



An-Najah National University
Faculty of Engineering & Information Technology

Graduation Project 2

**Photographer Robot
-Thikra-**



Students:

Dania Shashtare and Rama AbuGhdeeb

Supervisor:

Dr. Saad Tarabieh

Academic Year: 2025–2026

Presented in partial fulfillment of the requirements for the
Bachelor's Degree in Computer Engineering

Acknowledgment

We would like to express our sincere gratitude to our supervisor, Dr. Saad Tarabieh, for his guidance, support, and valuable feedback throughout this project. His expertise played a major role in the successful completion of this work.

We would also like to thank our families for their continuous encouragement and support during our academic journey.

Disclaimer

This report was written by Dania Shashtare and Rama Abu Ghdeeb at the Computer Engineering Department, Faculty of Engineering and Information Technology, An-Najah National University. It has not been altered other than editorial corrections.

Abstract

This project presents the design and implementation of **Thikra**, a smart photographer robot developed to automate the photo capturing process in an intelligent and interactive manner. The main objective of the system is to provide a hands-free photography solution that minimizes human intervention while maintaining accurate framing and high-quality image results.

The system employs computer vision techniques to detect users, determine their position, and measure the distance between the camera and the subject. Based on this information, the robot automatically adjusts the camera orientation and position to ensure proper alignment and balanced image composition.

User interaction is achieved through an integrated keypad, display screen, and audio feedback system, allowing users to operate the robot easily and intuitively. Captured images can be previewed in real time and printed instantly, enhancing the overall user experience.

Thikra integrates both hardware and software components, including camera modules, motors, sensors, and microcontrollers, which work together to ensure smooth camera movement, reliable user detection, and synchronized system control.

Compared to traditional photography systems and conventional photo booths, the proposed robot offers autonomous operation, real-time user interaction, and automatic adjustment capabilities, making it suitable for events, exhibitions, and indoor public environments where efficiency, accuracy, and user convenience are essential.

Contents

Acknowledgment	1
Disclaimer	2
Abstract	3
List of Figures	6
List of Abbreviations	7
Chapter 1 — Introduction	8
1.1 Statement of the Problem.....	8
1.2 Objectives of the Work.....	8
1.3 Significance of the Work	9
1.4 Organization of the Report	9
Chapter 2 — Theoretical Background and Previous Work	10
Chapter 3 — Methodology	11
3.1 System Design	11
3.2 Hardware	12
3.2.1 Main Processing Units	12
3.2.2 Vision System	13
3.2.3 Motion and Actuation	14
3.2.4 Sensors	16
3.2.5 User Interaction Components	17
3.2.6 Power System.....	20
3.3 Software	21
3.4 Workflow	22
3.4.1 Constraints.....	26
Chapter 4 — Results and Analysis	27
4.1 User Detection Accuracy.....	27
4.2 Photo Capture Performance	27
4.3 User Interaction and Photo Selection	27
4.4 Summary of Results	28
Chapter 5 — Discussion	29
5.1 Problem Resolution	29

5.2	Contribution	29
5.3	Logical Implications	29
Chapter 6 — Conclusions and Recommendations		31
6.1	Conclusion	31
6.2	Recommendations	31

List Of Figures

3.1	System design of the Thikra Photographer Robot	11
3.2	Raspberry Pi 4 Model B	12
3.3	Arduino Mega 2560.....	12
3.4	OAK-D Camera	13
3.5	Live Camera (USB Camera).....	14
3.6	DC Gear Motor used in the Thikra Robot	14
3.7	Aluminum Omni Wheels used for multidirectional movement	15
3.8	L298N Motor Driver Module used to control DC motors	16
3.9	HC-SR04 Ultrasonic Sensor used for obstacle detection.....	16
3.10	iPad display used for photo preview and user selection	17
3.11	Keypad interface used for password entry and session configuration	17
3.12	Speaker module used for audio guidance and user prompts.....	18
3.13	Relay-controlled flash system used for lighting enhancement	18
3.14	Seven Segment Display used to show remaining shots.....	19
3.15	Compact printer used for instant photo printing.....	19
3.16	Power bank used to supply stable power for the Raspberry Pi	20
3.17	Lithium battery pack used as the main power source for motors.....	20
3.18	LM2596 buck converter module used for voltage regulation	21

List of Abbreviations

- **AI** – Artificial Intelligence
- **SBC** – Single Board Computer
- **UART** – Universal Asynchronous Receiver/Transmitter

Chapter 1 — Introduction

Photography plays an important role in documenting moments during events, celebrations, and public activities. However, traditional photography methods often depend on human photographers or fixed photo booths, which may lead to limited flexibility, inconsistent image quality, and delays in capturing photos.

With the rapid development of robotics and computer vision technologies, it has become possible to automate the photography process and improve both accuracy and user experience. The Photographer Robot is a smart robotic system designed to capture photos automatically by detecting users, adjusting camera position, and interacting with them in an intuitive manner. This project integrates robotics, computer vision, and embedded systems to create an intelligent photography solution capable of operating autonomously. The robot aims to provide a modern, efficient, and interactive alternative to traditional photography systems, making it suitable for various indoor environments and public events.

1.1 Statement of the Problem

The main problems addressed in this project can be summarized as follows:

- Heavy dependence on human photographers in events and public places.
- Limited flexibility of traditional photography systems.
- Inaccurate camera positioning and inconsistent image quality.
- Long waiting times and crowding during photo capturing.
- Inability of conventional systems to adapt to different user positions.
- Negative impact on user experience due to poor image framing.
- Lack of an autonomous system capable of automatic photo capturing and camera adjustment.

1.2 Objectives of the Work

The main objectives of this project are:

- To design and implement an autonomous photographer robot capable of capturing images without human intervention.
- To detect users using computer vision techniques and determine their position and distance from the camera.
- To automatically adjust the camera orientation according to the user's position.

- To provide an interactive user interface using a keypad, display screen, and audio feedback.
- To allow users to preview and print captured images instantly.
- To integrate hardware and software components into a reliable and efficient robotic system.

1.3 Significance of the Work

The importance of this project can be summarized as follows:

- Introducing an innovative approach to automated photography by integrating robotics, embedded systems, and computer vision technologies.
- Reducing dependence on human photographers in events and public environments.
- Enhancing user experience through real-time interaction and automatic camera adjustment.
- Providing a practical solution that can be applied in events, exhibitions, malls, and indoor public spaces.
- Offering a scalable and modern photography system suitable for different environments.

1.4 Organization of the Report

This report is organized as follows:

- Chapter 1 introduces the project, presenting the problem statement, objectives, significance, and overall structure of the report.
- Chapter 2 provides the theoretical background and reviews related work relevant to robotics, computer vision, and automated photography systems.
- Chapter 3 describes the methodology of the project, including system design, hardware components, software implementation, and workflow.
- Chapter 4 presents the results and analysis of the implemented system.
- Chapter 5 discusses the achieved outcomes, system limitations, and project contributions. Finally,
- Chapter 6 concludes the project and provides recommendations for future improvements.

Chapter 2 — Theoretical Background and Previous Work

Recent research has highlighted the growing role of intelligent robotic systems in automated photography and human–robot interaction. Traditional photography systems rely heavily on human operators to control camera positioning, framing, and user guidance, which can limit consistency and efficiency, especially in public or self-service environments. Previous studies have explored the use of mobile robots and smart camera platforms capable of tracking users, adjusting camera orientation, and capturing images automatically using computer vision techniques. These systems often employ sensors, depth cameras, and AI-based algorithms to detect human presence, estimate distance, and ensure proper framing and image quality. Other research has focused on enhancing user interaction through integrated interfaces such as displays, keypads, and audio feedback, allowing users to interact naturally with the robotic photographer. Additionally, advances in intelligent image selection and automated triggering have reduced the need for manual intervention while improving accuracy and user experience. Building on these studies, our project Thikra integrates computer vision, autonomous robotic movement, and interactive user control into a single smart photographer robot. By combining real-time user detection, automatic camera alignment, and instant image preview and printing, the system extends previous work and provides a practical, hands-free photography solution suitable for real-world applications.

Chapter 3 — Methodology

The development of **Thikra** followed a structured methodology that combined system design, hardware integration, software development, and system-level testing. This approach ensured that each component was individually validated before being integrated into a complete robotic photography system capable of autonomous operation and real-time user interaction.

The following subsections describe the methodology in detail.

3.1 System Design

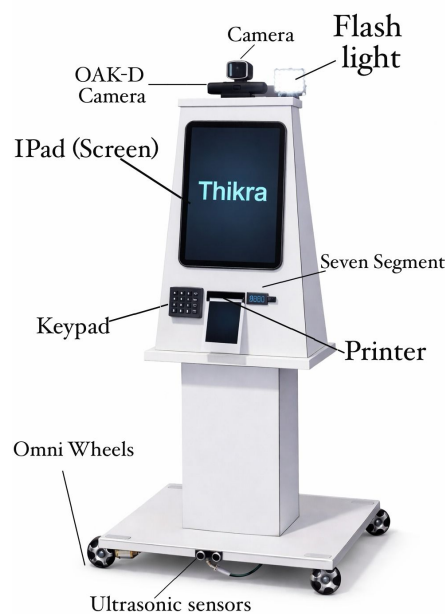


Figure 3.1: System design of the Thikra Photographer Robot

The Photographer Robot is built around a Raspberry Pi 4 as the main processing unit, supported by an Arduino Mega for motor control and hardware interfacing. The robot's mobility is achieved using four DC motors, allowing smooth movement in all directions.

An OAK-D camera is integrated for person detection and localization. The camera enables the robot to detect the presence of a user, determine their position, and move toward them. Once the user is detected and properly positioned, the system performs facial analysis to guide the user by providing voice instructions such as turning toward the camera, smiling, or opening their eyes before capturing the image.

In addition, During image capture, the flash is activated to enhance the quality of the photos taken by the camera.

3.2 Hardware

Each hardware component was selected based on performance, compatibility, and availability. A summary of the main components and their specifications is provided below:

3.2.1 Main Processing Units

Raspberry Pi 4



Figure 3.2: Raspberry Pi 4 Model B

A powerful, credit-card-sized single-board computer (SBC) that offers desktop-level performance, dual 4K display support, gigabit Ethernet, and USB 3.0 connectivity. It's a versatile device for hobbyists and professionals alike, functioning as a media center, smart home controller, server, robot brain, or educational platform, capable of running various operating systems, including Raspberry Pi OS (Raspberry Pi Foundation, n.d.). We used Raspberry pi documentation to install OS (Raspberry Pi Foundation, n.d.). The Raspberry Pi 4 model B Quad-core ARM Cortex-A72, 8 GB RAM, Wi-Fi and Bluetooth support was chosen in our project for its processing power and ability to run modern AI models.

Arduino Mega



Figure 3.3: Arduino Mega 2560

A microcontroller board based on the ATmega2560. It has 54 digital input/output pins (of which 15 can be used as PWM outputs), 16 analog inputs, 4 UARTs (hardware serial ports), a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started. The Mega 2560 board is compatible with most shields designed for the Uno and the former boards Duemilanove or Diecimila (Arduino, n.d.). The Arduino Mega was selected in our project for its large number of I/O pins and real-time performance in controlling multiple motors and sensor.

3.2.2 Vision System

OAK-D Camera



Figure 3.4: OAK-D Camera

Is the ultimate camera for robotic vision that perceives the world like a human by combining stereo depth camera and high-resolution color camera with an Neural Network inferencing and Computer Vision capabilities. It uses USB-C for both power and USB3 connectivity. 4 TOPS of processing power, two grayscale sensors for stereo depth, and a single center color sensor. It has been directly superseded by the 2nd generation OAK-D-S2 offering more CCM options in a sleeker body, or its Pro version with additional features (Luxonis, n.d.). The OAK-D was chosen in this project because it offers higher accuracy, faster processing, and easier integration compared to standard cameras.

Live Camera (USB Camera)



Figure 3.5: Live Camera (USB Camera)

In addition to the OAK-D camera, a separate USB live camera was integrated into the Thikra robot to capture the final high-quality photographs. While the OAK-D camera provides excellent performance for person detection, localization, and depth-based vision tasks, its image output is not optimized for producing high-resolution portrait photos.

Therefore, the USB live camera was selected as the primary photo-capturing device due to its improved image quality and sharper visual results. This camera is responsible for acquiring the final images displayed to the user and sent to the printing system, ensuring better overall photo quality and user satisfaction.

3.2.3 Motion and Actuation

DC Gear Motor



Figure 3.6: DC Gear Motor used in the Thikra Robot

DC gear motors were used in the Thikra robot to provide controlled and reliable mechanical motion. The selected motor model (25GA-370) operates at 12 V with a rated speed of 60 RPM and includes an integrated gearbox that increases torque while reducing rotational speed.

The gearbox allows the motor to deliver sufficient torque to move the robot smoothly and support its overall weight during operation. This makes the motor suitable for mobile robotic applications that require stable and precise movement rather than high speed.

In the proposed system, DC gear motors are responsible for driving the omni wheels and enabling the robot to move forward, backward, and rotate as needed. The motors are controlled through a motor driver module, which allows the system to regulate speed and direction according to user positioning and navigation requirements.

Aluminum Omni Wheels



Figure 3.7: Aluminum Omni Wheels used for multidirectional movement

Aluminum omni wheels were integrated into the robot to enable multidirectional movement. Unlike traditional wheels, omni wheels allow the robot to move laterally, diagonally.

This feature enhances the maneuverability of the robot, allowing it to adjust its position smoothly and accurately in confined indoor spaces. The use of aluminum material provides durability and mechanical stability while supporting the overall weight of the system.

L298N Motor Driver Module



Figure 3.8: L298N Motor Driver Module used to control DC motors

The L298N motor driver module was used to control the DC motors. This module allows the microcontroller to regulate both the speed and direction of the motors by providing adequate current and voltage levels.

The L298N driver serves as an interface between the Arduino Mega and the DC motors. It ensures reliable motor operation and protects the control circuitry from high-current loads, making it suitable for robotic motion control applications.

3.2.4 Sensors

Ultrasonic Sensors (HC-SR04)



Figure 3.9: HC-SR04 Ultrasonic Sensor used for obstacle detection

The HC-SR04 ultrasonic distance sensor is a low-cost and reliable module that provides non-contact distance measurements in the range of approximately 2 cm to 400 cm, with an accuracy that can reach up to 3 mm. The sensor operates by emitting an ultrasonic pulse through its transmitter and measuring the time required for the echo to return to the receiver.

Each HC-SR04 module consists of an ultrasonic transmitter, an ultrasonic receiver, and a control circuit. The sensor has four main pins: VCC (power supply), Trig (trigger input), Echo (echo output), and GND (ground) (SparkFun Electronics, n.d.).

In the Thikra photographer robot, multiple ultrasonic sensors were integrated to enhance safety and navigation. These sensors are used to detect nearby obstacles and surrounding walls, preventing collisions during robot movement and ensuring safe operation in indoor environments.

3.2.5 User Interaction Components

iPad Display



Figure 3.10: iPad display used for photo preview and user selection

An iPad screen was integrated to preview captured images and allow the user to select the desired photo before printing.

Keypad



Figure 3.11: Keypad interface used for password entry and session configuration

The keypad provides a simple interface for entering passwords, selecting the number of photos, and confirming printing options.

Speakers



Figure 3.12: Speaker module used for audio guidance and user prompts

Speakers were used to provide audio prompts guiding the user throughout the photo session.

Flash Control (Relay Module)



(a) Relay



(b) Flash

Figure 3.13: Relay-controlled flash system used for lighting enhancement

A flash unit ensures consistent lighting conditions and improves image quality during photo capture, especially in low-light environments.

In the Thikra robot, the flash is controlled through a one-channel relay module, which allows the microcontroller to switch the lighting on and off automatically when needed. This approach provides safe isolation between the high-power lighting circuit and the control electronics, ensuring reliable and synchronized flash activation during image capture.

Seven Segment Display

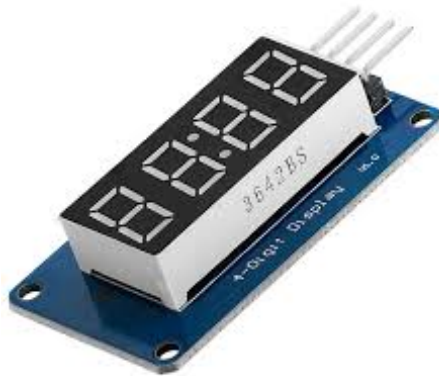


Figure 3.14: Seven Segment Display used to show remaining shots

The seven-segment display shows the remaining number of photos during the capture process.

Printer



Figure 3.15: Compact printer used for instant photo printing

A compact printer was integrated to produce instant physical copies of the selected image.

3.2.6 Power System

Power Bank



Figure 3.16: Power bank used to supply stable power for the Raspberry Pi

The Moxom PB-29 PD is a 20,000 mAh power bank with USB-C and USB-A ports that supports fast charging (PD and QC3.0). It gives a stable 5V, 3A output, which is exactly what a Raspberry Pi needs, and its large capacity lets the Pi run for many hours without recharging.

Lithium Battery Pack



Figure 3.17: Lithium battery pack used as the main power source for motors

Rechargeable lithium batteries were used as the primary energy source for powering the robot's motors and high-current modules. Lithium-based batteries offer high energy density, lightweight design, and sufficient discharge capability, making them suitable for mobile robotic platforms.

Buck Converter (LM2596)



Figure 3.18: LM2596 buck converter module used for voltage regulation

The LM2596 buck converter was integrated to step down voltage levels when necessary and provide regulated 5V power for sensors and control electronics. This ensures stable and safe operation for sensitive components.

Power Distribution

The system includes a power distribution mechanism that separates motor power from control electronics. Different voltage levels are regulated and supplied to the Raspberry Pi, microcontrollers, sensors, and actuators to ensure reliable performance and prevent voltage drops during movement.

3.3 Software

The software architecture of **Thikra** integrates computer vision processing, motor control, and communication between system modules to provide an autonomous and interactive photography experience.

- **AI Models (Person and Face Detection):** Pre-trained neural network models were deployed through the DepthAI framework, including person detection, face detection, and facial landmarks extraction. These models enable the robot to locate the user, ensure proper alignment, and verify facial readiness (eyes open, looking forward, and smiling) before capturing photos.
- **DepthAI SDK:** The DepthAI software development kit was used to interface the OAK-D camera with the Raspberry Pi. This allows real-time vision processing, user localization, and depth-based positioning, supporting accurate centering and tracking of the subject.
- **OpenCV Photo Capture Module:** OpenCV was used to operate the USB Live Camera, which serves as the main device for capturing the final high-quality images. The captured photos are saved and prepared for preview and user selection.

- **Arduino Firmware:** The Arduino Mega was programmed using the Arduino IDE and relevant libraries to control DC motors, ultrasonic sensors, keypad inputs, and relay-based flash activation. Arduino handles real-time hardware control tasks during the photography session.
- **User Interface and Web Display:** A lightweight local web-based interface was hosted on the Raspberry Pi to display captured images on the iPad. The interface refreshes automatically after each capture, allowing users to preview all photos, select their preferred image, and proceed with printing options.
- **Web Application** In addition to the physical keypad mounted on the robot, a lightweight web-based interface was developed and displayed on an iPad to provide a more convenient and interactive user experience. This interface allows users to preview the captured images in real time, browse through the session photos, and select their preferred shot easily.

Furthermore, a dedicated mobile printing application was integrated into the system workflow. Once the user selects the desired image on the iPad, the chosen photo is automatically transferred to the smartphone through the printing application, enabling instant printing without requiring manual file handling.

This approach enhances usability and accessibility in public environments by providing a smooth photo selection and printing process, while ensuring that all actions remain synchronized with the robot's photography session in real time.

- **System Communication:** Serial UART communication was established between the Raspberry Pi and Arduino Mega to synchronize movement commands, audio prompts, capture triggers, and user selections. HTTP requests are also used to instantly update the iPad display, ensuring smooth interaction and workflow continuity.

3.4 Workflow

The **Thikra Photographer Robot** is designed to provide a fully autonomous and interactive photography experience. Its operation follows a structured workflow that integrates computer vision, motor control, user guidance, and real-time photo preview. The overall session is executed through the following sequence:

1. Initialization and System Setup

At startup, the robot initializes all hardware and software modules.

- The Raspberry Pi loads the main Python application and starts the computer vision pipeline through the DepthAI framework.
- The Arduino Mega initializes the DC motors, ultrasonic sensors, keypad interface, flash relay, and audio modules.
- Serial UART communication is established between Raspberry Pi and Arduino to synchronize commands during operation.

Once initialization is complete, the system enters standby mode and displays a readiness message to indicate successful startup.

2. User Authentication and Session Configuration

Before starting the photography session, the user interacts with the system through the keypad mounted on the robot. The user enters a secure password to begin the session, then selects the required number of photos. This step ensures controlled access to the system and provides customized session settings according to the user's needs.

3. User Detection and Position Alignment

After the session is configured, the robot activates its vision system to detect and localize the user in real time. Using the OAK-D camera with DepthAI-based person detection, the system continuously analyzes each frame and extracts the user's bounding box. The horizontal center of the detected person is used to determine whether the user is positioned to the left or right of the camera frame, while the bounding box size is used as an indicator of distance. Based on these measurements, the Raspberry Pi sends motion commands to the Arduino Mega through serial UART communication, allowing the robot to move left or right to center the user, and forward or backward to maintain an appropriate shooting distance. This alignment process continues until the user remains stable within the desired frame region, ensuring proper framing and reliable positioning before proceeding to face verification and photo capture.

4. Face Readiness Verification

After the robot detects the user and adjusts its position in front of the camera, the system enters an important stage called **face readiness verification**. This step ensures that the captured images are clear and properly framed, avoiding common issues such as closed eyes or the user not facing the camera.

First, the system applies a pre-trained **face detection model** using the OAK-D camera. Once a face is successfully detected, the robot crops the face region from the frame to allow more accurate analysis.

Next, a **Facial Landmarks Model** is executed to extract a set of key feature points on the face, such as points around the eyes, nose, and mouth. These landmarks enable the system to understand the facial state of the user.

Based on the extracted landmarks, the robot verifies three main conditions before capturing a photo:

1. Looking Forward (Face Turn Detection)

To check whether the user is facing the camera, the system monitors the horizontal position of the nose relative to the eyes. When the face is straight, the nose appears near the center between both eyes. If the face is turned left or right, the nose shifts away from the center.

The following ratio is used:

$$ratio = \frac{nose_x - leftEye_x}{rightEye_x - leftEye_x}$$

where:

- $nose_x$ is the horizontal nose position,
- $leftEye_x$ is the left boundary of the eye region,
- $rightEye_x$ is the right boundary of the eye region.

If the ratio is close to 0.5, the face is aligned with the camera. Otherwise, the system detects a head turn and provides an audio prompt instructing the user to face forward.

2. Eye Open Detection (EAR)

To ensure that the user's eyes are open, the system calculates a simplified **Eye Aspect Ratio (EAR)** based on the eye landmark points.

$$EAR = \frac{eyeHeight}{eyeWidth}$$

where:

- $eyeHeight$ represents the vertical distance between the top and bottom eye landmarks,
- $eyeWidth$ represents the horizontal distance between the left and right eye landmarks.

When the eyes are closed, the vertical height becomes very small, causing the EAR value to drop. Therefore, if:

$$EAR < Threshold$$

the system plays a voice message asking the user to open their eyes.

3. Smile Detection (MAR)

To verify whether the user is smiling, the system measures the shape of the mouth using the **Mouth Aspect Ratio (MAR)**.

$$MAR = \frac{mouthHeight}{mouthWidth}$$

where:

- $mouthHeight$ is the vertical distance between the upper and lower mouth landmarks,
- $mouthWidth$ is the horizontal distance between the mouth corners.

During smiling, the mouth shape changes and the MAR value increases. Thus, if:

$$MAR > Threshold$$

the system considers the user to be smiling. Otherwise, it provides an audio prompt asking the user to smile.

Final Decision (Face Ready)

After evaluating the three conditions (face aligned, eyes open, and smiling), the system decides whether the user is ready for photo capture. The readiness condition is expressed as:

$$Ready = (Turn = False) \wedge (EAR > Threshold) \wedge (MAR > Threshold)$$

Once all conditions are satisfied, the robot automatically starts the countdown and captures the photo.

5. Countdown and Photo Capture

After readiness is confirmed, the robot starts the capture sequence.

- Arduino triggers an audio countdown.
- The flash module is activated through a relay
- The USB Live Camera captures high-resolution photos as final outputs.

All images are stored automatically inside a dedicated session folder.

6. Image Preview and User Selection

Once all shots are captured, the robot provides an interactive preview stage.

- Captured images are displayed instantly on the iPad interface.
- The user browses the session images and selects the preferred photo.
- The selected image is marked and prepared for printing.

7. Printing and Session Completion

After selection, the robot finalizes the session.

- The chosen photo is automatically transferred to the smartphone through a dedicated printing application, where it is sent to the printer for instant output.

8. Safety and Error Handling

Throughout operation, safety mechanisms are continuously applied.

- Ultrasonic sensors monitor nearby obstacles to prevent collisions.
- Movement and capture are stopped if unsafe conditions are detected.

This structured workflow enables **Thikra** to operate autonomously in indoor public environments while providing a smooth, safe, and interactive photography experience.

3.4.1 Constraints

During the development of the Thikra Photographer Robot, several constraints were identified:

- **Budget Limitations:** The project was developed with limited financial resources, which restricted the ability to use professional-grade photography systems or advanced robotic platforms. Therefore, affordable components such as standard DC motors, basic sensors, and lightweight camera modules were selected.
- **Processing and AI Performance:** Although the Raspberry Pi 4 provides strong embedded computing capabilities, it still has limited performance compared to dedicated GPUs. This constraint affected the complexity and speed of real-time computer vision tasks such as face detection and landmark extraction.
- **Image Quality Trade-off:** The OAK-D camera offers excellent AI processing features but its captured photo quality was not sufficient for high-resolution photography output. As a result, an additional USB Live Camera was integrated to provide better final image quality.
- **Mechanical Stability and Movement Accuracy:** The robot operates in indoor environments where slight vibrations, uneven floors, or wheel slippage can influence positioning accuracy. Maintaining stable alignment before capturing photos required careful speed control and motion calibration.
- **Lighting Conditions:** Poor or inconsistent lighting in public indoor spaces may reduce image clarity and affect face detection reliability. This constraint led to the inclusion of a flash and LED lighting system controlled through a relay module.
- **Power Supply Constraints:** The robot relied on lithium batteries and power banks, which provide limited current under high-load conditions. During simultaneous operation of motors, cameras, and flash modules, voltage drops and reduced performance were observed.

Chapter 4 — Results and Analysis

The performance of the smart photography robot **Thikra** was evaluated through a series of tests designed to measure its ability to detect users, capture photos automatically, and provide an interactive photography experience under various conditions. The evaluation focused on four main aspects: user detection accuracy, photo capture performance, user interaction and photo selection, and overall system reliability.

4.1 User Detection Accuracy

The robot relies on computer vision techniques using the OAK-D camera to detect the presence of users and estimate the distance between the camera and the person. During the tests, the system demonstrated a high success rate in detecting users in indoor environments such as university hallways and open rooms. The detection process remained stable across different user heights and positions, enabling the robot to initiate the photo capture process automatically. However, in conditions of low lighting or partial obstruction of the user, the confidence level in detection slightly decreased, occasionally resulting in a longer delay before starting photo capture. It was also observed that the system may make errors when multiple people are present in front of the camera simultaneously, or when the space is extremely narrow, making it difficult to accurately identify the target user.

Overall, the vision-based detection system proved efficient for reliable use in controlled indoor environments, with some limitations in multi-person scenarios or very tight spaces.

4.2 Photo Capture Performance

The performance of photo capture was evaluated based on the robot's ability to take the number of shots specified by the user. After selecting the desired number of photos, the system successfully captured all shots sequentially within a single session. The captured images were consistent in framing and focus, benefiting from the camera's fixed stabilization and distance estimation based on the vision system.

The absence of immediate retakes during the session simplified the system's behavior and ensured smooth photo capture without interruption, reducing user confusion and processing burden.

4.3 User Interaction and Photo Selection

User interaction was assessed through the system interface and workflow after photo capture. Once a session ended, all photos were sent and displayed on an iPad, allowing users to easily browse the entire set. Test results showed that users could navigate between images and select their preferred photo without difficulty.

Based on the user's choice, the system either prints the selected photo or ends the session without printing, providing the user with greater control over the final outcome while maintaining a simple and clear interaction process.

4.4 Summary of Results

Overall, Thikra demonstrated reliable ability to detect users, capture photos according to the specified number of shots, and provide effective interaction through photo review and printing options. The integration of computer vision with an interactive photography workflow reduced the need for a human photographer while maintaining a positive user experience.

While the system performed well in indoor environments, future development could focus on improving performance under challenging lighting conditions and expanding usability in more dynamic public spaces.

Chapter 5 — Discussion

The smart photography robot **Thikra** demonstrated a practical and innovative application of robotics, computer vision, and embedded control systems in interactive photography. In this chapter, we interpret the results of the Thikra project relative to the initial objectives, compare them to traditional photography practices, and discuss the robot’s strengths and limitations.

5.1 Problem Resolution

The main problem addressed by the project was the need to facilitate high-quality photo capture in a smart and interactive manner while minimizing direct human involvement. The goal of Thikra was to provide an automated photography experience capable of detecting users and capturing photos automatically, while allowing users to select their final images.

Based on testing and performance evaluation, the following results were achieved:

- The computer vision system using the OAK-D camera was able to accurately detect users in various indoor environments.
- Photos were captured smoothly according to the specified number of shots, maintaining consistent focus and framing.
- The user interface allowed users to review photos and select their preferred images for printing.

5.2 Contribution

The Thikra project provides an important prototype for developing robots in the field of smart and interactive photography. By integrating computer vision, motion control, and user interaction, the robot helps reduce human intervention while offering an easy and enjoyable photography experience.

Although the project is still in its early stages, it opens the door for further development, such as improving performance in low-light conditions and supporting more dynamic and public photography environments. This work demonstrates the potential of robots to play an effective role in simplifying and enhancing interactive photography.

5.3 Logical Implications

The project results indicate that robotics and AI can support the photography process in a more efficient and precise way. Thikra’s ability to detect users and capture photos automatically with minimal human involvement suggests that such systems can improve image quality and reduce common human errors in traditional photography.

Additionally, the photo review interface allows users to browse and select their preferred images, demonstrating that photography processes can be conducted more interactively and flexibly.

However, some limitations were observed, such as performance under low-light conditions or in crowded environments, indicating that further development is needed before the robot can be fully deployed in all real-world photography settings.

Chapter 6 — Conclusions and Recommendations

6.1 Conclusion

This project aimed to design and implement **Thikra**, a smart photographer robot capable of detecting users and moving autonomously in all directions to ensure that the person remains within the camera frame, while capturing high-quality images with minimal human intervention. The integration of AI-based computer vision, navigation sensors, and real-time interaction systems demonstrated the feasibility of developing an autonomous and effective photography platform.

Testing showed that Thikra can accurately detect and track a person in real time and adjust its movement to maintain proper framing. The system allows users to interact through an integrated keypad, display screen, and audio feedback, with the ability to preview and print images instantly. These results confirm that the use of robotics and artificial intelligence in photography can enhance image quality and reduce the need for human involvement.

However, several limitations were observed. The Raspberry Pi 4 restricts the processing speed of advanced AI models, and environmental factors such as low lighting conditions, crowded environments, and uneven surfaces can negatively affect detection accuracy and final image quality.

6.2 Recommendations

Based on the findings of this project, the following recommendations are proposed for future work:

- Use a higher-quality camera with improved resolution to enhance overall image quality.
- Train the AI model on larger and more diverse datasets to improve person detection and tracking under different conditions.
- Improve the robot's physical design to enable easier movement in tight or crowded spaces.
- Enhance the image editing process before printing to improve colors, lighting, and overall print quality.
- Integrate additional sensors such as LiDAR or depth cameras to improve navigation and person tracking in complex environments.
- Implement intelligent lighting solutions to maintain consistent image quality under varying lighting conditions.

Bibliography

- [1] Available: <https://arxiv.org/html/2401.11061v3>
- [2] Available: <https://www.mdpi.com/2673-4591/60/1/19>
- [3] Available: <https://www.sciencedirect.com/science/article/abs/pii/S0921889017308114>
- [4] Available: <https://www.sinoseen.com/robot-camera-self-directed-picture-taking-of-the-future>
- [5] Available: <https://news.cornell.edu/stories/2022/03/robot-photographer-has-eye-aesthetics>
- [6] Available: <https://news.cornell.edu/stories/2022/03/robot-photographer-has-eye-aesthetics>
- [7] Available: <https://www.diyphotography.net/this-photo-robot-knows-composition-and-is-coming>
- [8] Available: <https://skylinefba.com/automated-product-photography-solutions/>
- [9] Available: <https://www.youtube.com/watch?v=-q54cBmLN0A>
- [10] Available: <https://www.proquest.com/openview/3056b48ae368b61cefc030bfc4b9b605/1>
- [11] Available: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1775>
- [12] Available: <https://www.tandfonline.com/doi/abs/10.1080/17540763.2018.1445011>