products: (a) 800cc and (b) ½ gallon



(a) (b) (c) (d) (e)

Figure 8. Accuracy and recall plots for trained objects: (a) 500 cc, (b) 800 cc, (c) ½ gallon, (d) 1 gallon, and (e) 2 gallons

These five categories showed higher precision and recall values, indicating better generalization by the trained model. In Table 4, the confusion matrix for the classification results is presented, results now demonstrate a valid classification performance. The network's ability to correctly detect objects, especially for the 500 cc, 800 cc, and 1-gallon objects, is noteworthy, with accuracy percentages above 92%. Despite this, YOLO model has significantly improved classification performance, compared to the previous detection approaches, where only the 500cc object was successfully detected. This suggests that the model can now reliably classify most of the objects, though further tuning and additional data may be required to enhance detection for more complex shapes like the 2-gallon container.

Table 4. Confusion matrix for classification of 5 objects using YOLO v2 algorithm

	Predicted 500 cc	Predicted 800 cc	Predicted ½ gal	Predicted 1 gal	Predicted 2 gal
Actual 500 cc	384 (97.02%)	4	6	5	2
Actual 800 cc	5	376 (95.20%)	5	5	5
Actual ½ gal	4	6	340 (86.17%)	2	44
Actual 1 gal	5	7	10	364 (92.17%)	12
Actual 2 gal	5	6	52	8	325(82.07%)

Results shown in Table 5 summarize detection accuracy of different algorithms implemented for object classification. It is evident that the YOLO algorithm provides the best performance across all object types, achieving high accuracy, especially for objects with distinct shapes. The Hopfield algorithm is not recommended for applications where object shapes are very similar, as its performance is highly dependent on the chosen filter and number of pixels used. The algorithm's accuracy tends to degrade with high pixel values, as it loses convergence, and with low pixel values, the quality of the edge detection is compromised. While it does not require training, only the initial definition of patterns, its execution time averaged 237 ms for the images used, making it relatively fast but less effective for complex object shapes.

Table 5. Detection accuracy comparison of algorithms for classifying objects

Detection accuracy of algorithms for any object	Hopfield algorithm with Sobel filter	2D Cross-correlation with Sobel filter	YOLO v2 algorithm
500 cc	37.88%	90.91%	97.02%
800 cc	45.45%	62.12%	95.20%
½ gal	25.25%	40.91%	86.17%
1 gal	35.35%	81.06%	92.17%
5 gal	37.88%	44.95%	82.07%
Mean	36.36%	63.99%	90.53%

The 2D cross-correlation metric was proposed as a simple alternative for object classification. This method performed well for cylindrical objects, as their orientation does not significantly alter their silhouette. However, it is not recommended for solids with more complex geometries. Despite being the fastest in execution, with a processing time of just 18 ms, its application is limited to simpler shapes. Given limitations of the other algorithms, YOLO v2 algorithm was chosen. It leverages a pre-trained database of the environment in which objects will be used, allowing for focused training on object detection. Although the training process is time-consuming, taking 193 minutes, and its implementation is slower than the others, with an average execution time of 825 ms, the high accuracy in detection makes it the best choice for this application.

Given that the primary objective of this work was the classification of objects based on their appearance, rather than precise localization or multi-object detection, the focus was placed on comparing basic yet representative classification strategies. This choice was aligned with the simplicity of the simulated industrial scenario and the controlled conditions of the task. In this context, the implementation of methods such as SURF, 2D cross-correlation, and YOLO configured mainly for single-object classification—proved effectively. The results obtained validate the applicability of these strategies in structured environments, with the added benefit of rapid execution times. While more advanced deep learning models could offer improved performance in more complex scenes, the current approach provides a reliable and computationally efficient solution suitable for prototyping and experimentation in semi-controlled settings.

After integrating all the algorithms, the simulation environment was updated with the corresponding 3D models of the containers, enabling a complete test of the detection, classification, and manipulation pipeline. Figure 9 illustrates the final setup, where the UR5 robot identifies each object using the vision system, classifies it based on the trained models, and executes a pick-and-place task according to the assigned category. This integration confirms the coherence between visual recognition and robotic control within a fully simulated industrial scenario.

Figure 9. Simulation environment with integrated solids for transport and handling

4. CONCLUSION

This work demonstrated viability of combining classical image processing methods with deep learning techniques for object classification and robotic manipulation in a simulated industrial context. While traditional approaches such as SURF, edge filters, Hopfield networks, and cross-correlation offered fast execution and simplicity, they showed clear limitations in accuracy and adaptability when handling complex shapes or noisy environments. In contrast, YOLO v2 model exhibited significantly higher classification performance and robustness. Given the simplicity of the simulated setup and the goal of comparing classification strategies rather than full object detection pipelines, the applied methods proved effective, confirming that basic techniques can be useful for structured environments where computational resources are limited. To strengthen contribution and situate it within the current state of the art, future work must include comparisons with more recent pretrained models such as YOLO v5, Faster R-CNN, or EfficientDet. These models are likely to offer improved performance in terms of generalization and detection under more challenging conditions. Moreover, expanding the dataset, testing with real-world sensors, and evaluating against standardized benchmarks will be essential to validate system's scalability and relevance. The methodology lays a replicable foundation for integrating computer vision and robotic control in Industry 5.0-inspired environments and highlights importance of strategies that consider both accuracy and computational efficiency.

ACKNOWLEDGEMENTS

The authors would like to thank the Universidad Militar Nueva Granada, from where they are full-time associate professors.

FUNDING INFORMATION

Product derived from the research project titled "*Mejoramiento de procesos de control de inventario y alistamiento de pedidos en el marco de la industria 4.0*" INV-ING-3970 financed by the vice-rector for research of the Universidad Militar Nueva Granada, year 2024.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Juan Camilo Guacheta-Alba		✓	✓	✓	✓	✓		✓	✓	✓				
Anny Astrid Espitia-Cubillos	✓	✓					✓				✓	✓	✓	✓
Robinson Jiménez-Moreno	✓	✓		✓	✓					✓	✓	✓		

C	:	**C**onceptualization	I	:	**I**nvestigation	Vi	:	**Vi**sualization
M	:	**M**ethodology	R	:	**R**esources	Su	:	**Su**pervision
So	:	**So**ftware	D	:	**D**ata Curation	P	:	**P**roject administration
Va	:	**Va**lidation	O	:	Writing - **O**riginal Draft	Fu	:	**Fu**nding acquisition
Fo	:	**Fo**rmal analysis	E	:	Writing - Review & **E**diting			

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author.

REFERENCES

[1] H.-W. Lee, M.-I. Roh, Y.-M. Cho, and J.-H. Park, "Data association for autonomous ships based on virtual simulation environment," *Ocean Engineering*, vol. 281, p. 114646, Aug. 2023, doi: 10.1016/j.oceaneng.2023.114646.

[2] M. M. Alam and S. M. M. Rahman, "Affine transformation of virtual 3D object using 2D localization of fingertips," *Virtual Reality & Intelligent Hardware*, vol. 2, no. 6, pp. 534–555, Dec. 2020, doi: 10.1016/j.vrih.2020.10.001.

[3] H. Li *et al.*, "Virtual point tracking method for online detection of relative wheel-rail displacement of railway vehicles," *Reliability Engineering & System Safety*, vol. 246, p. 110087, Jun. 2024, doi: 10.1016/j.ress.2024.110087.

[4] Y. Ding and X. Luo, "A virtual construction vehicles and workers dataset with three-dimensional annotations," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 107964, Jul. 2024, doi: 10.1016/j.engappai.2024.107964.

[5] H. Zhang, T. Li, Y. Li, and Z. Wen, "Virtual electromagnetic environment modeling based data augmentation for drone signal identification," *Journal of Information and Intelligence*, vol. 1, no. 4, pp. 308–320, Nov. 2023, doi: 10.1016/j.jiixd.2023.08.002.

[6] Y. Shahedi, M. Zandi, and M. Bimakr, "A computer vision system and machine learning algorithms for prediction of physicochemical changes and classification of coated sweet cherry," *Heliyon*, vol. 10, no. 20, p. e39484, Oct. 2024, doi: 10.1016/j.heliyon.2024.e39484.

[7] M. Dutta and A. Ganguly, "Incremental-based YOLOV3 model with hyper-parameter optimization for product image classification in E-commerce sector," *Applied Soft Computing*, vol. 165, p. 112029, Nov. 2024, doi: 10.1016/j.asoc.2024.112029.

[8] J. Leng *et al.*, "Unlocking the power of industrial artificial intelligence towards Industry 5.0: insights, pathways, and challenges," *Journal of Manufacturing Systems*, vol. 73, pp. 349–363, Apr. 2024, doi: 10.1016/j.jmsy.2024.02.010.

[9] M. Jasiulewicz - Kaczmarek and A. Gola, "Maintenance 4.0 technologies for sustainable manufacturing - an overview," *IFAC-PapersOnLine*, vol. 52, no. 10, pp. 91–96, 2019, doi: 10.1016/j.ifacol.2019.10.005.

[10] E. Valette, H. Bril El-Haouzi, and G. Demesure, "Industry 5.0 and its technologies: a systematic literature review upon the human place into IoT- and CPS-based industrial systems," *Computers & Industrial Engineering*, vol. 184, p. 109426, Oct. 2023, doi: 10.1016/j.cie.2023.109426.

[11] K. Wan, C. Li, F.-S. Lo, and P. Zheng, "A virtual reality-based immersive teleoperation system for remote human-robot collaborative manufacturing," *Manufacturing Letters*, vol. 41, pp. 43–50, Oct. 2024, doi: 10.1016/j.mfglet.2024.09.008.

[12] M. Andersson and A. Syberfeldt, "Improved interaction with collaborative robots - evaluation of event-specific haptic feedback in virtual reality," *Procedia Computer Science*, vol. 232, pp. 1055–1064, 2024, doi: 10.1016/j.procs.2024.01.104.

[13] S. Mustary, M. A. Kashem, M. A. Chowdhury, and M. M. Rana, "Mathematical model and evaluation of dynamic stability of industrial robot manipulator: universal robot," *Systems and Soft Computing*, vol. 6, p. 200071, Dec. 2024, doi: 10.1016/j.sasc.2023.200071.

[14] T. Toner, M. Saez, D. M. Tilbury, and K. Barton, "Opportunities and challenges in applying reinforcement learning to robotic manipulation: an industrial case study," *Manufacturing Letters*, vol. 35, pp. 1019–1030, Aug. 2023, doi: 10.1016/j.mfglet.2023.08.055.

[15] A. Efendi, Y.-H. Shao, and C.-Y. Huang, "Technological development and optimization of pushing and grasping functions in robot arms: a review," *Measurement*, vol. 242, p. 115729, Jan. 2025, doi: 10.1016/j.measurement.2024.115729.

[16] V. Priyanka, Y. S. Rama, K. Sravani, and B. Kavya, "Implementation of sobel edge detection with image processing on FPGA," in *2024 2nd World Conference on Communication & Computing (WCONF)*, Jul. 2024, pp. 1–5, doi: 10.1109/WCONF61366.2024.10692301.

[17] L. L. Vaishnavi, M. Bhavana, D. C. M. Manideep, N. N, and V. Venugopal, "A fast and accurate object counting using sobel edge detection system for real – time star gazing," in *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*, Jul. 2024, pp. 1–5, doi: 10.1109/APCIT62007.2024.10673462.

[18] H. Gao, X. Peng, X. Pang, and X. Feng, "An efficient target recognition method based on mixed binary SURF descriptor with region segmentation," in *2024 43rd Chinese Control Conference (CCC)*, Jul. 2024, pp. 7339–7344, doi: 10.23919/CCC63176.2024.10662605.

[19] L. Ouyang, R. Lu, Y. Ye, M. Xing, and Q. Cai, "Feature point extraction and matching based on improved SURF algorithm," in *2023 China Automation Congress (CAC)*, Nov. 2023, pp. 1846–1851, doi: 10.1109/CAC59555.2023.10450357.

[20] R. Hamzah, L. Ang, R. Roslan, N. H. Ibrahim Teo, K. Abdul Samad, and K. A. Fariza Abu Samah, "Comparing modified Yolo V5 and faster regional convolutional neural network performance for recycle waste classification," in *2024 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, Jun. 2024, pp. 415–419, doi: 10.1109/I2CACIS61270.2024.10649835.

[21] V. S. Muraviev and P. E. Zhgutov, "The influence of transfer learning on the accuracy of object localization and classification in the YOLO detector," in *2024 26th International Conference on Digital Signal Processing and its Applications (DSPA)*, Mar. 2024, pp. 1–5, doi: 10.1109/DSPA60853.2024.10510060.

[22] A. Brey, J. J. Quintana, M. Diaz, and M. A. Ferrer, “Smartphone-based control system for universal robot UR5e: a tool for robotics education,” in *2024 XVI Congreso de Tecnología, Aprendizaje y Enseñanza de la Electrónica (TAEE)*, Jun. 2024, pp. 1–5, doi: 10.1109/TAEE59541.2024.10604994.
[23] J. Villalobos, I. Y. Sanchez, and F. Martell, “Statistical comparison of Denavit-Hartenberg based inverse kinematic solutions of the UR5 robotic manipulator,” in *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Oct. 2021, pp. 1–6, doi: 10.1109/ICECCME52200.2021.9591104.
[24] Industrias Novaquim SAS, “Cleaning products tailored to your needs,” (in Spanish). Industrias Novaquim SAS, https://www.novaquim.com/productos/categoria/limpieza-general/ (accessed Jul. 02, 2025).
[25] D. R. D. Varma and R. Priyanka, “Performance monitoring of novel iris detection system using sobel algorithm in comparison with canny algorithm by minimizing the mean square error,” in *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, Apr. 2022, pp. 509–512, doi: 10.1109/ICIEM54221.2022.9853127.
[26] H. Liu, “Fine-grained classification and anomaly detection system of motion injury images based on improved YOLO algorithm,” in *2024 International Conference on Data Science and Network Security (ICDSNS)*, Jul. 2024, pp. 1–5, doi: 10.1109/ICDSNS62112.2024.10690861.
[27] A. Kapoor, N. Singh, and V. Motilal, “Modelling social networks using modified hopfield neural network and identify leadership traits,” in *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, Feb. 2019, pp. 1–6, doi: 10.1109/ICACCP.2019.8882892.
[28] Y. Kinoshita, K. Yamaoka, and H. Kiya, “Maximization of 2D cross-correlation based on auxiliary function method for image alignment,” in *2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Oct. 2023, pp. 2043–2047, doi: 10.1109/APSIPAASC58517.2023.10317371.

BIOGRAPHIES OF AUTHORS

Juan Camilo Guacheta-Alba received his B.S. in mechatronics engineering from Universidad Militar Nueva Granada in 2020 and his M.Sc. in mechatronics engineering from Universidad Militar Nueva Granada in 2022. Currently, he is working toward the Ph.D. in mechanical engineering at Universidade Federal do Rio de Janeiro. His research interests include robotics, multi-agent systems and optimization. He can be contacted at email: juan.camilo@mecanica.coppe.ufrj.br.

Anny Astrid Espitia-Cubillos performed her undergraduate studies in industrial engineering in the Universidad Militar Nueva Granada in 2002 and M.Sc. in industrial engineering from the Universidad de Los Andes in 2006. She is an associate professor on industrial engineering program at Universidad Militar Nueva Granada, Bogotá, Colombia. She can be contacted at email: anny.espitia@unimilitar.edu.co.

Robinson Jiménez-Moreno is an electronic engineer graduated from Universidad Distrital Francisco José de Caldas in 2002. He received a M.Sc. in engineering from Universidad Nacional de Colombia in 2012 and Ph.D. in engineering at Universidad Distrital Francisco José de Caldas in 2018. He is associate professor of Universidad Militar Nueva Granada and his research focuses on the use of convolutional neural networks for object recognition and image processing for robotic applications such as human-machine interaction. He can be contacted at email: robinson.jimenez@unimilitar.edu.co.