Key Capabilities of Autonomous Mobile Platforms for Maintenance and Monitoring in Manufacturing Environments

Samantha Staudinger[*1], Yung-Ching Sun[*1], Hanna Chapin[2], Alyssa Carter[2], Kira Barton[1], and Dawn Tilbury[1]

***Abstract*—Quadruped robots are increasingly being deployed in manufacturing environments for autonomous inspection and manipulation tasks due to their unique advantages in mobility, such as navigating in tight spaces and stairs, which makes them ideal candidates for manufacturing plants. This paper investigates the potential of quadruped robots equipped with a 6-degree-of-freedom arm to perform complex tasks, including pick-and-place operations and identifying target objects in a cluttered environment. We conducted experiments focused on quadruped robots' ability to autonomously pick up, carry, and place a bucket in a structured workflow. We also investigated the use of computer vision and AprilTags [1] for object detection and tracking. The quadruped demonstrated robust performance in both tasks, although challenges arise when handling heavier loads or finding target objects in cluttered environments. These findings make clear the capabilities and limitations of autonomous quadruped robots with manipulators in manufacturing settings.**

I. Introduction

Manufacturing plants are using quadruped robots to automate a variety of operational tasks, particularly autonomous inspections [2, 3]. Previous work in [4, 5, 6] focused on the use of multiple sensors to inspect equipment for anomalies, such as air leaks and thermal problems, which is necessary to maintain operational efficiency. Recently, there has been interest in automating more complex tasks that involve human-robot interactions and safety, particularly pick-and-place, manipulation, and handover of objects.

Mobile manipulation robots have increasingly gained interest in industrial environments as they are capable of moving and performing dexterous manipulation tasks. They can navigate facilities, pick up and transport components, and even handle intricate tasks such as assembly or inspection, making them adaptable in a dynamic industrial setting [7].

In this paper, we use a Boston Dynamic's Spot robot [8] to conduct experiments. The mobility of Spot allows it to traverse industrial environments with stairs or dynamic objects to perform efficient manipulation actions such as picking up, carrying, and placing objects. One potential use case is a feeder test, which is used to compare the actual flow rate to the intended flow rate of various liquids. Specifically, a feeder test is currently performed by human operators, which requires precise actions such as locating and picking up a bucket, as well as opening and closing a valve to measure the liquid flow. Spot's 6-degree-of-freedom arm, with a reach of 984 mm (3.22 feet), a lifting capacity of 11 kg (24.25 lbs), and a carrying capacity of 5 kg (11 lbs) [9], makes it well suited for autonomously performing the feeder test. However, challenges remain in tasks such as locating objects in a cluttered environment or picking up and moving a weighted object like a bucket with liquid. As a case study, we explore the use of a Spot robot [8] with an integrated arm to investigate the capabilities and limitations of quadruped robots with manipulators for autonomous inspections within manufacturing environments.

Our main contribution focuses on experimental results used to investigate the ability and accuracy of Spot to pick up and move a bucket with varying weights as well as identify and approach a target bucket in a manufacturing-like environment. These experiments provide insights into the capabilities and limitations of quadruped robots with manipulators when identifying, safely navigating toward, and manipulating weighted objects in manufacturing settings with humans.

II. Background

Quadruped robots like Spot have the unique ability to navigate industrial environments with uneven terrain, narrow spaces, and other obstacles where wheeled robots struggle to traverse effectively. Most notably, the major advantage of using quadruped robots is the ability to traverse stairs, which enables multi-story inspections [10]. With recent advancements in quadruped robots, their application in autonomous maintenance and inspections has been gaining increasing attention and adoption. [11] and [12] designed quadruped robots for inspection in a substation featuring complex terrains such as steps, stone roads, and stairs. [13] proposed an autonomous quadruped robotic system capable of navigating and inspecting oil and gas platforms autonomously.

Most autonomous maintenance and inspection robots currently depend on multi-sensor perception systems for their daily operations [14]. For instance, mobile robots use a combination of infrared and visual sensors to detect equipment malfunctions before they result in significant delays. The ExR-2 robot, for example, uses optical and thermal imaging cameras to spot defects and anomalies in industrial environments [15]. Another application involves mobile robots equipped with depth cameras and visual feedback to monitor aircraft systems during assembly and provide assistance in manual processes [14]. Additionally, mobile robots with perception systems are used in hazardous environments to inspect critical infrastructure, such as pipelines or reactors, to identify issues like corrosion, wear, and leaks [16].

[*]Equal contribution.
[1]University of Michigan, Ann Arbor, MI 48109.
[2]Nestlé Purina, St. Louis, MO 63102.
This research was supported in part by a grant from Nestlé Purina.

Although various sensors can provide rich environmental information, certain inspection and maintenance tasks in manufacturing plants, such as checking for loose fasteners on assembly lines [17], performing feeder tests in automated production systems, or adjusting machine components, often demand precise physical interaction, which mobile robots with only sensing capabilities cannot perform. Equipping mobile robots with manipulators could bring more automation to maintenance and inspections. [18] introduced a mobile platform that can use the onboard manipulator to operate a handheld partial discharge (PD) sensor in substations to measure acoustic and transient earth voltage data for inspections. Ontario Power Generation has been using Spot and its arm to automate the process of circuit breaker tripping and racking, with a human issuing high-level commands to ensure safety [19].

While quadruped robots are widely used for autonomous maintenance and inspections, most tasks focus only on mobility and inspection. When equipped with a manipulator arm, there are still many opportunities and limitations to explore. Complex tasks requiring fine motor skills and physical interaction need both mobility and manipulation. However, the accuracy, lifting capability, reliability, and other limitations of these robots in delicate tasks require further investigation. In this paper, we begin to address these capabilities in the context of a potential new task that a quadruped robot with a manipulator arm may be able to accomplish in a manufacturing environment.

III. Manipulation with the Robotic Arm

In this section, we evaluate the capability of a quadruped robot to securely and safely pick up and transport a 5-gallon bucket as a payload. The key challenge is ensuring that the robot can accurately detect the bucket's position, securely grasp the payload, and maintain stability during transport, preventing any swinging or misplacement while walking autonomously. The goal is to develop a reliable workflow that allows the robot to consistently secure the bucket and transport it without risk of dropping or destabilizing the load. We consider three phases: bucket manipulation, pick and place maneuvering, and the addition of weight to the bucket. All phases in this section include an AprilTag on the bucket; identification of the bucket without and AprilTag is considered in Section IV

A. Bucket Manipulation with AprilTags

The workflow involves three main stages: 1) identifying the bucket's position in the world frame, 2) commanding the quadruped to walk to the front of the bucket, and 3) vertically picking up the bucket.

In the first stage, an AprilTag [1] detects and tracks the bucket's handle. The workflow checks all camera sources for the AprilTag, returning its bounding box. The 2D pixel coordinates are then converted into 3D camera coordinates using the intrinsic matrix and Perspective-n-Point algorithm [20], and transformed to the world frame. In the second stage, the go-to point for the quadruped is calculated with a 0.5-meter offset from the fiducial point. In the third and final stage, Spot unstows its arm and uses the grasp mechanisms to lift the bucket.

For this first stage, we assumed that: *The bucket handle stays upright throughout, with both the bucket and Spot starting in the same positions. The reachable workspace is limited to the area in front of Spot, with all goal locations within it.*

We found that Spot can always locate the bucket when fixed with an April tag and placed within the field of vision. Spot can sufficiently pick up the bucket from this location even when placed at any orientation. The results for workspace testing is shown in Fig. 2 (A) and (B) for empty and 5 lbs (2.27 kg) of weight in the buckets.

The empty bucket manipulation trials showed an average deviation of 3.8 cm from the goal location with a standard deviation of 1.1 cm, indicating good accuracy and precision, and a 21/21 success rate. The trials highlighted that the bucket sometimes swung after placement, causing deviations, especially in goal locations at the extremities of the reachable workspace. The worst performance was observed when placing the bucket directly in front of the robot, likely due to the swinging caused by the movement's perpendicularity to the bucket's handle.

We repeated the same tests with 5 lbs (2.27 kg) of weight in the bucket. The results from the 5 lbs (2.27 kg) weight in the bucket trial showed a mean deviation of 3.2 cm and a standard deviation of 1.4 cm, which are similar to the empty bucket trial. While the extremities of the reachable workspace still performed the worst, the added weight improved performance in the location directly in front of the robot by reducing swinging. Overall, adding the 5 lbs (2.27 kg) to the bucket did not significantly affect accuracy, precision, or reliability.

B. Autonomous Pick and Place Maneuvering onto Quadruped's Body

In the second stage, we develop a workflow for the quadruped's pick-and-place task to transport a payload.

The proposed workflow involves placing the bucket on the quadruped's back for stability, unobstructed movement, and efficient payload distribution. The back ensures the robot's center of gravity remains balanced, while keeping the bucket out of the mobility path.

The workflow begins with the robot arm unstowing and positioning itself. The gripper opens to prepare for the bucket, aligns with the fiducial marker, and closes to securely grasp it. The arm lifts the bucket slightly, then moves toward the back while maintaining the object in the kinematic frame. Finally, the arm follows a precise trajectory to center, rotate, and lower the bucket onto the back. The pictorial workflow is in Fig. 1.

For the first phase, we assumed that: *The bucket handle stays upright throughout, with both the bucket and Spot starting in the same positions. The reachable workspace is limited to the area in front of Spot, with all goal locations within it.*

The results, shown in Fig. 2 (C), reveal that all 8 trials were successful. A clear pattern emerged: positions directly in front

Fig. 1. Overall Workflow of Pick and Place of Bucket onto Back of Spot

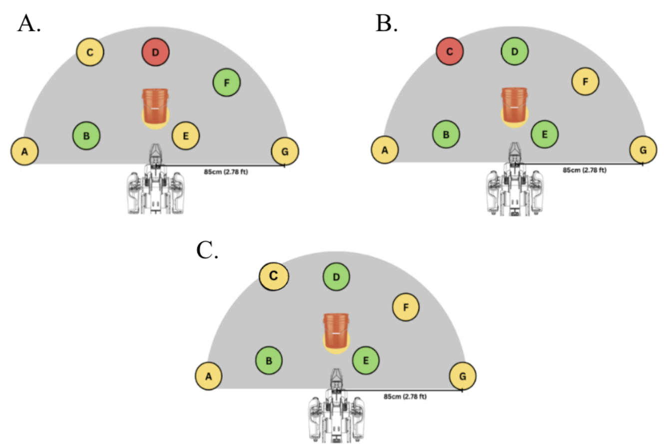

Fig. 2. Green goal locations if the average deviation of that location is below the overall mean, the yellow goal locations are within one standard deviation above the mean, and the red is more than 1 standard deviation above the mean. (A) Empty bucket trial test based on goal location (B) 5 lbs (2.27 kg) bucket trial test based on goal location (C) Pick and Place of Bucket on Back trial test.

of the robot or near its central coordinate frame provided the most stable and efficient workspace. Angled approaches, however, showed limitations, likely due to kinematic constraints related to the robot's central coordinate frame, affecting its stability during these approaches. Hence, the testing revealed that the best place to ensure proper pick and place of the object is limited to positions directly in front of the robot or within its central coordinate frame, where stability and efficiency are maximized.

C. *Weighted Pick and Place Testing*

In the third phase, we added weight to the bucket and repeated the tests, evaluating the limitations of the amount of weight the arm can carry to maneuver objects in pick and place tasks

The robotic arm successfully completed the pick-and-place task up to a weight of 17.5 lbs (7.94 kg), beyond which it failed to perform the operation effectively. The arm began dragging the bucket when the weight reached 7.5 lbs (3.40 kg), indicating the maximum load it could handle before experiencing issues with stability and control.

IV. Bucket Manipulation without AprilTags

While the use of Apriltags facilitated a focus on the manipulation capabilities of the robot, the requirement to add the fiducial markers may not be feasible in real-world industrial settings. Therefore, this section explores whether the robot, when integrated with computer vision techniques, can reliably detect and approach the target object within a manufacturing-like environment.

Our experiments employed Boston Dynamics' Spot robot, which features five optical and five depth cameras positioned at the front-left, front-right, left, right, and rear of its body. The optical cameras offer a full 360-degree view, while each depth camera covers approximately 90 degrees.

The overall workflow is shown in Fig. 3. The robot captures RGB images using its optical cameras and applies computer vision techniques to detect the target bucket. It then calculates the object's position in the world frame and navigates to its front. To simplify the task, we assume the target bucket is always the same type of orange bucket and is only one within Spot's field of view. Reliable detection is critical for this task. We adopted YOLO11 [21] model and implemented two approaches to detect our target bucket: (1) computer vision techniques with YOLO11n [21] pretrained model and (2) a bucket detection model trained on custom datasets.

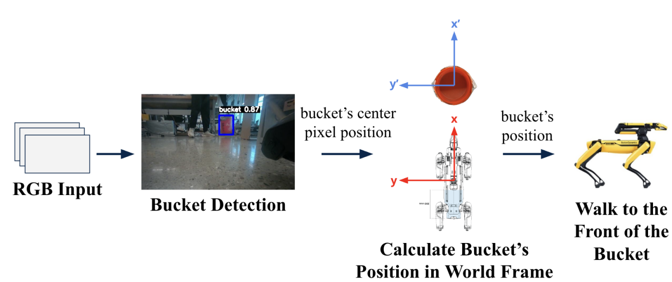

Fig. 3. The workflow of bucket manipulation without using Apriltags.

A. *Apply Computer Vision Techniques to Detect Target Bucket*

To allow for future extension to various objects, we initially utilized YOLO11n [21], a lightweight object detection model pretrained on the COCO dataset [22]. However, since COCO lacks a "bucket" class, YOLO11n [21] tends to classify the bucket as a "cup"; thus, we temporarily used "cup" as the target class during the tests. Given our assumption that the target bucket is always orange, we incorporated a color verification step following detection. Specifically, we extracted the region of interest (ROI) from the detection bounding box and applied Contrast Limited Adaptive Histogram Equalization (CLAHE) to the L-channel in LAB color space to enhance contrast and light and reduce noise. The image was then converted to HSV color space, and an orange mask was generated using predefined HSV thresholds. Morphological operations were then used to refine the mask by eliminating noise and small artifacts. Finally, we calculated the ratio of the orange mask area to the ROI to verify whether the detected object was an orange bucket.

B. *Bucket Detection Models Trained on Custom Datasets*

In cluttered and dark manufacturing-like environments, relying solely on the YOLO11n [21] pretrained model and color verification proved unreliable, see Fig. 4 (a). To address this limitation, we trained custom bucket detection models using images of our target buckets captured in target environments. Three datasets were collected for comparison: (1) 85 images captured using an iPhone14, (2) 213 images from Spot's onboard cameras, and (3) the combination of (1) and (2). Each dataset was split into training, validation, and testing sets with a $0.7 : 0.15 : 0.15$ ratio. We then used YOLO11n [21] model to train custom bucket detection models with the same training settings and hyperparameters on each dataset. After detecting the bucket, we evaluated the detection confidence. If the confidence exceeds a defined threshold, the bucket's position in the world frame is computed using camera models and coordinate frame transformations based on the detection's pixel position. Spot then navigates to the front of the bucket based on its world position.

C. *Results*

Table I and Fig. 4 (b)-(d) show the performance of bucket detection models trained on the three datasets. All models achieved high precision (P), recall (R), mean average precision at 50% IoU (mAP50), and mean average precision across IoU thresholds from 50% to 95% (mAP50-95). However, this good performance can largely be attributed to the controlled nature of the datasets, each containing only a single object class captured in consistent environments. As a result, the models are robust only within conditions similar to the datasets during training and may not be reliable in different scenarios. This limitation is evident in Fig. 4, where the model trained on the iPhone images fails to reliably detect the bucket when tested using Spot's onboard cameras.

TABLE I
PERFORMANCE OF CUSTOM BUCKET DETECTION MODELS

Dataset	P	R	mAP50	mAP50-95
85 images from an iPhone14	0.988	1.000	0.995	0.881
213 images from Spot	0.998	1.000	0.995	0.842
298 images from an iPhone14 and Spot	0.977	0.993	0.994	0.853

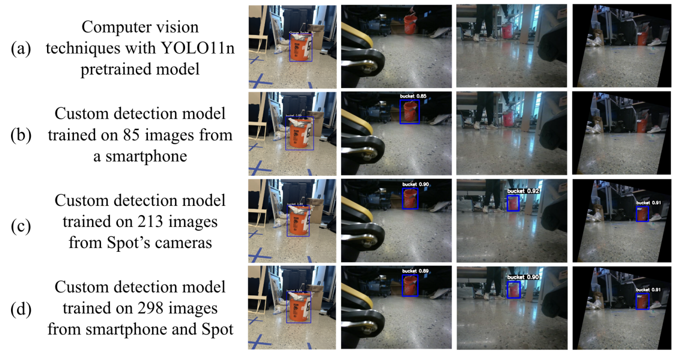

Fig. 4. Comparison of bucket detection results using different methods. The first column shows results tested on images from a smartphone, while columns 2 and 3 show results tested on images from Spot.

V. CONCLUSIONS AND FUTURE WORK

Experimental results show an average deviation of 3.8 cm for the empty bucket and 3.2 cm for the 5 lbs (2.27 kg) bucket. Deviations increase as the arm nears its workspace limit. To reduce collision risks, we placed the payload on the robot's body, and the robot transported it with high success. However, stability issues arise with payloads over 7.5 lbs (3.40 kg), failing at 17.5 lbs (7.94 kg). Overall, the robot is reliable for pick-and-place tasks but becomes unstable with heavier payloads, performing best within the central coordinate frame.

In the computer vision experiments, we demonstrated that utilizing a detection model trained on a custom dataset can reliably detect the target object in the environment, and Spot can autonomously navigate the target object using the detection results and perform the subsequent manipulation tasks. For others looking to apply this approach to customized needs, they should ensure that the custom dataset is collected using Spot's camera in the target environment for more reliable detection results. Additionally, in our current experiment, we assumed that the target object was within the robot's field of view. However, in real-world applications, a key research direction is exploring how Spot can autonomously and safely navigate narrow spaces with human and locate a target object outside its initial field of view in real-world environments.

Future work presents several research opportunities and challenges. For example, using the feeder test as a case study, in addition to identifying and moving liquid-filled buckets, the robot should also be able to open and close ball valves. Another challenge is ensuring the safety of quadruped robots with arms performing autonomous inspections in environments with narrow spaces, human workers, and hazardous machinery. Safety is crucial to prevent accidents, equipment damage, and injuries. A key issue in safe navigation is the robot's environment perception accuracy, affected by its oscillating body, changing posture, and non-smooth movement dynamics [23]. Additionally, we aim to explore human-robot interaction, such as human-robot handovers or response to human gestures.

References

[1] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*, pp. 3400–3407, 2011.

[2] Boston Dynamics, "The New Standard for Industrial Inspection," *Boston Dynamics*, 2024.

[3] A. Carter, "Néstle tech talk – peak into industrial revolution 4.0 by Nestlé Purina PetCare Company," in *Society of Women Engineers Annual Conference, WE24*, (Chicago), October 2024.

[4] G. K. Fischer, M. Bergau, D. A. Gómez-Rosal, A. Wachaja, J. Graeter, M. Odenweller, U. Piechottka, F. Höflinger, N. Gosala, N. Wetzel, *et al.*, "Evaluation of a smart mobile robotic system for industrial plant inspection and supervision," *IEEE Sensors Journal*, 2024.

[5] Boston Dynamics, "Spot at Michelin," *Boston Dynamics*, 2024.

[6] Boston Dynamics, "Spot Becomes Part of the Team at National Grid," *Boston Dynamics*, 2024.

[7] B. Hamner, S. Singh, S. Koterba, and R. Simmons, "An autonomous mobile manipulator for assembly tasks," *Autonomous Robots*, vol. 28, pp. 131–149, 2009.

[8] Boston Dynamics, "Spot - The Agile Mobile Robot," *Boston Dynamics*, 2024.

[9] Boston Dynamics, "Spot Arm Specifications and Key Concepts," *Boston Dynamics*, 2024.

[10] S. Halder and K. Afsari, "Robots in inspection and monitoring of buildings and infrastructure: A systematic review," *Applied Sciences*, vol. 13, no. 4, p. 2304, 2023.

[11] X. Hu, F. He, P. Xiao, T. Wang, D. Zhang, X. Zhou, and Y. Fan, "Design of a quadruped inspection robot used in substation," in *2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, vol. 4, pp. 766–769, IEEE, 2021.

[12] S. Yang, J. Liu, J. Meng, B. Zhang, Z. Sun, P. Xiao, and X. Dong, "Design of mechanical and hardware solutions for quadruped robot chassis based on substation inspection," in *2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC)*, vol. 7, pp. 1826–1829, IEEE, 2023.

[13] M. Ramezani, M. Brandao, B. Casseau, I. Havoutis, and M. Fallon, "Legged robots for autonomous inspection and monitoring of offshore assets," in *Offshore Technology Conference*, p. D011S006R005, OTC, 2020.

[14] M. Schneier, M. Schneier, and R. Bostelman, "Literature review of mobile robots for manufacturing," *Journal of Automation and Robotics*, 2015.

[15] R. Tomorrow, "Automating industrial inspection with autonomous mobile robots," *Robotics Tomorrow*, 2022.

[16] C. Gehring, P. Fankhauser, L. Isler, R. Diethelm, S. Bachmann, M. Potz, L. Gerstenberg, and M. Hutter, "Anymal in the field: Solving industrial inspection of an offshore hvdc platform with a quadrupedal robot," in *Field and Service Robotics: Results of the 12th International Conference*, pp. 247–260, Singapore: Springer Singapore, 2021.

[17] G. Jing, X. Qin, H. Wang, and C. Deng, "Developments, challenges, and perspectives of railway inspection robots," *Automation in construction*, vol. 138, p. 104242, 2022.

[18] E. Pearson, B. Mirisola, C. Murphy, C. Huang, C. O'Leary, F. Wong, J. Meyerson, L. Bonfim, M. Zecca, N. Spina, *et al.*, "Robust autonomous mobile manipulation for substation inspection," *Journal of Mechanisms and Robotics*, vol. 16, no. 11, p. 115001, 2024.

[19] Boston Dynamics, "OPG: Automating Circuit Breaker Tripping and Racking," *Boston Dynamics*, 2023.

[20] V. Lepetit and P. Fua, "Monocular model-based 3d tracking of rigid objects," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, IEEE, 2006.

[21] G. Jocher and J. Qiu, "Ultralytics YOLO11," 2024.

[22] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Doll'a r, and C. L. Zitnick, "Microsoft COCO: common objects in context," *CoRR*, vol. abs/1405.0312, 2014.

[23] G. Chen and L. Hong, "Research on environment perception system of quadruped robots based on lidar and vision," *Drones*, vol. 7, no. 5, p. 329, 2023.