2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Model-Free, Vision-Based Object Identification and Contact Force Estimation with a Hyper-Adaptive Robotic Gripper

Waris Hasan, Lucas Gerez, and Minas Liarokapis

***Abstract*— Robots and intelligent industrial systems that focus on sorting or inspection of products require end-effectors that can grasp and manipulate the objects surrounding them. The capability of such systems largely depends on their ability to efficiently identify the objects and estimate the forces exerted on them. This paper presents an underactuated, compliant, and lightweight hyper-adaptive robot gripper that can efficiently discriminate between different everyday life objects and estimate the contact forces exerted on them during a single grasp, using vision-based techniques. The hyper-adaptive mechanism consists of an array of movable steel rods that get reconfigured conforming to the geometry of the grasped object. The proposed object identification and force estimation techniques are model-free and do not rely on time consuming object exploration. A series of experiments have been carried out to discriminate between 12 different everyday life objects and estimate the forces exerted on a dynamometer. During each grasp, a series of images are captured that detect the reconfiguration of the hyper-adaptive grasping mechanism. These images are then used by an image processing algorithm to extract the required information about the gripper reconfiguration, classify the object grasped using a Random Forests (RF) classifier, and estimate the amount of force being exerted. The employed RF classifier gives a prediction accuracy of 100%, while the results of the force estimation techniques (Neural Networks, Random Forests, and 3rd order polynomial) range from 94.7% to 99.1%.**

I. Introduction

Grippers and hands allow robots to efficiently interact with their surroundings by grasping and manipulating different objects or parts of the environment. However, identifying an object whose the geometry is unknown, remains an important challenge that has not yet been addressed satisfactorily [1], [2]. The main issue with conventional and adaptive grippers is that they first need to detect the shape of the object and then based on the shape or the orientation of the object to execute the grasp using an appropriate grasp planning scheme. To do that, the gripper may have to approach objects with different geometries from different angles and to employ sophisticated sensing and control algorithms [1].

Over the last decade, many different techniques and sensors have been used to discriminate between different everyday life objects and measure the contact forces exerted. Typically, object identification is accomplished using image analysis techniques, tactile sensing solutions, or a combination of both [3]–[6]. However, the performance of a vision system may be hindered by occlusions (the gripper usually covers a significant part of the surface of the grasped object), poor lighting conditions (dark environments), or limited fields of view [7]. Furthermore, the most common approach to extract the 3D model of an object is by employing specialized vision systems that require complex computations and multiple bulky and expensive cameras, while the results are not always as accurate as desired [8].

Waris Hasan, Lucas Gerez, and Minas Liarokapis are with the New Dexterity research group, The University of Auckland, New Zealand. E-mails: whas626@aucklanduni.ac.nz, lger871@aucklanduni.ac.nz, minas.liarokapis@auckland.ac.nz

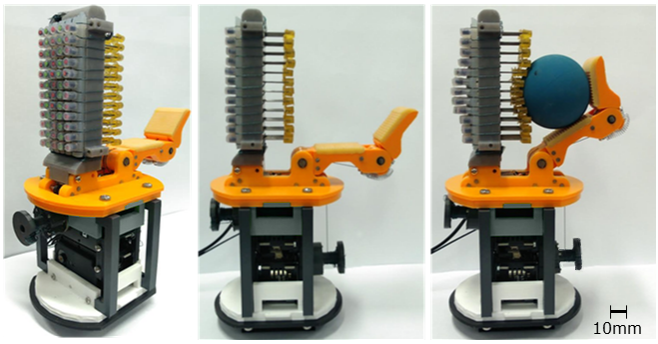

Fig. 1. The employed hyper-adaptive robotic gripper grasping an object and conforming to its geometry. The gripper conforms to the shape of the object due to its highly adaptive structure. A vision-based algorithm tracks the motion of each pin to estimate its displacement. The displacements of the pins are used in order to perform object identification for a set of everyday objects and to estimate the contact forces exerted during grasping.

In this paper, we employ a hyper-adaptive robotic gripper that can efficiently conform to the shape of the grasped object due to its highly adaptive nature (Fig. 1) and image analysis techniques to perform a simultaneous learning based object identification and contact force estimation during a single grasp. The image analysis algorithm tracks the motion of each pin of the hyper-adaptive design to discriminate between different everyday objects and estimate the contact forces applied to them. The efficiency of the proposed methods has been experimentally validated using a series of experiments that involve the identification of objects with a single grasp and the estimation of the forces exerted on a dynamometer (while it is grasped). This study tackles the object identification problem by employing the change in the shape of the hyper-adaptive grasping mechanism to discriminate between different object shapes instead of trying to identify them with a vision system (e.g., with point-cloud data). To do that, it employs a simple 2D camera that is used to extract information about the reconfiguration of the finger structure, effectively reconstructing the object shape. The use of the hyper-adaptive gripper together with the proposed vision-based scheme makes the system model-free and it allows the classification of the grasped object in an open-loop fashion during a single grasp.

978-1-7281-6211-9/20/$31.00 ©2020 IEEE

The rest of the paper is organized as follows: Section II presents the related work on object classification and identification, Section III presents the gripper and discusses the computer vision system proposed, Section IV details the experiments, Section V presents the results, while Section VI concludes the paper and discusses some future directions.

II. Related Work

Over the past few years, numerous studies have explored ways of identifying everyday life objects / tools, deriving their properties [3], [4]. The two most commonly used approaches employ vision-based sensing and tactile sensing. However, the choice is highly dependent on the environment, the type of application, and the type of robotic gripper being used [3]. In general, computer vision systems that have been employed in robotic grasping and dexterous manipulation depend on the camera type, the camera placement, and the number of cameras. In [9], the authors discuss the importance of these factors and their effect on the design of vision algorithms. They use eye-in-hand and stand-alone stereo camera systems to develop strategies to detect and grasp both known and unknown objects. In [10], the authors present a vision-based algorithm that uses two 3D cameras to generate a contour of the target object by combining partial point clouds from each camera. This algorithm detects the geometry of unknown objects and runs a virtual exploration on the partial point cloud to calculate the force balance computations on different planes, allowing it to determine the most stable grasp. In [11], two cameras and a linear laser generator were used to form a 3D model of the object under consideration, using point cloud data. Grasp planning was accomplished virtually by approaching the object using different grasp configurations. The drawback of this solution is the excessive amount of time that is required to compute the grasp configurations. In [12], the authors used a stereo vision system to determine the 3D geometry of objects and then used the localized model of an object to plan the grasp. Their vision system is limited to detecting rectangular prisms only. A 2D laser range finder was used in [13] to classify seven different objects, and the results show that this system is much more reliable and accurate than using a traditional laser beam scanner and a CCD (charge-coupled device) camera to classify objects. In [14], the authors used two cameras to find the distance to the object by triangulation instead of constructing a 3D model. The identification of the object and the identification of grasping points on the object surface had a success rate of 90% for known objects.

Other than detecting the geometry and the edges of rigid objects, computer vision has also been widely used to measure the applied forces and deformations of elastic objects being grasped. In [15], the authors used computer vision algorithms to monitor the characteristics of various soft, deformable objects, and then they modelled and predicted their behaviour when manipulated by a robotic gripper. In [16], the authors used data from RGB images and a Microsoft Kinect depth sensor to classify various non-rigid objects based on their material properties with classification success rate up to 98.3%. In [17], the authors propose a 3D vision algorithm that can monitor deformations in non-rigid objects. The experimental results show that such a system is robust in real-time tracking of the surface deformations.

The general problem that has been encountered in most of these studies is the selection of the most stable grasp from a wide range of possible grasps. Some of the aforementioned studies used 2D or 3D models of the objects for grasp determination, while others relied on the assumption that the objects under consideration belong to a particular set of shape primitives (cylinders, boxes, etc.). Using such an approach, they managed to minimize the number of possible grasps. However, these approaches only work well when the objects are both not occluded and well separated from the background [18]. This problem is highlighted in [18], where the authors propose a purely vision-based solution that is model-free. The proposed 3D algorithm can automatically select the most stable grasp for the objects from the potential grasps which are generated using surfels. In [19], a compliant, under-actuated robotic gripper with eight force sensors embedded in each finger is used. The authors propose a methodology that discriminates between everyday objects using a single grasp of the employed gripper (exploiting its reconfiguration) without any prior information about the gripper model or the object. The results show that such a methodology is able to identify objects with high accuracy.

The approach of this study is inspired by a previous work, detailed in [20], which presents a two-fingered, under-actuated gripper that has hyper-adaptive finger pads that can adapt to the shape of the object being grasped. This design allows model-free classification, which means that the controller parameters do not affect the classification process, unlike most of the studies discussed above. Apart from tactile sensors on the finger pads, they also used a Google Soli radar sensor to identify a set of different materials. More precisely, they used 20 different features to discriminate between 26 different objects with a classification accuracy of over 99%. The drawback of using tactile sensors in this particular design is that these sensors cover a considerable amount of area on the hyper-adaptive region, reducing the efficiency of the proposed design. Moreover, due to the size and wiring limitations, tactile sensors cannot be used to measure the contact force on each pin of the mechanism.

Thus, this study aims to estimate the force exerted by the pins and the displacement of each pin in the hyper-adaptive gripper surface by using computer vision algorithms, allowing also for identification of the grasped object. Hence, the proposed computer vision technique aims to address most of the aforementioned issues by processing 2D images from a single camera, which unlike 3D modelling does not require heavy computations [8]. Additionally, since this technique does not identify the objects by directly taking their pictures, but relies on the change in the shape of the hyper-adaptive gripper, it is free from any issues caused by occlusions during grasping. The camera can be enclosed within the hyper-adaptive region of the gripper since it only needs to capture images from the finger backside.

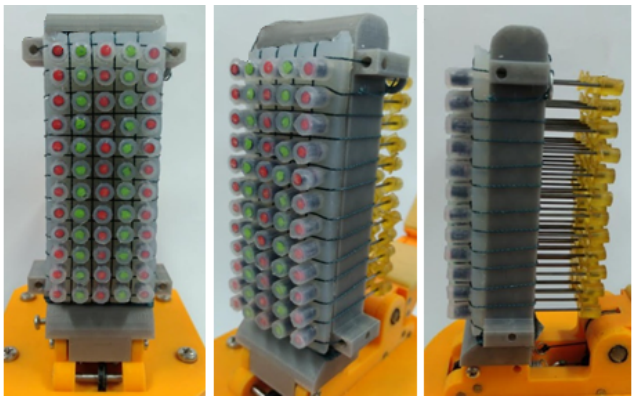

Fig. 2. The hyper-adaptive finger-pad consists of a 5x12 array of thin steel rods (60 in total) with a compliant rubbery tip made out of Smooth-On PMC-780 urethane rubber. The rubbery tip increases the friction between the finger and the object. These pins are displaced when the object is pressed against them. The steel rods are pushed passively back to their original positions by the silicone springs (made out of Smooth-On Ecoflex 00-30). The tips of the silicone springs have plastic caps that accommodate colored dots that are used for vision-based reconfiguration estimation.

III. Design and Methods

A. Hyper-adaptive Gripper

The term 'hyper-adaptive' refers to the ability of the gripper to efficiently conform to the shape of different objects because of its highly adaptive nature. In particular, this design improves the grasp stability when compared to the conventional adaptive grippers by increasing the number of contact points, as shown in Fig. 2. The hyper-adaptive gripper does not require the use of accurate grasp planning methods. The gripper was developed using 3D printed parts and simple materials. The gripper has a stationary hyper-adaptive thumb and an adaptive, tendon-driven, articulated finger. The design is based on the M2 gripper of the Yale OpenHand project [21], [22]. The hyper-adaptive pads on the thumb consist of a 5x12 array of thin steel rods (60 in total) with a compliant rubbery tip made out of polyurethane rubber (Smooth-On PMC-780) that increases the friction between the finger and the object. A higher number of steel rods means a higher resolution and a more stable grasp [23]. The proposed hyper-adaptive gripper has more than two times the number of pins of the previous version of the gripper described in [20]. These pins are displaced when the object is pressed against them by the articulated finger and pushed back to their original positions when the object is released. On the rear end of the thumb, a continuous block of silicone (made out of Smooth-On Ecoflex 00-30) was placed. It has a 5x12 array of deformable silicone molds (one for each steel rod). These silicone molds act as springs so that the displaced steel rods can return to their original positions when the articulated finger stops exerting forces on the object. Small 3D printed caps were placed in between the silicone springs and the steel rods to keep the steel rods from puncturing the silicone springs. The caps were printed to have bright colors so that they are easily visible through the silicone springs, and a camera can track their movement.

The design is completely modular and any part of the gripper can be easily disassembled for repair or replacement. The gripper operates using two Dynamixel MX-64AR motors that are connected to the base. A separate motor was used for each joint. The gripper was designed to have a fixed hyper-adaptive thumb opposing an articulated, adaptive finger, facilitating the camera and image processing set up. If an articulated, hyper-adaptive finger is selected, more than one camera will be required to extract pin displacement data from the different phalanges. The articulated finger was designed to have two spring-loaded pin joints (one at the base and one between the two finger phalanges), and the finger pads were made out of compliant material (Smooth-On Vytaflex 30) to have a firm grip on the object.

B. Vision-Based Reconfiguration Estimation

The methodology proposed in this paper uses a single camera with a fixed position and a field of view containing the back of the hyper-adaptive robot finger to detect and measure the slightest displacements / reconfigurations of the steel rods from their rest positions. The gripper is connected to the end-effector of a UR5 robot arm (Universal Robots, Odense, Denmark). A diffused light source is used to provide an even, soft illumination from all angles without any reflections. The proposed methodology consists of the following steps:

- A reference image is captured by a 12-megapixel camera when the hyper-adaptive robotic gripper is at rest (not grasping any object).
- The lens of the camera is fixed at an angle to the backside of the hyper-adaptive finger, detecting the coloured caps, as shown in Fig. 3 (this allows the 3D rectilinear movement of the steel rods to be captured in a 2D image).
- Pixels representing the centre of each coloured cap are located in the image taken in step 2.
- Images are captured during the whole process of object grasping, keeping the pose of the camera fixed with respect to the gripper. The centers of the caps are recorded during the reconfiguration of the fingers.
- The distances between the initial and the final positions of the centers of the colored caps are calculated. These distances represent the relative lengths by which the steel rods have been displaced by the object (white lines in Fig. 3).

The image processing algorithms have been developed in C++ using the Open-CV open-source computer vision library. The details of the image processing are as follows:

1) **Colored Caps Detection:** It is crucial to precisely detect all 60 caps from the captured image because all other steps of the algorithm depend on the accurate detection of the caps. The bright colors of the caps (red and green) are significantly distinctive from any other color in the image, therefore, the areas of the pixels representing these caps can be easily detected using simple color based detection. The input images

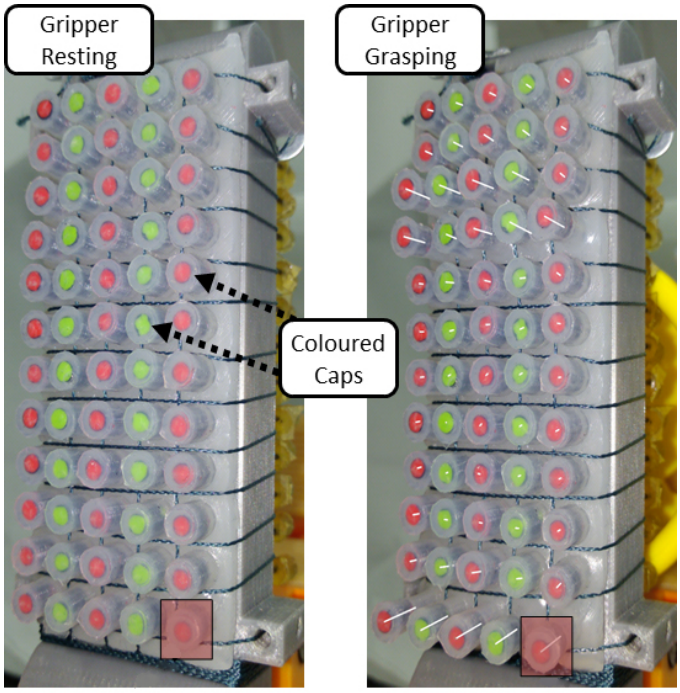

Fig. 3. The displacement of each pin is determined by calculating the centre pixel in the resting and final grasping position for each cap and measuring their distance. This distance represents the length by which the steel rod was displaced by the object (white line).

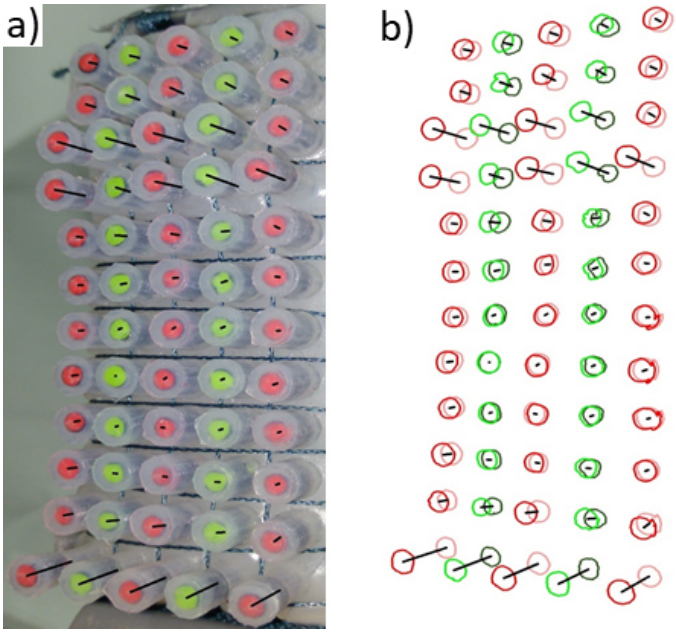

Fig. 4. Distance calculation between the centre pixels of the corresponding caps in the resting and final grasping images of the gripper. The black lines represent the displacements of the pins while grasping an object superimposed to the original image of the gripper. In subfigure b), the caps and their centre pixels are represented with light green and light red circles in the resting states and with dark green and dark red circles when they are in the final grasping states. The resting and grasping positions of the caps have been superimposed to demonstrate the process of calculating the distances between the corresponding centre pixels.

are converted into binary images through HSV color thresholding. These binary images are then sequentially passed through Erosion, Dilation, and Gaussian Blur filters to remove any remaining parasitic noise and to prepare them for Canny Edge Detection. Canny Edge Detection allows us to detect the exact location of each cap in the binary images.

2) **Detection of the Centers of Caps:** It can be seen in Fig. 4 that the edges of the caps that are detected by the Canny Edge Detector are not exact circles, but they have irregular shapes. Because of this, it is very difficult to find the pixels that are exactly at the centre of these irregular shapes. The solution is to draw a minimum enclosing circle around the detected edges of each cap and the centres of these minimum enclosing circles to be the centres of the caps. The algorithm predicts the displacement of the caps by calculating the displacement of these centre pixels.
3) **Sorting the Centers of Caps:** The centre pixels of the resting image (initial image) and the final grasping image of the gripper are stored in separate arrays and then the corresponding centre pixels are compared to calculate distances (centre pixels representing the same cap in both images). For example, the centre pixel of the top right cap in the resting image needs to be compared with the centre pixel of the top right cap in the final grasping image. To be able to do that, the centre pixels from resting and final grasping images of the gripper need to be stored in the same sequence in the arrays. This is achieved by applying a sorting mechanism on the arrays that store the centre pixels. Sorting is the main reason to choose two coloured caps (red and green) because different coloured caps can be detected and sorted independently making the sorting process modular, easier, and efficient. The sorting is done in three steps: i) Sorting the centre pixels from left to right (based on their x coordinates), allowing the separation of the centre pixels into five different columns, ii) adding the centre pixels of each column in a separate array to have better control over the sorting process, and iii) sorting the centre pixels in each array from top to bottom (based on their y coordinates).
4) **Calculating the Distances between Initial and Final Centers of Caps:** Once all the required centre pixels have been stored and sorted in the arrays, these arrays can be compared to find the Euclidean distances between the corresponding centre pixels in the resting and final grasping images of the gripper. This process can be seen in Fig. 4, where the black lines represent the displacements of the pins, the light green and the light red circles represent the centre pixels when the gripper is at a resting state and the dark green and the dark red circles represent the centre pixels when the gripper is at the final grasping state. The resting and final grasping states have been blended in Fig. 4-b) to demonstrate the process of calculating the distances between the corresponding centre pixels.

Fig. 5. Set of objects that were used in the grasping experiments. The dimensions of the everyday objects used can be found in [24]. The 3D printed objects have the following dimensions: 10x35x50 mm, 22x35x50 mm, 32x35x50 mm, and 50x35x50 mm.

C. Learning to Identify Objects and Estimate Contact Forces

In order to formulate object identification as a classification problem, a non-linear supervised machine learning method was employed, a Random Forest (RF) classifier. RF is a supervised learning algorithm and an ensemble classifier that consists of many decision trees and was originally proposed by Tin Kam Ho of Bell Labs [25] and Leo Breiman [26]. The final output of an RF classification model is the most popular class among all the trained decision trees. The RF model was trained with 100 trees. The feature space used consists of 60 features corresponding to the displacements of the 60 steel rods in each image. For force estimation, a RF regression model was used. The output of the RF regression model is the average output of all the trees. The RF regression model was trained again with 100 trees and the same features (the output was the force exerted). To avoid overfitting the 10-fold cross-validation procedure was used.

IV. Experiments

The experiments were conducted on a set of 12 different objects that included a mixture of everyday objects and 3D printed objects. All everyday objects used in the experiments were selected from the Yale-CMU-Berkeley (YCB) object set [24], an object set designed for facilitating benchmarking in robotic manipulation and grasping. These objects can be seen in Fig. 5. The objects mostly differ in shape, but some objects were deliberately chosen to have the same shapes but slightly different sizes to evaluate how well the gripper can discriminate similar objects.

Five grasping trials were performed for each object and images were captured after each grasp. The torque applied by the motors of the gripper was kept constant for every grasp by monitoring the supplied current. The robotic gripper was held at a fixed position and orientation by the Universal Robots UR5 robotic arm. All the objects were placed at the same position and orientation for all the trials. The lighting conditions were also kept the same throughout the experiments. These strict conditions were essential for the simple classifier that we used. The system can be made more robust to these environmental and positional changes by training it on a more complex classification method like Convolution Neural Networks.

One image was captured when the gripper was at rest and images of the final grasping configurations were captured for all trials (5 images for each of the 12 objects). Each of the 60 images that were taken when the gripper was at the final grasping configuration were compared with the image in which the gripper was resting, using the previously described algorithm. The algorithm derived the displacement values of the 60 steel rods for each of these 60 images. In order to evaluate how accurately this data can predict / identify the 12 different everyday life objects, the Random Forest (RF) classifier was employed. The prediction accuracy of the classifier was calculated over the 10 rounds of the cross-validation procedure (average accuracy).

The second experiment focused on estimating the amount of force exerted by the pins. As discussed earlier in this section, the images captured from the camera only provide information about the displacement of pins, therefore, to estimate the forces an equation or a model describing the relationship between the displacement and the contact force exerted has to be determined. Three different methods have been used for vision-based force estimation. First, a mathematical relationship between the pin displacements and the contact forces exerted on a SS25LA dynamometer (connected with Biopac MP36 data acquisition unit) was extracted. The best-fitting curve was provided by the following 3rd order polynomial:

$$F = 0.0008x^3 - 0.0126x^2 + 0.1175x + 0.0853 \qquad (1)$$

where F is the exerted contact force and x is the displacement of the pins. All the pins are identical and follow the same equation. The other two approaches focused on learning based force estimation. More precisely, the contact force estimation problem was formulated as a regression problem and two different regression techniques were employed for comparison purposes, the Random Forests regression model and an Artificial Neural Network (ANN). The ANN model was constructed with a single hidden layer with four hidden units and it was trained using the Levenberg-Marquardt backpropagation algorithm. The estimation accuracy of the regressors was calculated over the 10 rounds of the cross-validation procedure. The efficiency of the trained model was assessed using the percentage of the NMSE (Normalized Mean Square Error).

V. Results and Discussions

In this section, we present the results of the experiments conducted, focusing on vision-based object identification and contact force estimation.

A. Object Identification

The classifier identified the objects with 100% accuracy, meaning that the objects grasped by the gripper can be identified with complete certainty by exploiting the displacement

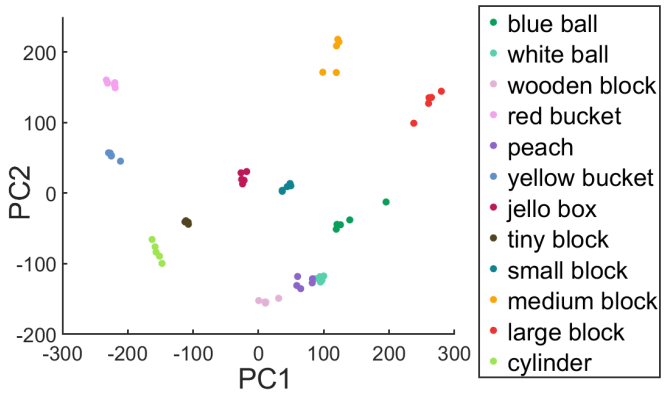

Fig. 6. Distribution of data after employing the Principle Component Analysis (PCA) method. There is no overlapping between data of different objects and the classes are well separated. This means that each class is easily distinguishable and the objects can be identified with high accuracy.

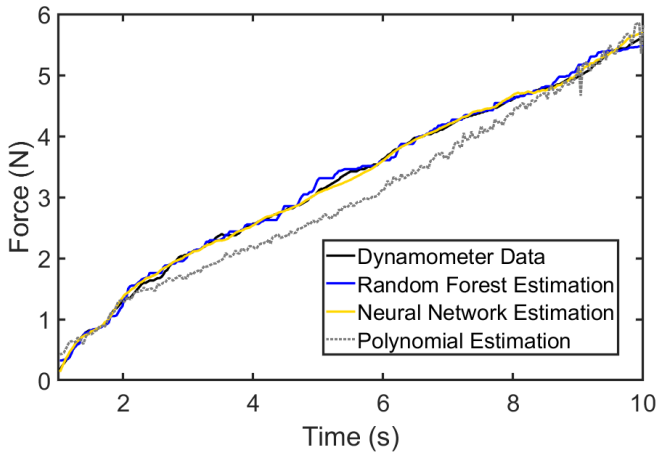

Fig. 7. Comparison between the real force exerted by the pins on a dynamometer and the estimated forces. The 3rd order polynomial estimated the force exerted with an average accuracy of 94.7%, while the Random Forest regression model and the Neural Network model achieved force estimation accuracies of 98.7% and 99.1% respectively.

data for each steel rod of the hyper-adaptive mechanism. To verify that the different objects are easily distinguishable due to the nature of the feature space used and not due to overfitting, a dimensionality reduction technique (Principle Component Analysis, PCA) was used to represent the 60-dimensional feature space in a low-dimensional (2D) manifold (retaining as much of the data variance as possible), where data manipulation and visualization can be facilitated. The outcome of the PCA algorithm can be seen in Fig. 6 where the first component (PC1) of PCA is depicted on the x-axis and the second component (PC2) on the y-axis. The colored dots in the plot represent the low dimensional data and each color represents a different object (12 objects in total). It can be seen that there are 5 dots for each color (the number of experimental trials). As it can be noticed, there is no overlapping between the data of different objects and the examined classes are well separated. This means that even after dimensionality reduction, where some information is lost, the data is easily distinguishable and can be classified with high accuracy.

B. Contact Force Estimation

A graph comparing the force exerted by the finger pins and the force estimated by the different methods employed, is shown in Fig. 7. The 3rd order polynomial estimated the forces with an average accuracy of 94.7%, while the Random Forest regression model and the Neural Network model achieved force estimation accuracy of 98.7% and 99.1% respectively. Hence, such a force estimation method can be easily and efficiently used during object grasping to derive the contact forces exerted and maintain the stability of grasps. According to [27], the quality of the grasp and the resistance to external disturbances is directly related to the number of contact points and the forces applied to the object. Besides the maximization of the contact area and the fine distribution of forces to the pins, estimating the amount of contact force applied to the pins allows grasping of delicate objects without damaging them.

VI. Conclusions and Future Directions

In this paper, we proposed a learning-based framework that employs a hyper-adaptive robotic gripper and vision-based reconfiguration estimation in order to: i) identify a range of everyday objects during the execution of single, open-loop, stable grasps and ii) to estimate the contact forces exerted. This can be easily accomplished with a simple 2D camera that extracts the shape of the grasped objects based on the reconfiguration of the pins of the hyper-adaptive robot gripper, while they are adapting to the surface of the grasped object. The object identification has been formulated as a classification problem employing a Random Forests classifier that provides a classification accuracy of 100%. The contact forces estimation is formulated as a standard regression problem employing three different techniques, Random Forests regression, Neural Networks, and a third-order polynomial. The results of the examined regression techniques range from 94.7% to 99.1%, demonstrating that the proposed methodology can be efficiently used as an alternative to force and tactile sensors that are typically employed for measuring the contact forces.

Regarding future work, the design of the hyper-adaptive robot gripper will be improved by including multiple hyper-adaptive, articulated fingers, increasing the gripping surface of the rubbery tips (also increasing the stability of the grasps), and increasing the density and resolution of the steel rod-based pins. In addition to that, micro-radar sensors, such as the Google Soli, can be combined with the gripper to discriminate between different everyday life objects not only based on their geometries but also based on their volume, material, and density to increase the robustness of the classification procedure for a large inventory of objects. The reliability of the system can be improved by enclosing the camera within the hyper-adaptive finger structure so that it can be isolated from the environment and have a

lighting system connected to the camera. The size of the enclosure depends on the distance of the camera from the finger structure and it can be reduced by using a wide angle lens or by having multiple cameras placed closed to the hyper-adaptive finger, capturing different sections.

REFERENCES

[1] T. Yoshikawa, M. Koeda, and H. Fujimoto, "Shape recognition and grasping by robotic hands with soft fingers and omnidirectional camera," in *2008 IEEE International Conference on Robotics and Automation*, 2008, pp. 299–304.

[2] D. Petković, A. S. Danesh, M. Dadkhah, N. Misaghian, S. Shamshirband, E. Zalnezhad, and N. D. Pavlović, "Adaptive control algorithm of flexible robotic gripper by extreme learning machine," *Robotics and Computer-Integrated Manufacturing*, vol. 37, pp. 170–178, 2016.

[3] T. Watanabe, K. Yamazaki, and Y. Yokokohji, "Survey of robotic manipulation studies intending practical applications in real environments-object recognition, soft robot hand, and challenge program and benchmarking," *Advanced Robotics*, vol. 31, no. 19-20, pp. 1114–1132, 2017.

[4] Z. Kappassov, J.-A. Corrales, and V. Perdereau, "Tactile sensing in dexterous robot hands," *Robotics and Autonomous Systems*, vol. 74, pp. 195–220, 2015.

[5] A. Yamaguchi and C. G. Atkeson, "Combining finger vision and optical tactile sensing: Reducing and handling errors while cutting vegetables," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, 2016, pp. 1045–1051.

[6] N. F. Lepora, K. Aquilina, and L. Cramphorn, "Exploratory tactile servoing with active touch," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1156–1163, 2017.

[7] A. J. Spiers, M. V. Liarokapis, B. Calli, and A. M. Dollar, "Single-grasp object classification and feature extraction with simple robot hands and tactile sensors," *IEEE transactions on haptics*, vol. 9, no. 2, pp. 207–220, 2016.

[8] A. Vàsquez and V. Perdereau, "Proprioceptive shape signatures for object manipulation and recognition purposes in a robotic hand," *Robotics and Autonomous Systems*, vol. 98, pp. 135–146, 2017.

[9] D. Kragic, M. Björkman, H. I. Christensen, and J.-O. Eklundh, "Vision for robotic object manipulation in domestic settings," *Robotics and autonomous Systems*, vol. 52, no. 1, pp. 85–100, 2005.

[10] Q. Lei, G. Chen, and M. Wisse, "Fast grasping of unknown objects using principal component analysis," *AIP Advances*, vol. 7, no. 9, p. 095126, 2017.

[11] B. Wang, L. Jiang, J. Li, H. Cai, and H. Liu, "Grasping unknown objects based on 3d model reconstruction," in *2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2005, pp. 461–466.

[12] G. Taylor and L. Kleeman, "Grasping unknown objects with a humanoid robot," in *Proc. 2002 Australasian Conference on Robotics and Automation*, vol. 27, 2002, p. 29.

[13] A. M. Pinto, L. F. Rocha, and A. P. Moreira, "Object recognition using laser range finder and machine learning techniques," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 1, pp. 12–22, 2013.

[14] A. Saxena, J. Driemeyer, J. Kearns, C. Osondu, and A. Y. Ng, "Learning to grasp novel objects using vision," in *Experimental Robotics*. Springer, 2008, pp. 33–42.

[15] A.-M. Cretu, P. Payeur, and E. M. Petriu, "Soft object deformation monitoring and learning for model-based robotic hand manipulation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 3, pp. 740–753, 2011.

[16] F. Hui, P. Payeur, and A.-M. Cretu, "Visual tracking of deformation and classification of non-rigid objects with robot hand probing," *Robotics*, vol. 6, no. 1, p. 5, 2017.

[17] F. F. Khalil, P. Curtis, and P. Payeur, "Visual monitoring of surface deformations on objects manipulated with a robotic hand," in *2010 IEEE international workshop on robotic and sensors environments*, 2010, pp. 1–6.

[18] U. Asif, M. Bennamoun, and F. Sohel, "Model-free segmentation and grasp selection of unknown stacked objects," in *European Conference on Computer Vision*. Springer, 2014, pp. 659–674.

[19] M. V. Liarokapis, B. Calli, A. J. Spiers, and A. M. Dollar, "Unplanned, model-free, single grasp object classification with underactuated hands and force sensors," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 5073–5080.

[20] Z. Flintoff, B. Johnston, and M. Liarokapis, "Single-grasp, model-free object classification using a hyper-adaptive hand, google soli, and tactile sensors," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1943–1950.

[21] R. Ma and A. Dollar, "Yale OpenHand project: Optimizing open-source hand designs for ease of fabrication and adoption," *IEEE Robotics & Automation Magazine*, vol. 24, no. 1, pp. 32–40, 2017.

[22] B. Ward-Cherrier, L. Cramphorn, and N. F. Lepora, "Tactile manipulation with a tacthumb integrated on the open-hand M2 gripper," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 169–175, 2016.

[23] C.-M. Chang, L. Gerez, N. Elangovan, A. Zisimatos, and M. Liarokapis, "On alternative uses of structural compliance for the development of adaptive robot grippers and hands," *Frontiers in Neurorobotics*, vol. 13, p. 91, 2019.

[24] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, "Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set," *IEEE Robotics Automation Magazine*, vol. 22, no. 3, pp. 36–52, 2015.

[25] T. K. Ho, "Random decision forests," in *Proceedings of the third international conference on Document analysis and recognition*, vol. 1, 1995, pp. 278–282.

[26] L. Breiman, "Random forests," *Machine learning, Springer*, vol. 45, no. 1, pp. 5–32, 2001.

[27] M. A. Roa and R. Suárez, "Grasp quality measures: review and performance," *Autonomous robots*, vol. 38, no. 1, pp. 65–88, 2015.