An Innovative Automated Robotic System based on Deep Learning Approach for Recycling Objects

Jaeseok Kim, Olivia Nocentini, Marco Scafuro, Raffaele Limosani, Alessandro Manzi, Paolo Dario and Filippo Cavallo
The Biorobotics Institute, Sant'Anna School of Advanced Studies, Viale Rinaldo Piaggio, Pontedera, Pisa, Italy

Keywords: Image Processing, Manipulation, Grasping, Deep Learning, Classification of Materials, Recycling System.

Abstract: In this paper, an industrial robotic recycling system that is able to grasp objects and sort them according to their materials is presented. The system architecture is composed of a robot manipulator with a multifunctional grasping tool, one platform, a depth and an RGB camera. The innovation of this work consists of integrating image processing, grasping, motion planning and object material classification to create a new automated recycling system framework. An efficient object recognition approach is presented that uses segmentation and finds grasping points to properly manipulate objects. A deep learning approach was also used with a modified LeNet model for waste objects classification, sorting them into two main classes: carton and plastic. Image processing and classification were integrated with motion planning that is used to move the robot with optimized trajectories. To evaluate the system, the success rate and the execution time for grasping and object classification were computed. In addition, the accuracy of the network model was evaluated. A total success rate of 86.09% and 90% was obtained for carton and plastic samples grasped using suction, while 86.67% and 78.57% using gripper. In addition, a classification accuracy of 96% was reached on test samples

1 INTRODUCTION

Nowadays, robotics has widely been developed in various fields (Bostelman et al., 2016). In particular, the industrial robotics is rapidly growing up to increase productivity and a new industrial paradigm as industry 4.0 is emerging (Lu, 2017). The goal of industry 4.0 is to reach a higher level of automation efficiently. Many industrial robotic applications (Lu, 2017) have been developed such as smart city, smart transportation, smart factory and etc. Especially, in order to build a smart city, several key technologies should be enhanced to improve the quality of human life. One of them is to develop an automated waste recycling system that could improve the quality of human life and protect the environment (Gundupalli et al., 2017).

Automated waste recycling system requires high functionalities such as object detection, object classification, motion planning and etc. In addition, complex manipulation scenario is necessary to develop the system. In order to develop a recycling system, all these functionalities should be integrated and each of them should be communicated with the others. However, the development of them is still an issue and in industrial environment considering also the interaction between workers and the robotic system.

Figure 1: Robotic recycling system using a multifunctional grasping tool.

In this paper, a system for recycling waste objects automatically was developed. For developing the system, a four major problem domains was considered: 1) object perception, 2) object classification, 3) motion planning with a multifunctional grasping tool and 4) integration of all components to obtain a reliability and acceptability industrial robotic system. The main contributions of this paper are the following:

- An efficient object recognition system is pre-

Kim, J., Nocentini, O., Scafuro, M., Limosani, R., Manzi, A., Dario, P. and Cavallo, F.
An Innovative Automated Robotic System based on Deep Learning Approach for Recycling Objects.
DOI: 10.5220/0007839906130622
In *Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2019)*, pages 613-622
ISBN: 978-989-758-380-3

sented that applies clustering and segmentation of waste objects, using a depth camera. It also supports grasping point detection and grasp affordance estimation for distinguishing objects using a multifunctional grasping tool.

- A deep learning approach is introduced with a modified LeNet model to classify materials of the wastes. This model can classify the objects into two main categories: plastic and carton. A dataset for training the model was collected and augmented applying rotation and illumination to the original dataset.
- A motion planning was applied to generate trajectories optimized for robot's arm movements and configure a proper pose of the multi-functional grasping tool. Moreover, motion planning was used to pick and place objects for recycling wastes in a specific area.
- Object perception, manipulation and object classification were integrated as main functionalities of a new automated recycling system framework.

In summary, the originality of the proposed system is to create an automated recycling system that not only picks the objects but also classifies them with multifunctional grasping tool using a deep learning approach.

2 RELATED WORK

In this section, works related to recycling system were reviewed, which covers three areas of interest: image processing, grasping, and classification of materials.

2.1 Image Processing using Depth Camera

Depth sensors are widely used to perceive a variety of environments and they are used to measure the distance data from a sensor to an object visualizing them using *point cloud data* (Masuta et al., 2016).

2.1.1 Point Cloud Data and Acquisition

Point cloud data is a collection of data points defined by a given 2D or 3D coordinates system and colour information. Point cloud becomes a common technique for image processing because it is easy to visualize and more accurate than traditional image processing techniques (Nurunnabi et al., 2012). Moreover, this method is often the only possible primitive for exploring shapes in higher dimensions (Donoho and Grimes, 2003), (Tenenbaum et al., 2000). Another benefit of this technique is to reduce computational time; in (Lei et al., 2017), Lei *et al.* used point cloud data acquired from a 3D camera justifying their choice as the fastest grasping approach.

Point cloud uses *segmentation* to process data; this technique is defined as the process of classifying point clouds into multiple homogeneous regions and this is helpful for analyzing the scene in various aspects such as locating and recognizing objects, classification, and feature extraction (Nguyen and Le, 2013), (Thilagamani and Moorthi, 2011). In (Vo et al., 2015), Vo *et al.* proposed an octree-based region growing algorithm for fast and accurate segmentation of terrestrial and aerial LiDAR point clouds. In (Ni et al., 2017), Ni *et al.* used segmentation method to process the acquired images. Based on the state of the art, we propose an approach that uses segmentation to decompose 3D data into meaningful regions functionally.

2.2 Grasping Strategy

The goal of this system is to robustly grasp objects without relying on their object identities or poses. As concern the grasping part, Principal Component Analysis (PCA) was used to choose the grasping tools (suction or gripper) to pick the objects, according to their dimensions. PCA is a standard tool in modern data analysis and it represents a simple method for extracting relevant information from confusing data sets (Xiao et al., 2013). PCA was used, in (Cruz et al., 2012) to accelerate the grasping process of unknown objects: a single-view partial point cloud was constructed and grasp candidates were allocated along the principal axis. In (Adnan and Mahzan, 2015), Adnan *et al.* used PCA in the grasping process to reduce the dimensional dataset of hand motion as well as measuring the capacity of the fingers movement. Another use of PCA is shown in (Dai et al., 2013), where the authors introduced a new PCA grasping motion analysis approach that captured correlations among hand joints and represented dynamic features of grasping motion with a low number of variables. The use of PCA in different grasping situations brought us to adopt this technique in the grasping part of our work.

2.3 Object Classification for Recycling

The environmental health in the world is bad influenced by an improper waste recycling management (Chu et al., 2018). To solve the problem, automated sorting and recycling waste materials system have been broadly investigated (Gundupalli et al., 2017).

In particular, the classification of industrial wastes is one of the core functions to develop the system. The recent search has proposed recycling system that has been widely developed based on computer vision and the use of deep learning algorithms for classifying waste materials. In (Simonyan and Zisserman, 2014) is highlighted and experimentally demonstrated how very deep convolutional networks can reach a high classification accuracy for large-scale image classification and can generalize well to a wide range of tasks and datasets. Awe *et al.* (Awe et al., 2017) used a faster R-CNN model, an object detection network with Region Proposal Networks (RPNs), in order to classify waste into three categories: paper, recycling and landfill. Rad *et al.* (Rad et al., 2017) have developed a computer vision based system for classification and localization of waste on the streets using GoogLenet. In particular, they had a significant improvement in classification accuracy splitting a class in two similar classes: leaves class and piles of leave class. This kind of approach allowed to perceive leaves grouped together and guaranteed a better generalization in classifying leaves. Mittal *et al.* (Mittal et al., 2016) introduced an android app, SpotGarbage, which uses an AlexNet model to detect and localize garbage in images. Based on the paper, a CNN network was designed with a modified LeNet model and training set, which has not complex images. It can classify waste objects in two main categories: carton and plastic.

3 SYSTEM ARCHITECTURE

The goal of this work is to develop a robotic recycling system that will be able to grasp objects in the actual environment and will be able to sort them according to their materials (carton or plastic). A Microsoft Kinect was attached under the platform and this sensor is used to process the point clouds. After the image processing, objects are grasped using UR5 robot arm with the grasping tools (Robotiq gripper and one big suction). Then, the objects are brought in front of an RGB camera that classifies them according to their material. Last, the objects that have to be recycled are collected in a box placed near the manipulator (see Figure 2). To develop the whole process described above, a platform was created to support the manipulator and to delimit its movements for safety reasons.

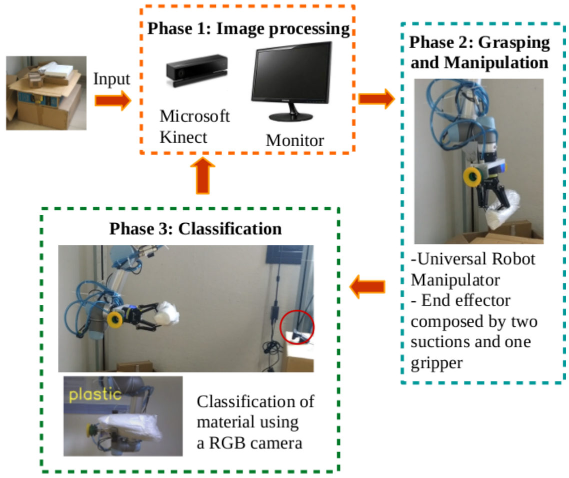


Figure 2: System structure composed of three phases: 1) Image processing, 2) Grasping and manipulation phase and 3) Classification of materials.

4 SYSTEM DESCRIPTION

This system can be divided into the following processes: a) image processing, b) grasping point detection, c) object material classification, d) motion planning.

In the proposed scenario, the Kinect camera acquires data as point cloud and processes them in order to obtain clusters representing each object inside the box. Then, the manipulator plans the optimal path to reach each object and chooses the grasping strategy between gripper and suction, according to the dimensions of the object grasped. After the grasping process, a modified LeNet model is trained to recognize the material of an object. The net classifies the objects more suitable between carton and plastic extracting their features with the RGB camera.

After material recognition, the object is moved to a delivery box by the arm (UR5 Manipulator) placed outside the structure. In details, two different delivery boxes were used: one that collects carton and the other one for plastic (see Figure 3).

4.1 Object Segmentation

In this work, the data acquired from the Kinect as point clouds were processed using the Point Cloud Library (PCL). PCL is first of all an open project for 2D/3D images and point cloud processing and in addition contains numerous state-of-the-art algorithms including filtering, feature estimation, surface reconstruction, model fitting, and segmentation (Rusu and Cousins, 2011). The data collected as point clouds are processed in order to extract the shapes of the ob-

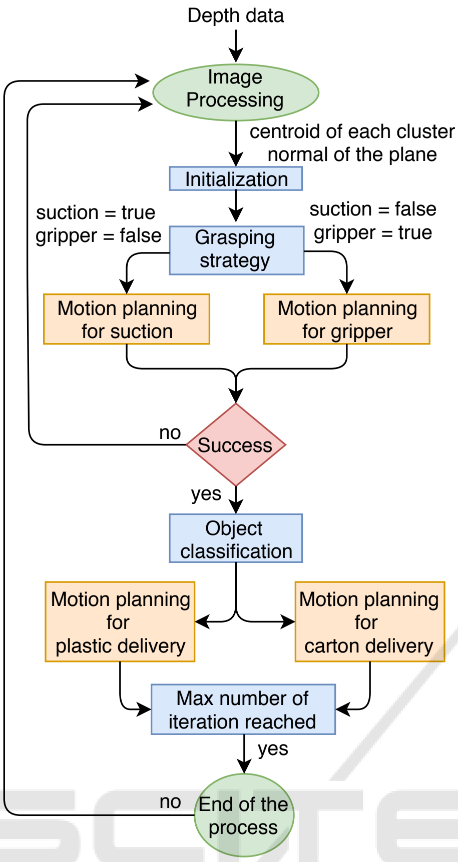


Figure 3: Flow chart of the robotic recycling system; It included image processing, grasping strategy object classification and motion planning.

jects that have to be recycled. In details, the image processing of these data is divided into the following steps: a) acquisition of data-set as point clouds, b) workspace filtering, c) clustering objects, d) plane model segmentation, e) extraction of the highest point of the object and normal of the plane. At first 3D point cloud data were acquired from the Kinect, attached to a support linked to the structure. After data acquisition, the workspace was set and three pass-through filters were implemented along the camera axis. The first one was applied along the z-axis of the camera frame in order not to detect the mobile platform as an object that had to be grasped. Then two other filters were applied, which set the x and y axes of the Kinect workspace to avoid the detection of the edges of the structure.

After the filtering process, clustering and segmentation processes were employed to take over each object of the scenario. Different methodologies have been suggested for 3D point cloud segmentation. They can be categorized into five classes: edge based methods, region-based methods, attributes

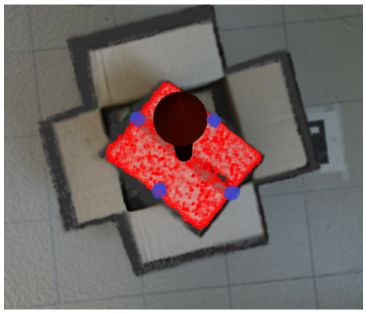


Figure 4: Image processing applied to an object. In red clustering and visualization of the normal and in blue PCA points are shown.

based methods, model-based methods, and graph-based methods (Nguyen and Le, 2013). In this work, a model-based method, that makes use of geometric primitive shapes for grouping points was employed; in details, a plane based model for the picking part was chosen. The main reason behind this choice has been that this model was shown as the most appropriate to extract a good surface for the grasping. Moreover, a plane model based is a very suitable choice because planes are one of the most important primitives since man-made structures mainly consist of planes, (Feng et al., 2014), (Xiao et al., 2013). Then, for each plane, the z coordinate of the highest point, and centroid x and y coordinates were sent as a goal state for the manipulation planning of the UR5 and the orientation of the plane was used to adapt to object shape using gripper/suction. The z-axis of the highest point was used instead of the centroid of the plane to prevent crashing the sample from the end-effector during grasping. If we had used the centroid, the robot would have pressed too much the sample and both of them would have damaged. Then, in order to work with world frame, points were converted from camera frame (see Figure 4).

4.2 Grasping Objects using Principal Component Analysis

The goal of the system is to perform robust grasping operations without predefined grasping pose estimation. To better achieve this goal,a multifunctional end effector was created that can use both suction and gripper tools. A similar end effector that had a retractable mechanism that enables quick and automatic switching between suction and gripper modalities was also used in (Chu et al., 2017) for recognizing and grasping objects. PCA was used in order to select the most suitable grasping tool, base on the size of the object. This technique finds the dimensions of the objects and compares them with the opening of the grip-

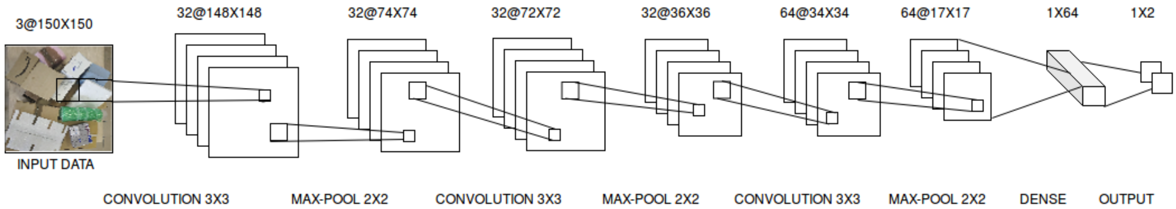


Figure 5: Convolutional neural network architecture for the classification of materials. The two outputs are for the probability of carton and plastic.

per: if the dimensions of the objects are bigger than the width of the gripper, we used suction, otherwise, we used gripper (see Figure 4).

4.3 Classification System

The goal of the proposed classification system is to separate waste objects in two different classes, after they have been picked up and positioned in front of an RGB camera. The chosen approach consists in recognizing and classifying each object separately in order to improve classification accuracy.

Based on (LeCun et al., 1998), a modified LeNet 5 model was developed that works with RGB images of 150x150 pixels as inputs (that are substantially bigger than the ones usually used for character recognition by standard LeNet models) and the output was modified as two classes. The first two convolutional layers learn 32 and last convolutional layer learns 64 filters, where each filter has size 3 x 3. Each convolutional layer is followed by the ReLU activation function and by a 2 x 2 Max-Pooling in both the x and y direction with a stride of 1. In order to avoid overfitting, regularization was applied choosing a dropout term of 0.5 after the first dense layer of 64 units. After another ReLU activation function, there is the last dense layer with 2 units, that are the number of the class labels in which the waste classification is performed. The proposed model was trained over a manually labeled dataset of normalized RGB images of the waste objects, with pixels values ranging from 0 to 1. The optimization algorithm used was stochastic gradient descent with learning rate of 0.01. Categorical-crossentropy, also called Softmax Loss, was selected as loss function (1):

$$-\frac{1}{N}\sum_{i}^{N}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right] \qquad (1)$$

The output of the trained CNN is represented with a probability over the 2 classes from an input image.

4.4 Motion Planning for Grasping Objects

For decades, motion planning has been developed for discovering optimal robot movement. In particular, Kinematics and Dynamics Library (KDL) and Open Motion Planning Library (OMPL) are broadly used to search movements of a robot arm. In ROS, Motion Planning Framework (Moveit!) (Chitta et al., 2012) was integrated with these libraries as plugins in the system architecture so that it can support self-collision avoidance with inverse kinematics to determine the feasibility of grasp. Moveit! also can generate several possible paths to reach the goal with sampling-based planning. In this work, before using Moveit!, collision areas were configured in URDF (the standard robot description format in ROS), to avoid crashing between obstacles and robot arm. Moreover, specific positions such as delivery places, initial robot arm position and etc. were defined in Semantic Robot Description Format (SRDF). During motion planning, it was allowed to apply replanning process because it supports searching more optimal path compared to the previous one generated. Furthermore, we used trajectory following method, which generates waypoints between the arm and goal that could avoid collisions and maintain defined end effector pose constantly.

(a) (b)

Figure 6: Total 30 samples (cartons and plastics) were prepared to test the recycling system.

5 EXPERIMENTAL SETUP

In this section, the experimental setup that is needed to demonstrate our system will be explained. A new

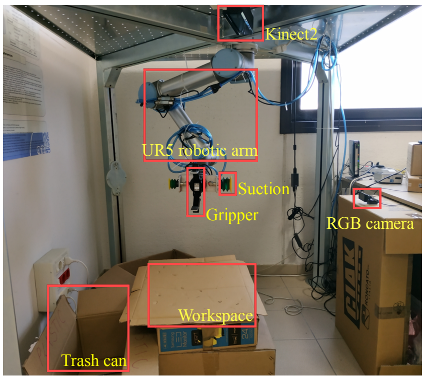


Figure 7: Experimental setup for the recycling system composed of five elements: UR5 manipulator, RGB and Kinect cameras, grasping tools, workspace, and trash can.

robotic platform is introduced, the procedure to collect the dataset is described, and the system is initialized. For the experiment, evaluation scenarios were organized.

5.1 Hardware System Description

The proposed robotic platform was developed to build a robotic recycling system during CENTAURO Regional Project - iSort (2016 - 2018). It consists of one cage with four steel bars, a robot arm (UR5), a depth camera (Microsoft Kinect v2), a Robotiq two-finger gripper, two suction cups (small and big) and a Logitech mono camera. The system was organized to allow implementation of major functionalities that recognize waste objects and classify them according to their material (see Figure 7).

The experimental setup was considered to build as one in the actual industrial environment. However, due to the safety issues with limited space, we could not build the same as one in the environment. Moreover, operation speed was reduced because of the protection of robot and human.

5.2 Collecting the Dataset

Before collecting images for the training part, information of cartons and plastics were recorded by a webcam mounted from different angles. Image frames were extracted automatically using ROS bag functionality. The dataset collected is composed of a total of 105 sample images, which are 51 carton and 54 plastic samples. However, the dataset had really few samples to exploit the real power of the CNNs. In order to overcome the limitation of training examples, data augmentation was applied to the training examples with a number of random transformations. As a result, the dataset increased the number of training samples (total 3002 samples) and never see twice the exact same images. This method helps to prevent overfitting and support the CNN model to generalize the situations that can be found in the actual environment. The augmented dataset will be publicly available[1].

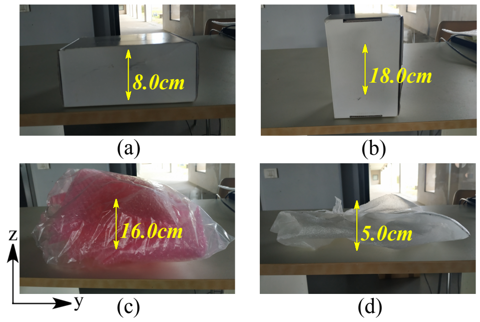


Figure 8: Object configurations for testing was selected based on material and its height (z-axis). (a) An short carton (SC), (b) An tall carton (TC), (c) An voluminous plastic (VP), (d) A non voluminous plastic (NVP).

5.3 System Initialization

The robot arm installed on the steel plate was placed out of workspace to be totally visible on depth camera system (Figure 1). Workspace was set in the middle of a platform with the four steel bar and it was placed on approximately 40 cm from the bottom; this position was considered an easy one to grasp objects. For object classification, a RGB camera was installed in front of the robot arm's initial position. Two boxes were prepared to collect waste objects from delivery of the arm. The arm initial position, the camera position for classification and the boxes for trash can positions were predefined during the recycling system operation. For the experiment, a set of total 30 objects (Figure 6) composed of 50% of carton and plastic was prepared to perform the following tests: grasp an object using a multifunctional grasping tool (suction and gripper) and classify object materials. In addition, if one carton was tested with different object configurations (Figure 8), it was counted as two samples. Moreover, if intuitively a grasping tool could not operate picking an object, the object was removed using the specific tool (e.g: if width of carton or plastic materials are bigger than gripper's width or thinner than suction's width).

The network was trained using 60% and 20% of the dataset as training set and validation set, and the remaining 20% as test set. The training process was

[1] The dataset will be released on github: https://github.com/Alchemist77/Centauro_Project

stopped after 250 epochs. The Stochastic gradient descent (SGD) was used as the optimizer to minimize the loss function. The batch size during training was of 16 elements and learning rate was 0.01. The networks were implemented using Keras with Tensorflow frameworks (Abadi et al., 2016)

5.4 Evaluation Scenarios

To evaluate the system, we measured the success rate and the execution time for grasping objects and object classification. In order to obtain the success rate for grasping objects, firstly a waste object (carton or plastic) was placed on the workspace letting the arm grasp it for 5 repetition trials without human intervention. To evaluate the performance of the success rate, the percentage of the number of trials was calculated:

$$G_s = \frac{G(r)}{5} 100(\%), \qquad r = 1, \ldots, N_r, \tag{2}$$

where G_s represents success rate, $G(r)$ is defined as the number of success at repetition r, and $N_r = 5$ is the number of repetitions during grasp process. The success rate of object classification has the same calculation procedure.

To measure the execution time for grasping objects and object classification processes, different initial states were proposed. For grasping objects process, object segmentation and clustering with motion planning were considered. In contrast, the object classification process started from the grasping of the object and the measurement ended when CNNs visualized the output of object material.

In addition, the different configuration of the shape of an object was considered during the evaluation of the two processes above.

6 EXPERIMENTAL RESULTS

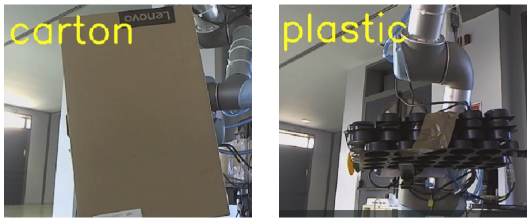
In this section, system results are shown analyzing first the success rate obtained for grasping and classification and then the execution time for each task. Furthermore, network performances in terms of loss and accuracy are discussed.

6.1 Success Rate

Tables 1 and 2 show the results obtained from grasping using suction (Figure 9a) and gripper (Figure 9b) respectively. Both tables are divided into 4 subgroups according to objects material and objects physical properties.

(a) (b)

Figure 9: Experiments for grasping objects using the multifunctional tool. (a) Grasping an object using suction, (b) Grasping an object using gripper.

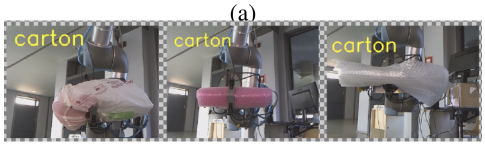


(a) (b)

Figure 10: Experiments for classification of objects with text sign. (a) carton classification (b) plastic classification.

First, a trial with grasping cartons was done using suction with 12 short carton (SC) and 9 tall carton (TC) samples (Figure 8 (a and b)) and each object was tested for five times. The same approach was used for plastic samples (4 voluminous plastic (VP) and 6 non-voluminous plastic (NVP) samples (Figure 8 (c and d))). The results show that suction has better performance with plastic (90% success rate) than with carton (86,09%). The main problem of grasping carton was to relate on the dimensions of the object: if the object was too thin, then the suction couldn't reach the grasping point because the segmentation could not find the object. When the object was too high there was the problem that the suction crashed because there was too much pressure on the tool. Other issues happened due to mechanical problems of the manipulator, the presence of holes and discontinuities

(a)



(b)

Figure 11: Examples of failed classification of objects. (a) carton classification (b) plastic classification.

Table 1: Success rate of grasping objects using suction.

Category	Object configuration	Number of objects	Number of attempts	Number of successes	Total success rate
Carton	SC	12	70	62/70 (88.57%)	99/115 (86.09%)
	TC	9	45	37/45 (82.22%)	
Plastic	VP	4	20	18/20 (90%)	45/50 (90%)
	NVP	6	30	27/30 (90%)	

Table 2: Success rate of grasping objects using gripper.

Category	Object configuration	Number of objects	Number of attempts	Number of successes	Total success rate
Carton	SC	5	25	21/25 (84%)	65/75 (86.67%)
	TC	10	50	44/50 (88%)	
Plastic	VP	5	25	19/25 (76%)	55/80 (68.75%)
	NVP	11	55	36/55 (65.45%)	

Table 3: Success rate of object classifications.

Category	Object configuration	Number of objects	Number of attempts	Number of successes	Total success rate
Carton	SC	6	30	30/30 (100%)	85/100 (85%)
	TC	14	70	55/70 (78.57%)	
Plastic	VP	7	35	30/35 (85.71%)	86/100 (86%)
	NVP	13	65	56/65 (86.15%)	

in the surface of objects and the presence of scotch tape in some parts of the carton samples. As regard to plastic samples, segmentation problems were occurred due to the flexible surface of plastic and due to some breakage of plastic during suction operation.

As regard to the grasping part using gripper, the results present the opposite situation: carton samples have a better performance (86.67%) than plastic samples (68.75%). The limitation of workspace is the main problem of grasping. In addition, rotational joint mechanical limitations blocked the robot or brought the manipulator to hit the object. Other two minor problems encountered were the wrong segmentation of an object and the height of a sample (if the object was too high, the gripper crashes).

Summarizing, from the experimental stage was understood that the best situation happened when plastic was taken using suction; on the contrary, plastic samples were discovered to be the worst samples to grasp using gripper.

Concerning the classification part, Table 3 shows the results obtained from classification experiments (Figure 10a and 10b). The percentage of success rate obtained for carton and plastic samples are quite similar: plastic samples have a slightly higher success rate (86%) than carton samples (85%). Both percentages are quite high and are similar to the success rates of the grasping experiments. The problems met during classification experiments were due to the colour of the samples. White and light blue cartons were detected as plastic due to the similarity of these colours with plastic colours (Figure 11a). A plastic bag was seen, instead, as a carton and these wrong classification happened because this object was grasped even if it was not part of the original dataset. The presence of the sunlight also affected the success of classification as it occurred for the packaging material (Figure 11b).

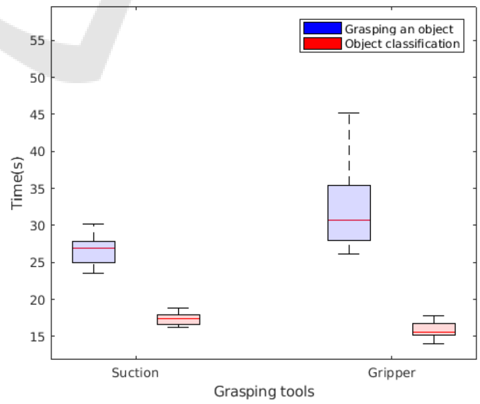


Figure 12: Visualization of execution times for grasping and classifying objects.

6.2 Execution Time

Figure 12 shows the execution time during classification and grasping tasks using gripper and suction respectively. Firstly, the time for grasping object is different, based on the grasping tool used: the mean value (26.50s) using suction is less than the one obtained with gripper (33.19s), therefore suction process is more fast. Also, standard deviation (SD) with suction (1.72s) is higher than SD as regarding of gripper (7.33s). Suction grasping performances are simpler and easier obtained compared to the ones using gripper. On the contrary, when the arm uses gripper, the variability of the execution time increased because extra processes were operated to detect the object highest point: robot adapts the end effector orientation in order to grasp the object. With regard to classification part, mean value using both tools (suction: 17.57s and 17.76s) are comparable. Moreover, SD using gripper (1.05S) and suction (1.06s) are not substantially different because the trajectories, which to the camera for classification were already predefined.

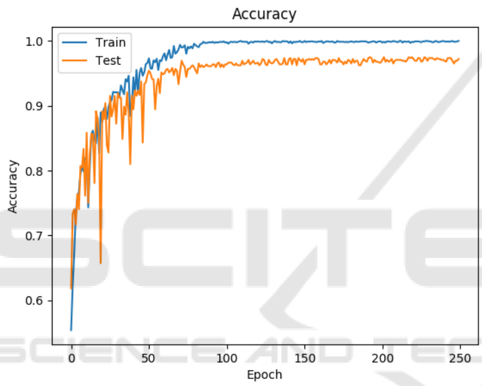


Figure 13: Accuracy value evaluation during 250 epochs for training and test (blue and orange solid lines).

6.3 Train and Test Set Results, Accuracy and Loss

In Figure 13 is showed the performance of the network in terms of accuracy, on training and test sets during 250 epochs. In blue and orange respectively, the accuracy of the network on training and test phases is represented. Accuracy reaches the value of 0.99 during training and 0.96 in the test phase. In parallel the optimizer algorithm leads loss to converge to a loss function of 0.01 in training phase and 0.08 in the test phase. Therefore, the model can perform and generalize well on new data and it is not affected by overfitting.

7 DISCUSSION AND CONCLUSION

In this work, the development of an autonomous robotic system was presented. This system is able to grasp objects and sort them according to their material compositions (carton or plastic), in order to foster recycling practices in industries. The mainly novelties of this work are two: building a preliminary framework for benchmarking industrial applications in sorting management and integrating functionalities as image processing, motion planning, grasping and classification in a unique robotic structure. Another challenging aspect of this work is the use of a multifunctional end effector equipped with both gripper and suction tools; this multifunctionality increased the success rate during the grasping process, reducing the probability of error. During the experiments, only the bigger suction was used. For future work, both types of suction will be applied for selecting the right one, according to the object dimensions. A limitation of the proposed work was that we did not considered grasping in cluttered environments. For this reason, this grasping part could be a challenging field to analyse in the future (ten Pas et al., 2017). Another issue of this work concerns the dataset: original dataset should be integrated in order to have a large-scale dataset, allowing the classification system to more generalize on new objects, furthermore a new approach based on learning more features should be found, to have a better classification of materials (Simonyan and Zisserman, 2014) . Finally, another challenging idea could be handling a greater variety of objects and new groups of material like glass and organic would generalize an automated recycling system completely. (Zeng et al., 2017).

ACKNOWLEDGEMENT

This work was supported by the Tuscany's regional research project CENTAURO: Colavoro, Efficienza, preveNzione nell'industria dei motoveicoli mediante Tecnologie di AUtomazione e RObotica. bando FAR-FAS 2014.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). Tensorflow: A system for large-scale machine learning. In *12th {USENIX} Sympo-*

sium on Operating Systems Design and Implementation ({OSDI} 16), pages 265–283.

Adnan, N. H. and Mahzan, T. (2015). Gesture recognition based on human grasping activities using pca-bmu. *Gesture*, 6(11).

Awe, O., Mengistu, R., and Sreedhar, V. (2017). Smart trash net: waste localization and classification. *arXiv preprint*.

Bostelman, R., Hong, T., and Marvel, J. (2016). Survey of research for performance measurement of mobile manipulators. *Journal of Research of the National Institute of Standards and Technology*, 121:342–366.

Chitta, S., Sucan, I., and Cousins, S. (2012). Moveit![ros topics]. *IEEE Robotics & Automation Magazine*, 19(1):18–19.

Chu, K., Zhang, Q., Han, H., Xu, C., Pang, W., Ma, Y., Sun, N., and Li, W. (2017). A systematic review and meta-analysis of nonpharmacological adjuvant interventions for patients undergoing assisted reproductive technology treatment. *International Journal of Gynecology & Obstetrics*, 139(3):268–277.

Chu, Y., Huang, C., Xie, X., Tan, B., Kamal, S., and Xiong, X. (2018). Multilayer hybrid deep-learning method for waste classification and recycling. *Computational Intelligence and Neuroscience*, 2018.

Cruz, L., Lucio, D., and Velho, L. (2012). Kinect and rgbd images: Challenges and applications. In *2012 25th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials*, pages 36–49. IEEE.

Dai, W., Sun, Y., and Qian, X. (2013). Functional analysis of grasping motion. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3507–3513. IEEE.

Donoho, D. L. and Grimes, C. (2003). Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, 100(10):5591–5596.

Feng, C., Taguchi, Y., and Kamat, V. R. (2014). Fast plane extraction in organized point clouds using agglomerative hierarchical clustering. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 6218–6225. IEEE.

Gundupalli, S. P., Hait, S., and Thakur, A. (2017). A review on automated sorting of source-separated municipal solid waste for recycling. *Waste management*, 60:56–74.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., et al. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Lei, Q., Chen, G., and Wisse, M. (2017). Fast grasping of unknown objects using principal component analysis. *AIP Advances*, 7(9):095126.

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of Industrial Information Integration*, 6:1–10.

Masuta, H., Motoyoshi, T., Sawai, K. K. K., and Oshima, T. (2016). Plane extraction using point cloud data for service robot. In *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on*, pages 1–6. IEEE.

Mittal, G., Yagnik, K. B., Garg, M., and Krishnan, N. C. (2016). Spotgarbage: smartphone app to detect garbage using deep learning. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 940–945. ACM.

Nguyen, A. and Le, B. (2013). 3d point cloud segmentation: A survey. In *RAM*, pages 225–230.

Ni, H., Lin, X., and Zhang, J. (2017). Classification of als point cloud with improved point cloud segmentation and random forests. *Remote Sensing*, 9(3):288.

Nurunnabi, A., Belton, D., and West, G. (2012). Robust segmentation in laser scanning 3d point cloud data. In *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on*, pages 1–8. IEEE.

Rad, M. S., von Kaenel, A., Droux, A., Tieche, F., Ouerhani, N., Ekenel, H. K., and Thiran, J.-P. (2017). A computer vision system to localize and classify wastes on the streets. In *International Conference on Computer Vision Systems*, pages 195–204. Springer.

Rusu, R. B. and Cousins, S. (2011). 3d is here: Point cloud library (pcl). In *Robotics and automation (ICRA), 2011 IEEE International Conference on*, pages 1–4. IEEE.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

ten Pas, A., Gualtieri, M., Saenko, K., and Platt, R. (2017). Grasp pose detection in point clouds. *The International Journal of Robotics Research*, 36(13-14):1455–1473.

Tenenbaum, J. B., De Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323.

Thilagamani, S. and Moorthi, S. (2011). A survey on image segmentation through clustering. *International Journal of Research and Reviews in Information Sciences*, 1.

Vo, A.-V., Truong-Hong, L., Laefer, D. F., and Bertolotto, M. (2015). Octree-based region growing for point cloud segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104:88–100.

Xiao, J., Zhang, J., Adler, B., Zhang, H., and Zhang, J. (2013). Three-dimensional point cloud plane segmentation in both structured and unstructured environments. *Robotics and Autonomous Systems*, 61(12):1641–1652.

Zeng, A., Song, S., Yu, K., Donlon, E., Hogan, F. R., Bauzá, M., Ma, D., Taylor, O., Liu, M., Romo, E., Fazeli, N., Alet, F., Dafle, N. C., Holladay, R., Morona, I., Nair, P. Q., Green, D., Taylor, I., Liu, W., Funkhouser, T. A., and Rodriguez, A. (2017). Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching. *CoRR*, abs/1710.01330.