Towards Autonomous Bricklaying Using Boston Dynamics Spot

Karl-Johan Sørensen[1*], Nico vom Hofe[2*] and Frank H.P. Fitzek[2,3]

[1]Department of Architecture, Massachusetts Institute of Technology, United States
[2]Deutsche Telekom Chair of Communication Networks, TU Dresden, Germany
[3]Centre for Tactile Internet with Human-in-the-Loop (CeTI), Dresden, Germany
∗authors contributed equally

karljo@mit.edu, nico.vom_hofe@tu-dresden.de, frank.fitzek@tu-dresden.de

Abstract -

Mobile legged robots offer unique advantages for on-site construction, including the ability to traverse unstructured terrain, climb stairs, and operate independently between spatially separated work zones. This paper presents a proof-of-concept system for autonomous brick stacking using a Boston Dynamics Spot quadrupedal robot equipped with a manipulator arm. The system operates between a supply area and an assembly area, using AprilTag fiducials for localization and autowalk missions for navigation. Target brick positions and orientations are specified in CAD software and exported directly for robotic execution. We evaluate placement accuracy through ten repeated assemblies of a curved four-layer wall comprising 14 wooden bricks. Results show mean translation errors below 10 mm and mean yaw errors below 1.5°, while the overall RMSE across all runs is 27.72 mm (translation) and 2.06° (rotation). The system was also tested outdoors with stairs between the two areas. The workflow is modular and easily extendable to different wall geometries.

Keywords -

Mobile robotic fabrication; legged mobile manipulation; visual fiducial localization; on-site construction; pick-and-place

1 Introduction

The construction industry is undergoing a transition towards greater automation, driven by rising housing demand, skilled labor shortages, and persistent safety concerns on construction sites. Yet, compared to sectors such as automotive and logistics, construction has seen limited productivity gains, with annual improvement rates of only 1 % against an industry average of 2.8 % [1]. On-site robotic automation has been proposed as one means to address these challenges [2]. Despite this potential, the deployment of robotic systems on construction sites remains limited [3]. Existing approaches face significant constraints. Large gantry systems require infrastructure that often exceeds the scale of the structure being built. Wheeled and tracked mobile platforms struggle with rough terrain, stairs, and cluttered site conditions. These constraints have confined most successful automation efforts to off-site prefabrication; yet construction fundamentally remains an on-site process where varying conditions resist complete pre-planning. Recent advances in commercial legged robotics have produced platforms capable of traversing stairs, uneven terrain, and human-centric spaces [4]. This makes them suited to the multi-level, cluttered environments of real construction sites, where wheeled and tracked platforms face fundamental access constraints. Prior work has demonstrated legged manipulation for general pick-and-place tasks [5, 6], but application to construction assembly has not been explored. To the best of our knowledge, no previous study has quantitatively evaluated the accuracy with which a legged robot can execute a brick-stacking task.

This paper presents a proof-of-concept for what we term the Mobile Agile Robotic Assembly (MARA) workflow: a distributed construction process in which a legged mobile manipulator autonomously traverses between spatially separated supply and assembly areas, picking and placing bricks according to a specified layout. The approach is evaluated using a Boston Dynamics Spot quadrupedal robot equipped with a manipulator arm. Target brick positions are modeled in Rhino, a Computer-Aided Design (CAD) program, and exported directly for robotic assembly. The system localizes by referencing each manipulation operation to a locally visible fiducial marker, thereby decoupling placement accuracy from global odometry drift. Placement accuracy is evaluated through ten repeated assemblies of a curved four-layer wall, reporting translational and rotational errors across layers.

2 Background

The concept of robotic masonry has evolved from early mechanized systems to the dedicated commercial bricklaying systems of today. Early mobile systems such as the BRONCO demonstrated the first robotic bricklaying on a mobile platform but required pre-surveyed sites, remote control, and continuous operator supervision [7]. Later, feasibility studies formalized the relationship between laying patterns, manipulator trajectories, and achievable precision through planning and optimization [8]. Although commercial systems such as the Semi-Automated Mason (SAM100) demonstrate viable high-throughput automation, they remain optimized for structured workspaces and require strict workflow integration and supervision by hu-

man crews [9]. These systems also share a spatial constraint: materials must be staged within or immediately adjacent to the robot's workspace. Deploying these systems in real-site contexts with environmental constraints highlights the gap between laboratory robotics and construction realities [10]. Competition-driven research further validates the autonomy necessary to assemble wall patterns with differently sized and colored bricks within time constraints [11, 12, 13]. Among these, Lenz et al. [13] provide the most detailed quantitative evaluation, reporting end-to-end translational placement errors of 13–19 mm and rotational errors below 1.3° using a wheeled omnidirectional platform with a UR10e arm. However, all of these systems depend on wheeled or tracked mobility, restricting their ability to navigate rough terrain, stairs, and cluttered environments. Additionally, despite their mobility, these platforms require close proximity to material supplies or human collaborators for resupply.

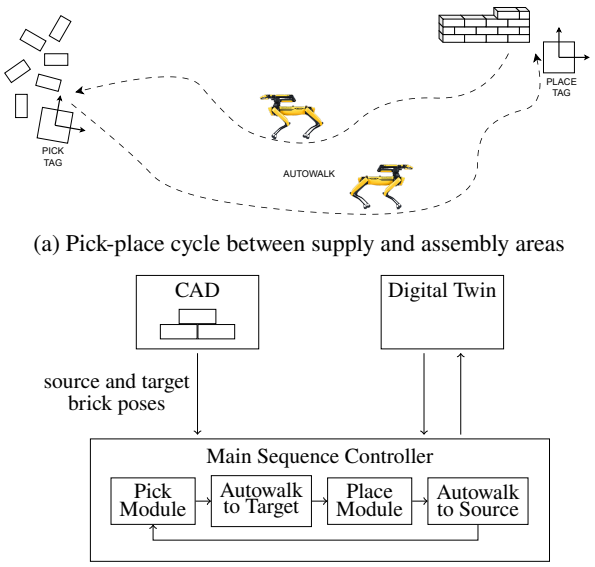
Legged robotic platforms can traverse stairs, uneven terrain, and constrained spaces, enabling autonomous operation between spatially decoupled supply and assembly zones without the site layout constraints imposed by material co-location. Early work has explored small-scale humanoid robots for on-site floor tiling, demonstrating perception-driven placement with a NAO biped [14]. More recent work on quadruped-mounted manipulators demonstrates pick-and-place pipelines integrating object detection with real-time control [5]. Learning-based approaches show that multi-stage, long-horizon pick-and-place tasks on legged mobile manipulators can be made robust through sim-to-real transfer [6]. These advances demonstrate the growing capability of legged manipulation; however, application to masonry assembly remains unexplored.

This paper addresses that gap, combining two domains that, to the best of our knowledge, have not previously been merged in this context: legged manipulation [5, 6] and mobile construction robotics [3, 15, 16, 17, 12]. We present the first demonstration of autonomous brick stacking on a legged mobile manipulator and provide the first quantitative evaluation of placement accuracy for a legged robot performing construction assembly. Our proposed system introduces a CAD-to-execution pipeline integrated with a legged manipulator that can operate across spatially decoupled supply and assembly areas connected by stairways, which are inaccessible to wheeled or tracked platforms.

3 Methodology

3.1 System overview

Our autonomous on-site construction system utilizes a *Boston Dynamics* Spot robot with a six-Degree of Freedom (DoF) manipulator arm, offering a reach of 984 mm at full extension and a lifting capacity of 11 kg [18]. This robot was chosen for its off-the-shelf availability, which makes deployment significantly more feasible than developing and certifying a custom platform. The system comprises two distinct areas: a supply area containing bricks and an assembly area where the target structure is constructed. Each area is defined by an AprilTag [19] fiducial from the *Tag36h11* set, as illustrated in Figure 1a. Specifically, the *PICK-TAG* is situated in the supply area, while the *PLACE-TAG* is located within the assembly area. Unlike systems that depend on active positioning infrastructure such as total stations, UWB, or long-range RFID, which require external power and increase installation effort, the localization system relies entirely on passive AprilTags that need neither a power supply nor geometric pre-calibration.



(a) Pick-place cycle between supply and assembly areas

(b) Software pipeline from CAD to robot execution

Figure 1. System overview

All brick poses are specified relative to these fiducial frames, providing robustness against cumulative drift in odometry. In the context of commanding arm motions, poses are transformed from the fiducial frame to Spot's vision frame. The latter is a world-fixed frame maintained by the robot's visual localization system [20]. By referencing each operation to the fiducial, this approach decouples placement accuracy from global localization drift.

The target brick positions are 3D modeled in the CAD software, Rhino. A custom script was developed to extract position and orientation data. The script then exports JSON configuration files that specify the poses of the bricks relative to the fiducials. This process establishes a direct connection between architectural design and robotic execution. Additionally, a human operator recorded two

autowalk missions for the supply-to-assembly and return journeys using *Boston Dynamics'* autonomous navigation system.

The process of repeating sequences for each brick configuration is illustrated in the software pipeline in Figure 1b. During execution, the system iterates through the configuration file and executes a four-step cycle for each brick: (1) pick a brick from the supply area, (2) walk to the assembly area, (3) place the brick at the target pose, and (4) return to the supply area. The digital twin component enables a remote operator to monitor the robot's state and perception in real time (Section 3.4).

The system was developed using Python and the *Boston Dynamics* Software Development Kit (SDK). The source code and experimental data are publicly available at `https://github.com/soerensenkarl/MARA-bricklaying`.

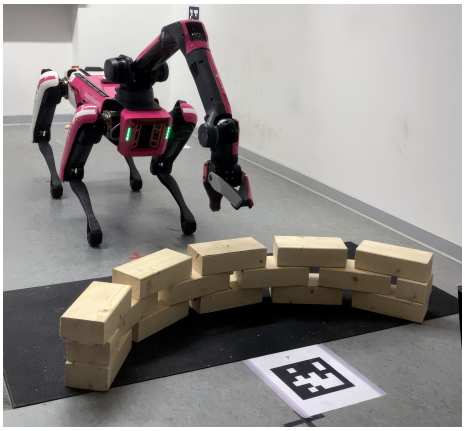


Figure 2. Spot after placing a brick on a curved wall

3.2 Manipulation

A common positioning strategy is employed in both pick and place operations. The robot first detects the relevant fiducial and moves its base to a position 0.8 m behind the target brick along the fiducial's x-axis, aligning its heading with the positive x-direction of the fiducial. This pre-positioning ensures the arm has sufficient reach to access the brick and adequate visibility of the fiducial markers. After reaching this position, the system reacquires the fiducial transform to establish a stable reference for arm commands.

The pick sequence proceeds as follows: The gripper opens and the arm approaches the brick, moving the hand frame to 20 cm above the target position specified in the configuration file. It then descends to 3.5 cm above the target, pauses briefly, closes the gripper to secure the brick, and lifts to a transport height of 80 cm. The gripper maintains a pitch angle of 95° (nearly vertical, pointing slightly backward) during picking. This angle was empirically determined to provide a reliable grasp.

The place sequence inverts this operation. The arm approaches from 30 cm above the target position, descends to 4.5 cm above the surface, pauses, opens the gripper to release the brick, retracts, and stows. During placement, the end-effector pitch is set to 90° (vertical) and the yaw is set according to the target orientation specified in the configuration file, enabling non-axis-aligned placements for curved geometries. A z-offset of −1.5 cm compensates for the distance from the hand frame to the center of the gripped brick. Hip height assist is enabled throughout both sequences to permit base height adjustments.

3.3 Navigation

The robot autonomously navigates between pick and place operations, executing pre-recorded autowalk missions to traverse the supply and assembly areas. Each autowalk consists of a graph of waypoints and edges, along with visual snapshots that were captured during the recording process. During operation, the module uploads the graph to Spot's GraphNav system and begins localization by searching for the nearest visible AprilTag. During navigation, GraphNav continuously localizes the robot by combining odometry with visual and geometric features and by comparing real-time sensor data against stored map snapshots [21]. The robot then autonomously navigates along the recorded path while avoiding obstacles.

3.4 Remote supervision

A digital twin component built on Robot Operating System (ROS) 2 Humble enables real-time remote supervision of the robot during operation. The wrapper streams the robot's internal state and visual perception to a remote operator via RViz. The streamed state includes the battery level and the position of each joint, allowing the robot's full pose to be reconstructed at each timestamp. The robot's visual perception of its environment is rendered as six point cloud streams from the five body-mounted RGB-D cameras and the gripper camera. This provides an operator with an intuitive, egocentric view of the robot's surroundings, as opposed to raw camera feeds. Environmental observations can be gradually integrated into a single 3D representation of the site using the mapping framework described in [22]. The digital twin achieves a mean frame rate of 3.5 Hz for the RGB-D streams and approximately 50 Hz for joint state updates. Each camera transmits a time-synchronized pair of RGB and depth ROS messages, with mean sizes of 921.7 kB and 614.46 kB, respectively.

3.5 Experimental setup

The experimental site comprised a supply area and an assembly area, separated by a distance of approximately 4 m. Fiducials were placed in opposite corners of the room. Figure 3 shows the experimental room-scale setup, including the three stages of picking up, walking with, and placing bricks. To avoid potential damage to the robot, wooden blocks corresponding to the standard brick proportions (200 mm × 100 mm × 60 mm) were used instead of actual bricks. These wooden bricks have a mean weight of 550 g.

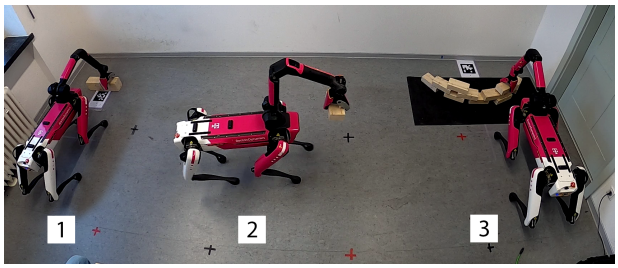


Figure 3. Superimposed video frames from the experiment showing Spot at pick position (1), autowalking (2), and at place position (3)

In the supply area, the bricks were arranged in two vertical stacks of seven bricks each. The target structure was a curved wall with four layers and a total of 14 bricks.

There were five bricks in the base layer and fewer in each subsequent layer as the wall tapered toward the top. A non-slip mat was positioned beneath the base layer to mitigate the risk of slippage during the process of placing the upper layers.

4 Results

The evaluation of placement accuracy was conducted by selecting planar position error (x and y) and yaw error as the primary metrics. Errors in the remaining pose dimensions (z, roll, and pitch) would indicate an unsuccessful build and were not analyzed. Any structural collapse resulting from substantial placement inaccuracies was designated as a failure. All ten experimental runs completed successfully without failure.

To record brick positions, a downward-facing camera was mounted at a height of approximately 1.33 m above the *PLACE-TAG*. Reproducing the camera position and parameters in a CAD model enabled the generation of ground-truth images for each layer of the target wall. A layer-wise image overlay between recorded images and ground truth was used to calculate the planar center-point offset between each placed brick and its target position. The same overlay was used to extract brick contours, enabling measurement of the angular deviation between desired and actual orientations, recorded as yaw error.

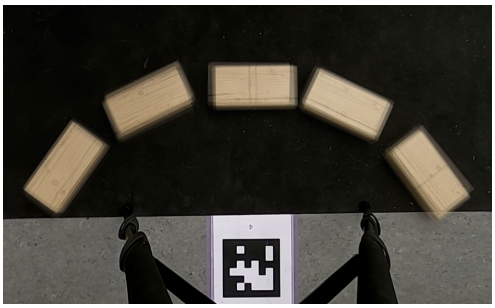


Figure 4. Superimposed overlay of all ten first-layer placements, showing variability across runs

The evaluation comprised ten fully autonomous constructions of the same structure. Figure 5 presents the resulting pose errors, grouped by construction layer. Error distributions are visualized using boxplots based on quartiles (Q_1, Q_2, Q_3) and the Interquartile Range (IQR), defined as $Q_3 - Q_1$. The whiskers extend to the most extreme data points within $1.5 \times$ IQR from the first and third quartiles. The single points represent maximum values that exceed the upper fence and constitute actual data values that surpass the established threshold. Consequently, these points are categorized as outliers.

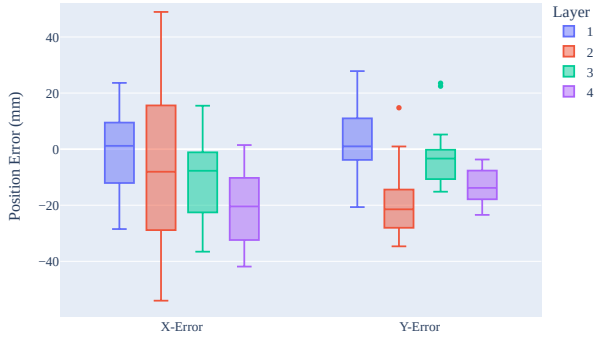
The initial layer, comprising five bricks, exhibits the highest number of measurements and the widest range of target yaw angles (107.8°). This range is necessary to establish the curved base of the wall. As shown in Figure 5a, this layer achieves the smallest median planar position errors, measuring 1.18 mm in the x direction and 1.01 mm in the y direction. Despite the large variation in the target yaw angles, Figure 5b shows a low median yaw error of 0.6° for the initial layer.

The second layer demonstrates a higher prevalence of pose errors compared to the base layer, showing median position offsets in the negative x and y directions and the most significant median yaw error among all layers, measuring at 3.32°. The third and fourth layers also exhibit slight negative planar offsets; however, their median yaw errors remain below 0.9°.

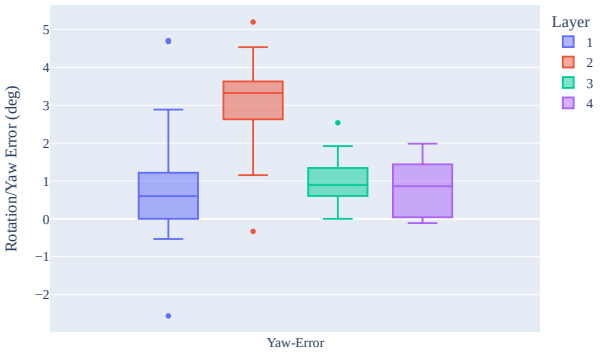
$$\mathrm{APE}_i = \left\| \widehat{\mathbf{p}}_i - \mathbf{p}_i \right\| \tag{1}$$

Equation 1 defines the metric used to evaluate the Absolute Pose Error (APE) of each brick. Aggregating all brick poses across the ten runs, the resulting Root Mean Squared Error (RMSE) is 27.72 mm for the translational component and 2.06° for the rotational component.

Several factors contribute to the observed error characteristics. The low error in the first layer can be attributed to the anti-slip mat, which increases friction at the ground



(a) Position error for each layer in x and y direction



(b) Yaw error for each layer

Figure 5. Error metrics for the individual layers

level. Friction is much lower in higher layers due to the smooth surfaces of the wooden bricks. Furthermore, Spot is designed for field mobility rather than factory-grade motion repeatability, contributing to inherent positioning variability. The apparent increase in errors at higher layers may also result from greater distance between the gripper and the visual fiducial anchor point, as the fiducial is fixed at ground level while the placement height increases with each layer.

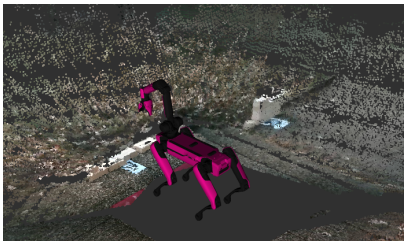
Execution times were recorded by logging timestamps at the start and end of each operation during the assembly sequence. Table 1 summarizes the mean durations for each phase of the pick-place cycle.

Table 1. Mean cycle times for each operation phase

Operation	Duration (s)
Pick	9.18
Walk to assembly area	15.46
Place	9.68
Walk to supply area	16.15
Total cycle	50.47

The mean cycle time of 50.5 s per brick results in a total assembly time of approximately 12 minutes for the 14-brick structure with a supply-to-assembly distance of 4 m. Navigation accounts for 63% of the total cycle time. These times reflect conservative speed parameters set during development; timeouts, motion durations, and arm trajectories were not optimized for speed, leaving considerable potential for efficiency improvements.

(a) Superimposed frames from outdoor brick-stacking video



(b) Digital Twin - Real-Time supervision using the ROS2 Wrapper and visualization in RViz

Figure 6. Autonomous construction scenario in an outdoor setting including stairways

To contextualize these results, the most directly comparable system is the NimbRo UGV [13], a wheeled platform with a UR10e arm that reported mean translational placement errors of 13–19 mm and rotational errors of 0.45–1.26° in autonomous brick-placement trials. Our mean translational errors below 10 mm are comparable to these figures, while the overall RMSE of 27.72 mm reflects sensitivity to occasional outliers at higher layers. The *In situ Fabricator* [16, 17], a 1.4 t tracked platform with an industrial arm offering 2.55 m reach and 40 kg payload, achieved a maximum assembly error of 7 mm over 1600 bricks with substantially different hardware. Other mobile manipulators report grasping accuracies of ±30–50 mm [12, 11] but do not quantify final placement error. No prior work reports placement accuracy for a legged robot in a con-

struction assembly context; the ten repeated assemblies therefore serve as both an absolute accuracy evaluation and an internal repeatability baseline. The manipulation-only cycle time (pick and place, ≈18.9 s) is comparable to other mobile manipulators, while the total cycle time of 50.5 s is dominated by navigation between spatially decoupled work zones, a phase absent in systems where materials are staged within the robot's workspace.

The indoor experiments evaluate placement accuracy under controlled conditions, but the primary motivation for a legged platform is access to terrain that wheeled and tracked systems cannot traverse. To validate this, a second experimental setup was created outdoors with the supply and assembly areas at different elevations, connected by stairways with climbing steps, as shown in Figure 6a. This configuration demonstrates operation in an environment where wheeled robots would fail. The robot autonomously descended stairs to pick bricks and climbed back to the assembly area to place them. Although individual pose errors were not measured for this setup, repeated runs successfully stacked a planar wall. A continuous video of the outdoor assembly is available at `https://youtu.be/p-ExWFgYa-I`.

For remote supervision, the robot's state and perception were continuously streamed to a remote operator via the ROS 2 digital twin (Section 3.4), as shown in Figure 6b.

5 Conclusion and future work

This paper demonstrates a proof-of-concept for autonomous bricklaying using a legged mobile robot. We introduce a MARA workflow deploying a mobile quadrupedal robot with an integrated manipulator arm. This allows the robot to navigate autonomously between a supply area and an assembly area while picking and placing bricks. Unlike stationary or wheeled industrial manipulators, the mobile robot can traverse unstructured environments and climb stairs, making it well-suited for on-site construction.

During our evaluation of the robot's accuracy, it autonomously constructed a curved four-layer wall. It successfully completed all ten experimental runs without human intervention. The system achieved mean translation errors below 10 mm and mean yaw errors below 1.5°, demonstrating the viability of legged mobile manipulation for construction tasks.

While the current system focuses on a well-defined operational setup, several aspects provide opportunities for further extension. The use of a single AprilTag fiducial per work area enables reliable localization within the tested scale. Extending to larger structures would require a multi-fiducial configuration. A quantitative evaluation was performed in a controlled indoor environment to ensure reproducibility. This evaluation was complemented by outdoor experiments that demonstrated qualitative feasibility. The current implementation uses dry stacking to provide a simplified baseline for the assembly of the structure. The system design does not preclude the integration of mortar or other bonding materials. With the arm installed, the robot retains a remaining body payload capacity of approximately 6 kg [18], which may accommodate a reservoir for bonding agents, though this has not been tested. Validating this extension with clay bricks and mortar application is a priority for future work.

The system uses a general-purpose, off-the-shelf robotic platform, which offers advantages in terms of availability, robustness, and flexibility. While task-specific design modifications could improve efficiency and practicality in construction scenarios, this level of specialization would increase development and production costs, particularly given the limited use of construction robots. Currently, navigation between supply and assembly areas relies on pre-recorded autowalk missions. Replacing these with fully autonomous path planning remains an open challenge. Additionally, developing recovery strategies for failed grasps or misplacements is a future area of development.

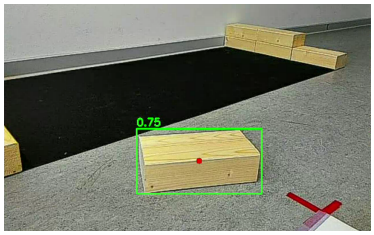


Figure 7. Brick detection model detecting a brick with 75 % confidence; showing the bounding box and center point

Gripper and body cameras could in principle provide visual feedback for alignment verification. Computer vision-based object detection offers additional possibilities. Preliminary tests with an ATSS-MobileNetV2 model demonstrate brick identification in camera images (Figure 7). This suggests potential for adaptive component recognition and placement verification. This approach would generate a pick vector transformable from the camera frame to the gripper frame, potentially enabling grasping of components in unpredictable positions. Additionally, the computer vision model could be trained to identify construction tools to perform inventory management and tool tracking on construction sites.

Scaling to larger structures would require a multi-fiducial localization system or an alternative approach. Rather than increasing the effort required for placing fiducials, one option is to replace fiducial-based localization with visual Simultaneous Localization and Map-

ping (SLAM), though this introduces its own challenges in dynamic construction environments. Such a framework would form the basis for environmental modeling necessary for robust navigation. This would allow the robot to localize using a sparse map representation of the environment [22]. This would reduce or eliminate the need for pre-recorded navigation, though autonomous path planning in active construction sites remains an open problem.

The system provides an initial step towards real-world masonry construction, with validation through clay bricks and mortar application as a direct extension. Further autonomous validation scenarios should include realistic construction environments featuring harsh terrain such as gravel and hard-to-reach areas. To support the MARA workflow capabilities, diverse wall designs should be evaluated to test the robustness of the CAD-to-execution pipeline.

Multi-robot deployment could increase construction efficiency, supported by a complete ROS 2 Humble-based system enabling collaborative map sharing for tasks such as site monitoring and inspection [23]. Efficiency could also be improved by stowing multiple bricks on the robot's back, utilizing its remaining 6 kg payload capacity.

The ability to traverse stairs may extend to movement on temporary scaffolding, which would decouple build height from the manipulator's reach. Possible further directions include real-time human-robot collaboration with live design updates, as-built verification against the target design model, and autonomous off-hours preparation of tools and materials.

The system reliably completed all experimental runs and demonstrated robust stair traversal between work zones. The overall performance highlights the feasibility of legged mobile manipulation for construction tasks, with occasional errors in the range of centimeters. These results suggest that such systems are a viable approach to automating on-site construction in unstructured, human-centric environments, where traditional robotic platforms have limitations.

Acknowledgement

Funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) as part of Germany's Excellence Strategy – EXC 2050/2 – Project ID 390696704 – Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of TUD Dresden University of Technology. The authors would like to thank the Federal Ministry of Research, Technology, and Space (BMFTR) for its support as part of the research program Communication Systems "Souverän. Digital. Vernetzt.". Joint project 6G-life, project identification number: 16KIS2413K

References

[1] Filipe Barbosa, Jonathan Woetzel, Jan Mischke, Maria João Ribeirinho, Mukund Sridhar, Matthew Parsons, Nick Bertram, and Stephanie Brown. Reinventing construction: A route to higher productivity. Technical report, McKinsey Global Institute, 2017. URL https://www.mckinsey.com/capabilities/operations/our-insights/reinventing-construction-through-a-productivity-revolution.

[2] Nathan Melenbrink, Justin Werfel, and Achim Menges. On-site autonomous construction robots: Towards unsupervised building. *Automation in Construction*, 119:103312, 2020. ISSN 0926-5805. doi:10.1016/j.autcon.2020.103312.

[3] Lingdong Zeng, Shuai Guo, Jing Wu, and Bernd Markert. Autonomous mobile construction robots in built environment: A comprehensive review. *Developments in the Built Environment*, 19:100484, 2024. ISSN 2666-1659. doi:10.1016/j.dibe.2024.100484.

[4] Boston Dynamics. Spot. https://bostondynamics.com/products/spot/. Accessed: 2025-11-29.

[5] Abhishek Jagadeesh Kasaragod. Dynamic Pick-and-Place System for a Manipulator on a Quadruped Using Object Detection. 12 2024. doi:10.25417/uic.29048471.v1.

[6] Haichao Zhang, Haonan Yu, Le Zhao, Andrew Choi, Qinxun Bai, Yiqing Yang, and Wei Xu. Learning multi-stage pick-and-place with a legged mobile manipulator. *IEEE Robotics and Automation Letters*, 10(11):11419–11426, 2025. doi:10.1109/LRA.2025.3608425.

[7] G. Pritschow, M. Dalacker, J. Kurz, and M. Gaenssle. Technological aspects in the development of a mobile bricklaying robot. *Automation in Construction*, 5(1): 3–13, 1996. ISSN 0926-5805. doi:10.1016/0926-5805(95)00015-1. 12th ISARC.

[8] Seung-Nam Yu, Byung-Gab Ryu, Sung-Jin Lim, Chang-Jun Kim, Maing-Kyu Kang, and Chang-Soo Han. Feasibility verification of brick-laying robot using manipulation trajectory and the laying pattern optimization. *Automation in Construction*, 18(5):644–655, 2009. ISSN 09265805. doi:10.1016/j.autcon.2008.12.008.

[9] Andrew Madsen. The SAM100: Analyzing labor productivity. *Construction Management*, 2019. URL https://digitalcommons.calpoly.edu/cmsp/243.

[10] Weiguo Xu, Dan Luo, and Yuan Gao. Automatic brick masonry system and its application in on-site construction. In *4th International Conference on Computer-Aided Architectural Design Research: Intelligent and Informed,(CAADRIA)*, pages 83–92, 2019. doi:10.52842/conf.caadria.2019.1.083.

[11] Meysam Basiri, João Gonçalves, José Rosa, Alberto Vale, and Pedro Lima. An autonomous mobile manipulator to build outdoor structures consisting of heterogeneous brick patterns. *SN Applied Sciences*, 3(5):558, 2021. ISSN 2523-3971. doi:10.1007/s42452-021-04506-7.

[12] Petr Štibinger, George Broughton, Filip Majer, Zdeněk Rozsypálek, Anthony Wang, Kshitij Jindal, Alex Zhou, Dinesh Thakur, Giuseppe Loianno, Tomáš Krajník, and Martin Saska. Mobile manipulator for autonomous localization, grasping and precise placement of construction material in a semi-structured environment. *IEEE Robotics and Automation Letters*, 6(2):2595–2602, 2021. doi:10.1109/LRA.2021.3061377.

[13] Christian Lenz, Jan Quenzel, Arul Selvam Periyasamy, Jan Razlaw, Andre Rochow, Malte Splietker, Michael Schreiber, Max Schwarz, Finn Süberkrüb, and Sven Behnke. Autonomous wall-building and firefighting: Team NimbRo's UGV solution for MBZIRC 2020. *Field Robotics*, 2:55–74, 2022. doi:10.55417/fr.2022003.

[14] De Hong Jung, Jaeheung Park, and Mathew Schwartz. Towards on-site autonomous robotic floor tiling of mosaics. In *2014 14th International Conference on Control, Automation and Systems (ICCAS 2014)*, pages 59–63, 2014. doi:10.1109/ICCAS.2014.6987959.

[15] Volker Helm, Selen Ercan, Fabio Gramazio, and Matthias Kohler. Mobile robotic fabrication on construction sites: Dimrob. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4335–4341, 2012. doi:10.1109/IROS.2012.6385617.

[16] Kathrin Dörfler, Timothy Sandy, Markus Giftthaler, Fabio Gramazio, Matthias Kohler, and Jonas Buchli. Mobile robotic brickwork. In *Robotic fabrication in architecture, art and design 2016*, pages 204–217. Springer, 2016. ISBN 978-3-319-26376-2. doi:10.1007/978-3-319-26378-6_15.

[17] Markus Giftthaler, Timothy Sandy, Kathrin Dörfler, Ian Brooks, Mark Buckingham, Gonzalo Rey, Matthias Kohler, Fabio Gramazio, and Jonas Buchli. Mobile robotic fabrication at 1:1 scale: the in situ fabricator. *Construction Robotics*, 1(1):3–14, 2017. ISSN 2509-8780. doi:10.1007/s41693-017-0003-5.

[18] Boston Dynamics. Spot arm. `https://bostondynamics.com/products/spot/arm/`, 2024. Accessed: 2026-01-25.

[19] Edwin Olson. Apriltag: A robust and flexible visual fiducial system. In *2011 IEEE International Conference on Robotics and Automation*, pages 3400–3407, 2011. doi:10.1109/ICRA.2011.5979561.

[20] Boston Dynamics. Geometry and frames. `https://dev.bostondynamics.com/docs/concepts/geometry_and_frames`, 2024. Accessed: 2026-01-25.

[21] Boston Dynamics. GraphNav localization. `https://dev.bostondynamics.com/docs/concepts/autonomy/localization`, 2024. Accessed: 2026-01-25.

[22] Nico Vom Hofe, Peter Sossalla, Johannes Hofer, Christian L. Vielhaus, Justus Rischke, Jannek Steinke, and Frank H. P. Fitzek. Demo: Robotics meets augmented reality: Real-time mapping with boston dynamics spot and microsoft hololens 2. In *2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 358–360, 2023. doi:10.1109/WoWMoM57956.2023.00062.

[23] Nico Vom Hofe, Pit Hofmann, Frederic Benken, Johannes Hofer, Pengjie Zhou, Ruifeng Zheng, Juan A. Cabrera, and Frank H. P. Fitzek. Sniffing out the path: Olfactory localization based low-cost slam for disaster scenarios. In *2024 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, 2024. doi:10.1109/GCWkshp64532.2024.11100702.