# Active sensing in a bioinspired hand as an enabler of implicit curriculum learning for manipulation

Romina Mir[a*], Pegah Ojaghi[b*], Ali Marjaninejad[a], Michael Wehner[c], Francisco Valero-Cuevas[a,d,f §]
[a]Biomedical Engineering Department, University of Southern California ,Los Angeles
[b]Computer Science and Engineering Department, University of California Santa Cruz, Santa Cruz
[c]Electrical and Computer Engineering Department, University of California Santa Cruz, Santa Cruz
[d]Division of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles
[f]Computer Science Department, University of Southern California, Los Angeles
*Denotes equal contribution, §Corresponding author

## 1 Introduction

Autonomous learning of in-hand manipulation is a litmus test for bioinspired robots and AI algorithms [1]. Currently, several groups have shown it is possible to use the fingers of a simulated or robotic hand to rotate an object [2]. However, that work was done with the object resting on the upward-facing palm and the object did not have to be held against gravity. Here we extend that work by exploring the role of the reward function used on the ability to learn a combined in-hand manipulation task against gravity. In particular, the task was to autonomously learn to lift a ball against gravity while spinning it against a torsional spring. Our question was to see if the algorithm would learn best by being rewarded for lifting the object, vs. by being rewarded for both lifting and spinning. We found that, somewhat counterintuitively, being rewarded for both features of the task are necessary for learning. We discuss why we believe that exploring how to spin the ball (a more "complex" task) was in fact critical to learning how to lift the ball (a "simpler" task).

## 2 Methods

Our agent is a 3-fingered bio-inspired hand with the two-joint, three-servo motor fingers simulated in the MuJoCo physics environment [3]. We also use MuJoCo's built-in features to record contact force magnitude (1D) ('touch') on the fingertips of all three fingers [2, 3]. 'Touch' sensor sites at the soft finger tips provide a nonnegative scalar-value indicating the cumulative normal contact forces on the sensor area. The friction coefficient for dynamically generated contact pairs is also specified for all fingertips (a.k.a. soft contact with friction) [3]. The in-hand manipulation task was rewarded for (i) lifting and holding a ball against gravity (Figure 1) while (ii) spinning it about a floating horizontal axis (Figure 2).

We used the end-to-end Proximal Policy Optimization (PPO) autonomous learning algorithm as per the PPO1 implementation from OpenAI's stable baselines repository

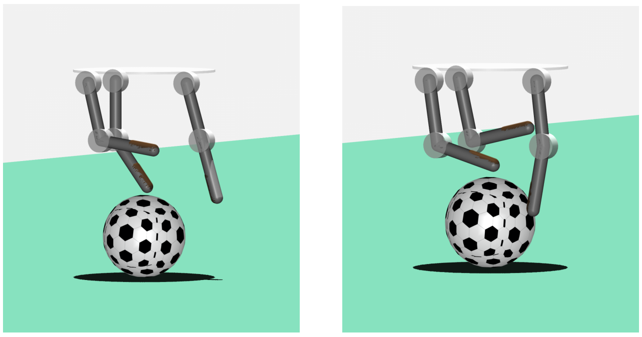

**Figure 1:** The three-fingered hand in the MuJoCo environment.

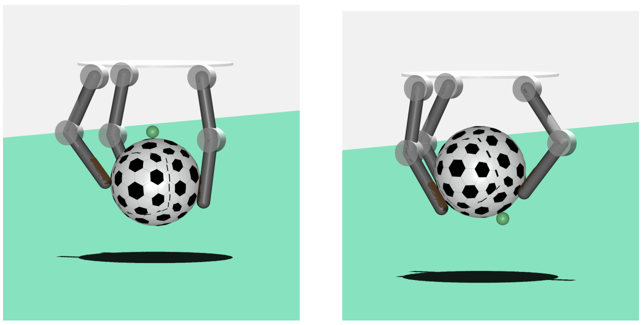

**Figure 2:** The three-fingered hand in the MuJoCo environment, holding a ball against gravity and rotating it about the Y axis (The small green dot on the ball is a point of rotation).

with MultiLayer Perceptron (MLP) Artificial Neural Network (ANN) as for the actor-critic map. We ran the training in independent Monte Carlo runs for the task, each simulating 2.77 hours of training. Each Monte Carlo run had $1,000$ training episodes lasting 10s (sampled at 0.01s).

The simulated hand consists of a palm and 3 identical servo-driven fingers: two adjacent fingers (analogous to the 'middle' and 'index' fingers) and one opposing them (analogous to the 'thumb') (Figure 1). Each finger consists of two joints. The size of the palm and length-ratio of each 'pha-

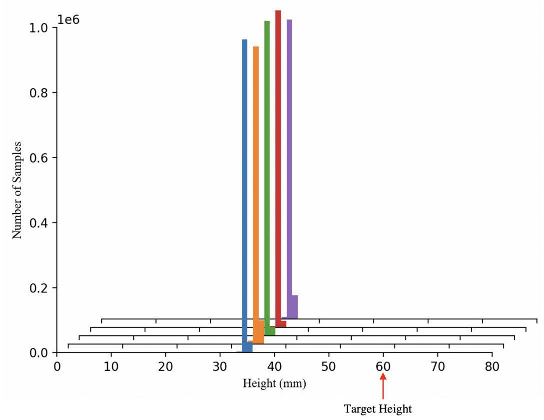


**Figure 3:** Results for lifting the ball beyond the rest position of 32 mm when only lifting is rewarded using the PPO algorithm. We show the histograms from 5 independent 10s Monte Carlo runs, each with 1e6 samples (i.e., time points).

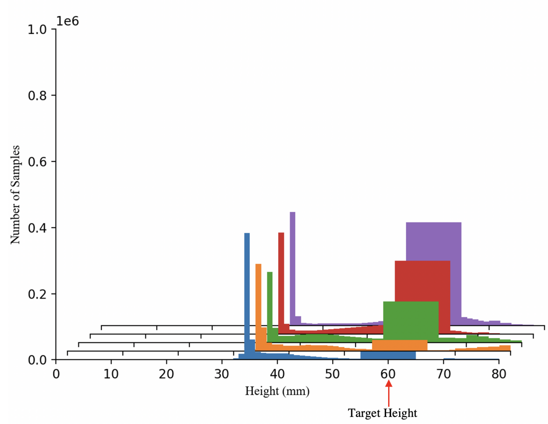


**Figure 4:** Results for lifting the ball beyond the rest position of 32 mm when both lifting and spinning are rewarded using the PPO algorithm ($\pm 5$). We show 5 independent Monte Carlo runs.

lanx' was based on the human hand ratios [4,5]. The fingers do not have an abduction-adduction degree of freedom at their base and can only flex or extend at the two joints. The sensory sites are only used on the internal side (i.e., the 'pads of the fingertips') of the distal phalanx of each finger.

We use a ball with friction as our experimental object in the task (see Figure 1). The ball has a mass of 5 gr and it has a total of 3-degrees of freedom (DoF): 2 translational DoFs ( $X$ and $Z$) and 1 rotational DoF along the $Y$-axis with a built-in damping with 0.0005Ns/m damping coefficient.

## 3 Result

In Figure 3, the reward is 100% for lifting the ball against gravity and 0% for spinning the ball. In Figure 4 the reward is now 60% for lifting and 40% for spinning. The histograms show the height of the ball for each independent Monte Carlo run (lasting 2.77 hours with 1e6 samples of 0.01 s. See Methods). The target height of the center of the ball was 60mm while the resting height of the center of the ball on the ground is 34mm (the radius of the ball). The task is learnable only when the reward includes both lifting and spinning (Figure 4). Any heights below 34 mm indicate penetration of the ground by the ball in this elastic medium.

## 4 Discussion and Conclusions

Our results suggest the implicit emergence of curriculum learning on the basis of random exploration of the full dynamics of the task. That is, while lifting the ball seems like a "simpler" task, it does not seem to be possible before the system "understands" the dynamics of manipulation. By rewarding (and therefore encouraging) exploration of the full dynamics of manipulation (i.e., the grasp matrix of the system [6]), the implicit model being built approximates full rank and is therefore more useful. This is also perhaps analogous to the concepts of observability and persistence of excitation in control theory [7].

In the context of Machine Learning, a human usually specifies a curriculum to be followed by the agent where prior knowledge provides a rank ordering of the functional components of the task based on assumed complexity. In our case, we did not specify such rank ordering as the agent learned autonomously without human intervention. In retrospect, one would have thought that lifting the ball by learning form closure is "simpler" than lifting and spinning the ball which requires force closure [1, 6]. However, we find that exploring how to spin the ball (a more "complex" task) seems to be, in fact, critical to learning how to lift the ball (a "simpler" task).

By demonstrating the natural emergence of learning for grasp and manipulation in biorobotic systems, we shed light on the implicit sensorimotor processes [8] in biological systems that may grant humans unparalleled dexterity.

### References


[1] FJ. Valero-Cuevas and M. Santello. On neuromechanical approaches for the study of biological and robotic grasp and manipulation. *Journal of neuroengineering and rehabilitation*, 14(1):1–20, 2017.

[2] M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, et al. Learning dexterous in-hand manipulation. *The International Journal of Robotics Research*, 39(1):3–20, 2020.

[3] E. Todorov, T. Erez, and Y. Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012.

[4] A. Melnik, L. Lach, M. Plappert, T. Korthals, R. Haschke, and H. Ritter. Tactile sensing and deep reinforcement learning for in-hand manipulation tasks. In *IROS Workshop on Autonomous Object Manipulation*, 2019.

[5] CC. Norkin and DJ. White. *Measurement of joint motion: a guide to goniometry*. FA Davis, 2016.

[6] RM. Murray, Z. Li, SS. Sastry, and SS. Sastry. *A mathematical introduction to robotic manipulation*. CRC press, 1994.

[7] M. Green and JB. Moore. Persistence of excitation in linear systems. *Systems & control letters*, 7(5):351–360, 1986.

[8] JA. Berry and FJ. Valero-Cuevas. Sensory-motor gestalt: Sensation and action as the foundations of identity, agency, and self. In *Artificial Life Conference Proceedings*, pages 130–138. MIT Press, 2020.