International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Multimodal Intrusion Detection In Restricted Zones Using Thermal-RGB Fusion And RFID Authentication

Swati Shilaskar[1], Shripad Bhatlawande[2], Jyoti Madake[2], Pallavi Mulmule[3], Anup Barde[2], Pratik Bhakare[2], Shrivardhan Baviskar[2]

[1]Department of Engineering Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
[2]Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, India
[3]Department of Electronics and Communication Engineering, School of Engineering and Technology, DES Pune University, Pune, India
*Corresponding author: swati.shilaskar@vit.edu

Abstract
Traditional single-sensor monitoring systems are not effective in low-visibility situations, like darkness, fog, and smoke, and often produce false alarms, thus restricting their useful application in critical security situations. Previous studies have investigated visual fusion based on thermal cues or RGB cues separately but none of the existing solutions can be deployed at the edge and combine dual-stream visual fusion and hardware identity authentication for real-time intrusion classification. Fusing thermal imaging (MLX90640: 24x32) with RGB video (IP camera: 640x480) streams, this paper proposes a multimodal intrusion detection system that uses two independently operating YOLO detection models, whose results are checked for consistency by spatial cross-checking through the Intersection-over-Union (IoU) method and trigger MFRC522 RFID-based personnel authentication. No server side processing is performed and the whole pipeline runs on a Raspberry Pi 5 which makes it suitable for resource constrained industrial, defence and warehouse applications. The proposed system is tested with 1000 pairs of synchronized thermal-RGB images under daytime, nighttime, and low-light conditions for its classification accuracy, precision, recall, and F1-score results are 98.23%, 0.9910, 0.9888, and 0.9899 respectively. The IoU-based fusion mechanism can significantly reduce false positives due to environmental heat sources and non-human thermal signatures from single-modality baselines. This system is light in weight, has multiple means of redundancy and is designed with RFID-based authentication and verification, which is a feasible and deployable solution for autonomous perimeter security.

Keywords - Intrusion detection, multimodal sensor fusion, thermal imaging, YOLO, RFID authentication, edge computing, Raspberry Pi, IoU verification.

I. Introduction

Whether it's a factory, military installation, or secure warehouse, a surveillance system in such areas should be able to operate 24/7 and in all weather conditions. In foggy and smoky weather the RGB camera system is highly sensitive to low light, and has a high false negative rate in low light conditions, thus compromising the security integrity. [1]. Deep learning object detection has also played an important role in improving the detection performance of humans in controlled environments, especially in the real-time based system, YOLO. [2]. But, in reality, single-modality systems are quite susceptible to the sensor noise, occlusion and variability in illumination. To address these challenges, multimodal sensor fusion has been proposed as a solution to the problem, and multiple data sources including RGB, thermal and depth data for improving the robustness of the detection in complex scenarios and etc. have been used. [4]. Thermal sensors can detect infrared radiation or heat emitted by the target in the dark and in low visibility situations making them ideal for night and low visibility surveillance. [1]. In recent studies, it has been demonstrated that thermal-RGB fusion is effective for object detection tasks [5][6] but most of the systems are based on server grade GPUs, thus they are not appropriate to be used in a field in resource-limited environments.

The one aspect of current IDSs that is not covered in the fusion approach is identity verification. In environments with high security requirements, besides detecting humans, it is also important to know if the detected person is authorized or not or if the detected person is an intruder. Access control based on RFID can be used as the identification mechanism [7] commonly used, and the reliability and low cost are good, but, so far, the use of the above in conjunction with AI image detection on the embedded machine has not been studied regularly. However, the single-sensor architectures of existing IDSs are not resilient in adverse scenarios [2][3], existing ones rely on cloud or GPU resources which limits their deployment on the edge [5][6] and they lack integrated identity authentication mechanisms [7][8]. There is no previous work that has all three challenges of multimodal visual fusion, RFID based authentication and real-time edge processing in a single embedded system. This article aims to solve this issue by suggesting a three-stage process for detecting and authenticating the device which can be implemented fully on the Raspberry Pi 5. The key contributions of the paper are a dual-stream YOLO framework for the detection that is applicable to the synchronized thermal frames and RGB frames, and a cross validation of detections by consistency check using IoU, which eliminates false positive detections caused by environmental heat sources, animals and partial occlusion. Based on successful visual detection (not achieved in any of the visual-only fusion systems studied), a binary classification between detected persons as authorized or not, was then carried out using an integrated RFID authentication module. In addition, the entire detection, fusion and authentication process is performed in an edge-deployable embedded solution, meaning it processes the entire process without any out-of-band compute requirements for 1,000 pairs of synchronized images during the day, at night and in low-light situations.

II. Literature Review

Thermal imaging has been widely discussed because it is able to detect objects by their emitted infrared radiation, not reflected light, and is usable in low-light or night-time conditions, without difficulty. Today's thermal imaging devices create a map of the temperature distribution of a scene, and each pixel in the map corresponds to a thermal signature of objects in the scene. This has opened the door for applications such as surveillance, healthcare, industrial inspection and autonomous systems, particularly when used in conjunction with machine learning techniques for automated analysis [1]. With the advancement of the deep learning-based object detection algorithm, the human detection systems have been greatly improved. The YOLO family of CNN architectures is widely used for real-time object detection with a high accuracy and low computational requirements. They provide automatic detection of people in video streams and have the potential to enhance security monitoring in the real world, for example in surveillance systems [2]. Thermal imaging has also found extensive use in aerial surveillance and search and rescue on occasions where visibility is poor. A UAV thermal camera can be used to find human heat signatures in forested or mountainous areas. Deep learning models based on thermal data have shown to perform well and detect humans from aerial data, which can be deployed on embedded platforms for real-time monitoring. [3]. To enhance the detection performance, multisensor fusion techniques have been developed by fusing data from various sensors including RGB cameras, LiDAR and thermal sensor. A combination of several sensing modalities can provide complementary data such as visual texture, distance and temperature. This fusion enhances the perception of the environment and increases the reliability of the detection systems of robotic and autonomous systems [4]. The fusion of thermal and RGB imagery has been successful in addressing the challenges associated with each specific sensor type. RGB cameras give detailed visual characteristics under good lighting conditions and thermal sensors can detect under dark conditions or bad weather. Multi-modal fusion approaches yield better object detection results than single-sensor approaches [5]. Recently, more sophisticated combinations of sensors have been investigated for enhancing the accuracy of human detection. Detection systems can use a combination of environment features by incorporating thermal, optical, and depth data. Experimental results verify that accuracy of detection is improved when both the multi-modal approach is used, rather than the single-sensor approach, and that it is more robust in different environmental conditions.

[6] IDSs have been extensively analyzed as a crucial part of security architectures which aim to detect unauthorized activities. It is a system that monitors the behaviour of the system or traffic on the system and notifies in case of abnormal behaviour or violation of the policy. In the past few years, the application of the machine learning techniques to improve intrusion detection systems has been increasing [7]. In today's digital world, intrusion detection is an important component in safeguarding data and network resources from malicious attacks. The best intrusion detection system keeps a track on the activities of the system and if it is found to be abnormal or not. Evaluation of such systems involves using metrics based on confusion matrices like true positives, false positives and detection accuracy [8]. Infrared thermography has also been much investigated as a sensing method which can yield detailed thermal data at various spatial scales. The thermal imaging technique is able to identify the temperature differences in the

environment, and has been applied to different fields, such as building detection and urban monitoring. The capabilities demonstrate the flexibility of infrared sensing ability for performing environmental observation and analysis [9]. The advances in thermal imaging technology and embedded hardware platforms have also made it easy to deploy the intelligent surveillance systems in real time. Thanks to edge computing devices and deep learning models, processing the thermal images and detecting objects directly in the embedded system can be realized, which is helpful for achieving efficient and low-power security monitoring solutions. [10] TIR has also been studied in the context of affective/emotional state recognition in multimodal human–robot interaction systems. Combining facial expression with skin temperature change, machine learning algorithms can detect emotions such as frustration, boredom and enjoyment. These multi-modal approaches allow the development of more flexible and intelligent robots which can respond to humans' emotional conditions [11]. Real-time processing of multimodal sensor data in intelligent systems has become possible thanks to the development of embedded computing and machine learning. The developments have contributed to making advanced detection and monitoring applications possible in different domains, such as robotics, surveillance, health and autonomous systems, where strong perception is required [12].

J. Jeyalakshmi et al. [13] have suggested a deep learning-based face recognition system using thermal imaging to improve the recognition performance of the traditional face recognition system in low light conditions. Similarly, Setjo et al. [14] used Haar-Cascade classifiers for human detection in thermal images which had achieved good detection accuracy and low computational cost. Kim et al. [15] tested both the RGB camera and the thermal camera, and concluded that the thermal camera has a higher performance in low light. Subramaniam et al. [16] tested a fused RGB-based and thermal based model, and showed that the use of multimodal fusion can greatly enhance the accuracy of human detection in surveillance and ADAS tasks. Ghodmare et al. [17] built a real-time surveillance system based on the YOLOv8 and NVIDIA Jetson Nano, highlighting the significance of edge computing for real-time object detection applications. Another study [18] presented a framework for detecting thermal anomaly using dual-sensor surveillance system based on context-encoder which improves the performance of the thermal anomaly detection system. Considering the contradiction between computing complexity and high accuracy, Li and Yao proposed a lightweight intrusion detection model using the multimodal fusion and attention mechanism [19]. The research conducted in smoky environment on the body-part detection with thermal camera [20] pointed out the significance of thermal vision for rescue operation and firefighter assistance. Peng et al. [21] proposed an edge enhanced and frequency aware fusion network for object detection that can improve the detection performance in various light conditions. To detect the face in the thermal image while preserving the spatial accuracy and temporal stability for healthcare and monitoring purposes, Mozafari et al. [22] addressed the above problem using deep learning technique. Moreover, Haar-Cascade based thermal human detection method [23] was proven to be economical and efficient in resource limited application. To sum up, the studies emphasize the potential that the integration of the technologies of thermal imaging, RGB fusion, deep learning, and edge computing can enable the development of reliable, robust, and efficient surveillance systems and human detection systems in the age of modern technology. A multimodal fusion algorithm based on CNNs was proposed to improve the detection accuracy of human from both visible and thermal images. For image fusion, the method is based on an encoder–decoder architecture, and for classification, it is based on the ResNet-152 model. The experimental results show that the fused model is more accurate than the singe visible model and singe thermal model, with accuracy of 99.2% on the KAIST dataset, proving the effectiveness of the multimodal fusion. However, the method could be more complex and not be practical for real-time embedded application[24]. Embedded vision systems have been developed to meet the real-time image processing demands for severe performance and energy requirements. There are numerous applications such as autonomous vehicles and healthcare. Model based design and heterogeneous programming are approaches that have been carried out recently in order to optimize on platforms like NVIDIA Jetson TX2. While these systems are very efficient and are able to be deployed on the edge, achieving high accuracy whilst also being low-power is a big challenge. [25] A current problem with the thermal surveillance systems is that they do not offer strong identity verification, which makes it hard to tell the difference between authorized and unauthorized persons. Meanwhile, many of the newer fusion-based methods rely on GPUs to perform the processing, and that type of processing is energy intensive, making it impractical for use in real-world, edge deployments. Moreover, existing intrusion detection systems are mostly network based and do not even consider physical threat detection. Hence, it is necessary to design a system to be optimized for efficient edge deployment and to integrate multimodal fusion with reliable identity authentication.

III. METHODOLOGY

A. System Overview

The proposed system will be a multi-authentication system based on thermal, RGB and RFID to detect unauthorized human access in restricted areas. The architecture can combine data of multiple sensing modalities to increase the accuracy of the detection under various environmental conditions e.g. low light, fog, and smoke. The thermal sensor can be used to sense the heat sources of the objects regardless of lighting conditions, while the RGB camera can obtain structural and visual information, which can be identified by a human being. The object detection models are based on the YOLO and each stream of images is processed. Fusion module on a Raspberry Pi users uses Intersection-over-Union (IoU) consistency based detections of both modalities as shown in equation (1).

$$\mathrm{IoU} = \frac{\mathrm{Area}(B_T \cap B_R)}{\mathrm{Area}(B_T \cup B_R)} \qquad (1)$$

where, IoU is Intersection of Union, B_T is Thermal Bounding Box and B_R is RGB Bounding Box.

When an RFID is identified it displays a window confirming that it is a person and the person detected is recognized as an authorized or intruder. This multimodal design is usable for detection and will not result in false alarm due to the noise in the environment, animals or thermal reflection.

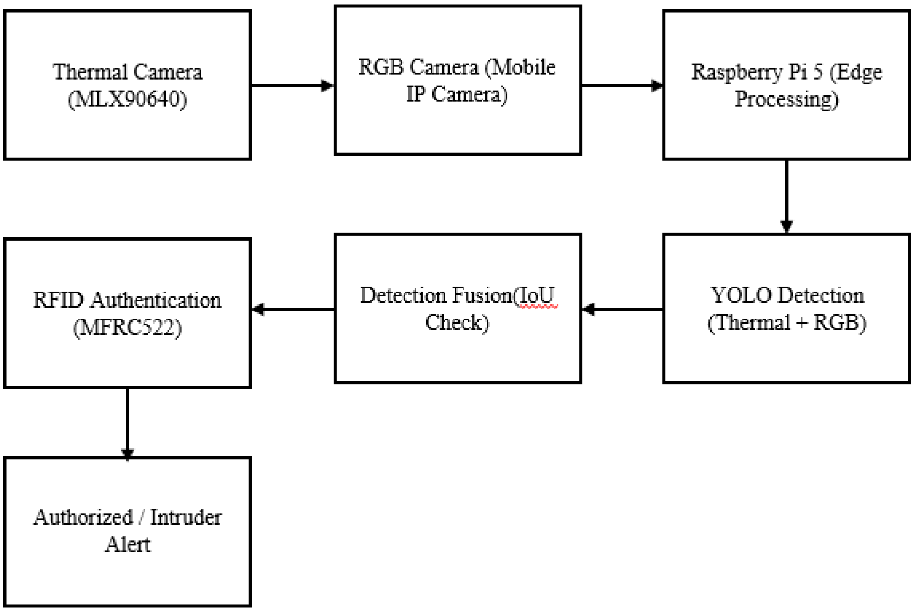

Fig.1. System Architecture of the Proposed Intrusion Detection System

B. Thermal Image Preprocessing

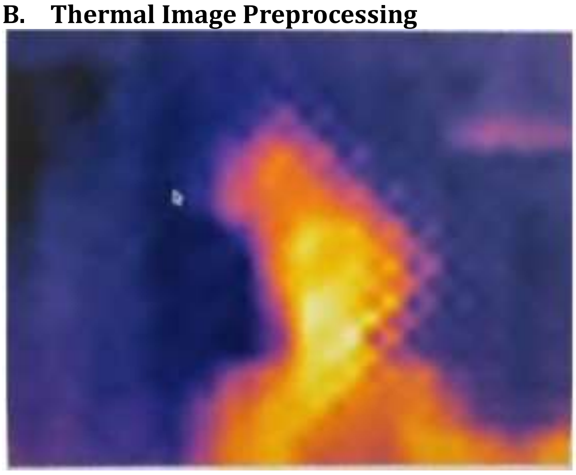

Fig.2. Thermal image of a human

The thermal sensor, MLX90640, forms a 24 × 32 matrix of temperature readings which depict the distribution of heat around the frame of view as equation (2) illustrates. The example image of a human is shown in Fig.2. Due to the nature of the deep learning models, the thermal matrix is scaled between the minimum and maximum values of the available data, known as min–max scaling..

$$I_{norm}(x,y) = \frac{T(x,y)-T_{min}}{T_{max}-T_{min}} \times 255 \quad (2)$$

where, $T(x,y)$ is the temperature value at pixel location, T_{min} is the minimum temperature value in frame, T_{max} is the maximum temperature value in frame and I_{norm} is the normalized grayscale pixel intensity

The normalized thermal frame is then upsampled to 640 × 640 resolution to match the input requirement of the YOLO detection network. An Inferno colormap is applied to enhance temperature contrast and improve human detection performance.

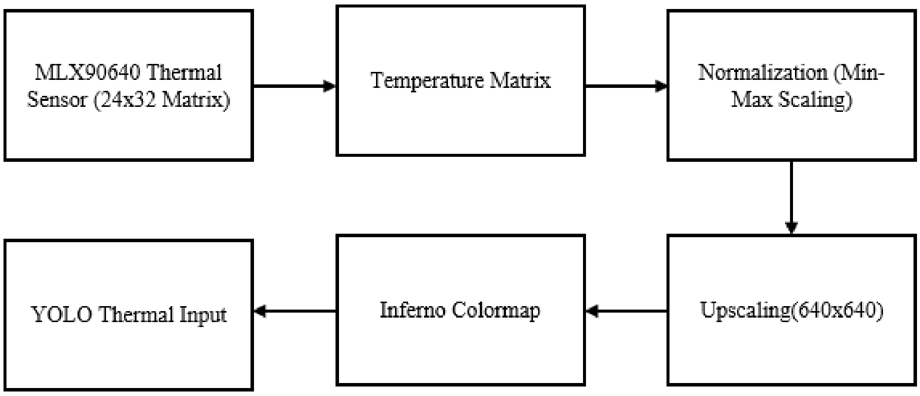

Fig.3. Thermal Image Preprocessing

C. RGB Image Acquisition and Preprocessing

The RGB video stream is obtained from an IP-based mobile camera connected to the Raspberry Pi through a wireless network. The captured frames are resized and normalized before being passed to the YOLO detection network. RGB detection complements thermal detection by providing visual features such as human body shape and motion patterns. This helps reduce false positives that may occur due to non-human heat sources such as machines or animals.

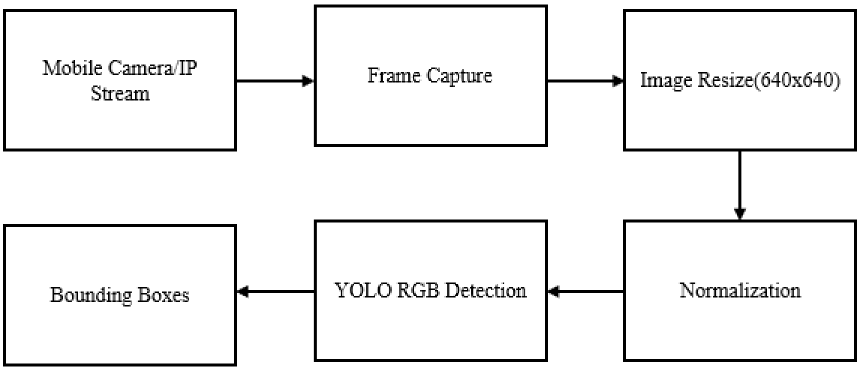

Fig.4. RGB Frame Acquisition and Processing Pipeline

D. YOLO Model Configuration

The RGB detection component employs YOLOv8 (nano variant) initialized with weights pretrained on the COCO dataset, retaining only the "person" class for downstream fusion. The thermal detection component uses a YOLOv8 model trained on thermal human signature data collected using the MLX90640 sensor under the operational conditions described in Section IV. Both models produce bounding boxes, class labels, and confidence scores as output. A per-model confidence threshold of 0.50 is applied to filter low-certainty detections before the fusion stage. Both models are optimized for inference on the Raspberry Pi 5 ARM Cortex-A76 processor using ONNX Runtime quantization.

E. RFID Based Authentication

Once a human target is confirmed through multimodal fusion, the system activates a 15-second authentication window. During this interval, the MFRC522 RFID module scans for RFID tags within the vicinity. If the scanned UID matches an entry in the authorized personnel database, the individual is classified as Authorized Personnel. Otherwise, the system classifies the individual as an Intruder and triggers an alert.

F. Intrusion Detection Authentication

Algorithm 1: Multimodal Intrusion Detection and Authentication

Input: Thermal frame T_t, RGB frame R_t, authorized UID database D_auth, IoU threshold θ = 0.4, confidence threshold δ = 0.50

Output: Classification label ∈ {No-Human, Authorized, Intruder}

Initialization

```
1.  I_norm ← MinMaxNormalize(T_t) × 255    //Thermal preprocessing
2.  I_upscaled ← Upscale(I_norm, 640×640)  //Resize for YOLO input
3.  B_T, C_T ← YOLO_Thermal(I_upscaled) //Thermal detection
4.  B_R, C_R ← YOLO_RGB(R_t)       //RGB detection
5.  if B_T = ∅ OR B_R = ∅ then
6.      return No-Human
7.  end if
8.  iou ← ComputeIoU(B_T, B_R)   //Spatial consistency
9.  if iou < θ then
10.     return No-Human            //False detection (noise)
11. end if
12. StartTimer(15 seconds)          //Trigger auth window
13. uid ← RFID_Scan(MFRC522)
14. if uid ∈ D_auth then
15.     return Authorized
16. Else
17.     TriggerAlarm()
18.     return Intruder
19. end if
```

End

The IoU threshold θ = 0.40 was selected based on empirical validation on the held-out development set, balancing suppression of false positives against sensitivity to partially overlapping detections caused by sensor misalignment. The 15-second RFID authentication window was determined through usability testing to allow sufficient time for tag presentation without introducing unacceptable security latency.

IV. Results

A. Experimental Setup

The proposed system was tested and validated in a Raspberry Pi 5 (ARM Cortex-A76, with 8 GB RAM) with Raspberry Pi OS (64-bit) operating system. The thermal stream was acquired by the MLX90640 sensor (24×32 native resolution, upscaled to 640×640 for YOLO input) and the RGB stream was acquired using a mobile IP camera at 640×480. Thermal and RGB frames were synchronized at the software level, under a maximum tolerance time offset of 50 milliseconds, and all experiments were conducted in an indoor controlled environment with three different light conditions: daytime (500–800 lux), night (0–5 lux) and low light (5–50 lux). We used the evaluation set of 1000 synchronized thermal RGB images (700/150/150 train/validation/test).

Table 1. Setup Description of the proposed system

Parameter	Description
Platform	Raspberry Pi 5, 8 GB RAM
Thermal sensor	MLX90640 (24x32 -> 640x640)
RGB camera	IP mobile camera, 640x480
Total image pairs	1000(600 Human/400 No-Human)
Human Samples	~600
Non-human samples	~400

Annotation format	YOLO bounding boxes
Preprocessing	Resize, normalize, noise filter
IoU fusion threshold	0.4
Confidence threshold	0.50

B. Performance Evaluation

The baseline comparison or ablation study of the project is shown in the following table

Table 2. Performance comparison of RGB system, Thermal System and Fusion System

Method	Accuracy	Precision	Recall	F1 Score
RGB Only	89.2%	91%	88%	89%
Thermal Only	91.5%	93%	90%	91%
Proposed Fusion	98.23%	99.1%	98.8 %	98.9 %

C. Qualitative Observations Across Conditions

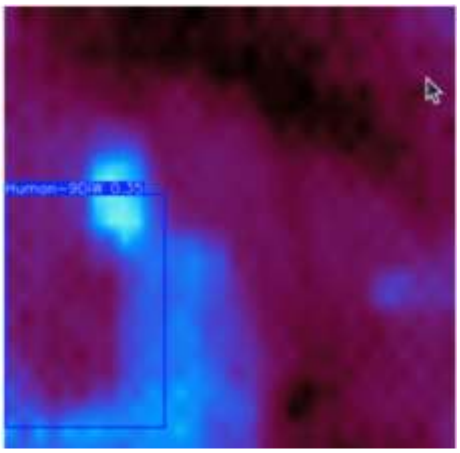

Fig.5. Human detected on Camera

Condition-wise disaggregation is reserved as a direction for future quantitative assessment in terms of a larger number of annotated images, but qualitative observations from testing sessions suggest that similar behaviour is observed across conditions. The RGB detection stream was more dominant in daytime scenes with thermal as the confirmatory modality. During the night and at low light levels, the thermal stream was used as the primary detection source and RGB was used to confirm spatial alignment. Such complementary behaviour is consistent with the results obtained from similar multimodal fusion results literature and has validated the design purpose of operating both the YOLO models simultaneously without any conditional switchover. While at low light scenarios, thermal camera was able to detect humans as shown in Fig.5 .
For a clear understanding of the observations and conditions, the analysis of latency of the system is shown in the Table 3.

Table 3. Latency Analysis

Parameter	Value
Inference Time(RGB)	~45 ms
Inference Time(Thermal)	~50 ms
Total Processing Time	~110 ms
Frame Rate	~9ms

V. Discussion

The experimental results presented demonstrate the high level of reliability that can be obtained using the proposed method to detect intrusion, compared to a single modality approach, through thermal imaging, RGB vision and RFID authentication. From the test results, the proposed fusion framework was able to achieve an overall accuracy of 98.23%, which outperforms the detection model based on RGB and the thermal detection model. The complimentary of the two sensing modalities is the reason for this improvement. RGB cameras offer rich information about the structure and texture of a scene under normal lighting conditions, while thermal imaging cameras can still be used in the dark or in areas where there is limited visibility, where the RGB cameras have poor performance. The IoU based fusion mechanism is also made robust for detection by verifying detections of both modalities before authentication. This approach can help reduce false alarms from environmental heat sources, moving machinery or partial occlusion that can trigger a false alarm in one sensor. The qualitative results for various lighting conditions also indicate that the system is dynamically advantageous of the strengths of the different modalities of sensing without any manual switching from one detection model to another. The RFID authentication layer gives additional security to the system as compared to most of the existing vision intruder detection systems. The suggested scheme is able to distinguish between the presence of a person and the presence of an unauthorised person, and therefore reduces unnecessary alarms and improves the usability of the system in industrial applications, warehouses or defense installations. The processing time measured is on the order of 110ms that is quite acceptable when the entire detection, fusion and authentication process cannot rely on cloud computing or GPU acceleration even on a Raspberry Pi 5. It captures the viability of the proposed architecture to be deployed at the edge branch environment with limited resources. While promising, there are some disadvantages to this. The native resolution of the thermal sensor MLX90640 is not optimal for the thermal feature extraction at a fine grain-level and the evaluation was conducted in a controlled indoor environment using a small thermal dataset. Future work is planned to be extended to a larger scale with outdoor testing under different environmental conditions, multiple subjects tested simultaneously, and detailed quantitative analysis, including the sensitivity of the IoU threshold for this test, and extended ablation tests to validate the contribution of each system component.

VI. Conclusion

This paper presented an edge-deployed pipeline to integrate thermal imaging, visual detection based on RGB images and person authentication based on RFID technology. A real-time multimodal intrusion detection system for restricted security areas was implemented. The dual-stream YOLO detection architecture and IoU spatial domain consistency checking cross-validated by single sensor can directly solve the main defects of single-sensor surveillance systems, such as the sensitivity of detection accuracy under poor light and environment. The system was tested on 1,000 synchronized images of color and thermal cameras with different light conditions (day, nite and low light) with an accuracy of 98.23%, precision of 0.9910, recall rate of 0.9888 and F1 score of 0.9899. The confusion-matrix analysis also demonstrates that the fusion mechanism with IoU removes false alarms caused by non-human sources of heat, and the second authentication mechanism, provided by the RFID authentication layer, removes false alarms caused by "authorized" users. A key characteristic of the proposed system is high accuracy intrusion detection system in the computing limitations of real world field installations, it is fully cloud independent and it does not require the presence of GPUs. This is the case for industrial security perimeters where infrastructure based solutions are often not feasible, and in environments such as access control and defense zones in warehouses. Current limitations include the limited native resolution of the MLX90640 sensor (24×32), evaluation only in controlled indoor environments, single person detection logic, no quantified latency profiling and no cross system baseline comparison. Section VI provides specific guidelines for future work under the stipulated restrictions.

VII. Future Scope

The proposed system will be developed further in four directions. This is partly due to higher spatial accuracy of higher resolution thermal sensors such as the FLIR Lepton 3.5 which removes "false negatives" in "edge of field" applications. Secondly, it will be tested outdoor, which could provide different conditions rain, wind, direct sunlight Last but not least, larger facilities (industrial campus, defense installation) will have the capability of multi-zone distributed sensing with nodes connected to a Raspberry Pi via IoT alert sent to security personnel in real time via GSM. In this work, next, a systematic ablation study will be performed, including the comparisons between the diversity of thermal-only, RGB-only, and fusion-only systems to quantify the contribution of each of these system components.

Acknowledgements
The authors would like to thank Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, to provide us with the opportunity to work on the project.
Funding
The research or project has received no external funding.
Conflict of Interest
The authors declare that there is no conflict of interest.

References
1. A. N. Wilson, K. A. Gupta, B. H. Koduru, A. Kumar, A. Jha, and L. R. Cenkeramaddi, "Recent advances in thermal imaging and its applications using machine learning: A review," IEEE Sensors Journal, vol. 23, no. 4, pp. 3395–3413, Feb. 2023, doi: 10.1109/JSEN.2023.3234335
2. S. P., V. Niranjani, A. R. H., and N. Soms, "Deep learning model for human detection – Enhanced YOLO comparative analysis," in Proc. Int. Conf. Computer Science and Engineering, 2023, doi: 10.1109/ICSES63445.2024.10763304
3. A.-A. Ulmămei, T. D'Adamo, C.-E. Vasile, and R. Hobincu, "Human detection in UAV thermal imagery: Dataset extension and comparative evaluation on embedded platforms," Journal of Imaging, vol. 11, no. 436, 2025, doi: 10.3390/jimaging11120436
4. W. Yang, H. Luo, K.-W. Tse, H. Hu, K. Liu, B. Li, and C.-Y. Wen, "Autonomous–targetless extrinsic calibration of thermal, RGB, and LiDAR sensors," IEEE Transactions on Instrumentation and Measurement, vol. 73, 2024, doi: 10.1109/TIM.2024.3480237
5. A. Nandakumar, L. D. Vanajakshi, and C. Lakshminarayanan, "Fusion of thermal and RGB images for traffic object detection," in Proc. IEEE Int. Conf. Image Processing Applications, 2023, doi: 10.1109/APSCON63569.2025.11144056
6. A. Ozcan and O. Cetin, "A novel fusion method with thermal and RGB-D sensor data for human detection," IEEE Access, vol. 10, pp. 68133–68145, 2022, doi: 10.1109/ACCESS.2022.3185402
7. A. Kiran, S. W. Prakash, B. A. Kumar, T. Sameeratmaja, Likhitha, and U. S. S. R. Charan, "Intrusion detection system using machine learning," in Proc. Int. Conf. Computer Communication and Informatics (ICCCI), 2023, doi: 10.1109/ICCCI56745.2023.10128363
8. C. Vij and H. Saini, "Intrusion detection systems: Conceptual study and review," in Proc. IEEE Int. Conf. Signal Processing, Computing and Control (ISPCC), 2021, doi: 10.1109/ISPCC53510.2021.9609481
9. M. Martin, A. Chong, F. Biljecki, and C. Miller, "Infrared thermography in the built environment: A multi-scale review," Renewable and Sustainable Energy Reviews, vol. 165, 2022, doi: 10.1016/j.rser.2022.112540
10. M. Brenner, N. H. Reyes, T. Susnjak, and A. L. C. Barczak, "RGB-D and thermal sensor fusion: A systematic literature review," IEEE Access, vol. 11, pp. 82410–82445, 2023, doi: 10.1109/ACCESS.2023.3301119
11. Y. Mohamed, A. Güneysu, S. Lemaignan, and I. Leite, "Multi-modal affect detection using thermal and optical imaging in a gamified robotic exercise," International Journal of Social Robotics, vol. 16, pp. 981–997, 2024, doi: 10.1007/s12369-023-01066-1
12. Q. Liu, M. Li, W. Wang, S. Jin, H. Piao, Y. Jiang, N. Li, and H. Yao, "Infrared thermography in clinical practice: A literature review," European Journal of Medical Research, vol. 30, 2025, doi: 10.1186/s40001-025-02278-z
13. J. Jeyalakshmi, A. J. Varma, and D. Saketh, "Robust face recognition in low-light environments using thermal imaging and deep learning techniques," in Proc. IEEE Int. Conf. Data, Computation and Communication (ICDCC), 2024, doi: 10.1109/ICDCC62744.2024.10961918
14. S. K. Ramareddy, "Integrating AI-Driven Data Visualization and Data Warehouse Analytics for Predictive Decision-Making in Financial and Stock Market Systems," 2026 International Symposium of Systems, Advanced Technologies and Knowledge (ISSATK), Hammamet, Tunisia, 2026, pp. 1-6, doi: 10.1109/ISSATK69667.2026.11566934.
15. C. H. Setjo, B. Achmad, and Faridah, "Thermal image human detection using Haar-cascade classifier", in Proc. 7th Int.l Annual Engineering Seminar (InAES), 2017, doi: 10.1109/INAES.2017.8068554
16. S. E. Kim, H. G. Lee, H. U. Ju, J. T. Rhee, T. W. Kang, and J. W. Jung, "A study on human recognition using general RGB and thermal imaging cameras in low-light environments", in Proc. 15th Int. Conf. on Information and Communication Technology Convergence(ICTC), 2024, doi: 10.1109/ICTC62082.2024.10826610
17. S. K. Subramaniam, M. S. Suryawanshi, N. M., and B. Marimuthu, "Evaluation of combined RGB and thermal imaging human detection model performance," in Proc. IEEE Int. Conf. Intelligent Processing and Advanced Computing Technologies (i-PACT), 2025, doi:10.1109/i-PACT65952.2025.11308029

18. V. Ghodmare, T. Kakade, S. Khan, S. Aghor, and P. Rane, “Edge computing-based real-time surveillance system with YOLOv8 object detection using NVIDIA Jetson Nano”, in Proc. 5th Int. Conf. on Pervasive Computing and Social Networking (ICPCSN), 2025, doi: 10.1109/ICPCSN65854.2025.11034845
19. X. Wangm W. Li and H. Li, “A context-encoder-based thermal anomaly detection network for dual-sensor surveillance”, in Proc. Int. Conf. on Machine Learning and Computer Application (ICMLCA), 2024, doi: 10.1109/ICMLCA63499.2024.10754432
20. Y. Li and J. Yao, “Research on a lightweight intrusion detection model based on multimodal fusion,” in Proc. Int. Conf. Artificial Intelligence and Security Systems, 2023, doi: 10.1109/ICNCIS67521.2025.11296132
21. S. Gelfert, “Body part detection in smoky environments with thermal camera using deep learning”, in Proc. 22nd Int. Conf. Control Automation and Systems (ICCAS), 2022, doi: 10.23919/ICCAS55662.2022.10003742
22. S. Gajula, "Intelligent Customer Churn Analytics in Digital Banking Using Advanced Machine Learning Models," 2025 1st International Conference on Emerging Trends in Information Systems and Informatics (ICETISI), Jakarta, Indonesia, 2025, pp. 1-6, doi: 10.1109/ICETISI67983.2025.11406030.
23. S. Peng, R. Xue, Y. Tong, Z. Wang, and H. Yang, “Multispectral object detection via edge-enhanced and frequency-aware fusion network,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 19, 2026, doi: 10.1109/JSTARS.2025.3648007
24. M. Mozafari, A. J. Law, J. R. Green, and R. A. Goubran, “Face detection in thermal images with improved spatial precision and temporal stability”, in Proc. IEEE Int. Conf. Biomedical and Health Informatics, 2024, doi: 10.1109/I2MTC53148.2023.10176080
25. El-Masry, J. (2026). Adaptive Explainable AI Framework for Trustworthy Decision Support in High-Stakes Intelligent Applications. Research Journal of Computer Systems and Engineering, 50–56.
26. S. Aldegheri and N. Bombieri, “Rapid Prototyping of Embedded Vision Systems: Embedding Computer Vision Applications into Low-Power Heterogeneous Architectures”, in 2018 International Symposium on Rapid System Prototyping(RSP), doi: 10.1109/RSP.2018.8631995
27. Sunkara, S. P. (2025). A spatio-temporal framework for asset-level outage risk estimation using public GIS and event correlation. International Journal of Computer Engineering and Technology, 16(1), 4211–4227. https://doi.org/10.34218/IJCET_16_01_286
28. B. Khalid, A. Khan, M. Akram and S.Batool, “Person Detection by Fusion of Visible and Thermal Images Using Convolutional Neural Network,” in 2019 2nd International Conference on Communication and Digital systems (C-CODE), doi: 10.1109/C-CODE.2019.8680991