http://www.jourmag.org/index.php/jma
Journal of Magnetism and Its Applications

Research articles

Design and Fabrication of Remote-controlled 4WD Robot

Muhammad Zain[1], Shoaib Ali[2], Rohail Abbas[3], Sajjad Bhangwar[4] Muhammad Kashif Abbasi[5] Muhammad Ramzan Luhur[6]
[1,2,3,4] Department of Mechanical Engineering Quaid-E-Awam University of Engineering Science and Technology Nawashah, Sindh Pakistan
[5,]Department Industrial and Manufacturing Engineering Quaid-E-Awam University of Engineering Science and Technology Nawashah, Sindh Pakistan
Corresponding author (muhammadkashif@quest.edu.pk)

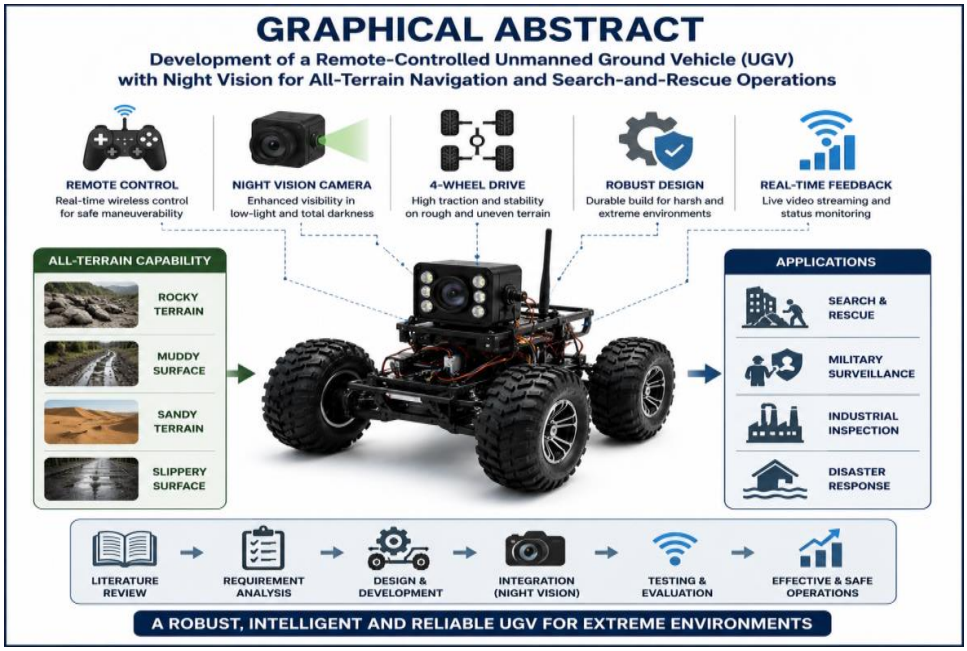
Article info

Keywords:
Unmanned Ground Vehicle (UGV),
Tracked Vehicle, Remote-Controlled Robot,
All-Terrain Navigation,
Rough Terrain,
Night Vision Camera,
Intelligent Sensing

Abstract

The rapid advancement of robotics and intelligent sensing systems has created significant opportunities for automating critical tasks in hazardous or inaccessible environments. Unmanned Ground Vehicles (UGVs) have increasingly become essential platforms for surveillance, inspection, and search-and-rescue operations. This study focuses on the development and application of a remote-controlled, tracked UGV designed to navigate rough, uneven, and slippery terrains where traditional wheeled platforms struggle. The tracked locomotion mechanism provides superior mobility, stability, and traction, ensuring minimal slippage across extreme landscapes. To extend operational capabilities beyond daylight hours, the system is equipped with an integrated night vision camera, ensuring reliable perception and visibility in low-light and nighttime conditions. By combining robust all-terrain navigation with real-time remote maneuverability and advanced sensing, this platform presents a highly effective and safe solution for a wide range of civilian and military applications in extreme environments.

GRAPHICAL ABSTRACT



1.1 INTRODUCTION:

The rapid development of robotics and intelligent sensing systems has created new opportunities for automating tasks in environments that are unsafe or difficult for humans. Unmanned ground vehicles (UGVs) equipped with advanced perception tools are increasingly being used in surveillance, inspection, and rescue operations. Among these platforms, tracked robots offer superior mobility, stability, and traction, enabling efficient movement across rough, uneven, or slippery terrains. Their ability to operate in extreme environments makes them valuable for both civilian and military applications.The tracked mechanism allows the robot to navigate obstacles with minimal slippage, while remote-controlled operation provides safe, real-time maneuverability. The night vision camera enhances visibility during low-light and nighttime conditions, allowing continuous monitoring even in total darkness [1], approximately 7,348 natural disasters were recorded globally during the last two decades, demonstrating the increasing frequency and severity of extreme events. Among these disasters, floods account for nearly 40% of all occurrences, followed by storms at 28% and earthquakes at approximately 8% [2]. Search and Rescue (SAR) operations are among the most challenging phases of disaster management. Rescue teams often operate in hazardous and unpredictable environments characterized by collapsed buildings, unstable terrain, fires, toxic gases, flooding, and poor visibility. These conditions not only delay rescue efforts but also place emergency responders at considerable risk. Consequently, robotics has emerged as a transformative technology capable of improving both the speed and safety of disaster response operations [3–5]. Search and rescue robots are specifically designed to assist human responders by exploring dangerous or inaccessible environments while collecting critical information about victims and structural conditions. These robotic platforms significantly reduce the exposure of rescue personnel to life-threatening hazards such as landslides, aftershocks, debris collapse, and chemical contamination [6–10]. Recent advances in artificial intelligence (AI), computer vision, and deep learning have substantially enhanced the capabilities of search and rescue robots. Machine vision techniques enable robots to autonomously recognize objects, navigate complex environments, and detect victims using visual information. [11–18]. Human bodies naturally emit thermal radiation that can be distinguished from surrounding objects, allowing victims to be detected even when visibility is severely compromised. These characteristics make thermal imaging particularly valuable for post-disaster environments where conventional vision systems often fail. Several studies have explored the use of thermal cameras and infrared sensors for human detection. Researchers have investigated methods based on low-resolution thermal cameras, infrared arrays, and hybrid sensing systems capable of recognizing human presence despite limited image quality [19–22]. Traditional inspection and surveillance robots with wheel-based systems often fail in off-road or uneven terrains. Additionally, standard cameras lack the capability to identify hidden heat-related issues such as overheating, leakage, or electrical faults, especially during night operations. Therefore, there is a need for a robust, terrain-independent robotic system with integrated night and thermal vision capabilities for efficient monitoring and inspection.

RESEARCH METHODOLOGY

The proposed methodology is described as under:

Step-1: (Mechanical Design)

Tracked wheel mechanism for all-terrain mobility.
Metal frame with linkages for tank-like motion.

Step-2: (Electronic Components)

Raspberry Pi as the main controller
Motor driver module for motion control.
RF transmitter-receiver pair for wireless communication.
Thermal and night vision cameras for dual-mode visual input.

Step-3: (Software Development)

Python-based programming for camera integration, motor control, and live video streaming.
Motion detection algorithm for identifying movement in surveillance mode

Step-4: (Control and Communication)

User commands transmitted via RF Joystick Controller (100 m range).
Onboard receiver connected to Raspberry Pi interprets commands and drives motors.
Real-time video streaming from cameras to Android device via Wi-Fi.

Testing and Validation:

- Terrain mobility testing on rough, inclined, and obstacle-filled surfaces.
  Thermal detection tests for heat/leakage identification.
- Night vision effectiveness in low-light conditions.

Wireless range and latency assessment for real-time control.

Main Components:

RF Transmitter & Joystick
RF Receiver
Raspberry Pi Controller
Motor Driver & DC Motors
Thermal Camera
Night Vision Camera
Live Video Transmission Module
Power Supply (Battery)

3.1 ESP32 DEVELOPMENT BOARD:

The ESP32 is a powerful microcontroller developed by Espressif Systems that serves as the brain of the robot. Compared to the Arduino Uno, it offers several advanced features, including built-in Wi-Fi and Bluetooth connectivity, a high-speed processor, larger memory capacity, and multiple communication interfaces. These capabilities make it well suited for robotics and IoT applications. In this project, the ESP32 is responsible for receiving movement commands, controlling the motor drivers, hosting WebSocket communication for real-time data exchange, and reading thermal camera data for monitoring and navigation.

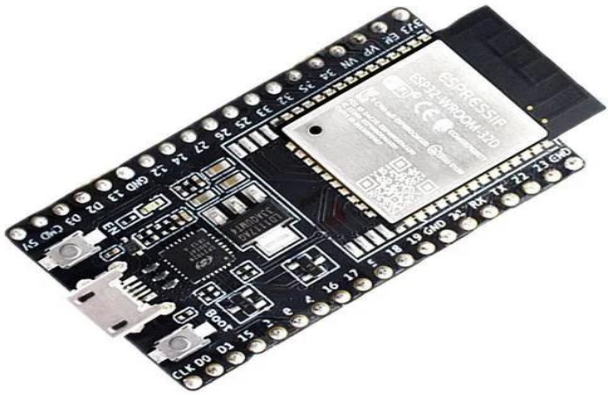


Figure 3.1Esp32 Development Board

3.2 DC GEAR MOTORS:

The robot is powered by four DC gear motors, each consisting of a DC motor coupled with a gearbox. The gearbox reduces the motor's rotational speed while increasing its torque, making the robot stronger, more stable, and easier to control. When a positive voltage is applied, the motor rotates in the forward direction, whereas reversing the voltage causes it to rotate backward. Using four motors provides improved stability, higher torque, and the ability to carry heavier payloads, making the robot suitable for demanding applications.



Figure 3.1DC Gear Motors

3.3 WHEELS:

The robot is equipped with four wheels that convert the rotational motion of the DC gear motors into linear movement. The wheels are arranged into left and right pairs, allowing each side to be controlled independently. This differential drive mechanism enables the robot to move forward, move backward, turn left, turn right, and rotate in place, providing excellent maneuverability in various environments.



Figure 3.2Wheels

3.4 L298N MOTOR DRIVER MODULES:

The robot uses two L298N motor driver modules to control its four DC gear motors. Since the ESP32 microcontroller can only generate low-current control signals, it cannot drive the motors directly. The L298N is an H-Bridge motor driver that amplifies these signals to provide the higher current and voltage required by the motors. Each module controls two motors, so two modules are needed for the four-motor system. In addition to controlling motor direction, the L298N supports Pulse Width Modulation (PWM) for speed control and provides electrical isolation between the motors and the ESP32, improving system reliability.

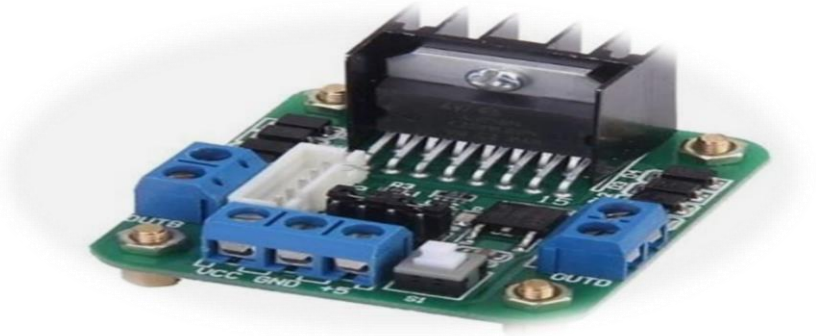


Figure 3.3L298N Motor Driver Modules

3.5 METALLIC LASER-CUT CHASSIC (MILD STEEL):

The robot is built on a laser-cut chassis made from mild steel (MS), which serves as its structural framework. Laser cutting provides precise dimensions and accurate component mounting, while the mild steel construction offers excellent strength, durability, and resistance to vibration. The chassis is capable of supporting heavy payloads and securely holds all major hardware components, including the ESP32, batteries, motor drivers, cameras, sensors, and wiring, ensuring a robust and organized mechanical design.

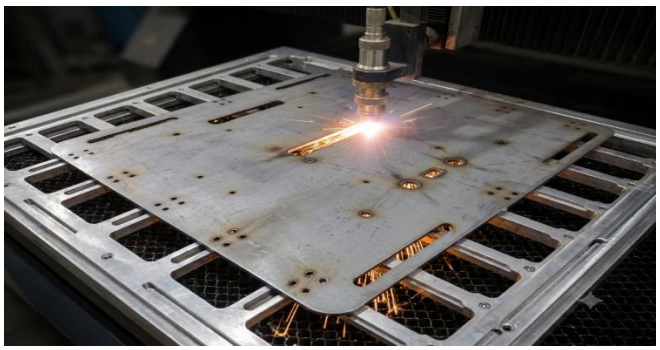


Figure 3.4Metallic Laser-Cut Chassis (Mild Steel)

3.6 IMOU CCTV CAMERA:

An IMOU wireless IP CCTV camera is mounted on the robot to provide live visual monitoring of its surroundings. Unlike sensors connected directly to the ESP32, the camera connects independently to the Wi-Fi network and transmits video using the Real-Time Streaming Protocol (RTSP). It offers high-definition video streaming, night vision capability, a wide viewing angle, and continuous surveillance. This enables remote observation of the robot's environment and supplies video input for computer vision applications.



Figure 3.5IMOU CCTV Camera

3.7 . RTSP (REAL-TIME STREAMING PROTOCOL):

The Real-Time Streaming Protocol (RTSP) is a network protocol used for transmitting live video streams. In this project, the IMOU camera continuously sends video frames over the Wi-Fi network using RTSP. A Python application running on a computer receives these video frames and processes them using the YOLO object detection algorithm. This workflow enables real-time object detection and monitoring without requiring a direct wired connection between the camera and the computer.



Figure 3.6RTSP (Real-Time Streaming Protocol)

3.8 THERMAL CAMERA (AMG8833 IR Thermal Imager Temperature Sensor Module):

The AMG8833 infrared thermal camera measures infrared radiation emitted by objects instead of capturing visible light. Every object emits heat, and the sensor detects temperature variations across its field of view. Rather than producing a conventional image, it generates an 8×8 temperature matrix in which each value represents the temperature of a specific pixel. Warmer objects appear brighter, while cooler objects appear darker in the thermal visualization. This thermal sensing capability is useful for applications such as human detection, fire detection, electrical equipment inspection, heat leakage analysis, and industrial monitoring.

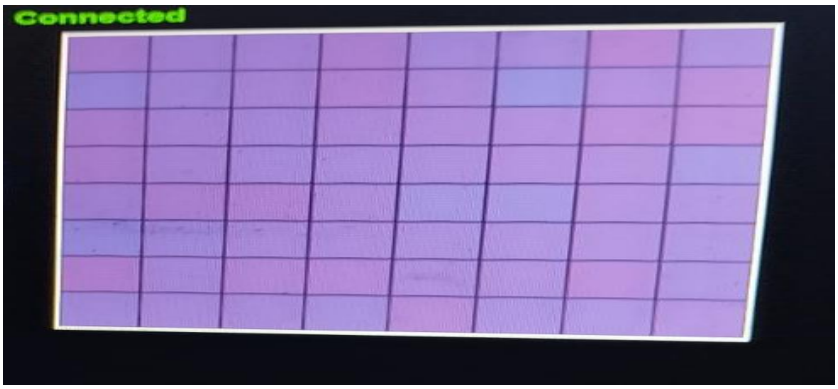


Figure 3.7Thermal Camera

3.9 Wi-Fi COMMUNICATION:

The robot relies on Wi-Fi communication to enable wireless interaction between its components. The ESP32, IMOU CCTV camera, thermal camera, and the monitoring computer are all connected to the same Wi-Fi network. This wireless architecture allows movement commands, live video, thermal data, and control signals to be transmitted in real time without requiring USB cables during operation. As a result, the robot can be remotely controlled while continuously monitoring its surroundings through both visual and thermal imaging.

3.10 WEBSOCKET:

WebSocket is a communication protocol that enables real-time, bidirectional communication between a client and a server over a single, persistent connection. Unlike traditional HTTP communication, where a new connection is established for every request and response, WebSocket keeps the connection open, allowing data to be exchanged instantly with minimal delay. This makes it highly suitable for applications that require continuous, low-latency communication. In this project, WebSocket is used as the primary communication channel between the robot and the control application. It transmits movement commands such as forward, backward, left, right, and stop from the user interface to the ESP32. In addition, the ESP32 sends thermal sensor readings, robot status information, and other real-time updates back to the monitoring application. This continuous two-way communication ensures smooth robot control and immediate feedback during operation.

3.11 PYTHON:

Python serves as the main processing and control software for the robotic system. It acts as an interface between the robot hardware, artificial intelligence algorithms, and the user. Due to its simplicity, extensive library support, and compatibility with machine learning frameworks, Python is widely used for robotics and computer vision applications. In this project, Python receives the live RTSP video stream

from the IMOU CCTV camera and processes each video frame using the YOLO object detection model. The processed video, including detected objects and bounding boxes, is displayed in real time for monitoring purposes. Python also reads temperature data from the AMG8833 thermal camera, processes the thermal information, and updates the WebSocket server with sensor readings and robot status. As a result, Python functions as the central processing unit of the software system by integrating video processing, thermal monitoring, artificial intelligence, and communication services.

3.12. YOLO (You Only Look Once):

YOLO (You Only Look Once) is a deep learning–based object detection algorithm designed for fast and accurate real-time detection of multiple objects within an image or video. Unlike traditional object detection methods that perform multiple processing stages, YOLO analyzes the entire image in a single pass through a neural network. This approach significantly improves processing speed while maintaining high detection accuracy, making it suitable for real-time robotics applications. YOLO is capable of recognizing a wide variety of object classes, including humans, vehicles, animals, furniture, bags, and many other everyday objects. In this project, the model is primarily configured for human detection. The live video stream received from the IMOU camera is processed by the YOLO model, which identifies people within the scene and draws bounding boxes around detected individuals. This enables the robot to monitor human presence in real time and supports applications such as surveillance, security monitoring, search and rescue, and autonomous navigation.

3.13 Human Detection Process

Camera Frame
↓
Python
↓
YOLO
↓
Human Found
↓
Draw Bounding Box
↓
Display Result

The detected person appears inside a colored rectangle with a confidence score.

3.14 SOFTWARE USED:

- Arduino IDE
- Python
- OpenCV
- YOLO
- WebSocket Server
- RTSP Client
- ESP32 Wi-Fi Library

3.15 ROBOT MOVEMENT:

The robot uses differential drive.

Forward

All four motors rotate forward.

↑ ↑
↑ ↑

Backward

All motors rotate backward.

↓ ↓
↓ ↓

Left Turn

Left motors stop or reverse. Right motors move forward.

Right Turn

Right motors stop or reverse. Left motors move forward

Rotate

Left wheels move forward.
Right wheels move backward. Robot spins around its center.

RESULTS AND DISCUSSIONS



Figure 4.1Final Project

4.1 COMPLETE WORKING OF THE PROJECT:

Step 1
Power is supplied to the robot. The ESP32 boots up.
Step 2
ESP32 connects to the local Wi-Fi network
Step 3
The WebSocket server becomes active. The control webpage connects to the robot.
Step 4
The user opens the control interface on a laptop connected to the same Wi-Fi network.
Step 5
When the user presses a movement button (Forward, Backward, Left, Right, or Stop), the command is sent through the WebSocket connection to the ESP32.
Step 6
The ESP32 receives the command and drives the appropriate L298N motor driver outputs, causing the four DC gear motors to move the robot in the requested direction.
Step 7
At the same time, the mounted IMOU CCTV camera continuously streams live video over the local network using RTSP. A Python application receives this video stream, processes each frame with the YOLO object detection model, and highlights detected humans by drawing bounding boxes with confidence scores.
Step 8
Simultaneously, the thermal camera captures infrared temperature data as a matrix of temperature values. This data is transmitted over Wi-Fi and displayed in real time on the WebSocket-based control interface, allowing the operator to observe temperature distribution alongside the normal video feed.
Step 9
The operator monitors both the AI-processed camera feed and the thermal display while driving the robot remotely, enabling safe inspection of the environment from a distance.

CONCLUTION AND FUTURE WORK

5.1 CONCLUTION:

The developed ESP32-based surveillance robot successfully integrates embedded systems, wireless communication, computer vision, thermal sensing, and artificial intelligence into a single mobile robotic platform. The ESP32 microcontroller serves as the central controller, managing robot movement through two L298N motor driver modules while maintaining real-time communication with the control interface using WebSocket technology. An IMOU IP camera provides continuous live video streaming via RTSP, which is processed by a Python application using the YOLO object detection algorithm to detect and identify humans in real time. Additionally, the AMG8833 thermal camera captures infrared temperature data, enabling the operator to monitor heat distribution alongside the visual camera feed. The integration of wireless robot control, AI-based human detection, and thermal imaging significantly enhances the robot's ability to perform remote monitoring and environmental inspection tasks. This combination provides improved situational awareness and allows the system to operate effectively in scenarios where direct human access may be difficult or unsafe. Overall, the project demonstrates the practical application of modern robotics, Internet of Things (IoT) technologies, artificial intelligence, and computer vision in developing an intelligent surveillance system. The robot can be applied in areas such as security surveillance, industrial inspection, search and rescue operations, hazardous environment monitoring, research, and STEM education. Furthermore, the modular design of the system provides a solid foundation for future enhancements, including autonomous

navigation, obstacle avoidance, path planning, cloud-based monitoring, and advanced AI capabilities, making it a scalable platform for next-generation intelligent robotic systems.

5.2 FUTURE IMPROVEMENTS:

- GPS-based outdoor navigation
- Cloud-based remote access over the internet
- Gas, smoke, and environmental sensors
- Automatic return-to-home function
- Battery level monitoring and low-power alerts

REFERNCES:

UNDRR Home. Available online: https://www.undrr.org/ (accessed on 23 June 2021).

Noticias ONU: Pese al Aumento de las Amenazas de Origen Natural en el Siglo XXI, los Países Siguen "Sembrando las Semillas de su Destrucción". 2020. Available online: https://news.un.org/es/story/2020/10/1482242 (accessed on 23 June 2021).

Drew, D.S. Multi-Agent Systems for Search and Rescue Applications. *Curr. Robot. Rep.* 2021, *2*, 189–200. [Google Scholar] [CrossRef]

Delmerico, J.; Mintchev, S.; Giusti, A.; Gromov, B.; Melo, K.; Horvat, T.; Cadena, C.; Hutter, M.; Ijspeert, A.; Floreano, D.; et al. The current state and future outlook of rescue robotics. *J. Field Robot.* 2019, *36*, 1171–1191. [Google Scholar] [CrossRef]

Queralta, J.P.; Taipalmaa, J.; Pullinen, B.C.; Sarker, V.K.; Gia, T.N.; Tenhunen, H.; Gabbouj, M.; Raitoharju, J.; Westerlund, T. Collaborative multi-robot systems for search and rescue: Coordination and perception. *arXiv* 2020, arXiv:2008.12610. [Google Scholar]

Pozniak, H. Robots… assemble! Although humans and highly trained dogs will always be critical in search-and-rescue operations, robots are being developed—Taking inspiration from the natural world—That are helping rescue teams save lives. *Eng. Technol.* 2020, *15*, 67–69. [Google Scholar] [CrossRef]

Shah, B.; Choset, H. Survey on urban search and rescue robots. *J. Robot. Soc. Jpn.* 2004, *22*, 582–586. [Google Scholar] [CrossRef]

Davids, A. Urban search and rescue robots: From tragedy to technology. *IEEE Intell. Syst.* 2002, *17*, 81–83. [Google Scholar]

Paoletti, M.; Haut, J.; Plaza, J.; Plaza, A. A Comparative Study of Techniques for Hyperspectral Image Classification. *Rev. Iberoam. Autom. Inform. Ind.* 2019, *16*, 129–137. [Google Scholar]

Chadwick, R.A. The impacts of multiple robots and display views: An urban search and rescue simulation. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Orlando, FL, USA, 26–30 September 2005; pp. 387–391. [Google Scholar]

John, V.; Mita, S.; Liu, Z.; Qi, B. Pedestrian detection in thermal images using adaptive fuzzy C-means clustering and convolutional neural networks. In Proceedings of the 2015 14th IAPR International Conference on Machine Vision Applications (MVA), Tokyo, Japan, 18–22 May 2015; pp. 246–249. [Google Scholar] [CrossRef]

Alsafasfeh, M.; Abdel-Qader, I.; Bazuin, B.; Alsafasfeh, Q.; Su, W. Unsupervised fault detection and analysis for large photovoltaic systems using drones and machine vision. *Energies* 2018, *11*, 2252. [Google Scholar] [CrossRef] [Green Version]

Tammana, A.; Amogh, M.; Gagan, B.; Anuradha, M.; Vanamala, H. Thermal Image Processing and Analysis for Surveillance UAVs. In *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*; Springer: Singapore, 2021; pp. 577–585. [Google Scholar]

Tamboli, M.S.M.J.; Desai, K.R. Approach of Thermal Imaging as a Facial Recognition. *JournalNX* 2016, *2*, 1–4. [Google Scholar]

Wang, W.; Zhang, J.; Shen, C. Improved human detection and classification in thermal images. In Proceedings of the 2010 IEEE International Conference on Image Processing, Hong Kong, China, 26–29 September 2010; pp. 2313–2316. [Google Scholar]

Ilikci, B.; Chen, L.; Cho, H.; Liu, Q. Heat-Map Based Emotion and Face Recognition from Thermal Images. In Proceedings of the 2019 Computing, Communications and IoT Applications (ComComAp), Shenzhen, China, 26–28 October 2019; pp. 449–453. [Google Scholar]

Park, J.; Chen, J.; Cho, Y.K.; Kang, D.Y.; Son, B.J. CNN-based person detection using infrared images for night-time intrusion warning systems. *Sensors* 2020, *20*, 34. [Google Scholar]

Portmann, J.; Lynen, S.; Chli, M.; Siegwart, R. People detection and tracking from aerial thermal views. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; pp. 1794–1800. [Google Scholar]

Cerutti, G.; Prasad, R.; Farella, E. Convolutional neural network on embedded platform for people presence detection in low resolution thermal images. In Proceedings of the ICASSP 2019-IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019. [Google Scholar]

Gomez, A.; Conti, F.; Benini, L. Thermal image-based CNN's for ultra-low power people recognition. In Proceedings of the 15th ACM International Conference on Computing Frontiers, Ischia, Italy, 8–10 May 2018; pp. 326–331. [Google Scholar]

Cerutti, G.; Milosevic, B.; Farella, E. Outdoor People Detection in Low Resolution Thermal Images. In Proceedings of the 2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech), Split, Croatia, 26–29 June 2018. [Google Scholar]

Jiménez-Bravo, D.M.; Mutombo, P.M.; Braem, B.; Marquez-Barja, J.M. Applying Faster R-CNN in Extremely Low-Resolution Thermal Images for People Detection. In Proceedings of the 2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applications (DS-RT), Prague, Czech Republic, 14–16 September 2020; pp. 1–4. [Google Scholar]