An IoT-based animal detection system using an interdisciplinary approach

Kiran Kumar Mamidi1*, Sai Nishwanth Valiveti1, Guru Charan Vutukuri1, Ashwin Kumar Dhuda1, Haider Alabdeli2, Rakesh Chandrashekar3, Sorabh Lakhanpal4 and Praveen5

1Department of AIMLE, GRIET, Hyderabad, Telangana, India.
2The Islamic university, Najaf, Iraq
3Department of Mechanical Engineering, New Horizon College of Engineering, Bangalore, Karnataka, India.
4Lovely Professional University, Phagwara, Punjab, India.
5Lloyd Institute of Engineering & Technology, Knowledge Park II, Greater Noida, Uttar Pradesh, India.

Abstract. Nowadays, educational institutions particularly colleges, engaged with students and staff, frequently confront various security challenges in their day-to-day activities. One prominent concern involves the threat of animal bites on the campus. In response to this issue, campus management has traditionally resorted to human patrols and physical barriers to deter animals. To address this multifaceted security challenge, the proposed method "An IoT-based Animal Detection System Using Interdisciplinary Approaches" introduces an innovative solution that leverages the power of IoT technology to enhance campus safety and security significantly. The system deploys a surveillance robot equipped with ultrasonic sensors and ESP32 cameras, employing the machine learning technique R-CNN for Animal Detection. This proposed method uses an interdisciplinary approach to develop an animal detection system capable of identifying and classifying various species. This proposed method aims to revolutionize campus security by seamlessly integrating advanced technology, mitigating risks proactively, streamlining processes through automation, and presenting a cost-effective alternative to traditional security approaches. Beyond the traditional methods, the proposed system achieves an impressive accuracy rate of animal detection approximately 97.6% enabling real-time alerts through push notifications to security personnel upon detection.

1 Introduction

In bustling university corridors, where the pursuit of knowledge intersects with the vibrant world of nature, animal encounters hold both wonder and potential concern. The presence of wildlife on campus can present both opportunities for enrichment and potential risks to

* Corresponding author: mamidi.kirankumar09@gmail.com
student safety. While traditional methods of animal detection exist, they often face limitations in real-world scenarios. Existing approaches may struggle with accuracy in challenging conditions like poor lighting, occlusions (objects blocking the view), and diverse animal appearances. Additionally, complex computational models used in some detection methods can strain resources, hindering real-time deployment on devices commonly used in the field. To address these shortcomings, an interdisciplinary IoT-based animal detection system leverages the combined power of Arduino Uno, ESP32 Camera, ultrasonic sensors, and GPS modules. This system aims to provide a new level of vigilance and responsiveness for animal monitoring in academic environments. The solution offers enhanced accuracy, discreet detection, real-time alerts through LCDs and buzzers, potential species recognition, and a detailed understanding of animal behavior patterns on campus. Beyond its immediate safety applications, the system could contribute to valuable ecological research on campus biodiversity and animal movement patterns. This innovative system promises not only to mitigate risks but also to foster a safer, more harmonious coexistence between humans and wildlife within the educational environment.

2 Existing Methods

S.K.L. and A. Edison [1] leverage depth-wise separable convolutions in their model, achieving an impressive 0.878 Intersection Over Union (IoU) for detection and 99.6% classification accuracy for identifying wild animals. Similarly, Ma and Yang [3] enhance YOLOv5 with weighted BiFPN and ECA modules, pushing accuracy to 95.5%. These studies showcase the power of deep learning for precise animal identification. N. Mamat et al. [4] demonstrate the effectiveness of YOLOv5 on Raspberry Pi for real-time farm protection. Their system detects four animal categories with a remarkable 94% mean average precision (mAP), highlighting the feasibility of IoT-based solutions for animal intrusion monitoring. Further afield, R. Viji et al. [5] combine YOLOv4 and DeepSORT for 98% mAP detection and animal tracking in forest corridors, promoting safe co-existence between humans and animals.

D. Yudin et al. [6] tackle the challenge of detecting large animals in road scenes. Their specialized data and various neural networks, with YOLOv3 leading the pack at 0.78 mAP and 35 fps for ten animal classes, pave the way for improved road safety. Meanwhile, N. Li et al. [9] address the difficulty of detecting animals behind cages. Their M2Det and transfer learning approach achieve improved accuracy with cage-specific training data, highlighting the importance of tailored solutions for unique scenarios. J.J. Daniel Raj et al. [7] utilize Raspberry Pi, camera, and YOLOv5 to detect animals and alert vehicles in forest corridors, potentially reducing accidents. P. Ji and Q. Zhu [8] embed a YOLOv1 model onto K210 chip for portable animal detection, aiding ecological research and animal protection efforts. These studies showcase the potential of IoT for real-time intrusion monitoring and enhanced wildlife protection.

B.K.N. and S.K. [10] implement a smart IoT farm system using convolutional neural networks (CNNs) for irrigation and animal intrusion detection. This combination not only enhances farm efficiency but also safeguards crops and livestock. P.K. Panda et al. [11] propose a similar system with an ultrasonic sensor and camera to detect wild animal intrusions on farms, sending alerts to farmers via an app, further securing agricultural environments. These studies reveal the immense potential of deep learning and IoT for accurate animal detection, real-time intrusion prevention, and improved human-animal interaction. Addressing specific challenges like cage occlusion and integrating advanced technologies like ESP32 cameras (as showcased by P. Manikandan et al. [2] can further refine these systems for robust and user-friendly monitoring solutions, benefitting both wildlife and human communities.
Existing animal detection methods often rely on cameras or sensors and may use deep learning models. However, in the dynamic university campus environment, these methods can face accuracy limitations due to factors like varying lighting, obstructions, and diverse animal species. Additionally, complex deep learning models may be computationally demanding for the types of portable or low-cost devices ideal for discreet campus monitoring. The proposed IoT Animal Detection seeks to address these challenges by emphasizing readily available components (Arduino Uno, ESP32 Camera) and incorporating multiple sensor types (ultrasonic, GPS). This approach aims to deliver greater accuracy and reliability for animal detection within the unique context of a university setting.

3 Proposed Method

In this section problem statement, objective, system model, and module descriptions are discussed.

3.1 Problem Statement

The existing security measures employed by colleges to prevent animal bites on campus, such as human patrols and physical barriers, are costly and ineffective. This poses a risk to the safety and security of students and staff. To enhance campus security and mitigate animal-related threats, there is a critical demand for the implementation of an IoT-based animal detection system. This system would employ surveillance robots equipped with motion sensors and cameras for image recognition, integrated with the machine learning technique R-CNN. The IoT system would identify and track animals on campus and provide real-time alerts to the college community. By deploying this system, colleges can significantly improve the safety and security of their campuses, safeguarding students, and staff from animal-related risks.

3.2 Objectives

- To develop a system capable of detecting and tracking animals in campus corridors using a surveillance robot.
- To achieve high accuracy in animal detection, minimizing false positives and negatives.
- To generate clear and concise SMS notifications informing the college community about detected animals.

3.3 Architecture Diagram

The Camera and Ultrasonic Sensors detect the presence of animals in the college corridors while the GPS Module tracks the location of the animals. The Arduino UNO Microcontroller receives the data from the Camera Ultrasonic Sensors and the GPS Module, and it sends the detected animal information to the Database and the GSM Module. The Database stores the detected animal information for future analysis and reporting. The GSM Module sends an SMS alert to the user with the detected animal information and its location as illustrated in Fig. 1.

3.4 Modules and its Description

The proposed work has a Hardware Module, Machine Learning Module, and Notification Module, and its connectivity is illustrated in Fig. 2.
3.4.1 Module 1: Hardware Module

The Hardware Module consists of Arduino Uno and ESP32 Camera modules which act as the physical foundation for the animal detection robot. It provides the essential hardware components, manages their interactions, and facilitates image capture, data transmission, and user notification. This module forms the critical link between the robot's sensors, the machine learning module, and its user interaction capabilities.
Arduino Uno [17] serves as the central hub for hardware interactions and control within the module. Its primary tasks include Pin Initialization setting up the communication channels with the ESP32 camera and user notification components. Communication Protocol Implementation establishes a serial connection to receive object data from the machine learning module. User Notification Control activates visual alerts based on the received data.

ESP32 Camera [15] module is responsible for image acquisition and transmission in which the camera initialization is set up for capturing images of the environment. Image Configuration adjust image settings to optimize image quality and transmission efficiency. Image Transmission acts as sending captured images to the Arduino Uno for further processing.

The robot's hardware duo captures and interprets its surroundings. The ESP32 camera acts as its keen eye, transmitting captured images to the analytical mind of the Arduino Uno. This central processing unit then receives crucial object data from the machine learning module, pinpointing the presence and species of any animals within the captured frames. In response, the Arduino Uno triggers a clear and concise notification protocol, utilizing an LCD for visual cues, a buzzer for audible alerts, and even SMS messages [18] for remote awareness. This seamless exchange of information ensures users stay informed without any embellishment or unnecessary distractions.

3.4.2 Module 2: Machine Learning Module

The image capturing and preprocessing begins with the ESP32 camera capturing raw images of the robot's environment. Module 2 then meticulously preprocesses these images, adjusting their dimensions, normalizing values, and performing any necessary transformations to ensure compatibility with the R-CNN [12] models. This preprocessing stage lays a solid foundation for accurate visual interpretation. General Object Detection with R-CNN leveraging the user-friendly Google's Teachable Machine framework as illustrated in Fig. 3, Module 2 implements an R-CNN model to analyze the preprocessed images. It identifies a wide range of objects within the scene, pinpointing their locations and defining their boundaries using bounding boxes. This creates a comprehensive visual map of the robot's surroundings, providing a foundational understanding of the environment's layout.

![Teachable Machine](image)

Fig. 3. Training Model in Teachable Machine
Animal Detection and Classification with Custom-Trained R-CNN to achieve a more refined level of visual understanding is done by Module 2 which deploys custom-trained R-CNN models specifically designed for animal detection [13]. These models meticulously scan the objects detected in the previous stage, searching for the presence of animals. If an animal is detected its species is accurately classified which provides valuable insights to the monitoring system as illustrated in Fig. 4.

Knowledge integration and action is the process of extensive visual information, encompassing both general object detection and animal classification, consolidated and relayed to the Arduino Uno. As the robot's central control unit, the Arduino Uno utilizes this integrated knowledge to make informed decisions and execute appropriate actions within its environment. This could involve navigating through obstacles, approaching, or avoiding specific animals, or interacting with their surroundings more adaptively and intelligently. The module's performance is continuously evaluated and refined through ongoing data analysis and model updates. New datasets and algorithmic advancements are incorporated to enhance accuracy, expand capabilities, and foster the robot's evolving visual perception. This ongoing process ensures that the robot's understanding of its environment continues to grow and improve over time, enabling it to adapt to new challenges and scenarios.

Fig. 4. Animal Detection and Classification

3.4.3 Module 3: Notification Module

User Notification acts as the robot's sound, seamlessly translating machine insights into clear and actionable alerts for users. Upon receiving object data from the Object Detection Module, the Arduino Uno, serving as the module's central hub, activates a range of notification mechanisms. Along with the notification LED illuminations for instant signal detection events, LCDs, or external screens present more detailed information such as animal type, location, and detection time, offering a visual overview of the robot's findings. A buzzer in the notification module sounds to capture attention in dynamic environments or when visual...
cues may be less noticeable, ensuring users remain informed even when their focus is elsewhere.

Remote Communication is provided in Module 3 for enhanced awareness and remote monitoring. A GSM module [19] is equipped in the notification module which empowers the robot to send SMS notifications directly to users' mobile devices, enabling them to stay connected to the robot's activities even when not in direct proximity. Through this versatile blend of notification methods, Module 3 ensures that users stay informed of the robot's animal detections, fostering timely responses and informed decision-making.

4 Results and discussions
In this section description of a dataset, experimental results, and the significance of the are discussed.

4.1 Dataset and its Description
The proposed method is built upon a meticulously curated dataset comprising 500 images sourced from various online platforms. Unlike publicly available datasets, this collection is custom crafted to align with the specific objectives of the paper, focusing on two distinct animal species: dogs and elephants. Each image in the dataset offers a unique visual perspective, contributing to a rich tapestry for model training. The dataset's strength lies in its diverse imagery, encompassing variations in resolutions, lighting conditions, and backgrounds. This intentional diversity fosters model resilience and adaptability, enabling it to effectively detect animals in a wide range of real-world environments. This holistic approach to dataset curation and annotation sets the foundation for a robust and versatile model capable of meeting the objectives.

Fig. 5. Sequence of Feature Extractor and Classifier
An essential aspect of the dataset is the precise annotation facilitated by the LabelImg tool. Each image undergoes meticulous labeling, involving the definition of accurate bounding boxes around individual animals and clear class labels that unambiguously identify the species. This annotation process ensures optimal model comprehension and facilitates
enhanced learning. To ensure robust model evaluation, the dataset is strategically divided into three subsets. The training set, comprising approximately 70% of the images, serves as the primary resource for model learning and refinement. The validation set, constituting approximately 15% of the images, aids in model evaluation during training, guiding parameter adjustments to prevent overfitting. The testing set, also around 15% of the images, provides an unbiased assessment of the model's real-world performance upon completion of training.

Fig. 5. below shows how the sequence of feature blocks and classifier [14] on top transfer the information from the raw image and predict requested target values.

4.2 Experimental Results

A brief explanation of the Experimental Result of An IoT-based Animal Detection System is explained in this section.

Robot movement and navigation represent another critical facet along with animal detection. Evaluating the robot's responsiveness to SMS commands for directional movement (left, right, stop, forward, backward) is paramount. The examination should encompass the smoothness and accuracy of turning and stopping maneuvers. Furthermore, the robot's ability to detect and circumvent obstacles using Ultrasonic Sensor data demands scrutiny, encompassing both reaction time and the efficacy of obstacle avoidance maneuvers. The Ultrasonic Sensor's accuracy in measuring distances to obstacles of distinct sizes and shapes is pivotal, considering potential interference from background noise or environmental factors. The robustness of the system's animal detection capabilities, coupled with the high accuracy rate, strengthens its overall performance and reliability.

It is also crucial to evaluate the connectivity and remote monitoring capabilities of the ESP32 Camera. The association of the camera with a designated Wireless SSID ensures seamless communication within the network. Moreover, the provision of a hotspot named 'iotserver' facilitates remote access to the ESP32 Camera's visuals. The ability to access the camera's IP address remotely enables real-time monitoring of animal detection activities, contributing to the overall assessment of the system's functionality and user convenience.

Concerning additional measurements, assessing power consumption is crucial to understanding the robot's energy efficiency. This involves examining battery life or energy usage during different operational modes and identifying optimization opportunities for extending battery life. Finally, evaluating overall system performance, including response time from sensor detection to notification or action, is essential for gauging the efficiency and reliability of the entire system in fulfilling its animal detection and communication goals. The exceptional accuracy rate of 97.6% underscores the success achieving its intended objectives.

The following are the steps of the experiment:

Initialization:

GPS Module Activation: The robot powers on and immediately activates the GPS module to acquire coordinates as illustrated in Fig.6.

![Fig. 6. GPS Module Activation](image)
Once coordinates are retrieved, they are displayed prominently on the LCD screen as illustrated in Fig. 7.

![GPS Coordinates Display](image1)

**Fig. 7.** GPS Coordinates Display

**Registration:**

The LCD screen instructs the sender to send a registration SMS as illustrated in Fig. 8. using the format "*SenderNumber" to the GSM SIM number associated with the robot.

![LCD Screen Instructs the Sender](image2)

**Fig. 8.** LCD Screen Instructs the Sender

Upon receiving the registration SMS, the robot stores the sender's number for authorized communication. The registered sender can control the robot's movements remotely by sending specific SMS commands.

**Remote Control:**

To enable live streaming from the ESP32 Camera, connect the robot to the WiFi network named "iotserver," which is configured for the camera. Once connected, obtain the IP address of the ESP32 Camera. Use the acquired IP address to access and view the live stream remotely as illustrated in Fig. 9.

![Remote Viewing](image3)

**Fig. 9.** Remote Viewing
Detection and Alerting:

The robot continuously scans for obstacles. If encountered, it sends an SMS alert with the format “Obstacle Detected: GPS Coordinates” to the registered sender as illustrated in Fig. 10, and 7 displays the GPS coordinates on the LCD screen. The robot likewise scans for animals. Upon detection, it sends an SMS alert with the format “AnimalName Detected: GPS Coordinates,” as illustrated in Fig. 10, 11, and 12 displays the animal name and coordinates on the LCD screen.

Fig. 10. Obstacle and Animal Detection along with Location

The success of the animal detection robot's experimental results hinges on a multifaceted evaluation, including an impressive 97.6% accuracy rate achieved after 30 epochs of training, as detailed in Fig. 13. Model Accuracy Graph. In terms of detection accuracy, the ESP32 Camera's proficiency in discerning diverse animal types under varying lighting and distance conditions is crucial. Assessments should still include evaluating false positives and negatives to gauge the camera's precision and potential areas for improvement. This remarkable accuracy enhances the credibility of the detection system, reinforcing its effectiveness in identifying animals accurately.

In addition to the comprehensive evaluation, specific attention was given to the performance of the animal detection robot under varying lighting conditions [20], acknowledging the crucial role of ESP32 Camera accuracy in low-light and high-light scenarios. Low-Light Condition Accuracy in the ESP32 Camera demonstrated a commendable proficiency in discerning animal types under low-light conditions, such as during nighttime or in shaded areas. The system exhibited an approximate accuracy rate of 95% in animal detection under low-light conditions, showcasing its reliability when faced with challenges related to image clarity and contrast in environments with limited illumination.

Fig. 11. Location of Robot when Animal Detected
High-Light Condition Accuracy in the ESP32 Camera's accuracy was also rigorously assessed under high-light conditions, including scenarios with bright sunlight or well-lit environments. The system excelled in maintaining an accuracy rate of around 98% in identifying animals under intense light, affirming its effectiveness in scenarios where image quality or glare could pose potential challenges.

**4.3 Significance of the Proposed Method**

The proposed animal detection system represents a significant advancement in robotics. At its core is the ESP32 Camera, ensuring precise detection and classification of various animal species even in challenging environments. A thorough analysis of false positives and negatives refines the camera's accuracy, enhancing reliability for real-world use with unpredictable animal appearances and behaviors. Integration of an Ultrasonic Sensor allows for skillful obstacle avoidance, enabling the robot to successfully navigate complex and changing environments. SMS commands and responsive maneuvering make the system user-friendly, further emphasizing its practical design.

Comprehensive user notifications are essential to the system's success. Features like LCD clarity, clear data representation, buzzer alerts, and GSM communication keep users informed of animal type, location, and robot status. Additionally, the ESP32 Camera's wireless capabilities and dedicated hotspot ('iotserver') enable remote monitoring. This allows users to observe animal detection in real-time. The focus on energy efficiency and reliable communication creates a complete and user-focused solution, demonstrating great potential within the field of animal detection robotics.

**5 Conclusion and Future Enhancements**

Conclusion and Future Enhancements of the proposed method are discussed in this section.

**5.1 Conclusion**

The proposed method, "An IoT-based Animal Detection System Using Interdisciplinary Approach" demonstrates the power of an interdisciplinary approach to solving complex
security challenges. By seamlessly integrating sensor networks, image recognition algorithms, and machine learning, it offers a remarkably effective solution for real-time animal monitoring on college campuses. At the heart of the system's success is its exceptional 97.6% accuracy rate in animal detection and species identification – a testament to the strength of this technological collaboration.

The integration of the ESP32 camera, R-CNN machine learning model, and Arduino Uno creates a robust and adaptive system. The camera acts as a high-resolution eye, the R-CNN model provides intelligent interpretation, and the Arduino Uno ensures swift, multi-layered notifications for comprehensive campus awareness. This approach not only safeguards campus communities but also promotes cost-effectiveness by automating aspects of security and generating valuable data for proactive prevention.

This proposed method marks a significant milestone in reimagining campus security. Its interdisciplinary design, outstanding accuracy, and focus on user-centric data delivery establish it as a valuable asset for any educational institution. As this technology advances, its potential extends beyond campus borders. This paper underscores the importance of cross-disciplinary collaboration in solving real-world problems, ultimately offering the blueprint for a future where humans and wildlife coexist safely and harmoniously.

5.2 Future Enhancements

The following are the future enhancements of the proposed method of an IoT-based Animal Detection

Enhancing Navigation and Localization through GPS/RTLS integration will enable precise location tracking for autonomous navigation, route planning, and geofencing. Location data will reveal patterns in animal activity on campus. Advanced Perception and Interaction using LiDAR will improve obstacle detection and 3D environment mapping, resulting in robust navigation. Object recognition will distinguish animal species, and audio analysis (microphones) will detect unique animal sounds.

Communication and Data Analytics provides Campus network integration that can facilitate real-time data, remote control, and cloud-based processing. Big data analytics will uncover animal behavior patterns. Campus-wide notifications will alert users to potential animal interactions. Focusing on creating a user-friendly interface (web or mobile) for data visualization and robot control. Implement strong security measures for data protection. Prioritize ethical practices for animal observation and data collection.

References


