2024 2nd International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC)

2024 2nd International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC) | 979-8-3503-6657-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICMACC62921.2024.10894649

Cygnus: A Vision-Based Drone System for Drowning Detection Using IoT

Dhumravarna Ambre
Department of Computer Engineering
Ramrao Adik Institute of Technology
Navi Mumbai, India
dhumravarnaambre36@gmail.com

Hariharan Sureshkumar
Department of Computer Engineering
Ramrao Adik Institute of Technology
Navi Mumbai, India
har.kum.rt20@dypatil.edu

Prasiddh Trivedi
Assistant Professor, Electrical Engineering
Gati Shakti Vishwavidyalaya
Vadodara, India
prasiddh.trivedi@gsv.ac.in

***Abstract*—With an estimated 236,000 fatalities each year and 7% of injury-related deaths worldwide, drowning is still a major global concern. In light of drone technology, this study presents a drowning detection system that makes use of IoT and AI/ML. Although drowning events can happen to anyone, our strategy concentrates on vision-based techniques because of their inherent benefits, which include their non-intrusive nature and real-time monitoring capabilities. Our focus on vision-based solutions arises from the limitations of traditional wearable systems, which include their dependence on life buoys and insufficient infrastructure. Our technology provides an aerial perspective of drowning incidents through a wireless drone network. By merging RGB and IR vision, it improves detection accuracy across a range of light, weather, and water conditions. We use edge computing to maximize efficiency and deploy machine learning models, particularly Convolutional Neural Networks (CNN), on drone microprocessors to enable real-time drowning detection. Furthermore, we use LoRa to create a Wireless Sensor Network (WSN) to enable better drone cooperation and communication. In summary, this work introduces a drone-based drowning detection system that makes use of IoT and AI/ML. Our solution intends to improve public safety, reduce drowning incidents, and maybe save lives by addressing the shortcomings of current approaches and embracing fresh methodologies.**

***Index Terms*—drones, surveillance, ML models, IoT, drowning.**

I. INTRODUCTION

Drones, also known as unmanned aerial vehicles (UAVs), are efficient surveillance tools that function as aerial sentinels as they often have cameras mounted on them. Applications for drone-based surveillance include road traffic monitoring, security, and precise geo-localization of targets for search and rescue [1], [2], [5], [6], [8], [9], [10], [11], [12]. An onboard object detection program [3] receives the video input from the drone's camera to identify nearby target objects.

A. Solution Approach

Our strategy involved utilizing a ResNet model trained on dataset from the [4] Vega research paper. Leveraging both test and train datasets, we determined an optimal altitude that strikes a balance between accuracy and the area covered by the drone's camera. As outlined in the Vega paper, we recognized that higher altitudes enable broader coverage but may compromise accuracy.

To enhance our methodology, we introduced sensor fusion, combining Infrared and RGB images to augment detection accuracy, particularly in distinguishing drowning from swimming incidents [18], [19], [20], [21], [22], [23].

Our primary model, the YOLOv8, trained on around 1000 images, forms the backbone of our approach. The workflow begins with the drone positioned at a safe height of 100m (compliant with the government's approved height of 120m). The model initiates by executing a sensor fusion process, generating a merged IR and RGB image. This composite image aids in identifying drowning or swimming scenarios. Should the accuracy fall below the specified threshold, our approach draws from the concept introduced in the Vega research paper known as the DroneZoom technique. This involves the drone lowering its altitude to obtain a clearer view of the target, thereby enhancing detection capabilities.

We summarize Cygnus's Contributions:

1) Comprehensive System Architecture: Cygnus encompasses a robust system architecture comprising various critical elements. Notably, it incorporates an MIS system, enabling lifeguards to monitor situations effectively through a dedicated application [24], [25], [26], [27], [28].

2) Advanced Drone Network: Cygnus boasts an adaptable drone network capable of scaling optimally in scenarios such as power surges. It includes features like drone waypoint navigation and autonomous drone docking, allowing seamless wireless charging [13], [14], [15], [16], [17].

3) Sensor Fusion: A pivotal addition to Cygnus is the implementation of sensor fusion, merging Infrared (IR) and RGB images, significantly enhancing detection accuracy.

4) Onboard Edge Computing: Cygnus leverages onboard edge computing, a critical feature that minimizes detection latency directly on the drone, thereby ensuring swift and efficient processing.

II. RELATED WORK

A. Drowning Detection Systems

In the domain of drowning detection and prevention, various innovative approaches have been explored. These include the use of Unmanned Aerial Vehicles (UAVs) for deploying flotation devices in harsh sea conditions, shore-based cameras integrated into comprehensive systems for timely intervention and proactive monitoring, and image analysis techniques in swimming pool environments. Drones have been leveraged for live surveillance over bodies of water using computer vision for real-time recognition and location of individuals at risk. A novel proposal incorporates Edge Computing devices with a lightweight convolutional autoencoder for unsupervised drowning detection. Additionally, focused efforts on high-risk groups, such as children and the elderly, integrate deep learning and object recognition, autonomously deploying drones and notifying local authorities [4-12]. Collectively, these contributions offer a diverse landscape of drowning detection methodologies, providing valuable insights into the field.

B. Wireless Senor Network

In the realm of WSN and mobile networks, significant advancements include a mobile gateway system integrating a quadcopter drone with a WSN for data collection. Addressing WSN challenges, another proposal employs Thread, a Low-Powered IPv6-based wireless protocol, with UAVs for extended range, incorporating secure data transmission through Elliptic Curve Cryptography. Research explores UAV communications, emphasizing cellular-connected UAVs and UAV-enabled aerial communication platforms. In IoT-based LoRa networks, investigations provide insights and recommendations. A comprehensive exploration defines nodes and devices in UAV-based data collection, featuring various routing algorithms [13-17]. These contributions collectively form a multifaceted backdrop, covering mobile gateway design, WSN architecture, UAV communications, IoT-based LoRa networks, and UAV-based data collection, serving as a rich source of knowledge and inspiration.

C. Sensor Fusion

Research in RGB and Infrared image processing introduces a deep learning-based matching method using a densely connected Convolutional Neural Network (CNN) to extract common features from diverse spectral bands. Additionally, a novel RGB-IR cross-input and sub-pixel upsampling network enhances the spatial resolution of infrared images. A survey categorizes RGB-infrared trackers into distinct groups. In sensor calibration, a method facilitates external calibration of a thermal infrared camera and a LiDAR sensor through the design of a 3D calibration target. Advances include a multi-scale transformation and norm optimization-based approach for infrared and visible image fusion, a difference maximum loss function to enhance CNN performance, and a combination of Thermal Infrared Camera and LiDAR sensors for better accuracy in autonomous vehicles. Lastly, a study introduces an effective method for the fusion of RGB-D and thermal sensor data to improve human detection accuracy [18-25]. These diverse contributions encompass various facets of RGB and Infrared image processing, providing valuable context and insights for the present study.

III. PRELIMINARIES

In this section, we describe the preliminaries of our setting and all the components involved in the system.

Drone: It is assumed that the drones employed in the system has the ability to navigate in three dimensions and have access to a GPS-based navigation system. In order to travel the area and establish waypoints, the drones are additionally equipped with 3D waypoint navigation. We'll assume for analytical simplicity that the drone travels between waypoints at a constant velocity of v.

Cameras: We utilize a dual-camera setup consisting of two cameras placed perpendicular to each other. One camera is oriented to capture a front view, while the other is positioned to obtain a downward view, enabling a comprehensive coverage area for detection purposes. Each camera employs a wide-angle lens to encompass a larger field of view and mitigate issues associated with rotational motion that might cause blurriness or unclear imagery.

For analytical purposes, let's denote the field of view (FoV) angle of each camera as θ_f for the front-facing camera and θ_d for the downward-facing camera, respectively. The area covered by the image captured by the front-facing camera at height h_f from the ground plane is given by:

$$A_f = k_f . h_f^2$$

Similarly, for the downward-facing camera at height h_d above the ground plane:

$$A_d = k_d . h_d^2$$

Here, the constants k_f and k_d are derived from the respective FoV angles (θ_f, θ_d) and the aspect ratios of the cameras. The area covered increases quadratically with the height *h*.

Target: We assume that individuals appear in the area according to a time-variant homogeneous Poisson process with a rate λ and a uniformly random location distribution. In our scenario, these individuals represent people who may be swimming or in distress, allowing us to detect and differentiate between regular swimming and potential instances of drowning. The parameter λ serves as an input to our system, governing the rate at which these individuals appear within the monitored area.

Similar to modeling natural disasters, accidents on highways, or intrusions in an area, a Poisson process provides a framework to simulate the appearance of individuals—specifically, people engaged in swimming activities or facing potential danger in the water. This model enables us to analyze their behavior and detect critical situations, such as instances of drowning, within the observed area.

Object Detection and quality metrics: The drone has an inbuilt object detection program OD. The detector is based on Transfer Learning which is implemented on YOLO Object Detection Algorithm this program has been trained with dataset with sufficient examples of the class of objects to detect. OD outputs two parameters as part of a detection: a) Confidence and b) A rectangular box that contains the object called the bounding box. High confidence is often necessary in an application like surveillance to definitively identify targets without detecting false positives.

Detection Latency: If a target appears at time t in the area and the drone detects the target at time t+Δ, then we consider the detection latency to be Δ.

1) Δ- Maximum detection latency within which detection must be reported
2) λ- Object appearance rate or Input to our model (Unit: m^{-2} sec^{-1})
3) θ_f - Observation altitude from the front-facing camera of the drone θ_f or upper altitude.
4) θ_d - Observation altitude from the down-facing camera of the drone θ_f or upper altitude.

Fig. 1. Notations used in derivations.

IV. OBJECT DETECTION TRENDS

In this section, we detail the trends in detection confidence and Intersection over Union (IoU) of an energy-efficient version of the Faster R-CNN object detector paired with the InceptionResNet v2 backbone [9]. We empirically assess the detector's confidence and IoU by executing inference on a test dataset comprising targets of diverse sizes. Our goal is to pinpoint the optimal height for the drone, ensuring a balance between detection accuracy and Area coverage.

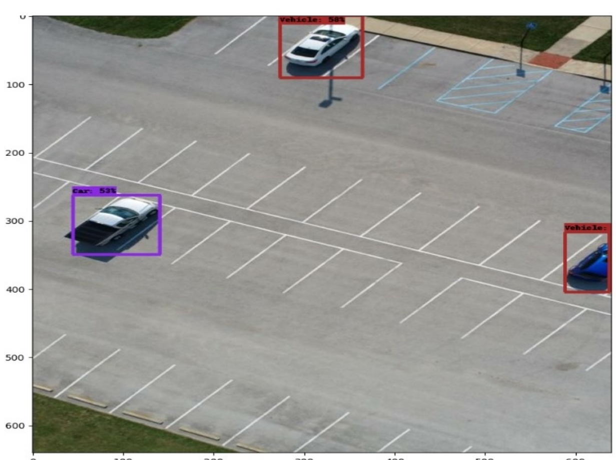

Fig. 2. Cars observed from height of 200ft

Initially, we trained the Faster R-CNN model with the InceptionResNet v2 backbone using a dataset comprising images depicting outdoor scenes and including target objects such as pedestrians, cars, and trucks. Subsequently, we curated a testing dataset specifically composed of bird's eye view images capturing cars from various drone perspectives. The observed trend reveals a decline in both detection confidence and Intersection over Union (IoU) values as the target size decreases, indicating that smaller targets are detected with reduced precision.

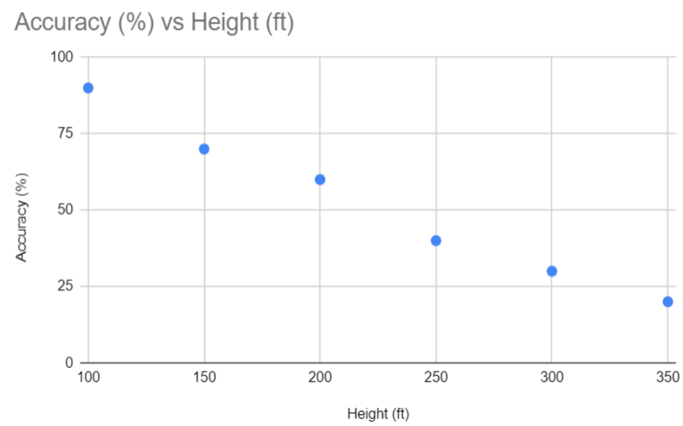

Fig. 3. Graph for plotting Object detection accuracy vs Height.

In this section, we elucidate the operational mechanism of the drones. We strategically deploy drones at a specific altitude, ensuring an optimal balance between the covered area and detection accuracy. The utilization of [4] DroneZoom facilitates the enhancement of accuracy by adjusting the drone's height, thereby allowing for a trade-off between optimal coverage and increased precision.

This sensor fusion technique integrates both visible RGB and infrared (IR) images through a complex fusion strategy employing the VGG19 neural network. The fusion process encompasses multiple phases: beginning with low-pass filtering applied to both RGB and IR images to extract their respective 'low frequency' components. Subsequently, a pre-trained VGG19 model captures activations from specific layers, followed by feature extraction to represent diverse attributes within the images. The Sensor fusion generates saliency maps accentuating crucial regions, effectively emphasizing relevant details from both the visible RGB and infrared domains. The final fused output amalgamates the low-frequency elements with high-frequency components obtained from the saliency maps, resulting in an enriched fused image that incorporates informative content from both the RGB and IR sources. We also use an image restoration method for motion-blurred objects [27] due to the motion of the drones.

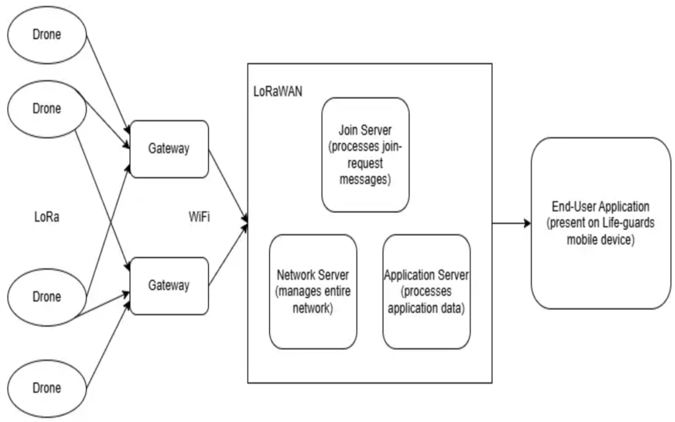

Fig. 4. LoRaWAN Network Architecture.

V. SYSTEM OVERVIEW

A. Drone Charging

Optimizing drone delivery energy efficiency requires strategic placement of charging stations. Using simulations and a battery-aware power model, stations are located based on optimal flight parameters like delivery distance, payload weight, and speed. This approach aims to minimize recharging travel distance, reduce energy consumption, and enhance overall network efficiency.

B. Wireless Sensor Network

Our focus is on optimizing LoRa network performance by adjusting transmission power for End Devices (EDs) communicating with multiple gateways. We organize the network into cells, each with EDs connected to a gateway, ensuring proximity for optimal communication. Strategic gateway placement, guided by a Greedy Gateway Placement algorithm, balances installation costs with network performance for an efficient setup.

VI. IMPLEMENTATION

The Cygnus Drowning Detection System is designed around a LoRa WAN network architecture. Each drone is equipped with a GAP8 microprocessor responsible for merging RGB and IR images and executing YOLOv8 for drowning detection. The flight operations are managed by a STM32F4-powered Flight Controller, while LoRa technology facilitates efficient location transmission. The drone also features a GPS module for capturing coordinates, and RGB and IR cameras are strategically placed for comprehensive imaging.

Within the LoRa WAN network, a Network Server serves as the Management Information System (MIS), storing data in InfluxDB and forwarding it to the Application Server. The Application Server hosts an end-user web application on the lifeguard's phone, providing real-time tracking of the drone's GPS location.

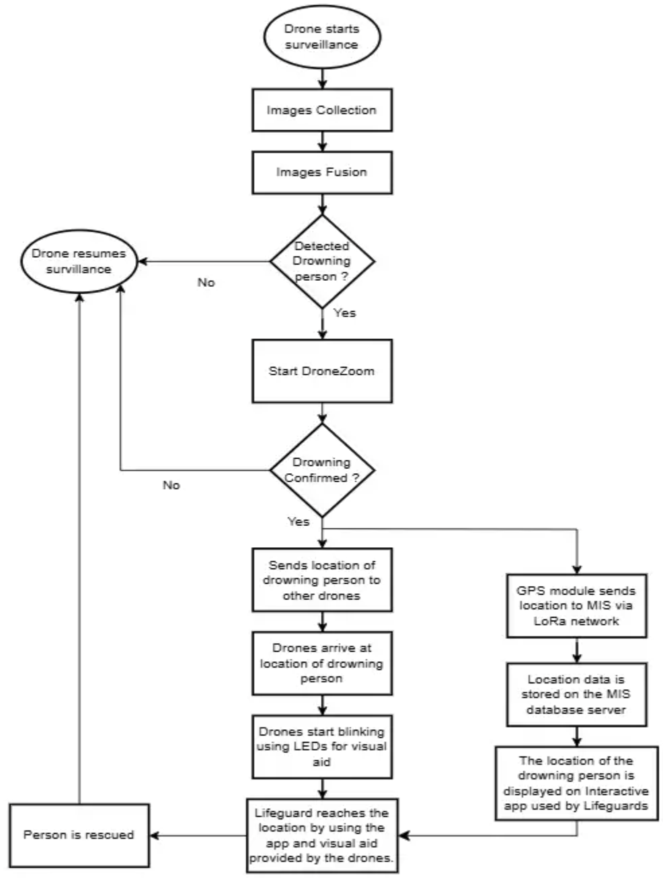

Fig. 5. Flow Chart of the proposed system.

VII. EXPERIMENTS AND PERFORMANCE ANALYSIS

The graph depicting ”train/box_loss” (Fig. 6) provides a visual insight into the training process of a model designed for object detection. With the x-axis representing the progression of training, denoting epochs, and the y-axis reflecting the corresponding value of the ”train/box_loss,” the graph encapsulates the model's performance over time. At the outset, as the x-value stands at 0, the y-value is notably high at 1.8, indicating a substantial loss associated with bounding box predictions. This signifies the initial stage of training, where the model's predictions are relatively inaccurate. However, as training progresses and the x-value increases to 20, the y-value steadily decreases to 0.8. This decline in the ”train/box_loss” implies significant improvement in the model's ability to precisely predict bounding boxes as it learns from the training data. Thus, the graph serves as a visual testament to the model's evolving competency in accurately identifying objects within images as the training process advances.

The precision graph of the YOLO V8 model spans 25 epochs on the x-axis, with precision values ranging from 0.5

Algorithm 1 Enhanced Drone Routing Algorithm

```
Require: Drones, Δ, λ
Ensure: Efficient drone routing and rescue operation
  procedure MAINDRONEROUTING
     InitializeDronesAndPaths()          ▷ Initialize drone positions and fixed paths
     while true do
        for each drone in Drones do
           if DrowningPersonDetected(drone) then
              if DistressSignalReceived(drone) then
                 DynamicAStarRouting(drone)
              else
                 FixedPathRouting(drone)
                 DroneZoomOperation(drone, Δ, λ)
                 PerformRescueMission(drone)
              end if
           end if
        end for
        ResumeSurveillance()
     end while
  end procedure
  procedure DYNAMICASTARROUTING(drone)   ▷ Implement Dynamic A* for rerouting
     // Update drone's path dynamically based on real-time information
     // Consider the distress signal location and optimize the path accordingly
     // Update drone's path using the Dynamic A* algorithm
  end procedure
  procedure FIXEDPATHROUTING(drone)      ▷ Implement Fixed Path Routing
     // Generate a fixed path for the drone based on its current position and destination
     // This can be done using algorithms such as Dijkstra's or A*
     // Update the drone's path to follow the fixed path
  end procedure
```

at epoch 0 to 0.9 at epoch 25. Starting with a moderate accuracy level, the model steadily improves its precision over subsequent epochs. By epoch 25, it achieves a commendable precision of 0.9, showcasing significant enhancement in object detection accuracy. This graph serves as a clear visualization of the model's iterative refinement and increasing reliability in accurately identifying objects within images throughout the training process.

In this research, we use a dataset of about 1000 images to conduct a thorough supervised learning experiment with the YOLOv8 model. The binary labels used to meticulously annotate the dataset are '1' for swimming and '0' for drowning scenarios (Fig. 7). The YOLOv8 model is trained on this fully annotated dataset, which allows it to distinguish and classify swimming and drowning actions with accuracy. We assess the model's memory, precision, and overall accuracy in differentiating between drowning and swimming cases through extensive testing. Upon completing the training of the YOLOv8 model on our custom dataset, we achieved a precision rate of 95%, a recall rate of 68%, and a mean Average Precision (mAP) score of 83%.

Fig. 6. Confusion Matrix for YOLOv8 model.

Fig. 7. Detection of Drowning person.

VIII. CONCLUSION

In conclusion, Cygnus has successfully met its objectives, establishing itself as an innovative aquatic safety solution. Its comprehensive system architecture includes an integrated Management Information System (MIS) and a lifeguard application, improving situational awareness. The advanced drone network

Fig. 8. Precision of YOLOv8 Model.

in Cygnus, with adaptive scaling and features like waypoint navigation, shows resilience even in challenging situations. Sensor fusion, combining IR and RGB images, enhances detection accuracy for improved water safety understanding.

REFERENCES

[1] He, X., Yuan, F., Zhu, Y. (2021). Drowning detection based on video anomaly detection. Lecture Notes in Computer Science, 700–711. https://doi.org/10.1007/978-3-030-87361-5$_{5}$7.

[2] He, Xinyu, Yuan, Fei, Liu, Tingzhuang, Zhu, Yi. (2023). A video system based on convolutional autoencoder for drowning detection. Neural Computing and Applications, 35, 1-13. https://doi.org/10.1007/s00521-023-08526-9.

[3] Shatnawi, M. et al. (2023) ‘Deep learning and vision-based early drowning detection’, Information, 14(1), p. 52. doi:10.3390/info14010052.

[4] A. Bandarupalli, S. Jain, A. Melachuri, J. Pappas and S. Chaterji, ”Vega: Drone-based Multi-Altitude Target Detection for Autonomous Surveillance,” 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), Pafos, Cyprus, 2023, pp. 209-216, doi: 10.1109/DCOSS-IoT58021.2023.00044.

[5] Seguin, Celia, Blaquière, Gilles, Loundou, Anderson, Michelet, Pierre, / Markarian, Thibaut. (2018). Unmanned aerial vehicles (drones) to prevent drowning. Resuscitation, 127. https://doi.org/10.1016/j.resuscitation.2018.04.005.

[6] Zhao, Junhao, Xiao, Gang, Zhang, Xingchen, Bavirisetti, Durga. (2019). A Survey on Object Tracking in Aerial Surveillance. https://doi.org/10.1007/978-981-13-6061-9$_{4}$.

[7] Rakotondraibe, M., Fang, T., Saniie, J. (2023). Drowning Recognition for Ocean Surveillance using Computer Vision and Drone Control. In 2023 IEEE International Conference on Electro Information Technology (eIT), Romeoville, IL, USA (pp. 394-398). https://doi.org/10.1109/eIT57321.2023.10187281.

[8] S. Sivakami, K. Janani, K. Janani and R. Ranjana, ”Drowning prevention system-at sea level,” 2017 2nd International Conference on Computing and Communications Technologies (ICCCT), Chennai, India, 2017, pp. 370-373, doi: 10.1109/ICCCT2.2017.7972308.

[9] Bäckman A, Hollenberg J, Svensson L, Ringh M, Nordberg P, Djärv T, Forsberg S, Hernborg O, Claesson A. Drones for Provision of Flotation Support in Simulated Drowning. Air Med J. 2018 May-Jun;37(3):170-173. doi: 10.1016/j.amj.2018.01.007. Epub 2018 Mar 9. PMID: 29735229.

[10] Handalage, U., Nikapotha, N., Subasinghe, C., Prasanga, T., Thilakarthna, T., Kasthurirathna, D. (2021). Computer Vision Enabled Drowning Detection System. In 2021 3rd International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka (pp. 240-245). https://doi.org/10.1109/ICAC54203.2021.9671126.

[11] Q. He, Z. Mei, H. Zhang and X. Xu, ”Automatic Real-Time Detection of Infant Drowning Using YOLOv5 and Faster R-CNN Models Based on Video Surveillance,” in Journal of Social Computing, vol. 4, no. 1, pp. 62-73, March 2023, doi: 10.23919/JSC.2023.0006.

[12] P, D. et al. (2023) ‘AI-driven drowned-detection system for rapid coastal rescue operations’, Spatial Information Research [Preprint]. doi:10.1007/s41324-023-00549-7.

[13] I. Deaconu and A. Voinescu, ”Mobile gateway for Wireless Sensor Networks utilizing drones,” 2014 RoEduNet Conference 13th Edition: Networking in Education and Research Joint Event RENAM 8th Conference, Chisinau, Moldova, 2014, pp. 1-5, doi: 10.1109/RoEduNet-RENAM.2014.6955319.

[14] Jawhar, Imad Mohamed, Nader Al-Jaroodi, Jameela. (2015). UAV-based data communication in wireless sensor networks: Models and strategies. 2015 International Conference on Unmanned Aircraft Systems, ICUAS 2015. 687-694. 10.1109/ICUAS.2015.7152351.

[15] N. Ravi, R. Chitanvis and M. El-Sharkawy, ”Applications of Drones using Wireless Sensor Networks,” 2019 IEEE National Aerospace and Electronics Conference (NAECON), Dayton, OH, USA, 2019, pp. 513-518, doi: 10.1109/NAECON46414.2019.9057846.

[16] Song, Q. et al. (2021) ‘A survey of prototype and experiment for UAV Communications’, Science China Information Sciences, 64(4). doi:10.1007/s11432-020-3030-2.

[17] Ghazali, Mohamad Teoh, Kelvin Rahiman, Wan. (2021). A Systematic Review of Real-Time Deployments of UAV-Based LoRa Communication Network. IEEE Access. 9. 10.1109/ACCESS.2021.3110872.

[18] Ranjit, S., Kang, K. Kim, W. Investigation of lock-in infrared thermography for evaluation of subsurface defects size and depth. Int. J. Precis. Eng. Manuf. 16, 2255–2264 (2015). https://doi.org/10.1007/s12541-015-0290-z

[19] Zhu, R. et al. (2019) ‘Matching RGB and infrared remote sensing images with densely-connected convolutional neural networks’, Remote Sensing, 11(23), p. 2836. doi:10.3390/rs11232836.

[20] Xingchen Zhang, Ping Ye, Henry Leung, Ke Gong, Gang Xiao, Object fusion tracking based on visible and infrared images: A comprehensive review, Information Fusion, Volume 63, 2020, Pages 166-187, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2020.05.002.

[21] Farahnakian, Fahimeh Heikkonen, Jukka. (2020). Deep Learning Based Multi-Modal Fusion Architectures for Maritime Vessel Detection. Remote Sensing. 12. 2509. 10.3390/rs12162509.

[22] Guofa Li, Yongjie Lin, Xingda Qu,An infrared and visible image fusion method based on multi-scale transformation and norm optimization, Information Fusion, Volume 71, 2021, Pages 109-129, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2021.02.008.

[23] Xiao, X. et al. (2021) ‘Infrared and visible image object detection via focused feature enhancement and cascaded semantic extension’, Remote Sensing, 13(13), p. 2538. doi:10.3390/rs13132538.

[24] Choi, Ji, Kim, Min. (2021). A Sensor Fusion System with Thermal Infrared Camera and LiDAR for Autonomous Vehicles: Its Calibration and Application. 361-365. 10.1109/ICUFN49451.2021.9528609.

[25] A. Ozcan and O. Cetin, ”A Novel Fusion Method With Thermal and RGB-D Sensor Data for Human Detection,” in IEEE Access, vol. 10, pp. 66831-66843, 2022, doi: 10.1109/ACCESS.2022.3185402.

[26] Y. Chen, D. Baek, A. Bocca, A. Macii, E. Macii and M. Poncino, ”A Case for a Battery-Aware Model of Drone Energy Consumption,” 2018 IEEE International Telecommunications Energy Conference (INTELEC), Turino, Italy, 2018, pp. 1-8, doi: 10.1109/INTLEC.2018.8612333.

[27] L Ge and L Dou 2023 J. Phys.: Conf. Ser. 2478 062004

[28] B. Ousat and M. Ghaderi, ”LoRa Network Planning: Gateway Placement and Device Configuration,” 2019 IEEE International Congress on Internet of Things (ICIOT), Milan, Italy, 2019, pp. 25-32, doi: 10.1109/ICIOT.2019.00017.