Smart Squishy Sorting: A Tactile-Property-Driven Robotic System for Toy Cleanup and Sorting

Xander Backus
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA
xabackus@mit.edu

Anushka Aggarwal
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA
anushka6@mit.edu

Rachael Cai
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA
rachcai@mit.edu

***Abstract*—This project presents a robotic sorting system designed to autonomously classify and organize toys scattered on the floor into designated bins based on their tactile properties. The toys are categorized as either hard or soft, with examples ranging from rigid toy cars and Legos to soft dolls and teddy bears. Leveraging a vision-based perception system, the robot detects and localizes toys, then executes pick-and-place operations using an IIWA robotic arm while determining its trajectory based off each object's hydroelastic properties, classified through force-sensing feedback. The system evaluates object elasticity by measuring deformation under controlled force, allowing accurate classification. Motivated by the repetitive nature of toy cleanup tasks and the broader applicability of tactile-based robotic sorting in industrial and logistics environments, this project seeks to bridge the gap between vision-based object recognition and force-based material characterization. While our initial goal was a sorting accuracy of 80%, through these methods near 100% accuracy in object classification is achieved.**

***Index Terms*—robotic sorting system, vision-based perception, force sensing, autonomous sorting.**

I. Introduction

The recent rapid advancements in robotics and automation address a various real-world challenges, ranging from industrial applications to household chores. One persistent problem faced in many homes is the repetitive and tedious task of organizing toys scattered by children. This project addresses this issue by designing a robotic system that autonomously sorts toys based on their physical properties, specifically tactile properties such as hydro-elasticity.

The system focuses on dividing toys into two broad categories: hard toys, including items such as toy cars, Legos, and rubiks cube; and soft toys, such as dolls, dog toys, and teddy bears. The distinction between these categories is based on the tactile properties of the toys, specifically their elasticity, which determines how much they deform under applied force.

The robotic system consists of a vision-based perception system integrated with motion planning and tactile sensing capabilities. The robot's perception system identifies and localizes the toys, analyzing their point cloud data to understand their spatial configuration. The IIWA robotic arm, equipped with a WSG gripper and force sensors, carries out the task of picking and placing the toys. Force sensing enables the system to determine whether an object is hard or soft based on how it deforms when gripped. This integration of vision and tactile feedback is critical to achieving high accuracy in sorting and classification.

The overall environment for the system is designed to provide controlled testing conditions. Toys are initially placed in the "play area", ensuring sufficient spacing between them to facilitate perception and prevent overlap. Two empty bins are positioned within the robot's reach, one designated for hard toys and the other for soft toys. All toys are modeled in simulation with separate visual and contact geometries for the greater computational efficiency compared to more detailed meshes and also for convenient experimentation with different geometries. This setup allows the robot to perform pick-and-place operations efficiently without requiring excessive mobility. The project aims to achieve at least 80% classification and sorting accuracy as its minimum viable product.

A key aspect of the project is its adaptability and scalability. While the initial implementation targets a controlled environment with limited toy variety and reach, the system design has been structured to support future enhancements. These include camera systems above both bins to calculate point clouds within each bin, introducing the possibility for extended work like moving objects between bins and reorienting toys like dolls and cars to be stored upright or in some other specific way.

The motivation behind this project extends beyond the convenience of automating household chores. Sorting by tactile properties is a crucial capability for robotics, particularly in industrial and logistics settings. For instance, factories and warehouses often require robots to distinguish between items based on their material properties, such as hardness or elasticity. The ability to integrate vision and force sensing into a cohesive system has implications for a range of applications, from packaging and quality control to agricultural sorting tasks. This project serves as a proof of concept for such systems, demonstrating how the fusion of perception and tactile feedback can enable intelligent robotic sorting in constrained environments.

Despite the progress in robotics, challenges remain in integrating multiple capabilities into a seamless system. The primary challenges in this project include combining visual perception with motion planning, enabling the robot to identify toys, plan grasping motions, and execute placement actions. Another significant challenge is the determination of toy elasticity using force sensors. Designing an effective tactile sensing algorithm to distinguish between hard and soft toys based on their deformation is critical to the project's success.

The project contributes to advancing the field of robotic manipulation and sorting by addressing these challenges. It leverages concepts such as motion planning, force control, and object recognition, all of which are essential in developing intelligent robotic systems. Moreover, the project emphasizes a modular approach to system design, enabling individual components like vision processing, tactile sensing, and motion planning to work cohesively.

II. Related Work

A. Blind Manipulation of Deformable Objects Based on Force Sensing and Finite Element Modeling [1]

Recent advancements in robotic manipulation of deformable objects have explored alternatives to vision-based methods, addressing challenges such as occlusion and lighting sensitivity. A notable approach leverages force sensing combined with finite element modeling (FEM) to estimate and control deformations without visual input. While the system we're building will not have access to the initial meshes of objects ahead of time for practical purposes and generalization, it will be acting similarly, using applied force as an input and outputting an estimation for whether an object is rigid or soft based on deformation.

B. Robot Autonomous Sorting System for Intelligent Logistics [2]

This system captures frame images, transforming them into coordinates for path planning and grasping. The strategy of linking visual data to robot movement provides a good starting foundation for our sorting system's design.

C. Lessons from the Amazon Picking Challenge [3]

This reports on a robotic arm that picks up objects using a system that focuses on finding a feasible balance in four key aspects: tight integration and constant revision of modularity, combining computation and enmeshment, weighing feedback over planning, and using reasonable assumptions to simplify general problems. This resulting arm can accurately pick objects from a shelf, winning the Amazon Picking Challenge. The consideration behind creating a system that works smoothly alongside hardware and the findings from the four key aspects mentioned in the paper should be noted when designing a system where both processing and physical movements have a strong impact on its success.

D. A Comprehensive Review of Robot Intelligent Grasping Based on Tactile Perception [4]

This review compares various grasping techniques, evaluating their speed, accuracy, and adaptability. The findings will inform our choice of algorithms to handle toys with diverse textures and forms.

E. Learning the elasticity parameters of deformable objects with a manipulation robot [5]

This robot evaluates the elasticity properties of deformable objects by applying controlled forces and observing object reaction to those forces. This approach provides a possible strategy for our system to determine and categorize an object's elasticity during the grasping stage.

III. Technical Approach

The technical approach for this project focuses on designing and implementing a robotic sorting system that integrates vision-based perception, tactile sensing, and motion planning to autonomously classify and organize toys based on their physical properties. This section details the system architecture, methodologies, tools, and algorithms employed to achieve the project's objectives, emphasizing the steps taken to ensure accuracy, efficiency, and scalability.

A. Environment Simulation in Drake

The project employs Drake to simulate the environment, allowing for the testing and refinement of the sorting system. The simulation setup includes:

1) Toy Models: Each toy is modeled with realistic geometric and physical properties, including elasticity.

2) Bin Placement: Two bins are positioned within the robot's reach to minimize unnecessary movement.

3) Toy Placement: Toys are positioned on the floor in the play area, within view of several cameras and within reach of the arm, while also ensuring sufficient spacing to simplify perception tasks.

We chose to model each toy using separate visual and contact geometries that occupy the same volume instead of a mesh-based approach for greater speed, stability, and ease in geometry experimentation at the loss of visual deformation. This trade-off was made keeping in mind the scope of the project and the amount of object experimentation required to confirm a consistent method for classification.

B. System Architecture

The robotic sorting system is built around the simulation environment using Drake, which provides tools for physics-based modeling and simulation. The primary components include:

1) Vision System: Responsible for detecting and localizing toys in the environment.

2) IIWA Robotic Arm: Equipped with a WSG gripper for grasping and tactile sensing.

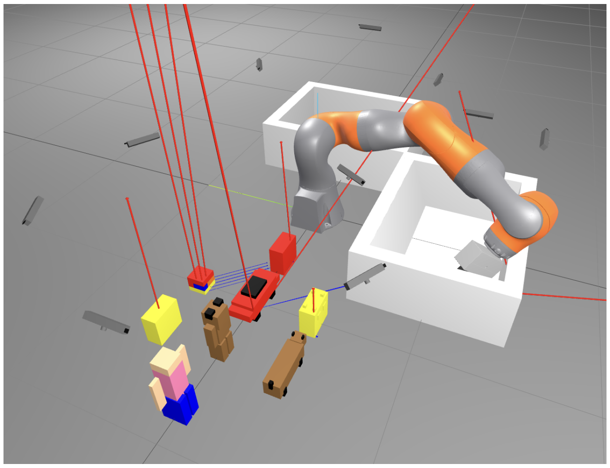

Fig. 1. Initial simulation setup with objects at locations within reach of the arm and within view of the cameras

3) Control System: Manages communication between vision, motion planning, and tactile sensing components.

4) Environment Setup: Includes spaced-out toys and designated bins for sorted objects.

The workflow begins with the vision system identifying toys on the floor, followed by motion planning for the robotic arm to grasp a toy, analyze its elasticity using tactile sensing, and place it in the appropriate bin.

C. Vision-Based Perception System

The vision system is responsible for detecting toys and estimating their positions and orientations. A point cloud-based approach is used to analyze the environment. The vision pipeline consists of the following.

1) Toy Detection: Using object detection algorithms to identify toys based on their shapes, sizes, and spatial distribution.

2) Point Cloud Processing: Using information from depth cameras near the arm, the perception system generates a 3D point cloud of the environment, enabling precise localization of toys.

3) Point Cloud Cropping: We limit the point cloud that the arm should use for gripping to the play area. This is done to keep it from grabbing cameras or bins.

4) Segmentation and Classification: The point cloud is segmented to isolate individual toys.

The vision system uses a depth camera to generate a 3D point cloud, segment toys, and localize them.

D. Grasping and Tactile Sensing for Hydroelasticity Measurement

Classifying toys into hard and soft categories requires analyzing their deformation under force. The system employs force sensors integrated into the robotic gripper to measure toy elasticity. The process involves:

1) Antipodal Grasping and Motion Planning: The system takes in point clouds from the vision system to determine the robot arm's trajectory to optimally grasp a nearby toy and estimate the distance between the toy's antipodal points.

2) Grasp Calibration: The gripper applies a constant force to the toy during the pick stage of the pick-and-place operation, and monitors the separation between the ends of the gripper at the start of the postpick step.

3) Deformation Analysis: The deformation is recorded using the separation feedback. Hard toys exhibit slight deformation, whereas soft toys deform significantly.

4) Classification Algorithm: A threshold is defined in ratio of the grip-distance to antipodal-width. Soft toys exhibit lower grip-separation/antipodal-distance ratios, signifying significant deformation under force. Hard toys exhibit higher grip-separation/antipodal-distance ratios, signifying negligible deformation. After running many experiments, we chose a separation threshold ratio of 0.8.

This approach ensures that the system can handle a wide range of toy materials, shapes, and sizes while maintaining high classification accuracy.

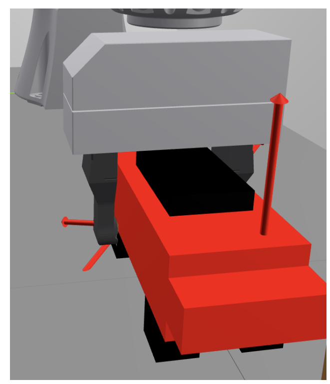

Fig. 2. Arm grips car, taking into account gripper spacing

E. Bin Placement

Since the frames passed into MakeGripperPoseTrajectory must include a "place" frame before the pick-and-place operation can begin, and ending one pick-and-place operation results in the gripper opening before the next operation can start, pausing in the middle of the operation to classify an object as hard or soft and then continuing the operation based on the result of that classification is very difficult.

To get around these limitations, we do the following:

1) Trajectory Planning: Once an object has been selected for gripping and an antipodal grasp candidate has been selected, we create a common pick trajectory composed of a "prepick" frame, a "pick" frame, a default "postpick" frame, and a dummy "place" frame.

We simultaneously precompute the two separate place trajectories for placing the object into the hard bin and the soft bin, respectively, consisting of "preplace", "place", and "postplace" frames.

2) Object Classification: During the "postpick" frame point in the trajectory, the system calculates the deformation ratio and classifies the object in the gripper as hard or soft.

3) Trajectory Switching: Based on the result of the classification in the "postpick" step, we dynamically choose which placement trajectory to follow for the placement stage. The pick-and-place operation goes on to follow the "preplace", "place", and "postplace" frames from the new trajectory, placing the object into one of the bins at the end.

```
Press Escape to stop the simulation
Planned full pick-and-place (both hard and soft) trajectories at time 1.1.
The object is soft. Ratio: 0.51
Planned full pick-and-place (both hard and soft) trajectories at time 31.700000000000003.
The object is hard. Ratio: 0.94
```

Fig. 3. Deformation ratio determined (teddy bear and model car)

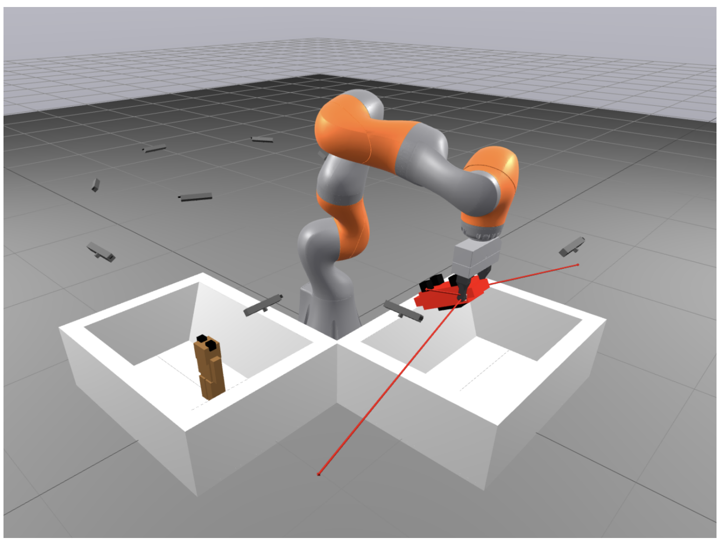

Fig. 4. Arm places car in the bin for hard objects

F. Motion Planning and Grasp Execution

Motion planning is critical for enabling the IIWA robotic arm to execute pick-and-place operations. The process involves:

1) Path Planning: Using inverse kinematics, the arm generates collision-free paths to approach the toy.

2) Antipodal Grasping: The system evaluates potential grasp points on the toy, ensuring a stable grip by analyzing surface curvature and friction properties.

3) Pick-and-Place Execution: The arm executes the planned path to grasp the toy, analyze its properties, and place it in the appropriate bin.

The motion planning algorithm incorporates real-time feedback from the vision and tactile systems to adjust the arm's movements dynamically. For example, if a toy is misaligned during an attempt to grasp, the arm recalculates its trajectory to ensure a successful operation.

G. Integration of Vision, Tactile Sensing, and Motion Planning

The integration of multiple subsystems is achieved through a modular architecture. Each subsystem communicates through a central control framework, which manages the flow of data and commands. The workflow is as follows:

1) Detection and Localization: The vision system identifies toys and provides their coordinates to the control framework.

2) Motion Planning: The control system generates a trajectory for the robotic arm to approach the toy.

3) Tactile Sensing: The gripper measures the toy's deformation during grasping, and the control system classifies it as hard or soft.

4) Placement: The arm places the toy in the designated bin based on its classification.

In the initial structure of the system, we tried separating picking the object and placing the object into two separate pick-and-place operation, with categorization happening during the first and influencing the second. However, this led to the arm constantly dropping each object immediately after classifying it up, due to the ending of each pick-and-place object causing the arm's gripper to open.

We switched to a modified version of the system by adding new cameras and using two types of pick-and-place operations, with the first having the arm pick up an object and place it in an area away from the other objects while assigning it a classification based off the deformation ratio, and the second having the arm look for that recently placed object in that specific area, pick it back up, and place it in a bin based off its classification. However, this added unnecessary complexity with new cameras and a new grasp selector, as well as introducing more room for error and taking twice as long.

Ultimately, we decided on the current approach, which combines speed, accuracy, and efficiency.

Our modular approach ensures flexibility, allowing individual subsystems to be upgraded or replaced without disrupting the entire system. This was very useful, as both our vision system and pick-and-place system went through multiple revisions, but changes to one of these systems did not negatively impact the other systems.

IV. PRIMARY CHALLENGES

The development of the robotic sorting system involved addressing several key challenges to achieve accurate and efficient toy classification and placement. These challenges stemmed from the integration of diverse subsystems, the need for precise sensing, and the physical manipulation of objects with varying properties.

A. Adding Arbitrarily Many Cameras

The first major challenge we encountered was creating a system to create and integrate many new custom cameras. The documentation regarding camera configurations in Drake is opaque, and there are many variable inputs for a CameraConfig object, each of which can cause a camera to be essentially useless if not assigned a good value.

B. Integrating Visual Perception with Motion Planning

One of the primary challenges was combining visual perception with motion planning to enable seamless grasping and placing of toys. The vision system needed to accurately detect toys, localize them within a noisy environment, and provide reliable input for motion planning. The robot arm's movements required precise coordination with the perception data to ensure stable and efficient pick-and-place operations, particularly in dynamic or cluttered settings. This was especially difficult for small objects (objects which are deficient in size along at least one dimension such as the Rubik's cube or the Teddy bear), because when the margin of error for the motion planning is large relative to the size of an object, it is very easy for the gripper to miss the object or get an unstable grip on the object's edge.

C. Determining Hardness and Softness of Objects

Classifying toys as hard or soft based on their deformation response to force was another challenge. Since the classification of an object as hard or soft depends on the gripper squeezing the object at the grasping location where an antipodal grasp candidate was selected. For objects which have unstable grasp geometries, such as the original car design we had planned, the gripper is likely to slip to a more narrow grip, resulting in a gripper-separation/antipodal-distance ratio that is too low and a subsequent potential misclassificatin of a hard object as a soft object.

D. Smoothly Changing Trajectory after Classification

An unexpected challenge during the motion planning part of the project was setting up the system to grip an object, classify it, and then set it on the right trajectory to one of the bins. This had to be done in one smooth pick-and-place operation, since each pick-and-place movement involved opening up the gripper at the end and dropping the object. This challenge was a time-consuming part of the project as we were unclear on whether it would be possible for the robot arm to perform an immediate trajectory switch in the middle of a pick-and-place operation without opening its grip and dropping the object. A lot of time was spent working to restructure the system, create potential workarounds, and finally, find a way to swap trajectories without dropping the object.

Addressing these challenges was critical for building a robust system that could sort toys with high accuracy and reliability in diverse conditions.

V. Evaluation Criteria

The system's performance is based on its classification and sorting accuracy. The primary objective of the system is to accurately determine whether whether an object is soft or hard, which we prioritize in the evaluation.

The metric for success is the system's ability to classify each toy as hard or soft and move it into the correct bin with at least 80% accuracy.

To assess performance, we created four simulations designed to test the classification and sorting accuracy of the system. Each simulation had two of eight unique objects within reach of the robotic arm for sorting. We limited each simulation to two objects to prevent potential issues caused by overcrowding in the simulation space, like the arm accidentally knocking objects out of frame or getting stuck due to limited space. While these issues are potential cause for concern in any grasping system, this setup helped ensure that most of the errors observed would come from issues in the classification algorithm rather than mechanical failures or collisions.

Each simulation was run five times to account for variability in performance. For each trial, we recorded whether the robotic arm was able to correctly follow through for each stage, from picking to classification to placing.

VI. Results

We counted the cases without failures across each stage. With the picking stage, we considered gripping failures in cases where the object was knocked out of play area or the arm never attempted to grip the object. With the classification stage, we considered cases where the object was given the wrong classification after being picked properly. With the placing stage, we considered movement failures where the gripper dropped the object on the way to the bin.

After 20 runs of 2 objects each, the results are as follows:

Stage	Before	After	Accuracy
Picking	40	33	82.5%
Classification	33	33	100%
Placing	33	30	90.9%
Overall	N/A	N/A	75%

TABLE I
EVALUATION RESULTS BY SYSTEM STAGES

We calculate the accuracy for each stage using the proportion of objects that correctly move onto the next stage out of all the objects that reach the current stage, with the accuracies as follows:

Gripper Accuracy: 82.5%

Classification Accuracy: 100%

Movement Accuracy: 90.9%

Sorting Accuracy: 100%

The combined accuracy of our classification and sorting system, i.e. the probability that we correctly classify and sort an object given that the gripper manages to grab it and does not accidentally drop it or bump into a camera on the way to the bin, is near 100%. Given this, it is safe to say that we have succeeded in making a working classification and sorting system.

VII. Future Work

In addition to the primary objectives, the project incorporates two stretch goals to enhance the robotic sorting system's functionality and scalability. Future work could focus on expanding the robot's operational range and improving its precision during placement tasks.

A. Mobile Base for Enhanced Coverage

The first stretch goal involves equipping the robot with a mobile base, enabling it to traverse the room and access toys beyond the initial reach of its stationary arm. This upgrade requires the integration of navigation and mapping capabilities, such as Simultaneous Localization and Mapping (SLAM), to allow the robot to move efficiently within its environment while avoiding obstacles. A mobile base significantly expands the robot's operational range, making it adaptable to larger spaces and more complex configurations of scattered toys. To achieve the first reach goal, the robot must demonstrate the ability to move beyond its initial static position and drive to access toys further away.

B. Organized Placement and Orientation

The second stretch goal focuses on developing an algorithm for precise placement of toys within bins, particularly for orientable objects such as dolls, teddy bears, and toy cars. The system will ensure these objects are placed in their correct orientation, such as upright or in a predefined alignment. Achieving this requires enhancing the perception system to identify toy orientation during grasping and integrating motion planning adjustments for accurate placement. This feature improves storage organization and replicates human-like sorting behavior.

These potential future goals enhance the system's functionality, ensuring greater adaptability and precision in diverse settings such as factories and warehouses. By meeting these criteria, the robotic sorting system demonstrates its capability to efficiently automate the classification and sorting of rigid and deformable objects, paving the way for broader applications in industrial and domestic settings.

VIII. Conclusions

This work demonstrates an autonomous robotic sorting system that combines vision-based perception, force sensing, and motion planning to organize toys based on their hydroelastic properties. The system achieved high classification accuracy for grasped objects, confirming that tactile classification and feedback can be integrated with a vision perception system to successfully sort objects based on tactile properties.

However, there were some failures that occurred in the sorting system happened outside of the classification stage. In the system, classification success depends on reliable motion planning to first accurately grasp the object before classification and sorting. In the results, primary sources of error came from grasping failures due to collisions or unstable picks.

Outside of more robust motion planning, potential future improvements include expanding the robot arm's range using a mobile base to navigate around a larger scenario with toys scattered further apart. Another possibility is focusing on enhancing the vision system and the system's placement precision with object-specific orientation algorithms. These developments would extend the sorting system's use to tasks in warehouses and manufacturing that require precise placement during sorting of deformable and rigid objects. With continued refinement, this system holds promise for automating material-handling tasks in structured environments.

IX. Acknowledgment

We would like to express our gratitude to Professor Lozano-Perez, Dr. Kalodner-Martin, and the rest of the 6.421 course staff. We would particularly like to thank Quincy Johnson for his help in brainstorming solutions to the many issues we encountered throughout the course of this project. Finally, we would like to thank Professor Russ Tedrake and his associates for inventing Drake and making this project possible.

X. Contribution

Xander created the initial scene with a table and cameras around the play area filled with SDFs, integrated camera functionality so that all new cameras contributed to the point cloud, designed the vision- and gripping-based classification system, designed the dynamic pick-and-place system (as well as several of its predecessors), and fine-tuned the SDF files to give them more stable contact geometries and make them less top-heavy. Rachael set up the custom objects and their positions and welded the table for the new play area scene, and designed and implemented alternative planner staging and pick-and-place operations based on separate point clouds to suit the classification system. Anushka designed and managed the SDF files, assisted Xander in setting up the visual system, created a point cloud system and took the lead in the communication component of the project.

References

[1] J. Sanchez, K. Mohy El Dine, J. A. Corrales, B.-C. Bouzgarrou, and Y. Mezouar, "Blind Manipulation of Deformable Objects Based on Force Sensing and Finite Element Modeling," Frontiers in Robotics and AI, vol. 7, Jun. 2020, doi: https://doi.org/10.3389/frobt.2020.00073.

[2] M. Song, S. Xin, "Robot Autonomous Sorting System for Intelligent Logistics," *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2021.

[3] C. Eppner, et al., "Lessons from the Amazon Picking Challenge," *Robotics: Science and Systems*, 2016.

[4] T. Li, et al., "A Comprehensive Review of Robot Intelligent Grasping Based on Tactile Perception," *Robotics and Computer-Integrated Manufacturing*, 2024.

[5] B. Frank, et al., "Learning the Elasticity Parameters of Deformable Objects with a Manipulation Robot," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010.