

Towards Learning Boulder Excavation with Hydraulic Excavators

Jonas Gruetter¹, Lorenzo Terenzi¹, Pascal Egli² and Marco Hutter¹

Abstract—Construction sites frequently require removing large rocks before excavation or grading can proceed. Human operators typically extract these boulders using only standard digging buckets, avoiding time-consuming tool changes to specialized grippers. This task demands manipulating irregular objects with unknown geometries in harsh outdoor environments where dust, variable lighting, and occlusions hinder perception. The excavator must adapt to varying soil resistance—dragging along hard-packed surfaces or penetrating soft ground—while coordinating multiple hydraulic joints to secure rocks using a shovel. Current autonomous excavation focuses on continuous media (soil, gravel) or uses specialized grippers with detailed geometric planning for discrete objects. These approaches either cannot handle large irregular rocks or require impractical tool changes that interrupt workflow. We train a reinforcement learning policy in simulation using rigid-body dynamics and analytical soil models. The policy processes sparse LiDAR points (just 20 per rock) from vision-based segmentation and proprioceptive feedback to control standard excavator buckets. The learned agent discovers different strategies based on soil resistance: dragging along the surface in hard soil and penetrating directly in soft conditions. Field tests on a 12-ton excavator achieved 70% success across varied rocks (0.4–0.7m) and soil types, compared to 83% for human operators. This demonstrates that standard construction equipment can learn complex manipulation despite sparse perception and challenging outdoor conditions.

Keywords: Construction Robotics, Field Robotics, Reinforcement Learning.

I. INTRODUCTION

Autonomous boulder removal remains a critical gap in construction robotics. During site preparation, operators extract rocks using standard digging buckets, relying on visual and tactile feedback to adapt to varying rock geometries and soil conditions.

This capability is essential for practical automation. While grippers offer better control, tool changes interrupt workflow and are impractical when the bucket is needed for subsequent excavation. Manipulating irregular objects with standard buckets would enable seamless task transitions.

Boulder extraction requires coordinating multiple joints to align, scoop, and secure irregularly shaped rocks while adapting to soil resistance—from hard surfaces where rocks slide to soft conditions requiring different penetration depths. Outdoor environments compound these challenges through dust, variable lighting, long-range sensing requirements (7 m to 10 m), and arm-induced occlusions.

¹The authors are with Robotic Systems Lab, ETH Zurich, 8092 Zurich, Switzerland.

²The author is with Gravis Robotics AG, 8050 Zurich, Switzerland.

Corresponding author: Jonas Gruetter
jonasgrutter@gmail.com

This work was supported by the NCCR Digital Fabrication (DFAB) and the Swiss National Science Foundation (SNF) through project No. 188596.

Recent work demonstrates autonomous excavation of continuous media [1], [2] and precise assembly with discrete objects using grippers [3], [4]. Learning-based methods adapt trajectories using hydraulic feedback [1], while classical approaches use specialized grippers to calculate stable grasp points on detailed 3D models of objects [5], [6].

These approaches fail for boulder extraction with standard tools. Classical methods need accurate 3D models and specialized end-effectors. Continuous media controllers lack reactive behaviors for individual object manipulation. No existing method handles the combination of irregular objects, standard buckets, and harsh outdoor conditions.

We present a reinforcement learning approach enabling autonomous boulder excavation using standard buckets and sparse perception—just 20 LiDAR points per rock.

Policies trained in simulation with rigid-body dynamics and analytical soil models learn adaptive strategies from sparse point clouds and proprioceptive feedback. Different behaviors emerge without explicit soil estimation: surface dragging in hard soil versus direct penetration in soft conditions.

Field deployment on a 12-ton excavator achieved 70% success across varied rocks (0.4–0.7m) and soil types, compared to 83% for human operators, demonstrating complex manipulation without specialized tools.

Our contributions are:

- 1) A reinforcement learning framework combining rigid-body simulation with analytical soil models to discover adaptive boulder pickup strategies using standard buckets.
- 2) A sparse perception pipeline using SAM2 segmentation with LiDAR, demonstrating boulder excavation with only 20 points per rock.
- 3) Field validation achieving 70% success rate compared to 83% for human operators across diverse conditions.

A. Related Work

The convergence of advanced perception systems and learning-based control has opened new possibilities for autonomous manipulation across scales. While dexterous manipulation has seen rapid progress through deep learning, transferring these advances to heavy machinery presents unique challenges in perception, control, and safety. This work bridges these domains, presenting the first demonstration of robust boulder manipulation on construction equipment using only a standard digging bucket, guided by reinforcement learning and sparse 3D perception.



Fig. 1: Example excavation scene with rocks of varying shapes, sizes, and surface conditions (target rocks highlighted in blue). Reliable detection and removal of such objects is critical for safe and efficient downstream tasks like terrain leveling or continued excavation.

1) Perception-based Learning for Manipulation:

Learning-based manipulation has progressed from RGB-based policies [7], [8] to 3D representations better suited for real-world robotics. Early approaches using pre-trained vision encoders like CLIP [9], [10] often fail in outdoor settings due to depth ambiguity and poor generalization. For construction machinery operating at long range, point clouds provide more reliable spatial information.

LiDAR is ideal for construction sites due to its robustness to lighting conditions and its ability to capture large-scale 3D structures. Macaro et al. [11] employ dual LiDARs on an excavator to enable rock segmentation via geometric clustering; however, performance degrades on uneven terrain where boundaries become ambiguous or individual rocks are over-segmented.

Point cloud processing has advanced from PointNet’s [12] unordered input handling to hierarchical models [13] and recent RL-tailored designs. PointPatchRL [14] introduces patch-based transformers, while ANYmal parkour [15] uses 4D convolutions on LiDAR data, favoring this modality over depth cameras for its superior range—a key advantage in our setup.

Many complex manipulation policies rely on demonstration data. Imitation learning approaches such as RVT-2 [16] and Act3D [17] employ large 3D encoders trained on expert rollouts, but collecting such data is often impractical for heavy machinery. Distillation-based methods [18], [19] also depend on high-quality visual inputs, which prove unreliable in outdoor environments due to dust, occlusion, and variable lighting conditions.

2) *Autonomous Heavy Machinery and Construction Robotics:* Construction automation has traditionally focused on continuous media tasks such as trenching [20], grading [21], and loading [22], [23]. More recently, efforts have extended to manipulating discrete, irregular objects. Classical approaches rely on gripper-based grasp

planning with force-closure guarantees [4] and detailed perception pipelines for precise placement, as shown for autonomous dry stone wall assembly [5], [6]. While effective in structured settings, these methods require accurate 3D models, hand-crafted grasp and placement strategies, and specialized end effectors.

In contrast, tasks like site preparation demand fast, robust boulder excavation using standard tools. Here, strategies must adapt to uncertain rock shapes and varying soil resistance, making reactive behavior more critical than geometric precision.

Learning-based methods address this challenge by enabling adaptive behavior to emerge from interaction, eliminating the need for explicit modeling or manual heuristics. For example, soil-adaptive excavation [1] demonstrated that reinforcement learning (RL) policies can adjust trajectories based on hydraulic pressure feedback, while autonomous bucket filling [2] showed that proprioceptive signals support material-agnostic adaptation.

Building on these insights and the perception advances discussed in Section I-A.1, the application of RL to point-cloud inputs proves crucial. For instance, recent work [24] trains RL policies using compact geometric point-cloud embeddings and lightweight MLPs for sequential excavation of cluttered rigid objects, thereby reducing both training time and policy complexity. Unlike Lu et al. [24], who focus on sequential excavation of small rigid objects (1-7cm) using 7-DOF robotic arms with rich geometric representations, our approach tackles boulder excavation—manipulating massive irregular rocks (0.4-0.7m diameter) on industrial-scale 12-ton excavators using extremely sparse perception (only 20 LiDAR points) and analytical soil interaction models for field deployment on construction equipment.

II. METHOD

We train the policy in simulation using a soil resistance model and sparse point cloud observations, then deploy it on a 12-ton hydraulic excavator. The full system—including perception, simulation, and control—is summarized in Fig. 2.

A. Simulation

The training environment for our RL agent was developed within NVIDIA Isaac Lab [25]. Isaac Lab is used to simulate the rigid-body dynamics of the excavator and its interaction with rocks and terrain, excluding soil forces which are modeled separately.

Excavator Model: The excavator model was simplified to include only the key actuated joints—base rotation, boom, stick, telescopic extension, and bucket pitch—to enable efficient simulation (Fig. 3). Remaining joints were fixed, and hydraulic dynamics were omitted. Hydraulic dynamics are handled by low-level PID controllers on the real hardware that track velocity commands, as detailed in Section III-B. The excavator stands on three static support bases to emulate a fixed, stable operating position with a friction coefficient of 0.8 [2].

Rocks: A dataset of 50 real rock geometries was generated

by 3D scanning physical rocks, capturing natural variation in shape and surface features. Each rock was assigned a nominal density of 2500 kg m^{-3} , with mass randomized by $\pm 10\%$ per episode. Dimensions were varied along the principal axes (x, y, z) within the ranges 0.3 m to 1.0 m, 0.1 m to 0.5 m, and 0.2 m to 0.7 m, respectively. Friction coefficients were sampled uniformly from 0.35 to 0.6, based on reported values for dry rock–soil interfaces. Rock instances were randomized across parallel environments to promote generalization.

Environment Setup: Rocks are instantiated on a flat, horizontal support surface. Collision properties are defined such that the excavator’s bucket can collide with the rocks, and the rocks can collide with both the bucket and the support surface. However, the excavator body itself does not collide with this support platform. This configuration creates a designated manipulation surface, akin to a testbed table, where rocks are presented for the excavation task.

Soil Model: The bucket–soil interaction is simulated using a quasi-static 2D excavation force model, implemented following [1], and based on the classical Fundamental Equation of Earth-Moving [26]. The model estimates resistive forces using key geotechnical parameters—soil cohesion, internal friction angle, and soil-tool adhesion—computed in the plane of excavation. The 2D simplification enables computationally efficient training across thousands of parallel environments required for RL, while parameter randomization bridges the fidelity gap by exposing the policy to diverse soil conditions. These parameters are randomized during training to encourage generalization across soil types. The soil is modeled as a planar layer beneath the rock, co-located with the support platform. Interaction forces are computed in a vertical slice aligned with the current arm configuration, and applied to the shovel in Isaac Lab as state-dependent external forces opposing the digging trajectory. The “soft” and “hard” soil conditions referenced in evaluation correspond to the specific parameter ranges shown in Table I.

Simulated Perception: LiDAR-like point clouds are generated via Isaac Lab’s dynamic raycasting, using a virtual sensor mounted atop the cabin to mimic the real hardware’s

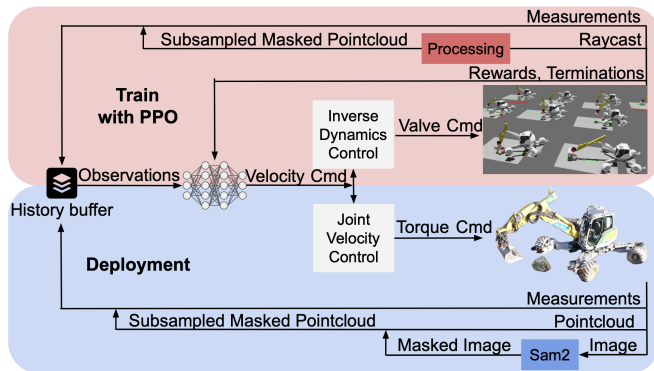


Fig. 2: Overview of our autonomous boulder removal system: (a) Perception pipeline using SAM2 segmentation on RGB images combined with sparse LiDAR points, (b) Simulation environment showing varied rock geometries and soil conditions.

TABLE I: Geotechnical parameters for soft and hard soil conditions used in simulation and evaluation

Parameter	Soft Soil	Hard Soil
Cohesion, c [Pa]	0	105,000
Internal friction angle, ϕ [°]	31.5	32.0
Unit weight, γ [N/m ³]	19,500	21,000
Soil-metal friction angle, δ [°]	23.0	23.0
Cavity expansion factor, CP	1	300
Adhesion factor, c_a/c	0	0.5

position and field of view. Rays intersect with both rock and bucket meshes, simulating occlusions. The simulator tags each hit with its source mesh, enabling the extraction of points belonging to the target rock, which are then used as policy input.

B. Baseline

The baseline is the continuous soil excavation policy from Egli et al. [1], [2], which outputs joint velocity commands using identical soil resistance modeling. However, it was trained for continuous media excavation with observations focused on soil height fields and rewards targeting soil volume collection.

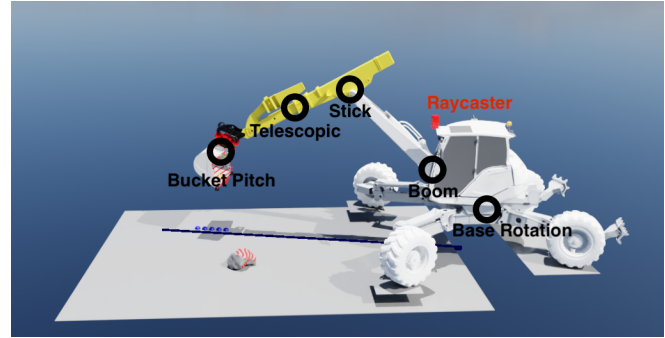


Fig. 3: Simulation setup. Red points indicate LiDAR raycast hits, and the dark blue line represents the soil surface model aligned with the excavator arm.

C. Reinforcement Learning Formulation

The agent learns a policy $\pi(a_t|s_t)$ that maps observations s_t to joint velocity commands a_t to acquire a target rock. This involves: (1) aligning the bucket with the rock, (2) executing a scooping motion to capture it, and (3) lifting and curling to secure it. The actuated joints include base turn, boom, stick, telescopic extension, and bucket pitch, denoted $q = [q_{\text{turn}}, q_{\text{boom}}, q_{\text{stick}}, q_{\text{tele}}, q_{\text{pitch}}]^T$.

1) *Observation Space (s_t):* The observation vector concatenates the following components:

- Proprioceptive state: joint positions, velocities, and torques as in [1]
- Turn joint position and its velocity history of length L
- Bucket pose and velocity in the rotating cabin frame
- Downsampled point cloud of $N = 20$ 3D points representing the target rock in the rotating base frame
- Previous action a_{t-1} , with turn joint actions stored in a history buffer of length L

All inputs are normalized to $[-1, 1]$.

2) *Action Space* (a_t): The policy outputs joint velocity commands:

$$a_t = [\dot{q}_{\text{turn}}, \dot{q}_{\text{boom}}, \dot{q}_{\text{stick}}, \dot{q}_{\text{tele}}, \dot{q}_{\text{pitch}}]^T \quad (1)$$

Turn commands include a deadband threshold ϵ_{turn} :

$$\dot{q}_{\text{turn}} = \begin{cases} 0 & \text{if } |\dot{q}_{\text{turn}}| < \epsilon_{\text{turn}} \\ \dot{q}_{\text{turn}} & \text{otherwise} \end{cases} \quad (2)$$

To emulate real-world delays, a control delay $\tau_{\text{delay}} \in [0, 1.2]$ s is randomly sampled at the start of each episode and applied throughout, enabling online delay compensation during training [27]. The turn joint is particularly affected by delays due to its different mechanics compared to traditional hydraulic cylinders and strong dependence on excavator inertia from varying arm configurations. The observation history length L is chosen such that $L \cdot \Delta t_{\text{control}} \geq \max(\tau_{\text{delay}})$, ensuring sufficient context for compensation.

3) *Rewards*: The reward function combines dense shaping and sparse terminal signals to guide efficient, safe, and task-completing behaviors, as summarized in Table II. Rewards R1–R6 encourage proper alignment, successful scooping, curling beyond $\theta_{\text{target}} = 0.5$ rad, lifting to a target height $h_{\text{desired}} = 0.5$ m above the soil. Penalties P1–P2 discourage abrupt actions and excessive bucket velocities above $v_{\text{max}} = 0.6$ m s⁻¹, while P3–P4 penalize unsafe or unnecessary turning, particularly during soil interaction. R7 and P6 serve as terminal rewards or penalties based on task outcome.

4) *Termination Conditions*: Episodes terminate on task success, safety violations, or timeout, enabling clear supervision and correct advantage estimation with GAE [28]. Terminations fall into three categories: T1 (timeout after 29 s), T2–T6 (safety and physical constraints), and T7 (success: the boulder is lifted above h_{desired} and curled beyond θ_{target} s). Safety thresholds include a maximum base speed of $v_{\text{max, base}} = 0.1$ m s⁻¹, bucket velocity limit $v_{\text{max, term}} = 1.2$ m s⁻¹, and angle-of-attack threshold $\alpha_{\text{threshold}} = 0$ rad. The boulder is considered dropped if it falls below h_{min} , defined by platform height. See Table III for full criteria.

5) *Curriculum and Initialization*: To accelerate training and improve generalization, we employ a five-stage curriculum that progressively increases task complexity. Each episode begins from a pre-generated, collision-free configuration sampled from a cached dataset of valid initial states, ensuring safe and consistent resets.

The curriculum is structured to guide the agent through increasingly difficult behaviors—starting with lifting, then scooping, alignment, and ultimately generalizing across diverse rock and soil configurations.

- **Level 0**: Rock is placed directly in front of the bucket (within a $2.5 \times 0.6 \times 0.4$ m region); soil is soft; angle-of-attack termination (T6) is disabled to focus on learning lifting without caring how the soil is penetrated.
- **Level 1**: T6 is activated, enforcing correct bucket orientation during scooping.
- **Level 2**: Soil properties (cohesion, adhesion factor, internal friction angle, soil-tool friction angle, weight

unit per volume, cutting resistance) are randomized, exposing the agent to varying excavation dynamics.

- **Level 3**: Rock and bucket initial positions are sampled from a wider $4.5 \times 3 \times 0.4$ m volume, requiring learning alignment strategies.
- **Level 4**: P5 penalty is enabled to discourage misaligned digging; rock longitudinal position is further randomized to train depth-aware scooping.

Transitions between levels are triggered automatically based on training performance: the agent advances to the next level once the success rate on the current level exceeds 80% over a rolling window of 1000 episodes. This staged approach enables efficient skill acquisition and supports robust sim-to-real transfer.

6) *Training*: The policy is trained using Proximal Policy Optimization (PPO) [29] with Generalized Advantage Estimation (GAE) [28]. We use an actor-critic architecture with two hidden layers of 256 units, leaky ReLU activations, and an initial noise standard deviation of 0.4. Optimization uses a fixed learning rate of 10^{-4} , clip ratio 0.2, value loss coefficient 0.5, entropy coefficient 0.005, and 5 epochs per update over 4 minibatches. The agent is trained for 9000 iterations (24 steps per environment per update), requiring approximately 8 years of experience in simulation. Full training code is based on RSL-RL [30].

D. Hardware Platform and Perception Pipeline

We deploy our policy on a Menzi Muck M545, a 12-ton hydraulic walking excavator [31]. The learned joint velocity commands from simulation are executed on the real system using PID and feedforward valve-flow controllers, which generate solenoid currents for the hydraulic valves. Autonomous control is enabled by retrofitting electrically actuated pilot-stage valves. Arm joint states are measured via IMUs, leg joints with magnetostrictive sensors, and joint torques are estimated from hydraulic cylinder pressures. An onboard IMU provides global orientation.

Perception is performed by a cabin-mounted Ouster OS1 Rev 6 LiDAR and a ZED 2i stereo camera. We employ SAM2 [32], a foundation model with video tracking capabilities, to segment the target rock. A custom ROS 2 wrapper enables real-time tracking within a moving window [33]. Segmentation is manually initialized by the user, who selects a small number of pixels (5) on the target rock in the camera image. The resulting mask is back-projected onto the LiDAR point cloud to extract a 3D surface representation. Inference runs onboard on an NVIDIA RTX 4080 GPU with a latency of $\simeq 95$ ms. Given our control loop operates at 6 Hz, this allows one inference per policy step.

SAM2 is well-suited for the harsh outdoor conditions encountered (mud, shadows, diverse rock textures) due to its strong generalization capabilities. Its robustness across lighting and occlusion scenarios enables zero-shot segmentation in challenging real-world scenes (Fig. 4). The full deployment pipeline is summarized in Fig. 2.

TABLE II: Reward and penalty terms used during training. *Notation.* We denote positions and velocities using $\mathbf{p}_{\text{object}}^{\text{frame}}$ and $\mathbf{v}_{\text{object}}^{\text{frame}}$, respectively, where the superscript indicates the reference frame (e.g., base, rotbase, world). For instance, $\mathbf{v}_{\text{bucket}}^{\text{rotbase}}$ is the linear velocity of the bucket in the rotating base frame. The bucket’s curl angle is θ_{curl} , with θ_{target} the desired value for securing a rock. The digging penalty P5 depends on the current bucket depth d_{bucket} , and thresholds d_{soft} and d_{hard} , along with lateral alignment y_{bucket} with bounds y_{min} and y_{max} .

ID	Description	Expression	Weight
Rewards			
R1	Alignment with boulder (lateral)	$\exp(- \mathbf{p}_{\text{boulder},y}^{\text{rotbase}} ^2)$	0.005
R2	Bucket proximity to boulder	$\mathbf{1}_{\ \mathbf{p}_{\text{boulder},y}^{\text{rotbase}} - \mathbf{p}_{\text{bucket},y}^{\text{rotbase}}\ ^2 < 1.5}$	0.01
R3	Bucket beneath boulder	$\mathbf{1}_{R2 \neq 0} \cdot \mathbf{1}_{\mathbf{p}_{\text{bottom plate},z}^{\text{base}} < \mathbf{p}_{\text{boulder},z}^{\text{base}}}$	0.01
R4	Rock inside shovel	$\mathbf{1}_{\text{boulder in shovel}}$	0.075
R5	Rock secured via curl	$\mathbf{1}_{\text{in shovel}} \cdot \mathbf{1}_{\text{curled enough}}$	0.05
R6	Lifted to target height	$\mathbf{1}_{\Delta z > 0} \cdot \mathbf{1}_{\text{in shovel}} \cdot \exp(-\ \mathbf{p}_{\text{bucket},z}^{\text{base}} - h_{\text{desired}}\ ^2)$	0.05
R7	Task success (terminal)	$\mathbf{1}_{\text{success}}$	20
Penalties			
P1	Action smoothness	$\ \mathbf{a}_t - \mathbf{a}_{t-1}\ ^2$	-0.005
P2	Excessive bucket velocity	$\max(0, \ \mathbf{v}_{\text{bucket}}^{\text{rotbase}}\ - v_{\text{max}}) \cdot 10^{\frac{\max(0, \ \mathbf{v}_{\text{bucket}}^{\text{rotbase}}\ - v_{\text{max}})}{-v_{\text{max}}}}$	-0.1
P3	Unnecessary turn	$\mathbf{1}_{ \dot{q}_{\text{turn}} \geq \epsilon_{\text{turn}}}$	-0.005
P4	Turning while digging	$\mathbf{1}_{\text{in soil}} \cdot \mathbf{1}_{ \dot{q}_{\text{turn}} \geq \epsilon_{\text{turn}}}$	-0.025
P5	Misaligned shallow scooping	$\frac{d_{\text{bucket}} - d_{\text{soft}}}{d_{\text{hard}} - d_{\text{soft}}} \cdot \frac{ y_{\text{bucket}} - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}}$	-0.0125
P6	Task failure (terminal)	$\mathbf{1}_{\text{failed}}$	-0.5

TABLE III: Termination conditions used during training and evaluation. *Notation.* α denotes the angle between the bucket’s bottom plate orientation and its instantaneous velocity vector in the xz plane, capturing the alignment between the cutting edge and its direction of motion. Time limit is 29 s.

Condition	Criterion
T1: Episode timeout	$\text{time} \geq \text{time_limit}$
T2: Excessive base velocity (-)	$\ \mathbf{v}_{\text{base}}^{\text{world}}\ > v_{\text{max, base}}$
T3: Joint velocity limit exceeded (-)	$\max_i \dot{q}_i > \dot{q}_{\text{max}}$
T4: Excessive bucket speed (-)	$\ \mathbf{v}_{\text{bucket}}^{\text{rotbase}}\ > v_{\text{max, term}}$
T5: Boulder dropped (-)	$\mathbf{p}_{\text{boulder},z}^{\text{base}} < h_{\text{min}}$
T6: Invalid angle of attack (-)	$\alpha_{\text{bucket}} > \alpha_{\text{threshold}}$
T7: Task success (+)	$\text{boulder in shovel} \wedge (\theta_{\text{curl}} > \theta_{\text{target}}) \wedge (\mathbf{p}_{\text{boulder},z}^{\text{base}} > h_{\text{desired}})$

III. RESULTS

We evaluate our approach through extensive simulation experiments and field deployment on a 12-ton Menzi Muck M545 excavator. Results demonstrate effective boulder extraction across varied conditions, with simulation comparisons highlighting the necessity of specialized learning for this task.

A. Simulation Experiments

Test conditions encompassed 50 scanned rock geometries divided into small (0.3-0.5m) and large (0.8-1.0m) cate-



Fig. 4: SAM2 segmentation in a difficult outdoor setting: low light, occlusion, and cluttered terrain.

gories, with soil parameters ranging from soft to hard. Our method demonstrated consistently high performance (88-97%) across all conditions, while the baseline approach succeeded only with small rocks in soft soil (86%), where standard digging trajectories naturally capture rocks along with surrounding soil. The baseline’s performance degraded significantly with larger rocks even in soft soil (44%) and failed entirely in harder soil conditions.

This performance gap reveals the fundamental complexity of boulder manipulation. In hard soil, effective scooping becomes impossible as rocks slide along the surface rather than entering the bucket when penetration depth is insuffi-

cient. These challenging conditions demand highly reactive behaviors—the ability to dynamically adapt trajectory and depth based on real-time rock pose feedback—capabilities that traditional continuous media controllers inherently lack. Large rocks present additional challenges: they require precise reactive maneuvers to fit properly within the bucket, which explains our method’s slightly reduced but still robust performance (88%) in the most demanding scenario.

To evaluate generalization, we tested on 25 previously unseen rock geometries (0.8-1.0m) in hard soil conditions. The system achieved 87% success compared to 88% on training rocks, demonstrating strong robustness to novel shapes and confirming that our approach learns generalizable manipulation strategies rather than memorizing specific geometries.

B. Hardware Deployment

We conducted 40 hardware trials across four conditions: two rock sizes (approximately 0.4m and 0.7m diameter) and two soil types (soft and hard). An operator with approximately 300 hours of general excavator experience performed the same tasks for comparison. We compare against the continuous excavation baseline from Egli et al. [1], [2], which lacks boulder-specific manipulation capabilities. Tables IV and V summarize the results.

TABLE IV: Simulation Results

Condition	Boulder RL (Ours)	Digging RL Baseline
Small/Soft	96%	86%
Big/Soft	97%	44%
Small/Hard	94%	4%
Big/Hard	88%	0%
Average	94%	33%

TABLE V: Hardware Results

Condition	Trials	Boulder RL (Ours)	Human Operator
Small/Soft	10	90%	90%
Big/Soft	10	80%	90%
Small/Hard	10	50%	80%
Big/Hard	10	60%	70%
Overall	40	70%	83%

The autonomous system achieved 70% success rate compared to 83% for the human operator (Table V). Both performed best on small rocks in soft soil (90% success), while small rocks in hard soil proved most challenging for the autonomous system (50% vs 80% human).

1) *Comparative Failure Analysis:* Human operators achieve their 83% success rate using vision alone, interpreting soil deformation, bucket motion, and rock movement to infer resistance and adjust strategies accordingly. The autonomous system achieves 70% success using sparse point clouds and proprioceptive feedback.

Both systems exhibited similar failure modes in hard soil conditions: rocks sliding along the ground surface rather than entering the bucket when insufficient penetration depth prevented proper scooping. In soft soil, failures occurred when the bucket was overfilled, causing rocks to fall during lifting.

The largest performance gap (30%) occurred with small rocks in hard soil. Human operators maintained visual tracking through occlusions, compensating with spatial memory when rocks disappeared behind the bucket. In contrast, our vision system loses tracking during such occlusions, and small rocks are particularly susceptible to complete visual obstruction during the final approach. The cabin-mounted sensors compound this issue, as the arm itself blocks the camera’s view during close manipulation. Additionally, the autonomous system struggles to recover when debris accumulates or soil height varies around the rock—conditions under which humans detect changes and intuitively reposition for cleaner approaches.

C. Soil-Adaptive Behavior

The policy exhibits different excavation strategies based on soil resistance, consistent with behaviors observed in prior soil-adaptive excavation work [1]. Fig. 5 shows representative trajectories from hardware deployment. In hard soil, the bucket follows an extended horizontal path (approximately 40% longer), dragging along the surface to minimize penetration resistance. In soft soil, the bucket penetrates more directly, reducing the horizontal scooping distance.

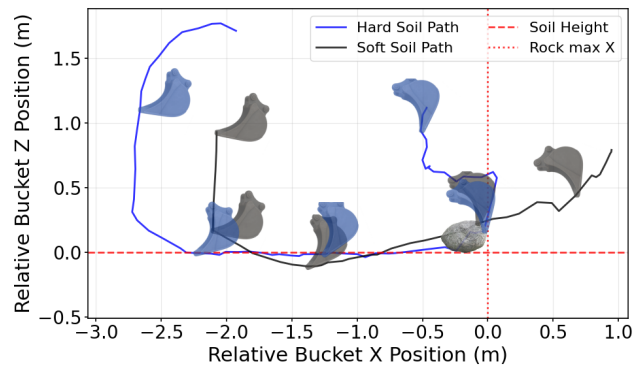


Fig. 5: Shovel trajectories in hard vs soft soil conditions. The policy adapts its strategy based on resistance: extended surface dragging in hard soil versus direct penetration in soft soil. Trajectories aligned at soil level for comparison.

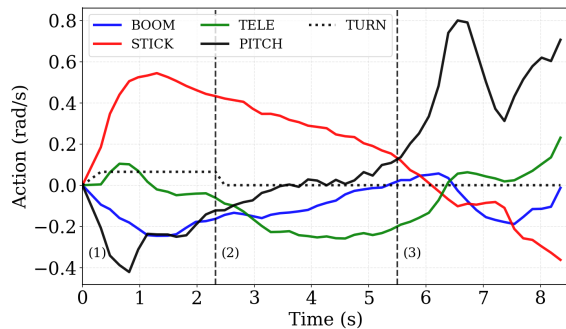
These strategies emerge from the interaction between proprioceptive feedback and the reward structure, without explicit soil property estimation. Dragging behavior is observed when torque readings are high during initial penetration attempts, while more direct approaches occur when resistance is low. Human operators develop similar strategies, but based almost entirely on visual cues, observing how soil deforms and how the bucket progresses to infer resistance levels.

D. Control Behavior Analysis

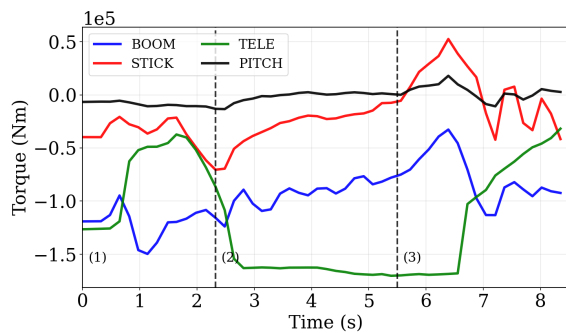
Fig. 6 presents joint commands and torque measurements during successful execution. The control sequence shows three distinct emergent phases:

- 1) **Alignment:** Turn joint actively positions the rock in front of the bucket
- 2) **Scooping:** Coordinated boom, stick, and bucket motion to capture the rock
- 3) **Lifting:** Upward motion with bucket curl to secure the load

Once aligned, turn commands consistently enter the dead zone (Eq. (2)), effectively constraining the task to 2D manipulation. This mirrors human operation where operators typically align first, then execute scooping in a vertical plane. The torque profiles reveal how soil conditions influence execution: elevated stick and boom torques in hard soil (Fig. 6b) correlate with the extended dragging trajectories.



(a) Joint velocity commands



(b) Measured joint torques (turn joint excluded)

Fig. 6: Control signals during successful rock extraction. Vertical lines indicate phase transitions: alignment → scooping → lifting.

E. Failure Modes and Real-World Challenges

Analysis of failed trials reveals three primary categories where simulation assumptions break down:

- **Perception and occlusion:** The fixed cabin-mounted sensors create blind spots during close manipulation. This manifests in two ways: (1) complete tracking loss when small rocks disappear behind the bucket, and (2) inability to detect partially buried rocks where only the visible portion is captured by LiDAR. These perception

failures account for the majority of unsuccessful trials, particularly with small rocks.

- **Terrain and soil dynamics:** Real environments introduce unmodeled boulder-soil coupling that affects extraction dynamics, while accumulated soil progressively fills the bucket, reducing space for rock capture and causing spillage failures observed in soft soil conditions.
- **Control constraints:** Imprecise turn control (due to coarse dynamics modeling) sometimes leads to suboptimal alignment with large rocks, occasionally causing lateral entrapment against the bucket.

These limitations highlight fundamental trade-offs in sim-to-real transfer: while higher-fidelity simulation, such as particle-based methods, could address some issues, the computational cost currently prohibits large-scale RL training.

IV. CONCLUSION

We demonstrated autonomous boulder extraction on a 12-ton hydraulic excavator, achieving 70% success rate compared to 83% for human operators across rocks (0.4–0.7m) and varying soil conditions. The reinforcement learning policy, trained with rigid-body simulation and analytical soil models, adapts its strategy to soil resistance by dragging along hard surfaces and penetrating directly in soft conditions, using only standard digging buckets.

The system operates with 20 LiDAR points per rock and proprioceptive feedback, showing that boulder manipulation does not require dense geometric reconstruction. Adaptive behaviors emerge from the interaction between proprioceptive sensing and reward structure without explicit soil property estimation. Primary failures occur when small rocks become fully occluded during approach, causing tracking loss. The fixed cabin-mounted sensors create blind spots during close manipulation, and the analytical soil model cannot capture all real-world boulder-soil coupling effects.

This work shows that boulder manipulation, which previously required tool changes or manual operation, can be automated using standard excavator buckets through learning-based control. The approach demonstrates that practical construction automation often requires adaptive behavior rather than sophisticated hardware. Future work should address tracking through occlusions using temporal models and extend to multi-rock scenarios common in site clearing.

REFERENCES

- [1] P. Egli, D. Gaschen, S. Kerscher, D. Jud, and M. Hutter, “Soil-adaptive excavation using reinforcement learning,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9778–9785, 2022.
- [2] P. Egli, L. Terenzi, and M. Hutter, “Reinforcement learning-based bucket filling for autonomous excavation,” *IEEE Transactions on Field Robotics*, vol. 1, pp. 170–191, 2024.
- [3] R. L. Johns, M. Wermelinger, R. Mascaro, D. Jud, I. Hurkxkens, L. Vasey, M. Chli, F. Gramazio, M. Kohler, and M. Hutter, “A framework for robotic excavation and dry stone construction using on-site materials,” *Science Robotics*, vol. 8, no. 84, p. eabp9758, 2023.
- [4] M. Wermelinger, R. Johns, F. Gramazio, M. Kohler, and M. Hutter, “Grasping and object reorientation for autonomous construction of stone structures,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5105–5112, 2021.

- [5] R. L. Johns, M. Wermelinger, R. Mascaro, D. Jud, F. Gramazio, M. Kohler, M. Chli, and M. Hutter, "Autonomous dry stone: On-site planning and assembly of stone walls with a robotic excavator," vol. 4, no. 3-4, pp. 127–140.
- [6] M. Wermelinger, R. L. Johns, F. Gramazio, M. Kohler, and M. Hutter, "Grasping and Object Reorientation for Autonomous Construction of Stone Structures," p. 9392354.
- [7] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen, "Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection," in *The International Journal of Robotics Research*, vol. 37, no. 4-5. SAGE Publications, 2018, pp. 421–436.
- [8] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, and S. Levine, "QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation," in *Conference on Robot Learning*. PMLR, 2018, pp. 651–673.
- [9] M. Shridhar, L. Manuelli, and D. Fox, "Perceiver-Actor: A Multi-Task Transformer for Robotic Manipulation," in *Conference on Robot Learning*. PMLR, 2023, pp. 785–799.
- [10] Physical Intelligence Team, " π 0.5: A Vision-Language-Action Flow Model for General Robot Control," *arXiv preprint arXiv:2501.00211*, 2025.
- [11] R. Mascaro, M. Wermelinger, M. Hutter, and M. Chli, "Towards automating construction tasks: Large-scale object mapping, segmentation, and manipulation," *Journal of Field Robotics*, vol. 38, no. 5, pp. 684–699, 2021.
- [12] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 652–660.
- [13] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [14] B. Gyenes, N. Franke, P. Becker, and G. Neumann, "PointPatchRL – Masked Reconstruction Improves Reinforcement Learning on Point Clouds," in *8th Annual Conference on Robot Learning*, 2024.
- [15] D. Hoeller, N. Rudin, D. Sako, and M. Hutter, "ANYmal Parkour: Learning Agile Navigation for Quadrupedal Robots," *Science Robotics*, vol. 9, no. 88, p. eadi7566, 2024.
- [16] A. Goyal, V. Blukis, J. Xu, Y. Guo, Y.-W. Chao, and D. Fox, "RVT2: Learning Precise Manipulation from Few Demonstrations," in *Robotics: Science and Systems*, 2024.
- [17] T. Gervet, Z. Xian, N. Gkanatsios, and K. Fragkiadaki, "Act3D: 3D Feature Field Transformers for Multi-Task Robotic Manipulation," in *Conference on Robot Learning*. PMLR, 2023, pp. 3949–3965.
- [18] T. Chen, M. Tippur, S. Wu, V. Kumar, E. Adelson, and P. Agrawal, "Visual Dexterity: In-Hand Reorientation of Novel and Complex Object Shapes," *Science Robotics*, vol. 8, no. 84, p. eadc9244, 2023.
- [19] Y. Ze, Z. Yan, N. Aubouin, M. Guo, T. Haarnoja, and X. B. Peng, "H-GAP: Humanoid Control with a Generalist Planner," in *International Conference on Learning Representations*, 2024.
- [20] H. N. Cannon, "Extended Earthmoving with an Autonomous Excavator." Master's thesis, Carnegie Mellon University, Pittsburgh, PA, May 1999, tech. Rep. CMU-RI-TR-99-10. [Online]. Available: <https://www.cs.cmu.edu/afs/cs/Web/People/hcannon/thesis.pdf>
- [21] D. Jud, I. Hurkxkens, C. Girot, and M. Hutter, "Robotic embankment," vol. 5, no. 2, pp. 101–113.
- [22] A. Stentz, J. Bares, S. Singh, and P. Rowe, "A robotic excavator for autonomous truck loading," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, vol. 3, Victoria, BC, Canada, 1998, pp. 1885–1893.
- [23] L. Zhang, J. Zhao, P. Long, L. Wang, L. Qian, F. Lu, X. Song, and D. Manocha, "An autonomous excavator system for material loading tasks," vol. 6, no. 55, p. eabc3164.
- [24] Q. Lu, Y. Zhu, and L. Zhang, "Excavation reinforcement learning using geometric representation," *IEEE Robotics and Automation Letters*, vol. 7, 04 2022.
- [25] M. Mittal, C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, J. L. Yuan, R. Singh, Y. Guo, H. Mazhar, A. Mandekar, B. Babich, G. State, M. Hutter, and A. Garg, "Orbit: A unified simulation framework for interactive robot learning environments," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, p. 3740–3747, Jun. 2023.
- [26] A. R. Reece, "Paper 2: The fundamental equation of earth-moving mechanics," *Proceedings of the Institution of Mechanical Engineers, Conference Proceedings*, vol. 179, no. 6, pp. 16–22, 1964. [Online]. Available: https://doi.org/10.1243/PIME.CONF.1964.179_134.02
- [27] F. A. Spinelli, P. Egli, J. Nubert, F. Nan, T. Bleumer, P. Goegler, S. Brockes, F. Hofmann, and M. Hutter, "Reinforcement learning control for autonomous hydraulic material handling machines with underactuated tools," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 12 694–12 701.
- [28] J. Schulman, P. Moritz, S. Levine, M. Jordan, and P. Abbeel, "High-dimensional continuous control using generalized advantage estimation," 2018.
- [29] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017.
- [30] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to Walk in Minutes Using Massively Parallel Deep Reinforcement Learning," in *Proceedings of the 5th Conference on Robot Learning*. PMLR, 2022, pp. 91–100.
- [31] D. Jud, S. Kerscher, M. Wermelinger, E. Jelavic, P. Egli, P. Leemann, G. Hottiger, and M. Hutter, "Heap - the autonomous walking excavator," *Automation in Construction*, vol. 129, p. 103783, Sep. 2021.
- [32] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryal, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson *et al.*, "SAM 2: Segment Anything in Images and Videos," *arXiv preprint arXiv:2408.00714*, 2024.
- [33] J. Grütter, "SAM2 ROS2 Wrapper for Real-time Tracking," <https://github.com/leggedrobotics/segment-anything-2-tracking-ros-2>, 2024.