ROBOTIC MASONRY ARCH CONSTRUCTION CONSIDERING OBJECT RECOGNITION AND HUMAN-ROBOT INTERACTION

A thesis submitted

To Kent State University in partial

Fulfillment of the requirements for the

Degree of Master Science in

Architecture and Environmental Design

by

Vahid Koliyaee

Nov 2023

Thesis written by

Vahid Koliyaee

B. Architectural Engineering, Shahid Beheshti University, 2017

M.S. in Architectural Technology, University of Tehran, 2021

M.S. in Architecture and Environmental Design, Kent State University, 2023

Approved by

Dr. Rui Liu , Advisor

Dr. Jong-Hoon Kim, , Committee Members

Professor Bill Lucak,

Dr. Lameck Onsarigo

Accepted by

Dr. Reid Coffman ________________________, Program Coordinator, MSAED

Sung Ho Kim ___________________________, Director, Architecture

Mark Mistur ____________________________, Dean, College of Architecture and Env. Design

Abstract

In the midst of a technological revolution, where industries are rapidly evolving to embrace cutting-edge technologies, AEC faces a persistent challenge. It grapples with a noticeable disparity between state-of-the-art technologies and their practical implementation. This research takes a deep dive into the transformative potential of HRI within the AEC context, with a primary goal of bridging this profound gap. The central aim of this research is to investigate the feasibility of empowering robots with advanced object detection capabilities, primarily focusing on the YOLO algorithm, to autonomously recognize and select construction materials, such as bricks, during construction tasks. This innovative approach marks a significant departure from conventional construction practices, where instructions typically flow in a one-way direction from humans to robots. The introduction of the ArchiTech project, which is built upon the ROS and Python, exemplifies an unprecedented shift in making robots smarter and more responsive to human operators.

The integration of robots into human-oriented tasks offers a plethora of benefits, and these extend beyond the AEC sector. First and foremost, it leads to enhanced precision and quality. Robots, being inherently precise and consistent, have the capacity to perform repetitive tasks with an unparalleled level of accuracy, which translates into improved product quality and consumer satisfaction. Moreover, robots enhance safety and ergonomics in the workplace. They can undertake tasks that are perilous or physically demanding for humans. By doing so, they improve worker safety and the overall ergonomics of the workplace. This is particularly notable in industries like construction, where robots can be deployed to handle heavy loads and operate in challenging terrains, thereby minimizing the risk of accidents and injuries. Robotic fabrication also holds the promise of boosting efficiency and productivity. Robots work tirelessly without succumbing to fatigue, leading to increased output and reduced production time. In sectors such as logistics, robots expedite the sorting and delivery of goods, thereby shortening lead times and ensuring timely deliveries. This heightened efficiency results in cost savings and a competitive edge in various industries.

The adaptability and customization afforded by robotic systems are essential for addressing individual requirements. Robots can be programmed and configured to cater to diverse tasks and production needs.

The introduction of the ArchiTech project, underpinned by the YOLO algorithm for object detection, underscores the importance of harnessing cutting-edge technology to address the challenges associated with one-way communication between humans and robots. By enabling robots to perceive and select construction materials independently, this project heralds a new era in which robots become active collaborators in construction projects, revolutionizing the field and opening new avenues for precision, adaptability, and efficient communication in construction processes.

The recommendations for future work are multifaceted. Enhancing the robot's semantic understanding in natural language processing is crucial, enabling it to comprehend construction-specific terminology and context effectively. The integration of machine learning for informed decision-making will further elevate the robot's autonomy, allowing it to analyze situational factors, constraints, and make informed decisions in real-time. Establishing a human feedback loop for construction workers to correct and guide the robot during HRI tasks is an essential step towards ensuring adaptability to site-specific requirements and the preferences of the human workforce.

Keywords: Human-Robot Interaction, robotics, Object detection, YOLO, Robotic operating system

Table of Contents

List of Figures

List of Tables

Acknowledgment

I wish to extend my heartfelt appreciation to my research advisor, Dr. Rui Liu, for his unwavering guidance, exceptional encouragement, and valuable insights throughout this research endeavor. His continuous support, remarkable motivation, and insightful advice have been instrumental in shaping the course of this study. My gratitude also goes to my committee members: Dr. Kim, Professor Lucak, and Dr. Onsarigo, whose invaluable suggestions and feedback significantly contributed to this study. Their expertise and perspectives enriched the depth and quality of the research, making it a more comprehensive and insightful endeavor. Last but not least, I want to convey my deep thanks to my parents, who offered continuous mental and emotional support, even from a distance. Their unwavering encouragement, belief in my abilities, and unwavering faith in my pursuit of knowledge were instrumental in making this academic journey possible. Their love and support were the bedrock upon which this research endeavor was built.

Chapter 1: Introduction

1.1 Background

In recent years, the rapid growth of advanced computing and artificial intelligence has sparked a profound transformation in various industries. In particular, the field of Human-Robot Interaction (HRI) has emerged as a focal point of interest among academics, technical corporations, and the media. This paradigm shift towards HRI in design and construction is driven by several compelling factors that underscore the significance of this research area. First and foremost, the construction industry faces a persistent challenge of labor shortage (Glaser & Molla, 2017). Skilled labor shortages have become a prevalent issue, impacting the timely completion of projects and overall productivity. As traditional methods alone struggle to meet the growing demands of the construction sector, the integration of smart robots and automation technologies becomes increasingly imperative. HRI offers a promising avenue to address this labor scarcity by enhancing the capabilities of both human workers and robots, ultimately optimizing efficiency, and mitigating the effects of workforce limitations. Moreover, the inherent dangers associated with construction activities pose a constant threat to the well-being of workers. Construction sites are rife with hazards, ranging from heavy machinery operation to working at great heights. By introducing robots into the construction environment, particularly in tasks that are hazardous for humans, there is a potential to significantly reduce workplace accidents and injuries.

However, since the industrialization in the 18th century, the interaction between individuals and machines has been strained (Liu et.al 2023). In the mid-1950s, some people started to see automation as having a possible detrimental effect, endangering the health and lives of industrial workers who faced a variety of growing pressures, both physically and psychologically (Beach, 1971). However, this concern over the relationship between humans and machines was not new. During the early days of the Industrial Revolution in the United Kingdom, textile workers damaged machines they believed jeopardized their livelihoods (Smith, 2007). This historical episode illustrates the tensions between humans and machines and serves as a precursor to the debates and anxieties surrounding automation in the mid-20th century. In the early 1950s, popular opinions of automation were dominated by a negative view of it as a possible "nightmare". Negativity was fueled in part by representations of the operational site shown in current popular culture, specifically in the film industry. Charlie Chaplin's 1936 movie, Modern Times, a comical but harsh assessment of industrialization's dehumanizing effects, reflected these negative attitudes. Chaplin's movie was a political act about the influence of modernization on people's working life and other parameters like mental health, and so it connected with a wider, long-standing concept of "the human motor", or the leveraging of human effort into industrial output. While this unflattering analogy was prevalent during the interwar period and continued into the postwar period, by the 1950s, it had been supplemented by a more favorable impression of automation as a method to "liberate tasks from the form and content and physical skills of the body" (Rabinbach, 1990). However, Chaplin's visual concept became an enduring depiction of industrial automation, towards the point that Chaplin's film was widely invoked and referenced through the 1960s, and over twenty years after its premiere.

There is a fundamental social demand for buildings that is safe, economical, sustainable, resilient, and smart. The world's population is urbanizing, needing a new method to develop dense and adaptable infrastructure. Nowadays, 54% of people are living in cities, and so this figure is expected to rise to 66% around 2050, over 90% of that expansion concentrated in developing nations (Ritchie & Roser, 2018). Even to this day, 1.6 billion individuals lack sufficient accommodation, and an additional 60 million are forced to evacuate their homes each year due to starvation, conflict, and other natural disasters. Furthermore, the construction industry is inherently dangerous, accounting for around 20% of all worker accidents in the United States (U.S.D, 2014). Automation in construction may reduce these issues while also allowing for totally new uses. In the wake of Hurricane Katrina in 2005, the 2010 Haiti catastrophe, or the 2011 Fukushima Daiichi nuclear tragedy, robots may have built temporary shelters, containment buildings, or barriers to hide radiation. Remote automated buildings might possibly find uses in interplanetary habitats, such as assisting forthcoming human Mars expeditions (Khoshnevis, 2004).

The Fifth Industrial Revolution, often referred to as Industry 5.0, presents various perspectives distinct from Industry 4.0, which primarily focuses on connecting devices. In the context of Industry 5.0, there is a growing belief in the potential for humans to collaborate with machines and robots within manufacturing settings (Kadir, 2019). Rather than relying on traditional programming languages and compilers, this concept involves instructing robots directly through spoken commands, akin to guiding a human worker. This collaboration envisions having a partner or assistant on the shop floor who can efficiently handle labor-intensive tasks with exceptional precision and accuracy, without experiencing fatigue. To facilitate human-robot cooperation, it is essential to leverage automation and robotic technologies. The notion of a fully automated "lights-out factory," where machines are solely programmed and maintained, has proven to be impractical. Running a factory entails a significant amount of human ingenuity, continuous learning, and adaptability (Kadir, 2019). The “Fourth Industrial Revolution," characterized by advancements in genetics, artificial intelligence (AI), robotics, and self-driving cars (Schwab 2016). This "robot revolution" is reshaping our planet and ushering in an era of unprecedented productivity. However, it is essential to contextualize this revolution by considering the lessons learned from the first three industrial revolutions and examining the evolving relationship between humans and machines. The First Industrial Revolution, which began in the late 18th century, was marked by mechanization, particularly in textiles and manufacturing. It fundamentally changed the nature of work, moving people from agrarian lifestyles to factory jobs. The Second Industrial Revolution, in the late 19th and early 20th centuries, introduced electricity and mass production, further altering work dynamics and urbanizing societies. The Third Industrial Revolution, often referred to as the Digital Revolution, emerged in the mid-20th century with the advent of computers and automation. Throughout these revolutions, the relationship between humans and machines evolved as technology advanced. Initially, there was resistance and concern as workers feared job displacement and dehumanization. However, over time, societies adapted and found ways to coexist with machines, ultimately improving productivity and living standards. As we embark on the Fourth Industrial Revolution, we must draw lessons from history. Industry automation, including the use of machines to perform tasks previously done by humans, predates this era. The key lesson is that while automation can lead to job displacement in certain sectors, it also has the potential to create new opportunities and enhance overall productivity. To navigate this revolution successfully, it is crucial to strike a balance between harnessing the benefits of technology and ensuring the well-being of the workforce. Just as in previous revolutions, the relationship between humans and machines will continue to evolve, and our ability to adapt will be paramount.

Accordingly, many recent research studies have investigated the possibility of combining robotics with digital fabrication (Retsin & García, 2016). This innovative approach holds the promise of revolutionizing construction by seamlessly integrating advanced robotics into the fabrication and assembly processes. Academic institutions and industry leaders have begun to adopt and adapt human-robot interaction criteria in the development of what is now termed the "Collaborative Construction" procedure (Vasey et al., 2016). This framework envisions a future where humans and robots, equipped with advanced computing capabilities, work in harmonious coordination to achieve unprecedented levels of efficiency, precision, and safety in the construction industry. Based on these groundbreaking advancements, the robotic production and assembly process enters a new domain of human-robot-computer interaction. The synergy between humans, robots, and computing technologies promises to redefine the way people conceive and execute construction projects. As human delve deeper into this transformative era, it becomes increasingly apparent that the fusion of human expertise and robotic precision will shape the future of design and construction, providing solutions to longstanding challenges while ushering in a new era of creativity, productivity, and safety in the built environment.

1.2 Problem Statement

The rapid advancement of technology, particularly in the realms of robotics and computer science, has ushered in a new era of possibilities and capabilities in various industries. However, the architecture, engineering, and construction (AEC) manufacturing sector often find themselves grappling with a growing gap between these state-of-the-art technologies and their practical implementation. While cutting-edge tools like robot arms and advanced computing systems have been developed, they remain less accessible to architects, engineers, and construction managers who are at the forefront of shaping our built environment. In stark contrast to sectors like automotive manufacturing, where automation and robots are commonplace, the AEC fields still largely adhere to traditional methodologies and approaches. Despite the emergence of computational design software and the introduction of robotic fabrication as innovative methods for enhancing efficiency and enabling the construction of complex architectural forms, the industry has been slow to fully embrace these technological advancements. Consequently, the design and manufacturing processes in AEC have evolved at a slower pace than the rapid technological advances available to them.

This research endeavors to bridge the existing gap between cutting-edge technology and the AEC industry. It seeks to understand the driving forces behind the resistance to adopting advanced robotic and computational tools in design and construction, and explores the transformative potential of Human-Robot Interaction (HRI) in this context. The primary motivation is to unravel the complexities surrounding the integration of robots into AEC and to redefine the conventional perception of robots as mere operators executing human commands. In the current landscape of robotics in construction, where human instructions are translated into robotic actions without real-time feedback, can robots be developed to perceive, differentiate, and accurately select building materials such as bricks during construction operations? Can robots be empowered with visual perception and recognition capabilities to autonomously detect variations in building materials and notify human operators of any anomalies or issues encountered during construction processes? Is it possible for robots, equipped with advanced vision systems, to participate in complex tasks, such as constructing arches, while actively detecting and responding to changes in the environment, ensuring precision and adaptability throughout the construction process? These research questions delve into the core issue of one-way communication between humans and robots in building projects and explore the potential for robots to possess visual perception, adaptability, and feedback capabilities. The aim is to investigate how robotics can evolve from being mere executors of pre-defined commands to active collaborators with the ability to perceive and respond to dynamic construction environments in real-time.

1.3 Obstacles

Although the growing inclination and considerable interest in using robotics in AEC, there is a significant gap between HRI and AEC industry. In recent years, most of the research projects regarding the use of robots in construction area explored in the digital fabrication, which assumes robots as an agent should operate the humans' orders/codes (Prado et.al 2014; Parascho et.al 2017). These findings and considerable development in HRI in other fields like computer science and robotics, leave us with multiple questions. What is the current level of autonomy in robotic construction projects and is it possible to level up this collaboration to be closer to higher level of autonomy in construction? What are the HRI applications and possibilities in construction industry? Will this interaction between human and robot enhance construction productivity, safety, or other important factors?

1.4 Current Approach and Gap

According to prior investigations, there are various software toolboxes and plugins such as KUKA-PRC, trying to ease the connection between robots and architects. But they are not suitable and sufficient for HRI which demands a two ways connection between human/s and robot/s. There is a significant barrier that prevents researchers to go further in this research area. This study's primary objective is to explore whether it is possible to provide a platform in which robots and humans can interact with each other through robotic operating system (ROS) in construction. It intends to integrate various requirements of design, analysis, and construction into one coherent framework.

1.5 Research Structure

This investigation is arranged as follows to address the gap identified in the preceding section.

1.5.1 Research Questions

This study aims to answer the following key research questions:

1. In the context of current robotics practices in design and construction, where instructions flow solely from humans to robots without reciprocal communication, can robots be

developed to visually recognize and accurately select construction materials like bricks during construction tasks?

2. Is it feasible to equip robots with advanced object detection capabilities to autonomously detect variations in construction materials, such as bricks and promptly alert human operators to any deviations or challenges encountered during construction processes?

3. Can robots, armed with state-of-the-art vision systems, actively participate in intricate endeavors, such as building arches, while simultaneously perceiving and responding to changes in the construction environment, ensuring precision and adaptability throughout the construction undertaking?

These research questions are framed to address the central issue of one-way communication between humans and robots in design and construction projects and to explore the potential for robots to possess object detection, adaptability, and feedback capabilities. The research structure outlined in this study aims to shed light on the evolution of robotics from passive executors of predefined commands to active collaborators capable of perceiving and reacting to dynamic construction conditions in real-time.

1.5.2 Intended Contribution

The field of HRI in construction has gone through significant growth in recent years, with increasing interest in developing robotic solutions to improve efficiency, productivity, and safety on construction sites. One area of HRI in construction that has yet to be fully explored is the use of object recognition technology to enhance human-robot collaboration.

The main objective of this research is to develop a framework for HRI in construction that utilizes object recognition to improve the performance and safety of human-robot collaboration. Specifically, this research aims to:

- Investigate the use of object recognition technology in construction applications and evaluate its potential benefits.

- Develop a framework for HRI in construction that integrates object recognition technology to improve human-robot collaboration.
- Evaluate the performance and safety of the proposed framework through simulation and field tests.

The proposed research is expected to contribute to the field of HRI in construction by providing a deeper understanding of the potential benefits and challenges of using object recognition technology in human-robot collaboration. Additionally, the framework developed in this research is expected to serve as a valuable tool for practitioners in the construction industry, as it will provide a means for improving the efficiency, productivity, and safety of human-robot collaboration on construction sites.

1.5.3 Thesis Organization

This thesis is divided into five chapters, which are as follows:

The first chapter is an introduction that comprises an explicit explanation, framing of the questions, and a description of the research objectives. The second chapter is a literature review that provides an in-depth analysis of the history of HRI, object recognition, Level of autonomy, robotic masonry arch construction, and architectural robotics and how robotics they have been utilized collaborated in various projects. The comprehensive literature review helps identify the state-of-the-art in object recognition technology, HRI in construction and the challenges that both of them face. The research method is explained in depth in Chapter three. A framework is proposed in this study by integrating the object recognition technology into the HRI in construction. The fourth chapter is to evaluate the proposed framework through simulation and field tests about data analysis. It discusses the methodology used to create the HRI platform, describes its iterations, and analyzes the experiment's findings. The fifth chapter presents the conclusions of the research and recommends the next stages.

Chapter 2: Literature Review

2.1 Introduction

In contemporary times, robotics and automation have become increasingly prevalent within numerous industries, including architecture and construction. The integration of robotic systems into construction operations has the potential to enhance productivity, precision, and safety significantly. However, to effectively implement these systems within the architecture and construction fields, a profound comprehension of two critical components of HRI is essential: object recognition and human-robot communication. These two components serve as the backbone of successful HRI in construction. Object recognition is indispensable because it allows robots to identify and understand their environment, ensuring they can work seamlessly alongside human workers while comprehending the layout and composition of the construction site. Meanwhile, human-robot communication, including advanced techniques like natural language processing (NLP), is vital for facilitating real-time collaboration and information exchange between humans and robots, ultimately leading to improved teamwork, efficiency, and safety on the construction site. Together, these elements empower robots to not only perceive their surroundings but also to interact effectively with their human counterparts, making them invaluable assets in the modern construction industry.

By the mid-1950s, significantly as industrial progress accelerated, some introduced automation as a fascinating, desirable, and probably inevitable development in human's utilization of technical and scientific information; from this vantage point, automation was a critical instrument for attaining economic progress, improved living standards, greater freedom, and better lives for all (Wingo, 1963). On the other hand, those who saw automation as a possibly dangerous trend highlighted methods in which automation would endanger the health and lives of the country's industrial workers; this side of the discussion predicted that individual employees would face various physical and psychological growing pressures (Liu et al., 2023). While robot technology was mainly created in the early twentieth centuries, the concept of machine behavior and its ramifications for humans has existed in religions, myths, logic, and literature for millennia. The term "robota" was derived from the Czechoslovakian term robot a, which means "worker" (G. & C. Merriam Co, 1965). The term "robot" seems to have first appeared in Karel Chapek's 1920 drama Rossum's Universal Robots, although that was far from the first instance of a human-like device to explore possible connections between computational design systems and the inherent properties of

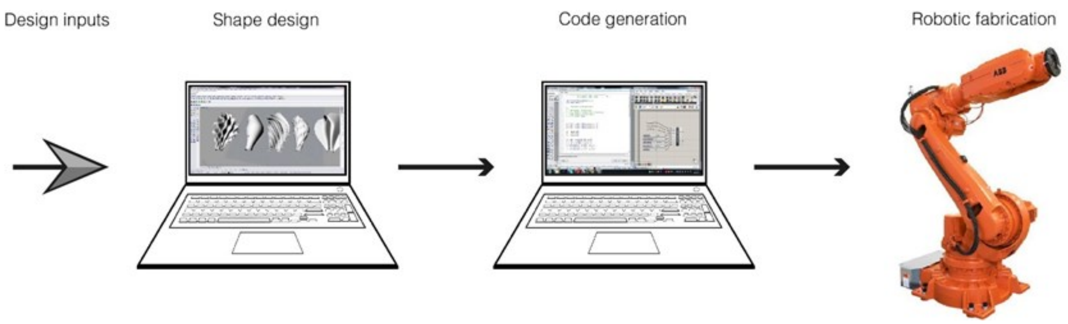


Figure 1: Digital Fabrication Process

material systems (Gramazio & Kohler, 2008). Prior to recent, most research on robotics in architecture focused on a file-to-factory approach (Johns et al., 2014), where all discrete fabrication stages were established prior to the entire fabrication process (Figure 1).

The interaction between humans and robots, known as human-robot interaction HRI is critical in ensuring the successful deployment and acceptance of robotic systems in construction processes. In the realm of masonry arch construction, HRI assumes even greater significance due to the collaborative nature of tasks. Robots must communicate effectively and coordinate their actions with human workers, ensuring seamless integration and synchronization of efforts. Understanding the roles of humans and robots in HRI is crucial to design efficient and intuitive interfaces and establishing effective communication channels between the two parties.

This chapter explores the existing body of knowledge and research related to object recognition and HRI in the context of robotic construction. By examining the relevant literature, we gained insights into the current state-of-the-art technologies, methodologies, and practices employed in these areas. Additionally, this study investigated the impact of automation and robotics on the construction industry, mainly focusing on assessing autonomy levels and applying Boyd's OODA

(Observe, Orient, Decide, Act) loop framework in automation in construction. Furthermore, this chapter presented case studies highlighting successful applications of robotics in architecture and construction. Specifically, this exploration analyzed their contributions to advancing the field and identifying key lessons learned. Additionally, this study explored the Robotic Operating System (ROS) and its advantages in the AEC domains, shedding light on its potential for enhancing the capabilities of robotic systems. By comprehensively examining the existing literature and synthesizing key findings, this chapter laid the foundation for the subsequent chapters of this thesis, which delved deeper into the development and implementation of a robotic masonry arch construction system that leveraged object recognition to facilitate HRI. This research aims to contribute to the body of knowledge in this field and provide valuable insights and recommendations for the future design and deployment of robotic systems in design and construction applications.

2.2 HRI

2.2.1 Introduction to HRI

HRI is a discipline dedicated to understanding, designing, and evaluating robotic systems intended for use with or by humans (Huang, 2016). It is a subfield of Human-Computer Interaction (HCI), and scientific research plays a significant role in shaping HRI (Yanco & Drury, 2002). HRI focuses on establishing principles and techniques that enable robots to interact with humans safely and efficiently (Feli-Siefer & Mataric, 2010). Effective interaction, characterized by successful human-robot communication and the establishment of trust, is at the core of HRI. The communication interface between humans and robots, including the communication medium, communication indicators, and the technologies used for transmission, dramatically influences the performance of HRI (Steinfeld et al., 2006). This study emphasizes the various communication channels and the development of expressive visual communication cues in HRI. Furthermore, incorporating human-like designs and natural user interfaces can enhance robotic systems' overall adoption and interactivity (Richert et al., 2018).

The study of HRI is driven by the need to design robots that can seamlessly integrate into human-centric environments and perform tasks in collaboration with humans. It requires a deep understanding of human cognition, behavior, and social dynamics and the development of advanced robotic technologies. By investigating the principles and mechanisms underlying successful human-robot interactions, researchers in HRI aim to improve robotic systems' usability, acceptance, and effectiveness across various domains. Furthermore, HRI explores the ethical, psychological, and societal implications, paving the way for responsible and human-centered design practices.

2.2.2 Background

In HRI, the physical appearance of a robot plays a crucial role in how it is perceived and engaged with by humans (Luptetti, 2017). Anthropomorphic features like human-like designs and expressive facial expressions enable robots to convey trustworthiness and engage in substantial social interactions with humans (Stoeva & Gelautz, 2000). The physical appearance of a robot during the initial encounter significantly influences how users perceive it and subsequently affects the nature of human-robot engagement (Luptetti, 2017). Trust is critical to human acceptance and effective robot communication (Billings et al., 2012).

The field of HRI has its roots in early research and science fiction, but it has gained significant momentum with the advancements in robotics and artificial intelligence. The idea of machines interacting with humans has existed in religious texts, myths, and literature for centuries. However, the formal study of HRI began to emerge in the late 20th century with the advent of robots and automation.

As technology progressed, HRI researchers explored various aspects of the interactions. One crucial aspect is the role of communication in facilitating effective interaction. Communication channels between humans and robots involve visual cues, auditory signals, haptic feedback, and NLP. Developing intuitive and efficient communication interfaces is essential for enabling seamless collaboration and understanding between humans and robots. Furthermore, robots' physical appearance and embodiment play a significant role in HRI. The design choices regarding the robot's form, structure, and features can influence how humans perceive and interact with the robot. Anthropomorphic designs, mimicking human-like faces, gestures, or body structures have been explored to enhance robots' human-like qualities and social engagement. However, the design of robots must carefully balance anthropomorphism with functional requirements to avoid falling into the uncanny valley, where excessive human likeness can lead to discomfort or unease in human users.

2.2.3 Importance of HRI in AEC

HRI holds significant importance in AEC, where the integration of robotic systems can revolutionize the industry. The collaboration between humans and robots in these fields can improve efficiency, safety, and precision in construction operations. Effective HRI plays a crucial role in ensuring seamless coordination, communication, and task allocation between human workers and robots, ultimately enhancing productivity and project outcomes (Al-Hussein et al., 2017; Christensen & Grönvall, 2012; Staub-French & Khanzode, 2007; Whelan et al., 2013; Yang & Chang, 2018). One of the key aspects of HRI in AEC is the division of labor between humans and robots. Robots can be leveraged to perform repetitive and physically demanding tasks, such as heavy lifting, bricklaying, and welding. At the same time, humans can focus on more complex and creative aspects of the construction process. This division allows for optimal resource utilization and increases overall productivity. For example, robots can be programmed to position accurately and lay bricks in masonry arch construction, while skilled human workers can focus on design considerations and quality control.

Furthermore, HRI in AEC involves the development of intuitive and user-friendly interfaces to facilitate effective communication and control of robotic systems. Human operators need to be able to interact with robots naturally and efficiently, whether through physical interfaces, voice commands, or augmented reality systems. These interfaces should be designed with the specific needs and capabilities of the construction environment in mind, ensuring seamless integration of robots into the existing workflows and enabling easy collaboration between humans and machines.

HRI also plays a critical role in addressing safety concerns in the construction industry. By delegating hazardous tasks to robots, the potential for human injuries can be significantly reduced.

Robots equipped with advanced sensing capabilities can navigate complex construction sites, detect potential hazards, and alert human workers to potential dangers. Moreover, HRI allows for real-time monitoring of robotic systems, enabling human operators to intervene or provide guidance when necessary. To fully realize the benefits of HRI in AEC, interdisciplinary collaboration between robotics engineers, architects, construction professionals, and human factors experts is essential. Moreover, the importance of HRI in AEC cannot be overstated. Through effective collaboration and interaction between humans and robots, construction processes can be streamlined, efficiency can be enhanced, and safety can be improved. Developing intuitive interfaces, task allocation strategies, and real-time monitoring systems are key research areas that contribute to the successful implementation of HRI in these industries. By embracing HRI principles and leveraging robotic technologies, the AEC fields can embrace a future where human-robot collaboration unlocks new possibilities and transforms how we build and create.

2.2.4 Key Challenges in HRI

HRI in AEC meets many substantial challenges that must be overcome before it can be employed successfully. Designing appropriate interfaces for controlling systems which allow seamless interaction and collaboration among humans and robotic devices is one of the main challenges (Marti et al., 2018). The interface should be intuitive, user-friendly, and capable of supporting the different skill sets and experience of human workers (Tenneti et al., 2016). Establishing an equilibrium among automation and human supervision is critical to ensuring that humans can successfully oversee and intervene when needed while maximizing robot efficiency and capabilities (Bishop & Hudson, 2018).

2.2.5 Communication and Collaboration

Human-robot communication and cooperation are critical to efficient HRI in architecture and construction. Robots must properly grasp and interpret human orders and intents, whether via voice recognition, gesture recognition, or other means of interactions. NLP and interpreting human purpose are current research areas for improving human-robot communication (Thomaz & Breazeal, 2008). Human-robot collaboration necessitates the smooth sharing of information and tasks. Real-time feedback, clarifying inquiries, and adaptation to changing situations should be possible for robots (Haddadin et al., 2016).

2.2.6 Perception and Understanding

Robots with enhanced perception and comprehension skills are required for efficient HRI in AEC. They should be able to appropriately observe and evaluate their surroundings, recognize items, and comprehend the connections between them (Bentivegna et al., 2018). Recognition techniques that include computer vision, depth sensing, and LiDAR are critical in allowing robots to navigate complicated building sites while still interacting securely with human employees (Jin et al., 2020). The development of reliable methods and machine learning approaches is critical for increasing robot perception and comprehension in dynamic and unstructured manufacturing contexts.

2.2.7 Safety and Trust

Safety and trust are critical factors in HRI in AEC. When robots are employed in construction operations, it is critical to ensure the safety of both human employees and the general population (Chen et al., 2017). To prevent accidents and other risks, robots must be trained in monitoring and adjusting to their environment (Li et al., 2019). Trust between people and robots is also essential for efficient collaboration. Clarity in robot behavior, explicit communication of goals, and the capacity to comprehend human cues and social conventions all aid in the development of trust (Sabelli et al., 2011). Safety standards, laws, and guidelines all play a significant role in the research and development and implementation of construction robots (ISO, 2011). In conclusion, HRI in AEC has numerous major obstacles that must be solved. Effective communication and teamwork, enhanced perception and comprehension, and providing safety and trust are all significant success components. Current research and development activities aimed at overcoming these issues will lead the way for the broad use of autonomous systems in the building industry, ultimately improving efficiency, security, and overall project results.

2.2.8 Role of Humans in HRI

Despite being necessary for the development of autonomous manufacturing and robotic techniques, the involvement of humans in these procedures needs to be noticed. Even in the most basic robotic procedures, humans are directly engaged. They frequently serve as more than task executors by observing the robotic operation, making required modifications, and continuously refining the building operations. There are two ways to look at robotic building processes: (1) full automation intended to replace humans and (2) collaborative procedures involving humans and robots that extend potential design and construction options (Han et al., 2021). While automation may result in more rapid and effective building processes, employing human actors in the manufacturing process is a significantly more worthwhile and sustainable aim. Instead of eliminating the presence of human employees on building sites, it is beneficial to investigate how people may be assisted by robots and vice versa (Liu et al., 2023). As a result, the interaction between people and robots, as well as the control of many machines, becomes a critical study topic for advancing robotic building called- HRI. HRI and machine-machine collaborations are not novel fields of investigation. Even though they are new to design and construction, they are developing a considerable collection of information from other areas, such as robotics, mechatronics, engineering, and data science. Through novel techniques and applications, architecture can link many such fields. As a result, it is critical for architects, engineering, and construction managers to have information available in robotics. We may find new topics of investigation at the interconnections of robotics-related fields and AEC by providing current methodologies and breakthroughs available to AEC professionals, which would not be possible if we addressed each issue individually (Han et al., 2021). Over the last few decades, the discipline of robotic systems has evolved into two separate applications: industrial and creative. Industrial robots are designed to replace people in repetitive, filthy, and dangerous activities. They are designed to boost speed and production while lowering expenses. Lower-skilled employees, on the other hand, regard industrial robots as effectively snatching their jobs. On the other hand, the creative use of robots strives to supplement humans by providing superhuman accuracy and re-enabling complexity (ARUP, 2020).

2.2.9 Role of Robots in HRI

Robotic masonry structure building was first demonstrated at the architectural dimension in the Gantenbein Winery construction, when robotic arms were utilized to build the complex's undulating brick walls (Bonwetsch et al. 2006; Bonwetsch and Kohler, 2007). While new robotic tools have influenced modern masonry manifestations operational performance results in architecture and construction method has also influenced the development of correlating robotic fabrication methods and machinery (Bonwetch et al. 2006; Abdelmohsen et al. 2019; Piškorec et al. 2018). Integrative design techniques proposed co-development of design formulation, material testing, and robotic production strategy to speed iterative advancement among equipment and ideas (Parascho et al. 2015). However, when implemented in contexts other than their original goal, tools and procedures created in this method may encounter issues owing to overspecialization.

2.3 Computer Vision / Object Detection

2.3.1 Introduction

In recent years, there has been a growing interest in applying deep learning i.e., computer vision (CV) and NLP to AEC industry. Object detection, a crucial task in CV, has a wide range of applications, including security and surveillance, autonomous driving, and robotics, which enables computers to identify and locate objects within digital images or videos. This technology has a rich history, with early methods dating back to the 1960s (Horn, 1986). Despite the recent significant progress in CV and object detection techniques, limited practices have demonstrated their potential applications in the field AEC industry. This gap has resulted in missed opportunities for improving safety and efficiency in the AEC industry, where fatalities and injuries remain a significant concern (Hinze, 2018). However, recent studies have shown promising results in using CV in forensic inspection and construction sites, where it can help identify distresses, monitor worker behavior, and prevent accidents (Kang et al., 2020; Anani-Manyo and Liu, 2022). Additionally, object detection could also be applied in robotic construction, where it has potential to assist with tasks and address the current labor shortage in the construction industry (Wang et al., 2019).

2.3.2 YOLO

YOLO (You Only Look Once) is a renowned real-time method and a deep learning-based CV system for object detection (Sultana et al., 2020) that was first introduced by Joseph Redmon et al. (2016) in their article *'YOLO: Real-Time Object Detection'*. Scholars have produced various YOLO later versions with the latest version of YOLOv8 (Zhao et al, 2019; Zou, 2019; Laroca et al. 2018; Tian et al. 2019; Jamtsho et al. 2021; Han et al. 2018; Lin & Sun, 2018; Lu et al., 2018; Li et al. 2022; Wang et al. 2022; Zhang et al. 2023). YOLO is quick and precise, making it suited for a wide range of real-time applications such as security systems, autonomous vehicles, and robots. YOLO splits a picture into grid cells and detects items in each one, allowing several objects to be recognized in a single frame. This method makes YOLO extremely efficient, because of its rapid speed and excellent accuracy. Overall, YOLO is a robust and adaptable object recognition system that has made significant contributions to deep learning and CV.

Various object detection techniques, such as ImageNet, TensorFlow, and the YOLO family of algorithms, exist in this field. ImageNet is a dataset of more than 14 million images, is a popular resource for developing and testing object detection algorithms (Deng et al., 2009). TensorFlow, an open-source machine learning platform, includes object detection APIs that allow developers to create and train object detection models (Pattanayak, 2023). In general, object detection algorithms may be divided into two categories: classification-based algorithms and regression models. Classification-based algorithms are incorporated in two stages. They begin by identifying areas of focus in a picture. Second, they use convolutional neural networks (CNN) to categorize these areas. This technique may be time-consuming since the computer must perform forecasts for each specified location. The region-based convolutional neural network (RCNN) and its relatives Fast-RCNN, Faster-RCNN, and the newest addition to the family, Mask-RCNN, and Retina Net are well-known examples of this sort of algorithm (Girshick et al. 2014; Girshick, 2015; Ren et al. 2015; He et al. 2017; Lin et al. 2018).

Regression models predict classes and bounding boxes again for entire picture in a single run of the algorithm, rather than picking important areas of an image (Arya & Rawat, 2020). The

YOLO family algorithms are one of the well-known applications from this category. They are often employed for real-time object identification since, in general, they exchange a bit of accuracy for significant speed benefits. YOLO algorithms determine an object's classification and the bounding box that specifies the object's position. Four characteristics may be employed to characterize any bounding box, including the center of a bounded box, its width and height, and a label indicating the object's classification.

2.3.3 Potential applications of object recognition in the AEC Industry

Object detection technology may help the AEC industry in a wide range of applications, including:

- Quality assurance: Object recognition can be used to examine building sites, discover errors or changes from the blueprint and verify that quality requirements are fulfilled (Wang et al. 2021).
- Inspections for safety: It can help discover potential dangers such as loose cables or tripping hazards, making the building site safer for both employees and visitors (Li et al. 2021).
- Progress tracking: It can be employed to track the progress of building projects in real-time, allowing stakeholders to notice any delays or concerns that need to be resolved fast (Zhou, 2020).
- Monitoring of equipment: Object detection has the potential to track the usage and management of heavy machinery and tools on construction sites, assisting in ensuring that the equipment is utilized safely and effectively (Wang et al. 2020).
- Inventory control: It could aid inventory management by tracking building materials and supplies, ensuring that the appropriate supplies are in the correct location at the right time (Braun & Borrmann, 2020).
- Verification and confirmation of BIM (Building Information Modeling): It can be utilized to guarantee that the construction process matches the computerized BIM model, reducing errors and ensuring that the building is going as planned (Liu & Li, 2021).

- Non-destructive testing (NDT): Technologies, such as X-ray imaging, can employ object detection to find interior faults and abnormalities swiftly and reliably in structures and materials (Kuchma & Golovko, 2021).
- Robotics and automation: Object detection may be used to provide robots and automation processes with the capacity to see and respond to their surroundings, allowing for more advanced and adaptable automation solutions (Khan et al. 2021).

2.3.3.1 Safety

Object detection methods have the possibility of significantly enhancing construction site safety by recognizing and alerting employees to possible risks. Cameras, for example, can be deployed to detect employees in dangerous zones and inform them if they are not carrying the necessary personal protective equipment (PPE) (Wu et al., 2021). Similarly, object detection can be employed to track the motion of heavy machinery and notify employees if they encounter it, lowering the probability of an accident (Xie et al., 2018). Using object detection to avoid accidents and save lives is an important topic of study for the construction.

Object detection algorithms may be used on construction sites to track worker movements and highlight possible risks such as employees entering risky zones or committing dangerous acts. This data may then be utilized to increase safety regulations and develop more efficient safety training programs. Object detection can also aid in the identification of worker weariness, which is a significant contributor to incidents in the construction sector. Companies can take proactive actions to lower the probability of accidents and enhance worker security through tracking worker movements and behavior.

Furthermore, in the construction industry, object detection may be utilized to enhance machine and material tracking. Construction organizations may effectively control their inventory and avoid theft or loss of valuable equipment by employing object recognition sensors and technologies to monitor the movements of materials and equipment. Moreover, real-time equipment and material tracking can assist optimize logistics and scheduling, decreasing the time and expense involved with transferring items and machinery to the job site. Overall, the use of object recognition technology in buildings has the possibility of enhancing worker safety, boosting productivity, and saving costs for industrial firms.

2.3.3.2 Security

Object detection methods have also been used for security considerations in the construction industry. Video-based object detection systems, for example, may be used to detect and monitor suspicious users on building sites, improving security, and lowering the risk of theft and damage. These technologies have been proved to identify intruders and unwanted entry to restricted areas on building sites (Suresh et al., 2020). Object detection systems' real-time recognition and tracking capabilities allow security personnel to respond quickly to possible security threats, improving safety and security on construction sites.

Another application of object detection in construction regarding security is the detection of dangerous or hazardous situations. For example, object detection systems can be used to detect the presence of hazardous materials or equipment that may pose a risk to workers or the environment. In addition, object detection can be used to detect potential safety hazards such as loose scaffolding or equipment that has been left unattended. By quickly identifying and alerting personnel to these hazards, object detection systems can help prevent accidents and ensure a safer working environment for construction workers (Wang et al., 2018).

Furthermore, object detection systems can also be used to monitor and enforce compliance with safety regulations and procedures in construction sites. For instance, object detection can be used to track whether workers are wearing the appropriate PPE such as hard hats, safety goggles, or harnesses. This can help ensure that workers are following safety regulations and can alert management if there are any violations. Object detection systems can also be used to monitor construction sites for compliance with environmental regulations, such as detecting and tracking the release of pollutants or hazardous materials. This can help construction companies avoid costly fines and penalties while also promoting a safer and more environmentally friendly workplace (Jiang et al., 2021).

2.3.3.3 Robotic Construction

Object detection algorithms have also been used to enable robotic operation in the construction industry. For example, a vision-based object identification system has been presented to increase the precision and effectiveness of robotic excavations in building sites (Li et al., 2020). To recognize and categorize objects in real time, the system relies on a combination of algorithms from deep learning and computer vision techniques. The robotic system can accomplish the excavation work more efficiently and precisely by properly recognizing the position and shape of items in the excavation site. Similarly, another study developed an object detection system using stereo vision for the guidance of robotic welding in construction. The system was able to accurately identify the location and shape of workpieces and welding tools in real-time, which enabled the robot to adjust its motion and welding parameters accordingly. This led to improved welding accuracy and reduced welding defects (Wang et al. 2018).

According to a National Association of Home Builders survey (2019), the construction industry has been facing a significant labor shortage in recent years. An aging workforce, a lack of vocational training programs, and harsher immigration rules have all been implicated for the labor shortage. As a result, the construction industry has experienced project delays and increased expenses because of the necessity for overtime compensation and increased recruitment attempts to fill unfilled jobs.

Also, as technology and automation continue to advance in construction, there is an inherent tension between human labor and machines and the role of human and robot has been changing continually (Liu et al., 2023). Object detection techniques have the potential to address this challenge by increasing the efficiency of construction using robots. Robotic systems equipped with object detection capabilities can automate a wide range of construction tasks, including bricklaying, excavation, and concrete pouring. By reducing the need for human labor, such systems can help to overcome the labor shortage and improve productivity in construction (Li et al., 2020). One use of object recognition in construction is the employment of robotic systems in a completely automated method for choosing and putting the correct bricks. This system can recognize the orientation and positioning of bricks and decide the best arrangement for each block in a wall using object recognition. Jiang et al. (2021) established this method in a case study by constructing a robotic system capable of correctly and effectively laying bricks in a completely automated way. The method achieves a high degree of precision, with just a tiny percentage of bricks performing manual changes, illustrating how object identification in robotic systems has the potential to transform building processes and alleviate manpower shortages (Jiang et al., 2021). Overall, object detection techniques have the potential to significantly improve the efficiency and safety of robotic operations in construction sites and can help address the challenges of labor shortages and increasing demand for construction projects.

2.4 Automation in Construction

Technology is beneficial to businesses. For organizations all around the world, it has enhanced productivity, decreased operational expenses, improved quality, and expanded production flexibility. Automation and robotics also have advanced tremendously in recent times, with equipment that is quicker, smaller, and more inexpensive. To boost their competitiveness and cut production costs, an increasing number of businesses are utilizing innovative technology in their warehouses and factories. Businesses' labor expenses have been lowered because of automated manufacturing methods. While human labor normally performs 8-hour shifts, robots may operate continuously. Once programmed, a machine may simply transition among procedures and programs to generate and develop a product with minimum effort. Robots, unlike human labor, do not have physical or emotional constraints, which can reduce expenses. Machines do not require health insurance, retirement plans, or breaks to relax or rest. Therefore, the Boston Consultancy Firm estimates that a robot replacing a welder may save a firm $17 per hour (Cocco, 2016).

Artificial intelligence and robots, for example, will provide major benefits to people, enterprises, and economies, increasing economic growth and productivity. The magnitude to which these techniques replace employees will be determined by the rate at which they develop and are adopted economic growth and rise in labor demand. Even as technology causes job losses in some fields, automation will affect many more—60% of jobs have at least 30% of constituent labor tasks that might be automated. It will also generate new vocations that do not present now, much like previous technologies did. While currently shown technologies have the technical ability to automate around half of all labor activities globally, the percentage of work replaced by 2030 is anticipated to be lower due to technological, economic, and social variables influencing acceptance. A study of 46 countries implies that between practically zero and one-third of workplace, activities might be replaced by 2030, with a mean of 15% (McKinsey Global Institute, 2017). The amount varies greatly among nations, with advanced economies being more affected than emerging economies.

2.4.1 Autonomy Level Assessment Level

HRI challenges, according to Goodrich and Schultz, are "to comprehend and shape the relationships among one or more individuals and one or more robots" and may be divided into the following basic sections (Atanasova et al. 2020):

- Autonomy level and behavior
- Information exchange's nature
- Team structure
- Flexibility, education, and training of humans and the robot
- Task shape

All inventions may be categorized into distinct stages. This study suggests three existing metrics to increase communication among digital producers and designers: NASA's Technology Readiness Level (TRL)-US Department of Defense Manufacturing Readiness Level (MRL) to demonstrate public acceptance-Fabrication Arup's Autonomy Levels to demonstrate the disruptiveness of each technology.

2.4.2 Arup’s Fabrication Autonomy Levels

The International Association of Automotive Engineers has characterized levels of automation ranging from "hands-on" Level 1 in which the operator and the automated process share the vehicle's control (e.g., Cruise Control) to "eyes-off" Level 2 in which the vehicle and the automated process share the control of the car (e.g., Autopilot). Level 3: The driver must still be prepared to intervene if the vehicle requires it. When no human assistance is necessary, Level 5 is "steering wheel optional." Arup suggests six stages of digital fabrication autonomy, influenced by all these stages and a demonstration by Daniel Prohasky, Swinburne Innovations Fellow, at the Arup Explores event in Melbourne. Arup created these tiers to expand dialogue beyond specialists and enthusiasts. Different levels provide designer possibilities and freedom, as well as obstacles (ARUP, 2017).

2.4.3 Boyd's OODA Loop

John Boyd, a military expert, devised the observe-orient-decide-act (OODA)loop. Boyd applied the notion to the warfighting process, which occurred frequently at the functional level throughout military conflicts. It is currently frequently used to comprehend business procedures and learning processes. In encountering human opponents, the method illustrates how agility may defeat brute force. It is particularly relevant to computer security and cyberwarfare (Clarke, 2019).

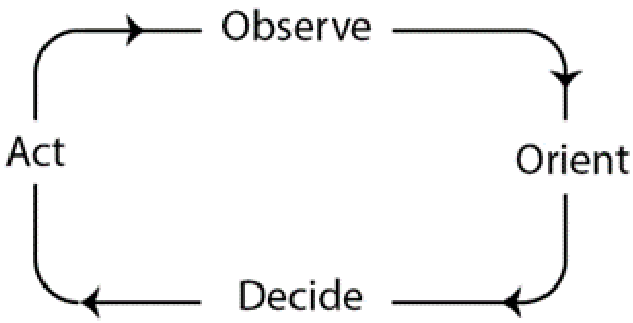


Figure 2: OODA Loop (Richards, 2020)

The OODA loop is a key idea in law, industry, as well as military matters. As is shown in Figure 2 and according to Boyd, decision-making follows a cyclical pattern of OODA. An entity (person or institution) that can execute this cycle swiftly, monitoring and responding to events unfolding faster than an opponent, might therefore "get in" the opposition's decisions cycle and gain an edge (Richards 2020). All the robotic projects in AEC can be analyzed not only based on all autonomy models which discussed above, but also their process has an ability to be discussed based on OODA model, see Table 1. Consequently, there would be a project that states in low level of autonomy based on one of models such as Arup, however by scrutinizing the project process can recognized that observation states in level four or higher, however, the orientation level states the first level. Consequently, to increase their autonomy level, all the robotic projects in construction not only should discuss based on their autonomy level but also, they are supposed to scrutinize themselves based on the OODA model.

2.5 Robotic Operating System (ROS)

2.5.1 Introduction

ROS is a renowned open-source framework that has contributed significantly to the area of robotics in design and construction. ROS is a versatile and modular framework for designing robotic systems, with a variety of devices, libraries, and features that enable smooth communication and integration between various robotic components. This section will go through the important characteristics and benefits of ROS in the architectural and construction industries.

One of ROS's most significant advantages is its ability to foster interoperability and collaboration across various robotic systems and devices. ROS has a distributed architecture, which allows several robots and sensors to interact with one another and share data throughout an organized manner. This functionality is especially useful in design and construction, where several robots and sensing equipment may need to collaborate on complicated tasks like cooperative assembling or environment mapping. ROS provides a standard protocol for communication and message-passing system that allows for smooth interaction and data exchange throughout various components, independent of their hardware or software variances. In 2007, Stanford University's Artificial Intelligence Lab created the ROS.

Table 1: Boyd's OODA (Proud et al., 2013)

Level	Observe	Orient	Decide	Act
1	Humans are the primary designers and supervisors in design and construction (characterized as information filtering, prioritization and comprehension).	Humans are in charge of assessing and interpreting the information of design and construction, and generating predictions.	The machine or robot does not execute or help in rating duties. Humans should accomplish everything.	Only humans could carry out decisions.
2	Humans are the primary designers and supervisors for obtaining and reviewing all information related to design and construction, with machine or robot backup in case of an emergency.	Humans are the primary source of forecasting and analysis for contingencies. To interpret the design and construction is the duty of the humans.	All decisions are made by humans, however machines or robots can be employed to assist.	Humans are the main operators with machine or robots covering contingencies.
3	The machine or robot is in charge of obtaining and showing unstructured, unprioritized data to humans. The humans are still the primary observer of data for the purpose of design and construction.	The machine or robot is the main analyzer for predictions with humans shadowing for emergencies. But humans still have the responsibility in interpreting data related to design and construction.	Analyses and Evaluations are performed by both humans and computers. Human outcomes are seen as superior.	With human confirmation, the machine or robot performs the decisions. Human shadows contigency.
4	The machine or robot is in charge of obtaining and presenting design and construction data for the humans. However, it emphasizes non-prioritized, pertinent data for the human.	The machine or robot evaluates the information and forecasts the activities in design and construction. Despite the fact that the human is in charge of interpreting the data.	Evaluations are performed by both humans and computers. The computer's output is supeiror to human's.	Prior to actually executing the decisions in design and construction, the machine and robot provides the human a pre-programmed limited time to veto.
5	The machine and robot is in charge of collecting data for the humans and it only shows filtered non-prioritized items to human.	The machine or robot evaluates the informatin related to design and construction and overlays predictions with evaluation. For contingencies, the humans shadows the interpretations.	The machine or robot is in charge of evaluations, showing the humans all outcomes, including reasons.	Before execution, the machine or robots permits the humans context-dependent limited time veto. Humans shadows for contingencies.
6	The machine or robot collects, filters, and organizes data related to design and construction that is shown to the human.	The machine or robot evaluates the design and construction, and overlays analysis with prediction, and display entire outcomes to the human.	The machine or robot performs evaluations and shows a reduced set of rated possibilities in design and construction to the human while explaining "why" choices were reached.	The machine and robot operates autonomously, notifies the human, and permits for overriding following operation in design and construction. Humans are contingency shadows.
7	The machine or robot collects, filters, and prioritizes information in design and construction without presenting it to the human, only displaying a "program operating" flag.	The machine or robot analyzes, anticipates, interprets, and documents comprimizations to provide a conclusion that is only shown to humans if the conclusion matches the specified environment.	The machine or robot is in charge of evaluations. The machine or robot conducts overall rankings and shows a limited subset of ranked possibilities to the humans without explaining "why" judgments were chosen.	The machine or robot performs autonomously in design and construction and just tells the human when necessary. It enables override capability following operation. Humans are contingency shadows.
8	The machine or robot collects, filters, and ranks information related to design and construction without presenting it to the user.	The machine or robot anticipates, analyzes, and integrates data related to design and construction to provide a conclusion that is not visible to humans.	The machine or robot is in charge of evaluations, making the final decisions in design and construction. However, the findings are not displayed to the humans.	The machine and robots runs autonomously with no human intervention.

It arose from a partnership between scientists and engineers who wanted to overcome the difficulties of constructing sophisticated robot systems. ROS was first supported by the Stanford Artificial Intelligence Robot (STAIR) project before becoming open source, resulting in widespread acceptance and expansion. The ROS community grew swiftly, garnering contributions from researchers, amateurs, and industry experts all around the world. ROS has advanced significantly over the years, with frequent upgrades and new features introduced to improve its abilities (Maldonado, 2012).

The PR2 (Personal Robot 2) produced by Willow Garage, a robotics research center, is one famous example that demonstrates the capabilities of ROS (Kalakrishnan et al., 2012). The PR2 is a multipurpose mobile manipulator robot that can execute a variety of jobs. It was created with ROS as its primary software framework in mind, allowing for the easy integration of numerous sensors, actuators, and algorithms. The PR2 project highlighted ROS's capacity to aid complicated robot system development by providing a strong and adaptable software framework, allowing scientists and programmers to concentrate on a high degree of tasks. Another notable use of ROS is in the realm of self-driving cars. ROS was used as the primary foundation of the RobotCar research at the University of Oxford's autonomous driving platform. ROS supplied the researchers with the necessary equipment and libraries for awareness, planning, and management, allowing them to develop a viable autonomous car. ROS's usefulness in handling the complexity of autonomous systems and aiding the development of complicated algorithms for perception, mapping, and movement was proved by the RobotCar project (Milford, 2012).

2.5.2 Concepts of ROS

The main concept in ROS includes:

- Nodes and Communication
- Topics and Messages
- Services and Service Calls
- Packages and Libraries
- Launch Files and Parameter Server
- Visualization and Debugging Tools
- Simulation and Testing:

Nodes and communication are at the foundation of the ROS. A node is an application module that executes a particular duty inside a robotic system. These nodes are modular in nature, containing functions such as input from sensor processing, control algorithms, and actuator control. ROS communication is supported via a publish-subscribe messaging model. Nodes can publish data-containing messages on specified subjects, and other nodes are able to subscribe to specific topics to obtain the messages. This decoupled communication approach encourages loose coupling and allows multiple nodes to be integrated into a coherent robotic system.

Topics in ROS serve as pathways for nodes to exchange information. A topic denotes a particular subject or information stream, including sensor data or control directives. Nodes can send messages on a certain topic, containing data structures described by their message type. Other nodes that want to receive such data can subscribe to the topic while processing the incoming messages. Node decoupling through topics enables flexible and scalable communication patterns. Messages are specified in a structured manner, usually in ROS message documents, to enable consistent data transmission between nodes.

Through services and service calls, ROS also offers a request-response communication pattern. Nodes are able to utilize services to deliver and request specific capabilities from other nodes. A service provider determines the service category, which determines the request and response message formats. Other nodes can then submit requests for services, referred to as service calls, to the relevant node, causing the requested functionality to be executed. The replying node handles the request and sends back a suitable response message. This synchronous communication method allows nodes to connect in a more complicated and interactive manner, making activities like configuration, parameter setting, and robot state queries easier.

ROS packages software components, which serve as a controlled system unit of code arrangement. Packages are used to group together similar nodes, files for configuration, launch files, and dependents. They enable developers to properly organize and share their code. Furthermore, ROS includes many libraries and tools, including navigation, understanding, and manipulation libraries, that extend the capabilities of ROS and aid developers in the development of complicated robotic systems.

ROS leverages launch files to make it easier to start many nodes and configure their settings. Launch files identify the nodes that need to launch and identify the variables, topics, and services connected with them. This functionality simplifies the setup of a robot system and allows for rapid deployment. Furthermore, ROS has a parameter server, which serves as an integrated worldwide storage space for variables. Nodes may read and write variables from the server, enabling robot system runtime setup and dynamic reconfiguration.

ROS enables visualizations and debugging features that assist in the investigation and analysis of robotic system behavior. Users can explore data from sensors, robot simulations, and trajectories using Rviz, a 3D visualization tool. It aids in the verification of perception techniques and overall system behavior. Furthermore, ROS includes logging and debugging applications that allow for data recording and playback as well as continuous tracking of system signals and troubleshooting data.

ROS facilitates modeling and testing, both of which are essential for creating and evaluating robotic systems. Gazebo, a robust physics-based simulator included in ROS, enables developers to construct simulated spaces and test algorithms before deploying them on actual hardware. This skill contributes to the reduction of costs and hazards involved with real-world testing. ROS also includes a testing framework known as rostest, which allows the development of unit and integration tests for ROS nodes, assuring the accuracy and reliability of the provided capabilities.

2.5.3 Goals of ROS

The following is a list of ROS's philosophical objectives (Quigley, 2009):

Peer-to-peer: A system created using ROS comprises several processes that may operate on various hosts and are connected in real time via a peer-to-peer topology. The advantages of the multi-process and multi-host architecture may also be realized by frameworks based on a central server, such as CARMEN (Montemerlo et al., 2003), but a central data server is troublesome if the computers are linked in a heterogeneous network.

Multi-lingual: Numerous individuals prefer certain programming languages over others when developing code. These preferences are the outcome of individual trade-offs between programming time, debugging simplicity, syntax, runtime effectiveness, and a variety of other factors, both technical and cultural. We created ROS to be language-neutral for these reasons. C++, Python, Octave, and LISP are the four very diverse languages that ROS now supports; further language ports are in varying stages of completion. The messaging layer is where the ROS standard is located. Most major languages have adequate implementations of XML-RPC, which is used for peer-to-peer relationship negotiation and configuration. We prefer implementing ROS directly in every intended language rather than providing a C-based implementations with stub interfaces produced for all popular languages in order to better adhere to each language's standards. The Octave client is constructed by encapsulating the ROS C++ library, although there are times when it is more practical to add integration with a new language by using an already-existing library. ROS uses an easy-to-use, language-neutral interface defining language (IDL) to specify the messages transmitted between modules to facilitate cross-language development.

Tools-based: Instead of creating an isolated development and runtime environment, we have chosen a microkernel design to manage the complexity of ROS, which makes use of many small tools to build and run the many ROS components. These tools carry out a variety of activities, such as navigating the source code tree, getting and setting configuration parameters, visualizing the peer-to-peer relationship topology, calculating bandwidth usage, plotting message data graphically, and creating documentation automatically, among other things.

Thin: Most current robotics software projects have drivers or algorithms that could be used outside of the project, similar to those put forth in (Makarenko et al., 2007). Unfortunately, for a number of reasons, a large portion of this code is now so intertwined with the middleware that it is challenging to "extract" its functionality and reuse it elsewhere. Since almost all complexity is contained in libraries, and only tiny executables that expose library functionality to ROS are produced, code can be extracted and reused more easily than it was intended to. Additionally, when code is incorporated into libraries, unit testing is frequently much simpler because standalone test programs can be created to test various library features. ROS re-uses code from multiple other open-source projects, such as the drivers, navigation system, and simulators from the Player project (Vaughan & Gerkey, 2007) vision algorithms from OpenCV (Bradski & Kaehler, 2008), and planning algorithms from OpenRAVE (Diankov & Kuffner, 2008), among many others. In each instance, ROS is used with the least amount of wrapping or patching possible, merely displaying several configuration choices and to move information into and out of the relevant applications. The ROS development system may continually update the source code from outside repositories, integrate patches, and other things to take advantage of the ongoing community advancements.

Free and Open-Source: The complete source code for ROS is accessible to everyone. This is essential to making the entire software stack's debugging process easier. Although private systems like Microsoft Robotics Studio (Jackson, 2007), and Webots (Michel, 1998) have many positive qualities, we believe a fully open platform is incomparable. This is especially true when hardware and several software tiers are created and tested concurrently. The BSD license, which permits the creation of both commercial and non-commercial enterprises, is used to distribute ROS. ROS doesn't demand that components connect together in an identical executable; instead, it uses inter-process connections to send data across modules. As a result, systems created using ROS are able to use fine-grain licensing for all of its separate parts: individual modules can contain software covered by a variety of licenses, such as the GPL, BSD, or proprietary, but license "contamination" ceases at the module boundary.

2.5.4 Advantages of ROS in AEC

ROS uses a modular design strategy and nodes, which are software components that encapsulate different capabilities. These nodes exchange messages to enable software modules to be reused and seamlessly integrated across various robot systems. The design process is sped up by this modular architecture, which also makes quick prototyping possible. Additionally, Since ROS is open source, a robust developer community is fostered, encouraging cooperation and information exchange. Users can share code, algorithms, and techniques within the ROS structure, accelerating the creation of original robotic applications. The development of complicated robot systems has been made easier thanks to the abundance of libraries and packages that have been produced because of this collaborative environment.

Developers are freed from worries about particular hardware specifications thanks to ROS, which abstracts away underlying hardware issues. This lack of dependence on specific hardware makes it easier to design software that can run on several robotic systems. Further easing the entire design process, ROS offers standard connections for the seamless integration of sensors and actuators. Moreover, the "Gazebo" simulator and the "rostest" framework are just a couple of the simulations and testing tools that ROS integrates. With the help of these technologies, engineers may build virtual environments to test robot design and algorithms prior to actual deployment. Simulation increases the effectiveness of the manufacturing cycle by reducing expenses and hazards related to real hardware testing.

ROS facilitates distributed computing, allowing for the development of massive robot systems. Developers can divide computational jobs over a number of machines or robots by utilizing the ROS communication architecture, enabling simultaneous execution and the best possible resource use. Applications combining multi-robot systems and manufacturing automation benefit most from this scalability. In addition to, rapid prototyping and iteration are made possible by ROS's benefits of versatility, hardware independence, and simulation capabilities. Rapid construction and evaluation of various robot arrangements, algorithms, and control schemes is possible for developers. Advanced robotic systems can be developed more quickly and with greater inventiveness thanks to this iterative development process.

2.6 Robotics in AEC: Case Studies

2.6.1 Introduction

Robotics integration in architecture and construction has made considerable advances in recent years, influencing traditional construction industry, and opening up new opportunities to improve construction productivity and efficiency. Robotic systems provide several advantages to design and construction, including enhanced precision, increased safety, and more rapid building processes. This section seeks to offer an overview of robotics applications in architecture and construction, emphasizing important achievements and investigating their consequences.

Automation and robotics have been employed by the design and manufacture through assembly and servicing. A robot is a reprogrammable, multifunctional manipulator designed to move material, parts, tools, or specialized devices through variable programmed motions for the performance of a variety of tasks. Robots are used for computer-aided design and generative simulation throughout the design process, allowing architects to explore complicated geometries and new design approaches (Gramazio & Kohler, 2008). They can also help with building component manufacturing, utilizing sophisticated robotic fabrication techniques to construct complex structures (Johns, Kilian, & Foley, 2014). The application of robotic technologies in construction enables accurate assembly of architectural components (Huang, 2016). Furthermore, robots may also be used for repair and examination employment opportunities, assuring the long-term durability and the sustainability of built buildings (Cote et al., 2013).

The implementation of robots in architecture and construction has been motivated by the advancements in robotic technology, the demand for increased efficiency and production, and the need to solve labor shortages and safety problems. The increased availability of low-cost and versatile robotic platforms, combined with advances in sensing and control technologies, have permitted the completion of complicated building duties that were previously difficult or time-consuming for human laborers by themselves (Haddadin et al., 2016). Robotics can alleviate worker shortages in the construction industry by boosting human employees' skills, enabling them to focus on more sophisticated and strategic parts of the building process (Tenneti et al., 2016). Furthermore, incorporating robots into the building process assists enhancing safety by reducing human employees' exposure to dangerous situations and repetitive jobs (Chen et al., 2017).

In conclusion, the advent of robots in architecture and construction has resulted in dramatic transformations, opening up new avenues for efficient design, production, assembly, and maintenance of completed buildings. This section goes further into the current literature to investigate the cutting-edge technologies, approaches, and practices used in integrating robotics in architecture and construction. Gaining an understanding of the accomplishments and problems in this industry allows for significant recommendations for directing the future development and implementation of robotic systems in design and construction.

Enclosed herewith are a series of meticulously documented case studies, each distinctly embodying a unidirectional mode of communication between human operators and robotic systems. It is imperative to note that these case studies exemplify a paradigm wherein the robots' functionality is confined to the sequential interpretation and execution of programmed instructions. Consequently, the level of automation achieved in these specific scenarios is significantly constrained, necessitating a thorough and in-depth exploration in forthcoming sections to shed light on this limitation and its potential ramifications.

2.6.2 LightVault

The Light-Vault project, completed by researchers from Princeton University and sponsored by Skidmore, Owings, and Merill, provides a novel robotic construction strategy. As it is shown in Figure 2-3, the researchers devised a technique for the two robots to work together to construct a vault without scaffolding. They employed two industrial Universal Robot arms to construct a magnificent vault which is 7 feet tall, 12 feet in diameter, and 21 feet long, composed of 338 translucent glass bricks. Since this inquiry, prototyping, and final installation occurred in different locations, one of the most significant aspects of the research was to develop a production method that could be easily adaptable to different robotic setups. As it is shown in Figure 3, LightVault used less materials. It avoided the requirement for buttresses or scaffold all throughout construction, and it increased the vault's fundamental levels of performance. They established the idea of interchangeability in robotic construction. They then built on it by addressing the four key issues in the development of the Light-Vault construction process: (1) prototype scalability, (2) gripper development, (3) path design and ordering, and (4) fabrication tolerances.

The impact of the LightVault project extends beyond its successful construction. It serves as a pioneering example of how innovative robotic techniques can be employed in construction, potentially reshaping the future of design and construction practices. This project opens new avenues for research and development in the field, with the potential to improve efficiency, reduce costs, and enhance the sustainability of construction projects worldwide. However, in this project, human engineers played a pivotal role in crafting the code that precisely guided the robots. It was a one-way direction, where humans wrote the code, and the robots executed it flawlessly, showcasing the synergy between human ingenuity and machine precision. Even in this project, where the code was meticulously designed by human engineers, the LightVault project serves as a pioneering example of how innovative robotic techniques can be employed in construction, potentially reshaping the future of design and construction practices.

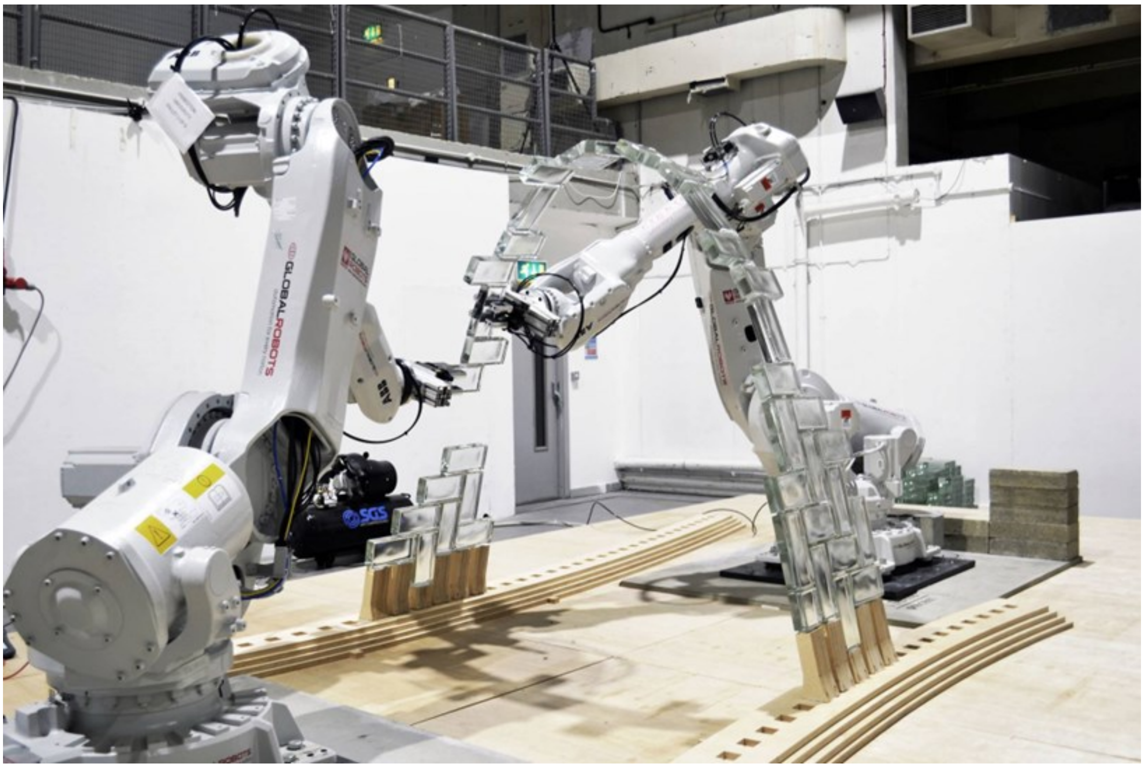


Figure 3: LightVault (Xi Han, 2020)

2.6.3 Robotic Timber Construction

The pioneering project in robotic timber construction, known as the Robotically Fabricated Structure (RFS) pavilion, was successfully completed in 2021. This project introduces a new process for robotically fabricating customized timber subassemblies using standard timber elements. The research focuses on a nonstandard timber structure consisting of four bespoke subassemblies, including three vertical supports and a Zollinger roof structure, to validate the proposed process. The fabrication setup comprises two robotic cells designed for specific assembly routines. As shown in Figure 4, the first cell features a 4-axis portal robot with a telescopic base mounted on a 2-axis gantry system. It includes a picking station, assembly platform, and automatic tool-changing station. The second cell consists of a six-axis robotic arm mounted on an external linear axis. It incorporates a picking station, table saw, and assembly platform. Both setups utilize custom end-effectors such as a pneumatic gripper with a nailing gun and a circular saw for trimming timber elements (Adel et al., 2021).

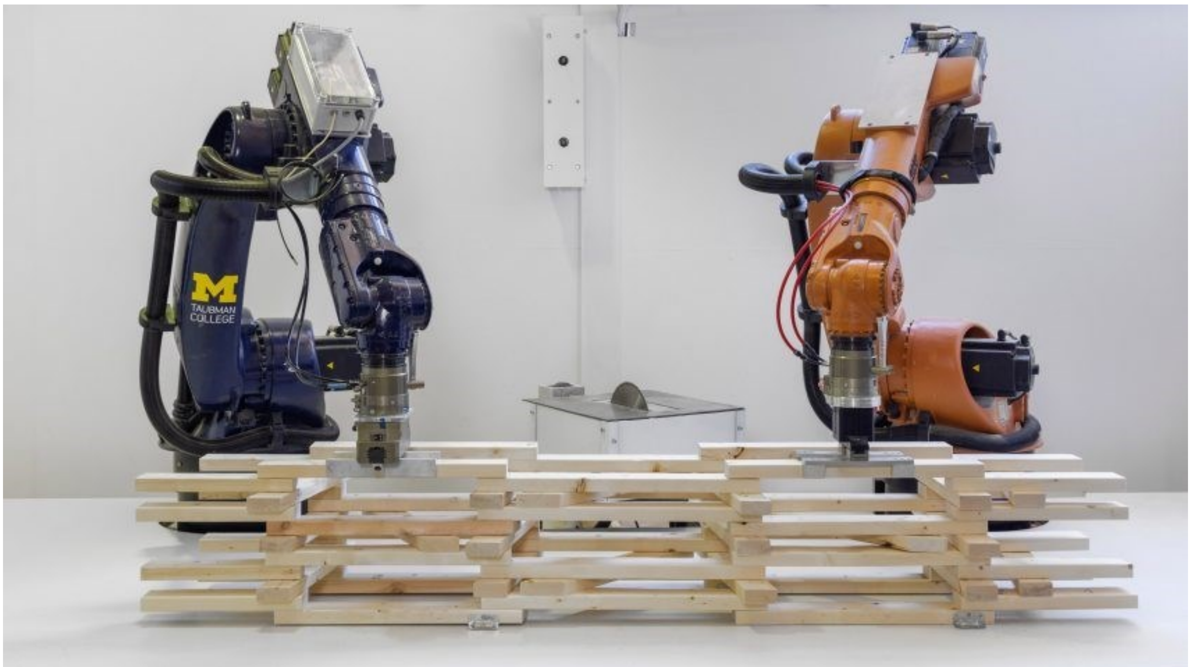


Figure 4: Robotic Timber Construction (Adel et al., 2021)

Each fabrication setup has specific constraints that influence the design. The case study, a pavilion, demonstrates how the proposed approach can be applied to design and assemble different building components within these constraints. For the roof structure, the first fabrication cell is used. The timber elements are precut, and the automated assembly process involves the robot picking up elements, placing them on the assembly platform, nailing them to the structure, and adjusting their orientation. The portal robot also attaches the long edge slats and trims their corners using the circular saw. The industrial robotic arm is utilized for the assembly of the vertical supports. Due to the arm's reach limitations, each support is divided into two halves, fabricated separately, and then connected. The process involves the robotic arm cutting timber slats with the table saw, placing them on the assembly platform, and joining them to the existing timber layers using screws. This approach demonstrates the feasibility of the proposed robotic fabrication process for creating bespoke timber subassemblies, allowing for efficient and precise assembly within the defined constraints of each setup.

The RFS pavilion project demonstrates the feasibility of the proposed robotic fabrication process for creating bespoke timber subassemblies. It not only showcases efficient and precise assembly within the defined constraints of each setup but also offers valuable insights into the future of sustainable and technologically driven construction practices. By merging human expertise with robotic precision, this project sets a pioneering example, potentially reshaping the landscape of timber construction and design methods, ultimately promoting sustainability and precision in the construction industry. It's noteworthy that the robots in this project were not equipped with the ability to sense or provide feedback from human operators during the operation, like many other task-based robotic operations, they followed pre-programmed instructions precisely. Additionally, the robots were specifically programmed to pick up wood pieces in a specified location with a predefined orientation, highlighting the project's meticulous attention to detail and precision in timber assembly.

2.6.4 The Programmed Wall

The Informed Wall Project was created in 2006 at ETH Zurich. The idea was to evaluate the architectural potential of brick walls by building them with robot arm. Students conducted this experiment using a six-axis robot with a 3×3×8m intervention area that can construct architectural elements at scale (Figure 5). Another experiment objective was utilizing various materials, procedures, and shapes during the construction process. The build was made to exclude any exogenous interferences, which include wind, spatial obstructions, and human interaction, and to allow the robot arm to travel to any location in three dimensions and do all duties specified in the Edeffector software (Bonwetsch et al., 2017).

Using the MAYA program, it created a computer script that could convert CAD data into coordinates. Numerous wall prototypes were created using the software and building materials, concluding that the robot arms may be used to build brick walls with common mistakes. The team concluded that a greater expenditure on hardware and software would be necessary to perform more complicated geometric patterns, as using bricks limits the intricacy of the ways to be made (Bonwetsch et al., 2017).

Figure 5: The Programmed Wall (Bonwetsch et al., 2017)

By removing the constraints of traditional construction methods, it opens doors to new possibilities in design, enabling architects to envision and create structures that were once deemed unattainable. Moreover, this project underscores the role of technology in shaping the future of construction, where robotic arms serve as tools of precision, executing tasks with unwavering accuracy and consistency. It's important to note that the Programmed Wall project operated in a one-way direction, where human engineers coded the instructions without real-time feedback or sensing from the robot arm. The robot executed these meticulously crafted codes with precision and reliability, showcasing the potential for pre-programmed robotic construction in architectural projects.

In conclusion, the Programmed Wall stands as a pioneering example of how robotics can redefine the boundaries of architectural design and construction. Through innovation, experimentation, and the fusion of digital and physical realms, it provides a glimpse into a future where robotic technology plays an increasingly pivotal role in shaping our built environment.

2.6.5 The brick Labyrinth

The inception of "The Brick Labyrinth" project emerged from a collaborative three-month endeavor at ETH Zurich in 2016, marking a significant milestone within the Master of Advanced Research in Design and Digital Fabrication program. This project represented the studio's inaugural substantial undertaking and a distinctive achievement as the first large-scale architectural product constructed through multi-robotic automation. "The Brick Labyrinth" project embodies an extension of these groundbreaking concepts (Piskorec et al., 2018). As shown in Figures 6 and 7, the project unfolds across three pivotal phases: a research and design stage, a computational setup stage, and a fabrication stage, as meticulously detailed by Piskorec and colleagues in 2018 (Piskorec et al., 2018).

The major objective of this task was to create a sizable labyrinth structure at the RFL using no more than 12,000 dry stack bricks and an area of about 100 square meters. The team did this by mounting two ABB robotic arms on one bridge (Piskorec et al., 2018). Two spiraling walls, approximately eight by 10 meters and made of more than 10,000 bricks, were the design chosen to begin the construction process. The method produced a tortuous path that led to various spatial sensations (Piskorec et al., 2018). A Python setup built on COMPAS, a framework based on open source created by Philippe Block's research team at ETHZ, was used to develop the computational setup step. The software is divided into three main sections: (1) the computational design stage configures the final geometry based on each wall axis curve to resolve shift problems. (2) the structural stability analysis involves importing the building's mesh data and brick assembly into COMPAS to conduct a straightforward equilibrium examination, followed by a multiple scenarios analysis. (3) the robotic fabrication simulations and online robot control section analyzes the building sequence, robot movement, and stability (Piskorec et al., 2018).

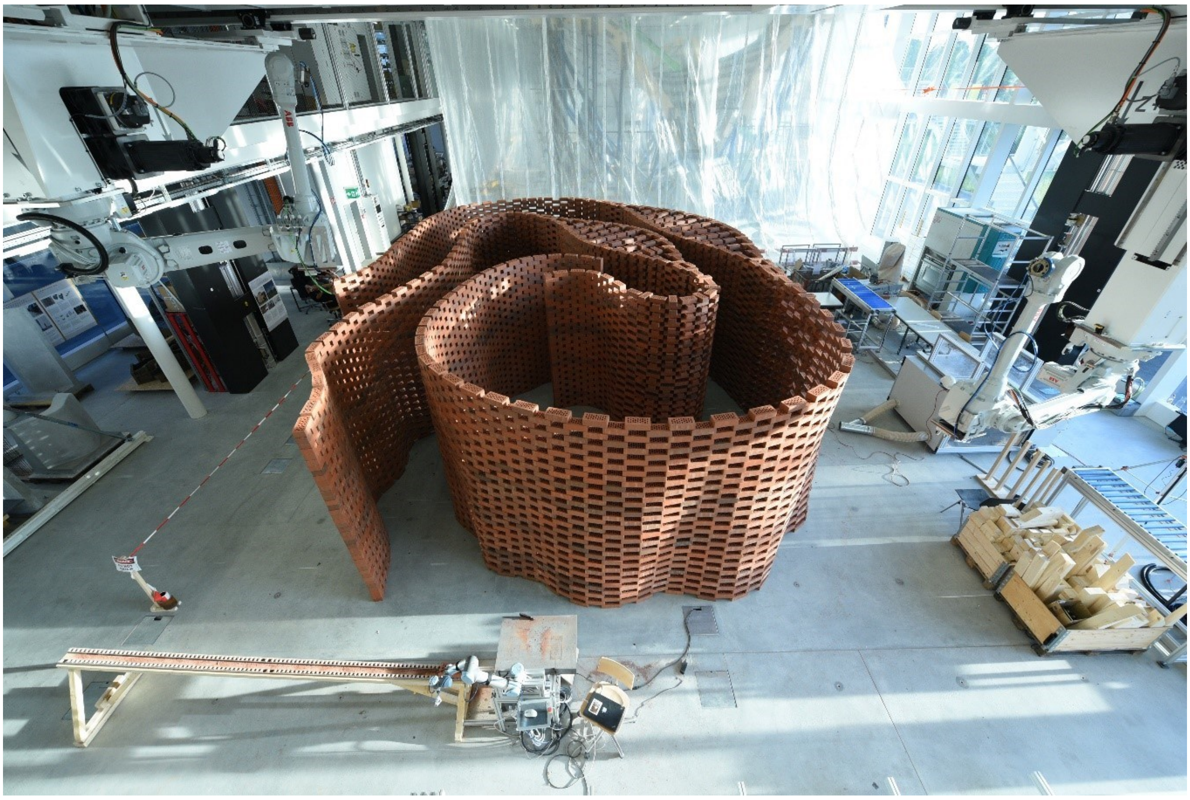


Figure 6: brick Labyrinth project (Piskorec et al., 2018)

Figure 7: Brick Labyrinth project (robotic process)(Piskorec et al., 2018)

The final step was the fabrication procedure, during which the outcome took shape. ABB IRB 4600 2.5 robot arms were mounted upside down on both sides of the employed bridge's two telescoping Z-axes, which allowed for independent three-dimensional flexibility. To speed up the construction process, it was required to create a brick dispenser known as a brick magazine that could pick up to eight brick stacks and drop them one at a time. The robot arm builds the desired construction by placing each brick in set order. This project introduces methods for a multi-robot automatic brick layering procedure for large-scale, custom-made architectural components to produce complicated shapes in a fully digital environment (Piskorec et al., 2018).

2.6.6 Robotic Masonry Arch

2.6.6.1 Introduction

Historically, some vaults have been constructed without formwork. However, enough information for the actual process of making these masonry vaults is rarely available. There are some researches that examined the historic documents found so far and discovered that the basic principle for constructing masonry vaults without temporary support relies on the self-supporting courses which are stable on themselves by forming the arch shape (Wendland, 2007). There are different techniques for constructing masonry vaults that do not even require a minimum guideline during construction, like the leaning bricks technique and the corbelled vault construction technique. The leaning brick technique was first introduced in Egypt and northern Sudan when brick architecture was flourishing. The technique was revived by an Egyptian architect Hassan Fathy and implemented in his works; a Craft School and Agriculture Cooperative Center at New Baris Egypt, and the Cultural Center of Garagus (Serageldin, I. 2007; Miles, M., 2006). In this technique, the first course is laid leaning against the wall or a previously constructed arch using mortar. Other courses are laid one after another leaning against the previous one. The vaults are built in the shape of a catenary to reduce the risk of buckling due to the longer courses. This type of vaulting does not require any temporary support during assembly (Gargiulo et al., 2006; Ponce et al., 2004).

Form-finding is a process in which new forms of a statically equilibrated structure are generated by controlling the parameters like site conditions, height, shape of the plan, etc. Before the invention of computer-based computational tools, designers used analog models of hanging chains and hanging cloth to study forms of arches and shells. In this study, Algorithms will be tested to build masonry corbel arches that can be built without centering based on the principle of equilibrium conditions for each masonry unit in Grasshopper with python scripting. A variety of masonry vaults could be generated from the algorithm (Neupan and Liu, 2021).

Moreover, there are some factors to Consider in Arches with Robotics in a Practical Aspect:

- Types of Robots: The selection of an appropriate robot type plays a crucial role in achieving successful arch construction. Considerations should be given to factors such as the robot's reach, payload capacity, and accuracy.
- End Effector Size and Limitations: The end effector, or the tool attached to the robot arm, should be carefully chosen based on the desired functionality and constraints of the arch construction. Factors such as rotation angle and weight limitation are important to ensure the end effector can perform necessary tasks like picking up and placing bricks effectively.
- Material Properties and Dimensions: The size and weight of the bricks used in the arch construction must be taken into account when designing the robotic system. Understanding the properties of the materials, such as their dimensions and weight, is essential for proper programming and accurate positioning of the bricks.
- Safety Factor: To account for variations in brick sizes, it is recommended to include a safety factor of around 10 percent in the robotic arch construction process. This ensures that any inconsistencies in brick dimensions do not compromise the structural integrity and stability of the arch.
- Distance of Bricks in Each Layer: In the programming phase, the generated code in platforms like Grasshopper determines the distance between bricks in each layer of the arch. It is essential to carefully define and control this distance to achieve the desired structural stability and aesthetic appearance of the arch.

- Number of Bricks and Arch Height: The total number of bricks required for constructing the arch, as well as the desired height of the arch, should be considered during the planning stage. These factors affect the overall scope and complexity of the robotic arch construction process and should be accounted for in the system design and programming.
- Clash Detection in Simulations: Clash detection is a critical aspect when using KUKA PRC (Parametric Robot Control) for arch construction. Clash detection algorithms help identify potential collisions or interferences between the robot arm, end effector, and the arch structure. By incorporating clash detection into the programming, collision risks can be minimized, ensuring safe and efficient operation of the robot during the construction process.

Overall, successful arch construction with robotics requires careful consideration of robot type, end effector specifications, material properties, safety factors, brick distances, arch height, and clash detection algorithms. By addressing these factors in a comprehensive manner, architects and engineers can optimize the robotic system's performance and achieve accurate and aesthetically pleasing arch structures. Another important project which is important to mention is “Exploring Masonry Vault Forms Without Centering” by Babita Neupane, Kent State University in 2021.

2.6.6.2 Exploring Masonry Vault Forms Without Centering

The masonry vault forms constructed without the traditional use of centering takes center stage was explored by Neupane (2020) (Figures 8, 9, and 10). Drawing inspiration from historical construction techniques and contemporary form-finding methods, this research delves into the possibilities of designing and building masonry arches, vaults, and domes that rely on self-supporting courses. (Neupane, 2020). Each masonry unit is meticulously considered within its equilibrium condition, forming the foundation of this innovative approach.

Central to this thesis is the development of algorithms that enable the creation of self-supporting masonry structures. The research leverages the principles of equilibrium to formulate these algorithms, addressing two distinct scenarios: one where bricks remain unbonded with mortar and another where mortar plays a role in securing the structure. Additionally, the investigation extends to the behavior of these structures when subjected to additional vertical loads, further enriching the understanding of their stability and design possibilities. Additionally, beyond theoretical exploration, this research translates its findings into practical applications. Through the utilization of the Python component in Grasshopper, a specialized tool is developed, employing the algorithms proposed in this thesis. This tool empowers architects and designers with the capability to engage in parametric design, generating new and innovative forms for masonry vaults that can be constructed without the need for centering. By facilitating the creation of these novel architectural elements, this research opens up exciting avenues for the integration of self-supporting masonry structures in contemporary design and construction practices.

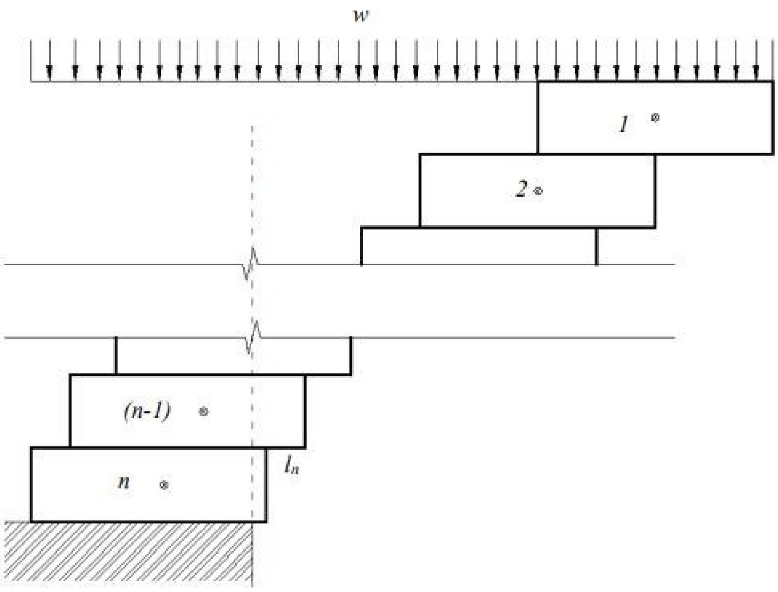


Figure 8: Equilibrium of 'n' number of bricks when not bonded with mortar with additional load case (Neupane, 2020)

The exploration of self-supporting masonry structures, driven by principles of equilibrium, holds promise for revolutionizing robotic operations in construction. By understanding the inherent stability of masonry elements and formulating algorithms that enable their construction without centering, this research paves the way for the integration of robotics in masonry work. The ability to generate precise and stable architectural elements without the need for extensive support structures aligns with the goals of robotic operations—efficiency, accuracy, and adaptability. This project sets the stage for potential advancements in the use of robots in construction, where robotic arms can be programmed to execute intricate masonry tasks with greater efficiency, reduced material waste, and enhanced precision. Consequently, the impact of this project extends beyond architecture, potentially reshaping the future of robotic operations in the construction industry.

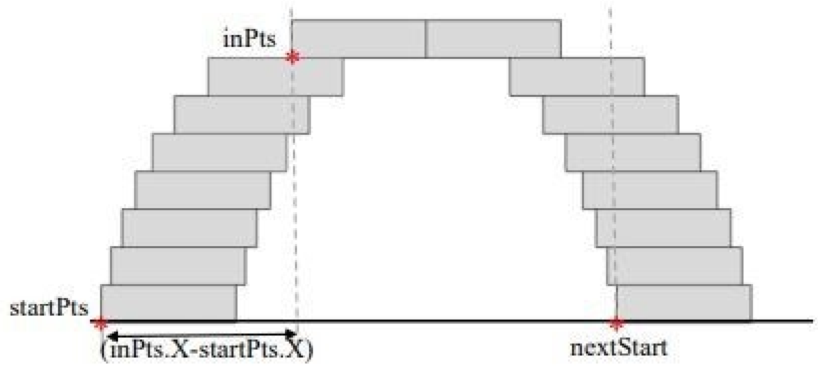
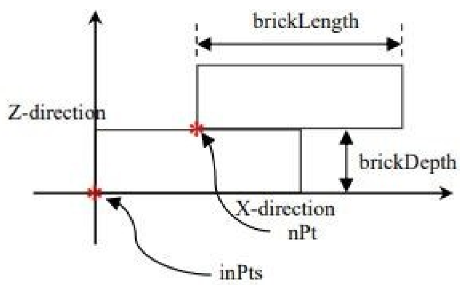


Figure 9: Image showing terminologies used in coding (Neupane, 2020)




```
#Section-1: method to create the arch on the left side of the structure
#creating first brick
#taking start point as inPts
startPts = inPts
newBrick = createBrick (startPts, brickLength, brickWidth, brickDepth)
brickList.append(newBrick)
#Method to stack bricks
for i in range(N-1,0,-1):
    x=(1/(2*i)*brickLength)
    offsetList.append(x)

for i in range(1,N):
    x = offsetList[i-1]
    print x
    nPt = rs.VectorAdd(inPts, [x,0,brickDepth])
    nBrick = createBrick(nPt,brickLength,brickWidth,brickDepth)
    brickList.append(nBrick)
    inPts = nPt
```

Figure 10: Determining a starting point for the right half of the arch (Neupane, 2020)

Chapter 3: Research Methods

3.1 Introduction

This chapter presents the research method for constructing a masonry arch using human-robot interaction and object detection techniques. This study explored the feasibility and effectiveness of integrating robotics with building construction. The research focused on leveraging the capabilities of a robot, namely the DofBot, to assist in the fabrication process while incorporating object detection using the YOLO (You Only Look Once) algorithm. The selection of the DofBot as our primary robotic platform for this research project was a result of careful consideration. While the Kuka robot is a well-established industrial robot known for its versatility, the DofBot offered advantages in terms of accessibility and affordability, making it an ideal tool for this study within the constraints of our research environment. Additionally, our choice of the DofBot provided an excellent opportunity to delve into ROS, which is widely recognized as a key technology for robotics research and development, which was not directly accessible in KUKA arms. This allowed us to enhance our proficiency in ROS while conducting this research.

Additionally, this section discusses utilizing Roboflow for dataset creation. Roboflow offers a user-friendly platform that simplifies the generation of annotated datasets by providing various tools for data preprocessing, augmentation, and labeling. This study explains how to create a comprehensive dataset of images and corresponding annotations for training the object detection model. This investigation delves into implementing the YOLO (You Only Look Once) algorithm for object detection. YOLO is a state-of-the-art deep learning algorithm known for its real-time object detection capabilities. We describe YOLO's architecture and working principle and the specific modifications made to adapt it to the masonry arch construction scenario. The trained YOLO model enables the robot to identify and locate bricks and other relevant objects within the construction site.

As shown in Figure 11 this study commenced with a thorough examination of the landscape of robotic applications within the fields of architecture and construction. Upon identifying research gaps and formulating specific research questions and objectives, the investigation progressed to the physical construction of an architectural structure using a robotic arm. Subsequently, the study extended its scope by incorporating object detection, specifically utilizing the YOLO framework, into the construction process. Moreover, the research introduced user-friendly interfaces to facilitate enhanced human-robot interaction, enabling human operators to make dynamic decisions and modifications during the construction process. This approach is a departure from the prevalent one-way direction typically observed in most robotic projects within the architecture and construction domains.

Figure 11: Thesis Workflow

3.2 System Requirements and Algorithm Workflow

In the realm of modern technological advancements, understanding the system requirements and algorithm workflow is paramount to the successful development of any technical solution. This section meticulously delineates the prerequisites and procedures that underpin the proposed system, elucidating the intricacies of its functioning. The algorithm workflow represents the orchestration of various components and processes to achieve a desired outcome. It signifies the step-by-step progression of the implemented algorithm, from data input to final result. This section aims to provide a comprehensive overview of the algorithm's inner workings, fostering clarity and transparency in its operation. The system requirements stand as the foundation upon which the entire system is built. This encompasses both hardware and software prerequisites. By identifying these requirements, potential bottlenecks or limitations can be anticipated and mitigated, ensuring a seamless execution of the algorithm. As depicted in Figure 12, the meticulous progression of dataset creation through to the final testing phase becomes readily apparent.

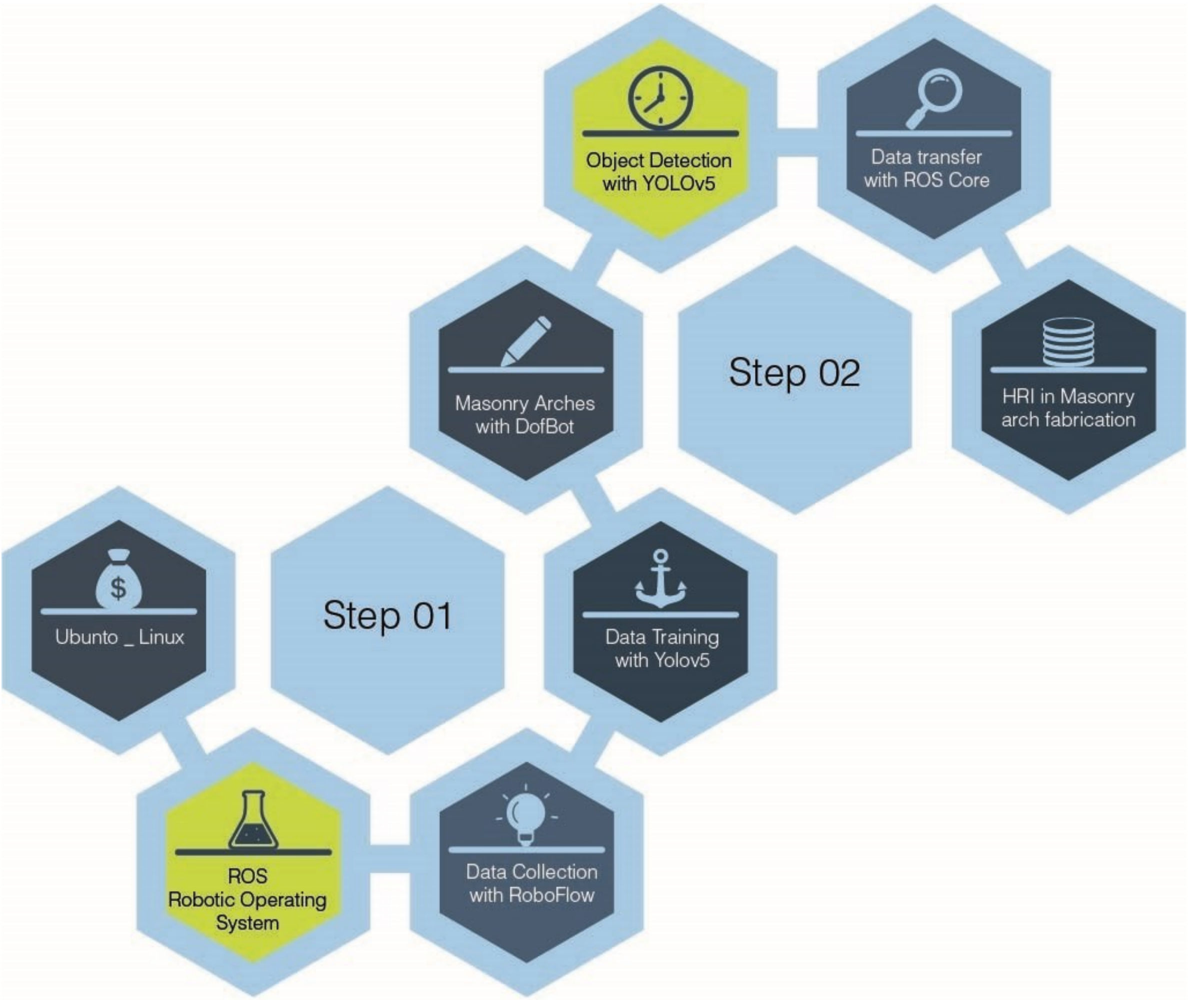


Figure 12: Detailed Workflow of thesis

3.2.1 Hardware and Software Requirements

A pivotal aspect of any technical system's deployment lies in the accurate identification and allocation of hardware and software resources. This subsection meticulously delves into the intricacies of the prerequisites that form the backbone of the proposed system, ensuring a holistic understanding of the infrastructure needed for its successful operation.

3.2.1.1 ROS Requirements

The ROS is the linchpin of our project's software framework. ROS not only facilitates communication between various system components but also enables the seamless integration of diverse algorithms. The version compatibility, libraries, and dependencies need to be satisfied for optimal ROS functionality. Moreover, a breakdown of the required configuration settings and communication protocols is also presented, ensuring a comprehensive grasp of ROS's role in the system.

3.2.1.2 DofBot Setup

The heart of the proposed system beats within the DofBot (Figure 13) – a versatile degree-of-freedom robot platform. To bring the DofBot to life, a meticulous configuration of both hardware and software elements is imperative. Concurrently, the software facet is explored, encompassing motor control algorithms, sensor interfaces, and calibration procedures that collectively govern the

Figure 13: DofBot

DofBot's behavior within the system. By engaging with these comprehensive guidelines for hardware and software prerequisites, the subsequent phases of implementation and testing are poised for success. These sections collectively reinforce the robustness of the proposed system, underpinned by a well-defined workflow and the necessary components to drive its execution.

3.2.2 Dataset Creation with RoboFlow

The evolution of technology has propelled the field of computer vision to new heights, demanding constant innovation in training methodologies and data acquisition techniques. Roboflow is a popular online platform for managing and preprocessing computer vision datasets, offering a variety of features such as annotation, image augmentation, and data conversion (Alexandrova, 2015). Worldwide developers and researchers have used Roboflow to accelerate their computer vision projects. Zheng et al. (2021) utilized Roboflow to preprocess a dataset of satellite imagery and reported significant improvements in model performance and efficiency with enhanced quality and usability of large-scale datasets. RoboFlow was used to generate the case study dataset in this project (Figure 14).

Figure 14: RoboFlow Diagram (dataset process)

3.2.2.1 Dataset for Bricks

Raw images of bricks were initially extracted from videos at a rate of 8 frames per second, a choice we made to balance computational efficiency and data richness, but this number can be increased to produce more images. These images were subsequently classified into five distinct categories representing different brick types (labeled as BrickTypes A to E, as illustrated in Figure

15) using RoboFlow's annotation tools. For each image, precise bounding boxes were manually annotated according to its corresponding category, as depicted in Figure 16.

In order to effectively prepare the dataset for training and evaluation, we partitioned it into three subsets: a training set, a validation set, and a testing set, allocating 70%, 20%, and 10% of the data to each subset, respectively. This partitioning strategy was meticulously designed to achieve a balance between various factors. The 70% allocation for the training set provides sufficient data for the model to learn the intricate details and features of the brick types, utilizing a total of 2,245 images. The 20% validation set serves to fine-tune the model's hyperparameters and prevent overfitting, which is crucial for optimizing model performance. The remaining 10% of the data, comprising 339 images, forms the testing set, allowing us to rigorously assess the model's accuracy and precision on unseen data, thereby providing a robust measure of its overall performance.

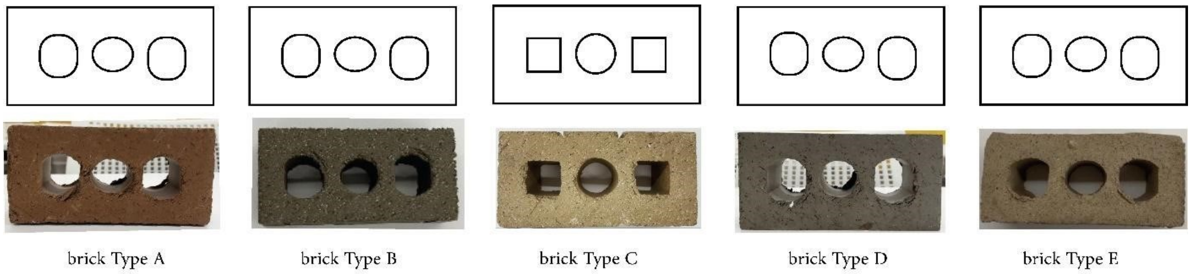


Figure 15: Brick types

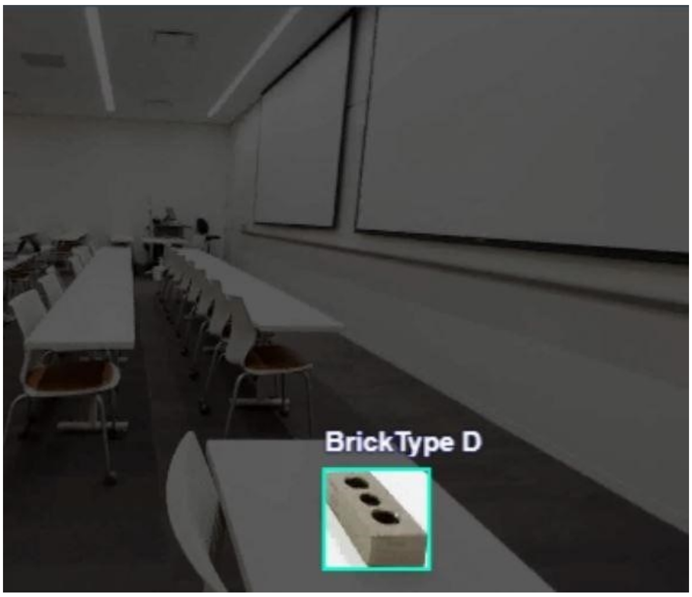


Figure 16: Example of bounding box for each brick

The resizing of all images to a uniform 640 × 640-pixel dimension was a step taken to ensure consistency in the dataset, resulting in a total of 3,220 source images across the five classes. This decision simplifies the model's architecture and enhances its efficiency. Figure 17 provides a visual overview of the workflow employed in the development of the model for this case study.

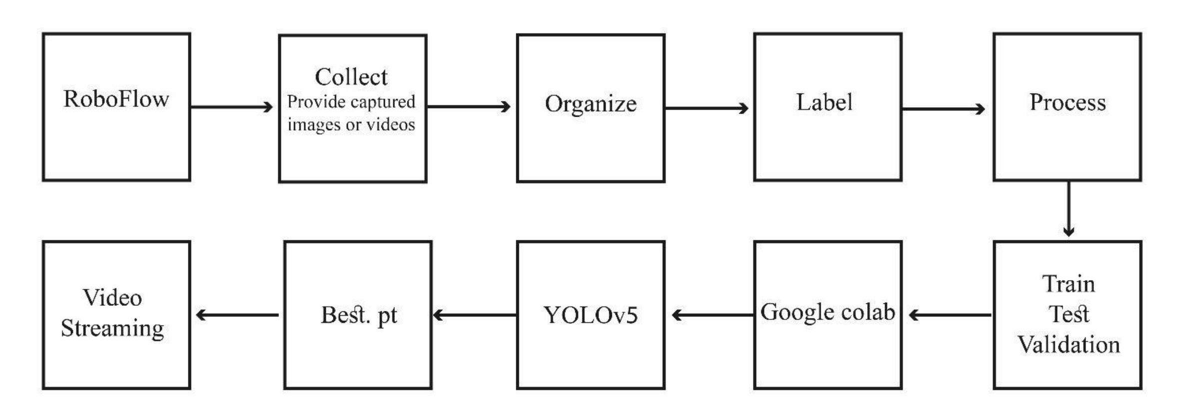


Figure 17: Dataset workflow using Roboflow and Yolo

3.3 ArchiTech Project

3.3.1 System Overview

In the realm of contemporary robotics, the ArchiTech system represents a significant advancement. Developed atop the acclaimed ROS framework, the system predominantly utilizes Python—a widely recognized and flexible programming language in the computational domain. The ArchiTech platform is architecturally structured around three salient modules:

- Human Interface:

  A nuanced interface designed for optimized human-computer interaction, ensuring that operators can effectively relay commands and decisions to the system.
- Object Detection:

  Leveraging the latest in algorithmic design, this module facilitates high-precision object recognition, enabling the system to discern minute details with impeccable accuracy.
- Robot Controller:

The nerve center of the system, responsible for orchestrating the array of robotic activities and ensuring seamless task execution.

As elucidated in Figure 18, the system offers operators an intuitive front-end GUI coupled with the robust capabilities of Rhino. Through this interface, they can make judicious selections pertaining to the architectural design and brick type for construction. Upon finalizing these selections, the requisite data is transmitted to the Robot Controller, which in turn crafts a comprehensive construction plan.

In its operational phase, the robot proactively seeks access to an integrated camera system. Through this, it deploys a refined object detection algorithm, thereby pinpointing the exact spatial coordinates and rotational angle of the designated brick. Post this identification phase, it commences the execution of the predefined construction plan. To ensure unparalleled precision in brick placement, the Robot Controller refers to a 3D virtual model superimposed at the desired position. This acts as a reference guide to verify the alignment of the brick. Any discrepancies observed prompt the Robot Controller to immediately transmit an alert to the operator, suggesting the necessity for manual intervention to recalibrate the robot's operations. This sophisticated feedback loop underscores the system's commitment to maintaining the highest standards of construction accuracy.

Among its myriad features, ROS facilitates services catering to a heterogeneous computer cluster. This encompasses a gamut of functionalities: from hardware abstraction and meticulous low-level device control to the seamless implementation of ubiquitously employed operations. Furthermore, ROS excels in enabling efficient message-passing between disparate processes and streamlining package management.

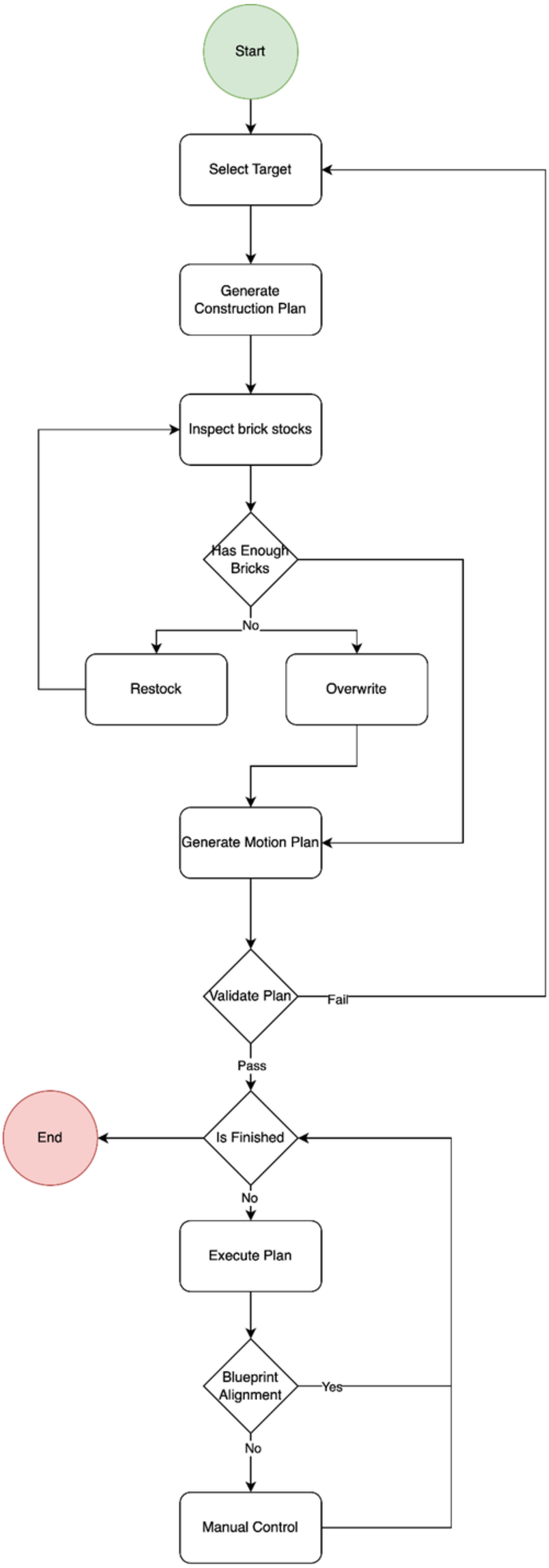


Figure 18: ArchiTech Project Flowchart

3.3.3.1 Graphical User Interface (QT framework)

A Graphical User Interface (GUI) is a visual platform that allows users to interact with software applications using graphical elements like icons, buttons, windows, and menus. GUIs are designed to simplify the user experience, making software more accessible and user-friendly. They play a crucial role in modern computing, enabling users to perform tasks without the need for complex text-based commands.

In the context of our research, the GUI is of paramount importance, and it is specifically implemented using the QT framework. The QT framework is a powerful and open-source toolkit for building graphical user interfaces. It is known for its versatility, enabling developers to create sophisticated and user-friendly interfaces for applications. What sets QT apart is its ability to offer cross-platform support, making it possible to develop GUIs that work seamlessly on various operating systems, including Windows, macOS, and Linux.

The choice of the QT framework for our GUI design reflects our commitment to creating an accessible and efficient user experience within the ArchiTech system. The framework's extensive library of tools and components allows us to craft a highly interactive and visually appealing interface that streamlines the interaction between the user and the system. Its adaptability across different platforms ensures that our GUI is accessible to a wide range of users, regardless of their operating system preferences. In essence, the implementation of the QT framework for our GUI is a pivotal component in our pursuit of enhancing human-computer interaction and usability within the ArchiTech system.

3.3.2.2 Rhino Computational Design Generator (RCDG)

The RCDG represents a pivotal component of our research, enabling the dynamic and practical generation of arches without the need for centralized designs. This innovative approach relies on a custom code developed within the Rhino and Grasshopper design environment, underpinned by Python scripting. The code operates based on brick dimensions and the desired number of bricks, allowing for the efficient creation of arches with remarkable adaptability.

The effectiveness of this code is discussed in this thesis, showcasing its potential use in the real-world utility. The code serves as a foundational element of the ArchiTech project, seamlessly integrated into our user interface. This integration empowers users to select the number of bricks and, in the background, leverages the code's capabilities to drive subsequent stages of the ArchiTech project, orchestrating robotic operations.

The synergy between Rhino, Grasshopper, and Python in this code is noteworthy. Rhino provides a versatile 3D modeling platform, while Grasshopper offers a visual programming interface for algorithmic design. Python, as the scripting language, bridges these components, enabling the code to take shape. This triad of technologies amalgamates to provide a sophisticated and flexible solution for the on-demand generation of arches within the ArchiTech system, contributing to the overarching objective of enhancing architectural and construction practices through innovative automation.

3.3.2.3 Open Motion Planning Library (OMPL) and MoveIt Package

In the field of robotics, motion planning is a critical aspect that determines the ability of a robot to perform tasks efficiently and safely in complex environments. Open Motion Planning Library (OMPL) and the MoveIt package in the ROS are two key tools that play a crucial role in addressing the challenges of motion planning for robots.

OMPL is an open-source software framework designed to address the intricate problem of motion planning in robotics. It offers a comprehensive set of tools, data structures, and algorithms for motion planning, making it a valuable resource for researchers and roboticists (LaValle, 2006). OMPL provides a flexible and extensible platform for developing motion planning algorithms that can be customized to suit specific robotic applications. This extensibility makes it particularly well-suited for academic research and for addressing the diverse requirements of industrial applications. Researchers and developers can utilize OMPL to experiment with and implement a wide range of motion planning algorithms to efficiently navigate robots in complex environments.

MoveIt, integrated within ROS, is a powerful and widely used software framework for motion planning, manipulation, and control of robotic systems. MoveIt builds upon the capabilities of OMPL, providing a higher-level interface and adding functionalities related to manipulation and control. This package is designed to simplify and accelerate the development of robotic applications by offering a unified platform for perception, control, and motion planning.

MoveIt's functionality spans from grasping and manipulation planning to collision detection, kinematic analysis, and trajectory generation. It seamlessly integrates with various sensors, robot models, and end-effectors, making it versatile for a broad spectrum of robotic applications. MoveIt is compatible with a wide range of robotic hardware and is well-supported by the ROS community. MoveIt simplifies motion planning tasks, enabling robot operators to specify goals and constraints more intuitively. This makes it a valuable tool for designing, testing, and deploying robotic systems in diverse environments. The package's extensive documentation and community support further contribute to its widespread adoption within the robotics community.

OMPL and the MoveIt package in ROS collectively provide a comprehensive solution for addressing the intricate challenges of motion planning in robotics. OMPL's flexibility and algorithmic richness enable researchers to develop custom motion planning solutions, while MoveIt builds upon this foundation to offer a unified and user-friendly interface for robotic manipulation and control. These tools are instrumental in advancing the capabilities of robots and driving innovation in various fields, from manufacturing and logistics to healthcare and autonomous systems.

3.4 Summary

This chapter provides a comprehensive framework for the ArchiTech project, an innovative venture into masonry arch construction through human-robot interaction and advanced object detection techniques. The chapter commences with an insightful introduction outlining the study's objectives, which revolve around the application of the DofBot robotic platform and the You Only Look Once (YOLO) algorithm for real-time object detection. The selection of the DofBot is underscored due to its accessibility and affordability, making it an ideal choice for this research's unique constraints. Additionally, the chapter delves into the creation of an extensive dataset using Roboflow, a versatile tool for data preprocessing, augmentation, and labeling, paving the way for training the object detection model.

Furthermore, Chapter 3 delineates the system prerequisites and algorithm workflow, offering a deep understanding of the technical infrastructure essential for the project's success. The ROS takes center stage as the linchpin of the software framework, enabling seamless communication between system components and the integration of diverse algorithms. The chapter also scrutinizes the hardware requirements for the DofBot, emphasizing the precise configuration of its components, both hardware and software, to ensure accurate operation. The dataset creation process, including the classification of construction bricks and the partitioning of data, is portrayed as pivotal, facilitating efficient training and testing of the object detection model. Scaling down the blocks replicates real-world operational constraints and enhances the model's adaptability, especially within the DofBot's limitations. This approach emphasizes the flexibility and practicality of the You Only Look Once (YOLO) model for object detection.

In the context of the ArchiTech project, Chapter 3 introduces the three salient modules, including Human Interface, Object Detection, and Robot Controller, which together constitute the architecture of this innovative system. The human interface is illustrated as a critical component, encompassing the GUI built using the versatile QT framework, offering a user-friendly platform for interaction. The RCDG emerges as a key element, allowing for on-demand arch generation, thus providing adaptability and practicality to the system. Finally, the role of the OMPL and the MoveIt package in ROS are elucidated. OMPL's open-source framework and rich array of motion planning algorithms empower researchers to tailor solutions for various robotic applications. MoveIt, seamlessly integrated within ROS, simplifies motion planning tasks, allowing for intuitive goal specification and aiding in the development of robotic systems with wide-ranging applications.

In essence, Chapter 3 establishes the foundations of the ArchiTech project by detailing the research methods, system requirements, and tools. It emphasizes the fusion of robotics and computer vision for masonry arch construction, underlining the significance of human-robot interaction within a dynamic construction environment, ultimately showcasing the innovation and practicality of this cutting-edge project.

Chapter 4: Results and Discussion

4.1 Introduction

In the era of advanced technology, the fusion of artificial intelligence and robotics has opened new vistas of possibilities, reshaping the landscape of automation and interaction between machines and humans. This chapter delves into a comprehensive discussion of the outcomes and implications of employing the YOLOv5 model in our research, particularly in the domain of brick detection. Beyond mere object detection, the thesis explores the intricate world of HRI, where our robotic system communicates with human operators through intuitive interfaces. This dynamic exchange goes beyond traditional one-way robotic operations, introducing a two-way flow of information, feedback, and decision-making that enhances the synergy between technology and human expertise.

This study presents the intricate details of YOLOv5 results, including precision, recall, and mAP scores, revealing the model's proficiency in identifying and classifying various brick types. These results are not confined to the theoretical realm but have profound implications for practical applications. The ability of the YOLOv5 model to detect bricks with exceptional accuracy ignites a spark of innovation in the realm of construction, particularly in the context of robotic operations. This chapter not only explores the success of the YOLOv5 model in detecting bricks but also sheds light on its role as a cornerstone in the implementation of brick placement within the thesis project. The precision offered by the YOLOv5 model empowers the robotic system to pick and place bricks with utmost accuracy, transforming a vision into a tangible reality.

While robotic operations traditionally follow predetermined instructions, this research takes a leap forward by incorporating a profound dimension of HRI. The robotic system isn't merely executing tasks; it actively engages with human operators, seeking their input and providing real-time feedback during operations. This interactive paradigm enhances the adaptability of our robotic system, making it capable of responding to human operators' dynamic needs and the uncertainties of real-world construction. Through interfaces designed to mimic the simplicity of a window, the system initiates dialogues with human operators, asking crucial questions such as material preferences and handling scenarios when resources run scarce. The result is a seamless interaction between humans and robots that transcends conventional approaches, ultimately fostering a collaborative environment where technology and human expertise merge harmoniously.

YOLOv5 results and their real-world application in brick detection, paving the way for the transformative power of HRI in the realm of robotics and construction. The journey navigates the precision of code and the subtleties of communication, unveiling a future where robotics and humans operate as a unified force.

4.2 YOLOv5 results

4.2.1 Precision and Recall

In evaluating the results of the YOLOv5 model, two crucial metrics, precision and recall, offer insights into its performance. Precision, often denoted as 'P,' measures the accuracy of positive predictions made by the model. In the context of brick detection, a high precision score indicates that the model correctly identifies a significant number of true positive instances, which are bricks in this case. Precision is a crucial metric because it tells us how reliable the model's predictions are. In this study, a precision score of 0.982 for BrickType A suggests that the model is highly accurate in labeling the correct bricks. On the other hand, recall, often denoted as 'R,' assesses the model's ability to find all relevant instances. A recall score of 0.954 for BrickType A implies that the model can sensitively detect nearly all instances of this brick type, which is vital in ensuring that no bricks are missed during the detection process. Therefore, the YOLOv5 model's impressive precision and recall scores underscore its efficiency in accurately identifying and capturing the majority of bricks.

4.2.2 Mean Average Precision (mAP)

Mean Average Precision (mAP) is another key metric for evaluating object detection models like YOLO. The mAP score represents the model's ability to both identify and accurately localize objects within images. This means it not only detects objects but also accurately pinpoints their positions. A high mAP score indicates that the model excels in precisely localizing objects within images. For instance, with a mAP50 score of 0.992 the YOLOv5 model demonstrates remarkable accuracy in identifying and locating BrickType A objects. This result is particularly significant in applications where the spatial position of objects is crucial, such as robotics and automated quality control in manufacturing. The mAP score offers a holistic assessment of the model's object detection capabilities, reflecting its overall accuracy in localizing and identifying objects in a given dataset.

4.2.3 Confidence Rate and Thresholds

In the context of object detection, confidence rates and thresholds play a pivotal role in determining the model's robustness and adaptability to different scenarios. The confidence rate represents the level of certainty or confidence that the model has in its predictions. Typically, YOLO models provide a confidence score for each detected object, indicating how sure the model is about the object's presence. Adjusting the confidence threshold allows you to control the trade-off between precision and recall. A higher confidence threshold means the model will make predictions with greater confidence, potentially reducing false positives but possibly missing some true positives. Conversely, a lower threshold might capture more instances but could also introduce more false positives.

4.2.4 YOLO Result for Bricks

The YOLOv5 model achieved high accuracy in detecting bricks, as demonstrated by the results based on the validation source in Table 2. The model was able to quickly detect bricks with accurate labeling and a percentage of accuracy, which suggests the potential for utilizing object detection in construction, particularly in robotic construction. This could allow for precise brick detection and could aid in the construction process by assisting in the selection of the appropriate bricks for a given function or notifying humans about specific bricks in the construction process.

Overall, the results of this investigation highlight the importance of utilizing advanced computer vision techniques and tools, such as RoboFlow and YOLOv5, to enhance the quality and usability of large-scale datasets in various domains.

Table 2: YOLOv5 Results for Bricks

Class	Images	Instances	P	R	mAP50	mAP50-95
BrickType A	636	226	0.982	0.954	0.992	0.852
BrickType B	636	167	0.994	0.97	0.993	0.938
BrickType C	636	87	0.981	0.943	0.988	0.898
BrickType D	636	94	0.926	0.926	0.96	0.873
BrickType E	636	103	0.953	0.883	0.946	0.851

In the context of our YOLOv5 results, the term "Instances" holds particular significance. It signifies the total count of individual objects or items within each distinct class, such as BrickType A, BrickType B, and so forth, which our model successfully identified and localized in the dataset of images. For instance, when considering "BrickType A," our model excelled at detecting and accurately localizing 226 instances of this specific brick type across a set of 636 images. Similarly, in the case of "BrickType B," the model exhibited proficiency in detecting and localizing 167 instances of this particular brick type within the same 636 images. These "Instances" metrics are of paramount importance as they shed light on the model's capability to not only recognize objects but also quantitatively assess their prevalence within the dataset. This quantitative insight into object detection performance is integral to the thorough evaluation of our model's effectiveness and is a critical component of our study's findings.

As shown in Table 2, the model has performed well in identifying and classifying objects of BrickType A to E. The model was evaluated on 636 images containing a total of 226 instances of BrickType A, and achieved a precision score (P) of 0.982, indicating a high degree of accuracy in identifying true positives. The recall score (R) of 0.954 indicates that the model has a high level of sensitivity in detecting all relevant instances of BrickType A. Additionally, the model achieved an mAP50 score of 0.992, which is a measure of how well the model localizes and identifies objects within images. This indicates that the model has a high level of accuracy in identifying and localizing instances of BrickType A.

The mAP50-95 score of 0.8520, which is a measure of the model's accuracy across a range of confidence thresholds, suggests that the model may struggle with more difficult instances of BrickType A or in situations with low contrast or poor lighting. Overall, the results indicate that the YOLOv5 model is well-suited for identifying and classifying objects of BrickType A to E with a high degree of accuracy, although some potential limitations may exist in certain scenarios.

Figure 19: Real-time object detection with YOLOv5

In Figure 19, the post-training process of our YOLO model is presented, showcasing the utilization of Google Colab (Figure 20) for model training. Subsequently, the YOLOv5 model is executed within the YOLO folder via a Linux environment, on Linux system. This process entails real-time object detection, which is based on the dataset we meticulously trained and discussed in the preceding results section.

Figure 19 illustrates the YOLO model's dynamic operation, wherein it promptly generates bounding boxes around detected objects, labels these objects with their respective categories, and provides a confidence rating, all of which are superimposed onto the images in real time. In the figure's left-hand side, there is a terminal window within Linux, continuously displaying data for each frame of the process. This terminal data can be harnessed for various purposes, including potential integration with a master system using a specific communication protocol. For instance, this data can be transmitted to a robotic arm that can then selectively retrieve bricks based on user-defined criteria, enhancing the adaptability and utility of the system.

Figure 20: Screenshot of YOLOv5 model in Google-Colab

4.2.4.1 Brick Detection and Its Implications

The YOLOv5 model has demonstrated remarkable success in the domain of brick detection, with high accuracy and precise labeling. The implications of these results are far-reaching, particularly in the context of construction, where precise object detection can significantly improve efficiency and safety. The model's ability to identify bricks suggests potential applications quickly and accurately in robotic construction. By facilitating the selection of appropriate bricks and notifying human workers about specific bricks in the construction process, YOLOv5 has the potential to revolutionize the industry. This technology could lead to more precise and error-free construction processes, ultimately saving time and resources.

4.2.4.2 Applicability to Various Brick Types

The effectiveness of the YOLOv5 model is not limited to a single type of brick. In fact, it excels in identifying and classifying multiple types of bricks, as exemplified by the evaluation on BrickType A to E. The model's precision score (P) of 0.982 underscores its capability to identify true positives with a high degree of accuracy. Its recall score (R) of 0.954 emphasizes its sensitivity in detecting all relevant instances of BrickType A. Furthermore, the model achieves an mAP50 score of 0.992, indicating its competence in localizing and identifying objects within images. This implies that the YOLOv5 model can not only identify bricks but also accurately pinpoint their location within the images.

4.2.4.3 Thresholds and Challenges

While the model excels in various aspects of brick detection, the mAP50-95 score of 0.8520 hints at potential challenges. This score represents the model's accuracy across a range of confidence thresholds, indicating that it may face difficulties with more complex instances of BrickType A or in situations with low contrast or poor lighting. It's essential to acknowledge these limitations, as they provide insights into the practical use of the model. Further refinement and adaptation of the model may be necessary for scenarios that involve challenging environmental conditions.

4.2.4.5 Overall Significance

In summary, the results showcase the YOLOv5 model's high accuracy in identifying and classifying bricks, making it a valuable tool for the construction industry and beyond. Its performance across various brick types highlights its versatility. However, the model's performance under different confidence thresholds suggests the need for careful consideration of its application in real-world scenarios. As the field of computer vision continues to evolve, the YOLOv5 model's capabilities and limitations provide valuable insights for future research and applications. Overall, this study underscores the significance of harnessing advanced computer vision techniques to enhance the quality and usability of large-scale datasets across diverse domains, opening up new possibilities for automation and precision in various industries.

4.3 Constructing Masonry Arches without Centering

Our research endeavors encompassed the development and validation of complex algorithms designed to create masonry corbel arches that could be constructed without the need for traditional centering support structures. These algorithms operated on the foundational principle of ensuring equilibrium conditions for each masonry unit, a concept exemplified in Figure 22. This innovation opened up a world of possibilities as a multitude of masonry vault designs became accessible through the algorithm, as demonstrated in the work of Neupan and Liu (2020).

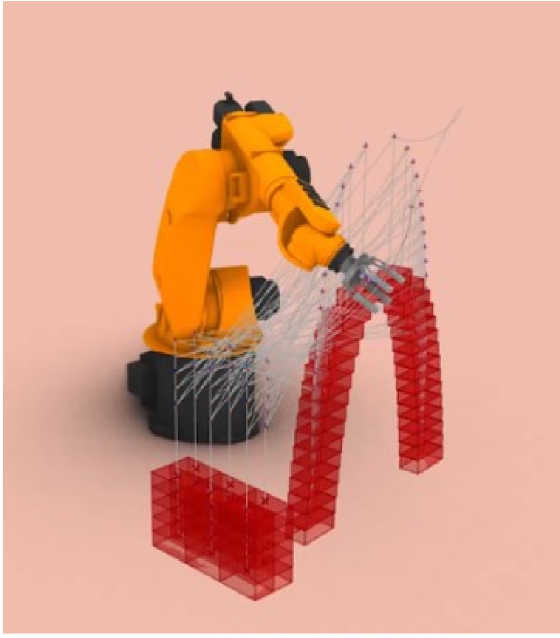


Figure 21: Kuka-Prc Model for Masonry Arch Construction

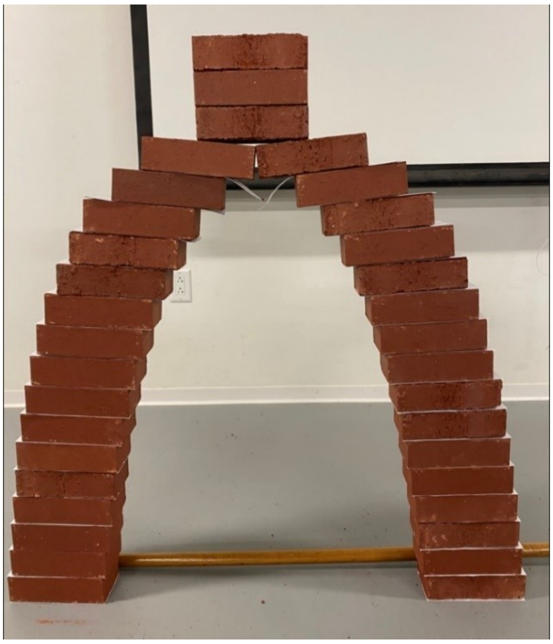


Figure 22: Constructing Masonry Arch based on the code's results.

The algorithms weren't solely confined to the realm of theoretical design; they were put into practice within the context of robotic construction using KUKA|PRC, as vividly depicted in Figure 21. However, the realities of real-world construction and the inherent uncertainties of hardware presented challenges. Automation in construction, when executed by robot arms, introduced potential pitfalls, as task executions occasionally failed, raising concerns of structural integrity and the safety of human workers. To address this critical issue, we recognized the necessity of early detection of near-miss failures before they escalated into critical incidents.

Here, our well-developed code, rigorously tested and refined over time, played a pivotal role. This code effectively harnessed data on brick dimensions, which served as a crucial foundation in our research. The code's accuracy and efficiency enabled designers to precisely determine the dimensions needed to create a resilient and structurally sound corbel arch. Furthermore, we leveraged this data in the context of the HRI process, where the robotic system collaborated closely with human workers. The code facilitated a nuanced understanding of human requirements and helped the robot adapt to these needs. In doing so, it transformed the HRI process into a dynamic and collaborative effort, where our robot responded seamlessly to human supervisors' real-time verbal guidance. This adaptability was a critical element of our approach, given the inherent limitations in the reasoning capacity of machinery. Ultimately, the project demonstrates not only the theoretical advancements in masonry corbel arch design but also the successful practical implementation in the domain of construction, emphasizing safety, precision, and the potential for transformative change in the construction industry.

4.4 Human-Oriented tests

In this project, we introduced a meticulously crafted approach where we input the exact dimensions of each brick were introduced into our code. The accuracy of the code and the subsequent physical representation set the stage for a practical implementation of the project. However, the real-world presented its own set of challenges. Not all bricks were uniform in their dimensions, and variations became evident during the physical assembly process. Some bricks had imperfectly straight edges, while others were composed of different materials, further complicating the task.

It was through these initial trials and errors that the project discovered an effective solution. By incorporating a 90% safety factor into the code, there is a way to mitigate the challenges presented by non-standard bricks. This safety factor was instrumental in handling the variability in brick dimensions and the inherent imperfections in the materials. As a result, the robustness of our approach became evident, as demonstrated in the images in Section 4.4.1, 'Human-Oriented Tests.' This section provides a visual representation of the challenges and the successful outcomes achieved through the adaptive coding approach serves as a testament to the adaptability of the code, allowing the project to handle real-world complexities and varying materials effectively, ultimately leading to the successful assembly of the project.

4.5 Robotic Operations

The robotic operation discussed in this section can be categorized into three distinct phases:

- Phase 1, preliminary tasks with the physical robot. The initial stage involves conducting practical tasks with the physical robot, such as object manipulation (e.g., picking and placing). Notably, the DofBot is utilized in the Co-Worker Project, enabling early engagement with the robotic system.

- Phase 2, data set definition for object detection (YOLOv5). As discussed earlier, a crucial aspect of robotic operation involves defining the dataset required for object detection. In this case, the YOLOv5 algorithm is employed to facilitate this process.
- Phase 3, ArchiTech project with simulation and integration of YOLOv5. The operation then transitions into a simulation environment, where efforts are made to integrate the YOLOv5 object detection framework with the robot arm.

4.5.1 Preliminary Tasks with the Physical Robot

4.5.1.1 Preliminary test with robot arm (DofBot)

In this pivotal phase of the research, comprehensive warm-up tests with the DofBot robotic arm were completed. These tests involved the manipulation of various objects, with precise source and target locations predefined. The DofBot executed code sequences to autonomously pick up objects and place them at designated target locations as shown in Figure 23. This stage was fundamental in priming the robot for more complex tasks and provided insights into its operational capabilities.

As depicted in Figure 22, a visual representation highlights the key elements of the operation. In this configuration, a designated target location (T) serves as the focal point for the DofBot robotic arm's activities. Further denoted as A, B, C, and D, these predefined locations represent the precise positions to which the robot must navigate, retrieve objects, and sequentially stack them atop one another at the designated target point. It's worth noting that Target places X and Y are initially identical but will increase as the operation progresses, primarily contingent on the height of the objects being handled.

This operation is structured into a systematic sequence of four steps, and after completing each step, the robot is required to return to its home base. Complementing this, Figures 23, 24, and

25 present a visual record of the DofBot in action, providing snapshots that illustrate its dynamic movements as it orchestrates the manipulation and transportation of objects. These visual aids serve as valuable documentation of the operational process in motion.

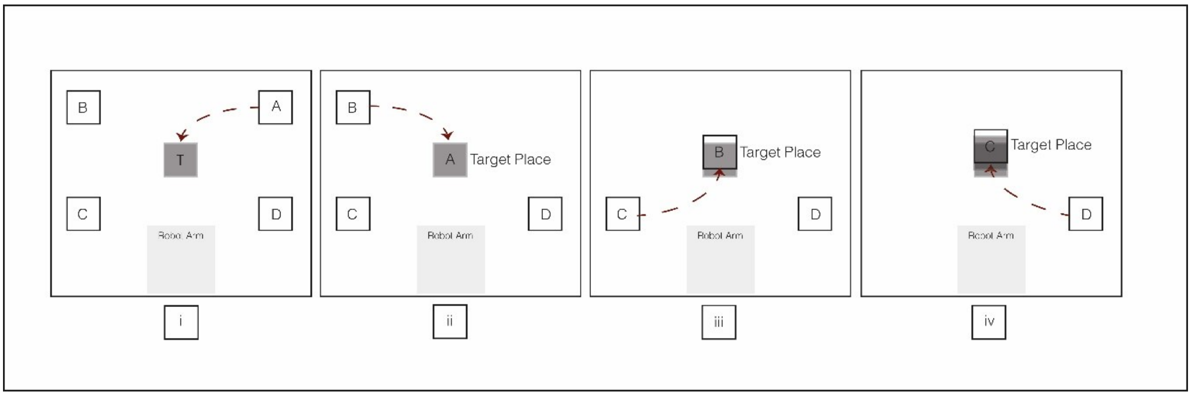


Figure 23: Phase 1, Pick up and Place operation with DofBot

Figure 24: Pick up and Place operation with DofBot

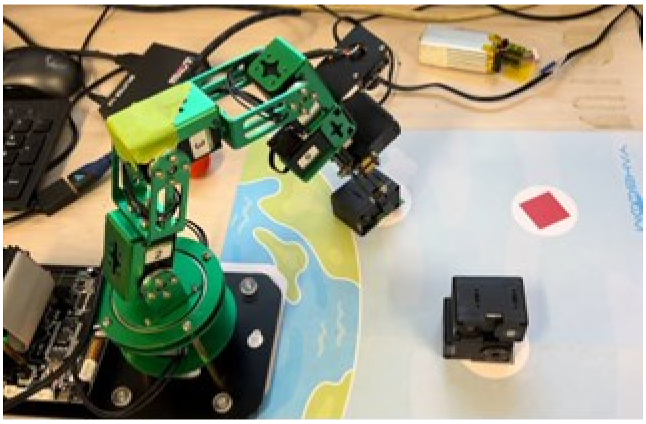


Figure 25: Phase I Operation with DofBot

4.5.2 ArchiTech project

4.5.2.1 System Overview

In the field of modern robotics, the ArchiTech system stands as a significant leap forward. Built on the renowned ROS framework, this system primarily employs Python, a highly regarded and adaptable programming language within the realm of computing. The architecture of the ArchiTech platform revolves around three distinct modules including (1) human interface: this module offers a refined interface that optimizes human-computer interaction, ensuring that operators can effectively communicate commands and decisions to the system; (2) object detection with cutting-edge algorithms: this module supports precise object recognition, allowing the system to discern even the smallest details with exceptional accuracy; (3) robot controller: Serving as the central hub of the system, the robot controller module is responsible for coordinating a range of robotic activities and ensuring the seamless execution of tasks (Figure 26).

4.5.2.2 Software Framework

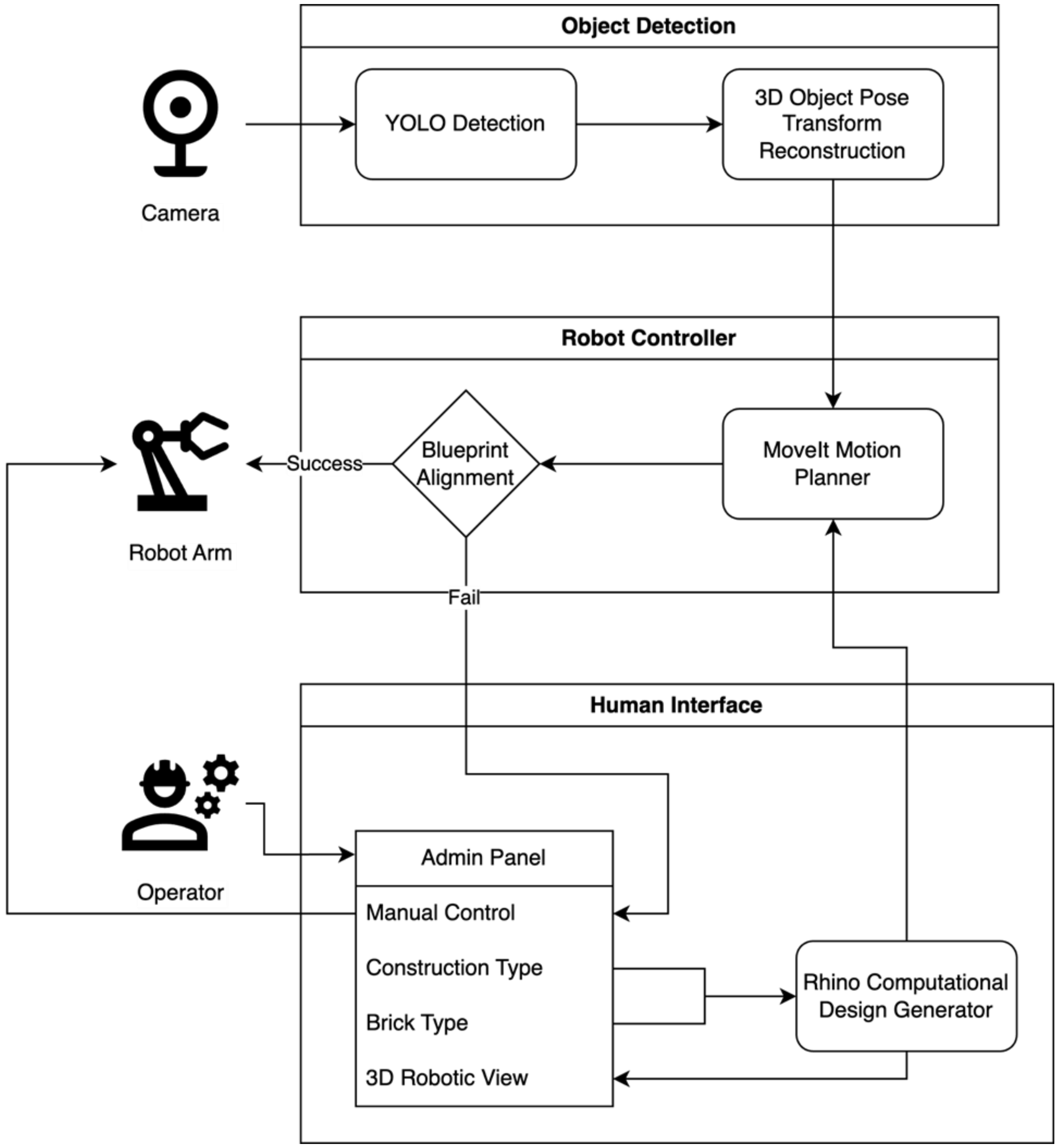


Figure 26: Detailed Workflow of ArchiTech Project

In the contemporary landscape of robotics research, this project is grounded on ROS—a renowned open-source robotics middleware suite. Contrary to what its nomenclature might suggest, ROS transcends the conventional definition of an operating system (OS). Instead, it encapsulates an extensive collection of software frameworks tailored explicitly for advanced robotic software development. Among its myriad features, ROS facilitates services catering to a heterogeneous computer cluster. This encompasses a gamut of functionalities, : from hardware abstraction and low-level device control to the seamless implementation of ubiquitously employed operations. Furthermore, ROS excels in enabling efficient message-passing between disparate processes and streamlining package management.

Central to the developmental architecture is the Python programming language. Renowned for its interpretative nature, Python is an object-oriented, high-caliber language enriched with dynamic semantics. Its intrinsic high-level data constructs, when amalgamated with its dynamic typing and binding capabilities, position Python as a quintessential tool for rapid application development. Moreover, Python's versatile nature makes it an ideal candidate for scripting purposes, serving as an adhesive to cohesively integrate pre-existing components, thereby enhancing the overall efficiency and interoperability of our system.

4.5.2.3 Human Interface

Within the purview of this research, the front-end interface bifurcates into two distinct segments, including the admin panel GUI and the RCDG. The admin panel's foundation lies in the QT framework—a robust, free, and open-source platform renowned for crafting intricate graphical user interfaces. This framework not only supports the development of cross-platform GUIs but also ensures compatibility across a diverse array of software and hardware ecosystems, as delineated in Figure 27.

Serving as the nexus for control and real-time feedback, the admin panel empowers operators with comprehensive oversight. Through this interface, operators can seamlessly embark on their architectural endeavors, selecting desired designs and specifying the brick type. Once these preferences are delineated, the GUI efficiently communicates this data to the RCDG. Subsequent to this, the RCDG harnesses its computational prowess to render a detailed 3D model of the envisioned structure, in this case, a bridge. This dynamic model allows users to iteratively adjust parameters and brick types, facilitating a meticulous fine-tuning process that hones the final architectural rendition.

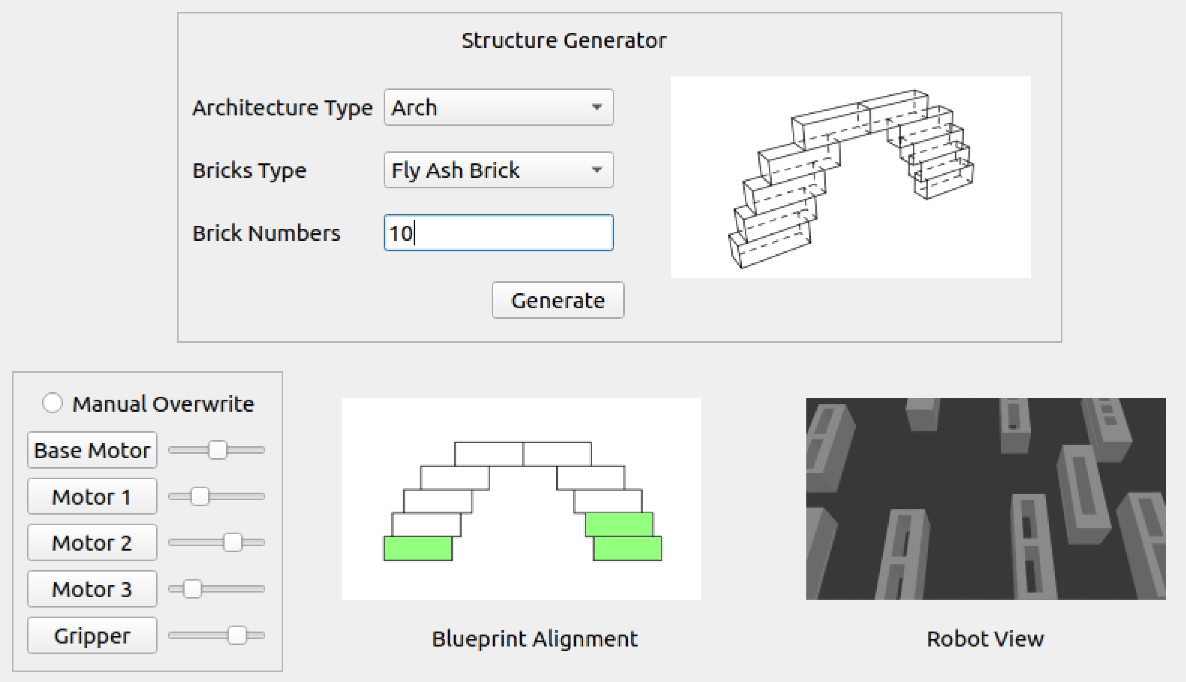


Figure 27: ArchiTech GUI design with QT

Upon achieving a satisfactory design, the RCDG meticulously catalogs the three-dimensional coordinates of each brick component into a text file. Concurrently, it captures a visual representation of the final architecture. This dual-output mechanism not only aids in the subsequent construction phase but also serves as a pivotal safety measure, ensuring that all architectural nuances are adhered to with the utmost precision.

4.5.2.4 Vision Postprocessing System

In the dynamic environment of an active construction site, an eclectic assortment of bricks and materials are scattered, reflecting the complexities inherent to the field. Given this context, the precision and efficiency of object detection become pivotal to accurately discern the spatial positioning and type of each distinct item. This innovative system, meticulously designed for real-world applications, springs into action once the user delineates the construction parameters.

Upon activation, the robot commences by initializing its integrated camera, systematically identifying, and cataloging the requisite bricks essential for the specified construction. At the heart of this process is the advanced YOLO algorithm. Distinctly designed, YOLO employs an end-to-end neural network that concurrently predicts bounding boxes and class probabilities. This innovative approach marks a departure from traditional object detection algorithms, which predominantly repurposed classifiers for detection tasks.

With YOLO's deployment, this system undertakes an initial assessment, determining if the present bricks on the platform suffice for the planned construction. In the event of a deficit, the robot apprises the user of the missing bricks, presenting two pragmatic courses of action, await restock and overwrite.

This option of await restock places the robot's operations in a temporary hiatus, resuming only post the user's replenishment of the requisite bricks.

Opting for overwrite: Opting for this, the robot persists with the available resources, halting operations once it depletes its brick supply. Subsequently, the onus shifts to the operator to either introduce a different brick type or momentarily halt the operation for restocking.

Transitioning to the execution phase, the system incorporates the BPMS to ascertain the precision of each brick's placement. BPMS functions by assimilating a 2D visual representation of the architectural design—originally generated by the RCDG—into its memory. It then embarks on edge detection, meticulously scrutinizing every brick's alignment. Should any misalignment exceed a tolerance of 2mm, the BPMS promptly alerts the operator, highlighting potential discrepancies and beckoning their intervention for judicious corrective action.

4.5.2.5 Robot Controller

Central to the functionality of the ArchiTech system is the sophisticated Robot Controller module (Figure 27, 28 and 29). This module adeptly synthesizes data received from both the Human Interface and the Vision Postprocessing System, channeling this amalgamated data into the motion planning system. The culmination is a systematic execution blueprint tailored to actualize the envisioned architecture.

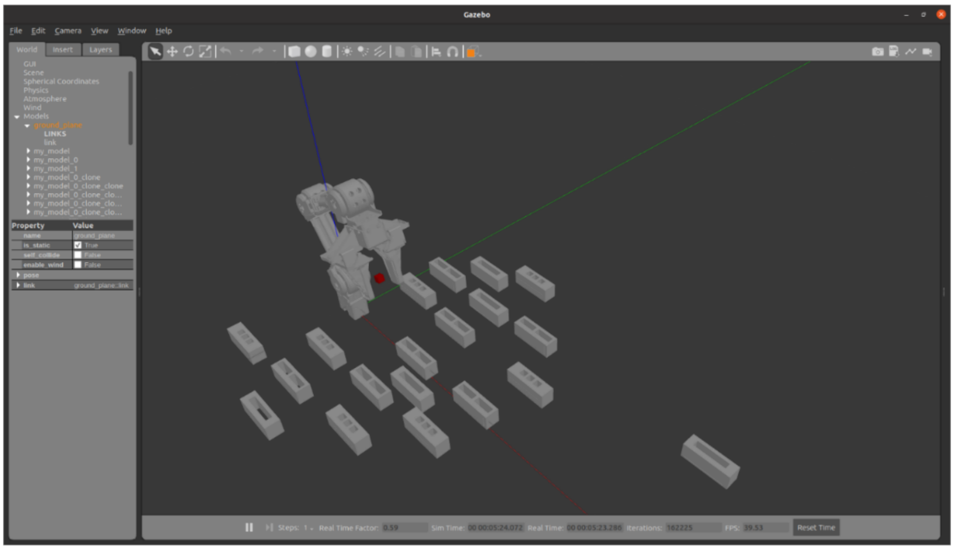


Figure 28: ArchiTech Project, Object detection with YOLOv5 for choosing the accurate brick

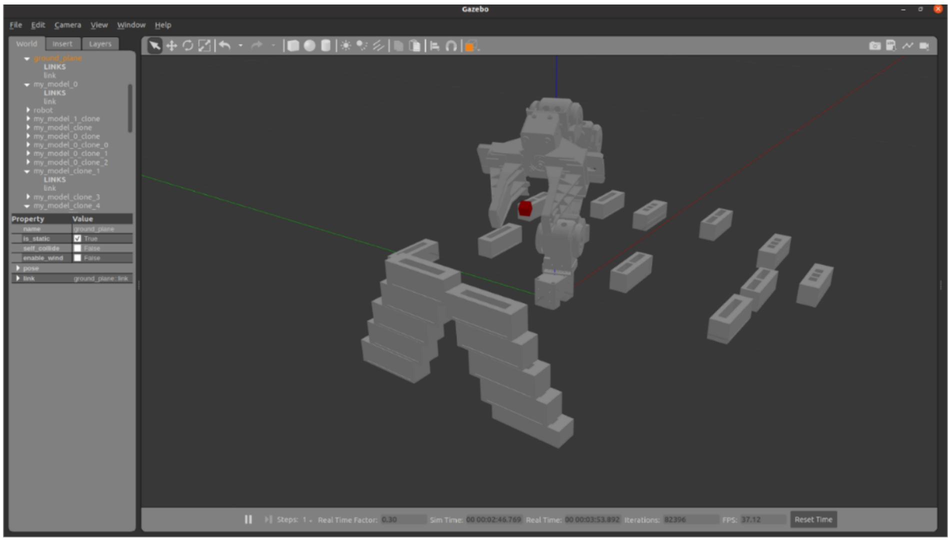


Figure 29: ArchiTech Project, pick up and place based on RCDG code

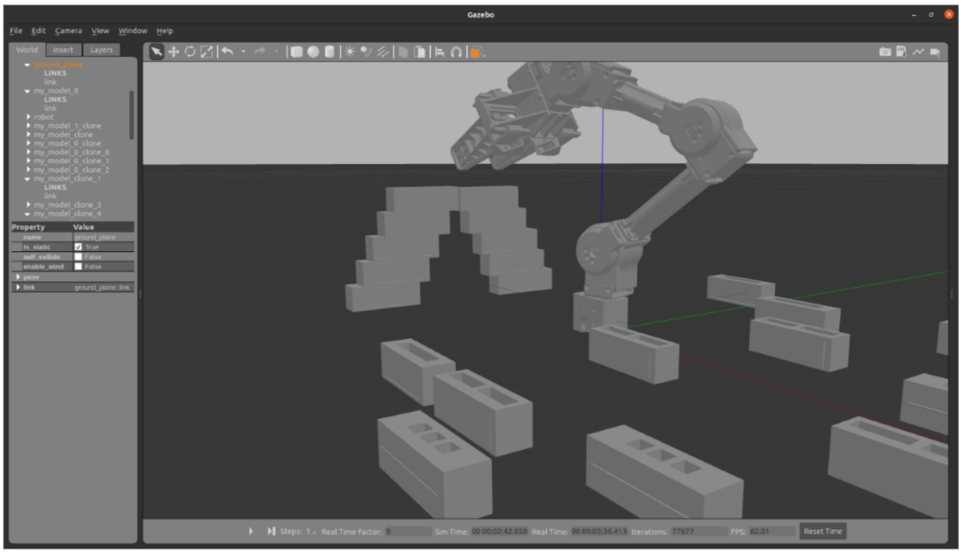


Figure 30: Phase III, ArchiTech Project

A pivotal component driving this process is the "MoveIt" framework—widely regarded as the quintessential robotic manipulation platform for ROS. Encompassing cutting-edge advances across diverse domains such as motion planning, manipulation, 3D perception, kinematics, control, and navigation, MoveIt embodies the future of robotic motion. Moreover, underpinning the planner within this framework is the OMPL. The synergy between MoveIt and OMPL is undeniable; MoveIt seamlessly integrates with OMPL, harnessing the motion planning algorithms intrinsic to OMPL as its foundational, go-to planners.

The motion planning workflow is initiated once the planner imbibes the granular architecture transformation data, along with the specifications of the source bricks. Priming the library for the planning phase is the preliminary step, ensuring readiness for seamless execution. Should the internal planner navigate this preparatory phase without hiccups, it propels the crafted planning data to the robot, dictating a sequence of actions. Execution fidelity is paramount; hence, any anomalies detected via the BPMS during the plan's unfolding instantly halt the robot's movements, awaiting further directives from the operator. Subsequent rectifications enable the planner to recalibrate, invoking a fresh planning phase attuned to the current conditions. This cycle—plan, execute, evaluate, recalibrate—ensures precision and adaptability.

In the realm of contingency planning, we've integrated a robust manual control mechanism. This feature, accessible via a GUI interface, grants operators the autonomy to manually manipulate the robot's arm, circumventing potential operational impediments that might impede automatic operations. This dual-layered approach—automated precision complemented by manual override—ensures uninterrupted progress, even in the face of unforeseen challenges.

4.6 Summary

Chapter 4 provides a comprehensive exploration of the outcomes and implications of employing the YOLOv5 model in the field of brick detection and Human-Robot Interaction (HRI). This chapter highlights the pivotal role of advanced technology, specifically the fusion of object detection and human-robot Interaction, in reshaping the landscape of automation and collaboration between machines and humans. The key points covered in this chapter are as follows:

- YOLOv5 results for brick detection: The YOLOv5 model exhibited remarkable accuracy in detecting various brick types. The model's precision, recall, and mAP scores were extensively discussed. These scores demonstrated the model's proficiency in accurately identifying and classifying different types of bricks. This high accuracy not only holds theoretical significance but also promises practical applications in the realm of construction. The YOLOv5 model's precision empowers robotic systems to pick and place bricks with exceptional accuracy, turning vision into reality.
- HRI: traditional robotic operations often follow predefined instructions, but this research introduces a new dimension of HRI, where robots actively engage with human operators. The robotic system initiates real-time dialogues with human operators through intuitive interfaces, asking crucial questions and seeking input during operations. This interactive paradigm enhances the adaptability of the robotic system, making it responsive to human operators' dynamic needs and the uncertainties of real-world construction. This collaborative environment fosters a harmonious synergy between technology and human expertise.
- Human Interface and Robotic Control: The chapter also delves into the components of the ArchiTech system, showcasing how the system integrates human interaction and object detection for the effective execution of architectural projects. The ArchiTech system incorporates a Human Interface, Object Detection, and Robot Controller modules, all built on the ROS. It uses Python for its programming language. The Human Interface allows operators to communicate with the system and select architectural designs, while Object Detection employs advanced algorithms for precise object recognition. The Robot Controller module synthesizes data from both the Human Interface and the Vision Postprocessing System: using the MoveIt framework for motion planning.
- Vision Postprocessing System: The Vision Postprocessing System is responsible for accurately detecting and cataloging the bricks required for the construction process. It utilizes the YOLO algorithm to assess the availability of bricks and, if necessary, prompt

human operators to restock. Additionally, the BPMS ensures the precision of brick placement, highlighting discrepancies for operator intervention.

This chapter emphasizes the promising results of YOLOv5 in brick detection, and the innovative integration of Human-Robot Interaction in the construction industry. It sets the stage for the transformative power of robotics and human collaboration in the world of construction and beyond.

Chapter 5: Conclusion and Recommendation

Leveraging the ArchiTech system presents a comprehensive, end-to-end solution that empowers operators to actualize their envisioned architectural constructs with precision and efficiency. However, it's imperative to highlight certain operational nuances that frame our current capacity. Owing to resource constraints, the experiments in this study predominantly employ a scaled-down robot arm and utilize rudimentary bricks to construct an arch model. This is representative of the initial phase, geared toward establishing foundational capabilities.

Furthermore, as this investigation expands the scale and complexity, it becomes increasingly evident that our robotic arm needs to evolve beyond its current capabilities. Introducing multifunctional tools is a logical progression. For instance, integrating a mechanism that automatically applies cement post brick-placement would be a significant stride towards emulating real-world building processes. This not only enhances the authenticity of the construction simulations but also paves the way for a more holistic and realistic representation of architectural construction in automated settings.

5.1 Beneficial aspects of robotic fabrication with human-oriented task

Robotic fabrication is a transformative technology that holds the potential to revolutionize numerous industries by enhancing efficiency, precision, and flexibility in manufacturing processes. When robots are integrated into tasks with a focus on human-centric goals, the resulting synergy leads to several significant benefits. These human-oriented robotic applications are not just about automating tasks but also about creating environments and systems that prioritize the well-being and safety of individuals. In this discussion, we will delve into the beneficial aspects of robotic fabrication in the context of human-oriented tasks.

- Enhanced precision and quality: One of the most notable advantages of robotic fabrication in human-oriented tasks is the level of precision and consistency it brings to the processes. Robots can perform repetitive tasks with an unparalleled level of accuracy, ensuring that the end products meet high-quality standards. In fields such as medicine, where precision is critical, robots assist in surgeries, reducing the margin of error. In manufacturing, the precision offered by robots leads to improved product quality, ultimately benefiting consumers.

- Improved safety and ergonomics: Robots can take on tasks that are hazardous or physically demanding for humans. This not only improves worker safety but also enhances ergonomics in the workplace. Robots can handle tasks in environments with extreme temperatures, toxic substances, or high-risk situations without the need for humans to expose themselves to danger. In industries like construction, robotic systems are used to lift heavy objects or work in challenging terrains, reducing the risk of accidents and injuries.
- Increased efficiency and productivity: Robotic fabrication excels in enhancing efficiency and productivity. Robots work around the clock without fatigue, leading to increased output and reduced production time. In warehouses and logistics, robots can swiftly sort and deliver goods, reducing lead times and ensuring timely deliveries. This increased efficiency results in cost savings and improved competitiveness in industries.
- Adaptability and customization: Robotic systems can be programmed and configured to adapt to various tasks and production needs. In human-oriented tasks, this adaptability allows for customization to suit individual requirements. For example, in the healthcare sector, robots can be tailored to provide specific care for patients with different medical conditions. In the food industry, robotic chefs can adjust recipes and serving sizes to meet dietary preferences and restrictions. This adaptability caters to the uniqueness of each human-centered task.
- Augmentation of human capabilities: Robotic fabrication can complement human abilities, rather than replacing them. In healthcare, robots can assist medical professionals in diagnosing and treating patients. In education, robots can enhance the learning experience by providing personalized guidance and support.

These collaborations between humans and robots result in the augmentation of human capabilities, improving the overall quality and efficiency of various tasks.

The integration of robotic fabrication into human-oriented tasks offers a multitude of benefits, ranging from precision and quality improvement to increased safety and productivity. As technology continues to advance, the synergy between humans and robots promises to redefine the way we approach tasks, ultimately leading to safer, more efficient, and more customized solutions in a wide range of industries. The key to realizing these benefits lies in thoughtful integration and collaboration between humans and machines.

5.2 Discussion

In conclusion, this thesis addresses critical questions which listed below, at the interfaces of architecture and construction, HRI and object detection.

- In the context of current robotics practices, where instructions flow solely from humans to robots without reciprocal communication, can robots be developed to visually recognize and accurately select construction materials like bricks during construction tasks?
- Is it feasible to equip robots with advanced object detection capabilities to autonomously detect variations in construction materials, such as bricks and promptly alert human operators to any deviations or challenges encountered during construction processes?
- Can robots, armed with state-of-the-art vision systems, actively participate in intricate endeavors, such as building arches, while simultaneously perceiving and responding to changes in the construction environment, ensuring precision and adaptability throughout the construction undertaking?

The research questions posed in the introduction underscore the significance of shifting from one-way communication in robotic projects to a paradigm where robots actively collaborate with human operators. The study focuses on equipping robots with state-of-the-art vision systems, notably YOLOv5, to recognize and select construction materials like bricks. The introduction of the ArchiTech system, based on the ROS and Python, exemplifies a groundbreaking approach to make the robot arm smarter than before, fostering seamless communication and cooperation between humans and robots. The system's vision postprocessing system, integrating YOLO, ensures precision and adaptability in construction, providing feedback to human operators in real-time. The use of the "MoveIt" package and the OMPL further enhances the system's precision and adaptability. The thesis introduces a dual-layered approach with a manual control mechanism for contingency planning, offering a robust solution for unforeseen operational challenges.

This research significantly contributes to the evolution of robotic projects in design and construction by not only enhancing the technical capabilities of robots but also redefining the way humans and robots collaborate in complex architectural endeavors. The ArchiTech system's innovative features provide an illustrative example of how technology can facilitate a more dynamic and interactive approach to architectural construction, resulting in higher precision and adaptability. Furthermore, applying YOLO, as an object detection approach, in the system highlights the importance of harnessing cutting-edge technology to address the challenges of one-way communication between humans and robots.

Moreover, this thesis comprehensively explores robotics, object detection, and human-robot interaction, with the ArchiTech system at its core. The research elucidates the potential for robots to become active collaborators in building projects, revolutionizing the field and opening new avenues for precision, adaptability, and efficient communication in construction processes. This work contributes to the broader field of robotics and human-robot interaction and paves the way for more advanced and interactive robotics applications in architecture, engineering and construction.

5.3 Limitations and challenges

While the preceding chapters have illuminated the innovative and promising aspects of the object detection and HRI in design and construction via ArchiTech project, it is essential to acknowledge its inherent limitations. This section provides a meticulous exploration of the constraints and challenges that have been encountered during this research. By understanding these limitations, a more comprehensive perspective on the scope and applicability of the ArchiTech project can be gained.

- Hardware Limitations: The ArchiTech project's hardware components, primarily the DofBot robotic platform, are subject to certain limitations. The DofBot, while chosen for its accessibility and affordability, has constraints on its load-carrying capacity and range of motion. These limitations affect the size and weight of bricks that can be handled, potentially restricting the practicality of the system in large-scale construction projects. Furthermore, hardware failures or inaccuracies in sensors can impede the precision and reliability of the construction process.
- The object detection aspect of the ArchiTech project relies heavily on the YOLO algorithm. While YOLO is renowned for real-time object detection, it may encounter difficulties in complex lighting conditions, occlusions, or when objects exhibit subtle variations in appearance. Additionally, the accuracy of the object detection model is contingent on the quality and diversity of the dataset. Insufficient or biased training data could lead to false positives or negatives, affecting the overall performance.
- The GUI designed for human interaction with the ArchiTech system may pose usability challenges. Users with limited technical expertise in architecture or robotics may find the interface complex, potentially hampering its adoption. Enhancing the user-friendliness of the GUI while maintaining its functionality is an ongoing challenge.
- Scalability and cost implications: scaling up the ArchiTech project for larger construction projects would involve increasing the number of robots and sensors. This expansion may result in higher implementation costs and complex coordination challenges. Achieving an optimal balance between scalability, cost-effectiveness, and operational efficiency is a concern for broader adoption.
- Ethical and safety Considerations: The integration of robots in construction raises ethical and safety concerns. Ensuring the safety of workers and the public when robots are

deployed on construction sites is of paramount importance. Ethical considerations include the potential displacement of human labor and the ethical use of data for object detection.

In conclusion, the ArchiTech project, while showcasing significant potential for innovation in the construction industry, is accompanied by a set of limitations. Recognizing and addressing these limitations is vital for the project's evolution and real-world applicability. As technology and research progress, these limitations may be mitigated, paving the way for more extensive and impactful deployments of the ArchiTech system.

5.4 Recommendations and Future Work

Semantic Understanding in Natural Language Processing (NLP): applying the natural language processing capabilities of the robotic system is essential. The development of a profound understanding of construction-specific terminology and context is imperative to facilitate more effective and precise communication between the robot and human counterparts. This advancement will elevate the robot's ability to comprehend and respond to construction-related instructions and queries.

Integration of machine learning for informed decision making: to elevate the autonomy of the robot, the incorporation of a decision-making framework into its AI system is recommended. Machine learning algorithms can be integrated to empower the robot to analyze situational factors, consider constraints, and make informed decisions regarding brick placement. This step will contribute to the efficiency and adaptability of the robot in real-time construction scenarios.

Establishment of a human feedback loop: the implementation of a feedback mechanism that enables construction workers to correct and guide the robot during HRI tasks is advised. This feedback loop will not only serve as a valuable learning tool for the robot but also ensure its adaptability to site-specific requirements and the preferences of the human workforce.

Real-world deployment and testing: transitioning from simulation environments to real-world construction sites is an imperative next step. The practicality, efficiency, and safety of the HRI system must be rigorously evaluated in live construction environments. This real-world deployment will provide insights into the system's readiness for large-scale adoption within the construction industry.

User experience assessment: conducting user experience assessments, involving feedback from construction workers and other stakeholders, is a critical undertaking. These assessments will gauge the ease of interaction with the robot, the overall efficiency of the HRI system in facilitating construction processes, and its tangible impact on the workflow and project outcomes.

Scalability and cost analysis: to ensure the viability of the HRI system on a broader scale, a thorough analysis of its scalability is warranted. This analysis should encompass the assessment of costs, workforce implications, and the potential return on investment for construction companies, providing a comprehensive understanding of the system's feasibility and economic advantages.

These recommendations collectively outline a strategic roadmap for advancing HRI in the construction sector, marking a significant step towards more efficient, adaptive, and collaborative construction processes.

References

Beach, E. (1971). Automation in Perspective: Applied Economics, 3:2, 147-152, DOI: 10.1080/00036847100000036Canning, C., Donahue, TJ., & Scheutz, M. (2014). Investigating human perceptions of robot capabilities in remote human–robot team tasks based on first-person robot video feeds: IEEE International Conference on Intelligent Robots and Systems.

Rabinbach, A. 1990. "The Human Motor: Energy, Fatigue and the Origins of Modernity." Berkeley, CA.

Retsin, G., & García, M. J. (2016). Discrete computational methods for robotc additive manufacturing. ACADIA 2016: Posthuman Frontiers: Data, Designers, and Cognitive Machines - Proceedings of the 36th Annual Conference of the Association for Computer Aided Design in Architecture, 332–341

Schwab, Klaus. 2016. "The Future of Jobs Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution."

Vasey, L., Grossman, T., Nagy, D., Kerrick, H., Atherton, E., Thomasson, D., … Menges Achim. (2016). Collaborative Construction: Human and Robot Collaboration Enabling the Fabrication and Assembly of a Filament-Wound Structure (pp. 184–195).

Prado, M., Dorstelmann, M., Schwinn, T., Menges, A., & Knippers, J. (2014). Core-Less Filament Winding. In W. McGee & M. Ponce De Leon (Eds.), Robotic Fabrication in Architecture, Art and Design 2014. Springer. https://doi.org/10.1007/978-3-319-04663-1

Parascho, S., Gandia, A., Mirjan, A., Gramazio, F., & Kohler, M. (2017). Cooperative Fabrication of Spatial Metal Structures. In A. Menges, B. Sheil, R. Glynn, & M. Skavara (Eds.), Fabricate 2017 (pp. 24–29). UCL Press.

Liu, R., Koliyaee, V., Liu, R., Lucak, B. (2023). Changing Roles of Human in the Human-Robot Interactions in Design and Construction. In Architectural research center consortium.

Kadir, Gözde, Bülent (2019) “Industry 5.0 and Human-Robot Co-working” In: 3rd World Conference on Technology, Innovation and Entrepreneurship (WOCTINE)

Wingo, W. (1963). Automation’s Impact: The Science Newsletter, 84(18), 278.

Liu, R., Koliyaee, V., Lucak, F., Guo, Y., & Liu, R. (2023). “Changing Roles of Human in the Human-Robot Interactions in Design and Construction.” ARCC 2023 International Conference in Dallas: The Research-Design Interface.

G. & C. Merriam Co. (1967). Webster's Seventh New Collegiate Dictionary: G. & C. Merriam Company.

Gramazio, F., & Kohler, M. (2008). Digital materiality in architecture. Lars Müller Publishers.

Johns, O., Kilian, M., & Foley, D. (2014). Robotic construction methods: Computational design, materialisation, and assembly. In Robotic Fabrication in Architecture, Art and Design (pp. 40-57). Springer.

Johns, R.L., Kilian, A., & Foley, N. (2014). Design Approaches Through Augmented Materiality and Embodied Computation.

Huang, J. (2016). Robotics in architecture: From concept to production. Architectural Science Review, 59(1), 20-29.

Cote, P., Sherman, C., & Sweet, T. (2013). Robots in architecture: Advances in architectural geometry. In ACADIA 2013 Adaptive Architecture Conference Proceedings (pp. 79-82).

Haddadin, S., Albu-Schäffer, A., De Luca, A., & Hirzinger, G. (2016). The role of robotics in developing the future of sustainable and smart construction. Robotics and Computer-Integrated Manufacturing, 37, 13-23.

Tenneti, R., et al. (2016). Towards Construction Robotics: An Agile, Mobile Robot for On-Site Construction. Automation in Construction, 69, 31-41.

Chen, Y., et al. (2017). On-Site Safety Assurance of Construction Robots in Human-Robot Interaction Scenarios. Automation in Construction, 80, 1-11.

Parascho, S., Han, I.X., Beghini, A., Miki, M., Walker, S., & Bruun, E.P. (2021). LightVault: A Design and Robotic Fabrication Method for Complex Masonry Structures.

Adel, A., Augustynowicz, E., Wehrle, T. (2021). Robotic Timber Construction: Case study structure. ACADIA 2021, 530-536.

Bonwetsch, T., Kobel, D., Gramazio, F., & Kohler,8 M. (2017). The Informed Wall applying additive digital fabrication techniques on architecture. ACADIA 2017: Synthetic Landscapes Digital Exchange, 491-494.

Piskorec, L., Jenny, D., Parascho, S., Mayer, H., Gramazio, F. and Kohler, M. (2018). The Brick Labyrinth. In: J. Willmann, P. Block, M. Hutter, K. Byrne and T. Schork, ed., Robotic Fabrication in Architecture, Art and Design 2018, 1st ed. Zurich: Springer International, pp.490-500.

Huang, A. S. (2016). Human-robot interaction: A survey. ACM Computing Surveys, 49(4), 1-43. doi:10.1145/2996794.

Yanco, H. A., & Drury, J. L. (2002). Classifying human-robot interaction: An updated taxonomy. In Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2002) (pp. 153-158). IEEE.

Feli-Siefer, E., & Mataric, M. J. (2010). Efficient human-robot teaming through adjustable autonomy. In Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010) (pp. 101-108).

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., & Goodrich, M. (2006). Common metrics for human-robot interaction. In Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI 2006) (pp. 33-40).

Richert, J., Lauckner, M., Frommberger, L., & Hägele, M. (2018). Enhancing user acceptance of service robots: A survey. In Proceedings of the 27th IEEE International Symposium on Robot and Human.

Luptetti. (2017). ML. Robots, aesthetics, and heritage contexts: Association for Computing Machinery. Interactions 24, 5 (September - October 2017), 6–7. https://doi.org/10.1145/3125389.

Stoeva, D., and M. Gelautz. 2000. "Body Language in Affective Human–Robot Inter-Action." ACM/IEEE International Conference on Human-Robot Interaction; Piscataway. , 606–8.

S. M. LaValle, "Planning Algorithms," Cambridge University Press, 2006.

Smith, John. "The Impact of Industrialization on Workers' Livelihoods." Journal of Industrial History, vol. 25, no. 2, 2007, pp. 123-140.

Hannah Ritchie and Max Roser (2018) - "Urbanization". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/urbanization'

Websites

MoveIt Documentation http://moveit.ros.org/

OMPL: The Open Motion Planning Library, https://ompl.kavrakilab.org/

Urbanization: https://ourworldindata.org/urbanization

Appendix

- **YOLO Mathematic Concepts and Methods**
- Interpreting YOLO's Output Predictions

Understanding YOLO's output predictions involves grasping how it interprets bounding box coordinates and class probabilities. It's important to note that YOLO calculates the x and y coordinates of a bounding box relative to the top-left corner of the grid cell in which it resides, not the image's top-left corner. Each grid cell predicts coordinates as offsets relative to its own position.

To illustrate, consider dividing an image into a 3x3 grid. If an object's center falls within the central grid cell and we assume each grid cell has dimensions represented by 'A,' the object's center's coordinates within the cell are (0.6A, 0.6A) relative to the cell's top-left corner. YOLO predicts these coordinates as values between 0 and 1, representing fractions of 'A.' Consequently, (0.6, 0.6) signifies 60% of 'A' to the right and 60% down from the cell's top-left corner. These coordinates can be converted relative to the entire image since we know which cell predicts the box and its relative coordinates.

Hence, YOLO predicts the width and height as 1/3 and 2/3 of the image's width (W) and height (H), respectively, resulting in width and height predictions of $\sqrt{(0.33 * W)}$ and $\sqrt{(0.66 * H)}$. Additionally, YOLO assigns a probability between 0 and 1 to indicate the presence of an object. This object probability, when multiplied by the Intersection over Union (IoU) of the predicted box with the ground truth, yields the confidence score. IoU measures the overlap between the predicted and ground truth boxes, ranging from 0 (no overlap) to 1 (complete overlap).

A confidence score of 1 represents complete confidence, while 0 indicates zero confidence. A higher confidence score suggests greater certainty of object presence. This confidence score is then multiplied by the conditional class probability to calculate the probability score for a specific class's presence.

In summary, the predicted bounding box parameters typically take the form (0.6, 0.6, $\sqrt{0.33}$, $\sqrt{0.66}$, 1), representing the x-coordinate, y-coordinate, width, height, and object probability, respectively. YOLOv1, introduced in 2015, employs a convolutional architecture with 24 convolutional layers, 4 max-pooling layers, and two fully connected layers. It takes input color images of size 448x448 for object detection and predicts a cuboidal output with dimensions (7x7x30) for tasks like PASCAL VOC. The network's feature extractor utilizes convolution layers with varying filter sizes, employing 1x1 convolutions to reduce computational complexity and introduce non-linearity. While inspired by GoogLeNet, YOLOv1 doesn't use inception blocks but leverages 1x1 convolutions to manage channel depth. These 1x1 convolutions are applied to reduce parameter explosion in subsequent layers, enhancing model efficiency. The YOLOv1 model undergoes pretraining on the first twenty convolution layers using the ImageNet dataset at a resolution of 224x224, with the subsequent layers adapted for object detection at a resolution of 448x448. Transfer learning plays a crucial role in improving YOLO's performance by leveraging patterns learned from ImageNet. Dropout is applied to prevent overfitting, and batch normalization is not used in the model. During training, YOLO employs a sum-squared loss function, balancing localization and classification errors. The loss includes parameters λcoord and λnoobj to mitigate issues related to cells without objects. This approach ensures that the model converges effectively while handling various detection scenarios.

- **The Loss Function in YOLO**

YOLO's loss function may seem intimidating but it's quite simple. The loss used is called sum-squared loss and is used for all the tasks in YOLOv1. The authors quote that it is easier to optimize sum-squared than, say, log-likelihood. But they have also mentioned that it has some drawbacks.

The sum-squared error weighs localization and classification errors equally. And, since many grid cells in an image don't contain any objects, the sum-squared error tries to make the confidence score of these cells to zero. This means that their loss will dominate the gradients and it doesn't let the model converge. To solve this issue, the authors introduce the parameters λcoord and λnoobj.

These two parameters weigh the different terms in the loss function to keep the loss due to the cells with no object low. And also weighs the coordinate loss more.

The mathematical formulas used in YOLO (You Only Look Once), focusing on the key aspects:

1.Bounding Box Coordinates:

YOLO predicts bounding boxes for each object in the following format:

a) x: The x-coordinate of the bounding box's center relative to the grid cell.
b) y: The y-coordinate of the bounding box's center relative to the grid cell.
c) w: The width of the bounding box relative to the entire image.
d) h: The height of the bounding box relative to the entire image.
e) confidence: A confidence score that indicates the likelihood that the bounding box contains an object.

2.Class Probabilities:

For each grid cell and each bounding box, YOLO predicts class probabilities using the softmax function:

P(class) = e^(score_class) / Σ(e^(score_i)) for all classes

Where:

- P(class): Probability of an object belonging to a specific class.
- score_class: The raw score for the class.
- score_i: Raw scores for all classes.

3.Intersection over Union (IoU):

The Intersection over Union (IoU) is used to evaluate the overlap between predicted bounding boxes and ground truth bounding boxes. It's calculated as:

- IoU = (Area of Intersection) / (Area of Union)

Where:

- Area of Intersection: The region where the predicted and ground truth bounding boxes overlap.
- Area of Union: The total region covered by both the predicted and ground truth bounding boxes.

4. Loss Function:

- YOLO uses a combination of loss functions during training. The primary loss components are: Localization Loss: Measures the accuracy of bounding box predictions, often calculated using the mean squared error (MSE) between predicted and ground truth bounding box coordinates.
- Confidence Loss: Penalizes confidence scores for false positives and false negatives. It is often calculated using a binary cross-entropy loss.
- Class Loss: Measures the accuracy of class predictions, typically calculated using categorical cross-entropy.

The total loss is usually a weighted sum of these components, and backpropagation is used to update the model's weights. These formulas are at the core of YOLO's object detection capabilities. The specific implementation details and variations can vary among different YOLO versions and implementations.

Yolo Requirements and Build-up functions

1. pip install ultralytics or git clone https://github.com/ultralytics/ultralytics.git

2. In Google colab notebook or any Python Environments (Ubunto recommended):

o from ultralytics import YOLO

Load a model

model = YOLO("yolov8n.pt") # load a pretrained model

Use the model

results = model.train(data="coco128.yaml", epochs=5) # train the model

results = model.val() # evaluate model performance on the validation data set

results = model("https://ultralytics.com/images/cat.jpg") # predict on an image

success = YOLO("yolov8n.pt").export(format="onnx") # export a model to ONNX

3. Editing the Sources location into the data.yaml

train: ../train/images

test: ../test/images

val: ../valid/images

nc: 5

names: ['fish', 'cat', 'person', 'dog', 'shark']

4. Traing data set based on Yolo models

!python train.py --batch 4 --epochs 50 --data /content/drive/MyDrive/Yolov5_colab/data.yaml --cfg /content/yolov5/models/yolov5s.yaml

5. Save the Best.pt file as the trained model and copy into Yolo/weights folder

6. Opening a terminal in Ubunto and YOLO folder

python detect.py --weights best.pt --source1

Source1 is defined a laptop's webcams, this can be changed by specific image: python detect.py --weights best.pt --source input_image.jpg