Estimation of Package-Boundary Confidence for Object Recognition in Rainbow-SKU Depalletizing Automation

Kento Sekiya[1][a], Taiki Yano[1][b], Nobutaka Kimura[2][c] and Kiyoto Ito[1][d]

[1]*Research & Development Group, Hitachi, Ltd., Kokubunji, Tokyo, Japan*

[2]*Research & Development Division, Hitachi America, Ltd., Holland, Michigan, U.S.A.*

Keywords: Object Recognition, Robot, Depalletizing, Boundary Confidence.

Abstract: We developed a reliable object recognition method for a rainbow-SKU depalletizing robot. Rainbow SKUs include various types of objects such as boxes, bags, and bottles. The objects' areas need to be estimated in order to automate a depalletizing robot; however, it is difficult to detect the boundaries between adjacent objects. To solve this problem, we focus on the difference in the shape of the boundaries and propose package-boundary confidence, which assesses whether the recognized boundary correctly corresponds to that of an object unit. This method classifies recognition results into four categories on the basis of the objects' shape and calculates the package-boundary confidence for each category. The results of our experimental evaluation indicate that the proposed method with slight displacement, which is automatic recovery, can achieve a recognition success rate of 99.0 %. This is higher than that with a conventional object recognition method. Furthermore, we verified that the proposed method is applicable to a real-world depalletizing robot by combining package-boundary confidence with automatic recovery.

1 INTRODUCTION

Rainbow-SKU depalletizing, which is the process of picking up various types of objects from a loaded pallet, is a strenuous manual task, so automating the task with robots is highly desirable. Many researchers have proposed depalletizing systems for automating robots by combining robot motion planning with image recognition (Nakamoto et al., 2016; Eto et al., 2019; Doliotis et al., 2016; Aleotti et al., 2021; Caccavale et al., 2020; Katsoulas and Kosmopoulos, 2001; Kimura et al., 2016).

Automated robots need to complete a series of picking tasks accurately and quickly in order to be applicable in warehouses. If robots pick incorrect objects, workers must perform a manual recovery, e.g., remote control, which increases downtime. There are several causes of incorrect picking, such as the shortage of adsorption power in the robot hand and false estimation of an object's position or pose. To address the hardware problem, robot hands have been developed which can grasp objects of various shapes (Tanaka et al., 2020; Fontanelli et al., 2020). In the software, a function is needed to estimate objects' boundaries from images and point clouds. However, there are few methods which have been successfully used to estimate the areas of all types of objects' in rainbow-SKU depalletizing. This is because object boundaries differ depending on the shape and material of the object, e.g., cardboard, bags, rolls of toilet paper, and shrink-wrapped packages containing multiple bottles or cans in transparent wrapping. It is also difficult to divide multiple objects placed adjacent to one another because of the missing the boundary between the objects. Without the boundary, multiple objects are recognized as one object and robots incorrectly pick multiple objects at the same time.

To estimate object boundaries with high accuracy, we introduce package-boundary confidence, which assesses whether the recognized boundary correctly corresponds to that of an object unit. When the confidence is high, the robots pick the object, and when the confidence is low, the robots do not pick the object and switch to automatic recovery mode. In this study, we use slight displacement as an automatic recovery, which is to pick the edge of object and move it a short distance. By doing this, the gap between multiple ob-

[a] https://orcid.org/0009-0007-1667-3081
[b] https://orcid.org/0000-0001-9433-0569
[c] https://orcid.org/0000-0001-5248-5108
[d] https://orcid.org/0000-0002-2243-5756

Sekiya, K., Yano, T., Kimura, N. and Ito, K.
Estimation of Package-Boundary Confidence for Object Recognition in Rainbow-SKU Depalletizing Automation.
DOI: 10.5220/0012307300003660

In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024) - Volume 3: VISAPP*, pages 309-316
ISBN: 978-989-758-679-8; ISSN: 2184-4321

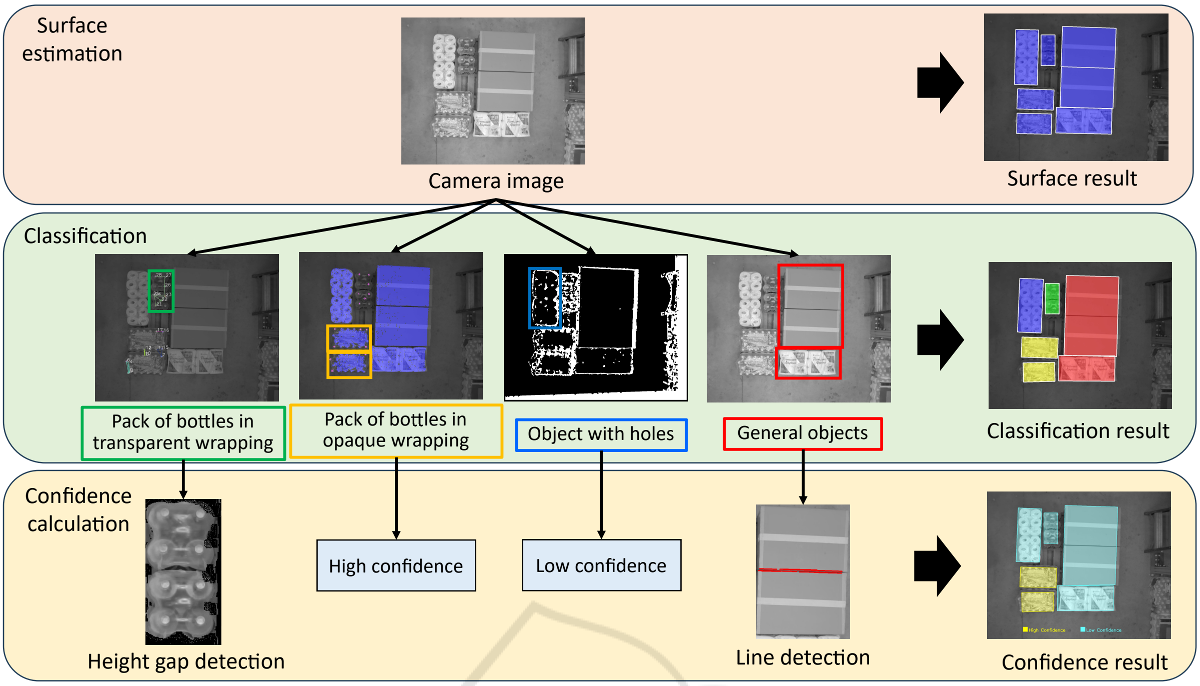


Figure 1: Overview of the proposed method, which consists of three steps: surface estimation, classification, and confidence calculation. By classifying the results of surface estimation into four categories depending on the objects' shape and calculating confidence for each category, the proposed method can achieve reliable recognition for rainbow SKUs.

jects will be correctly recognized as distinct boundaries. Although slight displacement is quicker than manual recovery, doing it too frequently causes the throughput to deteriorate. Therefore, a recognition method with both high accuracy and high throughput is necessary.

In this paper, we propose a method for calculating package-boundary confidence. Since robots need to recognize the boundaries of various types of objects, we classify objects into one of four categories and change how to calculate confidence depending on the object's boundaries. We conducted experiments to simulate a rainbow-SKU depalletizing process using a 3D vision sensor. The results indicate that the proposed method achieves a success rate of 99.0%, which is higher than that with a conventional object recognition method. We also evaluated the frequency of slight displacement, which was 37.5%. These results show that the proposed method is applicable for a wide variety of objects in rainbow-SKU depalletizing.

2 RELATED WORK

In this section, we discuss conventional object recognition methods used for depalletizing.

2.1 Deep-Learning-Based Segmentation

Deep-learning-based segmentation has been used for depalletizing recognition (Girshick, 2015; Liu et al., 2016; Redmon et al., 2016; He et al., 2017). This method estimates objects' areas and classifies the areas into classes simultaneously. In recent years, deep-learning methods are applied to object recognition for depalletizing (Buongiorno et al., 2022). However, there is no large dataset of rainbow-SKU object, so deep-learning methods are applied to a limited variety of objects such as cardboard.

2.2 Edge-Based Boundary Detection

Conventionally, edge-based boundary detection has been widely used in region estimation for boxed objects (Katsoulas and Kosmopoulos, 2001; Naumann et al., 2020; Stein et al., 2014; Li et al., 2020). By estimating the edges on the basis of the luminance gradient or degree of change in the normal direction, the boundary of each object can be recognized. However, due to the difficulty in detecting the edges of wraps that cannot be measured, there is a risk of detecting individual products as a single object rather than an entire shrink-wrapped package. The proposed method is an extended approach of this type of method; our method does not require a large dataset to train the model, and it can be applied to cardboard packages as well as shrink-wrapped objects in rainbow-SKU depalletizing.

3 METHODS

3.1 Concept of Object Recognition

This section describes the concept of the proposed method for recognizing various objects. In rainbow-SKU depalletizing, the shape of object boundaries is different from each other, such as a gap, a straight line, and part of an arc. Also, because shrink-wrapped objects have small gaps between individual objects, gaps must be distinguished from the correct boundary. Therefore, it is difficult to evaluate the recognized boundary consistently.

In the proposed method, we classify objects into four categories and calculate package-boundary confidence in different ways for each categories. The number of categories are determined taking into the varieties of rainbow SKUs in the warehouse. To recognize various objects including packs of wrapped bottles, we selected the object recognition architecture from (Yano et al., 2023) as the base architecture in this research.

3.2 Classification

This section describes how objects are classified into the four categories. The previous method (Yano et al., 2023) estimated object surfaces from gray-scale images and point clouds. Therefore, in the present study we use information based on object surfaces and classify them into one of the following four categories: pack of bottles in transparent plastic wrapping, pack of bottles in opaque wrapping, object with holes, and general object.

Multiple bottles shrink wrapped in transparent plastic are defined as a pack of bottles in transparent wrapping. The bottle caps are regarded as small surfaces. Because these surfaces are too small to be recognized as a single object unit, multiple surfaces are connected and recognized as a single object unit. Objects recognized from connecting areas are classified into this category.

Multiple bottles shrink wrapped in opaque plastic are defined as a pack of bottles in opaque wrapping. The opaque wrapping is recognized as a large surface in the middle areas of the object, and the bottles are recognized as small surfaces in the surrounding areas. As shown in Figure 2, to detect bottle caps, we detect circles with Hough transformation for gray-scale images and calculate the ratio of circles in surrounding areas to that in the middle areas (Yuen et al., 1990). If the ratio of circles is high, the object is classified into this category.

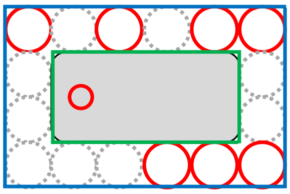


Figure 2: Circle detection for pack of bottles in opaque wrapping. Middle areas are internal green line and surrounding areas are between green and blue lines.

Objects which have gap areas in depth inside objects are defined as an object with holes. As shown in Figure 3, we focused on a fact that depth information of such objects have several holes corresponding to tube holes or empty areas between tubes that are touching. We make depth images from point clouds and calculate the depth of gap areas. If there are many gaps, the object is classified into this category.

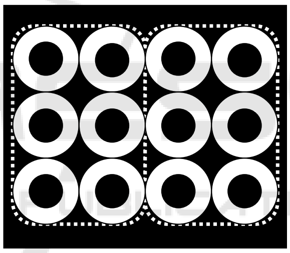


Figure 3: Depth image of object with holes. Black areas indicate tube holes or empty areas.

Finally, objects which are not classified into the previous three categories are defined as a general object.

3.3 Confidence Calculation

This section describes how package-boundary confidence is calculated for various objects. The method of calculating confidence changes depending on the category in which the object has been classified.

3.3.1 Pack of Bottles in Transparent Wrapping

The boundaries of this type of objects are the gaps between several connected areas. As shown in Figure 4, four objects (i.e., packs of bottles) are placed adjacent. Multiple bottles are connected by graphs, but several object units are over-connected. In this situation, we calculate the depth of the gap areas on the graphs. If there are large deep gaps, the recognition areas do not need to be divided, and we set the confidence to low. On the other hand, if the graphs do not have large deep gaps, we set the confidence to high.

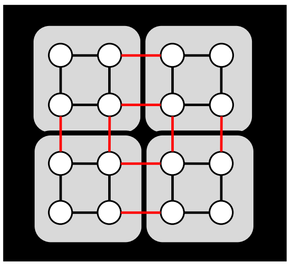


Figure 4: Four objects (packs of bottles in transparent wrapping) placed adjacent. Red graphs mean overlapping large deep gaps.

3.3.2 Pack of Bottles in Opaque Wrapping

The boundaries of this type of objects are bottle caps in the surrounding areas. Because the opaque wrapping consists of large surfaces, the recognition areas of large surfaces can be detected reliably. Also, if multiple objects are placed adjacent, large surfaces do not overlap with each other because there are bottle caps between the large surfaces. Therefore, the confidence is always set to high for this type of object.

3.3.3 Object with Holes

The boundaries of this type of objects are deep gaps. However, it is difficult to determine whether single object unit is really single object unit or be separated into multiple object units. This is because both correct boundaries and gaps inside objects are similar deep gaps. Therefore, the confidence is always set to low for this type of object.

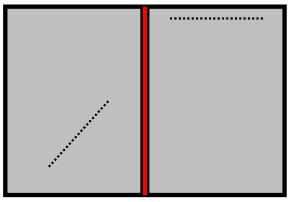


Figure 5: Line detection for a general object. Dotted lines are removed lines and a red line is used for confidence calculation.

3.3.4 General Object

The boundaries of this type of objects are straight lines. This type of objects do not include bottles or paper rolls, and one unit is square-shaped in depalletizing. If these objects are placed adjacent, boundaries can be detected as a pattern of straight lines even if there are no gaps between objects. As shown in Figure 5, we use line detection with Hough transformation for a gray-scale image and a depth image (Duda and Hart, 1972). We also remove detected lines which are not vertical to the side of recognition areas and are near the side of recognition areas. This is because these lines are not the boundaries which divide recognition areas into multiple objects. We calculate the number of lines in recognition areas which are not removed. If there are many lines, the recognition areas do not need to be divided and the confidence is set to low. On the other hand, if there are few lines, the confidence is set to high.

3.4 Two Parameter Sets for Various Objects

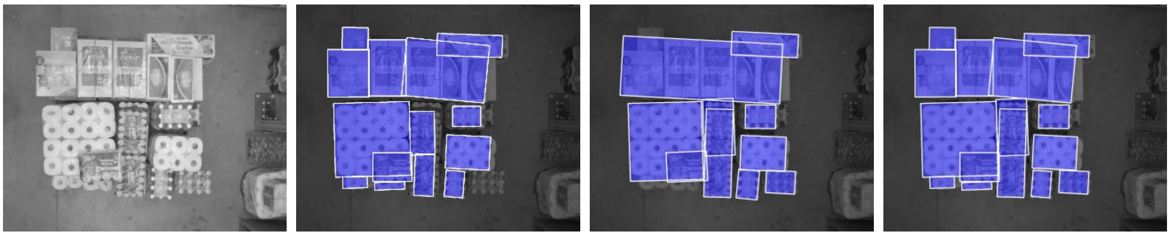
In the previous method (Yano et al., 2023), it was difficult to correctly detect both large top surfaces as well as small top surfaces such as bottle caps using only a single parameter set for object recognition.

The first parameter set is adjusted for detecting even small and thin edges. As shown in Figure 6b, when using the first parameter set, the algorithm correctly divides boxes which are touching, but it detects many edges from complicated measured data, such as bottles in transparent wrapping, and divides them into many small surfaces. As a results, it fails to detect packs of bottles.

The second parameter set is adjusted to ignore small and thin edges. As shown in Figure 6c, when using the second parameter set, the algorithm successfully detects bottles in transparent wrapping. However, it fails to divide boxes that are touching because it ignores the relatively thin boundary. Based on this preliminary trial, we apply each parameter set individually and integrate the two results as shown in Figure 6d.

3.5 Slight Displacement

This section describes how slight displacement is performed for automatic recovery. As mentioned in 1, package-boundary confidences are used to detect successful results of object recognition as well as to switch the process of robot motion. When the package-boundary confidence is low, slight displacement is conducted so that multiple objects placed adjacent have enough gaps between them (Figure 7). Then, the object surfaces are estimated again and the recognition is successful.



(a) Raw image (b) First parameter (c) Second parameter (d) Integrated results

Figure 6: Two parameter sets for surface estimation. Results of first parameter are divided into many surfaces. Results of second parameter are undivided for multiple objects.

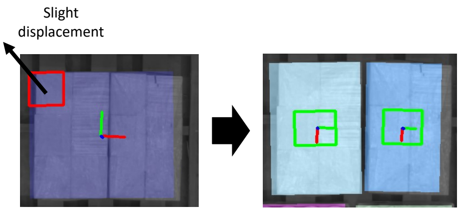


Figure 7: Improvement of recognition by slight displacement.

4 EXPERIMENTS AND RESULTS

4.1 Experimental Setups

In the experiment, we collect datasets of gray-scale images and point clouds using the vision system of the depalletizer. The vision system is a TVS 4.0 vision sensor, a 3D vision head with two cameras and an industrial projector, with a resolution of 1280 x 1024. The height of the vision sensor from the floor surface is 3,200 mm. The 32 types and eight groups of objects to be recognized are shown in Figure 8. In the proposed method, these objects are classified into four categories.

To evaluate the recognition rate of the proposed method, we selected five pairs of objects from the SKUs as shown in Figure 8 and arranged them so that the objects in each pair would be close to each other as shown in Figure 9. An example of the situation where the SKUs in one pair of objects are same is shown in Figure 9a, and that where the SKUs are different is shown in Figure 9b, respectively. We captured 300 images while changing the gaps between paired objects to 0 mm, 10 mm, and 20 mm, and applied our technique to these images.

4.2 Definition of Successful Recognition

A recognition is defined as successful when the robot picks the correct object. Whether the robot avoids incorrect picking is determined by the object area of surface estimation and the confidence whose threshold is set as 0.5 (Table 1). If the object area from the surface estimation is correct, the recognition is successful regardless of the confidence, because when the confidence is high, the robot picks the object directly, and when the confidence is low, the robot slightly displaces the object and the second attempt at recognition is successful. If the object area from the surface estimation is incorrect, the result can either be a success or failure. When the confidence is high, the robot picks the wrong object and the resulting recognition is a failure, whereas when the confidence is low, the robot slight displaces the object.

Table 1: Definition of success.

Object area	Confidence	Result
Correct	$\geqq 0.5$ (high)	Success (Direct picking)
Correct	< 0.5 (low)	Success (Slight displacement)
False	$\geqq 0.5$ (high)	Failure
False	< 0.5 (low)	Success (Slight displacement)

4.3 Results

Figure 10a and Figure 10b show example results of the tests where object pairs with the same and different SKUs were placed close to each other, respectively. There are two patterns of object arrangements in each figure, and we show the results of two conditions for each pattern in which the gap between a pair of objects is 0 mm or 20 mm.

Table 2 shows the resulting success rates for the same and different SKUs. Table 3 shows a comparison of conventional methods and the proposed method. As conventional methods, we used the surface estimation method (Yano et al., 2023), which does not consider package-boundary confidence. In

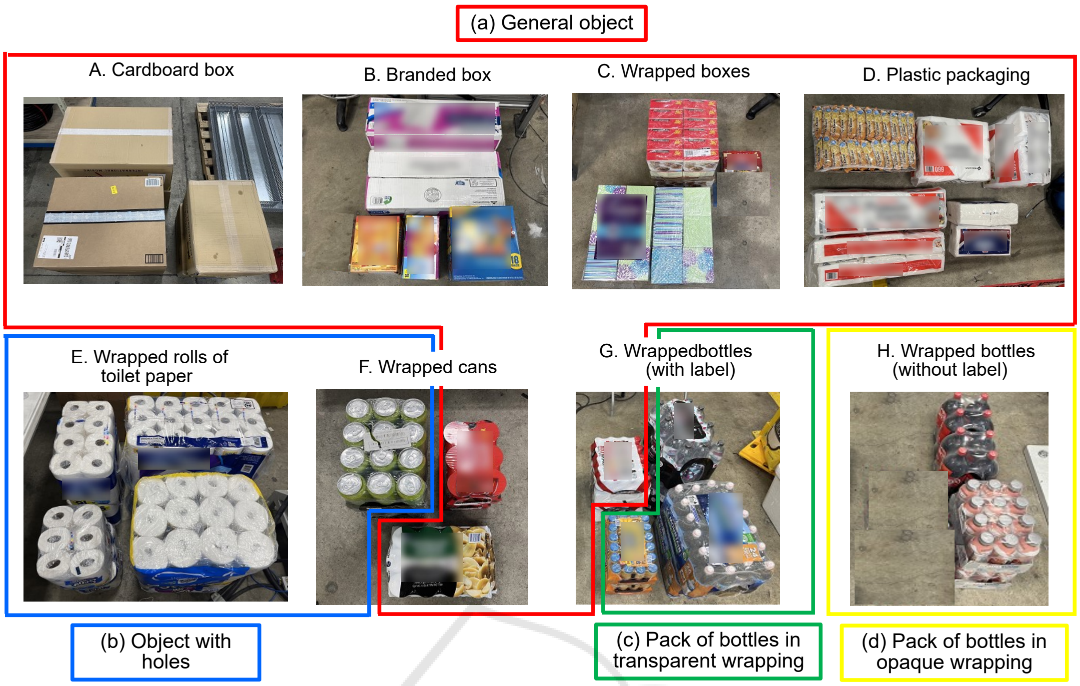


Figure 8: Rainbow SKUs used in experiments. The 32 types and eight groups of objects are classified into four categories.

(a) Same SKUs (b) Different SKUs

Figure 9: Scenes used for evaluation of the proposed method.

Table 2: Resulting success rates and frequency of slight displacement.

Condition of object pairs	Success rate	Frequency of slight displacement
Same SKUs with 0 mm gaps	99.4% (523/526)	67.9%
Same SKUs with 10 mm gaps	99.4% (523/526)	35.6%
Same SKUs with 20 mm gaps	99.0% (521/526)	28.0%
Different SKUs with 0 mm gaps	98.5% (403/409)	37.0%
Different SKUs with 10 mm gaps	98.8% (404/409)	25.8%
Different SKUs with 20 mm gaps	98.8% (404/409)	25.5%
Total	99.0% (2778/2805)	37.5%

this method, we use one of the parameters shown in Figure 6b and Figure 6c for each experiment, and if the estimation is false, the recognition would be a failure.

The proposed method had a high success rate for 32 types of objects, with a total success rate of 99.0 %, which is higher than that of the conventional methods. In addition, the frequency of slight displacement was 37.5 % in total.

Table 3 also shows estimation of the increase in depalletizing time for each condition. In case of task failure, human intervention is required for recovery and it takes five times the duration of a successful operation. Additionally, in case of slight displacement, there is no human intervention, but since two picking actions are performed for each object, it is estimated to take twice the time. The estimated overall operation time for all objects in the case of all successes can be expressed as Equation (1). In the conventional method, frequency of slight displacement is 0 %, while in the proposed method, frequency of slight displacement is calculated at 37.5 %. Increase of operation time is estimated to be reduced by 35 % compared to the conventional method.

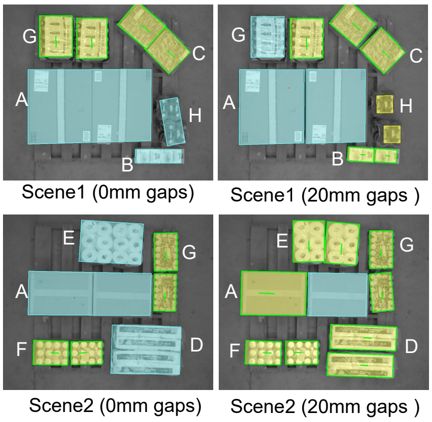
$$T = S \times 1 + (1 - S) \times 4 + D \times 1 \qquad (1)$$

where:

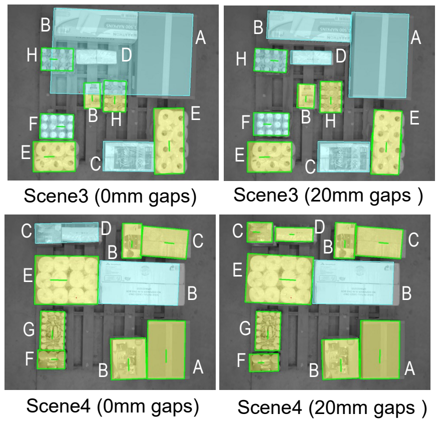
T : Increase of operation time [%]

S : Success rate [%]

D : Frequency of slight displacement [%]



(a) Same SKUs



(b) Different SKUs

Figure 10: Examples of confidence calculation results. Yellow and cyan objects have high and low confidences, respectively. A–H correspond to the eight groups in Figure 8.

Table 3: Comparison between conventional methods and the proposed method.

	Success rate	Increase of time (Estimated)
1st parameter	74.7 %	176 %
2nd parameter	74.8 %	176 %
Proposed	99.0 %	141 %

4.4 Discussion

The proposed method achieved a high success rate when two objects were placed adjacent to each other as shown in Figure 9. However, in a real depalletizing environment, various objects are stacked on top of each other as shown in Figure 11. Figure 12 shows the proposed method applied to the scene shown in Figure 11b. The top view shows that some objects were occluded and the confidence could not be calculated. Therefore, depth information must be considered when determining the order in which the objects are picked. By picking objects in order from highest to lowest, occluded objects are picked later. Also, as higher objects are picked, the objects at the bottom are no longer occluded, improving the accuracy of confidence calculation. Hence, by considering robot motion planning, the proposed method can be applicable in real depalletizing environments.

The proposed method also revealed the limitations of slight displacement. Slight displacement contributed to high recognition accuracy, which was 99.0 %. However, the high frequency of slight displacement still caused a decrease in throughput. High throughput is crucial for operating depalletizing robots in warehouses. Reducing the frequency of slight displacement will need to be addressed in the future.

(a) View from side

(b) View from top

Figure 11: Arrangement of stacked objects.

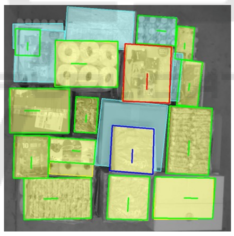


Figure 12: Resulting scene of stacked objects.

5 CONCLUSION

We proposed a package-boundary confidence estimation method that enables reliable recognition for various objects in rainbow-SKU depalletizing. The proposed method focuses on the differences in the package boundary of each type of object. Then we classified the results of surface estimation into four categories and calculated the package-boundary confidence using a different technique for each category.

In the experiment, the proposed method demonstrated a high success rate for 32 types of objects, with a total success rate of 99.0 %, which is higher than that of the conventional method.

We also determined that the proposed method is applicable when various objects are stacked. The introduction of slight displacement to the depalletizer system is expected to reduce the frequency of manual recovery performed by workers.

Our future work includes integrating boundary estimation with deep-learning methods to avoid results with low confidence regardless of correct object areas. Although our method reduced incorrect picking, the increase in the frequency of slight displacement caused the throughput of robot automation to decrease. We also aim to develop more short-time recovery methods, focusing on causes of failed recognition.

ACKNOWLEDGEMENTS

We are grateful to Mr. Takaharu Matsui for his fruitful discussions. We also thank Mr. Koichi Kato for his assistance in implementing the software for the proposed method.

REFERENCES

Aleotti, J., Baldassarri, A., Bonfè, M., Carricato, M., Chiaravalli, D., Di Leva, R., Fantuzzi, C., Farsoni, S., Innero, G., Rizzini, D. L., Melchiorri, C., Monica, R., Palli, G., Rizzi, J., Sabattini, L., Sampietro, G., and Zaccaria, F. (2021). Toward future automatic warehouses: An autonomous depalletizing system based on mobile manipulation and 3d perception. *Applied Sciences (Switzerland)*, 11(13).

Buongiorno, D., Caramia, D., Di Ruscio, L., Longo, N., Panicucci, S., Di Stefano, G., Bevilacqua, V., and Brunetti, A. (2022). Object detection for industrial applications: Training strategies for ai-based depalletizer. *Applied Sciences*, 12(22):11581.

Caccavale, R., Arpenti, P., Paduano, G., Fontanellli, A., Lippiello, V., Villani, L., and Siciliano, B. (2020). A flexible robotic depalletizing system for supermarket logistics. *IEEE Robotics and Automation Letters*, 5(3):4471–4476.

Doliotis, P., McMurrough, C. D., Criswell, A., Middleton, M. B., and Rajan, S. T. (2016). A 3D perception-based robotic manipulation system for automated truck unloading. *IEEE International Conference on Automation Science and Engineering*, 2016-Novem:262–267.

Duda, R. O. and Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15.

Eto, H., Nakamoto, H., Sonoura, T., Tanaka, J., and Ogawa, A. (2019). Development of automated high-speed depalletizing system for complex stacking on roll box pallets. *Journal of Advanced Mechanical Design, Systems and Manufacturing*, 13(3):1–12.

Fontanelli, G. A., Paduano, G., Caccavale, R., Arpenti, P., Lippiello, V., Villani, L., and Siciliano, B. (2020). A reconfigurable gripper for robotic autonomous depalletizing in supermarket logistics. *IEEE Robotics and Automation Letters*, 5(3):4612–4617.

Girshick, R. (2015). Fast R-CNN. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2961–2969.

Katsoulas, D. K. and Kosmopoulos, D. I. (2001). An efficient depalletizing system based on 2D range imagery. *Proceedings - IEEE International Conference on Robotics and Automation*, 1:305–312.

Kimura, N., Ito, K., Fuji, T., Fujimoto, K., Esaki, K., Beniyama, F., and Moriya, T. (2016). Mobile dual-arm robot for automated order picking system in warehouse containing various kinds of products. *IEEE/SICE International Symposium on System Integration*, pages 332–338.

Li, J., Kang, J., Chen, Z., Cui, F., and Fan, Z. (2020). A workpiece localization method for robotic depalletizing based on region growing and ppht. *IEEE Access*, 8:166365–166376.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). SSD: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, Proceedings, Part I 14*, pages 21–37. Springer.

Nakamoto, H., Eto, H., Sonoura, T., Tanaka, J., and Ogawa, A. (2016). High-speed and compact depalletizing robot capable of handling packages stacked complicatedly. *IEEE International Conference on Intelligent Robots and Systems*, pages 344–349.

Naumann, A., Dorr, L., Ole Salscheider, N., and Furmans, K. (2020). Refined Plane Segmentation for Cuboid-Shaped Objects by Leveraging Edge Detection. *Proceedings - 19th IEEE International Conference on Machine Learning and Applications*, pages 432–437.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 779–788.

Stein, S. C., Schoeler, M., Papon, J., and Worgotter, F. (2014). Object partitioning using local convexity. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (June):304–311.

Tanaka, J., Ogawa, A., Nakamoto, H., Sonoura, T., and Eto, H. (2020). Suction pad unit using a bellows pneumatic actuator as a support mechanism for an end effector of depalletizing robots. *ROBOMECH Journal*, 7(1):1–30.

Yano, T., Kimura, N., and Ito, K. (2023). Surface-graph-based 6dof object-pose estimation for shrink-wrapped items applicable to mixed depalletizing robots. In *VISIGRAPP (5: VISAPP)*, pages 503–511.

Yuen, H., Princen, J., Illingworth, J., and Kittler, J. (1990). Comparative study of Hough transform methods for circle finding. *Image and vision computing*, 8(1):71–77.