https://doi.org/10.2352/EI.2025.37.15.AVM-105

Enhancing Robotic Navigation with Large Language Models

Xunyu Pan and Jeremy Perando
Department of Computer Science and Information Technologies, Frostburg State University, Frostburg, Maryland, USA

Abstract

Robotics has traditionally relied on a multitude of sensors and extensive programming to interpret and navigate environments. However, these systems often struggle in dynamic and unpredictable settings. In this work, we explore the integration of large language models (LLMs) such as GPT-4 into robotic navigation systems to enhance decision-making and adaptability in complex environments. Unlike many existing robotics frameworks, our approach uniquely leverages the advanced natural language and image processing capabilities of LLMs to enable robust navigation using only a single camera and an ultrasonic sensor, eliminating the need for multiple specialized sensors and extensive pre-programmed responses. By bridging the gap between perception and planning, this framework introduces a novel approach to robotic navigation. It aims to create more intelligent and flexible robotic systems capable of handling a broader range of tasks and environments, representing a major leap in autonomy and versatility for robotics. Experimental evaluations demonstrate promising improvements in the robot's effectiveness and efficiency across object recognition, motion planning, obstacle manipulation, and environmental adaptability, highlighting its potential for more advanced applications. Future developments will focus on enabling LLMs to autonomously generate motion profiles and executable code for tasks based on verbal instructions, allowing these actions to be carried out without human intervention. This advancement will further enhance the robot's ability to perform specific actions independently, improving both its autonomy and operational efficiency.

Introduction

The advent of large language models (LLMs) has introduced a paradigm shift in artificial intelligence, driving significant advancements in natural language understanding and reasoning. Multimodal LLMs, such as GPT-4 [1] and Gemini [2], leverage transformer architectures with extensive training on internet-scale datasets, equipping them with the ability to generate contextually relevant and semantically rich outputs. Additionally, key innovations, including in-context learning [3], chain-of-thought prompting [4], and multimodal integration, have extended the utility of LLMs beyond traditional text-based applications, enabling more versatile and sophisticated reasoning. By combining textual and visual inputs, these models excel at interpreting complex, unstructured data, leading to improvements in tasks such as information retrieval, content generation, human-robot interaction, scene understanding, decision support, sentiment analysis, and context-based object recognition. The ability to process both structured and unstructured data enables LLMs to bridge the gap between perception and planning, opening avenues for novel applications across various domains.

In robotics, the three core capabilities of *perception*, *planning*, and *control* are essential for effective navigation and action. Traditionally, these capabilities have been enabled through specialized sensors and pre-programmed responses. Although such approaches perform well in structured environments, they often struggle in dynamic and unpredictable settings due to their limited adaptability and reliance on predefined rules. Recent surveys [5, 6] suggest that LLMs can overcome these limitations by allowing robots to interpret unstructured data, such as visual scenes and natural language instructions, and make real-time decisions with greater adaptability. In contrast to conventional sensor-heavy systems, integrating LLMs offers a more flexible framework for robotic navigation, enhancing perception, planning, and control through semantic reasoning and multimodal understanding.

In this work, we present a novel intelligent robotic system that integrates the advanced processing capabilities of LLMs to enhance navigation and environmental adaptability. Compared to traditional methods, this approach reduces dependence on multiple sensors and extensive programming, creating a more adaptable and efficient system for navigation in complex, real-world scenarios. The robot developed for this study is a custom-built platform powered by a Raspberry Pi 4 running Raspberry Pi OS. Its control software, written in Python, interfaces with the OpenAI GPT-4 API to perform environmental analysis. Equipped with a camera, the robot captures images and sends them to GPT-4, which provides feedback such as object recognition, weight estimation, and motion planning. The hardware also features a claw for obstacle manipulation and an ultrasonic sensor for precise distance measurement, compensating for GPT-4's limitations in estimating distance from a single image. Actions are then executed based on the responses from the LLM, allowing the robot to operate effectively in dynamic environments.

The performance of the proposed robotic system was evaluated with a variety of objects classified as either light and movable or heavy and immovable. For example, an apple was used to represent a light object, while a dictionary represented a heavy one. The robot accurately determined whether each object could be relocated and executed the appropriate action accordingly. Further testing with jar candles (heavy) and decks of cards (light) reinforced the system's accuracy and robustness. The experimental results demonstrate the potential of LLMs in developing intelligent robotic systems. By reducing reliance on specialized sensors and extensive pre-programming, LLMs enable robots to operate with greater intelligence and adaptability. These models enhance the dynamic planning capabilities of robotic systems, allowing them to respond flexibly to new and unpredictable situations, thereby significantly improving their autonomy and operational efficiency.

Figure 1. *Experimental setup showing the robot positioned in front of a crumpled paper towel and an apple, used to evaluate its ability to identify and interact with various obstacles.*

Related Work

Machine learning models, particularly deep learning frameworks, have long been studied for their potential to support robotics. Some task-specific robotic applications leverage end-to-end systems, where a unified model maps raw inputs (e.g., sensor readings, images, or text commands) directly to outputs (e.g., motor commands, navigation strategies) without intermediate steps. For instance, Vision-based Navigation Grounding (ViNG) [7], a Vision-Navigation Model (VNM), directly maps raw visual inputs to navigation actions, simplifying pipelines and enabling effective performance in open-world environments [8, 9]. Similarly, Contrastive Language–Image Pretraining (CLIP) [10], a Visual-Language Model (VLM), aligns images and text in a shared embedding space using contrastive learning, supporting robotics applications such as perception and navigation. In specific domains, Vision-and-Language Navigation (VLN) models map natural language instructions to navigation actions in simulated [11] or real-world [12] environments. Meanwhile, Vision-and-Language Action (VLA) models, such as PaLM-SayCan [13] and RT-2 [14] from Google DeepMind, process multimodal inputs directly into robotic actions, excelling in tasks like object manipulation and human-robot collaboration. These task-specific models are highly impactful for robotics tasks requiring seamless and rapid decision-making, but involve trade-offs in data requirements, interpretability, and generalization.

To address these challenges, Large Model Navigation (LM-Nav) [15] demonstrates the power of integrating pre-trained models for effective robotic navigation. It combines three key components: GPT-3 [16], an early-stage LLM, to parse user instructions into navigational landmarks; CLIP, a VLM trained to align textual landmarks with visual observations; and ViNG, a VNM designed to compute distances and generate navigation actions. A significant advantage of LM-Nav is its ability to generalize across diverse environments without requiring fine-tuning or human annotations. This capability arises from its modular design, where each component, including GPT-3, CLIP, and ViNG, is pre-trained on large, task-specific datasets. This design enables robust integration and performance in complex, real-world scenarios. The seamless integration of these components allows LM-Nav to execute natural language instructions effectively in dynamic and unstructured environments. This highlights the potential of combining LLMs with vision and navigation models to advance robotic autonomy.

Recent advancements in LLMs have unlocked new possibilities for robotics by enhancing core capabilities such as perception, planning, and control. These models enable robots to interpret unstructured data, including visual scenes and natural language commands, for dynamic decision-making and interaction. OpenAI's GPT-4 [1], with its multimodal processing and robust contextual reasoning, has shown significant potential in robotics, particularly in tasks involving natural language commands and environmental interpretation. Similarly, Gemini [2] represents the latest advancement in Google's AI models, building on the foundational strengths of PaLM 2 [17] and PaLM-E [18]. While PaLM 2 specializes in advanced language understanding and reasoning, PaLM-E expands these capabilities by integrating vision and language. Gemini takes these advancements further, seamlessly combining visual and textual inputs to deliver enhanced multimodal functionality, making it particularly well-suited for applications such as human-robot collaboration. Meta's LLaMA-3 [19, 20], known for its efficiency and scalability, offers a lightweight solution for robotics platforms with constrained computational resources, particularly for tasks requiring robust language processing.

LLMs are increasingly utilized in robotic navigation for their ability to interpret unstructured data and facilitate natural language interactions. Collectively, these advancements represent a transformative shift in robotics, enabling more intelligent, adaptable, and autonomous systems capable of operating effectively in complex and dynamic environments.

Methods

Robotics plays a crucial role in modern life, with applications spanning industrial automation, healthcare, transportation, smart agriculture, and home assistance. As robotic systems become increasingly prevalent, there is a growing demand for intelligent systems capable of effectively controlling robots to operate in dynamic and unstructured environments.

Built on the capabilities of advanced LLMs, this study introduces a novel approach by integrating GPT-4 into a robotic system for real-time environmental analysis and motion planning, as shown in Figure 1. Unlike traditional robotic systems that rely on predefined rules, requiring fine-tuning and significant programming to perform specific tasks, this approach leverages the generalization capabilities of pre-trained LLMs to dynamically interpret and adapt to new scenarios. The proposed system integrates

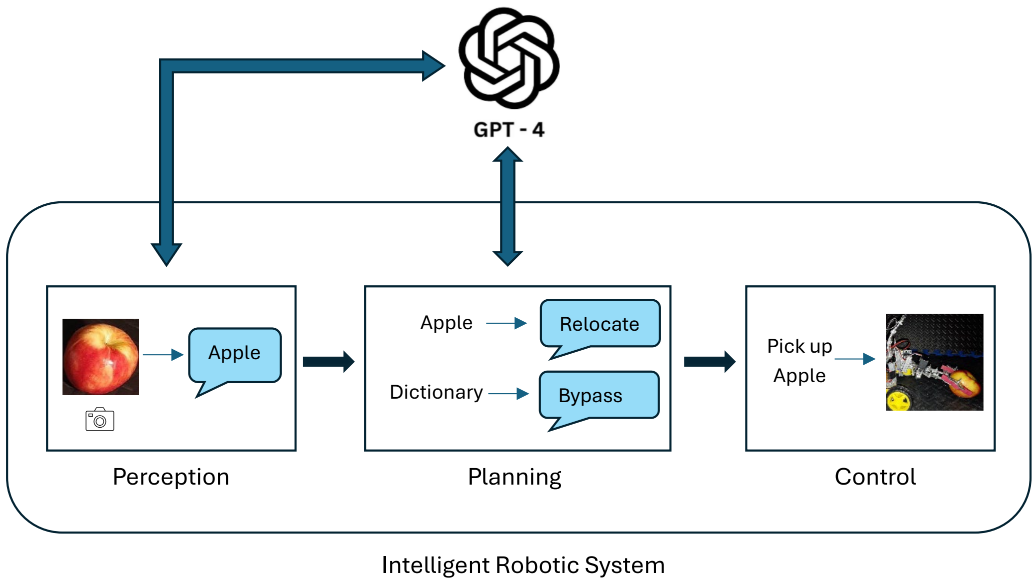


Figure 2. *Overview of the intelligent robotic system's decision-making process. The system integrates perception, planning, and control modules to identify obstacles and determine appropriate actions. Light objects, such as an apple, are relocated, while heavy objects, such as a dictionary, are bypassed. The robot executes these decisions to navigate efficiently toward the target object.*

multimodal inputs, which combine visual and textual analysis to enhance adaptability and enable complex interactions, such as object recognition, motion planning, and task execution.

The robot developed in this study is a custom-built system powered by a Raspberry Pi. Its control software, written in Python, interfaces with the OpenAI GPT-4 API for environmental analysis and motion planning. The system captures images through a camera and sends them to GPT-4, which responds with recommendations such as object identity, estimated weight, and planned motion. The hardware features a claw for object manipulation and an ultrasonic sensor for precise distance measurement, addressing GPT-4's limitations in estimating distance from a single image. The robot executes actions based on the feedback provided by the LLM.

As shown in Figure 2, the LLM-based intelligent robotic system performs navigation tasks through three successive modules: (a) *Perception Module*: captures and interprets the environment to extract analyzable data; (b) *Planning Module*: converts this data into navigation strategies and motion plans; and (c) *Control Module*: executes the plans by translating them into precise motor commands and real-time adjustments.

Perception Module

The perception module is responsible for capturing and interpreting the robot's surroundings to facilitate effective navigation, using a camera and an ultrasonic sensor as its primary inputs. At the start of navigation, the camera captures an image, which is processed through the OpenAI GPT-4 API for object recognition, aligning identified objects with textual descriptions, similar to the functionality of VLMs like CLIP. By leveraging GPT-4's advanced language and image processing capabilities, the system dynamically identifies objects and environmental features, improving adaptability in complex and unstructured settings.

The ultrasonic sensor complements the camera by providing precise distance measurements, addressing GPT-4's inherent limitations in estimating distances from a single image. It continuously measures the real-time distance between the robot and nearby objects to assist with navigation control. This hybrid setup ensures reliable perception, particularly in scenarios where visual inputs alone may be insufficient, such as in low-light conditions or cluttered environments.

Planning Module

The planning module bridges the gap between perception and control by translating interpreted data into actionable navigation strategies. Based on the identified object's image and the robot's hardware parameters, GPT-4 determines the most appropriate action for each scenario, ensuring precise and context-aware responses to environmental conditions. These actions are then structured into a motion plan aligned with task objectives, such as moving straight, turning left or right, or relocating obstacles. For instance, if the task involves retrieving a target object, the module identifies the most efficient path, either by bypassing heavy obstacles or by positioning the claw to manipulate light ones, to successfully approach the object. The integration of GPT-4's advanced reasoning capabilities enables the system to dynamically adjust its plans, allowing it to handle unexpected scenarios with minimal pre-programming.

Control Module

The control module executes actions from the motion plan created by the planning module, converting them into precise motor commands. It operates through a custom Python-based control framework that connects directly to the robot's hardware, including motors for movement and a claw for object manipulation. The ultrasonic sensor provides fine-grained distance control, essential for tasks requiring precise proximity adjustments, such as picking up obstacles or navigating tight spaces. Additionally, a gyro

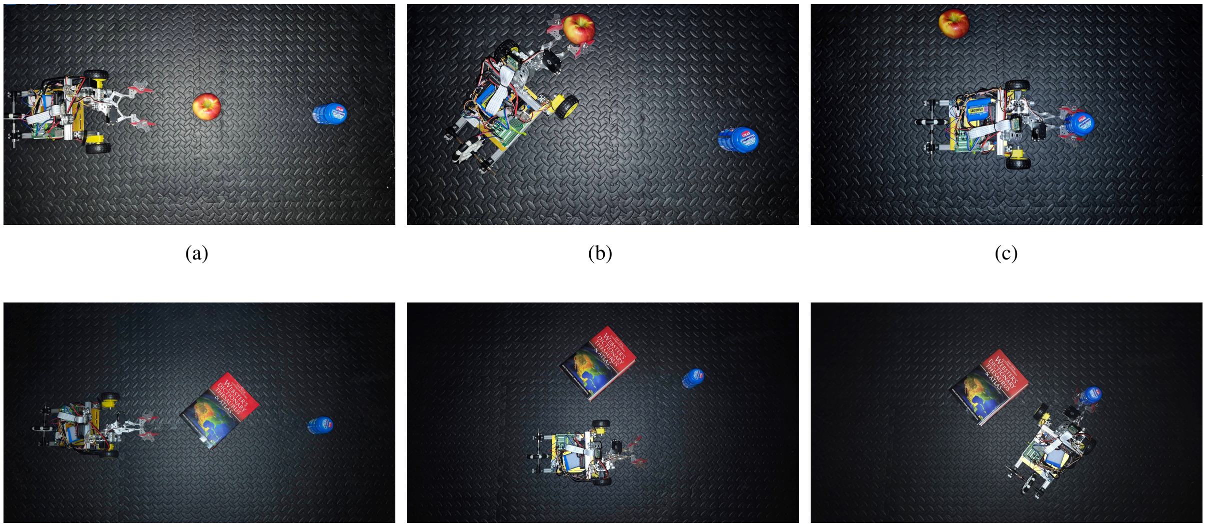


Figure 3. *Demonstration of the robotic system's motion planning process when navigating toward a target object while encountering obstacles. (a) The robot approaches an obstacle and identifies it as an apple. (b) It estimates the apple to be light, successfully grasps it, and relocates it to clear the path. (c) The robot continues forward and finally grasps the target object. (d) The robot encounters an obstacle and identifies it as a dictionary. (e) It estimates the dictionary to be too heavy to relocate and bypasses it. (f) The robot proceeds forward and ultimately grasps the target object.*

sensor monitors the robot's heading, keeping it on course and enhancing overall navigation accuracy.

Intelligent Robotic System

Together, the perception, planning, and control modules enable the proposed intelligent robotic system to perform complex tasks, including object recognition, weight estimation, and motion planning, even in dynamic and unstructured environments. By leveraging advanced LLM capabilities and integrating multiple sensing modalities, the system achieves a high degree of autonomy and adaptability with minimal reliance on task-specific programming.

Results

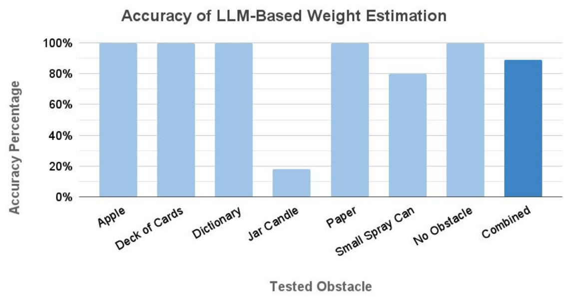
The proposed intelligent robotic system, developed in Python, integrates the OpenAI GPT-4 API for natural language processing and image analysis. It runs on a Raspberry Pi 4b with Raspberry Pi OS Lite (Debian 12) and features a high-mounted Raspberry Pi Camera v1 for image capture and a low-mounted ultrasonic sensor for precise distance measurement. The robot is equipped with two 5V DC motors for movement and a servo-controlled claw for object manipulation. Programs are launched remotely via SSH for secure command-line control. Designed for real-time interaction in dynamic environments, the system seamlessly integrates hardware and software for efficient task execution. Experiments were conducted in a controlled laboratory environment to simulate real-world challenges and validate the system's performance and reliability.

We first evaluated the robot's navigation capabilities through qualitative demonstrations. The robot was assigned the task of reaching the target object, a mint container, while navigating around various obstacles. In the first scenario, illustrated in panels (a) to (c) of Figure 3, the robot identified an apple as a light object, successfully picked it up, and relocated it to the designated area. In the second scenario, illustrated in panels (d) to (f) of Figure 3, the robot recognized a dictionary as a heavy object and bypassed it by selecting an alternative path.

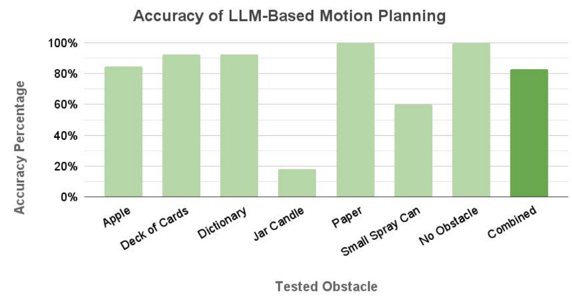
The robot's ability to classify objects as "light/relocate" or "heavy/bypass" was quantitatively evaluated across 100 trials with diverse test objects, including an apple (13 trials), a deck of cards (17 trials), a dictionary (13 trials), a jar candle (11 trials), a piece of paper (30 trials), a small spray can (10 trials), and no obstacle (6 trials). Each object was assessed for its weight classification using visual and textual data processed by GPT-4. As shown in Figure 4 (a), the robotic system demonstrated strong performance, achieving an overall weight estimation accuracy of 89%. It correctly classified most test objects, such as the dictionary, apple, deck of cards, and piece of paper, with 100% accuracy. However, misclassifications occurred with visually ambiguous items. For instance, the jar candle (18% accuracy) was frequently misidentified as an empty mason jar, while the small aluminum spray bottle (80% accuracy) was occasionally overestimated as heavier than it actually was.

As shown in Figure 4 (b), the robot correctly planned the optimal motion in 83% of trials. These motions included picking up and relocating light obstacles, bypassing heavy ones, and driving straight when no obstacle was present. Discrepancies occurred when the robot accurately estimated an object's weight but chose a suboptimal action. For example, in one trial, it interpreted "move object" as navigating around it rather than relocating it. In other instances, the robot was overly cautious, bypassing obstacles it could have relocated, as it was instructed to avoid objects in uncertain cases to ensure reliability.

Despite these limitations, the robot successfully reached and grasped the target object in 99% of trials, as shown in Figure 4 (c). This includes cases where it bypassed light obstacles instead of relocating them but still reached its destination. The only failure occurred when the robot attempted to move a heavy dictionary,

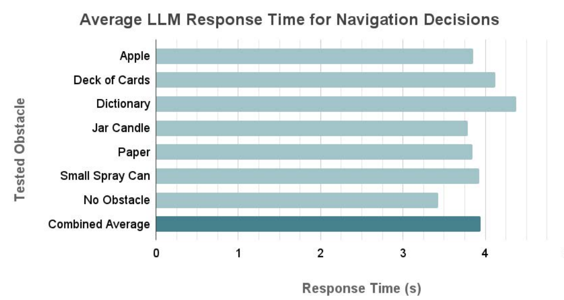


(a)



(b)

(c)



(d)

Figure 4. *A quantitative evaluation of the proposed intelligent robotic system was conducted over 100 trials using a diverse set of test objects as obstacles. The objects included an apple (13), a deck of cards (17), a dictionary (13), a jar candle (11), a piece of paper (30), a small spray can (10), and no obstacle (6). The evaluation metrics include: (a) Accuracy of LLM-based weight estimation, (b) Accuracy of LLM-based motion planning, (c) Success rate of achieved outcomes, and (d) Average LLM response time for navigation decisions.*

failed, and was unable to proceed. Finally, as shown in Figure 4 (d), the robot's decision-making time averaged 3.9 seconds per trial, suggesting the need for a faster model or alternative methods to mitigate this limitation in real-time operation.

Overall, the experimental results demonstrate that the system effectively integrates object recognition, weight estimation, and motion planning, achieving high reliability in completing navigation tasks. Future work will focus on refining the robot's decision-making process to increase the frequency of optimal actions and addressing edge cases, such as command misinterpretations, to further enhance its performance and adaptability.

Conclusions

The rapid advancement of Large Language Models (LLMs) has transformed various domains, including robotics, by enabling more intelligent and adaptable systems. Models such as GPT-4 offer advanced reasoning, natural language understanding, and multimodal capabilities, making them well-suited for robotics applications. In this work, we present an intelligent robotic system that integrates LLMs into its perception, planning, and control modules, allowing it to operate effectively in dynamic and unstructured environments. The perception module leverages GPT-4 to process visual and textual data, enabling the system to identify objects and understand the environment. The planning module uses this interpreted data to generate motion plans and navigation strategies aligned with task objectives. Finally, the control module executes these plans, translating high-level instructions into precise motor commands to ensure robust task performance. The system offers several advantages, including the ability to generalize across diverse scenarios, adapt to unforeseen situations, and reduce the need for extensive task-specific programming. Experimental results demonstrate high accuracy in object classification, effective decision-making, and the ability to handle complex interactions, highlighting the potential of LLM-based robotic systems for practical applications.

Despite its promising results, the proposed system has several limitations that present opportunities for improvement. The reliance on a camera and an ultrasonic sensor poses challenges in various scenarios. The camera is affected by low-light conditions, while the ultrasonic sensor struggles to accurately detect objects with complex geometries. Additionally, the system lacks advanced interaction capabilities, such as voice or text command inputs, which could enhance usability. Further optimization is also needed to improve decision-making speed and overall system efficiency for better real-time performance. To address these limitations, future work will focus on (a) developing a more robust and functional robotic platform to replace the current basic prototype, significantly improving the system's capability and reliability; (b) enabling voice or text command inputs to provide a more interactive and flexible user interface; (c) enhancing the system's control module to autonomously generate executable code for planned motions, further improving the robot's ability to perform specific actions without human intervention; (d) improving visual processing algorithms to minimize reliance on additional sensors like LiDAR, reducing the robot's construction cost and power consumption while ensuring precision in challenging conditions; (e) expanding the system's robustness to handle larger and more diverse environments, improving scalability and performance under practical scenarios. These advancements will further improve the system's reliability, adaptability, and practicality for real-world robotic applications.

Acknowledgments

This work was partially supported by a Frostburg State University (FSU) Student Engagement Project (SEP) award and the FSU Foundation Boxley Faculty Research Award (#34059).

References

[1] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, "GPT-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.

[2] R. Anil, S. Borgeaud, Y. Wu, J.-B. Alayrac, J. Yu, R. Soricut, J. Schalkwyk, A. M. Dai, A. Hauth, K. Millican *et al.*, "Gemini: A family of highly capable multimodal models," *arXiv preprint arXiv:2312.11805*, vol. 1, 2023.

[3] T. Brown, B. Mann, N. Ryder *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems (NeurIPS)*, vol. 33, pp. 1877–1901, 2020.

[4] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems (NeurIPS)*, vol. 35, pp. 24 824–24 837, 2022.

[5] Y. Kim, D. Kim, J. Choi, J. Park, N. Oh, and D. Park, "A survey on integration of large language models with intelligent robots," *Intelligent Service Robotics*, vol. 17, no. 5, pp. 1091–1107, 2024.

[6] F. Zeng, W. Gan, Y. Wang, N. Liu, and P. S. Yu, "Large language models for robotics: A survey," *arXiv preprint arXiv:2311.07226*, 2023.

[7] D. Shah, B. Eysenbach, G. Kahn, N. Rhinehart, and S. Levine, "ViNG: Learning open-world navigation with visual goals," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 215–13 222.

[8] D. Shah, B. Eysenbach, N. Rhinehart, and S. Levine, "Rapid Exploration for Open-World Navigation with Latent Goal Models," in *5th Annual Conference on Robot Learning*, 2021.

[9] D. Shah and S. Levine, "ViKiNG: Vision-Based Kilometer-Scale Navigation with Geographic Hints," in *Proceedings of Robotics: Science and Systems*, 2022.

[10] A. Radford, J. W. Kim, C. Hallacy *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763.

[11] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 3357–3364.

[12] P. Anderson, Q. Wu, D. Teney, J. Bruce, M. Johnson, N. Sünderhauf, I. Reid, S. Gould, and A. Van Den Hengel, "Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2018, pp. 3674–3683.

[13] M. Ahn, A. Brohan, N. Brown *et al.*, "Do as i can, not as i say: Grounding language in robotic affordances," *arXiv preprint arXiv:2204.01691*, 2022.

[14] A. Brohan, N. Brown, J. Carbajal *et al.*, "RT-2: Vision-language-action models transfer web knowledge to robotic control," *arXiv preprint arXiv:2307.15818*, 2023.

[15] D. Shah, B. Osinski, B. Ichter, and S. Levine, "LM-Nav: Robotic navigation with large pre-trained models of language, vision, and action," in *6th Annual Conference on Robot Learning*, Auckland, New Zealand, 2022.

[16] T. Brown, B. Mann, N. Ryder *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems (NIPS)*, vol. 33, pp. 1877–1901, 2020.

[17] R. Anil, A. M. Dai, O. Firat, M. Johnson, D. Lepikhin, A. Passos, S. Shakeri, E. Taropa, P. Bailey, Z. Chen *et al.*, "PaLM 2 technical report," *arXiv preprint arXiv:2305.10403*, 2023.

[18] D. Driess, F. Xia, M. S. Sajjadi *et al.*, "PaLM-E: An embodied multimodal language model," *arXiv preprint arXiv:2303.03378*, 2023.

[19] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar *et al.*, "LLaMA: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.

[20] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan *et al.*, "The LLaMA 3 herd of models," *arXiv preprint arXiv:2407.21783*, 2024.

Author Biography

Xunyu Pan received the B.S. degree in Computer Science from Nanjing University, China, and the M.S. degree in Artificial Intelligence from the University of Georgia. He earned the Ph.D. degree in Computer Science from the State University of New York at Albany (SUNY Albany) in 2011. He previously served as an instructor in the Department of Computer Science and Technology at Nanjing University. In 2012, he joined Frostburg State University (FSU), where he is a Professor of Computer Science and Director of the Laboratory for Multimedia Communications and Security. Dr. Pan received the 2011-2012 SUNY Albany Distinguished Dissertation Award and the 2016 FSU Faculty Achievement Award in Teaching. He has also received multiple Al and Dale Boxley Faculty Research Awards from the FSU Foundation and was inducted into the Honor Society of Phi Kappa Phi in 2023 as a faculty member in recognition of his outstanding research achievements. His publications span peer-reviewed conferences, journals, and book chapters in multimedia security, machine learning, robotics, communication networks, computer vision, and medical imaging. He is a member of ACM, IEEE, and IS&T. (Corresponding Author: xpan@frostburg.edu)

Jeremy N. Perando is a third-year undergraduate student at Frostburg State University, pursuing a B.S. degree in Computer Science. He has been actively involved in robotics in Western Maryland for over five years, assisting local robotics teams and organizations in promoting STEM education throughout the region. His research interests include robotics, artificial intelligence, and computer vision. After graduation, he plans to pursue a career in software development.