Construction & Robotics

RESEARCH DRIVEN PROJECT VOL. 7

Research Paper / SS 2024

Sigrid Brell-Cokcan
Thomas Adams

AUTHORS
Xiangcheng Li
Yufei Hu
Zeynep Sertkaya
Jiani Zhu Xingying
Li Zonghai Yang
Alperen Balcin
Merve Yanik
Damla Yigit
Nalan Yilmazarslan

2., corr. ed.

construction robotics
www.cr.rwth-aachen.de

international M.Sc.programme

www.cr.rwth-aachen.de
cr@ip.rwth-aachen.de

Editors:
Univ.-Prof. Dr.-techn. Sigrid Brell-Cokcan
Dr.-Ing. Thomas Adams

RWTH Aachen University
Chair of Individualized Production
Campus-Boulevard 30
52074 Aachen
Germany
www.ip.rwth-aachen.de

Layout:
David Lukert B.Sc.

Original cover:
Lukas Kirner M.Sc.

Funded by Production Engineering - Grant of RWTH Aachen University

The content of the book was created for teaching purposes as part of the Research Driven Projects course.

RWTH Aachen University
Templergraben 55
52066 Aachen

2., corr. ed., 2026

Available via the institutional repository of RWTH Aachen University:
DOI: 10.18154/RWTH-2026-07121

The authors are responsible for the content of the papers.

Preface

We are pleased to present the seventh volume of the Research Driven Project Book, showcasing the work of our students from the summer term of 2024. As digitalisation and automation continue to rapidly advance and transform the construction industry, the Construction & Robotics Master's programme at RWTH Aachen University remains committed to leading this evolution. Our mission is to equip students with the essential tools and knowledge required to drive innovation and improve efficiency, safety and productivity in construction.

This semester, our students have undertaken rigorous research projects exploring various facets of digitalisation and automation in the construction sector. This volume showcases their intellectual rigour, dedication and innovative problem-solving skills as they tackle complex challenges with cutting-edge solutions.

The topics covered in this volume reflect a wide range of interests and underline the vast potential within this dynamic field. Research areas include a semi-automated pavement fleet management system, spatio-temporal multi-agent path planning, SmartRebar: automated reinforcement monitoring system, BIM-integrated automated inspection of HVAC installations, SmartMotion - IoT-based energy efficiency tracking related to kinetic facade movement, and advanced sound detection for construction sites. These projects illustrate how advanced technologies can be used to improve construction processes and outcomes. Each project presented here is a testament to the intellectual curiosity and dedication of our students. They have skilfully integrated their academic knowledge with practical construction site experience to develop innovative concepts.

We are proud of our students' achievements and thank each and every one of them for their extraordinary efforts in making this collection a success. We look forward to their future contributions to the industry.

Sincerely,

UNIV.-PROF. DR.-TECH. SIGRID BRELL-COKCAN
INITIATOR OF CR-MASTER PROGRAM

DR.-ING. THOMAS ADAMS
LECTURER OF RESEARCH DRIVEN PROJECT

The article *"SmartMotion – IoT-Based Energy Efficiency Tracking Related to Kinetic Façade Movement "*by Alperen Balcin, Merve Yanik has been updated for this edition on page 68.

The authors are responsible for the content of the papers.

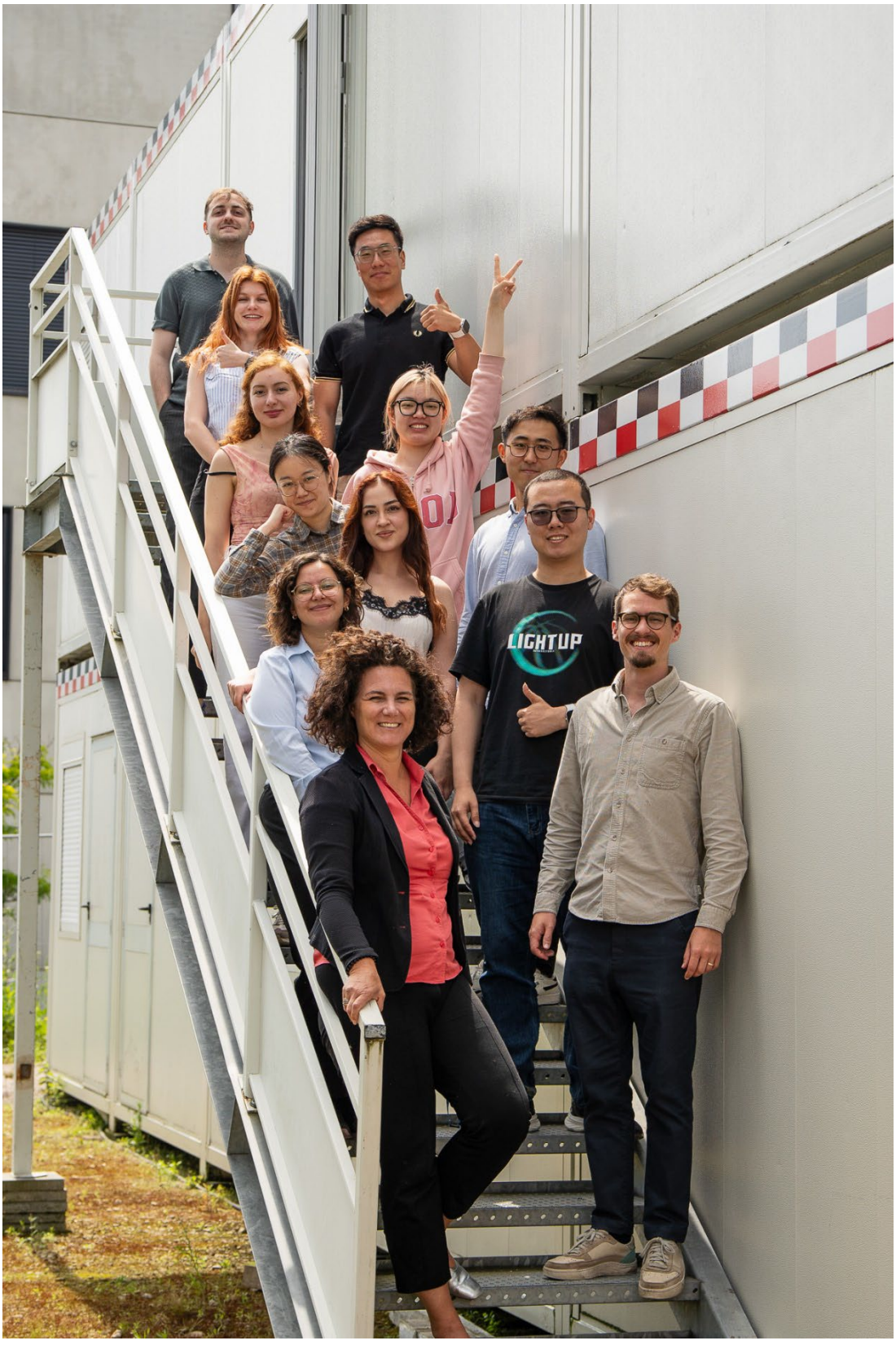


Figure 1: CR master students during RDP on Reference Construction Site in summer term 2024

Contents

Semi-Automated Pavement Fleet Management System 6
Xiangcheng Li, Yufei Hu

Spatiotemporal Multi-Agent Path Planning 24
Zeynep Sertkaya, Jiani Zhu

SmartRebar: Automated Reinforcement Monitoring System 37
Xingying Li, Zonghai Yang

SmartMotion – IoT-Based Energy Efficiency Tracking Related to Kinetic Façade Movement *(corrected version)* 53
Alperen Balcin, Merve Yanik

Advanced Sound Detection for Construction Sites 69
Damla Yigit, Nalan Yilmazarslan

XIANGCHENG LI, YUFEI HU

Semi-Automated Pavement Fleet Management System

ABSTRACT

The German Road Construction market is expected to grow significantly from 2024 to 2030, primarily driven by increased government investment in infrastructure and technological advancements. Innovations, including self-driving vehicles and automated paving systems, have significantly improved construction precision and efficiency. However, the deployment of fully automated systems presents several challenges, including extended setup times, as well as reliability concerns in unpredictable environmental conditions, which highlight the limitations of current technologies.

To address these issues, the Semi-Automated Pavement Fleet Management (SAPFM) System was developed, integrating advanced sensors with adaptive control strategies.

This system is demonstrated through pavement machinery equipped with features like Adaptive Cruise Control, Simulated Pavement Monitoring, and Adaptive Compaction. The SAPFM system balances automation with human oversight, ensuring safety and reliability by dynamically assessing automation feasibility and prompting manual intervention when necessary.

The SAPFM prototype has demonstrated that automation in road construction is feasible even on a small scale, utilizing low-cost sensors. Future development efforts will focus on expanding the SAPFM system's application across the entire pavement fleet, with the objective of optimizing automation and enhancing construction efficiency in complex operational environments.

.

Keywords: Upcycling; Object Detection; pd-AAC; Residue; Inspection Method

DOI: 10.18154/RWTH-2024-09254

Image created with Adobe Firefly

1. Introduction

The road construction industry is rapidly transforming, driven by technological advancements and increased government investment in infrastructure [1]. This growth is driven by the need to modernize and expand transportation networks to meet the demands of a growing population and economy. Innovations in automation and digitalization, including autonomous vehicles and automated paving systems, have significantly enhanced construction precision and efficiency. However, the deployment of fully automated systems presents challenges, including extended setup times for scanning and networking, as well as reliability concerns in unpredictable environmental conditions, which highlight the limitations of current technologies [3].

To address these challenges, this search proposes the Semi-Automated Pavement Fleet Management (SAPFM) System, designed to bridge the gap between full automation and manual operation. The SAPFM System integrates advanced sensors with adaptive control strategies. The prototype vehicle, part of a simulated pavement fleet, incorporates features such as Adaptive Cruise Control (ACC) with Path Tracking, Simulated Pavement Monitoring, and Adaptive Compaction, enabling dynamic assessment of automation feasibility and prompting manual intervention when necessary. This research aims to demonstrate the SAPFM System's potential to enhance the durability and precision of road construction, contributing to the optimization of automation in the pavement construction industry.

2. Literature Review

2.1 Market & Technology Analysis

The German Road Construction market is projected to expand at a compound annual growth rate (CAGR) of 9.5% from 2024 to 2030, primarily driven by increased government investment in infrastructure and advancements in construction technology [1]. The global autonomous construction equipment market is expected to grow at a CAGR of 6.6% from 2023 to 2036, increasing in market size from USD 10.66 million to USD 24.47 million, according to Research Nester. By 2036, setup costs may account for 10-20% of expenditures, while labor cost savings are projected at 30-40%. Equipment lifespan is expected to increase by 10-15%, with emission reductions of 15-20%, reflecting significant efficiency and environmental benefits [4].

Notably, as shown in Diagram 1, the Road Construction Machinery segment constitutes nearly one-third of the market share, underscoring its significance within the infrastructure construction [4]. The growing demand for infrastructure development, especially pavement construction, is the primary driving factor for the adoption of autonomous construction machinery. Key innovations, including self-driving construction vehicles, automated paving systems, and robotic arms, are enhancing precision and efficiency in construction tasks [5].

A notable application case of these technological advancements is illustrated by the Robot Road Construction 4.0 Project (2017-2021), led by STRABAG [3], which highlights significant advancements of autonomous in pavement construction technology. Conducted in collaboration with research and industry partners, the project aimed to develop an interconnected system for autonomous asphalt pavement construction. This system was successfully applied during the renovation of Federal Highway B-189 near Wolmirstedt, Magdeburg, Germany, demonstrating the effectiveness of autonomous asphalt pavers, material handling automation, and digital integration for precise operations. Key developments included the utilization of interconnected measurement and sensor technology, which facilitated real-time data exchange and remote monitoring through a centralized control system, thereby eliminating the need for a physical workstation at the construction site [6].

The project yielded significant outcomes, particularly in enhancing safety and health on construction sites. By distancing operators from traffic and harmful fumes, the project

The authors are responsible for the content of the papers.

adhered to mandatory safe distance guidelines, which prescribe a minimum distance of 80 cm between workers and moving traffic, with an additional speed-dependent safety margin, as stipulated by German safety standards (ASTA) [7]. Furthermore, the project reduced the need for workers in hazardous positions and allowed for the autonomous regulation of paving parameters. These advancements underscore the potential for future paving processes to be fully managed via a control center within the road paver, marking a substantial step forward in the field of autonomous road construction.

However, fully automated systems may face challenges in ensuring safety and reliability under all conditions, such as unexpected obstacles or environmental changes [5]. As noted by Sebastian Czaja, the Project Manager, "The lead time due to 3D scanning and geometry data collection, and the networking of the entire installation process increases the operating time considerably." Additionally, maintaining consistent quality in repetitive tasks remains a challenge [3].

Table 1: Global market forecast for autonomous construction machinery 2024 – 2036 [4]

2023-2036 Forecast	
CAGR	6.6%
Market Size (2023~2026)	10.66 ~ 24.47 Mio $
Setup Costs Percentage (2023~2026)	10-20%
Labor Cost Savings (2023~2026)	30-40%
Lifespan Increase (2023~2026)	10-15%
Emission Reduction (2023~2026)	15-20%

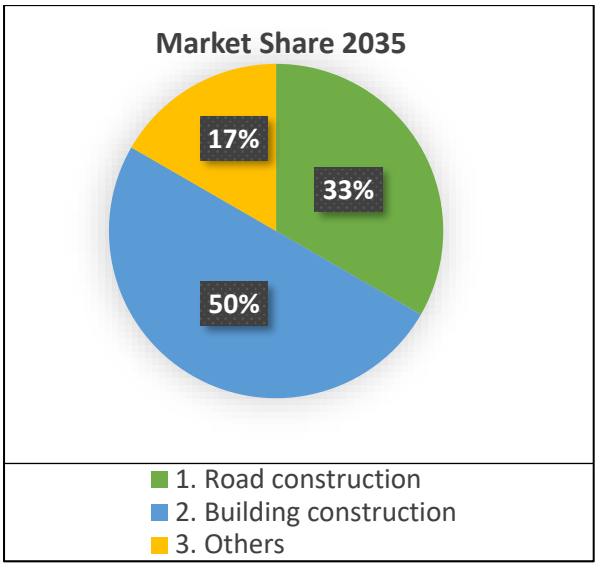


Figure 1: Market share forecast for autonomous construction machinery 2035 [4].

The authors are responsible for the content of the papers.

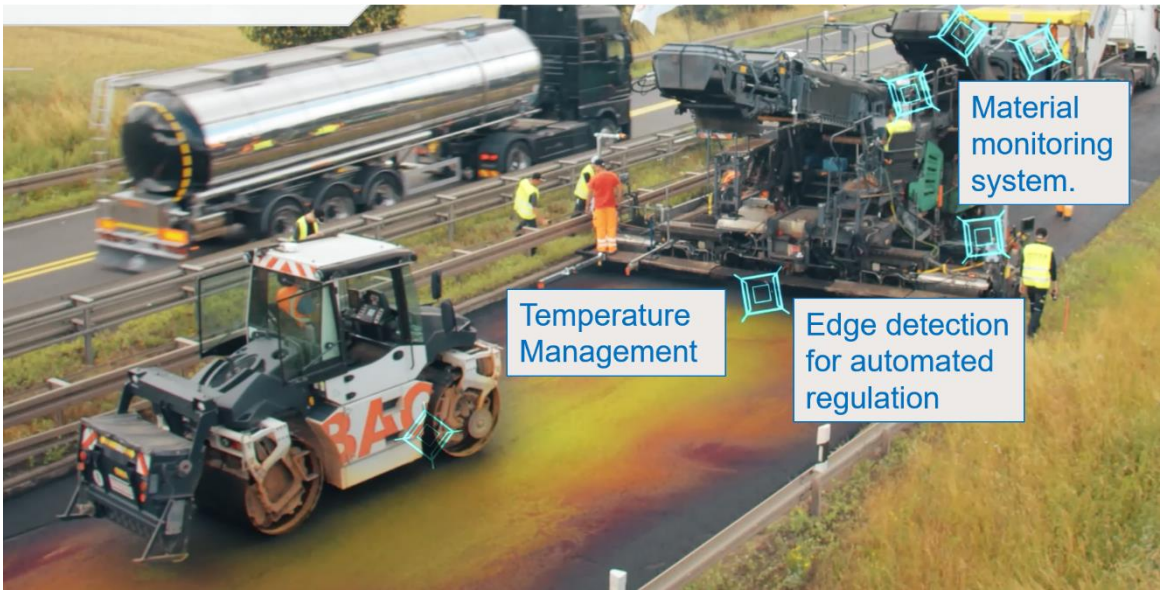


Figure 1: Robot Road Construction 4.0 ´Research Project [3] (adapted from STRABAG).

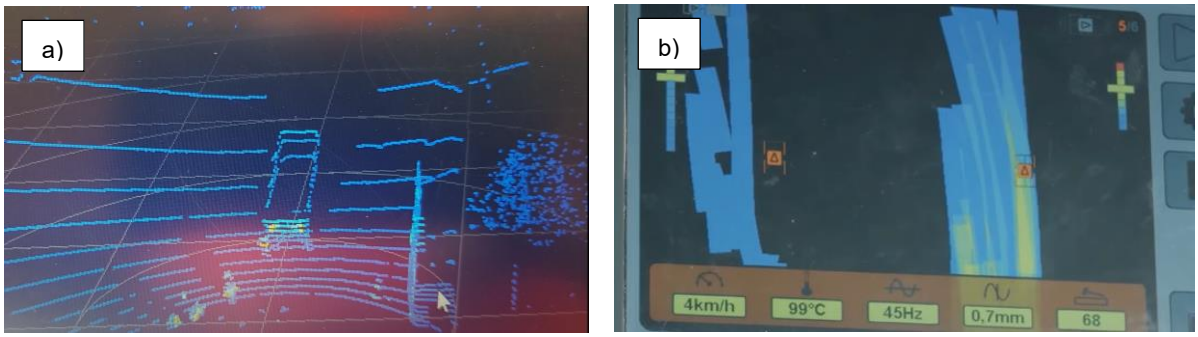


Figure 2: Robot Road Construction 4.0 BG BAU (Fig. 2A) LiDAR scanning (Fig. 2B) pavement geometry.

2.2 Functionality Analysis

The feasibility of automation in road paving has been effectively demonstrated by the Robot Road Construction 4.0 project powered by STRABAG. However, the extensive setup required for autonomous systems can diminish the time efficiencies achieved during the paving process [3]. Fully automated systems face significant challenges in ensuring safety and reliability under varying conditions [9], particularly when confronted with unexpected obstacles or environmental changes.

The implementation of full automation in pavement construction involves substantial challenges. The setup and maintenance of such systems are complex and demand considerable effort. The integration of advanced technologies, such as 3D scanning and real-time networking, can prolong project timelines, thus negating anticipated efficiency gains [3]. Additionally, ensuring reliability in complex and dynamic environments is challenging, as autonomous systems may struggle to adapt to unforeseen obstacles or environmental shifts [8]. The resource demands for deploying and sustaining these systems are also significant, potentially limiting their widespread adoption. Moreover, maintaining consistency and flexibility across diverse operational conditions remains a persistent challenge [9].

This section provides an in-depth analysis of the various aspects, advantages, challenges, functions, and technologies associated with the automation of road pavement construction.

2.3 Literature Conclusion

A critical concern and potential breakthrough in this field is determining the optimal transition point from automation to

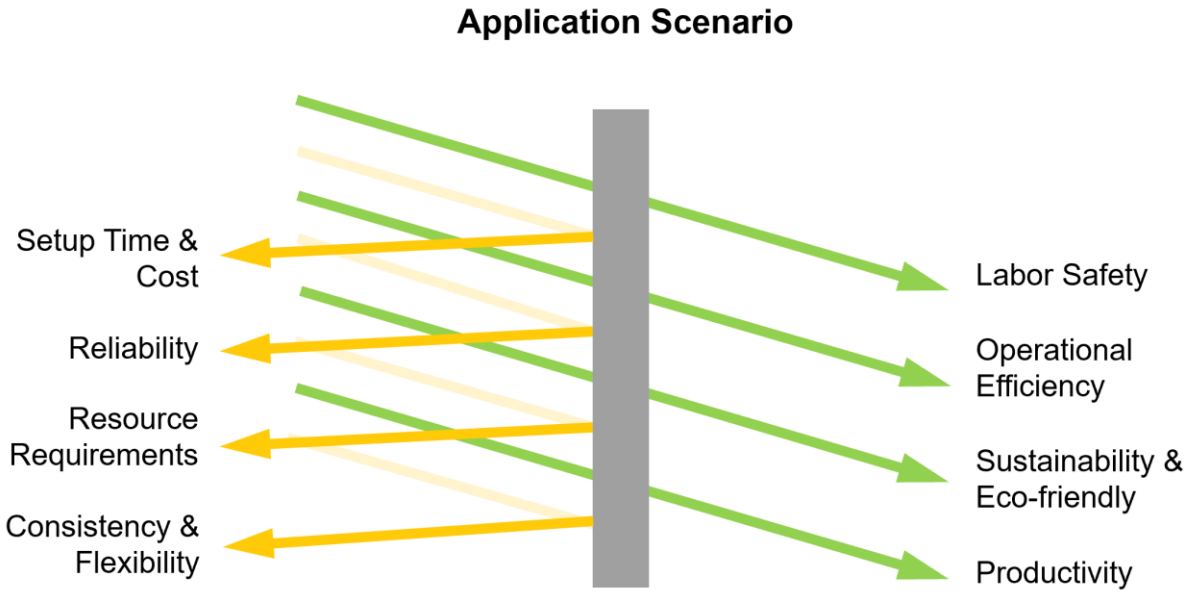


Figure 2: Advantages and challenges of automation's construction.

manual intervention in pavement machinery. Identifying this critical disengagement point is essential for optimizing the balance between full automation and human oversight.

The concept of semi-automation is well-illustrated by products like the SAM (Semi-Automated Mason) system powered by Construction Robotics, USA; and Teleo's Supervised Autonomy technology. SAM boosts productivity by automating repetitive tasks while allowing human operators to focus on critical decision-making, thus enhancing overall efficiency and quality [12]. Similarly, Taleo's technology retrofits existing construction equipment to enable semi-autonomous operations, where machines can be controlled remotely by a single operator. This system provides the flexibility to switch to manual control during critical operations, ensuring both safety and adaptability in complex environments [13]. These examples demonstrate the substantial benefits of semi-automation: increased efficiency, cost-effectiveness, and the ability to maintain human oversight where necessary, making it a highly effective approach in the evolving landscape of road construction.

Semi-automation allows machines to efficiently handle repetitive, labor-intensive tasks, thereby increasing productivity and reducing operational costs. At the same time, human operators remain engaged in critical decision-making roles, ensuring quality and adaptability to unforeseen challenges. In road construction, this approach offers the best of both worlds: the speed and consistency of automation with the flexibility and precision of human intervention, making it a highly effective solution in complex environments where full automation may not yet be feasible.

3. Methodology

To develop an effective and adaptive AI-based semi-autonomous road paver prototype, this study adopts a balanced approach between automation and manual control, ensuring efficient operation under varying road conditions. The system is designed to seamlessly switch between manual and fully automated modes, enhancing operational flexibility and safety.

Table 2: Key Aspects of Autonomous Road Pavement Construction

Aspect	Advantages	Challenges	Function	Technology
Labor & Functional Safety	• Enhanced safety by distancing operators from traffic and harmful fumes [3] • Reduced need for workers in hazardous positions [7]	• Setup & Maintaining effort [3] • Increased lead time due to 3D scanning and networking [8]	Autonomous driving Path guidance Disengagement	- Lidar and Radar Systems [8] - Adaptive Path Guidance [8]
Operational Efficiency	• Improved precision and efficiency in construction tasks [6]	• Reliability under complex environments [9]		- Sensor Fusion Technology [6] - Predictive Analytics [9]
Productivity & Quality control	• Autonomous regulation of paving parameters [8] Enhanced process control [5]	• Consistency & Flexibility [9]	Adaptive Compaction Rolling drum pattern Real- time Pavement monitoring	- Advanced Control Systems [5] - Automated Calibration [8] - Real-Time Data Processing [6]
Sustainability & Eco-friendliness	• Reduced environmental impact through efficient operation [2]	• Resource Requirements [3]	Fuel consumption Reduce Rework	- Emission Monitoring Systems [10] - Sustainable Material Utilization [10]



Figure 3: SAM 100 MAson Robot



Figure 4: Supervised Autonomy Machinery Fleet [13]

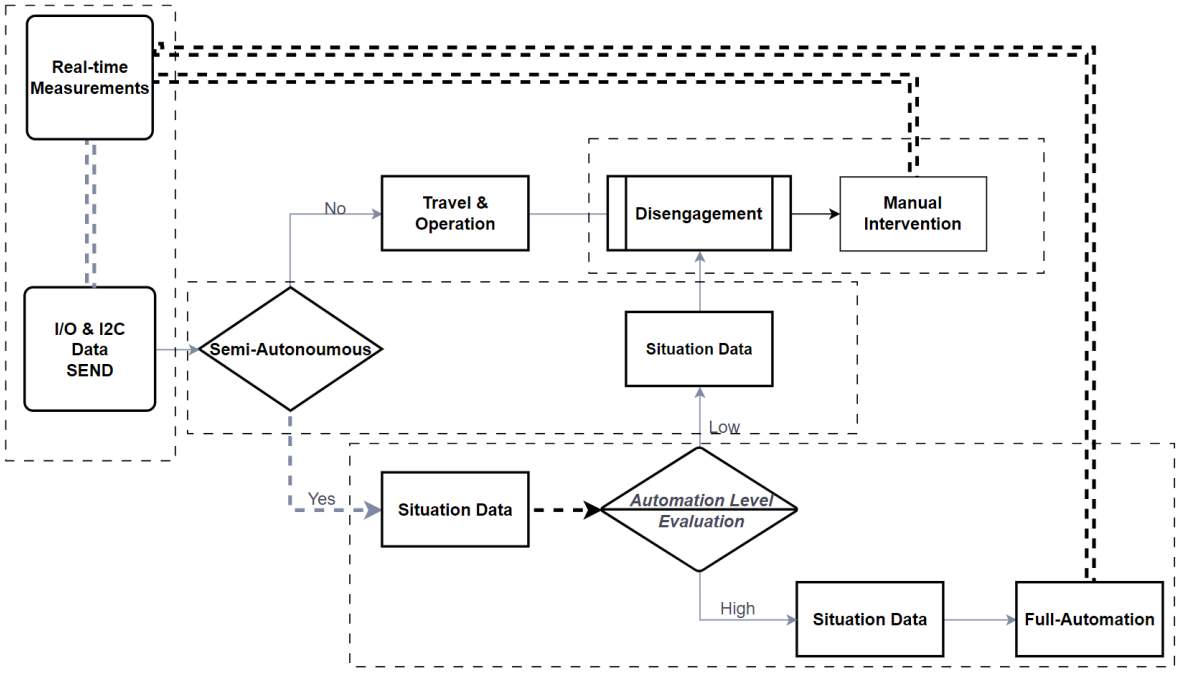


Figure 5: Concept of SAPFM: Automation Evaluation System

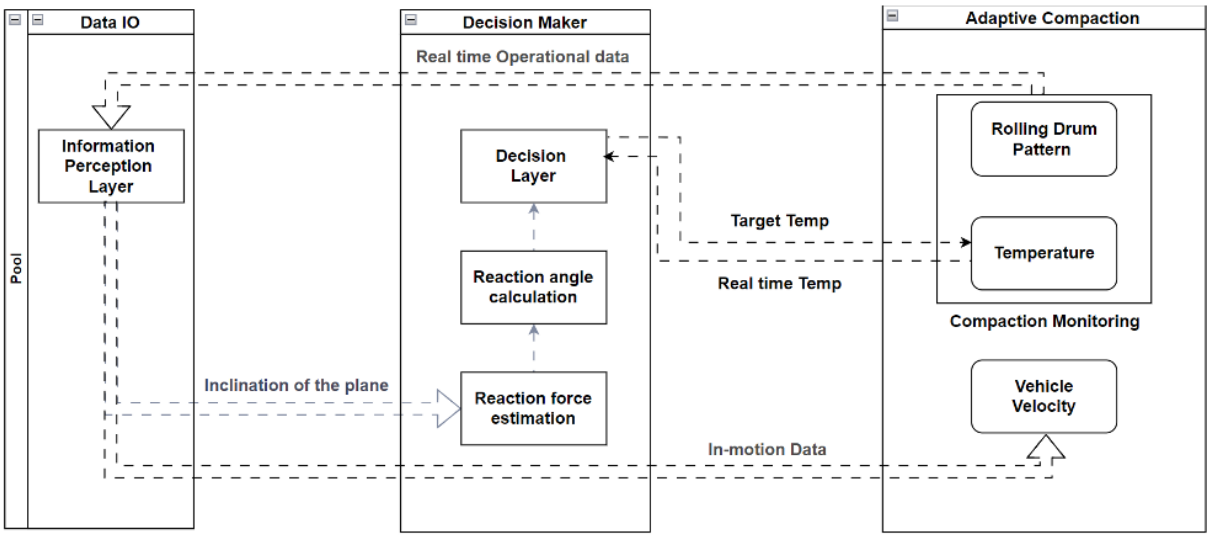


Figure 6: Concept of Adaptive Compaction

3.1 System Architecture

The system architecture is illustrated in Diagram 3, which provides an overview of the core functional modules and data flow.

A. Real-time Measurements and Data Input: The system initially collects real-time measurement data through various sensors integrated into the paver. These data are sent to the central control unit via I/O and I2C interfaces, crucial for assessing the current state of the machine and environment.

B. Automated Operation: The system evaluates whether it can enter automated operation mode based on real-time data. If conditions are favorable, the system transitions into automated operation mode, closely monitoring the situation data and adjusting operations as conditions change.

C. Decision-Making and Control: The decision-making process involves assessing the automation level based on situational data. If conditions are suitable for full automation, the system escalates to that mode; if conditions are less than ideal, the system uses indicators like LEDs to prompt

The authors are responsible for the content of the papers.

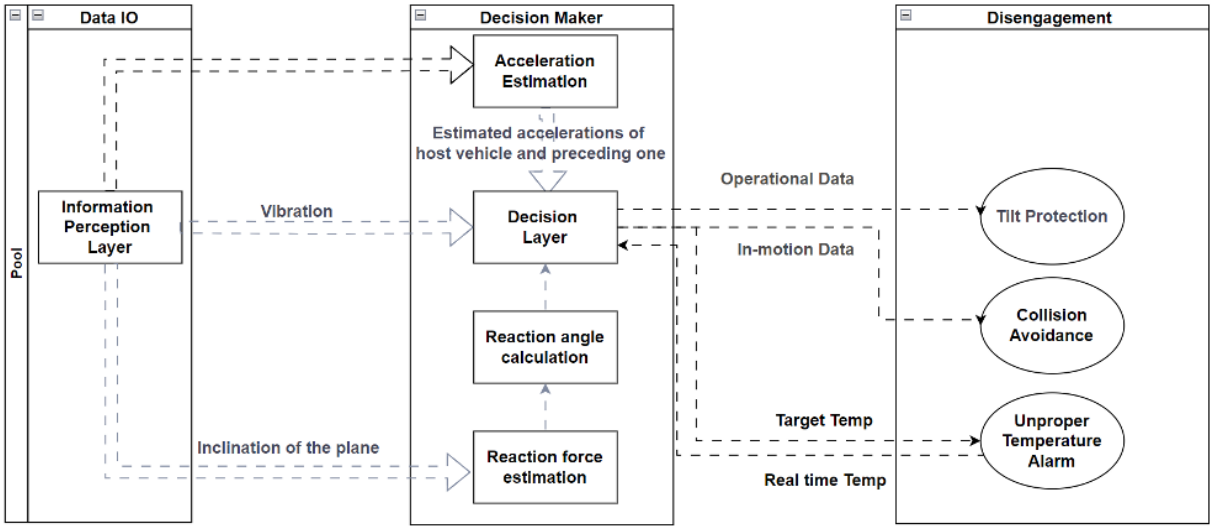


Figure 6: Concept of Adaptive Compaction

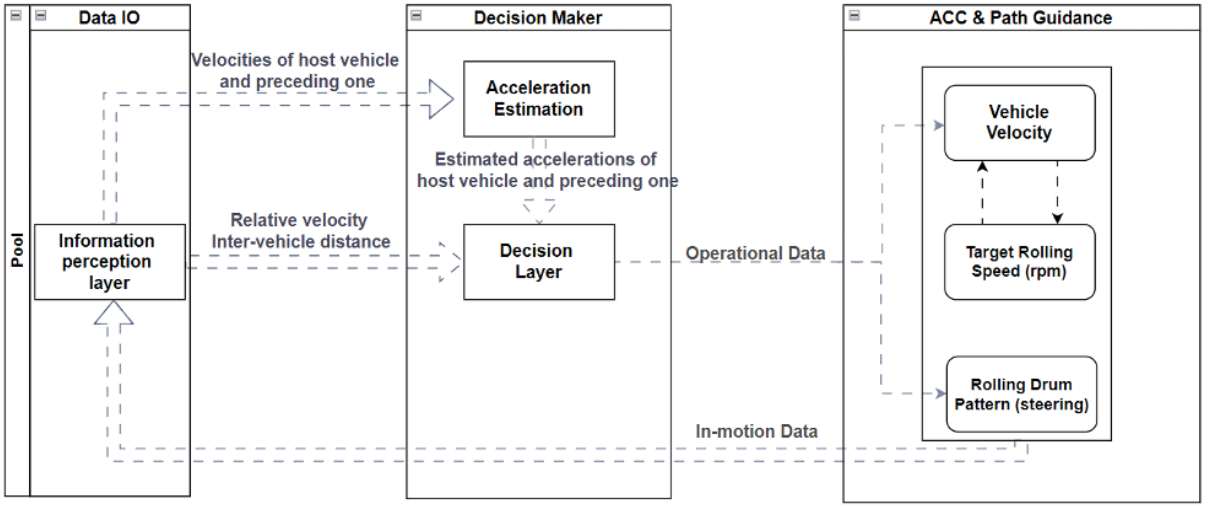


Figure 7: Concept of Adaptive Compaction

the operator to manually intervene. This ensures that the machine remains safe and controllable in complex road conditions.

D. Manual Intervention and Automation Disengagement: When suboptimal conditions are detected, the system prompts the operator via LED indicators to stop full automation and switch to manual mode. While this prompt mechanism does not achieve fully automatic disengagement, it ensures that the operator can respond promptly in critical situations, thereby enhancing operational safety and reliability.

3.2 Experiments and Validation

To validate the system's performance, several experiments were designed and conducted, focusing on the system's automated responses under different conditions.

1. Adaptive Compaction as shown in Diagram 4

The prototype was tested under various environmental conditions. The system automatically adjusted speed, rolling drum pattern, and operation trajectory based on real-time data received from the sensors to adapt to current working conditions.

The system dynamically adjusted the rolling drum pattern and temperature control to ensure optimal compaction quality under different conditions.

2. Disengagement as shown in Diagram 5

When the system detected unfavorable conditions (e.g., excessive tilt, improper

temperature, or potential collision risks), it prompted the operator via LED indicators to stop full automation and switch to manual mode.

This part tested the system's response speed and reliability in scenarios involving tilt protection, collision avoidance, and temperature anomalies, ensuring that the operator could manually take over in a timely manner.

3. ACC and Path Guidance as shown in Diagram 6

During testing process, the system automatically adjusted the paver's speed and guided its path based on real-time data to ensure consistent operation under various conditions.

The system's ability to dynamically adjust the target rolling speed and path, based on data such as the relative speed and distance between the host vehicle and preceding vehicles, was evaluated to ensure the paver could complete tasks stably and safely.

3.3 Performance Evaluation:

The evaluation primarily focused on the system's adaptive cruise control (ACC) performance under various conditions, particularly the interaction between ultrasonic sensor data, motor control, and LED indicators. Of course, this capability is not limited to the ACC function alone but also includes other key components, such as temperature sensors and triaxial sensors. The core of the assessment was how effectively the system could use LED prompts to signal when manual intervention was needed, thereby effectively automating or semi-automating the paver's operation. This demonstrated the system's semi-automation capabilities by accurately reflecting the need for human input based on real-time sensor data.

4. Prototype (Simulated Demonstrator)

The **Semi-Automated Pavement Fleet Management (SAPFM)** system prototype was developed as a conceptual demonstrator to integrate automation into pavement construction processes. Given their critical role in achieving precise compaction, road rollers are particularly well-suited for automation. Consequently, the road roller was selected as the prototype vehicle within the SAPFM framework. The prototype specifically focuses on the implementation and testing of an automated road roller within a series of simulated pavement operations.

4.1 Concept development

The concept of SAPFM centers on balancing automated processes with necessary human intervention. The prototype road roller is designed to adapt to various operational conditions, such as inclines and varying material conditions, using a combination of sensors, microcontrollers, and

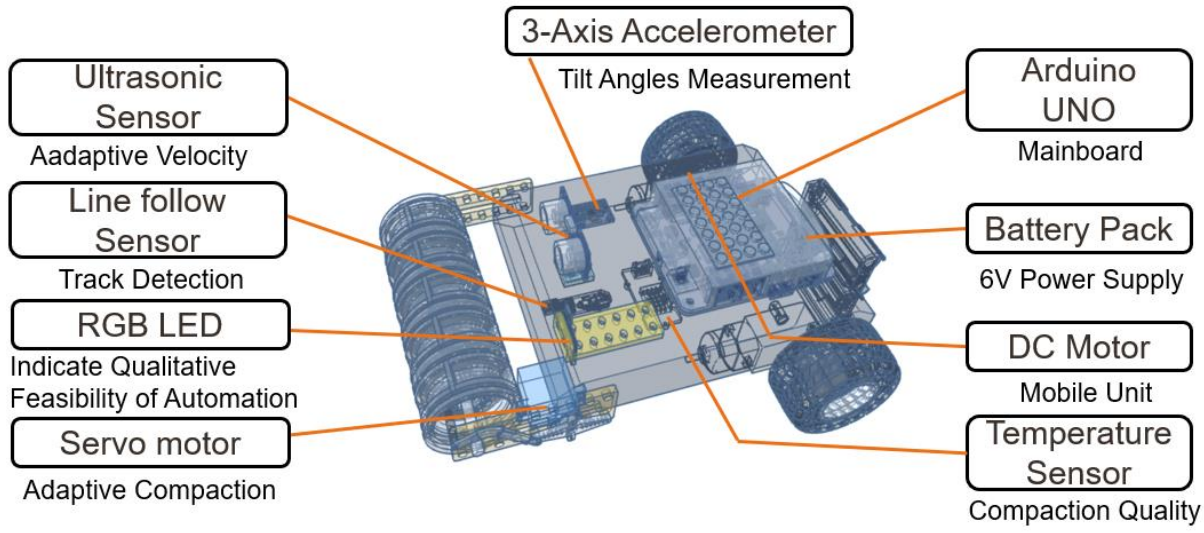


Figure 5: Prototype Hardware Components

The authors are responsible for the content of the papers.

control algorithms. As shown in Figure 5, the key components include:

- Arduino UNO: as the central microcontroller, mainboard.
- Ultrasonic sensors: for distance measurement and obstacle detection. 3-axis accelerometer: for monitoring the tilt and angle during operations.
  - Line-follow sensors: to ensure the roller maintains its correct path.
  - Servo motor: for simulating the compaction force adjustment.
  - Temperature sensor: for simulating the temperature difference of actual working conditions, to detect the compaction quality.
  - DC motor: for supporting the prototype mobile unit.
  - Battery pack: for supporting the prototype movement energy.

The prototype also integrates **RGB LEDs** as a visual indicator of automation feasibility, providing immediate feedback on the system's status based on real-time sensor data.

4.2 Prototype development

The development of the Semi-Automated Pavement Fleet Management (SAPFM) prototype was systematically structured into five key phases, each focusing on specific tasks to ensure a robust and functional system. Below is a detailed breakdown of each phase:

4.2.1 Hardware Preparation

Component Selection: The first step involved selecting the appropriate hardware components essential for the prototype. This included the Arduino UNO as the microcontroller, various sensors such as ultrasonic sensors for distance measurement, 3-axis accelerometers for detecting tilt and angle, and line-follow sensors for path tracking. Additionally, servo motors were selected to simulate compaction force adjustments.

- **Sensor and Actuator Preparation:** Each sensor and actuator was prepared and tested individually to ensure they met the project's technical requirements before being integrated into the system.

4.2.2 Connections and Wiring

- **Circuit Design:** A detailed circuit diagram was created to outline how each component would be connected to the Arduino board. The design ensured that all components could commnicate effectively and that the power supply was sufficient and stable.
- **Wiring and Assembly:** The components were then wired according to the circuit design. Exceptional care was taken to ensure that all connections were secure and that there was minimal interference between different signals. The wiring was also organized to facilitate troubleshooting and future modifications.

4.2.3 Software Development

- **Programming Environment Setup**: The software development began with setting up the programming environment. The Arduino IDE was used, and necessary libraries and dependencies were installed to support the sensors and actuators used in the prototype.
- **Code Implementation**: Custom firmware was developed for the Arduino UNO. The code was written to handle sensor inputs, process data, and control the actuators based on real-time conditions. This included implementing control logic for adaptive speed control, path correction, and automated compaction adjustments.
- **Sensor Data Processing**: Algorithms were developed to interpret data from the sensors accurately, ensuring the system could make informed decisions about speed, direction, and compaction force.

The authors are responsible for the content of the papers.

4.2.4 Testing and Debugging

- **Unit Testing**: Each component and module were individually tested to verify that it performed as expected. This included testing sensor accuracy, actuator responsiveness, and communication between the Arduino board and peripheral devices.
- **Integration Testing**: Once the individual components were verified, they were integrated, and the entire system was tested in a simulated environment. This phase focused on ensuring that all parts of the system worked together seamlessly, and that the system could handle various operational scenarios such as different road inclines and obstacles.

4.2.5 Optimization and Improvement

- **Performance Optimization**: After initial testing, the system's performance was analyzed, and optimizations were made to improve response times, reduce power consumption, and enhance overall stability. This included fine-tuning sensor sensitivity and refining the control algorithms.
- **Function Expansion**: Based on the results from testing and feedback from initial demonstrations, additional features were developed to expand the prototype's capabilities. This included improving the system's adaptability to more complex paving scenarios and enhancing the user interface for better manual control and monitoring.

4.2.6 Prototype demonstration

During the demonstration, the Semi-Automated Pavement Fleet Management (SAPFM) prototype exhibited its advanced capabilities in automated control, sensor integration, and user-friendly status indication, particularly aiding in the transition between fully automated and semi-automated modes.

Link to ''Semi-Automated Pavement Prototype''

https://github.com/Felix1199/Xiangcheng

The prototype leveraged the ultrasonic sensor to implement Adaptive Cruise Control (ACC). This sensor accurately measured the distance to obstacles and automatically adjusted the wheel's speed accordingly. This feature ensured safe and efficient operation under varying distance conditions, adapting the speed based on real-time data.

The 3-axis accelerometer allowed the prototype to detect changes in road incline. When an uphill or downhill slope was detected, the system automatically adjusted the roller's speed according to the detected angle, ensuring consistent compaction quality and operational safety. Additionally, the accelerometer data was utilized to control the servo motor, adjusting the roller's center of gravity under different slope conditions, further enhancing stability and compaction effectiveness.

The temperature sensor played a vital role by monitoring the road surface temperature in real-time. The prototype adjusted the wheel's speed based on this temperature data, ensuring optimal operation in various temperature conditions, thereby achieving the best possible compaction results.

Moreover, the prototype included line-follow sensors that enabled the roller to adhere to a predetermined path accurately. This ensured that the roller maintained the correct trajectory during operation, further enhancing the precision of the pavement construction process.

One of the standout features of the prototype was its ability to integrate data from multiple sensors—ultrasonic, 3-axis accelerometer, and temperature sensors—and use this information to control the RGB LED indicators. The RGB LED system provided a clear visual representation of the machine's working status, changing colors to reflect the current operational mode and sensor readings. This feature facilitated the transition between fully automated and semi-automated modes, allowing operators to make informed decisions quickly.

Link to ''Semi-Automated Pavement Prototype code''

https://github.com/Felix1199/Xiangcheng

The authors are responsible for the content of the papers.

Overall, the SAPFM prototype demonstrated its capability to respond not only to individual sensor inputs but also to synthesize data from all sensors, enabling synchronized operations and seamless mode transitions. The system's intuitive RGB indicators, responsive control mechanisms, and precise path-following functionality collectively ensure that the prototype meets the demands of modern, automated pavement construction.

5. Scientific Analysis, Results and Discussion

5.1 Scientific analysis

5.1.1 Ultrasonic sensor

The distance d to the target is calculated using the speed of sound v and the total time t it takes for the sound to travel to the target and return. The formula is:

$$d = \frac{v * t}{2}$$

Here, $\frac{t}{2}$ is the time for a one-way trip, and $v * t$ gives the distance traveled by the sound wave.

5.1.2 3-Axis Sensors:

3-axis sensors measure acceleration or magnetic field components along the x, y, and z axes. Key formulas include:

Acceleration Vector Calculation: For an accelerometer measuring acceleration components a_x, a_y, and a_z, the total acceleration magnitude a is:

$$a = \sqrt{{a_x}^2 + {a_y}^2 + {a_z}^2}$$

Tilt Angle Calculation: To calculate the tilt angle σ of a device using accelerometer data, assuming a_x and a_z are the acceleration components along the x and z axes, the angle is:

$$\sigma = \arctan\left(\frac{a_x}{a_z}\right)$$

Pitch Angle (θ): The pitch angle is the rotation around the device's x-axis (tilting forward or backward), and it can be calculated using the formula:

$$\theta = \arctan\left(\frac{a_y}{\sqrt{a_x^2 + a_z^2}}\right)$$

Roll Angle (φ): The roll angle is the rotation around the device's y-axis (tilting sideways), and it can be calculated using the formula:

$$\phi = arctan\left(\frac{a_x}{a_z}\right)$$

5.1.3 Temperature sensor:

Temperature sensors measure temperature using different principles. For thermocouples, the formulas is:

Thermocouple Temperature Calculation: Thermocouples generate a voltage V that relates to temperature T. The approximate formula for temperature is:

$$T = \frac{V}{S} + T_0$$

Where S is the sensitivity of the thermocouple (in mV/°C), and T_0 is the reference temperature (usually 0°C).

5.1.4 Sensor Fusion:

Combining data from multiple sensors can improve accuracy and reliability. A common fusion method is weighted averaging:

Weighted Average: When combining sensor readings X_i with corresponding weights w_i, the fused result $\bar{x}_{fused}$ is calculated as:

$$\bar{x}_{fused} = \frac{\sum_{I=1}^{N} w_i x_i}{\sum_{I=1}^{N} w_i}$$

The authors are responsible for the content of the papers.

Table 3: SAPFM simulation track-test results

Distance (cm)	Pitch (degrees)	Roll (degrees)	Sensor Value	Speed L/R	LED Color	Automation Feasibility %
23,79	6,27	-4,87	0	255/255	GREEN	80
22,5	-8,72	-8,43	1	80/255	GREEN	90
12,54	2,34	-4,29	2	255/80	BLUE	70
19,63	-3,46	-7,87	3	0/0	BLUE	75
19,25	9,32	-8,51	1	80/255	BLUE	65
6,72	-2,16	1,65	1	80/255	YELLOW	50
9,11	2,58	-3,01	1	80/255	YELLOW	50
8,82	-7,72	-9,6	0	80/80	YELLOW	30
0,23	4,69	4,66	3	0/0	RED	10
2,68	0,31	-3,61	1	80/255	RED	5

Table 4: Full automation simulation track test result

Distance (cm)	Sensor Value	Speed L/R	LED Color	Automation Feasibility %
9,68	0	0	GREEN	61
23,03	1	360	GREEN	71
6,74	1	0	GREEN	80
16,24	1	240	GREEN	81
22,96	0	360	GREEN	89
9,68	1	0	GREEN	97
26,96	0	360	BLUE	83
26,24	1	360	BLUE	81
33,52	1	510	RED	61
30,99	0	510	RED	92

The authors are responsible for the content of the papers.

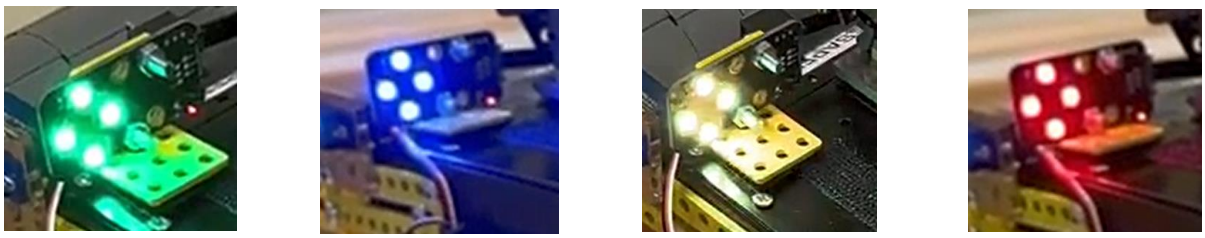


Figure 7: RGB indicate qualitative feasibility of automation under various conditions.

Table 5: Automation Feasibility Criteria

Color	Green	Blue	Yellow	Red
Automation Feasibility %	>70	55~70	25~55	<25

5.2 Analysis of the experimental results

This experiment analyzed the feasibility of both full automation and semi-automation modes under varying environmental conditions. Through a comprehensive assessment of factors such as distance, pitch angle, roll angle, sensor values, speed, and LED color, it was observed that the full automation mode performs well in simple and stable environments but experiences a significant decrease in feasibility when confronted with complex and dynamic conditions. In contrast, the semi-automation mode demonstrated greater adaptability, maintaining stable feasibility in challenging environments. Consequently, it is recommended to utilize the strengths of both full automation and semi-automation, selecting the most suitable mode based on specific environmental conditions to ensure both efficiency and safety in construction operations.

5.2.1 Results Analysis

- **Automation Feasibility:** The blue curve representing full automation feasibility illustrates how the system responds under varying conditions. When the sample number increases, representing different testing scenarios, the automation feasibility fluctuates significantly, especially in later stages. This pattern indicates that maintaining consistent automated operations is challenging in dynamic and complex environments. The pronounced dips in the curve highlight the difficulties of relying solely on full automation in construction settings, where environmental factors can drastically affect performance.

- **The Necessity of Semi-Automation:** The red curve emphasizes the importance of a semi-automated approach. While full automation shows promise in specific scenarios, such as stable and straightforward conditions, the semi-automated approach demonstrates better adaptability in complex situations. The semi-automation line remains more stable across various sample numbers, suggesting that combining automation with human oversight ensures greater safety and quality, especially in challenging environments. This approach allows for better handling of unexpected variables, making it a more viable solution for current and future road construction projects.

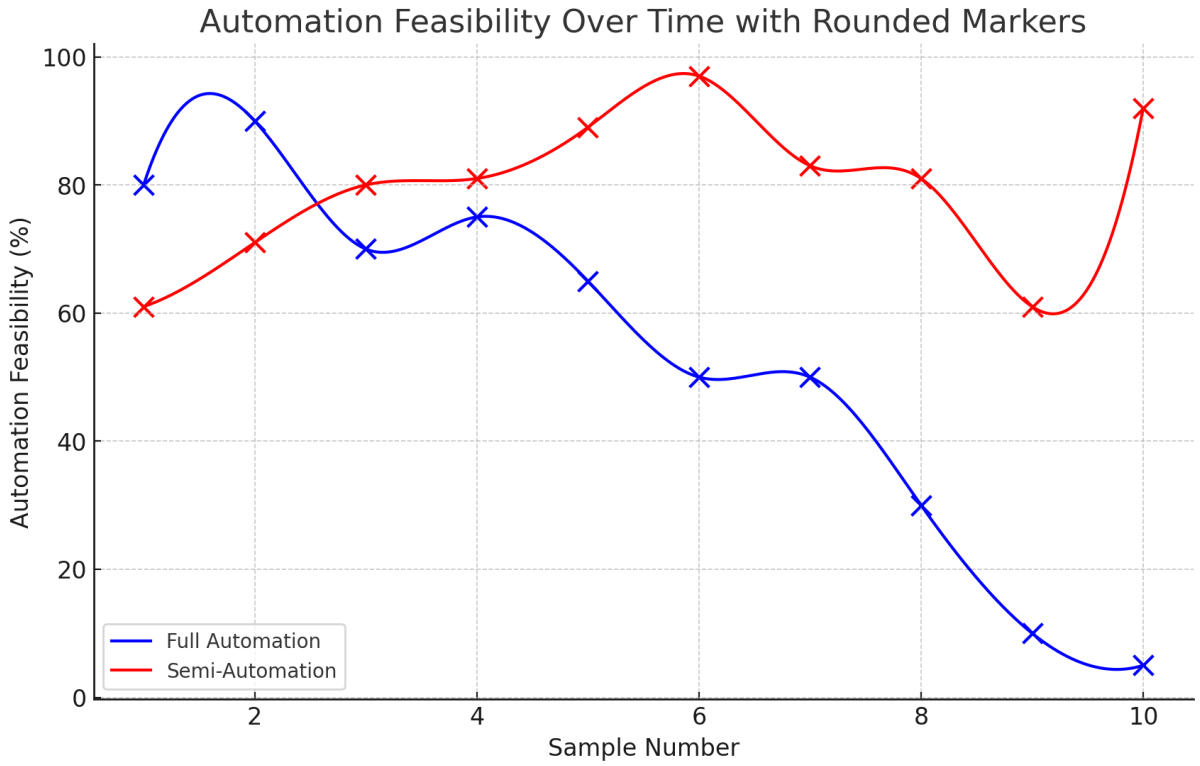


Figure 8: Simulated Track-Test Results

5.2.2 Discussion:

The SAPFM system has demonstrated high flexibility and adaptability under various test conditions, particularly when frequent adjustments to speed and operational strategies are required. While the system performs excellently in automated operations under stable conditions, its feasibility is challenged in more complex or rapidly changing construction environments. This indicates that a hybrid approach, combining full automation and manual control—semi-automation—can provide the optimal balance in practical construction scenarios. Future research could focus on further improving sensor stability and response time to enhance the system's performance in diverse construction environments and explore additional balance points between automation and semi-automation.

5.3 Outlook

The module explores the future potential of the SAPFM system and related technologies. It identifies emerging trends, potential research areas, and practical applications that could drive further advancements and improve the system's capabilities.

- **Advancements in Sensor and Data Processing Technologies:** This section outlines potential technological advancements that could enhance the SAPFM system. It includes discussions on emerging sensor technologies, data fusion techniques, and improvements in processing algorithms. These advancements are expected to address current limitations and enhance the system's performance in various operational scenarios.
- **Integration of AI and Machine Learning:** Explore how artificial intelligence (AI) and machine learning (ML) algorithms can be integrated into the SAPFM system to enhance its adaptability and decision-making capabilities [14]. The focus is on developing intelligent algorithms that can dynamically adjust system parameters and optimize performance based on real-time data and environmental changes.

The authors are responsible for the content of the papers.

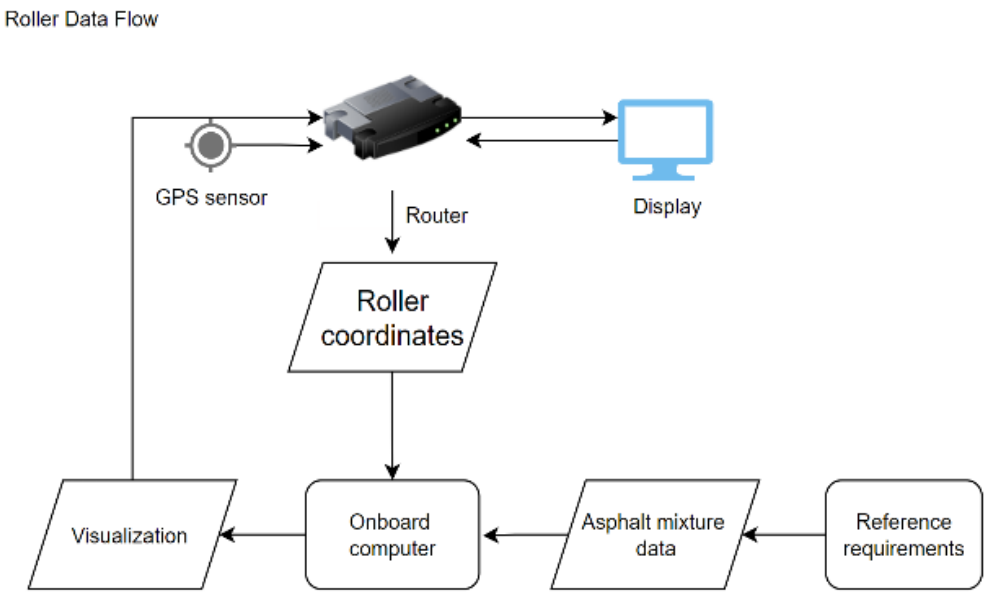


Figure 8: Road roller communication Data flow

- **Enhancing Semi-Automation Effectiveness:** It includes exploring how to better integrate manual intervention with automation, improving the flexibility and responsiveness of the system in various construction environments. The goal is to maximize the benefits of semi-automation and address any remaining challenges.
- **Communication:** Connected and autonomous vehicles (CAVs) is crucial for optimizing the Semi-Automated Pavement Fleet Management (SAPFM) system, enabling real-time data exchange between vehicles and the central processing center [15]. This ensures synchronization and enhances the effectiveness of semi-automated operations through dynamic response capabilities. Future advancements in CAV could further elevate the system's performance, solidifying its role in next-generation road construction technologies. For more detailed insights, you can explore further here.
- **Expansion and Real-World Testing:** Impact on the practical application of the SAPFM system in a wider range of construction environments. It includes plans for field trials and real-world testing to validate system performance and gather feedback for further refinement. Successful deployment in diverse settings will facilitate broader adoption and application of the technology.

6. Conclusion

The development and implementation of the Semi-Automated Pavement Fleet Management (SAPFM) System present a preliminary investigation into the feasibility of automation in road construction. SAPFM offers an initial exploration into small-scale automation in road construction. By low-cost sensors and adaptive control strategies, the prototype demonstrates feasibility in adaptive pavement settings. The SAPFM system effectively addresses the challenges posed by simulated fully automated systems, particularly in complex and unpredictable environments. The system's ability to dynamically assess the feasibility of automated operations and seamlessly transition to manual intervention when necessary, underscores its practical value in balancing automation with human oversight.

This research demonstrates that the SAPFM system not only enhances the precision and efficiency of road construction but also ensures the necessary flexibility to adapt to varying operational conditions. The hybrid approach of semi-automation proves especially beneficial in scenarios where full

automation may fall short, highlighting the critical role of human intervention in maintaining safety and adaptability. The system's performance in diverse test conditions confirms its potential to significantly improve the durability and precision of road construction projects.

Looking ahead, the SAPFM system presents significant opportunities for enhancement, particularly through advancements in sensor technology and data processing, which are crucial for enabling quicker and more accurate responses to environmental changes. These improvements will make the system more efficient and reliable across diverse construction scenarios. Future research should also aim to optimize the balance between automated and manual operations, refining the transition points to maximize operational efficiency while ensuring human oversight during critical decisions. Additionally, developing more sophisticated intervehicle communication systems will be vital for improving fleet coordination, enhancing safety, and streamlining operations. As the construction industry evolves, the SAPFM system is poised to become a key model for integrating automation with human expertise, driving safer, more efficient, and sustainable infrastructure development.

7. References

[1] Market Analysis Report, Construction Equipment Market Size & Share Report, 2020-2027 978-1-68038-118-4 www.grandviewresearch.com/ (Accessed June. 05, 2024).

[2] Research and Markets Analysis Report, "Autonomous Construction Equipment Global Market Opportunities and Strategies to 2033." www.researchandmarkets.com/ (Accessed June. 05, 2024).

[3] S. Czaja. "Autonomous asphalt paving in the real world" https://newsroom.strabag.de/news-robot-strassenbau-40-schafft-basis-fuer-den-autonomen-asphalteinbau?id=150641&menueid=28033&l=deutsch (Accessed June. 01, 2024).

[4] Research Nester. "Globale Marktgröße Prognose und Trendhighlights für autonome Baumaschinen im Zeitraum 2024 – 2036" https://www.researchnester.com/de/reports/autonomous-construction-equipment-market/4818/ (Accessed June. 15, 2024).

[5] P, Wolf, et al. "Towards High-Quality Road Construction: Using Autonomous Tandem Rollers for Asphalt Compaction Optimization."ISARC Proceedings, 14 Oct. 2020, pp. 90–97, www.iaarc.org/publications/2020_proceedings_of_the_37th_isarc/towards_high_quality_road_construction-using_autonomous_tandem_rollers_for_asphalt_compaction_optimization.html. (Accessed June. 15, 2024).

[6] BG BAU, Forschungsprojekt "Robot - Straßenbau 4.0" https://bauportal.bgbau.de/bauportal-12023/thema/tiefbau/projekt-robot-40-autonom-arbeitende-maschinen-im-strassenbau/ (Accessed June. 15, 2024).

[7] H. Leisering. "Die Arbeitsstättenregel ASR A5.2 Verbesserter Arbeits- und Gesundheitsschutz für Beschäftigte auf Straßenbaustellen, Arbeitsstellensicherung" ASTA (Accessed June. 18, 2024)

[8] W. Huang, et al. "Design and Implementation of PLC-Based Autonomous Construction System of Unmanned Vibratory Roller," 18, 2024)

[9] K. Sitaramanjaneyulu. " Smart Pavement Materials and Emerging Technologies." https://www.nbmcw.com/article-report/infrastructure-construction/roads-and-pavements/smart-pavement-materials-and-emerging-technologies.html/ (Accessed June. 28, 2024).

[10] A. Owais. "Robots in Construction: Bricklaying." AZoRobotics, 12 Dec. 2023, www.azorobotics.com/Article.aspx?ArticleID=654. (Accessed June. 28, 2024).

[11] V. Vince. "Precision Robotic Masonry - Vince Vasquez - Medium." Medium, Medium, 5 Dec. 2019, medium.com/@vincevasquez/precision-robotic-masonry-bb64a103559a. (Accessed June. 28, 2024).

[12] K. Jost. "Teleo Aims to Transform Off-Road Equipment Operations with Supervised Autonomy." Inside Autonomous Vehicles, Inside Autonomous Vehicles, 30 Aug. 2022, insideautonomousvehicles.com/teleo-aims-to-transform-off-road-equipment-operations-with-supervised-autonomy/. (Accessed June. 28, 2024).

[13] S. Burges. “Autonomous Vehicles: The Future of the Construction Industry.”, www.burges-salmon.com/news-and-insight/legal-updates/autonomous-vehicles-the-future-of-the-construction-industry. (Accessed 31 July 2024).

[14] S. Heeseung, et al. “Autonomous Condition Monitoring-Based Pavement Management System.” Automation in Construction, vol. 138, no. 0, 2022, trid.trb.org/view/1939875. (Accessed 31 July 2024).

ZEYNEP SERTKAYA, JIANI ZHU

Spatiotemporal Multi Agent Path Planning

ABSTRACT

This study investigates the use of spatiotemporal path planning to enhance autonomous navigation in dynamic environments. The spatiotemporal path planning methodology proposes an approach for multiple mobile robot navigation based on predicting agent's near-future trajectory. The predicted trajectories are converted into a time series of cost maps which are used to guide the agent to a desired point in space and avoid where collision in future might occur.

The primary research question focused on whether spatiotemporal path planning could improve navigation efficiency and safety. Initial tests showed that with the algorithms to incorporate the probability distribution of displacement, the system demonstrated a 36.75% reduction in motion time and improved collision avoidance capabilities. These results confirm the effectiveness of spatiotemporal path planning in optimizing autonomous movement in complex scenarios.

The project's findings are consistent with similar research, such as the work by Sato et al. (2023), which also demonstrated improved performance using trajectory prediction and spatiotemporal planning in environments with pedestrians. Future work will focus on incorporating advanced machine learning models, scaling the system for more complex environments, and conducting real-world tests to further validate the approach. Overall, this research shows that spatiotemporal path planning is a viable strategy for enhancing the efficiency and safety of autonomous systems, making it applicable to various fields including robotics, urban transportation, and industrial automation.

Keywords: multi-robot pathplanning; robot learning; dynamic environment; mobile robot

DOI: 10.18154/RWTH-2024-09256

Image created with Adobe Firefly

1. Introduction

Future developments of robotics require robots to coexist in a shared environment and operate with each other. Multi-agent systems explore the challenge to control the communication, coordination, and collaboration between multiple robots operating together in complex and dynamic environments. With the increased number of agents, or non-stationary obstacles, such as other robots or moving objects, the dimensionality of the problem and the complexity of computation for obstacle avoidance grows exponentially. Our main objective is to outline a methodology for mapping displacement predictions with time layer based on agent observations. This allows us to operate on a gridworld with the added time dimension, and track dynamic obstacles in possible near-future states [1]. Therefore, in this research, the term spatiotemporal path planning refers to the integration of spatial and temporal data to represent dynamic positions of agents and obstacles to map all near-future states and plan the optimal path for the agent.

We commence with the research questions

- How can spatial and temporal interpolation with Hermite cubic splines
- How are displacement predictions effectively converted into a time series of cost maps to navigate an agent in a dynamic environment?'

We built our research and this paper, on three foundations: (i) literature review and current developments, (ii) creating a research methodology, (iii) prototyping and conducting quantitative analysis with a test setup to experiment the proposed approach. Thus, the following result is derived: converting displacement predictions of dynamics obstacles into a time series of cost maps significantly reduces computation payload for agent's path plan and yield improved robustness in the navigation.

2. Literature review

Research on path planning methodologies for robots operating in dynamic environments has explored various solutions [1–5]. There are studies that propose methods for navigating environments with single or multiple dynamic obstacles, single or multiple agents, and finding optimal shortest path while considering various constraints. It must be remembered that the scenario proposed in each research is unique and therefore requires a specific solution.

2.1 Approaches for Dynamic Environments

Traditional motion planning algorithms in dynamic environments, such as A* search or dynamic window, often rely on currently observed information [6] and complex optimization problems [7] that struggle to effectively offer a solution when the number of dynamic agents increases [8], and their stationary model-based approach requires the path planning to be recalculated from scratch in the case that a deviation occurs in the environment. For example, the dynamic window approach usually assumes that all obstacles observed in the present remain stationary; consequently the agent may try to move into areas that will be occupied by a dynamic obstacle in the future. In this case, the agent would have to freeze or make a sharp turn, and restart fresh to reach the goal position. In order to realize smooth motion in a multi agent environment, future trajectories of obstacles must be considered [9].

2.2 Learning Based and Prediction Based Methods

Recent research has explored learning- and prediction-based methods to address the limitations of model-based planning frameworks in the complex and real-world scenarios [10]. Liu et al. [11] propose a decentralized structural recurrent neural network (DS-RNN) that stores the time-series data as a hidden state vector in the RNN [11] and used as the input to the controller. Dugas

The authors are responsible for the content of the papers.

et al. [12] proposes an end-to-end path planner using LiDAR data or local maps as inputs and outputs velocity commands. They represent the LiDAR data of robot's local environment as vectors in space by combining auto-encoders and transformers, to be able to transform the non-numerical data into numerical data, which is then used as the input to the controller supported with deep reinforcement learning (DLR). However, implementing a controller that handles all state probabilities takes significant computation and training, and in addition, the DLR method only determines the next action, and is therefore short-term.

2.3 Spatial and Temporal Analysis Implemented Methods

Time Enhanced A* search [13] is a multi-robot path planning using future occupied states, and it allows each agent to recognize the area it will occupy at each point by determining search priorities of the agents. While this approach provides a long-term planning strategy in the field of multi-agents, as the number of agents and missions increases, the prioritization of path search requires greater computation, resulting in suboptimal robustness.

D* (Dynamic A*) is an advanced path planning algorithm that extends the A* search algorithm to handle dynamic changes in the environment efficiently. It recalculates paths in real time when obstacles are encountered or when the environment changes [14–16], making it well-suited for applications where the map or terrain is not fully known in advance or is subject to changes. D* considers spatial aspects by finding the shortest path to the goal and adapting that path when obstacles appear. However, while D* can react to changes in real time, it primarily focuses on spatial adjustments rather than incorporating a predictive model that anticipates future states of the environment [2, 17]. In contrast, spatiotemporal path planning integrates both spatial and temporal analysis by predicting the future positions of obstacles [18] and dynamically adjusting the path based on those predictions. This allows spatiotemporal methods to proactively plan for potential future collisions rather than just reacting to present obstacles, offering a more comprehensive approach to navigation in highly dynamic environments.

Our research builds on existing research by trajectory prediction of dynamic obstacles, however, diverges with the implementation of Hermite cubic spline interpolation in spatial and temporal layers.

3. Research methods

Our method exploits the agent's observations to generate probability distributions for displacement of obstacles, which is then mapped as displacement predictions. In the initial step, the agent's position is estimated in gridworld coordinates from sensory input. After localization of the agent, obstacle observations in robot's coordinate system are transformed to gridworld coordinates with the added parameters: time t, obstacle ID h, position (x,y). From this point onward, the time value t is the representation of future time intervals with $t = 0$ being the current time. To predict trajectories of obstacles in future time intervals, we employ a combination of sparse graph convolution network (SGCN) and Hermite cubic spline interpolation (HCSI) in spatial and temporal dimensions.

3.1 Step 1: Finding Future Positions of Obstacles with Sparse Graph Convolution Network (SGCN)

The network focuses on capturing the interactions among obstacles by modeling them as a sparse graph where each node represents an obstacle [19], and edges represent interactions. The network generates multiple displacement probability distributions for each obstacle. These distributions are then mapped as discrete point estimates which represent future positions of obstacles at future time intervals. This step estimates where each obstacle will be located at a certain time step.

The authors are responsible for the content of the papers.

3.2 Step 2: Creating Trajectories by Interpolating Between Points with Hermite Cubic Spline

This step models the motion of obstacles by generating knots between two consecutive points of positions, and creating smooth curves of motion. After SGCN provides discrete point estimates for predicted future positions of obstacles, we must interpolate the discrete points into continuous curves for smooth transitions. To achieve this, first we define consecutive points and their first derivatives.

3.2.1 Defining Consecutive Points and Their First Derivatives

The sequence of predicted positions for each obstacle, which are denoted as $(x_0, y_0), (x_1, y_1), \ldots (x_n, y_n)$ where n represents the time step, are used to calculate the first derivative (or the tangents) at each point. This information gives us the direction of the movement.

3.2.2 Constructing the Hermite Cubic Spline

For each segment between consecutive points in the predicted paths generated by SGCN, Hermite Cubic Spline Interpolation (HCSI) is employed.

Hermite splines use four basis functions H_1, H_2, H_3, H_4 to interpolate between points.

$$H_1(t) = 2t^3 - 3t^2 + 1$$
$$H_2(t) = -2t^3 + 3t^2$$
$$H_3(t) = t^3 - 2t^2 + t$$
$$H_4 = t^3 - t^2$$

Using these basis functions, we compute the position $P(t)$ on the spline between P_i and P_{i+1} so:

$$P(t) = H_1(t)P_i + H_2(t)Pi + 1 + H_3(t)T_i + H_4(t)Ti + 1$$

where *t* varies from 0 to 1 for each segment. P(t) provides the curve between P_i and P_{i+1} and ensures the paths are continuous and there aren't abrupt or sudden changes in obstacle trajectories. Overall, the Hermite Cubic Spline *H(t)* for a segment between P_0 and P_1 with tangents T_0 and T_1 is given as:

$$H(t) = (2t^3 - 3t^2 + 1)P_0 + (t^3 - 2t^2 + t)T_0 + (-2t^3 + 3t^2)P_1 + (t^3 - t^2)T_1$$

3.3 Step 3: Spatiotemporal Mapping

3.3.1 Creating the Spatial Layer as Gridworld

Initialized with a 2D grid environment where each cell represents a specific area in the *x-y* plane and holds information about obstacle presence, uncertainties, etc. The resolution of this grid (i.e., the size of each cell) is chosen based on the agent's accuracy and the density of the obstacles.

3.3.2 Addition of the Temporal Layer

The 2D spatial grid is extended into a third dimension in *z*-direction which represents the layers of time. Each layer in z-direction corresponds to a specific time step. For instance, if predictions are made every 0.1 seconds, each layer would represent the state of the environment at that specific future time (*t=0, t=0.1, t=0.2*, ...). The maximum number of layers (or the depth of the grid) is determined by the prediction horizon (e.g., up to 2 seconds into the future).

3.3.3 Populating the Grid with Obstacle Position Predictions (Combining Spatial and Temporal Data with the Gridworld)

In the previous steps, we have already achieved these:

- Using the Sparse Graph Convolution Network (SGCN), predicted future positions of obstacles are generated. These predictions are made at discrete time intervals and provide spatial coordinates for each obstacle at each future time step.
- HCSI (Hermite Cubic Spline Interpolation) is applied to these

The authors are responsible for the content of the papers.

discrete points to smooth the trajectories and create continuous paths. The smoothed paths are then used to update the grid.

After continuous trajectories are generated with HCSI, each path segment is assigned to the corresponding grid map based on its predicted time frame. Each time step has its own grid layer, and the occupancy probabilities from smoothed paths are projected into these layers. For example, predictions for 1 second into the future would update the grid for *t=1*.

3.4 Step 4: Mapping Cell Occupancy Probability

For each time step, the predicted pedestrian positions are translated into probability distributions. This can be visualized as a heatmap where each grid cell is assigned a probability value indicating the likelihood of a pedestrian being present in that cell. The occupancy probability is higher near the predicted pedestrian positions and gradually decreases with distance. This distribution accounts for potential deviations in pedestrian movement, providing a buffer against sudden changes in direction.

Obstacle distribution O_t at time *t* in the gridworld coordinates is formulated as the sum of the O_{th}s for each obstacle with the equation below:

$$O_t = \sigma_1 \sum_{h \in H} O_{th} \quad (1)$$

$$O_{th} = \sigma_2 \sum_{s \in S} O_{pth}\,(x - x_{hs}, y - y_{hs}) \quad (2)$$

$$(x_{hs}, y_{hs}) \sim P_{(t-1)h}$$

Table 1: Experimental Settings

Number of dynamic obstacles	2
Agent	Mobile Robot
Microcontroller	Raspberry Pi 3
Motion Capture	MotionAnalysis Eagle x 18, Parallax 28015
Agent Motion	Stop, Straight, Autonomous Navigation

3.5 Step 5: Creating a Cost Map for Path Planning 3

Each cell in the spatiotemporal grid is assigned a cost value based on its occupancy probability. Occupied cells are assigned the highest cost (avoided by the path planner), caution cells are assigned intermediate costs (areas with lower probability of obstacle presence) and free cells are assigned the lower cost (preferred by the path planner). Uncertainties are treated as no-go buffer zones. The cells are defined by three cost $c(l)$ types as shown below:

$$c(l) = \begin{cases} C_0 \; if \; Occupied\big(d \le r_p + r_b\big) \\ C_c \; if \; Caution\big(r_p + r_b \le \mathrm{d} \le \mathrm{r_p} + r_b + b\big) \\ C_f \; if \; Free\big(r_p + r_b + b \le d\big) \end{cases}$$

3.6 Step 6: Searching Through Spatiotemporal Layers

The agent searches through these spatiotemporal layers to find a path from its current position to the goal. STP4 considers the costs across both spatial and temporal dimensions, ensuring that paths avoid future obstacle-occupied spaces. The path planner evaluates paths that extend into the future, checking each time layer for potential obstacles. This enables the robot to preemptively avoid areas that will be occupied by obstacles.

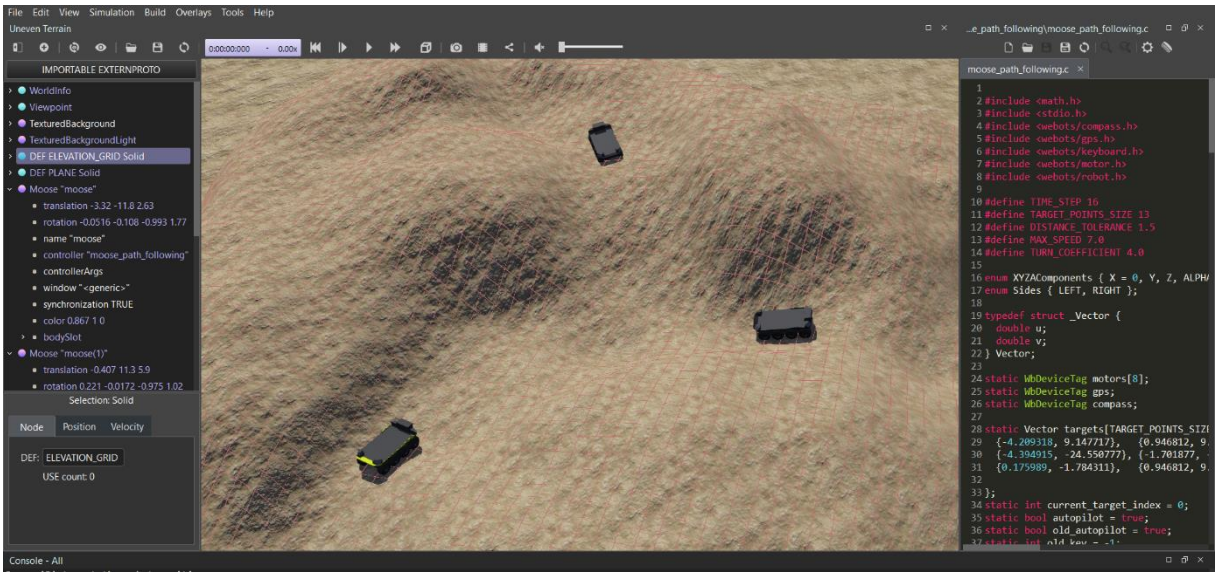


Figure 1: Simulation environment

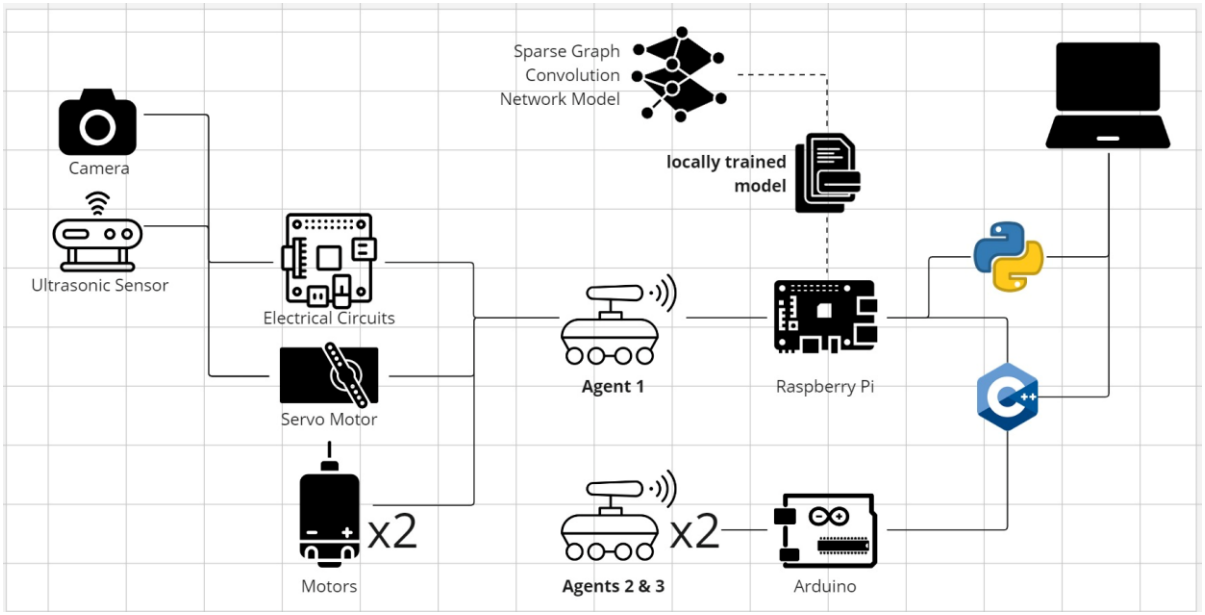


Figure 2: System Architecture

The prototyping process combines hardware and software components. Our system consists of three mobile robots and Raspberry Pi. The main algorithm runs on the Raspberry Pi on the main robot. The experimental settings are summarized in Table 1.

3.7 Concept development

The proposed experiment focuses on three mobile robots. The goal is to navigate the agent with a path plan which is updated in the presence of two dynamic constraints. The test was first conducted in a simulation environment and was designed in a mission area with three mobile robots, with one of them as the main agent and the other two as dynamic constraints. The mission area was confined to 65x65 fields in width and depth. Each individual field is 50x50 centimeter in size. Only the start and end goal were known to the agent.

3.8 Prototype development

The system architecture, as shown on Figure 2, integrates hardware components, microcontroller units, and software setup. A camera and ultrasonic sensor provide real-time environmental data, which is processed

through electrical circuits and servo motors to control the movement of multiple robotic agents. Raspberry Pi serves as the central processing unit, running locally trained models based on Sparse Graph Convolutional Networks to interpret sensor data and make path planning decisions using Python and C++ programming languages. Additional robotic agents are controlled via an Arduino microcontroller, demonstrating a coordinated multi-agent setup.

3.8.1 Hardware Assembly

The base structure of the robot includes a metallic chassis with mounting brackets for securing components. The main agent's chassis pieces were assembled, and Raspberry Pi was connected to the electrical circuit board. LiDAR, ultrasonic, and greyscale sensors were mounted on the robot. While the main agent is controller by Raspberry Pi, the secondary robots (that were used as dynamic constraints) were controlled by CyberPi and Arduino.

Raspberry Pi as central processor is mounted on the robot's chassis connected to the electrical circuit board that controls the electrical connections for motor, servo, camera, and so on. Raspberry Pi also handles sensor data processing and control signals for navigation and motion. Two direct current (DC) motors produce mechanical force, while three servo motors control rotation and angular movement, and control of the camera. Servos provide the robot with the ability to adjust its sensors dynamically based on environmental feedback.

The microcontroller board must be connected to a power supply unit, in the project a battery pack was used. The necessary circuitry includes various connections for interfacing sensors and actuators. Mounted at the front of the robot, the camera is used for visual input, capturing images or video for object detection, mapping and localization tasks. The camera is crucial for interpreting the environment and making informed path-planning decisions. The ultrasonic sensor is positioned beneath the camera, and is used for distance measurement, providing real-time feedback about obstacles or changes in the environment. It sends out ultrasonic waves and measures the time taken for the echo to return, allowing the robot to detect nearby objects and avoid collisions.

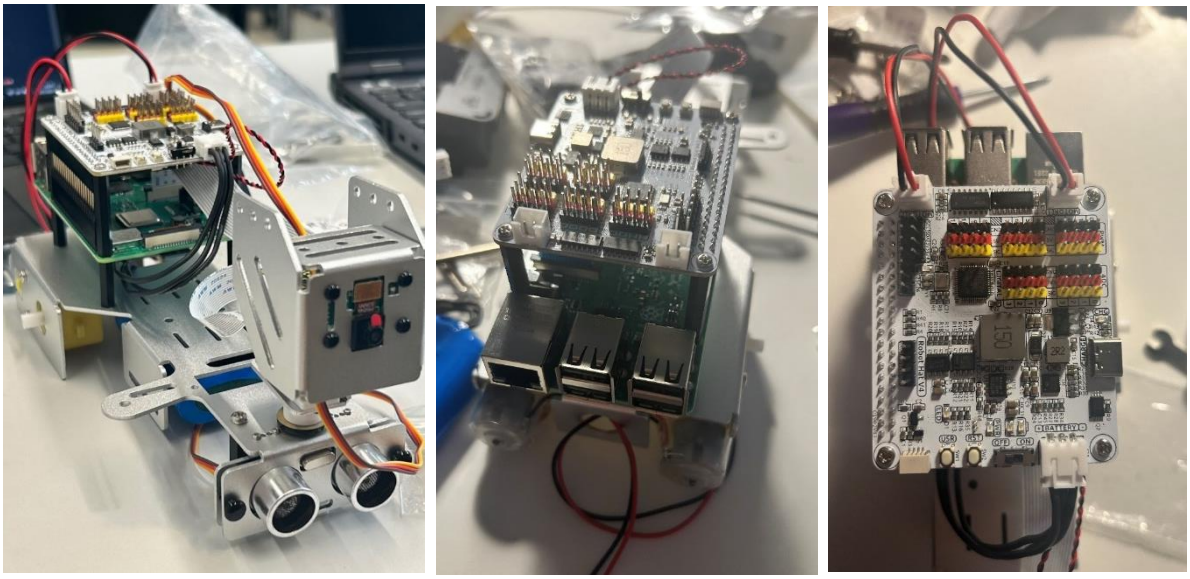


Figure 3: Robot chassis, camera, ultrasonic sensor, raspberry pi connection, electrical board connections

The authors are responsible for the content of the papers.



Figure 4: Raspberry Pi controls the primary robot



Figure 5: CyberPi controls the secondary robots

After installing the Raspbian OS on the Raspberry Pi, the camera module is connected to the Raspberry Pi's camera interface port. The ultrasonic sensor is attached to the GPIO pins, ensuring that the power, ground, and data lines are correctly wired. The Raspberry Pi will handle the main computational tasks, including image processing and running the trained model.

Arduino was connected to the Raspberry Pi via a USB connection for serial communication. Servo motors were wired to the Arduino's PWM-capable pins for control. DC motors were connected to motor drivers, which are then interfaced with the Arduino for movement control of the agents with motor control signals.

3.8.2 Software Setup

On the Raspberry Pi, install the necessary Python libraries as given in Table 3 below using pip, such as OpenCV for image processing, NumPy for data handling, and PyTorch for running the machine learning model.

3.9 Prototype demonstration

The test concept involves demonstrating the interaction between two robotic agents, Agent-1 and Agent-2, as they navigate towards a designated goal point. The objective is to evaluate how Agent-1 adjusts its speed and trajectory in response to the movement of Agent-2, which moves at a set motor speed. The test aims to observe whether Agent-1 can effectively use a spatiotemporal path planning algorithm to predict and avoid potential collisions based on the probability distribution of displacement. By analyzing the behavior of the agents under different conditions, the experiment seeks to refine and validate the effectiveness of the path planning and collision avoidance strategies employed. The experiment aimed to improve real-time responsiveness and ensure more accurate trajectory predictions.

4. Scientific results and discussion

4.1 Results and discussion

In the initial test, Agent-2 is put in motor motion of 60RPM. As demonstrated in Figure-6 that, it was observed that Agent-1 attempted to adjust its speed based on the probability distribution of displacement, however, did not manage to bridge the distance in time, and collided with Agent-2.

After adjustments in the algorithm and the mathematical calculations, as seen in Figure-7, the speed adjustment of Agent-1 was improved and the robot bridged the distance between start and terminal points, resulting the test positively.

Figure-10 shows the comparison of motion time and the kinematic profiles between regular motion and path planner motion. The top row shows the position over time, with the path planner achieving reduced motion time to 1.63s compared to regular motion of 2.57s.

Table 2: Hardware Components

Component	Amount
Robot Chassis Pieces	14
LiDAR Sensor	1
Ultrasonic Sensor	1
Greyscale Sensor	1
Raspberry Pi	1
Arduino	2
Servo Motors	3
DC Motors	2
Electrical Circuit	1
Wires	

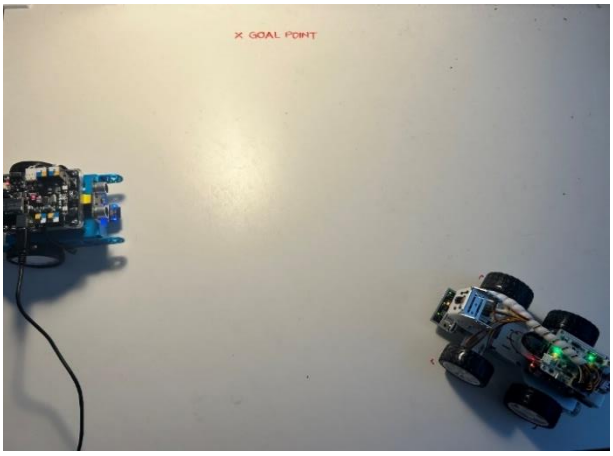


Figure 6: Testing (Agent-2 left, Agent -1 right)

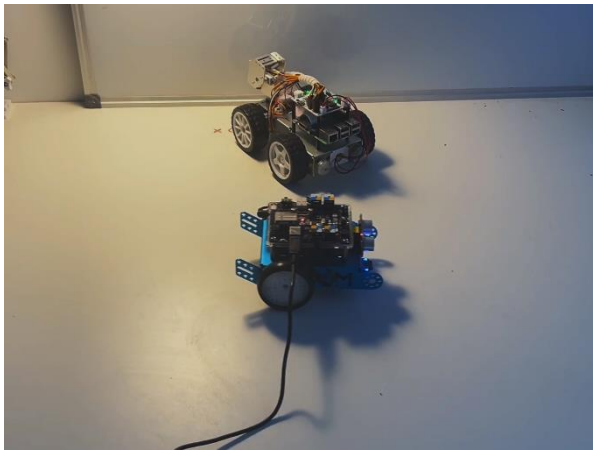


Figure 7: Positive result of the test

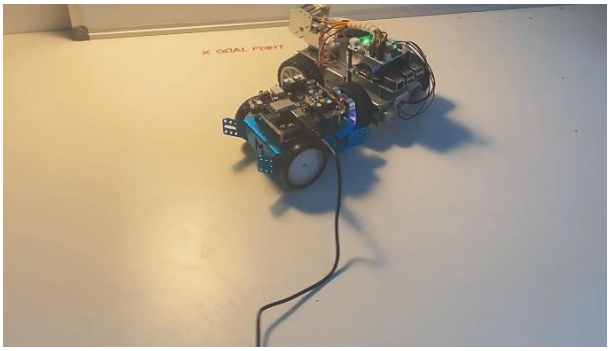


Figure 8: Due to faults in probability distribution, Agent-1's displacement was cut off before completion due to incorrect increase in its velocity and collision with Agent-2

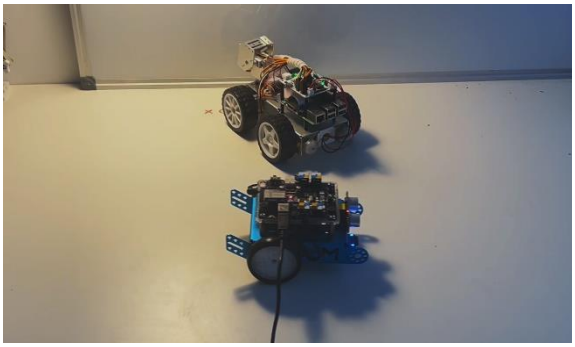


Figure 9: The probability distribution was corrected and Agent-1 managed to cross the distance in time to avoid collision with Agent-2

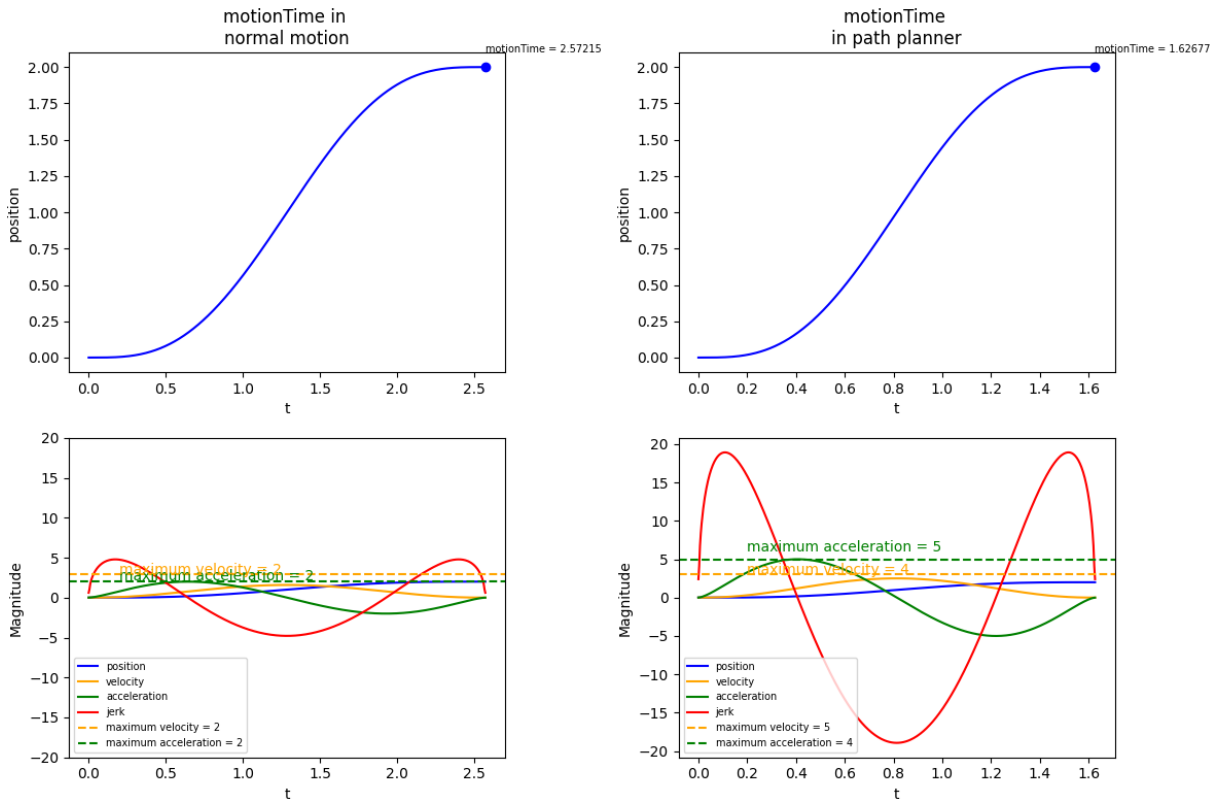


Figure 10: Comparison of motion time(top row) and kinematic profiles(bottom row) between regular motion(left side) and path planner motion(right side)

The bottom row displays position, velocity, acceleration, and jerk, indicating that the path planner allows higher velocity and acceleration limits, leading to faster completion. The results of the experiment show the difference between regular motion and the motion improved with the path planner. Over the course of 20 trials, the spatiotemporal path planning model achieved a reduction in motion time by 36.754% as calculated so:

$$Motion\ Time\ Improvement = \frac{2.57215 - 1.62677}{2.57215} \times 100$$

$$\frac{0.94538}{2.57215} \times 100 = 36.754466$$

In the given scenario and the specific positions of the agent and the constraint, the spatiotemporal path planned motion reached the terminal position in a shorter time frame. This indicates that the path planner can significantly reduce motion time by permitting higher velocity and acceleration within specified constraints, thereby optimizing the overall efficiency of the motion while still adhering to safety and performance limits.

The finding shows that optimizing path planning with the spatiotemporal method improves navigation efficiency and can adjust to real-time changes in dynamic constraint trajectories.

The application of path planning in environments with high dynamicity, such as spaces with non-stationary constraints or other robots, enable a mobile robot to navigate smoothly in dense crowds, avoiding abrupt movements and optimizing travel paths. our approach demonstrated that allowing higher velocity and acceleration constraints through path planning can lead to significant time savings without compromising safety or stability. These findings have direct implications for autonomous systems in urban mobility, robotics, and other applications where efficient and reliable navigation is critical.

The research supports that broader application of spatiotemporal path planning techniques, particularly in scenarios anticipating future positions of moving entities (e.g., public spaces, various active robots in the environment, or other forms of dynamic constraints) were helpful for optimizing

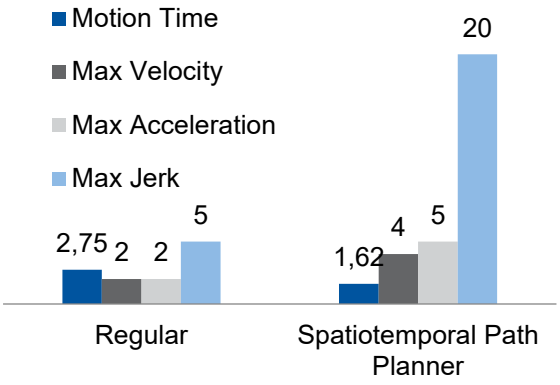


Figure 11 Comparison of motion time and kinematic profiles

navigation paths. By combining predictive models with path planning, as demonstrated in both studies, autonomous systems can achieve higher efficiency, reduce potential collision risks, and maintain smooth operation in complex environments. Future research may further explore integrating these methods to handle even more diverse and unpredictable scenarios, enhancing the robustness and adaptability of navigation systems.

5. Discussion

The spatiotemporal path planning project demonstrated the effectiveness of integrating advanced sensing, real-time data processing, and adaptive control strategies in autonomous robotic navigation. Through a series of tests and experiments involving multiple agents, the project successfully showcased the potential of combining hardware components like cameras, ultrasonic sensors, Raspberry Pi, and Arduino microcontrollers with sophisticated algorithms to achieve reliable and efficient navigation in dynamic environments.

The primary research question addressed in this project was: *Can* spatiotemporal *path planning techniques significantly improve the efficiency and safety of autonomous navigation in dynamic environments?* Based on the experimental results and analysis, the answer is positive. The use of spatiotemporal path planning, incorporating real-time adjustments to speed and trajectory based on environmental changes, led to a notable reduction in motion time and improved collision avoidance capabilities. These improvements confirm that spatiotemporal path planning can optimize the movement of autonomous agents, enhancing both efficiency and safety.

One of the critical findings of this project was the significant improvement in motion time and collision avoidance achieved by using a path planning algorithm that leverages spatiotemporal data. The comparison between normal motion and path-planned motion revealed that incorporating higher velocity and acceleration constraints led to a reduction in motion time by approximately 36.75%. These results validate the effectiveness of real-time path planning and prediction models in optimizing the performance of autonomous systems.

The project also highlighted the importance of fine-tuning the mathematical models and algorithms that control agent behavior. Initial tests showed that, without precise adjustments, Agent-1 struggled to avoid collisions with Agent-2, underscoring the necessity for accurate prediction and timely speed adjustments. Subsequent refinements to the algorithm improved the system's ability to predict movements based on the probability distribution of displacement, allowing Agent-1 to adjust its speed more effectively and avoid collisions. This iterative approach to algorithm development was crucial in enhancing the system's robustness and reliability.

Spatiotemporal path planning techniques could be applied in urban settings for autonomous vehicles, where predicting the movements of pedestrians and other vehicles is critical for safety and efficiency. In industrial automation, such systems could optimize the movement of robotic arms or mobile robots in manufacturing environments, reducing downtime and improving productivity.

Despite the successes, the project also faced several challenges. One primary challenge was the accurate prediction of dynamic obstacles, especially in scenarios involving sudden changes in speed or direction. While the improved algorithm showed better performance, there is still room for enhancing the prediction accuracy, particularly in highly dynamic or unpredictable environments. Future work could focus on integrating machine learning models that can learn from past experiences and adapt to new situations more effectively. Techniques such as deep reinforcement learning and neural

network-based prediction models could provide more robust solutions.

Recommendations for Future Work

Future research should focus on the following areas to enhance the capabilities and applicability of spatiotemporal path planning:

1. **Integration of Machine Learning Models**: Incorporating machine learning techniques, such as deep reinforcement learning or neural networks, could improve the system's ability to predict and respond to complex, dynamic environments. These models can learn from past experiences and improve their accuracy over time.
2. **Scalability**: Expanding the system to handle multiple agents in more complex environments is essential for practical applications. Future work should explore the coordination of multiple autonomous agents, ensuring they can navigate efficiently without collisions in crowded or unpredictable settings.
3. **Real-world Testing**: While the current project focused on controlled experimental setups, future research should include testing in real-world scenarios, such as urban environments with pedestrians and vehicles. This would provide valuable insights into the system's robustness and adaptability under real-world conditions.
4. **Improvement of Sensor Integration**: Enhancing the quality and integration of sensors, such as using more advanced LiDAR or depth cameras, can provide better environmental perception, improving the accuracy of path planning and collision avoidance.
5. **Long-term Adaptation and Learning**: Developing algorithms that allow the system to adapt and learn over the long term would be beneficial. This could involve the continuous updating of path planning strategies based on the evolving behavior of obstacles and agents within the environment.

Another area for future research involves scaling the system to handle more complex environments and multiple agents. As demonstrated in the research by Sato et al. (2023), handling dense crowds requires sophisticated coordination among multiple agents. Extending the current project to include more agents and more complex interactions could provide valuable insights into the scalability and flexibility of the proposed path planning techniques.

6. Conclusion

In conclusion, this project demonstrated the potential of spatiotemporal path planning in enhancing the performance and safety of autonomous robotic systems. By integrating real-time sensing, adaptive control, and predictive modeling, the system was able to navigate dynamically changing environments effectively. The findings provide a solid foundation for further research and development in autonomous navigation, with the potential to impact various industries and applications. Continued advancements in this field will likely lead to more intelligent, efficient, and safe autonomous systems, capable of operating in increasingly complex and dynamic real-world scenarios.

The knowledge gained from this project lays a foundation for further exploration and development in autonomous navigation. By addressing the identified areas for improvement and expanding the scope of testing, future research can build upon these findings to create more intelligent, responsive, and adaptable autonomous systems. As the technology continues to advance, the potential for safe and efficient autonomous navigation in increasingly complex real-world scenarios will become a reality, contributing to advancements in automation and the broader application of autonomous technologies.

The authors are responsible for the content of the papers.

7. References

[1] U. Orozco-Rosas, K. Picos, J. J. Pantrigo, A. S. Montemayor, and A. Cuesta-Infante, "Mobile Robot Path Planning Using a QAPF Learning Algorithm for Known and Unknown Environments," *IEEE Access*, vol. 10, pp. 84648–84663, 2022, doi: 10.1109/access.2022.3197628.

[2] I. Maurovic, M. Seder, K. Lenac, and I. Petrovic, "Path Planning for Active SLAM Based on the D* Algorithm With Negative Edge Weights," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 8, pp. 1321–1331, 2018, doi: 10.1109/tsmc.2017.2668603.

[3] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997, doi: 10.1109/100.580977.

[4] Y. Sato, Y. Sasaki, and H. Takemura, "STP4: spatio temporal path planning based on pedestrian trajectory prediction in dense crowds," *PeerJ. Computer science*, early access. doi: 10.7717/peerj-cs.1641.

[5] X. Deng, R. Li, L. Zhao, K. Wang, and X. Gui, "Multi-obstacle path planning and optimization for mobile robot," *Expert Systems with Applications*, vol. 183, p. 115445, 2021, doi: 10.1016/j.eswa.2021.115445.

[6] S. M. LaValle, *Planning Algorithms*. Cambridge University Press, 2006.

[7] Stachniss C. &Burgard W.(2002). An integrated approach to goal-directed obstacle avoidance under dynamic constraints for dynamic environments. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS).

[8] T. Kulvicius, S. Herzog, M. Tamosiunaite, and F. Worgotter, "Finding Optimal Paths Using Networks Without Learning-Unifying Classical Approaches," *IEEE transactions on neural networks and learning systems*, early access. doi: 10.1109/TNNLS.2021.3089023.

[9] L. Yang, L. Fu, P. Li, J. Mao, N. Guo, and L. Du, "LF-ACO: an effective formation path planning for multi-mobile robot," *Mathematical Biosciences and Engineering*, vol. 19, no. 1, pp. 225–252, 2022, doi: 10.3934/mbe.2022012.

[10] S. Daftry *et al.,* "MLNav: Learning to Safely Navigate on Martian Terrains," Sep. 2022. [Online]. Available: https://arxiv.org/pdf/2203.04563

[11] Liu, "Decentralized structural-rnn for robot crowd navigation with deep reinforcement learning," p. 3517, 2021.

[12] D. Dugas, J. Nieto, R. Siegwart, and J. J. Chung, "NavRep: Unsupervised Representations for Reinforcement Learning of Robot Navigation in Dynamic Human Environments," Aug. 2020. [Online]. Available: https://arxiv.org/pdf/2012.04406

[13] J. Santos, P. Costa, L. F. Rocha, A. P. Moreira, and G. Veiga, "Time enhanced A*: Towards the development of a new approach for Multi-Robot Coordination," in *2015 IEEE International Conference on Industrial Technology (ICIT)*, Seville, Spain, 2015, pp. 3314–3319, doi: 10.1109/ICIT.2015.7125589.

[14] S. K. Pattnaik, D. Mishra, and S. Panda, "A comparative study of meta-heuristics for local path planning of a mobile robot," *Engineering Optimization*, vol. 54, no. 1, pp. 134–152, 2021, doi: 10.1080/0305215x.2020.1858074.

[15] Y. Quan, H. Ouyang, C. Zhang, S. Li, and L.-Q. Gao, "Mobile Robot Dynamic Path Planning Based on Self-Adaptive Harmony Search Algorithm and Morphin Algorithm," *IEEE Access*, vol. 9, pp. 102758–102769, 2021, doi: 10.1109/access.2021.3098706.

[16] *Stentz A.(1995). The Focussed D*algorithm for real-time replanning. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.

[17] P. Hart, N. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968, doi: 10.1109/TSSC.1968.300136.

[18] K. Mangalam, Y. An, H. Girase, and J. Malik, "From Goals, Waypoints & Paths To Long Term Human Trajectory Forecasting," in 2021, doi: 10.1109/ICCV48922.2021.01495.

[19] Shi, "SGCN: sparse graph convolution network for pedestrian trajectory prediction," p. 8994, 2021.

The authors are responsible for the content of the papers.

XINGYING LI, ZONGHAI YANG

SmartRebar: Automated Reinforcement Monitoring System

ABSTRACT

The application of AI in construction sites has gained significant attention since the early 21st century, primarily focusing on non-productive functions such as helmet detection and material counting. This research explores the potential role of AI in the production process, specifically through the utilization of the YOLO (You Only Look Once) real-time detection model to identify and analyze construction reinforcements such as stirrups, longitudinal bars, and other elements tied to columns and beams.

The study investigates three primary research questions: (1) Can YOLO accurately identify various reinforcements in real-time? (2) Can the addition of segmentation and tracking layers enable YOLO to analyze area and count information effectively? (3) Can YOLO recognize the relative positions of stirrups or longitudinal bars in 3D space?

To address these questions, a dataset comprising 200 images was collected, 100 from actual construction sites and 100 from a laboratory model. The images were annotated using the "Segment Anything" format, specifically tagging different types of reinforcements, which were then converted into a YOLOv8-segment training format.

Through extensive training and fine-tuning, our research achieved several key outcomes. YOLO successfully identified all selected reinforcements and created bounding boxes encompassing the center points of all detected reinforcements in the images. The model also demonstrated the capability to recognize 3D information, determining whether reinforcements were positioned closer to the observer or camera. Additionally, YOLO could detect the absolute pixel differences in the y-coordinates between the center points of selected reinforcements and, using the width of the bounding box, identify the relative distances between them, such as the ratio of the diameters of adjacent reinforcements to the distance between them. Finally, the model was able to classify and count the different types of detected reinforcements.

Keywords: YOLOv8; Computer vision; 3D Recognition; Real-time Monitoring;

DOI: 10.18154/RWTH-2024-09257

Image created with Adobe Firefly

1. Introduction

The digitization process on construction sites is currently at a lower level compared to other industries. The review of construction components primarily involves quality inspections of rebar cages in beams and columns through random sampling on-site, with technicians manually measuring and compiling reports for construction operations.

- This approach lacks digital evidence storage, making the supervision quality of rebar cages highly dependent on the professional ethics of the technicians.
- The manual inspection process is time-consuming and labor-intensive, making it difficult to conduct comprehensive quality inspections across the entire site within limited timeframes.
- The absence of visualized rebar cage information means that future tasks, such as quality inspections, renovations, and demolitions of concrete structures, face additional testing challenges.

Simplifying and digitizing the manual supervision process in construction is crucial. Approaches like the use of point clouds and Building Information Modeling (BIM) are being explored, but both face similar challenges in terms of computational power and processing speed. In recent years, convolutional neural networks (CNN) have emerged as powerful tools for object recognition and information extraction from images. These models have demonstrated efficacy in automating tasks traditionally reliant on manual effort [1].

In this study, the performance of the recently popular real-time classification and detection model YOLOv8 is analyzed on a custom-built dataset to explore the potential of real-time detection AI models in the construction process. YOLOv8 is one of the computer vision modules in the artificial intelligence family. Computer vision enables computers to derive numeric information from digital images, videos, depth images, and 3D point clouds, process the information, and automate actions. Various aspects of applying computer vision to construction sites are relevant to this study, particularly object detection, recognition, labeling, learning processes, and 3D scene modeling. Although YOLOv8 has advantages in speed, it is slightly lacking in accuracy compared to other models [1] .

To explore the potential applications of YOLOv8 in the construction process, this study investigates three primary research questions:

1. Can YOLO accurately identify various reinforcements in real-time?
2. Can the addition of segmentation and tracking layers enable YOLO to analyze area and count information effectively?
3. Can YOLO recognize the relative positions of stirrups or longitudinal bars in 3D space?

2. Literature review

current famous usecase of computer vision in construction projects including using collected point clouds, images, and other sensor data to generate visualized models of construction site progress and then monitoring the construction progress. It shows great capabilities in visualization and abstruct information from construction site, automation, etc., but faces challenges such as high technical thresholds and complex data processing [2].

Figure 2 A result from Mask R-CNN.

2.1 Computer vision in point clouds

There are also studies in computer vision-based construction progress monitoring that using machine learning models to learn 2D image patches sampled per element and classify them into different material types. This enables the inference of different material types from point clouds, and progress tracking at the operational level based on reasoning about the observed frequencies of different material types [1].

2.2 AI applications for 2D images

The other example, Mask R-CNN models, used for image segmentation and information extraction, have found applications in detecting defects and segmenting surfaces [3].

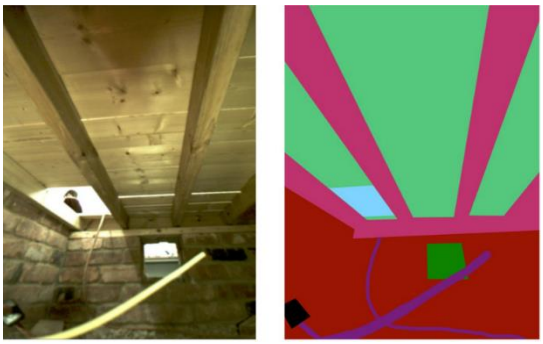


Figure 1: Encoding floorboards (light green), joists (magenta), walls (red), vents (dark green), cables (purple), and openings (cyan). Areas in black remained unclassified [3].

In the realm of construction sites, the utilization of the Mask R-CNN model is still in its nascent stages. However, there have been studies indicating its reliability in extracting various types of information from construction sites. Examples including automatically detect and segment defects on structure surfaces [4], identify regions such as floorboards, joists, air vents, pipework [3] and so on.

Here shows a result in the research conducted by Gary A. Atkinson and Wenhao Zhang in 2020, the Mask R-CNN model was found to reliably extract information from construction sites using only a small number of samples. They utilized the model to detect the underfloor scene, with results depicted in

2.3 AI applications for 3D images

There is also a lot of research in 3D imagery, such as ALPMS, which works by inputting images of a construction site and a four-dimensional Building Information Model (BIM) and outputting the percentage of completion of activities needed to update the project schedule. It creates as-built point clouds from the images, compares them with the planned BIM, generates orthographic views of the under-construction elements by projection transformation or Neural Radiation Field (NeRF), applies Deep Learning (AI)-based semantic segmentation for schedule inference, and estimates the percentage completion. Finally, the rich semantic information at the activity level is transferred to the as-built point cloud and BIM for 3D visualisation of the progress status. When applying the framework to two building construction projects, on average, ALPMS was able to report the activity percentage of completion with an average absolute error of less than 6%. Future research will focus on automating the mapping of project schedule activities and activity-level progress details for timely schedule updates and project completion dates [5]. Figure 2 shows the workflow.

To realise the application of Yolo model in the construction and supervision process, the movable camera position and ranging function are indispensable, however, most of the current research on Yolo is used for classification and counting, and the few ranging studies are based on the use of a specific space with a fixed camera, and the parameters are determined through on-site debugging. In this paper, the hoop spacing is dynamically obtained from the positioning frame of the hoops, which enables the use of a dynamic camera and the estimation of the relative position of the hoops.

3. Research method

The overall research two main experiments research process:

- Training model process: Analyze and find suitable training methods and hyperparameters to train and obtain suitable models.
- Output analysis process: Analyze the output of the model to get possible function based on the trained model.

The training model process mainly covers data collection, creation of the training set, and model tuning and training. This includes video collection, frame extraction and annotation, training set expansion, model training, and hyperparameter tuning, with the goal of finding a solution for enabling the model to accurately determine rebar types and their spatial locations.

Training model process:

Figure 3 Training model process

The authors are responsible for the content of the papers.



Figure 2: ALPMS workflow. Source: Pal et al. (2024) [5].

3.1 Data collection

The collected raw dataset is divided into two groups: 100 images collected from the construction site and 100 model images collected in the laboratory. The limited variety and small number of collected images could lead to a severe overfitting problem.

By altering the brightness, color temperature, and modifying the background, the original data set is expanded from 200 images to 1700 images and results in a better analysis output [6].

3.2 Datasets building

The creation of the Dataset is crucial. By labeling the same type of rebar located in the background as "back-ground," the model is enabled to recognize the 3D position and identify the rebar in the foreground.

3.3 Adjusting training methods

The creation of the Dataset is crucial. By labeling the same type of rebar located in the background as "back-ground," the model is enabled to recognize the 3D position and identify the rebar in the foreground. It has been shown that too few images and too many labelled columns and beams can lead to overfitting of the generated model. When a given training set is well labelled and there is enough features available for the model to learn, the model can accurately locate and classify the target by the pre-designed features, but when overfitting occurs, the model misses the pre-designed features and extracts other features to locate and classify the target. Overfitting results in a significant decrease in the reliability of the model's performance.

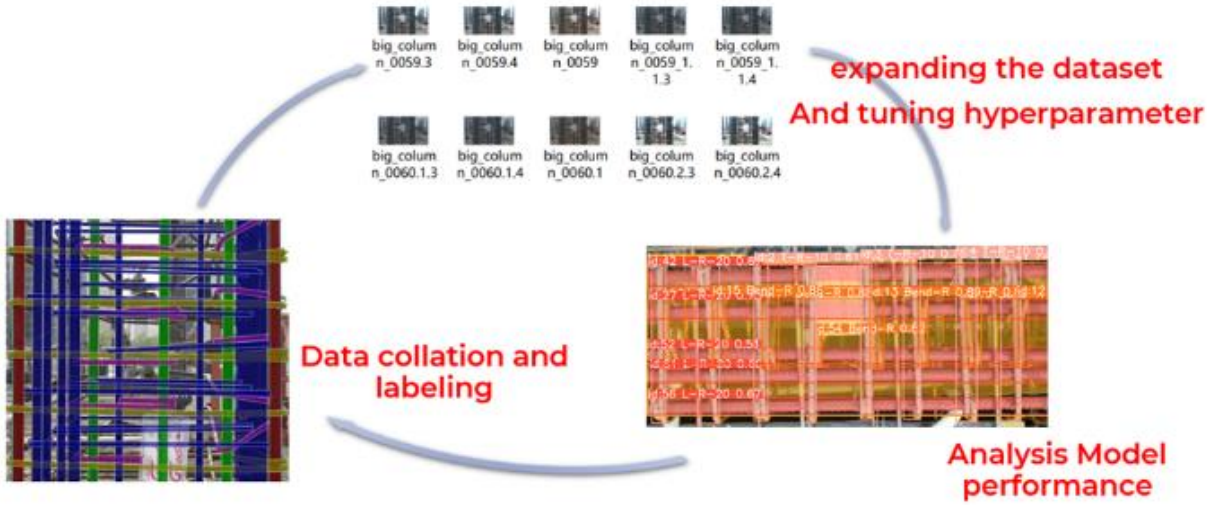


Figure 3: Training model process

3.4 Model performance analysis

For this model, the most difficult feature to obtain is the 3D position information of the longitudinal bars, because the longitudinal bars in the produced model are too distinguishable from each other, which leads to the difficulty of capturing the 3D information during the model training process, in addition to adjusting the correct hyperparameters, it is also necessary to periodically investigate the trained model to ensure that it does not lose the 3D information.

4. Prototype

4.1 Concept development

4.1.1 Concept of the Prototype

The project is designed to automate the monitoring of rebar placement in construction environments using sophisticated computer vision technologies. The fundamental concept revolves around leveraging the YOLOv8 model, a cutting-edge object detection algorithm, for real-time identification and classification of various rebar types, such as stirrups and longitudinal bars, within a three-dimensional spatial context. This system extends its capabilities by incorporating image segmentation, object tracking, and video processing technologies, which collectively enhance the detection accuracy and spatial awareness of the monitored construction elements. By integrating these advanced technologies, the project provides an intelligent and automated solution that not only improves construction quality and safety but also reduces the manual effort required to ensure precise rebar placement [7].

4.1.2 Tools and Platforms Used

The development of this prototype is based on several established open-source tools and platforms that are commonly used in computer vision and machine learning applications:

YOLOv8: YOLOv8 is the cornerstone of our system, providing real-time object detection capabilities that are essential for identifying rebar types and their positions within video feeds. The script my_detect.py demonstrates how YOLOv8 is utilized for continuous object tracking across video frames. The model is initialized with pre-trained weights optimized for detecting construction elements, and it processes each video frame to identify rebar, drawing bounding boxes around detected objects and classifying them by type. This setup allows for the precise localization of rebar in a 3D space, crucial for ensuring correct placement and spacing [8].

OpenCV: OpenCV is extensively used for handling video and image processing tasks within the prototype. In count_base.py, OpenCV manages the real-time input from

The authors are responsible for the content of the papers.

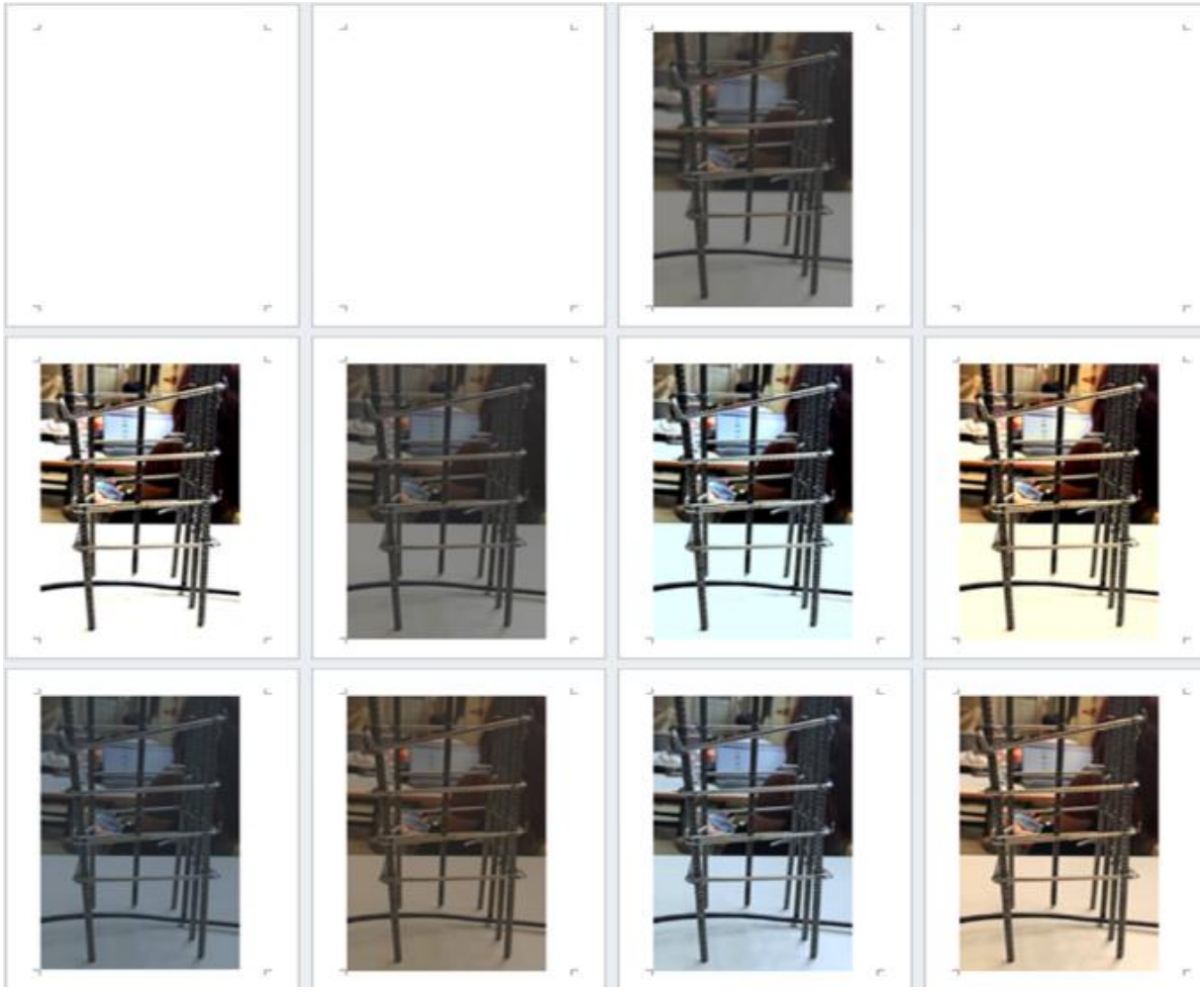


Figure 4: The process of data expansion is demonstrated, where the original image is expanded by adjusting brightness, color, and temperature to avoid model overfitting.

video feeds, captures frames, and converts them into formats suitable for processing by YOLOv8. OpenCV is also responsible for the visualization of detection results, drawing annotations and bounding boxes on frames, and displaying them in real time. This integration is vital for providing immediate feedback on rebar placement, helping to quickly identify any deviations from the desired configuration [9].

Segment Anything: This tool is employed to enhance the prototype's detection accuracy by performing image segmentation. By distinguishing rebar from the background, the system can more accurately identify and classify each element. The segmentation process is integrated into the detection pipeline to refine the regions of interest, improving the overall reliability of the rebar detection model. This step is critical for environments where the background might be complex or contain elements similar to rebar, thus preventing false positives [10].

BoT-SORT: Object tracking is another key component of the prototype, facilitated by BoT-SORT. This algorithm ensures that the detected rebar is continuously tracked across multiple frames, maintaining an accurate representation of their positions over time. The script my_detect.py incorporates BoT-SORT to link detected objects between frames, which is crucial for monitoring rebar that moves or changes appearance slightly due to camera angles or environmental conditions. This capability is essential for a comprehensive monitoring system that needs to ensure the correct placement and integrity of rebar throughout a construction project [11].

4.1.3 Research Process

The prototype design is visually represented through several diagrams and flowcharts that illustrate the system architecture and data processing pipeline. For instance, one

diagram may depict the flow of video data from the camera to the microcomputer, where it is processed by OpenCV, analyzed by YOLOv8, and then passed through segmentation and tracking modules. Another visualization could highlight the integration of different software components, showing how each tool (e.g., YOLOv8, OpenCV, BoT-SORT) interacts to deliver real-time detection and monitoring. This comprehensive visualization not only helps in understanding the system's operational flow but also aids in identifying potential areas for optimization or troubleshooting.

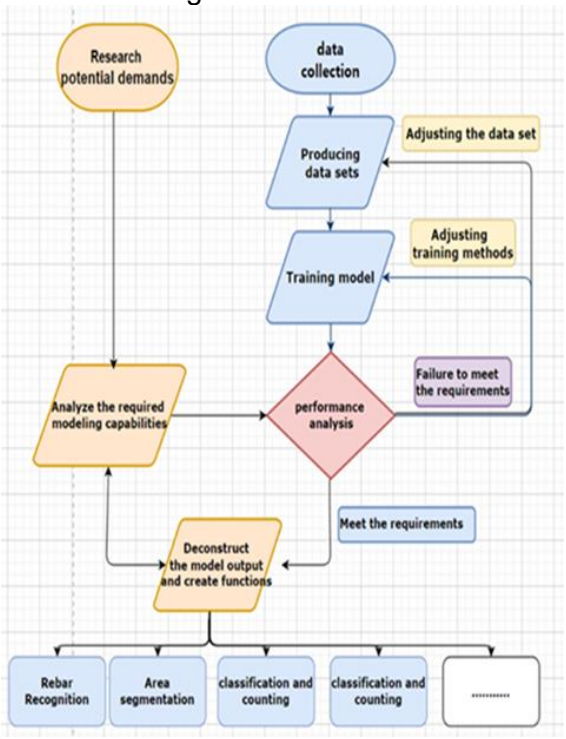


Figure 5: Overall research process

4.2 Prototype development

4.2.1 Realization of the Demonstration Sample

The prototype of "SmartRebar-Automated Reinforcement" is built to work with both simple and advanced tech. On the hardware side, it uses cameras and a microcomputer, like a Raspberry Pi or Arduino, to capture real-time video from construction sites. This setup is super flexible—you can use high-res cameras for a detailed view or just regular PC webcams, whether they're built-in or external. This means the system can easily fit into different setups, making it versatile for various construction needs.

For software, the system relies on YOLOv8, a powerful tool for detecting objects, which helps analyze video data instantly. Scripts like count_base.py and my_detect.py show how this model can be used to spot and track rebar in videos. With OpenCV, the system can process video from any camera, whether it's a high-end one on a construction site or a basic webcam plugged into a PC. This makes it adaptable and user-friendly, no matter what kind of equipment you have [12].

4.2.2 Hardware Configuration and Installation

For deploying YOLOv8 and OpenCV, a basic hardware setup is adequate. It is recommended to use high-definition cameras with at least 1080p resolution and 30 FPS frame rate for capturing detailed images. A microcomputer or workstation with a GPU such as the NVIDIA GTX 1050 or AMD Radeon RX 560, 15GB of RAM, and a mid-range processor like the Intel Core i5-10400 or AMD Ryzen 5 3600 is suitable for running YOLOv8 and OpenCV efficiently. For storage, an SSD or HDD with around 500GB is sufficient to manage video data and datasets. A standard internet connection with a bandwidth of at least 20 Mbps is enough for data transfer and remote monitoring. This setup provides a balanced and cost-effective solution for real-time object detection, ensuring smooth operation without the need for high-end components.

To set up the microcomputer for deploying YOLOv8 and OpenCV, start by installing the operating system, such as Ubuntu or Windows. If using an NVIDIA GPU, install the necessary GPU drivers and CUDA toolkit to ensure optimal performance. Next, create a Python virtual environment and install the required packages by running the following commands in the terminal.

Enter the command line in the prompt:

- conda activate (envspace_name)
- pip install ultralytics opencv-python
- sudo apt install ffmpeg

The authors are responsible for the content of the papers.

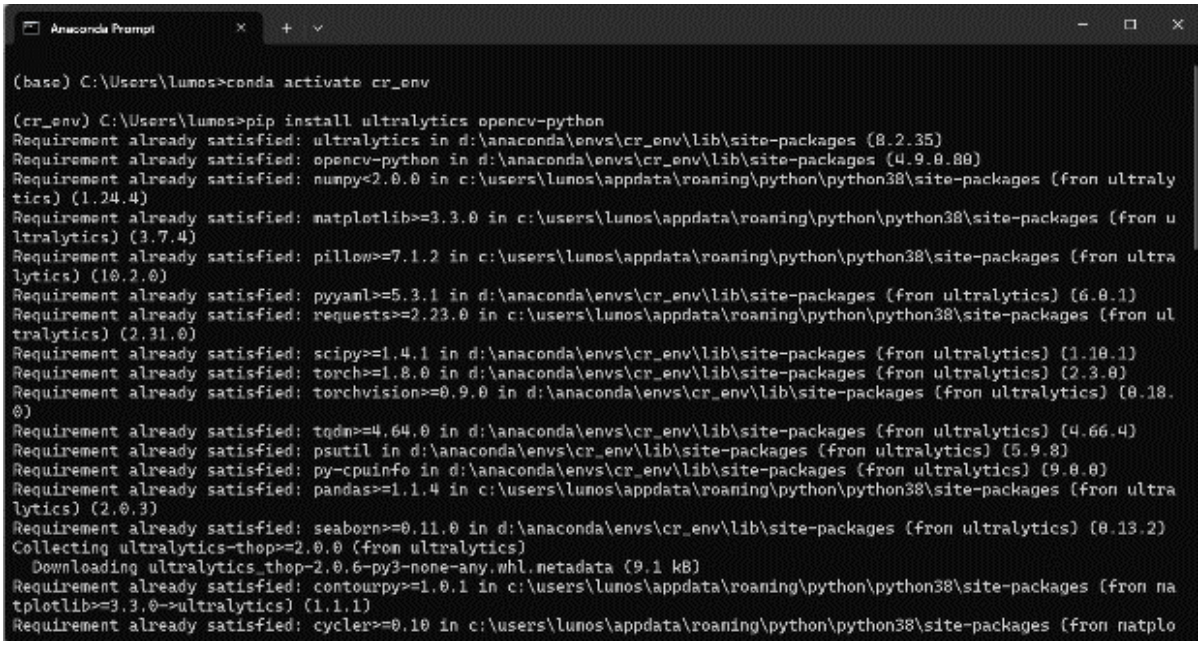


Figure 6: Enter the command in the prompt

```python
def adjust_darker(input_folder, output_suffix=".1", factor=0.8):
    # Get all file names in a folder
    for filename in os.listdir(input_folder):
        # Check if the file is a .jpg image and does not contain '.'
        if filename.lower().endswith('.jpg') and '.' not in filename[:-4]:          # Exclude the extension part for dot check
            # Opening an image file
            img_path = os.path.join(input_folder, filename)
            img = Image.open(img_path)
            random_number = round(random.uniform(-0.1, 0.1), 2)
            adjusted_factor = factor + random_number                                # introduce 20% randomness

            # Adjusting Brightness
            brightness_enhancer = ImageEnhance.Brightness(img)
            img_brightness_enhanced = brightness_enhancer.enhance(adjusted_factor)  # Reduce brightness by factor

            # Adjusting Contrast
            contrast_enhancer = ImageEnhance.Contrast(img_brightness_enhanced)
            img_enhanced = contrast_enhancer.enhance(adjusted_factor)               # Reduce contrast by factor

            # Constructing a new filename
            new_filename = f"{os.path.splitext(filename)[0]}{output_suffix}{os.path.splitext(filename)[1]}"
            new_img_path = os.path.join(input_folder, new_filename)
            print(f'adjust_darker: {filename} -> {new_filename}')

            # Saving the adjusted image
            img_enhanced.save(new_img_path)
```

Figure 8: Introducing Randomness

4.2.3 Model training

(1) Data Preparation Step

High-definition cameras are used to collect video data on construction sites. Then, the video frames are extracted using the FFmpeg tool in Anaconda Prompt. FFmpeg is an open-source multimedia framework that can handle audio and video data, including recording, conversion, and streaming operations.

The command used is as follows:

ffmpeg -i D:\0.8m_bar\25s_40s_beam.MP4 -ss 00:00:20 -t 00:00:20 -vf fps=4 output_dir/data-beam_%04d.jpg

- -i specifies the input file. D:\0.8m_bar\25s_40s_beam.MP4 is the path to the input video file.
- -ss specifies the starting point in the video. Here, 00:00:20 indicates that processing starts at the 20-second mark of the video.
- -t specifies the duration. Here, 00:00:20 means that processing will cover the next 20 seconds of video content starting from the specified start time.
- -vf indicates the use of a video filter.
- fps=4 sets the video frame rate to 4 frames per second. This means 4 frames will be extracted per second from the video.

```python
def adjust_warmer(input_folder, output_suffix=".4", factor=0.9):
    # Get all file names in a folder
    for filename in os.listdir(input_folder):
        # Check if the file is a .jpg image and does not contain '.'
        if filename.lower().endswith('.jpg') and '.3' not in filename[:-4]:          # Exclude the extension part for dot check
            # Opening an image file
            img_path = os.path.join(input_folder, filename)
            img = Image.open(img_path)
            random_number = round(random.uniform(-0.05, 0.05), 3)
            adjusted_factor = factor + random_number                                 # introduce 10% randomness
            # Adjusting Saturation
            color_enhancer = ImageEnhance.Color(img)
            img_color_enhanced = color_enhancer.enhance(factor)                      # Increase saturation by 20%

            # Adjusting Warmth by adjusting the image colors
            r, g, b = img_color_enhanced.split()
            r = r.point(lambda p: p * factor if p * factor <= 255 else 255)          # Increase red channel
            b = b.point(lambda p: p * (2 - factor) if p * (2 - factor) >= 0 else 0)  # Decrease blue channel
            img_warmer = Image.merge('RGB', (r, g, b))

            # Constructing a new filename
            new_filename = f"{os.path.splitext(filename)[0]}{output_suffix}{os.path.splitext(filename)[1]}"
            new_img_path = os.path.join(input_folder, new_filename)
            print(f'adjust_warmer: {filename} -> {new_filename}')

            # Saving the adjusted image
            img_warmer.save(new_img_path)
```

Figure 9 adjusting the brightness

• output_dir/ specifies the output directory.

• data-beam_%04d.jpg is the naming format for the output images, where %04d is a placeholder that indicates a 4-digit number sequence, resulting in filenames like data-beam_0001.jpg, data-beam_0002.jpg, and so on.

Subsequently, in Anaconda Prompt, Labelme is used to annotate the collected images in the "segment anything" format. After annotation, JSON files containing the annotation information. are generated, such as data-beam_0001.json, data-beam_0002.json, and so on.

Next, some functions are defined to augment the dataset by adjusting the brightness and warmth of the images. For example, starting with the image data-beam_0001.jpg, the process involves the following steps:

(2) **Brightness Adjustment**: The brightness of the image is modified to create variations. This results in:
   - data-beam_0001.1.jpg (the original image with increased brightness)
   - data-beam_0001.2.jpg (the original image with decreased brightness)
   - data-beam_0001.1.2.jpg (the brightness-increased image further adjusted to decrease brightness)

- **Introducing Randomness**: During the transformation process, random values are introduced, which means that each of the augmented images will have unique brightness distributions [13].

Next, the warmth and coolness of the images are adjusted. This process is similar to the brightness adjustments, but to enhance the randomness of the dataset, some images marked with certain labels are randomly selected to skip this transformation. For example, during the process of cooling down the images, those with names containing '.3' are excluded from the transformation.

After adjusting the brightness, the number of images was increased to four times the original amount. Following the warmth and coolness adjustments, the number of images was expanded to approximately fifteen times. the original quantity. However, it is important to note that while the number of images has been increased, the JSON files have not bee n proportionally expanded.

The transformed images only changed in terms of brightness and color, and did not alter the content or position of the images (such as flipping or rotating). Therefore, all the augmented images essentially share the same coordinate information, meaning their JSON files are identical. As a result, all the

```python
base = r"D:\Desktop\data_column"
json_dir = os.path.join(base, "json")
labels_dir = os.path.join(base, "labels")

print('Looking for JSON files...')
if not os.path.exists(json_dir):
    print('Error: JSON files directory not found:', json_dir)
else:
    print('JSON files directory found.')

    if not os.path.exists(labels_dir):
        print('Creating labels directory...')
        os.makedirs(labels_dir)

    for name in findAllFile(json_dir):
        print('Creating TXT file for', name)
        with open(os.path.join(base, "json", f"{name}.json"), 'r') as jf:
            data = json.load(jf)
        height = data["imageHeight"]
        width = data["imageWidth"]

        s = ""
        for item in data["shapes"]:
            label = item["label"]
            if label not in label_dict.keys():
                label_dict[label] = label_num
                label_num += 1

            s += str(label_dict[label]) + " "

            points = item["points"]
            for point in points:
                s += str(point[0] / width) + " "
                s += str(point[1] / height) + " "
            s = s[:-1] + "\n"

        with open(os.path.join(labels_dir, f"{name}.txt"), 'w') as tf:
            tf.write(s)


    yaml_path = os.path.join(base, "dataset.yaml")
    with open(yaml_path, "w") as f:
        print('Creating YAML file...')
        f.write(f"path: {base}\n")
        f.write("train: images\n")
        f.write("val: images\n")
        f.write("test: \n\n")
        f.write("names:\n")
        for key, num in label_dict.items():
            f.write(f"  {num}: {key}\n")

    print("Process completed.")
```

Figure 10 main modifications of YAML file

JSON files were duplicated and pasted until the number matched the expanded image set [14]. Then, the names of the JSON files were batch-renamed to correspond one-to-one with the expanded image set, completing the dataset expansion process.

An important step that follows is converting the "segment anything" format JSON annotation files into the TXT format files used for training YOLO. After a series of investigations and attempts, a Python script was written to achieve this. The script performs the following tasks:

- **Read and Parse JSON Content**: The script first reads and parses the JSON file to obtain the image's width and height.
- **Process Annotations**: For each annotation (in shapes), it retrieves the label and checks if it is already mapped in label_dict. If not, a new label ID is assigned.
- **Write TXT Files**: The script then constructs the YOLO format label content, which includes the label ID and coordinates of the annotation points relative to the image dimensions. This content is written into a corresponding TXT file, which has the same name as the input JSON file but with a .txt extension.
- **Create YAML File**: Finally, the script creates a YAML file to describe the basic information of the dataset.

Subsequently, according to YOLO's required format, the YAML file used for training the model is modified based on the corresponding content in the newly generated YAML file. The main modifications include:

```python
from ultralytics import YOLO
import multiprocessing

def train_model(epochs, imgsz, workers, batch_size, pretrained, device=0,
                mosaic=False, box=10.5, cls=5.5,  patience=50, verbose= True,
                lr0=0.01, lrf=0.01, optimizer='auto', overlap_mask=True,
                val=True, plots=True,  degrees=5, flipud=0.2, scale=0.1):#
    # Load a model
    model = YOLO(r"D:\Desktop\ultralytics\ultralytics\runs\segment\train4\weights\best.pt")
    # Train the model
    model.train(data=r"D:\Desktop\ultralytics\ultralytics\cfg\datasets\coco128-seg.yaml",
                epochs=epochs,
                imgsz=imgsz,
                workers=workers,
                batch=batch_size,
                pretrained=pretrained,
                patience=patience,
                device=device,
                mosaic=False,
                lr0=lr0,
                lrf=lrf,
                optimizer=optimizer,
                overlap_mask=overlap_mask,
                val=val,
                degrees=degrees,
                flipud=flipud,
                scale=scale,

                # Add mosaic parameter to augmentations
                )

if __name__ == "__main__":
    multiprocessing.freeze_support()
    # Define hyperparameters grid
    imgsz_list = [760]
    pretrained_list = [True]
    optimizer_list = ['auto']  # Add optimizer options here

    # Grid search
    for imgsz in imgsz_list:
        for pretrained in pretrained_list:
            for optimizer in optimizer_list:
                train_model(epochs=300, imgsz=imgsz, workers=1, batch_size=2,
                            pretrained=pretrained, mosaic=False, device=0,
                            patience=25, optimizer=optimizer, overlap_mask=True,
                            val=True)
                print(f"Training model with imgsz={imgsz}, pretrained={pretrained}, optimizer={optimizer}")
```

Figure 11 labels output and file paths

- **Mapping Output Labels to Actual Names**: This involves defining the correspondence between model output labels such as '0', '1', '2', '3', etc., and their actual names, for example, '20MM Longitudinal Rebar' as 'L-R-20', '10MM Stirrups' as 'T-R-10', and background as 'B-R', etc.
- **File Paths for Training**: The paths to the training files are specified.

Note that there is a space after the colon in the mapping between the output number and the real name of the model under names, and that not adding a space will result in a program error.

(3) Model parameters select process

Next comes the process of selecting hyperparameters for training the model.

In our training setup, all required hyperparameters were first identified and then tested multiple times. The hyperparameters that were most challenging to evaluate were divided into three groups:

- **imgsz**: 460, 760, 1080
- **Pretrained**: True, False
- **Optimizer**: Auto, SGD, Adam, AdamW, etc.

After training on a small subset of the data, the most suitable combination was identified by comparing box loss and test results: **imgsz: 760**, **Pretrained: True**, **Optimizer: Auto**.

It is important to note that the **Optimizer: Auto** option overrides the initial settings for the learning rates **lr0** and **lrf**. Through multiple trials, it was found that **Optimizer: Auto** was consistently the best option.

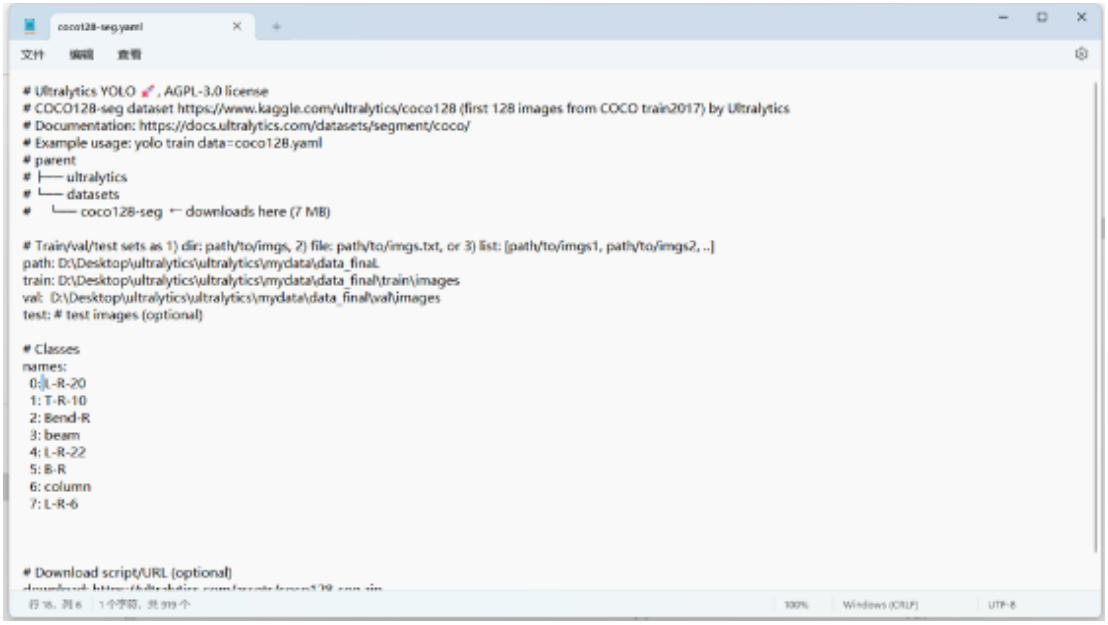


Figure 12: model's parameter

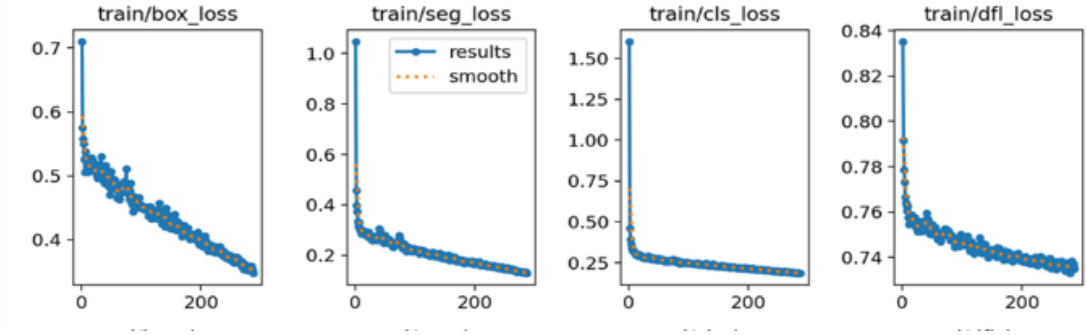


Figure 13: model performance

(4) Training supervisory process

During training, to mitigate the risk of overfitting, it is crucial to periodically check whether the trained model retains its ability to recognize 3D features [15]. If the model loses its ability to identify 3D features, training must be stopped promptly. In such cases, training should be resumed from the most recent suitable checkpoint or epoch where the model still demonstrated good performance in recognizing 3D features [16].

The code is available on Github: https://github.com/lumosconcerto3o/SmartRebarDetection.git

4.3 Prototype demonstration

4.3.1 How it works

- **Video Capture**: The system can use any camera—whether it's a high-quality site camera or just a basic webcam plugged into your PC. It captures live video feeds from the construction site to monitor the placement of rebar.
- **Detection and Tracking**: Using the YOLOv8 model, it analyzes the video frame-by-frame. It looks for different types of rebar, identifies them, and places a colored box around each detected piece [17]. This makes it easy to see exactly where each piece of rebar is on the screen. The ***my_detect.py*** script handles this process, running the YOLOv8 model to detect objects and track them across frames.
- **Real-Time Feedback**: The results are displayed instantly. You can watch the video feed with all the rebar labeled and tracked in real time, which helps in immediately identifying if something is misplaced or missing. The system even counts the number of rebar pieces and shows this data on the video output. For

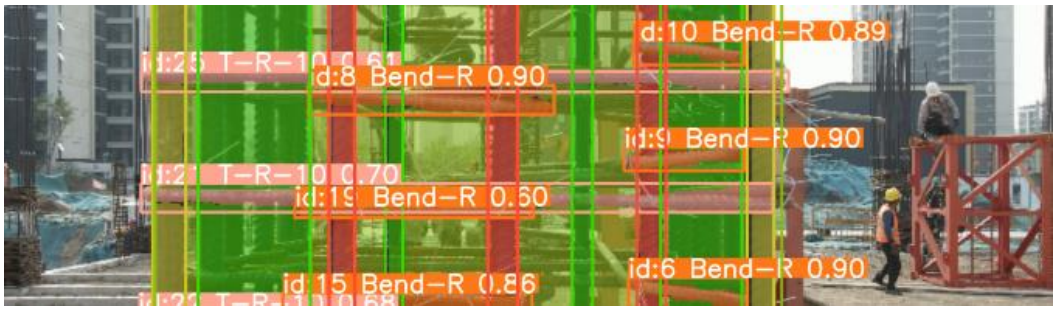


Figure 14: YOLOv8 model training for rebar detection, labeling positions and classifications

example, a small counter might pop up on the screen, showing the total number of detected rebar pieces and their types.

4.3.2 Outputs

Annotated Video: The output video will have boxes around detected rebars with labels showing what type they are. If you're using the system in a construction setting, you'll see these boxes moving around as workers handle the rebar, giving you a real-time update on placement accuracy [18].

Count and Classification: The system doesn't just show where the rebars are; it also counts them. You'll get a running tally of different types of rebar detected on the screen, making it easier to keep track of inventory and placement without manual counting.

Distance Measurement: In addition to detection and counting, the system can measure the distance between rebars, which is crucial for ensuring they are placed correctly according to construction guidelines. This data is overlaid on the video, so you can see exactly how far apart each piece is [19].

Alerts and Reports: If something is out of place—say, two pieces of rebar are too close together or one is missing—the system can send an alert. This helps the construction manager quickly address any issues, ensuring the site remains compliant with safety standards.

5. Scientific results and discussion

This study used the YOLOv8 model, a real-time object detection algorithm, to identify and classify different types of rebar used in construction. We focused on three main goals: (1) determining how accurately YOLO can detect reinforcements like stirrups and longitudinal bars in real time, (2) testing if adding segmentation and tracking features helps YOLO analyze the area and count of these reinforcements more effectively, and (3) assessing if YOLO can recognize the positions of these elements in 3D space [20].

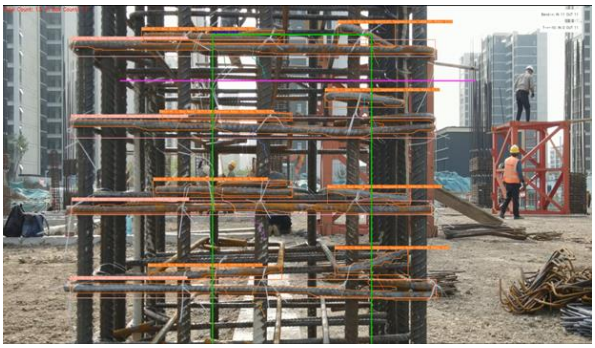


Figure 15 Correct and incorrect identification and distance

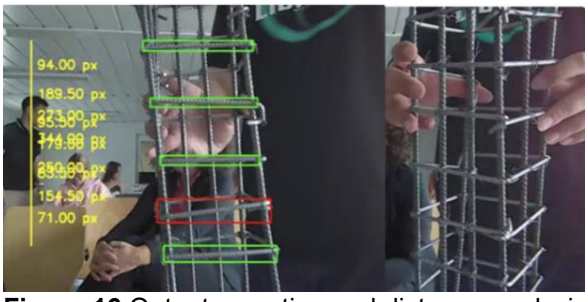


Figure 16 Output: counting and distance analysis



Figure 17 Output: counting and distance analysis

The authors are responsible for the content of the papers.

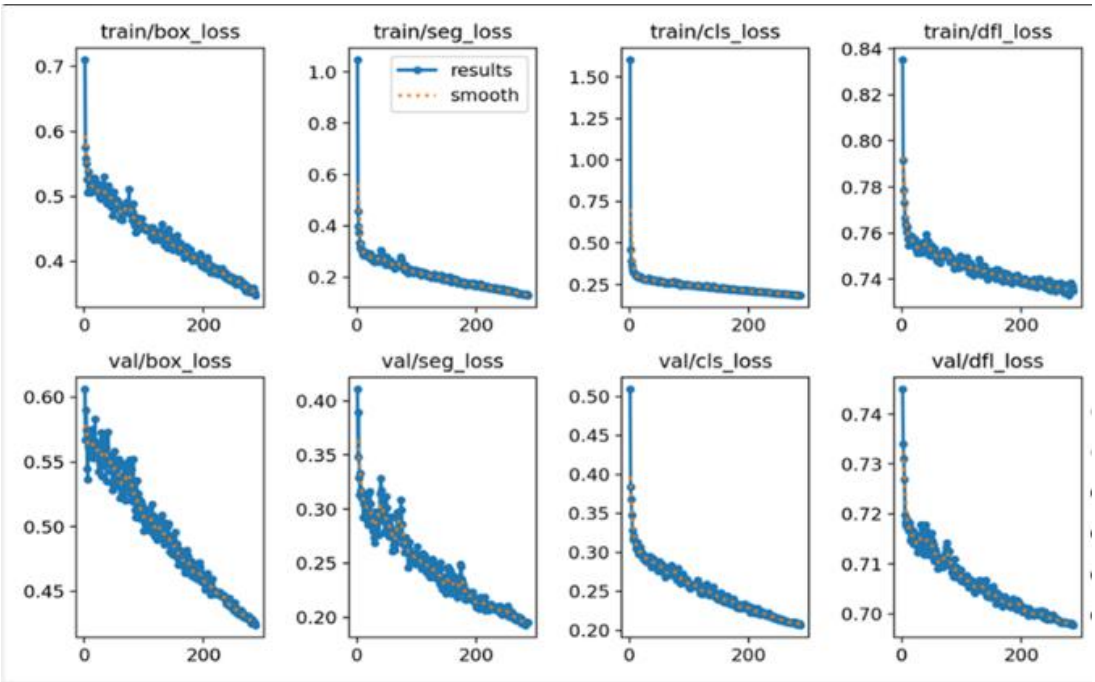


Figure 14: Model performance and comparison

Adapted, output that meets two goals, and a model that meets expectations is presented:

- **Accuracy in Detection and Classification:**

The results showed that the YOLOv8 model was effective in detecting and classifying all selected reinforcements, including stirrups and longitudinal bars. We initially used a dataset of 200 images but expanded it to 1700 images through data augmentation, such as changing brightness and color temperature and altering backgrounds [6]. This helped the model to achieve high precision in detecting these elements on both actual construction sites and in controlled lab settings. The model accurately placed bounding boxes around the center points of all detected reinforcements, showing its ability to correctly identify rebar types.

- **Analyzing Area and Count:**

With the addition of segmentation and tracking layers, the YOLOv8 model could not only identify reinforcements but also provide detailed information about their area and count. Segmentation improved the model's ability to distinguish rebar from the background, especially in complex environments where other objects might be present. The tracking feature ensured that the model could continuously detect rebar across multiple video frames, which is crucial for keeping an accurate count in dynamic construction scenarios [21].

- **Recognizing 3D Spatial Information:**

The model demonstrated the ability to understand 3D spatial information. It could determine whether reinforcements were closer to or further from the camera. This was done by calculating the pixel differences in the y-coordinates between the center points of detected reinforcements [22]. By using the width of the bounding boxes, the model could also estimate the distances between different reinforcements, like comparing the sizes of adjacent bars to the space between them.

6. Conclusion

- **Implications for Construction Automation:**

These findings suggest that using YOLOv8 in construction can greatly improve the automation and digitalization of inspection processes. The model's ability to detect, classify, and analyze rebar in real time can reduce the need for manual inspections, which are often slow and prone to errors. Adding segmentation and tracking features makes the model more reliable, especially in dynamic environments where objects are moving or partially hidden.

- **Challenges and Limitations:**

The model's performance depends on the quality of the input data and the precision of the annotations. In settings with poor lighting or a lot of visual clutter, the accuracy of detection might decrease. Moreover, running this model in real time requires substantial computing power, which may not be available on all construction sites.

- **Future Research Directions:**

Future research could focus on optimizing the YOLOv8 model to work on lower-end hardware, making it more accessible for a wider range of construction sites. Expanding the dataset to include more diverse construction scenarios and rebar types could improve the model's versatility. Additionally, integrating other AI models, like Mask R-CNN.

7. References

[1] Redmon, J. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

[2] Cheng, T., & Teizer, J. (2013). Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications. *Automation in construction*, 34, 3-15.

[3] Xu, Y., Li, D., Xie, Q., Wu, Q., & Wang, J. (2021). Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN. *Measuremen*t, 178, 109316.

[4] Tabernik, D., Šela, S., Skvarč, J., & Skočaj, D. (2020). Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing*, *31*(3), 759-776.

[5] Pal, A., Lin, J. J., Hsieh, S. H., & Golparvar-Fard, M. (2024). Activity-level construction progress monitoring through semantic segmentation of 3D-informed orthographic images. *Automation in Construction*, *157*, 105157.

[6] Gong, J., & Caldas, C. H. (2010). Computer vision-based video interpretation model for automated productivity analysis of construction operations. Journal of Computing in Civil Engineering, 24(3), 252-263.

[7] Li, Y., Lu, Y., & Chen, J. (2021). A deep learning approach for real-time rebar counting on the construction site based on YOLOv3 detector. *Automation in Construction*, *124*, 103602.

[8] Zhichao, H., Yi, W., Junping, W., Wanli, X., & Bilian, L. (2023). Improved lightweight rebar detection network based on YOLOv8s algorithm. *Advances in Computer, Signals and Systems*, *7*(10), 107-117.

[9] Frangez, V., Lloret-Fritschi, E., Taha, N., Gramazio, F., Kohler, M., & Wieser, A. (2021). Depth-camera-based rebar detection and digital reconstruction for robotic concrete spraying. *Construction Robotics*, 1-12.

[10] Ahmadi, M., Lonbar, A. G., Sharifi, A., Beris, A. T., Nouri, M., & Javidi, A. S. (2023). Application of segment anything model for civil infrastructure defect assessment. *arXiv preprint arXiv:2304.12600*.

[11] Aharon, N., Orfaig, R., & Bobrovsky, B. Z. (2022). BoT-SORT: Robust associations multi-pedestrian tracking. *arXiv preprint arXiv:2206.14651*.

[12] Zhang, J., Kang, X., Ni, H., & Ren, F. (2021). Surface defect detection of steel strips based on classification priority YOLOv3-dense network. *Ironmaking & Steelmaking*, *48*(5), 547-558.

[13] Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Albumentations: fast and flexible image augmentations. *Information*, *11*(2), 125.

[14] Jin, Y., Mishkin, D., Mishchuk, A., Matas, J., Fua, P., Yi, K. M., & Trulls, E. (2021). Image matching across wide baselines: From paper to practice. *International Journal of Computer Vision*, *129*(2), 517-547.

[15] Su, H., Qi, C. R., Li, Y., & Guibas, L. J. (2015). Render for cnn: Viewpoint estimation in images using cnns trained with rendered 3d model views. In *Proceedings of the IEEE international conference on computer vision* (pp. 2686-2694).

[16] Hou, Z., Qin, M., Sun, F., Ma, X., Yuan, K., Xu, Y., ... & Kung, S. Y. (2022). Chex: Channel exploration for cnn model compression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12287-12298).

[17] Khan, N., Zaidi, S. F. A., Yang, J., Park, C., & Lee, D. (2023). Construction work-stage-based rule compliance monitoring framework using computer vision (CV) technology. *Buildings*, *13*(8), 2093.

[18] Wang, S., Kim, M., Hae, H., Cao, M., & Kim, J. (2023). The development of a rebar-counting model for reinforced concrete columns: Using an unmanned aerial vehicle and deep-learning approach. *Journal of Construction Engineering and Management*, *149*(11), 04023111.

[19] Khow, Z. J., Tan, Y. F., Karim, H. A., & Rashid, H. A. A. (2024). Improved YOLOv8 Model for a comprehensive approach to object detection and distance estimation. *IEEE Access*.

[20] Jana, A. P., & Biswas, A. (2018, May). YOLO based Detection and Classification of Objects in video records. In *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)* (pp. 2448-2452). IEEE.

[21] Jiang, H., Sun, D., Jampani, V., Yang, M. H., Learned-Miller, E., & Kautz, J. (2018). Super slomo: High quality estimation of multiple intermediate frames for video interpolation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 9000-9008).

[22] Caicedo, J. C., & Lazebnik, S. (2015). Active object localization with deep reinforcement learning. In *Proceedings of the IEEE international conference on computer vision* (pp. 2488-2496).

The authors are responsible for the content of the papers.

ALPEREN BALCIN, MERVE YANIK

SmartMotion – IoT-Based Energy Efficiency Tracking Related to Kinetic Façade Movement *(corrected version)*

ABSTRACT

This paper investigates the development and optimization of a kinetic façade system created for the improvement of energy efficiency and indoor environmental quality via light-sensitive sensors and IoT-based control mechanisms. The Arduino microcontroller is used to tune the façade elements in real time to the ambient light intensity levels that are changing. Accordingly, the key research question that this paper attempts to address is: "How can adaptive building components use IoT enhance energy efficiency improve indoor comfort and reduce maintenance costs?". The primary objective is to show that adaptive components of the façade will save energy significantly by using daylight, thus in this way, natural light would be utilized and also the use of artificial equipment to control climate would also be minimized. It can be clearly seen from the findings of the experiment that the façade responds very good to the changing conditions of light, i.e., full opening of panels with high illumination and full closing with low light intensity, thus it maintains the desired temperature within the rooms and thus conserving energy, and most particularly, during the winter time. The study concludes that kinetic façades and IoT technology lead to better building performance and sustainable energy management, thus being applicable to future smart buildings. Further details and the source code of the project can be retrieved from GitLab at the following link: *https://git-ce.rwth-aachen.de/yankmrve/smartmotion_kinetic-facade-control.git*

Keywords: Kinetic Façade; Energy; IoT-Based Application; Smart House; Arduino Programming

DOI: 10.18154/RWTH-2026-07285

Image creat

1. Introduction

The global trend towards sustainability and energy efficiency in the built environment has been raising the bar for new advances in architectural design and building technologies. One of the recent innovations is in kinetic facades: dynamic envelopes of buildings are capable of responding to changing environmental conditions in real time with the aim of optimizing energy performance and indoor comfort. These include heat gain, daylight penetration, and natural ventilation control, which can make kinetic façades greatly energy-saving; in turn, they contribute to the support of sustainable architecture.

It's motivated mainly by the need to address the twin challenges of enhancing indoor comfort and reducing the energy footprint of buildings. With the current increase in the rate of urbanization, there is an increasing demand for buildings that are energy efficient without compromising the comfort of their human occupants. Traditional static façades are always limited in offering flexibility, which can be used to adapt to different environmental conditions during different periods of the day and even during different seasons. Kinetic façades offer great potential by changing their configuration to optimize energy use and indoor climate control in line with outdoor conditions.

This paper points out the fact that kinetic façades play an important role in indoor comfort improvement, keeping in mind energy saving. In particular, kinetics façades address issues related to temperature regulation, daylight optimization, and ventilation control. For kinetic façades, the components of a façade are changed in real time in view of externals such as sunlight, wind, and temperature. Using these mechanisms, ideal interior conditions are maintained by reducing reliance on artificial climate control systems like HVAC (Heating, Ventilation, and Air Conditioning). Therefore, interior spaces can remain comfortable for the occupants while real-time adaptability minimizes energy usage. For example, kinetic façades modulate the intake of solar heat that enters a building, which consequently reduces the need for air conditioning during hotter months. In colder periods, if more sunlight is allowed inside, then self-evidently the costs of heating are reduced. Moreover, due to changes in daylight levels, kinetic façades have the potential to reduce the use of artificial lighting, thus saving energy directly. Their efficiencies are further enhanced through smart integration of sensors and automated controls, hence making kinetic facades the advanced solution to achieve twin goals of sustainability with indoor comfort.

Motivated by the need to research the innovative design strategies that will be capable of operating advanced technologies, the paper presents light-sensoring systems and IoT-based monitoring systems within kinetic façade systems. By integrating the façade with light sensors, it will automatically adapt itself to available natural lighting intensity, therefore cutting down on the consumption of artificial light and saving energy. Real-time monitoring and control of performance are also possible with IoT, ensuring flexibility and control over the indoor environment by the users. This research will aim to develop a prototype for the kinetic façade system with integrated light sensors for IoT-based applications in order to improve energy efficiency and make indoor comfort better. The prototype aims to show how smart technologies could be integrated into building design for adaptive systems that will change according to the environment. This study, therefore, discusses the practical implementation of these technologies to contribute to the ongoing debate on sustainable architecture and ensure a viable solution toward energy-efficient building design. Of relevance is the growing body of literature that continues to highlight the potential of kinetic façades in reducing energy use and providing superior indoor comfort.

2. Literature review

The field of kinetic façades has garnered significant attention in recent years due to their potential to enhance building energy efficiency and optimize environmental control. These façades are designed to respond dynamically to external environmental conditions, such as sunlight and temperature, thus playing a crucial role in energy conservation.

2.1 Theories and Approaches in Kinetic Façades

The architecture of sustainability and kinetic façades requires buildings to dynamically adapt to changes in environmental conditions. They can alter their form or structure to relate to the effects of sunlight, temperature, and wind while assuring energy efficiency. As reported by Zanghirella et al., kinetic façades with advanced control systems can decrease the solar heat gain by 20% and, respectively, the cooling loads in buildings situated in hot climates [1]. Proceeding further, Yoon and Sun confirm that the use of kinetic shading systems is an effective way of reducing artificial lighting costs, which in combination with other benefits results in energy savings of about 18% for office buildings [2]. Another is developed for algorithmic optimization in "Designing a Kinetic Façade Using BB-BC Algorithm with the Focus on Enhancing Building Energy Efficiency": in this case, the form and behavior of kinetic façades are optimized using the Bang-Big Crunch (BB-BC) algorithm, where such systems have shown to consume 30% less energy compared to traditional static façades [3].

Performance optimization of kinetic façades often resorts to the direct use of computational tools and associated algorithms in their design. This sentiment is further supported by Zhang et al. in the 2019 study around the parametric design strategy, and it is stated that using simulation tools is very important to predict and enhance the energy performance of kinetic facades [4]. These studies highlight the importance of adaptive design to building-related energy-use reduction, particularly within urban areas where environmental conditions may vary.

2.2 Strengths and Weaknesses of Existing Research

One of the key strengths of existing research on kinetic façades is the detailed exploration of energy savings potential. Moloney et al., 2020 have shown structures with kinetic facades to save energy up to 30% compared to those with static facades mainly due to optimized solar gain and control of ventilation [5]. Also, Boud et al. (2018) further improve kinetic façades as far as energy savings are concerned and simultaneously contribute to occupant comfort by controlling the penetration of natural light and thermal conditions [6].

However, there are also notable weaknesses in the current body of research. For instance, while the energy-saving benefits are well-documented, there is limited research on the long-term durability and maintenance costs associated with kinetic façades. Pitts et al. (2021) suggest that the mechanical complexity of kinetic façades could lead to higher maintenance costs and potential reliability issues over time, which could offset some of the energy savings [7]. Further research is required to address these concerns by developing more robust kinetic systems.

2.3 Previous Kinetic Façades Implementations

The architectural landscape has been transforming substantially in recent years with innovative façade technologies that enhance aesthetic appeal and contribute to environmental sustainability. Of these, kinetic facades represent a very promising solution capable of dynamic reactions to changing environmental settings while optimizing energy use. Kinetic façades, with advanced materials and smart systems as integral features, bring a redefinition to the boundary of the traditional building envelope. At the same time, maintaining a balance between form and function, a good real-life example when one talks about kinetic facade is the Penumbra Shading System (Figure 1). This presents an innovative solution for window shading by the use of mechanical vents that control daylight effectively in three dimensions over the day. They, like indoor blinds, are designed to fold up and down along either side of the windows so as to correspond with the east-west direction of the sun for maximized efficiency. They fold upwards to serve as horizontal shades against the high afternoon sun; hence they provide not only vertical shading but also go further to portray a dynamic and adaptive approach to managing sunlight exposure on building façades.

The Kiefer Technic Showroom, Bad Gleichenberg, Austria (Figure 2), is one of



Figure 1: Kinetic façades move in three dimensions to screen sunlight from any angle [8]



Figure 2: The Kiefer Technic Showroom in Bad Gleichenberg [9]

these new dynamic architecture buildings where the façade involves 112 movable tiles that control sunlight and internal climate. Designed in 2007 by Giselbrecht & Partners, this building's façade opens out and folds back in accordance with the weather and eliminates the need for air-conditioning. Those 56 engines animate the movement of the perforated aluminum panels to create a dynamic façade that is in constant change, and serves both functional and aesthetic qualities.

Future Studies on Kinetic Façades

For all the immense potentials kinetic façades have in terms of enhancing the energy efficiency of buildings, there is yet very much that has to be discovered in a number of areas for it to be fully exploited. The future research will focus on the solutions in the durability of kinetic façades and their compatibility with advanced technologies, including its long-term environmental impact. One major area for further study is in the durability and maintenance of kinetic façades. Although current research focuses on the energy-saving benefits, the sustainability of these systems in the long term remains questionable. As opined by Pitts et al. (2021), kinetic façades are inherently mechanically complex and are hence bound to carry higher maintenance costs than conventional façades, which would undercut their sustainability over time. It is important to further research the development of more durable materials and innovative maintenance strategies that may extend the lifespan of these systems [7].

The integration of smart technologies with kinetic façades opens up another critical area for future research. It is true that real-time data analytics and predictive algorithms will provide much better adaptability to those systems. More advanced sensors and control systems could be integrated to allow kinetic façades to respond more preciously to the change of environment, for further energy saving and indoor environmental improvement, reducing the energy consumed for the building operation. Future studies should investigate the potential of such technologies to optimize the performance of the facade and how they can be integrated with the currently available systems for building management [10].

The environmental life cycle impact of kinetic façades also requires further research. Moloney et al. (2020) reveal that kinetic façades are capable of achieving a certain

degree of energy savings at the building operational level, even though the environmental costs at production and installation, and eventually at the end-of-life, have not yet been established. Future studies need to carry out a full Life Cycle Assessment (LCA) study in order to quantify the overall environmental impact and look for ways of minimizing it [11].

Table 1: Literature review summary and findings

No.	Authors	Title	Key Findings	Citation
[1]	Zanghirella, Babich, & Belleri	"Energy Performance of Kinetic Façades in Mediterranean Climates: A Case Study on Dynamic Sun-Shading Systems"	**20% reduction in solar heat gain**: Kinetic façades with advanced control systems can reduce solar heat gain by up to 20%, lowering cooling loads in hot climates.	Zanghirella et al. (2018)
[2]	Yoon and Sun	"The Impact of Kinetic Shading Systems on Daylighting and Energy Performance in Office Buildings"	**18% reduction in artificial lighting needs**: Implementation of a kinetic shading system can reduce artificial lighting needs by approximately 18% in office buildings.	Yoon, S., & Sun, S. (2020)
[3]	Babaee Tirkolaee, Goli, & Naderpour	"Designing a Kinetic Façade Using BB-BC Algorithm with a Focus on Enhancing Building Energy Efficiency"	**30% reduction in energy consumption**: Kinetic façades optimized with the BB-BC algorithm can reduce energy consumption by up to 30% compared to traditional static façades.	Babaee Tirkolaee et al. (2020)
[4]	Zhang, Du, & Chen	"Parametric Design Strategy for Kinetic Façades: Simulation and Performance Optimization"	**Importance of simulation tools**: Parametric design strategies and simulation tools are crucial for predicting and enhancing the energy performance of kinetic façades.	Zhang, Du, & Chen (2019)
[5]	Moloney, McLeod, & Phillips	"Energy Savings through Kinetic Façade Systems: A Comparative Study of Office Buildings in Different Climates"	**30% energy savings**: Buildings with kinetic façades can achieve up to 30% energy savings due to optimized control of solar gain and ventilation.	Moloney et al. (2020)
[6]	Boud, Rodi, & Benoy	"Occupant Comfort and Energy Efficiency in Buildings with Kinetic Façades: A Case Study"	**Improved occupant comfort**: Kinetic façades enhance energy efficiency and occupant comfort by regulating natural light and thermal conditions.	Boud et al. (2018)
[7]	Pitts & Mar	"Durability and Maintenance of Kinetic Façades: Challenges and Solutions"	**Concerns on durability and maintenance**: Mechanical complexity of kinetic façades could lead to higher maintenance costs and potential reliability issues over time.	Pitts & Mar (2021)
[8]	Frearson	"Kinetic louvres move in three dimensions to screen sunlight from any angle"	**Penumbra shading system**: Innovative window shading system with mechanical vents adjusting in three dimensions for optimized daylight control, providing both vertical and horizontal shading.	Frearson (2014)

[9]	Vinnitskaya	"Kiefer Technic Showroom / Ernst Giselbrecht + Partner"	**Dynamic architecture example**: Kiefer Technic Showroom with 112 movable tiles that adjust to control sunlight and internal climate, reducing the need for air conditioning.	Vinnitskaya (2010)
[10]	Lam & Cheng	"Smart Kinetic Façades: Enhancing Building Performance through Real-Time Data and Predictive Control"	**Integration of smart technologies**: Advanced sensors and control systems can enhance the adaptability and efficiency of kinetic façades.	Lam & Cheng (2020)
[11]	Moloney, McLeod, & Phillips	"Energy Savings through Kinetic Façade Systems: A Comparative Study of Office Buildings in Different Climates"	**Life cycle environmental impact**: While kinetic façades reduce operational energy use, the environmental costs of their production, installation, and disposal require further study.	Moloney et al. (2020)
[12]	Boud, Rodi, & Benoy	"Occupant Comfort and Energy Efficiency in Buildings with Kinetic Façades: A Case Study"	**Human interaction and user behavior**: More research is needed on how occupants interact with kinetic façades to ensure they are user-friendly and effective.	Boud, Rodi, & Benoy (2018)

Finally, Boud et al. (2018) draw attention to the general lack of empirical studies related to the human interaction dimension of kinetic façades. Consequently, what is required is the in-depth investigation of how occupants perceive and interact with such dynamic systems to provide very useful information for their design and operation. Much further research is needed in this area of studying user behavior in buildings with kinetic façades to ensure the effectiveness and user-friendliness of such systems [12].

By identifying these focus areas, future work can help overcome existing limitations of kinetic façades and enhance their contribution to sustainable architecture. Table 1 summarizes the state of the art by presenting the recent literature in the field.

3. Research methods

3.1 Research Design

3.1.1 Introduction to Research Design:

This study uses an experimental approach to carefully assess the performance and energy efficiency of kinetic façade systems that incorporate IoT sensors. The primary goal is to fine-tune the usability and energy-saving potential of these systems by closely monitoring how they respond to different environmental factors, particularly light levels. Although the current prototype uses acetate to mimic the flexibility of photovoltaic (PV) panels, it doesn't generate energy. Ultimately, the aim is to develop a system that includes real PV panels for energy production. The experimental method was selected because it allows for precise control and adjustment of key variables in a simulated environment, resulting in reliable and accurate data.

3.1.2 Rationale for Experimental Design:

The choice to adopt an experimental design is motivated by the need to clearly establish a cause-and-effect relationship between the independent variable, which is light intensity, and the dependent variables, including energy use, system responsiveness, and indoor comfort. This method enables the researchers to isolate specific factors, such as how the kinetic elements of the façade respond to changes in light levels and the potential energy savings from automation using sensors. By simulating real-world conditions within a controlled setting, the research aims to generate insights that are directly applicable to full-scale building projects.

3.1.3 Scope of the Research:

This research focuses on the development, implementation, and evaluation of a prototype kinetic façade system designed to optimize energy performance under varying light conditions. The prototype uses acetate material to represent the flexible nature of PV panels, with the expectation that future versions will incorporate actual PV panels for energy generation. The study centers on several key areas:

Energy Consumption:

The research looks into the energy required to operate the kinetic elements, such as moving panels with servo motors. Although the current prototype does not generate energy, integrating PV panels in future designs is a major goal for sustainable energy management.

Sensor-Driven Responsiveness:

The study examines how the façade system adjusts in real-time to data collected by IoT sensors, especially light-dependent resistors (LDRs). The sensors' ability to accurately track and react to changes in light intensity is essential for maximizing energy efficiency and enhancing indoor comfort.

System Optimization:

This research explores different strategies to improve the overall efficiency of the kinetic façade system. For example, it investigates the potential to reduce energy use by placing actuators in standby mode during low-activity periods, such as nighttime. The study thoroughly evaluates how various operating modes impact both energy savings and indoor comfort.

3.1.4 Justification of Research Design:

The justification for using an experimental research design lies in its ability to deliver concrete data on the operational efficiency and feasibility of kinetic façade systems in real-world scenarios. Conducting the study in a controlled laboratory environment allows for precise adjustments of environmental variables, ensuring that the results are both dependable and can be replicated. Furthermore, the combination of IoT technology with kinetic architecture introduces a novel approach to energy management within buildings, making this research particularly relevant to the fields of sustainable architecture and smart building design.

3.2 Data Collection Methods

3.2.1 Sensor Integration and Calibration

The integration and calibration of sensors are essential for the accurate and responsive operation of the kinetic façade system. The sensors chosen for this research were selected based on their precision in measuring environmental conditions, with a particular focus on light intensity, which is crucial for the functioning of the kinetic elements.

Sensor Selection:

The study primarily employs Light Dependent Resistors (LDRs) due to their high sensitivity to changes in light intensity. LDRs were chosen for their reliable performance in detecting ambient light levels and their

compatibility with the control systems of the prototype. Additionally, temperature sensors were included to monitor and manage the thermal conditions within the system, though the main emphasis remained on light intensity.

Sensor Placement:

The LDRs were strategically positioned across the façade, with one sensor on each panel and additional sensors placed between panels to ensure accurate light detection and comprehensive monitoring of light intensity. This arrangement was designed to capture both direct and diffused sunlight from various angles, providing a more thorough assessment of the environmental factors influencing the façade. The placement strategy was carefully planned to minimize shadow effects and ensure real-time detection of light variations.

Calibration Process:

Before the experiments commenced, all sensors underwent a detailed calibration process to ensure data accuracy. The LDRs were calibrated using a controlled light source with known intensity levels, enabling the adjustment of sensor readings to reflect actual light conditions. Calibration was carried out under different environmental conditions to account for potential variations in sensor sensitivity, which was crucial in minimizing measurement errors and improving the reliability of the data used in subsequent analyses.

Data Collection Precision:

The precision and reliability of the collected data can be further enhanced by applying noise filtering techniques and periodically recalibrating the sensors. Noise filtering algorithms can be used on the raw data to eliminate any false readings caused by transient light fluctuations or electronic interference. Additionally, regular recalibration helps to correct any drift in sensor sensitivity over time, ensuring that the data remains consistent throughout the experimental process.

3.3 Data Logging and Monitoring

To effectively evaluate the performance of the kinetic façade system, accurate data logging and real-time monitoring are essential. These processes are crucial for capturing consistent data and assessing system behavior under varying conditions.

Data Logging:

Initially, using an SD card for local data storage seemed like a good idea. However, switching to the Arduino IoT Cloud Remote app made things much smoother. This platform allowed for real-time monitoring and gave more flexibility in managing data. It also integrated seamlessly with the sensors and devices used in the project, ensuring continuous data logging throughout the experimental stages.

Real-Time Monitoring:

Monitoring in real-time was crucial. The Arduino IoT Cloud Remote app made it possible to track every change in light intensity, motor movement, and panel positioning as it happened. This immediate feedback ensured that the data collected was both accurate and reliable. The app's interface was designed to be user-friendly, giving a clear overview of the system’s status at any moment.

3.2.1. Experimental Conditions and Variables

The conditions and variables during testing were crucial in determining how the kinetic façade system performed under different scenarios. The experiments aimed to see how the system responded to changes in light, both indoors and outdoors.

Experimental Conditions:

Testing was done in both indoor and outdoor environments to get a complete picture of the system's performance. Indoors, tests were conducted under standard room conditions. Outdoors, the system was tested in sunny and cloudy weather. The lighting conditions included complete darkness, partial light, and

The authors are responsible for the content of the papers.

full sunlight. These scenarios helped assess how well the system adapts to various levels of light, which is important for optimizing its response to changes in the environment.

Controlled Variables:

To ensure that the results were consistent and accurate, several important variables were closely controlled:

- **Motor Power:** The force generated by the motors was adjusted to ensure that the panels opened and closed smoothly. Different motors were used to see how varying levels of power affected the system. For the project SG90 servo motor (x3) and MG996R servo motor (x1) were used for moving the kinetic façade. Two motors with different torque were used to see the effects of two different motors.
- **Distance Between Motor Arm and Panel:** The distance between the motor and the panel was kept the same throughout the experiments to ensure consistent movement.
- **Panel Opening Speed:** The speed at which the panels opened in response to light was carefully controlled. Different types of motors required different light sensitivity settings. In accordance with the usage of two different types of servo motors, different ranges of light scales were used. The LDR combined with the SD90 servo motor has maximum light sensibility of 850 as for the LDS combined with MG966R has maximum of 1023.
- **Panel Angle of Opening:** The angle at which the panels opened depending on the amount of light was critical for ensuring the system worked as intended.
- **Material Flexibility (Acetate):** The flexibility of the acetate material used for the panels was measured to simulate the behavior of more advanced materials like flexible PV panels.

Measurement Intervals:

Data was collected at different times during the testing phases, corresponding to the various lighting conditions. This approach provided a detailed look at how the system adjusted to changes in light.

3.4 Experimental Setup and Procedures

The experimental setup and procedures were carefully planned to evaluate how well the kinetic façade system performs under different light conditions. This section covers the mechanical and electronic components used and the methods applied to assess the system's functionality.

Motor Arm Movement and Panel Opening Mechanism

The performance of the kinetic façade system is directly influenced by the movement of the motor arm and the panel opening mechanism. These relationships are formalized as follows:

1.Motor Arm Displacement: The linear displacement of the servo motor arm (d_m) is directly related to the rotation angle (θ_m) of the motor and the radius (r) of the motor arm:

$$d_m = r.\theta_m$$

- This equation is essential for determining the extent of panel movement resulting from a specific rotational movement of the motor.

2. Torque and Force Relationship: The torque (τ) generated by the motor must be sufficient to exert the necessary force (F) to move the panel. The relationship is expressed as:

$$\tau = F.r$$

- This equation is critical for selecting appropriate servo motors (e.g., SG90, MG996R) to ensure that they generate enough torque to efficiently move the panels.

3. Energy and Power Consumption: The energy (E) consumed by the motor during

The authors are responsible for the content of the papers.

operation can be calculated using the torque and the rotation angle:

$$E = \tau \cdot \theta_m$$

- The power (P) consumed by the system is then given by:

$$P = \frac{E}{t} = \frac{\tau \cdot \theta_m}{t}$$

Light Intensity and Panel Opening

The relationship between light intensity (I), the motor's power output (P), and subsequently the panel's opening angle (θ_p) is crucial for understanding the system's response to varying environmental conditions. This relationship can be expressed in a simplified model as:

$$P = k_1 \cdot I$$

Where:

- P is the power output of the motor (measured in watts),
- I is the light intensity (measured in lux or another relevant unit),
- k_1 is a proportional constant that quantifies the relationship between light intensity and motor power. It indicates how responsive the motor is to changes in light intensity.

The motor's power output directly affects the opening angle (θ_p) of the kinetic façade panels. This relationship can be expressed as:

$$\theta_p = k_2 \cdot P = k_2 \cdot k_1 \cdot I$$

Where:

- θ_p is the panel's opening angle (measured in degrees),
- k_2 is another proportional constant that relates the motor's power output to the panel's opening angle.

These equations demonstrate that as light intensity increases, the motor generates more power, leading to a greater panel opening angle.

Application in the Experimental Setup

- **Servo Motors:** The SG90 and MG996R servo motors were selected based on their torque capacities, ensuring they provided sufficient force to move the panels as required.
- **Panel Movement:** The displacement and torque calculations ensured that the panels opened and closed within the desired angles in response to varying light intensities, controlled by the light sensors.

These procedures and calculations formed the foundation of the experimental setup, allowing for a precise and quantifiable analysis of how the kinetic façade system responds to different light conditions. The relationships defined between light intensity, motor power, and panel opening angle provide a solid basis for optimizing performance in real-world applications.

3.5 Data Analysis Methods

The data analysis in this study focused on applying the mathematical models established in the methodology to systematically evaluate the kinetic façade system's performance under varying light conditions. The primary objective was to quantify the relationships between light intensity, motor power output, and panel opening angle to assess the system's responsiveness and energy efficiency.

Application of Formulas: The formulas $P = k_1 \cdot I$ and $\theta_p = k_2 \cdot P$ were utilized to process the experimental data. These equations were applied to calculate the motor's power output at different levels of light intensity and determine the corresponding panel opening angles. The use of these formulas allowed for a structured analysis of how the system adjusts to environmental changes.

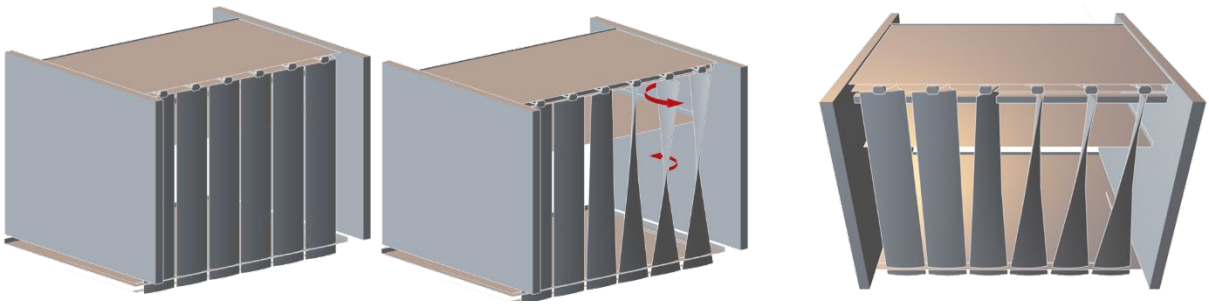


Figure 3: Prototype modeling views showing the panel layout and the red arrows indicating the opening direction.

Data Processing: The collected data were systematically processed to verify the theoretical models. Calculated values for motor power and panel angles were compared with the observed data to confirm the accuracy of the applied models. This comparison was crucial for validating the system's predicted behavior under varying light conditions.

Energy Consumption Analysis: Energy consumption was evaluated using the formula $E = \tau \cdot \theta_m$ to determine the energy required for panel movement under different light conditions. This analysis provided insights into the system's efficiency and identified potential areas for optimizing motor performance.

Statistical Methods: Descriptive statistics, such as mean and standard deviation, were employed to summarize the experimental data. Correlation analysis was conducted to explore the relationship between light intensity and motor performance, providing a quantitative assessment of the system's responsiveness.

Summary: The data analysis methods applied in this study provided a detailed understanding of the system's behavior under different environmental conditions. These methods laid the groundwork for further interpretation and discussion of the results, which will be addressed in the subsequent sections.

4. Prototype

In this project, a smart kinetic façade system was developed to optimize energy efficiency by adapting to varying weather conditions and the specific climate characteristics of the building's location. The prototype emphasizes a light-responsive façade design, wherein the façade elements automatically adjust based on light intensity, thereby enhancing the building's overall energy conservation.

The core concept involves the use of kinetic façade elements, constructed from acetate sheets, which respond dynamically to ambient light levels detected by integrated sensors. This adaptive system enables the façade to regulate sunlight exposure, contributing significantly to indoor energy management.

Concept development

The core concept focuses on designing a kinetic façade system that can automatically adapt to varying light conditions. The rectangular façade elements made from acetate sheets are programmed to respond to sensor data, opening or closing according to light intensity. This adaptive behavior effectively controls sunlight exposure, optimizing energy conservation within the building.

Prototype development

The development of the kinetic façade system was carried out through a series of systematic steps:

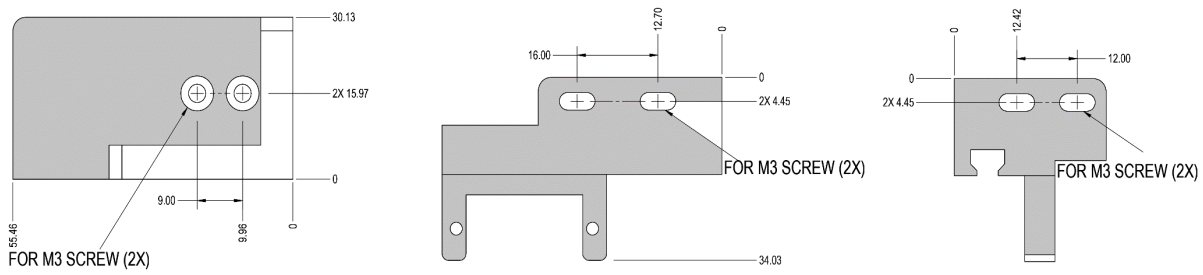


Figure 4: Above is the technical drawing of the 3D-printed motor housings used in the project, showcasing the dimensions and mounting points for the servo motors.

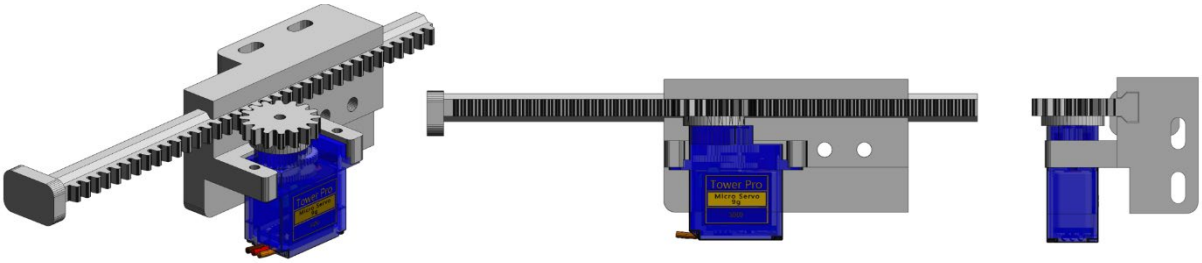


Figure 5: The gear and motor housing design for the façade movement

1.**Fabrication of Façade Elements:** Acetate sheets were chosen to model flexible photovoltaic panels with PET (Polyethylene Terephthalate) surfaces at the correct scale for achieving flexible, movable panels. The sheets were precisely cut into rectangular shapes, positioned at optimal intervals along the façade surface to allow responsive movement according to changing light levels.

2. Integration of Light Sensors: Sensors were strategically placed to detect environmental light levels and transmit signals to the Arduino microcontroller for processing.

3. Installation of Servo Motors: Servo motors were mounted behind each façade element to enable precise movement, controlled by the Arduino based on sensor data.

The 3D-printed motor housings were specifically designed to accommodate the *SG90* and *MG996R servo motors*, ensuring a secure fit and alignment with the façade panels. The motor mounts were custom-designed and fabricated using Blender 3D modeling software, followed by printing with a 3D printer to achieve the precise dimensions required for optimal performance.

The design process involved creating detailed 3D models tailored to the specific servo motors used in the project. These models were then translated into 3D-printed motor housings, ensuring that the servo motors were securely mounted and correctly aligned with the moving façade elements. The precision of these custom-made housings was critical for the reliable operation of the system.

4.Arduino Programming: The Arduino microcontroller was programmed to process real-time sensor data and control the servo motors, triggering the opening or closing of the façade elements based on predefined light thresholds. This systematic approach ensures that the prototype development process can be clearly replicated.

Prototype demonstration

The demonstration phase showcases how the façade elements dynamically respond to varying light intensities. The light sensors continuously monitor the environment and transmit data to the Arduino microcontroller when a specified threshold is exceeded. The microcontroller processes these signals and activates the servo motors, causing the façade elements to either close or open at the required angles.

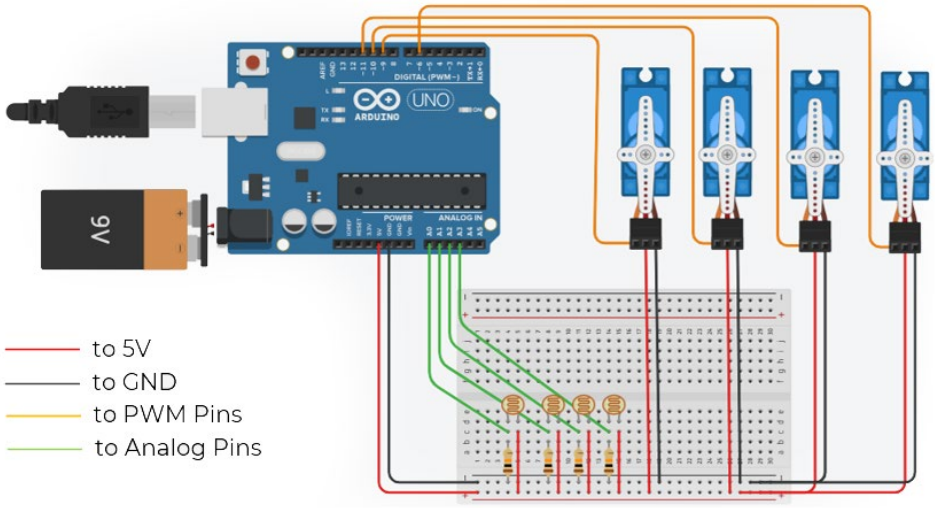


Figure 6: The layout of the components



Figure 7: Smart Motion Prototype

In scenarios of intense sunlight, the façade elements automatically close to mitigate excessive heat gain within the indoor space, thereby maintaining a cooler environment. Conversely, in low light conditions, the façade elements open fully to maximize natural lighting, reducing the reliance on artificial lighting and promoting energy efficiency. The façade's behavior can be customized according to seasonal requirements; for instance, during summer, closing the façade elements upon detecting high light intensity minimizes heat intrusion, while in winter, the system can be programmed to optimize natural lighting for warmth.

This prototype serves as a significant contribution to the development of smart kinetic façades, demonstrating how such systems can dynamically adapt to different climatic conditions, thereby improving energy management, enhancing building sustainability, and offering scalable solutions for modern architectural design.

5. Scientific results and discussion

5.1 Results

The experimental phase of our project focused on the behavior of kinetic façades equipped with light-sensitive sensors, observing their response to varying light conditions throughout the day. The study was divided into two primary observation periods: morning (6:00 AM - 12:00 PM) and afternoon (12:00 PM - 6:00 PM).

Morning Observations:

During the morning hours, the kinetic façade system demonstrated a clear response to increasing sunlight intensity. The façade's sensors detected the rising light levels, leading to a gradual opening of the façade panels. By analyzing the data, it was observed

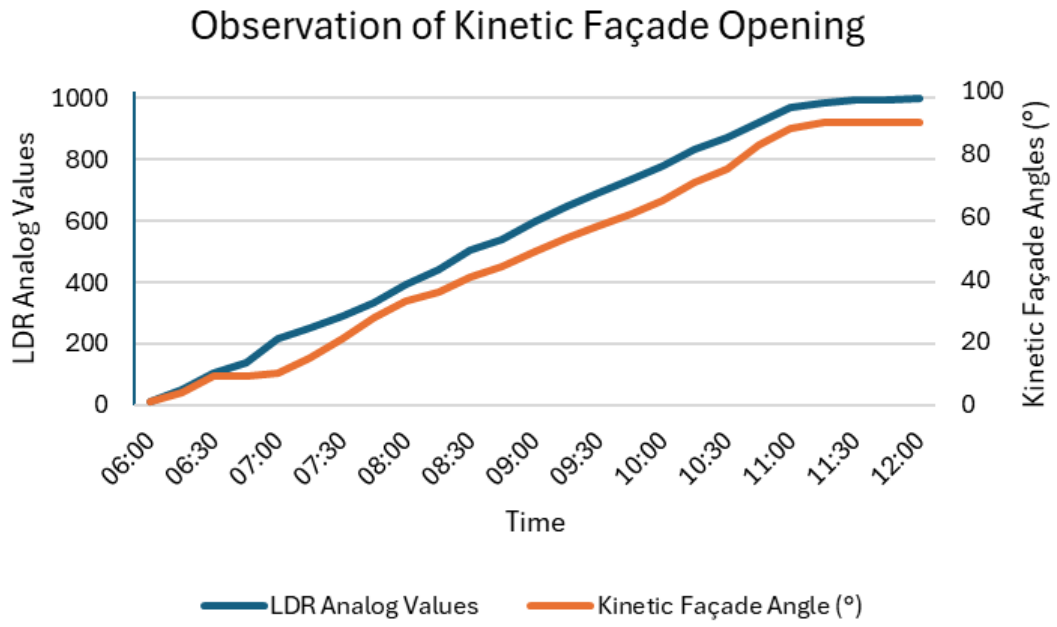


Figure 8: Observation of maximum kinetic façade opening

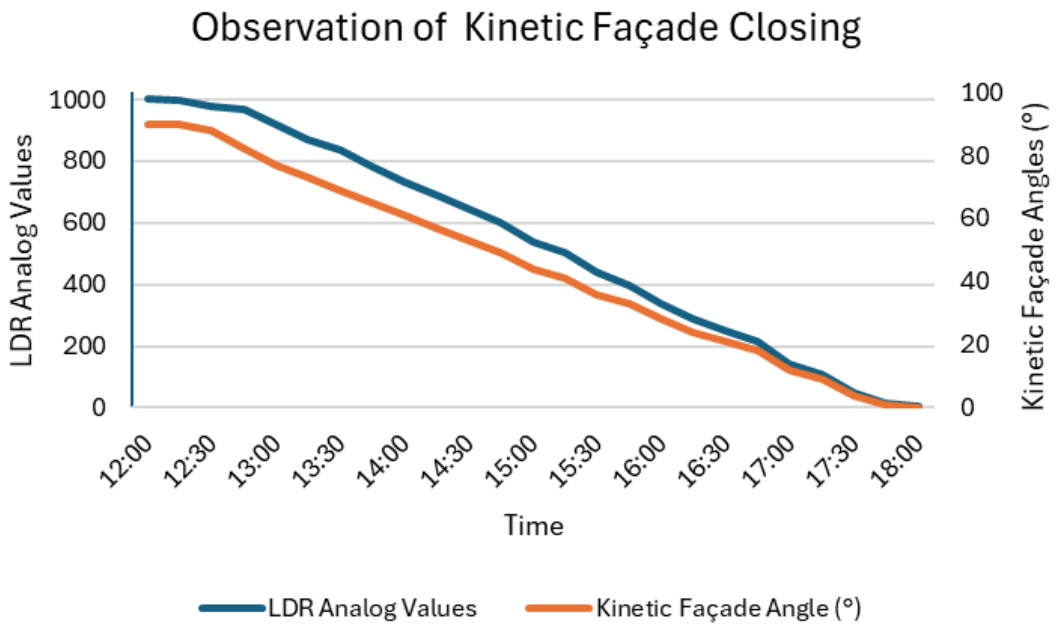


Figure 9: Observation of maximum kinetic façade closing

that the façade reached its maximum angle of opening during the peak sunlight hours between 10:00 AM and 12:00 PM. This behavior aligns with the system's design to maximize natural daylight penetration into the building, reducing the need for artificial lighting.

Afternoon Observations:

In contrast, the afternoon observations revealed a different behavior. As the sunlight intensity began to decrease after noon, the façade system responded by gradually closing the panels. This action was particularly prominent between 2:00 PM and 4:00 PM, indicating the system's ability to modulate based on both light intensity and the angle of sunlight. The closing of the façade during the afternoon helps to prevent excessive heat gain, contributing to the building's energy efficiency by reducing the load on cooling systems.

IoT Integration:

The kinetic façade system was integrated with an IoT-based interface using the Arduino Remote Cloud application. This interface allows for real-time monitoring and control of the façade system, providing users with the ability to adjust the façade position based on indoor and outdoor temperature data, as well as the façade's movement in response to light conditions. The interface also includes controls for indoor lighting, enabling further customization of the building's energy usage.

5.2 Discussion

The results of our study indicate that the kinetic façade system performs as intended, with the ability to dynamically adjust to varying light conditions, thereby optimizing both natural light usage and indoor comfort. The morning observations confirm that the system can effectively maximize daylight use, which is particularly beneficial for reducing reliance on artificial lighting during early working hours. The afternoon closing mechanism, on the other hand, helps to maintain cooler indoor temperatures by limiting direct sunlight exposure, thus supporting energy conservation efforts.

These findings demonstrate the potential of kinetic façades to significantly contribute to energy efficiency in buildings. By responding to real-time light conditions, the system not only enhances indoor comfort but also reduces the overall energy consumption of the building. The ability to integrate with an IoT platform adds an additional layer of functionality, allowing for remote monitoring and control, which is crucial for modern smart building management.

The implications of this study are twofold. First, the dynamic nature of the kinetic façade system shows promise for broader applications in various climatic regions, particularly those with significant daily variations in sunlight. Second, the integration with IoT technology underscores the importance of combining traditional architectural elements with modern digital tools to achieve optimal building performance.

One limitation of the study is the controlled environment in which the observations were conducted. The façade system's performance in real-world conditions, such as extreme weather events or irregular building orientations, was not tested. Additionally, the prototype's reliance on specific sensor calibration may limit its adaptability to different building types without further customization.

Future studies should focus on testing the kinetic façade system in diverse environmental conditions and on different building types to evaluate its broader applicability. Moreover, the development of more advanced sensor technologies that can adapt to a wider range of light intensities and environmental factors could further enhance the system's performance. Finally, exploring the integration of renewable energy sources, such as solar panels, with the kinetic façade system could provide a sustainable energy solution, making buildings not only energy-efficient but also energy-independent.

5.3 Conclusion

This study sought to answer the research question: "How can adaptive building components use IoT sensors and energy recovery technologies to enhance energy efficiency, improve indoor comfort, and reduce maintenance costs?"

To address this question, we focused on the implementation of kinetic façades integrated with IoT-based light sensors and controlled through Arduino programming. The primary goal was to explore how these adaptive systems could automatically respond to varying light conditions to optimize indoor environmental conditions and energy use.

Through the development and testing of our prototype, we found that kinetic façades could be effectively controlled to maximize energy efficiency and indoor comfort. The façades opened fully during periods of maximum daylight and closed during low-light conditions, thereby reducing the need for artificial lighting and maintaining a comfortable indoor temperature. This adaptive behavior, particularly when

The authors are responsible for the content of the papers.

combined with seasonally adjustable modes, demonstrated potential for significant energy savings, especially in heating during winter and cooling during summer.

Moreover, the IoT-based control system we developed allowed for real-time monitoring and manual adjustments via a user-friendly interface. This system not only enhanced the adaptability of the façade but also provided valuable data on indoor and outdoor conditions, which could be used for further optimization.

In conclusion, this research connects the realms of innovative architectural design and practical application through the creation of a functional kinetic façade prototype. By integrating IoT-based light sensors and adaptive control systems, the façade not only enhances indoor comfort but also plays a key role in energy conservation and the reduction of maintenance costs. These findings underscore the feasibility and effectiveness of kinetic façades as a sustainable solution for contemporary building design, with future research possibilities in advanced predictive algorithms and broader IoT integration.

6. References

[1] Zanghirella, F., Babich, F., & Belleri, A. (2018). "Energy Performance of Kinetic Façades in Mediterranean Climates: A Case Study on Dynamic Sun-Shading Systems." Energy and Buildings, vol. 172, pp. 290-302.

[2] Yoon, S., & Sun, S. (2020). "The Impact of Kinetic Shading Systems on Daylighting and Energy Performance in Office Buildings." Journal of Building Performance Simulation, vol. 13, no. 1, pp. 34-50.

[3] Soltani M., Atashi A. (2023). „Designing a Kinetic Façade Using BB-BC Algorithm with a Focus on Enhancing Building Energy Efficiency," 8th International conference on Civil Engineering, Architecture & Urban Development, vol. 8, pp. 249-262.

[4] Zhang, Y., Du, J., & Chen, Q. (2019). "Parametric Design Strategy for Kinetic Façades: Simulation and Performance Optimization." Sustainable Cities and Society, vol. 49, pp. 101-112.

[5] Moloney, J., McLeod, R., & Phillips, R. (2020). "Energy Savings through Kinetic Façade Systems: A Comparative Study of Office Buildings in Different Climates." Renewable Energy, vol. 159, pp. 301-312.

[6] Boud, M., Rodi, P., & Benoy, T. (2018). "Occupant Comfort and Energy Efficiency in Buildings with Kinetic Façades: A Case Study." Building and Environment, vol. 146, pp. 163-173.

[7] Pitts, A. C., & Mar, C. (2021). "Durability and Maintenance of Kinetic Façades: Challenges and Solutions." Energy and Buildings, vol. 235, pp. 110-127.

[8] Frearson, A. (2014). "Kinetic louvres move in three dimensions to screen sunlight from any angle". [Online]. Available: https://www.dezeen.com/2014/03/19/penumbra-kinestic-louvres-tyler-short-movie/.

[9] I. Vinnitskaya (2010). "Kiefer Technic Showroom / Ernst Giselbrecht + Partner," ArchDaily. [Online]. Available: https://www.archdaily.com/87552/kiefer-technic-showroom-ernst-giselbrecht-partner. Accessed: Aug. 13, 2024.

[10] Lam, S. Y., & Cheng, C. H. (2020). "Smart Kinetic Façades: Enhancing Building Performance through Real-Time Data and Predictive Control." Automation in Construction, vol. 113, pp. 103-121.

[11] Moloney, J., McLeod, R., & Phillips, R. (2020). "Energy Savings through Kinetic Façade Systems: A Comparative Study of Office Buildings in Different Climates." Renewable Energy, vol. 159, pp. 301-312.

[12] Boud, M., Rodi, P., & Benoy, T. (2018). "Occupant Comfort and Energy Efficiency in Buildings with Kinetic Façades: A Case Study." Building and Environment, vol. 146, pp. 163-173.

The authors are responsible for the content of the papers.

DAMLA YIGIT, NALAN YILMAZARSLAN

Advanced Sound Detection for Construction Sites

ABSTRACT

Monitoring and tracking of machines used in construction sites constitutes one of the important factors of project tracking. In this research paper, an innovative approach has been conducted on the automation of these processes by audio recording. This study demonstrates two main methodologies that were used to train and test the model as well as the features that were extracted. In the first approach, detection was performed using the Sport Vector Classifier by detecting and comparing the peak points of the sound data. In the second method, Random Forest classifier was used for the detection. Various spectral features were extracted including MFCC, chroma, contrast, tonnetz. With the features extracted from sound data, it was possible to obtain information about which machine was operated, which materials were worked on and how many times these machines were operated. After having a comparative analysis of several classifiers, a result was obtained achieving the highest accuracy up to 86% for the machinery-material detection and up to 100% for the repetitive sound counting. On the other hand, some challenges were faced including varying audio durations, background noises and limited machinery/material types. Despite all the challenges, the findings of the study suggest significant potential for audio-based monitoring systems in industrial applications.

Keywords: Sound Detection; Construction Machine Monitoring; Construction Automation; Activity Recognition; Feature Extraction

DOI: 10.18154/RWTH-2024-09260

Image created with Adobe Firefly

1. Introduction

In recent years, the way industries track and monitor operations has changed radically caused by the revolutionary advent of automation and machine learning, and their incorporation into several sectors, including the construction industry. This report describes research that addresses how to create a sound detection prototype that is able to identify different types of construction machinery and the materials they operate on, by using an advanced audio analysis. It also focuses on how to enhance the monitoring and analysis of machine operations in construction environments. Our motivation stems from the growing field of industrial automation and several studies that take audio data and advance machine learning for sound recognition.

The main goal of this research project is to find solutions for improving the machine operation monitoring and analysis in construction sites through sound detection technology. The necessity of such research is caused by the lack of automation processes and use of manual logging in machinery operations which are prone to errors and imprecisions. Our aim is to develop a more precise and more accurate solution using the sound recognition technology in an automated way. By this project, we can contribute to the optimization of industrial processes and the early detection of potential issues by also checking the comprehensive usage statistics for project and maintenance management.

This lack of automation makes it harder to make data-driven judgments in real time, in addition to requiring a large amount of manual labor. Therefore, inefficient project tracking, human labor dependency and lack of automation in construction sites are the main problems that we focused on for our research.

In this research we investigate how audio data can be used for machinery tracking and scheduling on construction sites. The research objective is to demonstrate whether it is possible to monitor some systems based on audio data reliably. Such investigation would allow us to see the possible future directions for technologies in construction applications by distinguishing between different types of machines, identifying the materials being processed and counting the times they work.

This paper will provide a comprehensive study of particular sound identification systems that were analyzed to test it as a prototype. Subsequently, we will examine how well such algorithms identify certain machines and materials while taking operating unpredictability, overlapping sounds, and background noise into account. Another important aspect of the study is that we will discuss the prototype's capability to monitor the repetitive sounds of machine operations using computer vision techniques. This feature allows us to determine how many times a machine was operated by identifying repetitive sound patterns and silent periods.

2. Literature Review

2.1. Automated machinery tracking and scheduling

The current approach to monitoring construction activities at construction sites relies on the manual collection and analysis of essential data. Automated activity recognition of construction machinery is a highly researched field given its potential to automate manual labor that is required in every construction site. In addition to assessing and analyzing productivity, activity tracking is valuable for safety analysis [1-2], hazard prevention [3-4], emission control and monitoring [5] and resource tracking [6]. Construction equipment activity recognition can generally be divided into two main approaches, depending on the utilization of primary sensors: (1) an active sensor-based approach and (2) a passive sensor-based approach.

2.1.1 Action detection of construction equipment using active sensors

The active sensor-based approach involves mounting sensors directly on the equipment being monitored. This includes the use of technologies such as the global positioning system (GPS), radio frequency identification (RFID), and inertial measurement units (IMU).

The authors are responsible for the content of the papers.

Rossi et al. [7] proposed a micro-controlled smart plug, which integrates technologies like GPS, power sensors, and Wi-Fi that allows for the monitoring and correction of machinery activities. Data from these machines are transmitted to the cloud, enabling real-time performance monitoring and supporting operational decisions. The system helps in predicting performance and planning maintenance by analyzing power consumption patterns specific to each machine. Langroodi et al. [8] used IMUs and GPS attached to the construction equipment for activity recognition and argued that current activity recognition methods face significant limitations because they rely heavily on extensive datasets, a challenge given the wide variety of equipment sizes and shapes. To overcome this, they introduced a Fractional Random Forest method, designed to create an accurate activity recognition model that remains effective and generalizable, even when working with smaller datasets. They focused on heavy construction machinery such as excavators in their case studies.

With this approach, the sensors are directly attached to the equipment, and it requires human involvement. However, this setup is not always practical, particularly for smaller equipment and construction tools like jackhammers, drills, sawing machines, concrete mixers and similar devices. These methods are not directly applicable to the subset of manual activities targeted in this paper, as the focus is on smaller construction equipment rather than heavy construction machinery.

2.1.1 Action detection of construction equipment using passive sensors

The passive sensor-based approach utilizes sensors that are external to the equipment being monitored. This approach includes the use of video and audio information for classification purposes. Using computer vision for activity detection is a widely researched topic.

Chen et al. [9] proposed a computer vision analysis to achieve real-time visual control of the work process in off-site construction. The framework can accurately acquire data from footage and provide process monitoring and problem feedback in real time. Bao et al. [10] employed long-sequence videos to automatically detect, track, and identify the activities of both an excavator and a dump truck.

Even though the computer vision systems are seen better than the traditional methods, cost of implementation, lack of expertise and resistance to change are major challenges in computer vision-based construction progress monitoring adoption.[11]

Many researchers have explored kinematic or visual sensing methods for automated equipment monitoring. Kim et al. [12] argued that combining these two approaches can increase the precision and experimented with a kinematic, visual and hybrid model. The results suggested that hybrid sensing could enhance recognition performance and improve the monitoring of construction equipment. Sherafat et al. [13] also mentioned that combining audio and kinematic data can yield up to 25% more accurate results compared to using single data sources alone.

2.2 Audio based activity tracking

Audio-based activity identification has seen recent applications on construction sites. General process that is used in existing research papers starts with preparing the data by recording and separating the audio. Data is segmented into smaller pieces and specific features are extracted to train the model. Existing approaches follow a similar process but are differentiated by the construction machines they focus on, features they extract from the audio to train the model and the machine learning algorithms they use.

Cheng et al. [14] proposed an audio-based system for activity analysis of single machines performing various tasks. The focus of the proposal was heavy construction machinery such as excavators, wheel loaders and dozers. A support vector machine (SVM) model was trained using time-frequency representations and the accuracy is increased by window filtering. Productive and non-productive activities of the construction equipment are detected.

Rashid and Louis [15] proposed using multi-class support vector machine (SVM) to detect nailing, hammering, sawing, drilling and no task. To achieve this time-domain,

The authors are responsible for the content of the papers.

Frequency-domain, Cepstral-domain and Wavelet- domain features are extracted from the audio. To eliminate irrelevant features from the dataset, a Relief-based feature selection algorithm, known as ReliefF, is implemented. Various window sizes are also tested to evaluate their impact on performance.

Mannem et al. [16] introduced a sound classification model for construction environments that presents a combination of Mel spectrograms and Mel-Frequency Cepstral Coefficient (MFCC) values. The model integrates deep neural networks, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, forming CNN-LSTM and MFCC-LSTM architectures. This approach allows for the extraction of both spectral and temporal features from audio data. The model was evaluated using audio data generated from construction activities in a real-time closed environment, achieving an overall Precision of 91%, Recall of 89%, and F1-score of 91%.

Scarpiniti et al. [17] proposed the use of Deep Belief Network (DBN) for audio-based and autonomous monitoring of construction sites. 97 min of clean audio from ten classes of various construction equipment and tools that are commonly used on construction sites (excavators of different size, bulldozers, compactors, concrete mixer and shovel) were collected and analyzed to validate the proposed approach. Several statistical measures such as the minimum and maximum values of MFCCs, and the first, second, third and fourth moments (i.e., mean, standard deviation, skewness and kurtosis) were derived from MFCCs and mel-scaled spectrograms and were provided as input to DBN.

3. Research Methods

3.1 Methodology Overview

In this section, the whole process will be thoroughly analyzed from the first concept design to the prototyping results by describing our sound recognition prototype's capabilities and limits.

In this project, a research methodology was used with the help of VDI 2221 design method as a guide. It shows the general process model for developing and constructing. The main objective was to create a sound detection system that could recognize construction machines and materials that are worked on. A mixed methodology strategy that combines both the qualitative and quantitative analysis was used to accomplish the project goals. Using mixed methodology made it possible to collect and examine data in a structured way and evaluate the effectiveness of proposed sound detection approaches in construction settings. In this section, the research methods that were used will be examined. The system functionalities of the project will be evaluated by checking all the VDI 2221 stages thoroughly.

3.2 Data Collection

The first step which the data had to be collected was critical in having a reliable and functioning dataset for training and testing the system. To train the model audio data from a drill and a grinder was collected. These construction machines were chosen for their ability to make specific and repetitive sounds. It was an important detail to have machines that made distinguishable sounds in order to simplify the task of machinery detection. Another important element was the choice of materials. Concrete and wood were chosen for the drilling machine tasks. For the grinder, concrete and brick materials were used. The sounds of the machines working on these materials were recorded with a telephone microphone in a controlled and minimal noise environment. More advanced microphone systems could also be used to obtain cleaner and higher quality sounds. The recorded sounds were then named and organized into folders according to the machines and materials used to create a dataset.

3.3 Data Preparation and Feature Extraction

As a starting point, we focused on single machine activities without overlapping sounds. The parts with no machine activity were trimmed from the audio and divided into 5-second sections, which was helpful to organize the data and easier to extract features from.

The authors are responsible for the content of the papers.

Feature extraction was one of the important steps applied to bring unprocessed audio data into a format suitable for machine learning. In this research paper, two approaches were examined and applied for sound detection.

In the first approach the audio data is transformed from the time domain to the frequency domain using the Fourier Transform to reveal the signal's frequency components. The resulting Fourier spectrum is normalized, and significant peak points are identified to create a "mean fingerprint," which captures the essential features of the audio. Using this process, audio recordings are converted into the mean fingerprint space. A Support Vector Classifier (SVC) is then trained on the data to classify new audio inputs based on their mean fingerprints.

In the second approach, four features were determined to be extracted and analyzed from the audio data. Firstly, the time-domain qualities that display the changes in a signal over a span of time were extracted. This feature shows the sounds' temporal properties and includes waveform aspects and energy-based metrics.

Secondly, the frequency domain was obtained, which involves analyzing signals or mathematical functions concerning their frequency components rather than their time-based characteristics.[18] In order to do that, the Mel-Frequency Cepstral Coefficients (MFCC) was extracted, which are usually used in sound detection due to their accuracy in short-term power spectrum of the sound.

Another important aspect was to check the spectral features of the audio signals to capture the harmonic content and the overall structure. The extracted features included; chroma, contrast, and tonnetz. To elaborate further, chroma refers to the quality of a color's purity, saturation or intensity. On the other hand, contrast represents the difference between different color hues. The last extracted feature was the tonnetz which is a graphical tool for analyzing and designing chord movement. [19]

Finally, we located and examined the sound waves' silent periods in order to find out the repetition and frequency of the machines' operations. Getting the silent periods was essential to check how frequently a machine is used throughout the time frame it was operated.

3.4 Classification and Model Training

In the first approach, SVC is used due to its ability to handle complex, non-linear data distributions, which are common in audio signals. The linear kernel used in this approach provides a balance between computational efficiency and accuracy, ensuring that the model can classify audio data quickly and reliably. Furthermore, SVC's resilience to overfitting, especially with small to medium-sized datasets like the one used in this study, makes it an ideal choice for creating a model that generalizes well to unseen data.

In the second approach, the RFC is used which is a commonly used algorithm in machine learning to classify and train the data. It combines the output of multiple decision trees to reach a single result. [20] The reason for its selection is based on its stability and capacity to manage a great number of characteristics without overfitting. All of the features that were extracted from the time and frequency domains were used to train the RFC. Here, we utilized 800 trees to establish a balance between model accuracy and computational efficiency.

The authors are responsible for the content of the papers.

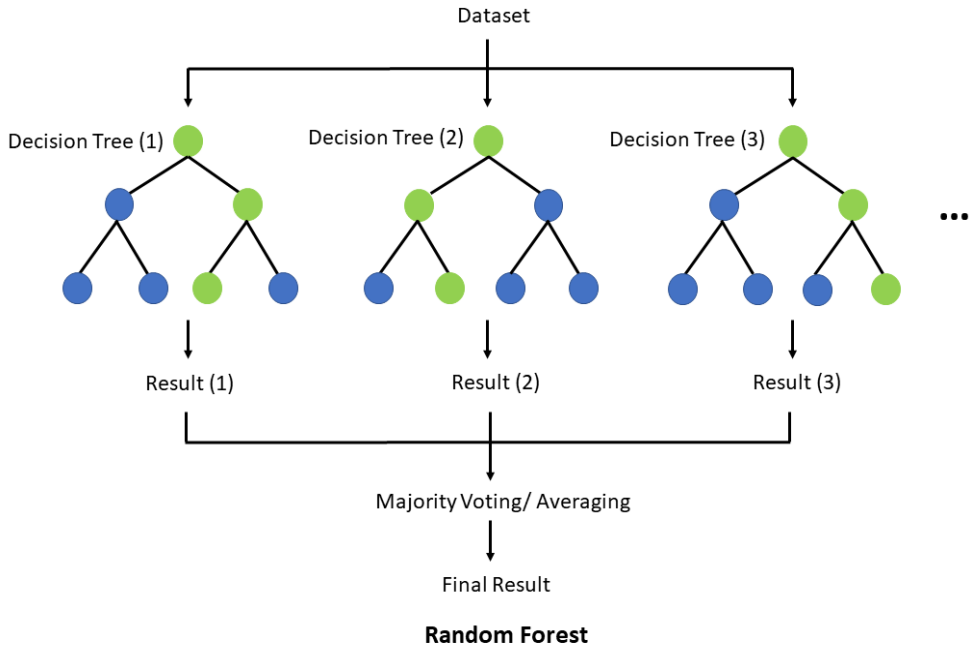


Figure 1. A demonstration diagram for Random Forest Classifier.

3.5 Evaluation

The basis of the evaluation was the classification accuracy of the trained model based on the audio recordings. The system's ability to detect which machine is working on which material and how many times the machine has been operated, based on sound patterns, were examined.

3.6 Application of VDI 2221 Design Method

The research methodology used in this study was selected by considering the practical constraints according with the study objectives. With the methodology that we have used, we made sure that our trained model consists of quality data by focusing on isolated sound events and employing meticulous methods for data preparation and feature extraction.

As mentioned, VDI 2221 design method was used as a method approach which provides an organized design development process with specific stages. It serves as the base for this project's study methodology as it is a general process model for developing and constructing. It is also an iterative procedure for a necessary control process and product improvement. [21] This design method was adapted to meet the needs of the realization of our sound detection research project. Below, all the stages of VDI 2221 method's application are examined.

The first stage was clarifying and establishing the task. The key point here was to identify the equipment, the material that the experiment would be done in a controlled environment. Accuracy, reliability, and functionality were the key performance metrics that were defined in this stage.

The function of the system and its capabilities were determined, including its capacity to record and process the audio data, extract the required features, categorize the audios based on machines and materials. These extracted features consist of peak points, MFCC, chroma, contrast, tonnetz along with RFC and SVC as machine learning tools to provide a potential solution for the functions and structures that are determined.

Use of advanced microphones in a controlled environment is one of the important factors that influences the quality of the data as it helps to reduce the background noise. In order to determine which combination of features and classifiers worked the best, several algorithms were tested and compared with each other. During this stage, the following classifiers were trained and tested; RF, AdaBoost, KNeighbors, MLP, SV.

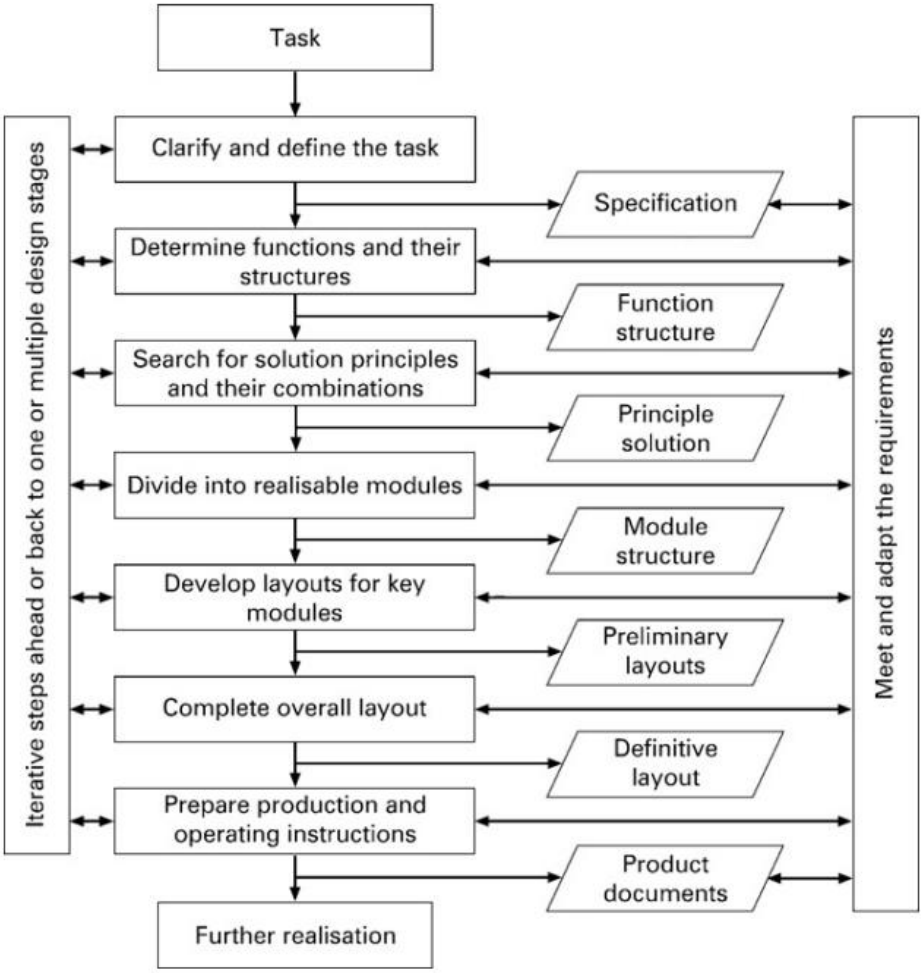


Figure 2: VDI 2221 design process flowchart.

The next stage was to divide the possible solutions into realizable modules by dividing the preprocessing, feature extraction, classification and data collection processes into manageable modules. Each module has been subjected to experimental tests to ensure that it works properly. For example, in the preprocessing module, we normalized the audio data, segmented it into smaller, manageable chunks, and extracted the relevant features according to that.

The divided elements were united in the initial design stage by creating a layout for the tested problems and found solutions. However, the layout plan with our solution included some future provisions to be solved later such as overlapping sound data or background noises.

Once the layout of key modules was developed and the final configuration was done, the system's accuracy and its integration into computer system was fine-tuned. The dataset which consists of 253 audio recordings that we have collected contained a variety of machines and materials and 80% of them were used for machine learning training purposes. On the other hand, the other 20% was used for testing.

Finally, a documentation was prepared, outlining the names and their functionality, along with the timestamps of each of the machines.

Such a planning with the help of VDI 2221 methodology has provided a clear road map in planning and realization of the research project from concept design to prototype that acts as a basis for further developments.

4. Prototype

4.1 Concept Development

Our project's focus is an audio-based machine monitoring system that can identify and detect individual machinery in operation



Figure 3: Illustrations from the design process.

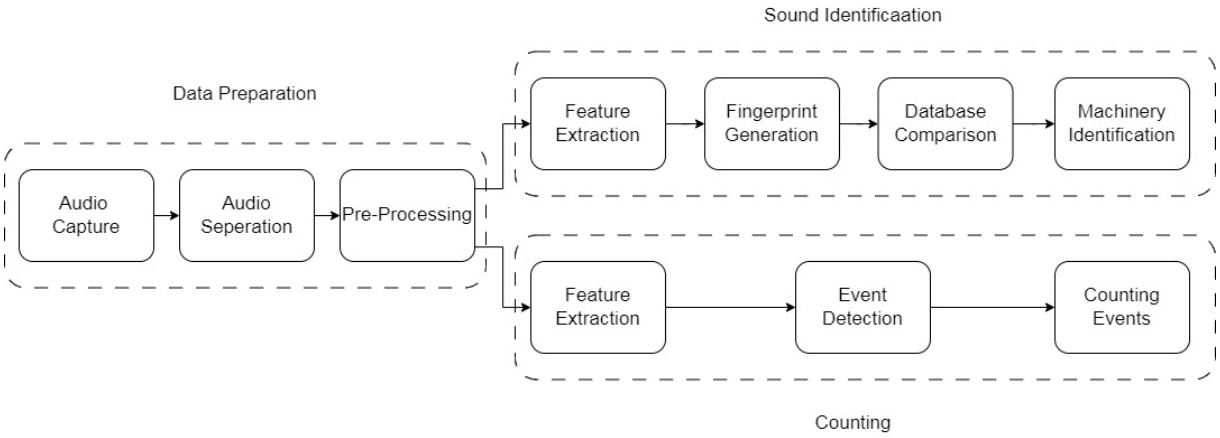


Figure 4: Process model for the prototype

as well as the materials being worked on. In addition to this, using audio data, we aimed to detect which machine has been running for how long and for machines that make a repetitive sound we detected how many times the machine has been operated.

After doing comprehensive research on construction equipment, we assessed the existing systems for tracking such equipment and sought for a gap in the monitoring processes. We have found that progress monitoring is one of the most challenging project management responsibilities and it plays a crucial role for accomplishing the financial and schedule goals of a project. *"This challenge results from the interdependencies between tasks and the fundamental complexity of projects. Therefore, one of the biggest challenges that project managers face is maintaining effective monitoring and control."* [22]

Although there are existing activity recognition systems in the industry, these methods face several challenges including cost of implementation, lack of expertise, digital noise, occlusion, and environmental effects, all of which need efficient solutions. Therefore, we developed our idea based on the increasing demand for effective, automated systems that can track and evaluate machine performance without the need for human involvement. The data collected on machinery activities is planned to be used for autonomously tracking work progress and monitoring on construction sites. Since the electrification of heavy construction machinery is becoming an

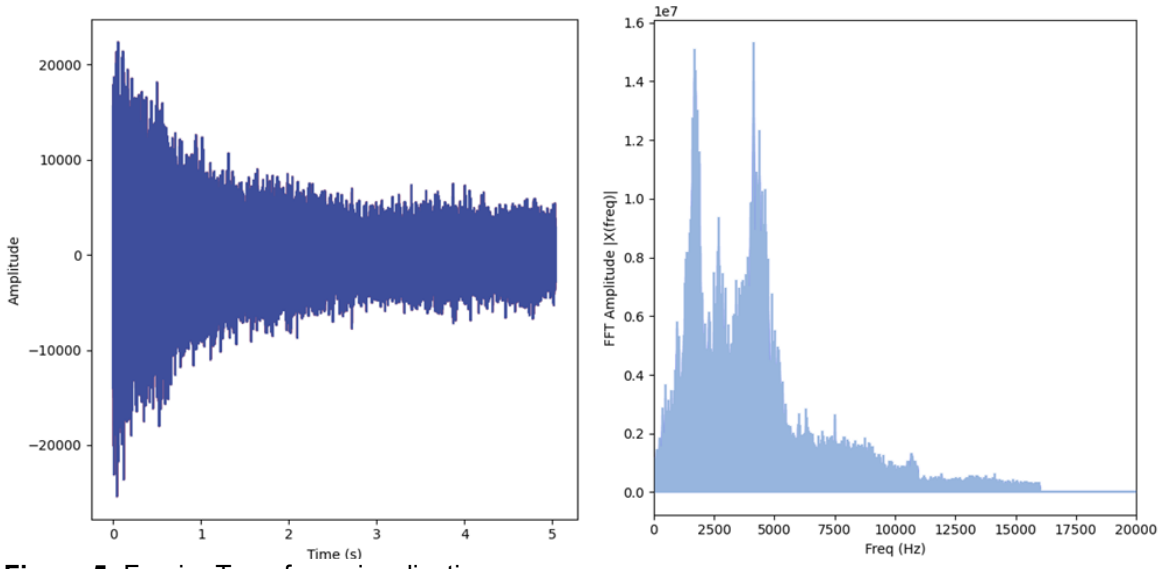


Figure 5: Fourier Transform visualization.

inevitable trend in development [23], our focus was on the frequently used smaller machinery.

As digital sensors and monitoring systems are successively integrated into heavy construction equipment day by day, audio-based tracking becomes less applicable to these types of machinery. Rather, we chose to focus on smaller machinery as they lack advanced digital tracking capabilities. This gap that we have detected inspired a keen insight on developing an effective solution for monitoring those smaller machines. Our solution consists of making use of a simple audio data acquisition setup to capture the sound emitted by construction machinery.

4.2 Prototype Development

To achieve the stated research goal, an audio-based machinery identification model was developed. The development consists of three major phases which are data preparation, sound identification and counting. As depicted in Fig. 1, sound identification and counting work independently.

4.2.1 Data Preparation

The first step of data preparation is audio capturing. In this research we choose to capture audio data from a grinder and a drill, each of which was used on wood and concrete. Single sounds of machine operations are considered to simplify the problem and audios are recorded in a closed environment with minimum background noise. A smartphone is used to record raw audio in mp3 format. Then the audio is converted to wav, parts with no activities are removed and manually separated into 5 seconds long files. The wave library is used to import audio into the code.

4.2.2 Sound Identification Using Peak Points and SVC

The prepared audio data is transformed from time domain into frequency domain using fourier transform. The Fourier Transform is a mathematical tool that transforms a time-domain signal into its frequency-domain representation, revealing the different frequencies that make up the original signal. By decomposing the signal into its constituent sine and cosine waves, the Fourier Transform provides insights into the signal's frequency components, amplitude, and phase.
The Fourier Transform is implemented by the numpy.fft.fft() function. The resulting fourier spectrum is normalized by dividing each value by the maximum value and the fingerprint function is applied. First the function transforms a fourier spectrum into the real value space via calling the absolute function. Later on the peak points of the absolute fourier spectrum which are greater than the threshold are detected. The threshold is used to lower the impact of noise and to simplify the

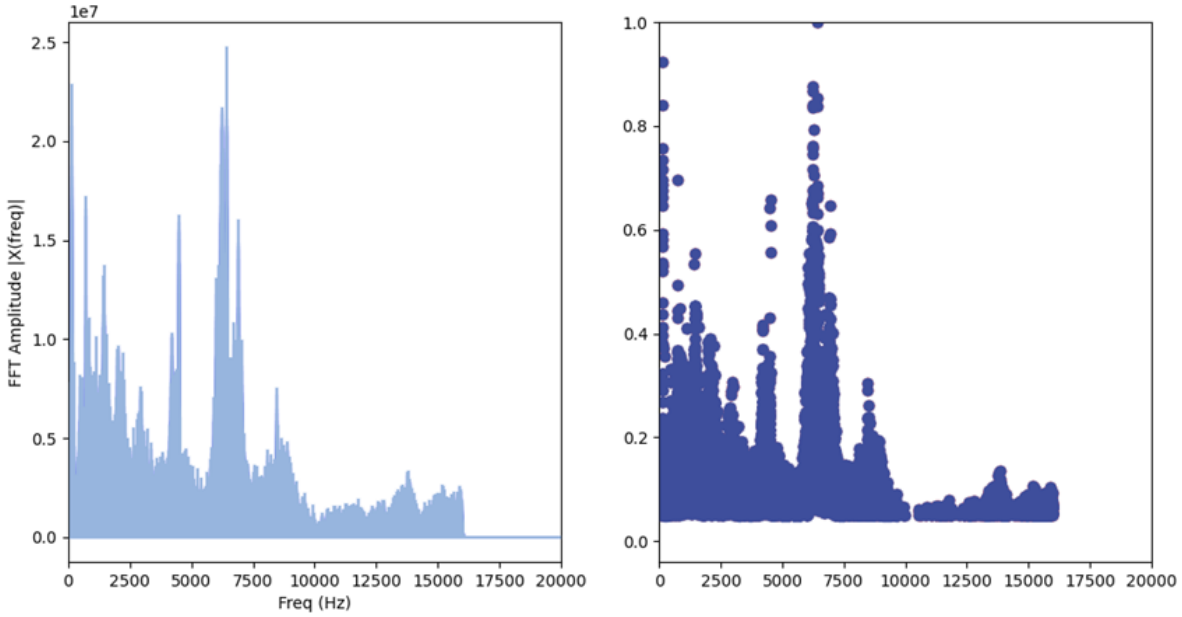


Figure 6: Getting the peak points to differentiate between audios.

data. Furthermore, the collected peaks are combined to calculate a mean fingerprint. The mean fingerprint represents the recorded audio in our "mean fingerprint" space, in which the classification will be applied.

To build classification model 132 five-second-long recordings are used. The recordings are transformed into the "mean fingerprint" space as mentioned and are split into training and testing sets with the ratio 80%/20%. The model is trained on randomly chosen 80% of the data and tested on 20%. This is done to evaluate the model on unseen data. As a classifier a Support Vector Classifier (SVC) is considered which is provided by the sklearn library. A SVC works by finding the optimal hyperplane that best separates data points from different classes in a feature space. The hyperplane is positioned to maximize the margin, which is the distance between the hyperplane and the nearest data points from each class, known as support vectors. SVC can handle complex, non-linear data distributions by using kernel functions that transform the data into higher-dimensional spaces where a linear separation is possible. The model then uses this hyperplane to classify new data points, determining which side of the boundary they fall on to assign them to a class.

The SVC uses a linear kernel, which results in linear hyperplanes. The decision function is one-vs-rest. Hereby, the multi-class classification problem is decomposed into multiple binary classification problems, with each classifier distinguishing one class from the rest.

The SVC is evaluated with the accuracy_score function imported from sklearn.metrics, which calculates the ratio of correct prediction to total predictions. The trained model is saved as a joblib file, thereby allowing it to be opened if needed.

To use the model, an unknown sound needs to be imported into the project. The length of the unknown sound differs, and it is seen that longer audios had lower detection accuracy. To counteract this, the unknown sound is cut into 5 second chunks and each chunk gets detected by the model.

Therefore, each audio chunk is transformed into the "mean fingerprint space" as described before. To identify the chunk, the mean fingerprint of it is placed into the space with the trained hyperplanes as shown in Fig. 5. The area where it is located defines the machinery and material combination. The most identified machinery and material combination in total is chosen. Detected machines and materials are saved on an excel sheet with the time they were detected, detected machinery and material, the duration of the audio file, the duration of activity and how many times the machine has been used.

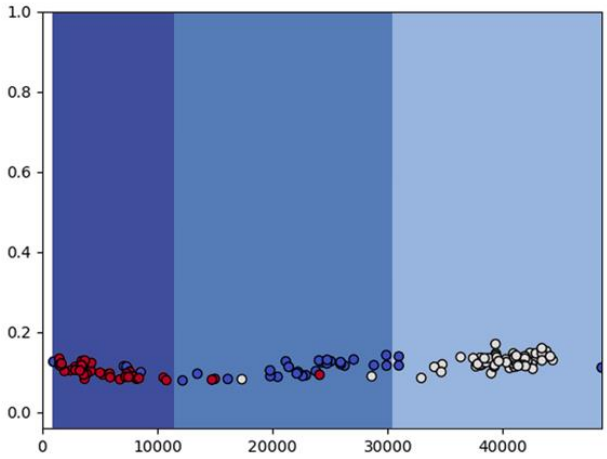


Figure 7: Visualization of the classification process

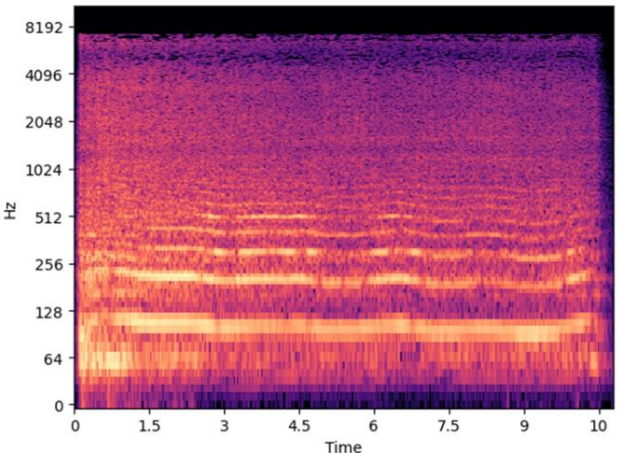


Figure 8: Visualization of the extracted features.

4.3 Sound Identification Using Spectral Feature Extraction and RFC

In another method that we have used for the Fourier Transform, specific spectral features were extracted in order to accurately identify and distinguish between different types of machinery sounds and the materials they interact with by analyzing the frequency content of the audio signals.

In this section of the code, we intend to analyze all the .wav files that are specified in the relevant directory. Therefore, the loop '*for wav_file in files_in_dir:'* iterates through each .wav file, as the first step. Thereafter, the function named '*preprocess_audio'* is called from the '*file_dir'* directory, as it returns the extracted features which are the mfcc, chroma, contrast and tonnetz. These features are important in terms of characterizing the audio signal by its frequency and harmonic content. Subsequently, the next line '*new_row'* creates a pandas dataframe with the extracted features as each of them are stored in separate columns in a dictionary. Finally, the last line appends the '*new_row'* that was created to an existing dataframe which is the *'df_drilling'*. Meanwhile, *'ignore_index=True'* argument ensures that the indices of the dataframe are in correct order.

4.3.4 Counting

One of the three main phases in this prototyping project is to count how many times a machine was used. For machines that are used in repetitive actions such as drills, the number of events is detected.

The event counting phase begins by loading and normalizing the audio file to ensure a consistent analysis environment. After trimming any silence from the beginning and end of the audio, the code computes a Mel-spectrogram via librosa.feature.melspectrogram, which captures the frequency content of the signal in a meaningful way. This spectrogram is converted to decibel units with librosa.power_to_db to enhance the visibility of frequency components, and the resulting image is saved.

For detecting periods of activity, the saved spectrogram image is processed using OpenCV. The image is converted to the HSV color space using cv2.cvtColor, which makes it possible to detect red areas that indicate machine activity.The code defines two red hue ranges to account for the dual nature of red in the HSV color spectrum, applying cv2.inRange to create masks that isolate these regions. Contours are then detected using cv2.findContours, representing the boundaries of active periods. These contours are filtered based on a minimum duration threshold, calculated from the pixel width and the total duration of the audio, ensuring that only significant periods of activity are considered. The code then visually marks these regions on the spectrogram and extracts the corresponding audio segments, returning key metrics such as the total audio duration, the number of detected active periods, and their cumulative duration.

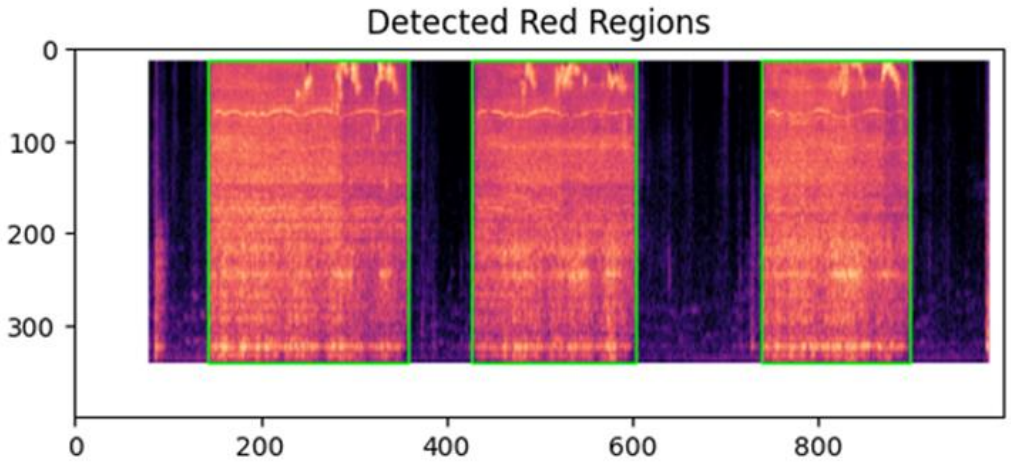


Figure 9: Detecting the red regions for counting repetitive sounds.

4.4 Prototype Demonstration

A demonstration of the prototype was carried out on reference construction site of RWTH Aachen University. A detailed prototype demonstration and project explanation can be found from the links below:

https://youtu.be/VL8qEQ0pDEs
https://youtu.be/GJ6F75krM8E

Furthermore, for those who wish to replicate and experiment with the project, the GitHub link provided below offers full access to the necessary resources:

https://github.com/damlayigit/Advanced-Sound-Detection-for-Construction-Sites.git

5. Scientific Results and Discussion

5.1 Results

The accuracy of the machine and material detection varied significantly depending on the classifiers that were used and the features that were extracted. The accuracy percentages were estimated by averaging the results obtained from executing the code for 20 times.

First method that was trained using SVC and the extracted peak points had the highest accuracy among the classifiers tested. The model is able to detect two machines and two materials with 86.66 % accuracy. This classifier was remarkably superior to the other classifiers that were used.

In the second method, spectral features from the audios were extracted. The final results among four of the classifiers were compared and, it is found that RFC has the highest accuracy, with a performance of approximately 67.35%. The other classifiers that were tested such as K-Nearest Neighbors (KNN) achieved an average accuracy of 61.45%, and Multi-Layer Perceptron (MLP) had a mean accuracy of 40.86%. The AdaBoost classifier demonstrated the lowest accuracy, with an average of 35.75%.





Figure 10&11: Prototype demonstration on site.

```
Mean accuracy for RF:  0.6735384615384615
Mean accuracy for AdaBoost:  0.3575384615384615
Mean accuracy for KNN:  0.6144615384615385
Mean accuracy for MLP:  0.4086153846153846
```

```
The average score is 0.8666666666666668!
```

Figure 12&13: Mean accuracy displayed on the output console.

The system was able to detect four different combinations of machinery and materials while using the RFC and SVC which had the highest accuracy. These combinations included tasks like drilling on concrete, drilling on wood, grinding on concrete, and grinding on brick. However, one of the results that stands out showed that the accuracy percentages have significantly decreased in the presence of background noise. In particular, it was observed that the accuracy of audio recordings longer than 5 seconds decreased by almost half. Despite these challenges we faced, the system effectively employed computer vision techniques to analyze sound waves, achieving near-perfect accuracy up to 100% in determining the number of times a machine was operated.

Throughout the testing process, the same brands and types of equipment and materials were used to ensure the accuracy and precision of the system, due to the lack of information on whether there would be a decrease in the accuracy of detection depending on the variations in the sound profiles of various brands or models.

Furthermore, during the training process, the accuracy of the system decreased as the sound samples' duration increased. A function was called to address this problem by dividing the unknown sound into 5-second segments, with each segment analyzed individually by the model.

These results suggest that while the RFC and SVC provide the best overall performance, improvements could be implemented for the future trials by addressing the challenges of background noise, sample duration, and variations in machine and material types.

5.2 Discussion

In this section of the report, the results are analyzed, and several factors that influence the detection accuracy of our project system are considered. These factors include machine types, material types, environmental conditions, and the length of audio files that were recorded in that environment. Enhancing noise reduction methods to effectively manage background interference could be the main focus of improvements. Furthermore, adding more machines and materials to the dataset would probably increase the robustness and generalization of the system.

The recording and monitoring of machine operation data is an additional area that is worth exploring. This is now handled by an Excel sheet, but using more sophisticated analytics could provide deeper insights and point out areas that require greater improvement.

As the convolutional neural network (CNN) models can capture the spatial characteristics of features, particularly for spectral, MFCC, or other acoustic features of speech signals, [24] adding neural networks might provide a way to achieve better accuracy, especially in noisy and complicated contexts in future developments. Therefore, we believe that neural networks have the potential to address some of the drawbacks associated with classical classifiers and are ideally suited for identifying complex patterns in audio data.

In conclusion, future improvements could include some solutions regarding the enhancing of noise robustness, supporting multi-machine detection, and expanding the database. In construction, this technology can optimize operations and enhance project tracking and maintenance.

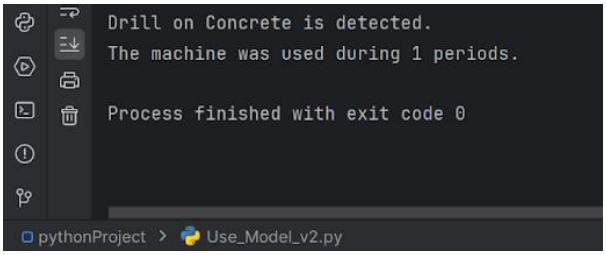


Figure 14: An example of machinery and material detection after running the code.

5.3 Outlook

Using audio data to detect construction machinery activities is a promising research area and it possesses the possibility to create an automated system to monitor work progress without the need for human involvement. However, several challenges need to be addressed for implementing an audio-based system on construction site conditions.

First of all, this prototype was trained with the audios of single machine activities that are recorded on optimal conditions. Nevertheless, on real-world construction sites it is not possible to obtain these optimal site conditions. Therefore, one of the next steps on this research would be to identify machine activities that are happening simultaneously, including the background noises of construction sites and training the model accordingly.

In this research training and testing is completed using the same machines and the trained model isn't tested on different machines of the same type. For this reason, data on how well the model would perform to detect the same machine type from a different brand or a manufacturer has not been acquired. This information can provide valuable insight on how much data needs to be gathered to implement audio-based monitoring systems on construction sites.

Implementation of a sound-based detection system on construction sites carries several challenges. For the system to work fully, microphones need to be correctly positioned, protected and the transmitting range must be determined. To have a realistic application plan, the way of transmitting audio data needs to be concretized.

In this research, the audio data is gathered using a smartphone. The impact that the recording device has on accuracy is one of the topics that can be further investigated.

6. References

[1] Alessandro Carbonari, Alberto Giretti, Berardo Naticchia, A proactive system for real-time safety management in construction sites, Automation in Construction, Volume 20, Issue 6, 2011, Pages 686-698, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2011.04.019.

[2] Tao Cheng, Jochen Teizer, Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications, Automation in Construction, Volume 34, 2013, Pages 3-15, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2012.10.017. (https://www.sciencedirect.com/science/article/pii/S0926580512001847)

[3] Tingsong Chen, Nobuyoshi Yabuki, Tomohiro Fukuda, Mixed reality-based active Hazard prevention system for heavy machinery operators, Automation in Construction, Volume 159, 2024, 105287, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2024.105287. (https://www.sciencedirect.com/science/article/pii/S0926580524000232)

[4] Yong-Cheol Lee, Moeid Shariatfar, Abbas Rashidi, Hyun Woo Lee, Evidence-driven sound detection for prenotification and identification of construction safety hazards and accidents, Automation in Construction, Volume 113, 2020, 103127, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2020.103127. (https://www.sciencedirect.com/science/article/pii/S092658051930038X)

[5] Lewis, P., Leming, M., Frey, H. C., & Rasdorf, W. (2011). Assessing Effects of Operational on Pollutant Emissions of Nonroad Diesel Construction Equipment. Transportation Research Record, 2233(1), 11-18. https://doi.org/10.3141/2233-02 [6] S. Bryson. "Tense tendencies in academic texts." https://www.scribbr.com/academic-writing/tense-tendencies/ (accessed May. 15, 2023).

[6] Reza Akhavian, Amir H. Behzadan, An integrated data collection and analysis framework for remote monitoring and planning of construction operations, Advanced Engineering Informatics, Volume 26, Issue 4, 2012, Pages 749-761, ISSN 1474-0346, https://doi.org/10.1016/j.aei.2012.04.004. (https://www.sciencedirect.com/science/article/pii/S1474034612000432)

[7] Andrea Rossi, Yanet Vila, Federico Lusiani, Ludovico Barsotti, Luca Sani, Pierfrancesco Ceccarelli, Michele Lanzetta, Embedded smart sensor device in construction site machinery, Computers in Industry, Volume 108, 2019, Pages 12-20, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2019.02.008.(https://www.sciencedirect.com/science/article/pii/S0166361518303944)

The authors are responsible for the content of the papers.

[8] Armin Kassemi Langroodi, Faridaddin Vahdatikhaki, Andre Doree, Activity recognition of construction equipment using fractional random forest, Automation in Construction, Volume 122, 2021, 103465, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2020.103465. (https://www.sciencedirect.com/science/article/pii/S0926580520310451)

[9] Chieh-Feng Cheng, Abbas Rashidi, Mark A. Davenport, David V. Anderson, Activity analysis of construction equipment using audio signals and support vector machines, Automation in Construction, Volume 81, 2017, Pages 240-253, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2017.06.005. (https://www.sciencedirect.com/science/article/pii/S0926580517305149)

[10] Bao, Ruxiao & Sadeghi, Amin & Golparvar-Fard, Mani. (2016). Characterizing Construction Equipment Activities in Long Video Sequences of Earthmoving Operations via Kinematic Features. 849-858.10.1061/9780784479827.086.

[11] Bozorgzadeh A and Umar T, Automated progress measurement using computer vision technology in UK construction, Proceedings of the Institution of Civil Engineers – Smart Infrastructure and Construction176(4): 165–182, 2023, https://doi.org/10.1680/jsmic.22.00026

[12] Jinwoo Kim, Seokho Chi, Changbum Ryan Ahn, Hybrid kinematic–visual sensing approach for activity recognition of construction equipment, Journal of

Building Engineering, Volume 44, 2021, 102709, ISSN 2352-7102, https://doi.org/10.1016/j.jobe.2021.102709. (https://www.sciencedirect.com/science/article/pii/S2352710221005672)

[13] Sherafat, Behnam & Rashidi, Abbas & Lee, Yongcheol & Ahn, Changbum. (2019). Automated Activity Recognition of Construction Equipment Using a Data Fusion Approach, Computing in Civil Engineering 2019, 1-8 https://doi.org/10.1061/9780784482438.001 (https://ascelibrary.org/doi/abs/10.1061/9780784482438.001)

[14] Chieh-Feng Cheng, Abbas Rashidi, Mark A. Davenport, David V. Anderson Activity analysis of construction equipment using audio signals and support vector machines, Automation in Construction, Volume 81, 2017, Pages 240-253, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2017.06.005.(https://www.sciencedirect.com/science/article/pii/S0926580517305149)

[15] Khandakar M. Rashid, Joseph Louis, Activity identification in modular construction using audio signals and machine learning, Automation in Construction, Volume 119, 2020, 103361, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2020.103361. (https://www.sciencedirect.com/science/article/pii/S0926580520309419)

[16] Karunakar Reddy Mannem, Eyob Mengiste, Saed Hasan, Borja García de Soto, Rafael Sacks, Smart audio signal classification for tracking of construction tasks, Automation in Construction, Volume 165, 2024, 105485, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2024.105485. (https://www.sciencedirect.com/science/article/pii/S0926580524002218)

[17] Michele Scarpiniti, Francesco Colasante, Simone Di Tanna, Marco Ciancia, Yong-Cheol Lee, Aurelio Uncini, Deep Belief Network based audio classification for construction sites monitoring, Expert Systems with Applications, Volume 177, 2021, 114839, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2021.114839. (https://www.sciencedirect.com/science/article/pii/S0957417421002803)

[18] Time Domain Analysis vs Frequency Domain Analysis: A Guide and Comparison. (2024, July 17). Advanced

PCB Design Blog. https://resources.pcb.cadence.com/blog/2020-time-domain-analysis-vs-frequency domain-analysis-a-guide-and-comparison

[19] FGJ, A. (n.d.). Frans Absil Music | Tonnetz Diagram | Riemannian Transformations. https://www.fransabsil.nl/htm/tonnetz_riemannian_transformations.htm, (accessed August. 10, 2024).

[20] What Is Random Forest? | IBM. (n.d.). https://www.ibm.com/topics/random-forest, (accessed August. 8,2024).

[21] Tomiyama, Tetsuo & Gu, P. & Jin, Yan & Lutters, Eric & Kind, Christian & Kimura, Fumihiko. (2009). Design methodologies: Industrial and educational applications. CIRP Annals-Manufacturing Technology. 58.543-565. 10.1016/j.cirp.2009.09.003., (https://www.researchgate.net/figure/Recom

mended-design-process-VDI-2221-151_fig17_280940066)

[22] Muhammad Ali Musarat, Abdul Mateen Khan, Wesam Salah Alaloul, Noah Blas, Saba Ayub, Automated monitoring innovations for efficient and safe construction practices, Results in Engineering, Volume 22, 2024, 102057, ISSN 2590-1230, https://doi.org/10.1016/j.rineng.2024.102057. (https://www.sciencedirect.com/science/article/pii/S2590123024003116)

[23] Xiaohui Huang, Wanbin Yan, Huajun Cao, Sujiao Chen, Guibao Tao, Jin Zhang, Prospects for purely electric construction machinery: Mechanical components, control strategies and typical machines, Automation in Construction, Volume 164, 2024, 105477, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2024.105477. (https://www.sciencedirect.com/science/article/pii/S0926580524002139)

[24] Chenjian Sun, Min Jiang, Linlin Gao, Yu Xin, Yihong Dong, A novel study for depression detecting using audio signals based on graph neural network, Biomedical Signal Processing and Control, Volume 88, Part B, 2024, 105675, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2023.105675. (https://www.sciencedirect.com/science/article/pii/S1746809423011084)

The authors are responsible for the content of the papers.

The authors are responsible for the content of the papers.



construction robotics

www.cr.rwth-aachen.de

international M.Sc.programme