Robots Work, People Rule: Human-Centered Pick-And-Place Automation.

Halit Bener Suay[1], Abhijit Majumdar[2] and Daniel H. Grollman[3]

Abstract— **We present a human-in-the-loop remote robot supervision platform that enables workers in warehouses and distribution centers to increase their throughput. In our platform humans take a central role in two ways: (i) through their contribution in graceful exception handling and (ii) in improving the machine learning model performance. A preliminary case study shows the feasibility of our approach using data from a deployed pick-and-place system in a production setting. We argue for a more human-centric view of automation that elevates human labor instead of replacing it.**

I. INTRODUCTION

Increasing demand in e-commerce, in particular due to the recent Covid-19 outbreak [1] and warehouse labor shortage has been forcing supply chain companies to provide better incentives for their workers and automate almost every step of the logistics operations for the last decade [2]. These steps, however, have not been sufficient to fully address the labor shortage the industry is still experiencing [3].

Despite the popular perception that automation replaces the manual workforce, the truth is that automated solutions in logistics today act as a transforming factor and a force multiplier to workers and enterprises [4] to help meet the increasing demand. This crucial *helper* role of today's automation comes with the responsibility of mindfully designed systems that improve the meaning and the experience of the work.

Commercialization of recent research in various sub-fields in robotics boosted startups' ability to solve various automation problems in logistics. With robotic systems that are able to handle greater uncertainty in their environment, successfully deployed systems today help move totes, shelves, pick-and-place items for shipping and fulfilling orders more efficiently [5].

Despite the increased capability of commercial robotic pick-and-place systems, particularly in warehouse applications, the variance in types of items a system is expected to handle still provides a challenge. The variability can be high enough that at the time of engineering, it is impractical, if not impossible, for a fully automated solution to reach human levels of robustness in material handling without any supervision. Thus, we argue that to be commercially viable, the future of work in automated warehouse logistics must incorporate human supervision input into its core design.

[1]Sr. Developer, Plus One Robotics, Pittsburgh, PA, 15213, USA `bener.suay@plusonerobotics.com`

[2]Sr. Developer, Plus One Robotics, San Antonio, TX, 78226, USA `abhijit.majumdar@plusonerobotics.com`

[3]Principal Engineer, Plus One Robotics, Boulder, CO, 80301, USA `dan.grollman@plusonerobotics.com`

In our domain of warehouse logistics, we focus on pick and place operations, where individual items or packages must be moved from one place to another. Software for these systems typically consist of several layers of model/heuristic based decision making components. Detecting and classifying objects and evaluating reachable pick poses are a few of the decisions the software needs to output. For the purposes of this paper, we focus on the *picking* problem. We call what is observable to the system during its engineering as its training input for the entire decision making pipeline and what the system observes during its lifetime after deployment as its test input.

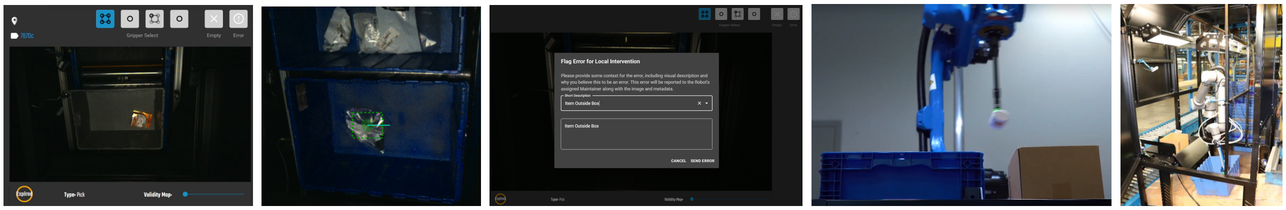
We draw similarities from the definition of covariate shift and heavy-tailed distributions [6], [7] to define our two core problems: the inequality of the distribution of the training data and the test data for a machine learning model, and/or the lack of training samples that can accurately capture the entire distribution. Therefore we propose a human-in-the-loop pick-and-place system that addresses these two issues by enabling below:

- A single worker, a *crew chief* to direct the work of many robots for graceful exception handling,
- The input of the crew chief to continuously adapt the decision making components' underlying models/heuristics to better match the observed input data during operation.

II. HUMAN-CENTERED AUTOMATION

We present a platform that enables multiple crew chiefs to direct the work of many robots. The inherent variability of items to be handled in the warehouse require a graceful exception handling approach. Our platform enables a remote worker to provide input by simply looking at the same image as the robot and clicking on the item most appropriate to be picked. We aim to maximize the number of robots per crew chief so that the users can achieve the workforce multiplication factor they need and obtain satisfactory throughput increases.

Fig.1 shows images from our deployed system where a crew chief is prompted when an exception occurs during the operation of a deployed pick-and-place system. A particular exception we focus on in this paper is a *pick* exception where the vision model or the end-effector of the robot has performed in a way that is different than expected, resulting in a failed pick. Pick-and-place system detects a failure and sends a request to the waiting crew chief (Fig.1a). Crew chief addresses a request with a pick response using the web interface (Fig.1b). Robot proceeds to pick the item based on



(a) A pick request (b) Pick response (c) Exception menu (d) Corrected pick (e) Test System

Fig. 1: Images from a deployed human-in-the-loop exception handling process flow to pick-and-place for shipping.

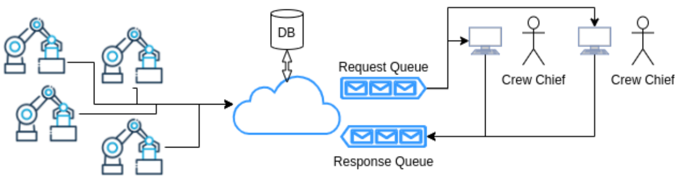


Fig. 2: Multi-robot, multi-crew chief platform.

the pick response and stores all the data the lead to this pick request and the successful pick.

With the help of the crew chief, the system is able to rapidly (within 5-6 seconds) [1] handle the exception and return to autonomous operation. Fig.1c shows a crew chief triggering an intervention exception. Meanwhile, the platform stores the request and the response on a secure, encrypted, database with access control privileges for future reference and use in improving the machine learning model for future picks. Fig.2 depicts the multi-robot, multi-human aspects of the cloud enabled data flow. All request and response data is stored in databases for future reference and use in various contexts such as system improvements and continuous machine learning to adapt to tail/edge cases and address covariate shift.

A. *Case Study*

We tested our continuous learning with human-in-the-loop paradigm on an actual production system we have deployed (Fig.1e). We started with an initial deep learning model trained with $\sim$9K images for object detection and the system ran for 15 weeks. During the initial test period, our system handled $\sim$36K pick cycles.

Our continuous learning paradigm works as below:

- We deploy our system with an initial set of models/heuristics.
- As the system operates, observed (test) data results in either a successful pick-and-place cycle or a failure.
- Failures get stored in a database and pushed into a request queue for crew chief(s).
- Crew chief(s) provide input, which are also stored and used for incremental updates to the learned model.

Following these steps results in moving more of the observed "tail" cases to be incorporated into model updates iteratively.

Of these $\sim$36K cycles, $\sim$23K triggered human supervision, indicating that the input were sufficiently different than images we have used for training our vision model for initial testing. Analysis of the model performance on the test case data shows that the initially deployed vision model's average precision was 0.9523 for the original test set vs. 0.5793 for the new test set; upon updating the model using the collected human supervision data, the average precision became 0.9454 for the original test set vs. 0.8153 for the new test set. These numbers show that the new test data is sufficiently different from the original training data to have degraded performance and that re-training our vision model using the new dataset, enables the model to achieve similar performance on both the pre-deployment and post-deployment data, hinting at a reduced covariate shift. We conclude that the human-in-the-loop system enabled a sustained production throughput and resulted in a better machine learning model despite the many tail cases observed during testing.

III. CONCLUSIONS

We presented a human-in-the-loop remote robot supervision and exception handling platform where robots work and people rule. [2] We believe that this approach is better and more feasible than a "lights-out" facility *thanks to* humans. Workers are a critical part of the system to ensure a sustainably higher throughput and robustness under varying distribution of inputs.

REFERENCES

[1] "Who director-general's opening remarks at the media briefing on covid-19 - 11 march 2020," https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020, accessed: 2022-05-05.

[2] "Q3 robot orders put 2021 on track for biggest year yet," https://www.automate.org/news/q3-robot-orders-put-2021-on-track-for-biggest-year-yet.

[3] "Understanding america's labor shortage - 08 february 2022," https://www.uschamber.com/workforce/understanding-americas-labor-shortage-the-most-impacted-industries.

[4] E. Dahlin, "Are robots stealing our jobs?" *Socius*, vol. 5, p. 2378023119846249, 2019. [Online]. Available: https://doi.org/10.1177/2378023119846249

[5] N. Correll, K. E. Bekris, D. Berenson, O. Brock, A. Causo, K. Hauser, K. Okada, A. Rodriguez, J. M. Romano, and P. R. Wurman, "Analysis and observations from the first amazon picking challenge," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 172–188, 2016.

[6] H. Shimodaira, "Improving predictive inference under covariate shift by weighting the log-likelihood function," *Journal of statistical planning and inference*, vol. 90, no. 2, pp. 227–244, 2000.

[7] Y.-X. Wang, D. Ramanan, and M. Hebert, "Learning to model the tail," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30, 2017.

[1] Physical manual intervention often requires having a worker walk out to the robot cell, which can take minutes or even longer.

[2] We would like to underline that deploying one of our systems has never resulted a dismissal of employment.