Embodied AI for Smart Robotic Cells in Manufacturing Applications

Satyandra K. Gupta

Smith International Professor, Viterbi School of Engineering, University of Southern California
Co-Founder and Chief Scientist, GrayMatter Robotics

Abstract

Many manufacturing companies are facing an acute shortage of qualified workers. Deploying robotic cells is a potential solution to address this challenge. Historically robots have been deployed only in mass production applications in manufacturing. A large fraction of manufacturing is classified as high-mix manufacturing where a large variety of products are produced. Manually programming robots is not a viable solution in high-mix manufacturing applications. Robotic cells need to be powered by embodied AI to make them useful in high-mix manufacturing applications. This paper aims to build a bridge between smart manufacturing and AI communities to enable AI researchers to develop methods and tool that can be successfully deployed to realize smart robotic cells for high-mix manufacturing applications. This paper highlights key requirements for developing embodied AI for powering robotic cells for high-mix manufacturing applications. It also makes the case for approaches that combine model-based and data-driven methods to meet the needs of embodied AI in manufacturing applications and describes the role of generative AI approaches in smart manufacturing applications. Finally, it describes how AI can be used to enhance digital twins and augment human-machine interfaces in manufacturing applications.

Introduction

A large number of manufacturing companies are facing an acute shortage of qualified workers. Many manufacturing tasks such as sanding, grinding, coating, blasting, and spraying involve holding a tool in hand and moving it over a surface to make value-added changes to the surface. Most such operations can be ergonomically challenging and have a very high labor churn. This is creating a major problem for manufacturing companies in terms of maintaining capacity and completing production on time. Robotic automation technologies are needed to speed up production and achieve consistent quality.

Manufacturing applications can be broadly classified into two categories: (1) mass production, (2) high-mix. In mass production applications, the identical part are produced repeatedly. High-mix applications have the following characteristics: (1) quick part changes in production, and (2) variability in part dimensions and geometry. Robots have been successfully used in mass-production applications. However, high-mix applications have posed significant challenges for robotic automation.

In mass production applications humans program the robots. Robots simply executed pre-programmed motions. Human programming effort is amortized over a large number of parts and therefore it can be economically justified. Moreover, by carefully controlling the variation in the upstream processes, the need for online process monitoring and adaptation can either be eliminated or minimized. High-mix manufacturing applications require the part changeover to be accomplished within a few minutes, therefore, we cannot rely on humans to program robots when a new part arrives. Unfortunately, at present, the use of robots in high-mix manufacturing applications is limited, requiring humans to perform ergonomically challenging and physically demanding tasks. Using robots in these applications requires robots to autonomously execute tasks on high-level task descriptions and deliver human competitive performance. This is a challenging problem and addressing it requires leveraging the latest advances in AI. The last decade has seen significant advances in AI such as reinforcement learning, deep learning, large language models, and generative AI that are endowing robots with new capabilities (Kusiak 2020; Bhatt et al. 2021; Zonta et al. 2022; Florence et al. 2022; Stocking, Gopnik, and Tomlin 2022; Huang et al. 2022; Zhang et al. 2023; Singh et al. 2023; Malhan and Gupta 2023; Chi et al. 2024; Mishra and Chen 2024; Gregory and Gupta 2024; Al-Hussaini et al. 2024).

For the last thirty years, I have been working in area of developing robotics technologies for manufacturing applications. We have leveraged recent advances in AI to create smart robotic cells for manufacturing applications (Manyar et al. 2022, 2023b,a; Kang et al. 2024b,a; Patel et al. 2024; Shukla et al. 2024; Manyar et al. 2024; Nemlekar et al. 2024). I also have considerable prior experience in transitioning technologies developed in my lab to the industry. I am co-founder and Chief Scientist at GrayMatter Robotics, a company focused on robotic automation solutions for high-mix manufacturing applications. GrayMatter Robotics' smart robotics cells are installed in many factories in the US serving aerospace & defense, specialty vehicles, marine & boats, metal fabrication, sports equipment, and furniture & sanitary ware sectors. These cells are helping companies to increase human productivity, reduce ergonomically challenging work, compress cycle times, expand capacity and improve sustainability.

This paper is based on a series a blog posts that I have written for Forbes Tech Council to discuss challenges and opportunities in the area of AI for enabling smart manufacturing.[1] This paper aims to build *a bridge between smart manufacturing and AI communities* so that AI researchers can develop methods and tool that can be successfully deployed in manufacturing applications. This paper highlights key requirements for using robots in high-mix manufacturing applications. It also describes how AI is being used to power smart robotic cells in manufacturing applications.

Need for Embodied AI

Most of the AI that we experience in our daily lives is digital AI. It produces digital artifacts, decision recommendations, or predictions that will either be used by a human or some other digital agent. Examples include generating a cover letter for a job application using ChatGPT, recommendations for watching a movie on NetFlix, creating a painting using Dall.E, and detecting a tumor in a medical image.

A different kind of AI is being developed to manage the behavior of physical systems. For example, a robot performing sanding needs AI to operate autonomously. This AI is called embodied AI. It is tasked with one or more goals and it uses the sensor data to produce a sequence of actions that the physical system executes to achieve the goal. The embodied AI monitors the cell state using sensors and generates instructions for the robot to perform the task. Digital AI and embodied AI share some similarities and utilize many underlying techniques. However, understanding the differences between these two types of AI is critical to successfully adapting digital AI approaches for use in the context of embodied AI applications.

The risk profile of embodied AI applications is often fundamentally different from that of digital AI applications. If the accuracy of a digital AI tool is 99%, then it can tremendously boost human productivity in many applications. For example, if you are using AI to generate a 1000-word cover letter that only requires you to manually edit 10 of those words, then you will be saving a lot of time compared to writing that letter from scratch. And for the case of a recommendation engine, you would not mind if it gave you a poor suggestion for a movie once every couple of months.

Accuracy requirements for the embodied AI system are often very different due to risk considerations. For example, if a robot has a success rate of 99% on processing steps and it works on a part that requires 200 steps, then every part made by the robot will contain two errors. As a result, the part would get scrapped or would require repairs. In most manufacturing applications, a technology that does not exhibit extremely high process quality will not be considered viable.

Risk consists of two aspects: (1) probability of making an error and (2) the consequence of making errors. When the consequence of making an error is not significant, then a much higher probability of error can be tolerated. That is why an error probability of 1% will be acceptable in many digital AI applications. Conversely, many embodied AI applications demand errors probabilities better than one in a million. Reducing error probability using a purely data-driven approach requires using enormous amounts of data. In most cases, the need for data grows exponentially. Unfortunately, acquiring data from physical systems is expensive. Therefore, a different approach needs to be followed when dealing with embodied AI applications.

To address the requirements outlined above, embodied AI for manufacturing applications needs to have the following characteristics:

- *Composable from Pre-trained Modular Components:* Physical systems can have a wide variety of configurations to support their intended requirements. For example, robotic cells in manufacturing applications can come in many different configurations based on the size of the part being processed and the process being performed (e.g., sanding or blasting). Different cells may include robots with different capabilities (mobile platform mounted robot vs. gantry mounted robot), different types of sensors (depth camera vs thermal camera), and different tools (orbital sanders vs blasting nozzle) . Therefore developing universal embodied AI that works for all manufacturing applications out of the box is not likely to perform well. The AI for the system needs to be quickly synthesized from the modular components to match the sensing and actuation capabilities of the system and the work environment.
- *Adaptable Based on New Data/Context:* As new data becomes available by deploying the system, it should be possible to improve the performance of the AI by using new data. AI should be able to adapt autonomously with minimal human supervision.
- *Upgradable with Minimal Effort:* The performance of a physical system may change over time because of wear and/or updates to physical components. This may require refinement of AI to ensure that it can keep up with system evolution over time. Therefore, the embodied AI system needs to be designed to ensure that it can be updated with minimal disruption to the system operation.
- *Risk-Informed Action Recommendations:* The system should be able to estimate its confidence in recommended actions. It should be able to analyze the sensor data being used and make sure that the data being used in making prediction/recommendation comes from the same distribution that is used to train the underlying AI component. If the data appears to be out of the training distribution, then the system should lower its confidence level in its recommendations. This should prompt it to perform risk analysis to analyze the consequences of failure. If the risk is too high, then it should seek help from humans.
- *Explainable:* If the system recommended actions do not match user expectations, then the system should be able to explain the rationale used to select actions.

[1]Please see https://www.forbes.com/councils/forbestechcouncil/people/skgupta/

- *Distributed Architecture to Enable Partitioning of Computation between Edge and Cloud:* It is not possible to perform all AI computation in the cloud in the context of embodied AI. The system should be designed to ensure that computation that is sensitive to network latency can be performed on the edge. The use of edge computing may also be needed to support applications where sensor data cannot be transmitted to clouds due to privacy/security reasons.

Integrating Model-Based and Data-Driven Approaches

Data-driven AI has delivered very impressive results in wide-ranging applications such as recommendation engines, playing games, face recognition, text translation, text synthesis, and fraud detection. This type of AI uses a vast amount of data to train the system. Collecting high-quality data in many manufacturing applications takes significant time and incurs prohibitively large costs. Therefore, unfortunately, a purely data-driven AI approach is not a viable model in many manufacturing applications.

Manufacturing has a lot of known models and useful process knowledge. Rediscovering these models and knowledge using a purely data-driven approach does not make sense. However, all known models make simplifying assumptions to reduce complexity and, therefore, are approximate in nature. Embodied AI needs to exploit the known models and use a data-driven approach to augment the known models and existing knowledge based on experimental data to fill the missing gaps. This type of approach is often called Physics-informed AI. It enforces known physics-based process models (or knowledge) as a constraint in the AI system to ensure that it does not learn anything that contradicts existing models/knowledge. For example, the system can enforce a constraint that increasing pressure on the sanding tool will increase the deflection of the part being sanded. We don't need to conduct a large number of tests to learn this already-known fact. If the measured data contradicts this constraint, then it is highly likely either the sensor is malfunctioning, or the part/tool is not clamped properly.

On one hand, the physics-informed AI approach restricts the solution space and, therefore, makes the problem much more tractable from the data requirement point of view. Let us consider the problem of predicting process output based on the input. If the output is expected to increase with an increase in the input, then the underlying model space is limited and it can be trained by a smaller amount of data because we don't need to consider arbitrarily complex models. On the other hand, using physics-informed AI requires more complex representations and associated methods to handle constraints to produce acceptable computational performance. We cannot use a simple neural network and train it with observed input and output data. If we did not explicitly enforce process model constraints, then there is no guarantee that the learned model would preserve the process constraint if the output used during training is noisy.

Here are a few representative use cases for using Physics-informed AI in manufacturing applications.

- Defect detection is an essential ingredient of smart manufacturing. Machine learning has emerged as a powerful technique for analyzing and classifying images. However, collecting a large number of images of physical defects needed to train a machine learning system is not possible. An alternative is to develop a pipeline for generating photo-realistic images. In this case, a physics-based process model can be utilized to generate realistic defects in virtual models. These virtual models can be used to generate photo-realistic images. Recent work has demonstrated that a training process that utilizes a combination of photo-realistic synthetic images and real images of defects works well in practice.
- To efficiently and accurately build part models, a machine learning-based approach is being utilized for predicting sensor performance (e.g., measurement errors). This approach needs to use known models of sensor performance during the training process. For example, if the sensor is expected to produce a higher amount of error when imaging from a larger distance, then this information is used during the training phase to ensure that the image is acquired from the right distance. This approach enables the acquisition of good-quality data for improved decision-making.
- Robots often require body-mounted hoses and cables for the cell to function. These appendages may restrict the motion of the robot. Applying overly conservative restrictions makes the robotic cell inefficient. On the other hand, not applying appropriate limits may result in damage to cables/hoses and force the cell to shut down. A learning framework can enable users to estimate robot motion limits when flexible components are attached to the robot. This approach can use a physics-based model of cables and hoses bending, twisting, and stretching to accelerate the learning process.
- A smart cell should be capable of building process models for new materials by autonomously conducting experiments. While the exact quantitative relationship between the input process parameters and process performance may not be known, often qualitative relationships between many variables are known (i.e., increasing torch velocity decreases weld height). We can utilize loss functions during the training phase that penalize deviations from known process constraints. This approach can enforce known knowledge and accelerate the model-building process.
- Smart robotic cells need to use advanced prognostics and health management to ensure a higher level of reliability. We need to utilize known causal relationships during the inferencing process to ensure that we can accurately predict the system state based on the process monitoring data. For example, we can utilize the force and vision data to hypothesize the cause of accelerated tool wear in robotic sanding.

Leveraging Generative AI in Manufacturing

Generative AI methods have been successfully used to generate a variety of outputs, including images, texts, audio, and videos. Recent advances in Generative AI (e.g., ChatGPT, DALL-E 2) are expected to revolutionize the future of work in many industry sectors.

Labor shortages are requiring robotic solutions to be deployed rapidly. Unfortunately, deployment of robotic takes a significant amount of human effort due to time needed to write software and test the system. Increasing complexity of robotic systems is aggravating this problem. Unfortunately, the availability of human expertise can become a bottleneck in robotic cell deployment.

The robotics community has always been at the forefront of leveraging the latest AI advances. So, a natural question is: how will generative AI concepts and tools can be used by the robotics community to accelerate the deployment of the next generation of smart robotic cells? Recent efforts are showing early signs of success in using generative AI in robotics applications. The list below highlights opportunities for using generative AI in smart robotic cells.

- *Generating Robot Motion from Natural Language Descriptions:* Robots often need to perform complex motions to successfully execute a task. Consider the example of sanding where the robot needs to move the sanding tool in a complex motion pattern to produce a scratch-free surface finish. Many process experts would prefer a new modality to generate robot motion based on natural language description of the motion. Generative AI now offers the capability to generate code from the text description, which enables humans to communicate with robots in a more natural, time-efficient manner and automatically create robot motion. This means human experts with no programming experience can get the robots to perform the right kind of motion. The elimination of robot programming is expected to remove bottlenecks and can significantly speed up this task.
- *Task Planning:* Many applications require robots to perform complex tasks. For example, consider the task of replacing a motor of a cooling fan in a control box. This requires the top-level task to be decomposed into much simpler subtasks and to determine the sequence of tasks. Once the task sequence is determined, the robot can generate motions for the simple tasks and execute this task. Traditionally, a task planner needs to be developed for each specialized domain and the addition of new objects or processes requires an update to the planner. Tasks planners often struggle to deal with new failure modes and therefore recovering from failures becomes challenging during task planning. Large Language Models (LLM) have become the foundation of many Generative AI techniques and can be used to identify sequences of atomic tasks needed to perform complex tasks. With the latest advancements in LLMs we can pose a query such as, "Provide step-by-step directions to obtain a tool from a locked shelf." and generate a sequence of various subtasks necessary to perform the overall task. Once atomic tasks have been identified, the robot can use a motion planner to generate the motion to execute the task. LLMs can be extremely useful in automatically generating task sequences based on common sense knowledge and using them can eliminate the need for developing domain-specific task planners.
- *Generating Synthetic Scenarios for Utilizing Reinforcement Learning:* Reinforcement learning has emerged as a useful tool for robots to acquire new skills. This type of learning requires the robot to train in a simulation using a trial-and-error approach. Manually generating a large number of scenarios for reinforcement learning is highly time-consuming and may still not cover all the relevant cases or provide adequate diversity. Generative AI technology can be used to generate distinct, synthetic scenarios to aid reinforcement learning in simulations. By enabling the robots to train on a diverse set of examples and scenarios, Generative AI can enable robots to learn new behaviors, strategies, or responses based on the learned patterns and context. This allows robots to develop robust and adaptive behavior when faced with new situations and perform tasks efficiently.
- *Failure Detection and Recovery for Resilient Operation:* The ability to recover from failure is needed on challenging tasks. Whenever testing reveals that the system is not able to recover from failure, the system needs to be improved. This requires an ability to detect failures and take recovery actions. Take, for example, an assembly operation where a previously-installed component is damaged due to a subsequent assembly operation. To proceed with the assembly, the damaged part will need to be replaced through a series of many disassembly steps. Manually coding all possible contingency actions is intractable. By learning the patterns and characteristics of normal operation, the AI model can generate synthetic data that deviates from the norm. These generated anomalies can be used to train a failure detection and recovery system to detect and respond to unexpected or abnormal situations, helping to improve its robustness and fault tolerance.
- *Automating Tool Design:* Many material handling tasks in manufacturing require the use of end-effectors and, depending on the task, these end-effectors need to be customized for optimal performance. In the case of picking and transporting a highly-compliant large sheet, designing a complex end-effector can take multiple days. If significant human effort is needed during the end-effector design process, the availability of humans can become a bottleneck. It will be much better for a Generative AI system to design the end-effector and a human expert to approve the design. Generative AI is showing promise in a wide variety of design tasks, and therefore, it can be utilized to automate the design of an end-effector for a new task. By reducing the human time required to design new end-effectors tailored to a specific task, autonomous robots can be deployed more rapidly.

Enhancing Digital Twins

A digital twin is a digital counterpart of a real-world system. The digital representation used in digital twins is created using data from sensors and Internet of Things devices, and it mimics the physical object or system in real-time. A digital model (e.g., 3D model of an object) created during the design stage to perform simulations is not a digital twin. Simulations prove invaluable during the design phase, allowing for the exploration of various design options before making a decision. However, digital models used during design represent the idealized state and not the actual physical system state. Similarly, 3D models created by reverse engineering processes (sometimes referred as digital shadows) are not digital twins because they cannot influence the physical system. The digital twin performs continuous model updates using sensor data to mirror the current state of the physical system. Therefore, information flows from the physical system to the digital twin. A digital twin is also used to influence the operation of the physical system. Therefore, information also flows from the digital twin to the physical system. This two-way information flow makes digital twins different from purely digital models.

Digital twins are being used in manufacturing in the following ways:

- Digital twins are being used to provide information to task planners and schedulers to make decisions about the next tasks to perform based on the current state of the manufacturing system.
- Digital twins monitor the condition and performance of machines and equipment in real-time and use this data to predict when maintenance is needed, reducing unexpected downtime and preventing machine breakdowns.
- Digital twins are being used for identification of defects and perform real-time quality control.
- When the system enters an error state, digital twins can be used to diagnose the problem and recommend the necessary actions to bring the system back to the normal state.
- By analyzing manufacturing process data, digital twins are able to identify areas for optimization and recommend changes to realize improvements.
- Digital twins can be used to optimize the manufacturing operations in real-time to support on-demand production of personalized products in a cost-effective manner.

AI is increasingly being used to augment capabilities of digital twin technology and create new opportunities. Here are a few examples:

- Sensors in a robotic surface finishing cell can be used to build a model of the part placed in the cell. This eliminates the need for using expensive part holding devices. To eliminate the possibility of collision, tool paths need to account for uncertainties in part models created using sensors. Therefore, the digital twin needs to use AI for predicting uncertainties in part models.
- Simulations are necessary to generate optimal plans for performing finishing operations. Traditional simulations lack the speed required when dealing with part models with uncertainties. Machine learning is being used to create fast simulations based on neural networks, endowing digital twins with new planning and prediction capabilities. Tool motion cannot be safely executed using position control in the presence of uncertainties. An AI-based controller needs to be used that uses force feedback to dynamically adapt tool motion to prevent part damage and maximize tool life.
- Robots performing surface finishing often require hoses and cables connected to the tool. These appendages may restrict the motion of the robot. A digital twin enables building models of all the peripherals (e.g., cables, and hoses) mounted on the robot. Using an AI-based approach enables the system to estimate robot motion limits based on the estimated states of flexible peripherals attached to the robot. These models can be used by the planner to ensure that the robot does not collide with these peripherals.
- AI-based prognostics and health management can be used by digital twins to ensure that the onset of adverse events can be automatically detected and corrective actions can be taken. For example, the digital twin can utilize the force and vision data to determine the cause of rapid tool wear in robotic finishing and take corrective measures to prevent it. The need for continuous improvements and process optimization requires the system to build its own process model. A digital twin can utilize machine learning to safely conduct autonomous experiments and build a process model for selecting the right process parameters for sanding a new material.

Augmenting Human-Machine Interfaces

Humans need to interact with machines in manufacturing applications to task the machines, initiate production, monitor the process, and troubleshoot problems. Historically, human-machine interfaces in the industrial setting have not been very user friendly. Humans often interact with industrial machines by pressing buttons, turning knobs, and typing on keyboards. These traditional interfaces are hard to master and can be quite frustrating for a new user. Even an experienced user may find such interfaces inefficient and not very helpful when trying to troubleshoot a failure.

Improved human-machine interfaces have potential to change the user experience and improve efficiency of the manufacturing operations. Well-designed interfaces can also reduce the probability of errors and accelerate the training process for new users.

Recent advances in AI are providing new ways for humans to interact with machines. Here are possible ways AI is expected to revolutionize human-machine interfaces.

- *New Interaction Modalities:* Recent advances in natural language processing and human speech understanding are enabling new modalities for humans to interact with machines. It is now possible for humans to talk to machines in natural language. This can reduce the need for people to learn new interfaces. Recent advances in image and video understanding is also enabling machines to understand human gestures. Machines can not only understand hand gestures, but also decipher the meaning of nodding of heads and small body movement. These can be very useful in understanding human intent. Recent advances are also showing that machines can also decode facial expressions. These advances will enable machine designers to offer new nodes of interactions. New interaction modalities can also improve accessibility and user experience, especially for individuals with limited mobility. Some people believe that new advances in brain-machine interfaces will enable humans to interact with machines by just thinking about the tasks. This technology is showing promise, but a lot of advances will be necessary before the promise of this technology can be harnessed in actual factory settings.

- *Preventing Cognitive Overloading by Delivering Relevant Information:* Modern machines deploy a lot of sensors and generate a lot of data. Providing all of this information to a user can easily overwhelm them by inducing cognitive overload. It is important to be selective about what information is presented to the user based on the interaction context. For example, if a user is trying to troubleshoot a quality issue, then it will be useful for them to only focus on the information that is likely to play a role in the failure. AI can play an important role in sifting through all the information generated by sensors and only delivering the relevant information to the user in the right format. AI can also play a role in monitoring the cognitive load of the machine operator and make suitable adjustments to prevent cognitive overloading.
- *Generating Alerts and Warning:* Sometimes humans might make mistakes and ask the machine to perform an unsafe operation. By monitoring human behaviors and the task state, the machine can predict occurrences of future unsafe situations and alert humans. AI can be used to simulate possible futures and perform risk assessment by accounting for uncertainties. AI can be used to assess risk in an objective manner and issue appropriate alerts and warnings.
- *Improving Training Efficiency:* Training new users to effectively utilize machines takes significant time and resources with traditional interfaces. This is becoming a major challenge in today's era of high labor churn. Most traditional interfaces were not designed with ease of training in mind. AI-powered interfaces can provide real-time feedback, guidance, and assistance to users during the training phase, helping them navigate complex tasks or troubleshoot problems effectively. Active learning can monitor user's mastery of a task and enable them to focus on areas where they need more practice. Moreover, virtual assistants equipped with AI can offer interactive support and tutorials, improving user productivity and learning outcomes during the training phase.
- *Learning Human Preferences and Generating Customized Interfaces:* Different people have different preferences. Therefore, to deliver the best experience to the operators, we need interfaces that match the user preferences. This requires moving away from the concept of fixed interfaces and instead focus on generating customized interfaces for different users based on their preferences. By analyzing user behavior and historical data, AI algorithms can predict users' preferences, needs, and intentions, allowing interfaces to anticipate user actions and provide personalized recommendations or assistance. This proactive approach can significantly enhance user satisfaction and efficiency. Recent advances in AI can be leveraged to create new interfaces based on individual user preferences, abilities, and past interactions. Using AI to deliver custom interfaces ensures that user-friendly and efficient interfaces can be delivered for a diverse user demographic in a cost-effective manner. This technology can also be leveraged to support workers with physical and intellectual disabilities. Customized interfaces can accommodate an individual's accessibility requirements by supporting appropriate interaction modalities.

Conclusions

Smart robotic cells are expected to transform high-mix manufacturing and address the labor shortage problem. Realizing such robotic cells will require new advances in embodied AI to address the risk considerations in manufacturing applications.

Embodied AI needed in manufacturing applications cannot be realized as a monolithic system running on the cloud. Embodied AI in the context of manufacturing should be viewed as a complex system that involves interactions among multiple AI components. The system should use the right functional decomposition to ensure that it is able to achieve the desired trade-off in performance and modularity. Many different AI approaches exist. It is unlikely that a single approach will suffice to deliver the desired performance. Therefore, each functional block should use the right AI approach by carefully considering pros and cons. Therefore, having the right system architecture in the embodied AI system is the key to success in manufacturing applications.

Generating a large amount of data is not possible in manufacturing applications from time and cost perspective. The embodied AI should be designed such that it can be trained with limited data generated by physical experiments. An approach that combines model-based and data-driven method is needed to successfully deploy embodied AI in manufacturing applications.

Deploying robotic cells in complex applications currently requires significant human effort. The availability of human resources needed to get this accomplished often emerges as a bottleneck and can cause delays in deployment. Generative AI is offering new tools to reduce the human expertise needed to deploy robots in manufacturing applications.

AI-powered digital twins are ushering a new era of smart manufacturing by lowering costs, reducing errors, improving quality, increasing performance, and reducing the environmental footprint.

Humans are important parts of manufacturing operations and therefore human-system integration issues need to be proactively addressed during the system design. AI can be used to revolutionize human-machine interfaces in manufacturing applications by fostering flexibility in processes and promoting more intuitive interactions for workers.

Acknowledgments

I would like thank all of my collaborators and students who have inspired me and shaped ideas presented in the article.

References

Al-Hussaini, S.; Guan, Y.; Gregory, J. M.; Pollard, K.; Khooshabeh, P.; and Gupta, S. K. 2024. Assessing the Impact of Alerts on the Human Supervisor's Decision-Making Performance in Multi-Robot Missions. *ACM Transactions on Human-Robot Interaction*, 14(1): 1–40.

Bhatt, P. M.; Malhan, R. K.; Rajendran, P.; Shah, B. C.; Thakar, S.; Yoon, Y. J.; and Gupta, S. K. 2021. Image-Based Surface Defect Detection Using Deep Learning: A Review. *Journal of Computing and Information Science in Engineering*, 21(4).

Chi, C.; Xu, Z.; Feng, S.; Cousineau, E.; Du, Y.; Burchfiel, B.; Tedrake, R.; and Song, S. 2024. Diffusion Policy: Visuomotor Policy Learning via Action Diffusion. *The International Journal of Robotics Research*.

Florence, P.; Lynch, C.; Zeng, A.; Ramirez, O. A.; Wahid, A.; Downs, L.; Wong, A.; Lee, J.; Mordatch, I.; and Tompson, J. 2022. Implicit behavioral cloning. In *Conference on Robot Learning*, 158–168. PMLR.

Gregory, J. M.; and Gupta, S. K. 2024. Opportunities for Generative Artificial Intelligence to Accelerate Deployment of Human-Supervised Autonomous Robots. *Proceedings of the AAAI Symposium Series*, 2(1): 177–181.

Huang, W.; Abbeel, P.; Pathak, D.; and Mordatch, I. 2022. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In *International Conference on Machine Learning*, 9118–9147. PMLR.

Kang, J. H.; Dhanaraj, N.; Manyar, O.; Wadaskar, S.; and Gupta, S. K. 2024a. A Task Allocation and Scheduling Framework to Facilitate Efficient Human-Robot Collaboration in High-Mix Assembly Applications. In *International Manufacturing Science and Engineering Conference*.

Kang, J. H.; Dhanaraj, N.; Wadaskar, S.; and Gupta, S. K. 2024b. Using Large Language Models to Generate and Apply Contingency Handling Procedures in Collaborative Assembly Applications. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 15585–15592.

Kusiak, A. 2020. Convolutional and generative adversarial neural networks in manufacturing. *International Journal of Production Research*, 58(5): 1594–1604.

Malhan, R.; and Gupta, S. K. 2023. The Role of Deep Learning in Manufacturing Applications: Challenges and Opportunities. *Journal of Computing and Information Science in Engineering*, 23(6): 060816.

Manyar, O. M.; Cheng, J.; Levine, R.; Krishnan, V.; Barbic, J.; and Gupta, S. K. 2022. Physics Informed Synthetic Image Generation for Deep Learning based Detection of Wrinkles and Folds. *ASME Journal of Computing and Information Science in Engineering*, 1–18.

Manyar, O. M.; McNulty, Z.; Nikolaidis, S.; and Gupta, S. K. 2023a. Inverse Reinforcement Learning Framework for Transferring Task Sequencing Policies from Humans to Robots in Manufacturing Applications. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 849–856. London, UK.

Manyar, O. M.; Narayan, S. V.; Lengade, R.; and Gupta, S. K. 2023b. Physics-Informed Learning to Enable Robotic Screw-Driving Under Hole Pose Uncertainties. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2993–3000.

Manyar, O. M.; Ye, H.; Sagare, M.; Mayya, S.; Wang, F.; and Gupta, S. K. 2024. Simulation-Assisted Learning for Efficient Bin-Packing of Deformable Packages in a Bimanual Robotic Cell. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Abu Dhabi, UAE.

Mishra, U. A.; and Chen, Y. 2024. Reorientdiff: Diffusion model based reorientation for object manipulation. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 10867–10873. IEEE.

Nemlekar, H.; Sivagnanadasan, N.; Banga, S.; Dhanaraj, N.; Gupta, S. K.; and Nikolaidis, S. 2024. Selecting Source Tasks for Transfer Learning of Human Preferences. *IEEE Robotics and Automation Letters*, 9(8): 6896–6903.

Patel, R.; Kanyuck, A.; McNulty, Z.; Yu, Z.; Carlson, L.; Heng, V.; Johnson, B.; and Gupta, S. K. 2024. Automated Plan Refinement for Improving Efficiency of Robotic Layup of Composite Sheets. In *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*, 3132–3139.

Shukla, R.; Yu, Z.; Moode, S.; Manyar, O. M.; Wang, F.; Mayya, S.; and Gupta, S. K. 2024. Performing efficient and safe deformable package transport operations using suction cups. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Abu Dhabi, UAE.

Singh, I.; Blukis, V.; Mousavian, A.; Goyal, A.; Xu, D.; Tremblay, J.; Fox, D.; Thomason, J.; and Garg, A. 2023. ProgPrompt: program generation for situated robot task planning using large language models. *Autonomous Robots*, 1–14.

Stocking, K. C.; Gopnik, A.; and Tomlin, C. 2022. From Robot Learning To Robot Understanding: Leveraging Causal Graphical Models For Robotics. In *Conference on Robot Learning*, 1776–1781. PMLR.

Zhang, J.; Zhang, J.; Pertsch, K.; Liu, Z.; Ren, X.; Chang, M.; Sun, S.-H.; and Lim, J. J. 2023. Bootstrap Your Own Skills: Learning to Solve New Tasks with Large Language Model Guidance. In *7th Annual Conference on Robot Learning*.

Zonta, T.; da Costa, C. A.; Zeiser, F. A.; de Oliveira Ramos, G.; Kunst, R.; and da Rosa Righi, R. 2022. A predictive maintenance model for optimizing production schedule using deep neural networks. *Journal of Manufacturing Systems*, 62: 450–462.