Real-Time Human Detection Using YOLOv11 for Early Warning in Beach Safety Zones

Kusrini
Informatics Study Program, Computer Science Faculty, Universitas Amikom Yogyakarta, Sleman, Indonesia
kusrini@amikom.ac.id (corresponding author)

Bayu Setiaji
Department of Informatics, Computer Science Faculty, Informatics Study Program, Universitas Amikom Yogyakarta, Sleman, Indonesia
bayusetiaji@amikom.ac.id

Elik Hari Muktafin
Application Software Engineering Study Program, Politeknik Prestasi Prima, Jakarta Timur, Indonesia
elixropas@gmail.com

Anna Baita
Informatics Study Program, Computer Science Faculty, Universitas Amikom Yogyakarta, Sleman, Indonesia
anna@amikom.ac.id

Bima Pramudya Asaddulloh
Informatics Study Program, Computer Science Faculty, Universitas Amikom Yogyakarta, Sleman, Indonesia
bima@students.amikom.ac.id

Dwi Ahmad Dzulhijjah
Department of Cyber Physical System, Applied Graduate School, Politeknik Elektronika Negeri Surabaya, Surabaya, Indonesia
dwiahmaddzulhijjah@gmail.com

Faris Khairil Murtadho
Communication Science Study Program, Economic and Social Faculty, Universitas Amikom Yogyakarta, Sleman, Indonesia
faristadho1@gmail.com

Received: 27 September 2025 | Revised: 14 December 2025 and 17 January 2026 | Accepted: 19 January 2026

 DOI: https://doi.org/10.48084/etasr.15025

ABSTRACT

This study proposes a novel approach for beach safety monitoring by implementing a real-time human object detection system using the state-of-the-art YOLOv11 model. The main objective is to detect human subjects—specifically their heads and shoulders—and classify their locations into predefined safety zones: safe, caution, and danger. The system is designed to trigger an early warning protocol when individuals are identified within the danger zone, providing a crucial intervention window. The study utilizes a meticulously pre-processed custom dataset, which has been tailored for this specific application. The performance of the YOLOv11 model is then quantitatively evaluated using standard metrics, including Precision, Recall, and mean Average Precision @ 0.50-0.95 Intersection over Union (mAP@0.50-0.95 IoU). The findings demonstrate the feasibility of employing deep learning models for proactive risk

management, with the proposed system achieving a mAP@0.50-0.95 of 67.6% and an impressive inference speed of 2.1 ms per image. The study offers a scalable and efficient solution to enhance visitor safety and prevent potential accidents in dynamic beach environments.

Keywords-maritime surveillance; YOLOv11; artificial intelligence

I. INTRODUCTION

Coastal zones are significant for global tourism and recreation, yet they present inherent safety risks to visitors, including strong currents, rip tides, and other environmental hazards [1-3]. Traditional surveillance and warning systems often rely on human lifeguards and static signage, which can be limited by human fatigue, visibility constraints, and the large size of monitored areas [4, 5]. The increasing demand for enhanced public safety and proactive risk management in these environments has driven the need for more advanced, automated solutions [6, 7]. Object detection technology, particularly with the rise of deep learning models, contributes to addressing these challenges by providing real-time, continuous monitoring and early warning capabilities that can supplement or even augment traditional safety protocols. The present research aims to leverage these technological advancements to develop a robust, intelligent surveillance system for beach safety.

This study proposes a novel approach to proactive beach safety by implementing a real-time human object detection system based on the state-of-the-art YOLOv11 deep learning model. The system is designed to continuously monitor beachgoers, specifically detecting human subjects by their heads and shoulders, and to classify their locations into three predefined safety zones: safe, caution, and danger. The primary innovation lies in the automated issuance of an early warning when an individual enters a designated danger zone, enabling intervention by safety personnel. To achieve this, the model is trained on a custom-processed dataset, and its performance is evaluated using a comprehensive suite of metrics, including Precision, Recall, and mAP@0.50-0.95 IoU. The research addresses a critical gap in existing solutions by focusing on a scalable, automated system capable of operating effectively in dynamic and challenging coastal environments.

Authors in [8] introduced YOLO-SS, an enhanced version of the YOLO framework specifically designed to improve small object detection in remote sensing imagery. The model incorporates an optimized backbone, a restructured loss function, and an asymmetric sample weighting strategy to better capture high-quality positive samples while reducing background noise. Applied to the AI-TOD dataset, YOLO-SS achieved an AP50 score of 0.535, surpassing YOLOv6L by 13.4% and outperforming several other state-of-the-art detectors. However, limited ablation studies were provided to isolate the contribution of each proposed component, which reduces clarity on what drives the performance gains.

Authors in [9] proposed YOLO-EDGE, an improved object detection algorithm tailored for traffic scenarios by enhancing the YOLOv8 framework with GhostConv, a novel C2FEMA module, and a lightweight SlimNECK structure. These modifications reduced computational complexity while preserving rich feature extraction, making the model more efficient for real-time deployment in autonomous driving systems. The experimental results on the PASCAL VOC dataset showed that YOLO-EDGE reduced model size by 7% to 20.9 million parameters, while improving mAP@0.5 by 2.2% and mAP@0.5:0.95 by 3.2%. Furthermore, on the UA-DETRAC dataset, the model achieved an mAP@0.5 of 64.9%, demonstrating strong performance in multi-object traffic detection. However, the study primarily benchmarked against YOLOv8s and did not provide extensive comparisons with other state-of-the-art lightweight detectors, which limits the scope of its claims.

Authors in [10] introduced a real-time human object detection system for beach safety monitoring using the YOLOv11 model, focusing on detecting heads and shoulders of beach visitors and classifying them into safe, caution, and danger zones. The dataset consisted of 502 original images collected at Drini Beach, which were augmented to 1,906 images and split into 70% training, 20% validation, and 10% testing sets. The YOLOv11s model achieved an mAP of 47.0 at 640-pixel resolution, with an inference speed of 2.5 ms on TensorRT-based devices and a parameter size of only 9.4 million, rendering it efficient for edge deployment. Nevertheless, inconsistencies in annotation styles were reported, which limited the model's optimal performance, reducing the reliability of its evaluation.

Authors in [11] presented an improved lightweight object detection model, YOLOv7-CA, which integrates a coordinate attention mechanism and depth-wise separable convolutions to enhance feature extraction while reducing computational cost. The model was evaluated on the VisDrone2019 dataset, achieving an mAP@0.5 of 47.6% and an mAP@0.5:0.95 of 26.4%, which represents improvements of 2.8% and 1.9%, respectively, compared to the baseline YOLOv7-tiny. In addition, the parameter size was reduced by 38.2%, and FLOPs decreased by 34.6%, making the model more suitable for deployment on resource-constrained devices. However, the study primarily focused on VisDrone2019 and did not validate performance across other benchmark datasets, which limits the generalizability of its findings.

Authors in [12] proposed an improved object detection framework based on YOLOv8, introducing a lightweight attention mechanism and feature fusion strategy to enhance detection accuracy while maintaining efficiency. The model was evaluated on the VisDrone2019 dataset, where it achieved an mAP@0.5 of 52.6% and an mAP@0.5:0.95 of 31.4%, outperforming the baseline YOLOv8s by 3.1% and 2.4%, respectively. In addition, the parameter count was reduced by 15.7% and FLOPs decreased by 18.9%, demonstrating the model's suitability for deployment on edge devices with limited computational resources. Nonetheless, limited ablation experiments were provided, making it difficult to isolate the contribution of each proposed component.

Recent YOLO-based approaches have also been explored within the context of beach safety monitoring. Authors in [13] employed YOLOv8 to detect drowning risks in coastal environments, demonstrating that lightweight detectors can effectively operate in dynamic outdoor scenes and offering insights into early-warning implementations in real beach conditions. Authors in [14] provided a comprehensive review of computer-vision techniques for drowning detection, highlighting key challenges, such as partial occlusion, wave interference, and small-object visibility, while emphasizing the need for more robust and real-time detection frameworks suitable for aquatic environments. These studies underscore the growing relevance of deep learning models for coastal safety applications and further justify the adoption of the more advanced YOLOv11 architecture in the present work, which aims to improve detection robustness and real-time performance in similarly challenging settings.

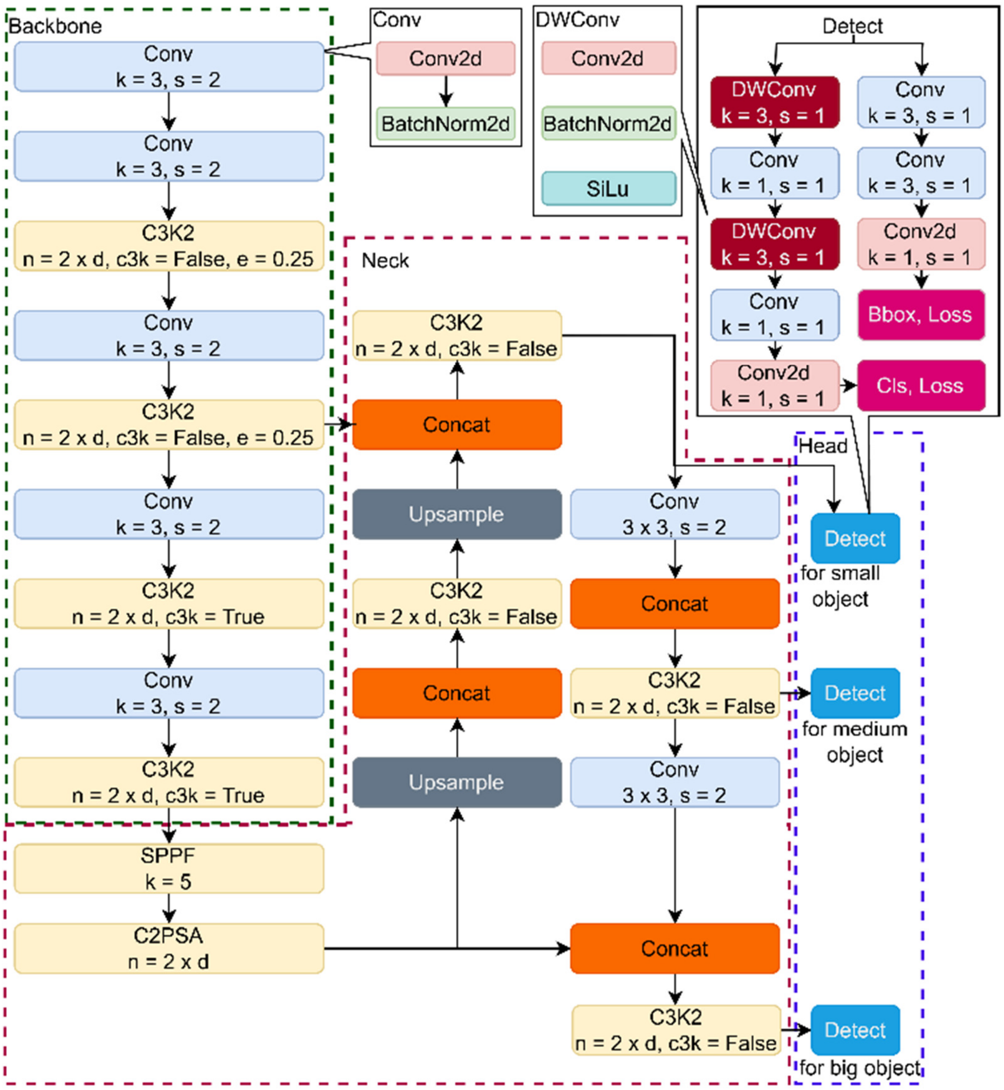

Fig. 1. Yolo 11 architecture.

Research on object detection in maritime environments often focuses on specific and controlled scenarios, including search and rescue operations using drones or human detection with older model versions such as YOLOv3, YOLOv4, and YOLOv5. There is a significant gap in the literature regarding integrated systems capable of continuous, real-time, land-based surveillance in dynamic coastal environments. Furthermore, these studies often do not present comprehensive performance metrics, such as the mAP@0.50-0.95s, which are important for assessing model robustness. To fill this gap, the present study proposes and implements an advanced solution for fully integrated coastal safety. Its main contribution is the first-time use of the YOLOv11 architecture for real-time human detection in coastal safety zones. Not only was a detection model developed, but also a complete integrated system was built, including a specialized hardware prototype with autonomous resources and an easy-to-use monitoring dashboard. Through this approach, the proposed model achieved outstanding performance with an mAP@0.50-0.95 of 0.676 and an extremely fast inference speed of only 2.1 ms per image, setting a new standard for efficiency and reliability in the field.

II. PROPOSED METHOD

The current study follows a comprehensive methodology, with the sequential steps from data acquisition to model evaluation. The process is systematically divided into several key phases: dataset preparation, model initialization, model training, and performance evaluation. This structured approach ensures reproducibility and clarity in understanding the experimental framework.

The system architecture integrates an outdoor surveillance camera, on-device inference using an embedded edge-computing device, and an audio-based alert mechanism for real-time early-warning activation. Real-time inference was deployed on an edge device configured to operate autonomously in outdoor conditions, enabling continuous monitoring without cloud connectivity.

A. Dataset Preparation

The research process began with collecting image and activity data from beach visitors as the primary data. The dataset was collected exclusively at Drini Beach, which may limit generalization to other coastal environments with different visual characteristics, such as wave dynamics, sand color, and crowd density. The collected dataset was then annotated using the Roboflow platform, with a primary focus on human objects within the water area [15-17]. This annotation process is illustrated in Figure 2, which provides examples of bounding boxes.

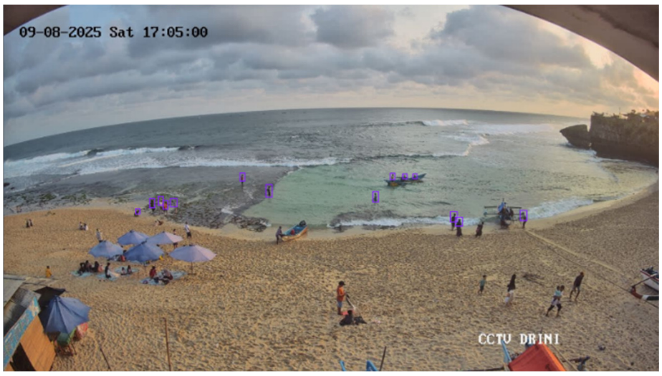

Fig. 2. Annotations on bright lighting.

After standardizing annotation styles, the model exhibited improved confidence scores and reduced variance in bounding box predictions, indicating more stable detection performance. The original dataset consisted of 502 images, which were then augmented to a total of 1906 images. The data were divided into three sets with specific proportions: 70% for training, 20% for validation, and 10% for testing [18, 19]. This division aims to ensure that the model is trained and evaluated effectively before being applied in a real-world environment.

B. Object Detection

In this study, object detection was performed using the YOLO11s model, which is a lightweight and efficient version of the latest generation of YOLO. YOLO11s is designed to detect objects in real time with high accuracy and low latency, making it suitable for applications that require speed and computational efficiency [20]. This model uses an optimized backbone and neck architecture to enhance feature extraction, enabling it to recognize objects with greater precision even under complex lighting or background conditions [21, 22].

YOLO11s was trained on an NVIDIA A100 40 GB GPU using Python 3.10, PyTorch 2.5.1+cu124, and CUDA 12.4, following the Ultralytics YOLOv11 training framework. Essentially, YOLO11s was trained using the COCO dataset, which comprises 80 object classes. In this experiment, the model was trained on a dataset specifically prepared for detecting beach visitors [23]. Inference was executed on an NVIDIA Jetson Orin NX (16 GB) running Ubuntu 22.04, JetPack 6.1, PyTorch 2.3.0 (ARM64), and TensorRT acceleration. Based on the validation results, YOLO11s obtained an mAP value of 47.0 at a resolution of 640 pixels, with an inference time of around 2.5 ms [24]. With a parameter efficiency of only 9.4 million, this model is ideal for deployment on edge devices or systems with limited resources.

C. Performance Evaluation

To quantitatively assess the effectiveness of the trained YOLOv11 model, a comprehensive performance evaluation was conducted using the model.val() method [22]. This critical stage involved feeding the model with the previously unseen validation dataset, as precisely defined within the data.yaml configuration file. The primary metrics extracted from these detailed evaluation results offered a robust measure of the model's accuracy and reliability. Foremost among these is the mAP@0.50-0.95 IoU, which stands as the most comprehensive and widely accepted metric in the field of object detection. This value represents the average of Average Precision (AP) scores calculated across a range of IoU thresholds, specifically from 0.50 to 0.95. A higher mAP across this range inherently indicates superior overall detection accuracy, accounting for varying degrees of localization strictness. Complementing this, mAP @ 0.50 IoU (mAP50) was also computed; this metric specifically assesses the mAP when a less stringent IoU threshold of 0.50 is applied, often referenced as the PASCAL VOC mAP [25].

Beyond these aggregate measures, the model's performance is further characterized by Accuracy, which represents the proportion of correctly classified instances [26]. Additionally, the AP at 0.50 IoU (AP50) is reported for each individual class, directly correlating with the mAP50 but specific to that category. These detailed, individual class metrics are invaluable for understanding the specific strengths and any weaknesses the model might exhibit across different object types, ultimately serving as a crucial benchmark for the model's efficacy in real-world application scenarios [27]. Accuracy and mAP_{50} are calculated by:

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (1)$$

$$\mathrm{mAP}_{50} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_{50,i} \quad (2)$$

III. RESULTS AND DISCUSSION

The initial test results of the beach visitor object detection system showed varying performance at the early implementation stage. The YOLOv11 model was trained with a data ratio of 70% for training, 20% for validation, and 10% for testing. Overall, the trained model successfully detected beach visitors.

A. Model Performance

The trained model successfully detected beach visitors, although the initial test results showed that the system had not fully achieved the expected targets. Inconsistent annotation styles among annotators, some providing very tight bounding boxes while others tended to be loose, caused the object detection system to struggle in achieving optimal performance. Annotation inconsistencies, particularly between tight and loose head–shoulder bounding boxes, were identified and corrected to improve training stability. Environmental factors, particularly tidal variations, were observed to reduce detection accuracy by approximately 2–3 percentage points due to decreased visible object scale and wave interference. During high tide, reduced visibility of the subject's head and increased wave reflections contributed to a higher false-negative rate.

The model's detection performance is illustrated in Figure 3, which displays an example image with detected visitor objects. Furthermore, the detection output not only includes labels and confidence scores but also entails the object's spatial coordinates. These coordinates are extracted after the bounding box is combined with the object tracking module (ByteTrack) and are projected onto a geographic coordinate system. The visualization of coordinate extraction and mapping to safety zones is shown in Figure 4.

Fig. 3. Beach visitor detection.

B. Performance Metric Analysis

Table I presents the comprehensive performance results. The model achieved a high mAP@0.50-0.95 value of 0.676, indicating its superior detection accuracy across a range of IoU thresholds. The mAP@50 metric reached an impressive 0.894. The Precision value of 0.959 shows that nearly all detected objects were correct, significantly minimizing false positives. A Recall value of 0.753 confirms the model's robust capability to both correctly identify objects and find a high proportion of them. Additionally, the model demonstrated remarkable efficiency with an average inference speed of only 2.1 ms per image. This speed, equivalent to approximately 476 Frames Per Second (FPS), confirms its suitability for real-time applications that demand both speed and computational efficiency.

Fig. 4. Extracting bounding box coordinates.

TABLE I. PERFORMANCE METRICS OF THE PROPOSED YOLOV11-BASED DETECTION SYSTEM

Metric	Value
Precision	95.9%
Recall	75.3%
mAP@50	89.4%
mAP@50-95	67.6%

As depicted in Figure 5, the loss function values, including train/box_loss, train/cls_loss, and train/dfl_loss, gradually decreased, indicating the model's effective learning from the ground truth data. The steady decrease in validation loss values also confirms that the model did not suffer from overfitting.

C. Comparative Analysis and Discussion

The proposed YOLOv11-based system offers a robust solution for real-time human object detection in dynamic beach environments, effectively addressing key limitations found in prior works. As presented in Table II, the proposed model demonstrates superior performance compared to existing schemes. Authors in [8, 11] focused on improving small object detection using YOLO-SS and YOLOv7-CA, respectively. While their models demonstrated improved performance on specific datasets, the mAP@0.50 scores were 53.5% and 47.6%, which are significantly lower than the proposed model's 89.4%. The research on traffic and drone datasets also yielded lower mAP@0.50 scores of 64.9% and 52.6%, respectively [12]. This highlights a notable performance gap between the present model and these previous lightweight detection frameworks.

The use of the latest YOLOv11 model in the present study incorporates architectural advancements that improve efficiency and accuracy. This is particularly evident in the mAP@0.50-0.95 value, which reached 67.6% for the proposed system. In contrast, authors in [11, 12] only achieved 26.4% and 31.4%, respectively, on this metric. This demonstrates the proposed model's superior capability to maintain performance across a wide range of detection strictness, a key factor for reliable early warning systems. Furthermore, the inference speed is a significant metric for real-time systems. The model demonstrated efficiency with an average inference speed of only 2.1 ms per image. This sets a new standard for on-device deployment, as it is equivalent to approximately 476 FPS, ensuring that the system can provide instant alerts. In contrast, authors in [10] reported an inference time of 2.5 ms, while other studies did not specify such metrics, limiting the scope of their claims for real-time applications.

TABLE II. COMPARISON BETWEEN THE PROPOSED METHOD AND EXISTING SCHEMES

Study	Best model	Precision	Recall	mAP@50	mAP@50-95
[8]	YOLO-SS	-	-	53.50	-
[9]	YOLO-EDGE	-	-	64.90	-
[10]	YOLOv11			47.00	
[11]	YOLOv7-CA			47.60	26.40
[12]	YOLOv8			52.60	31.40
Proposed method	YOLOv11	95.90	75.30	89.40	67.60%

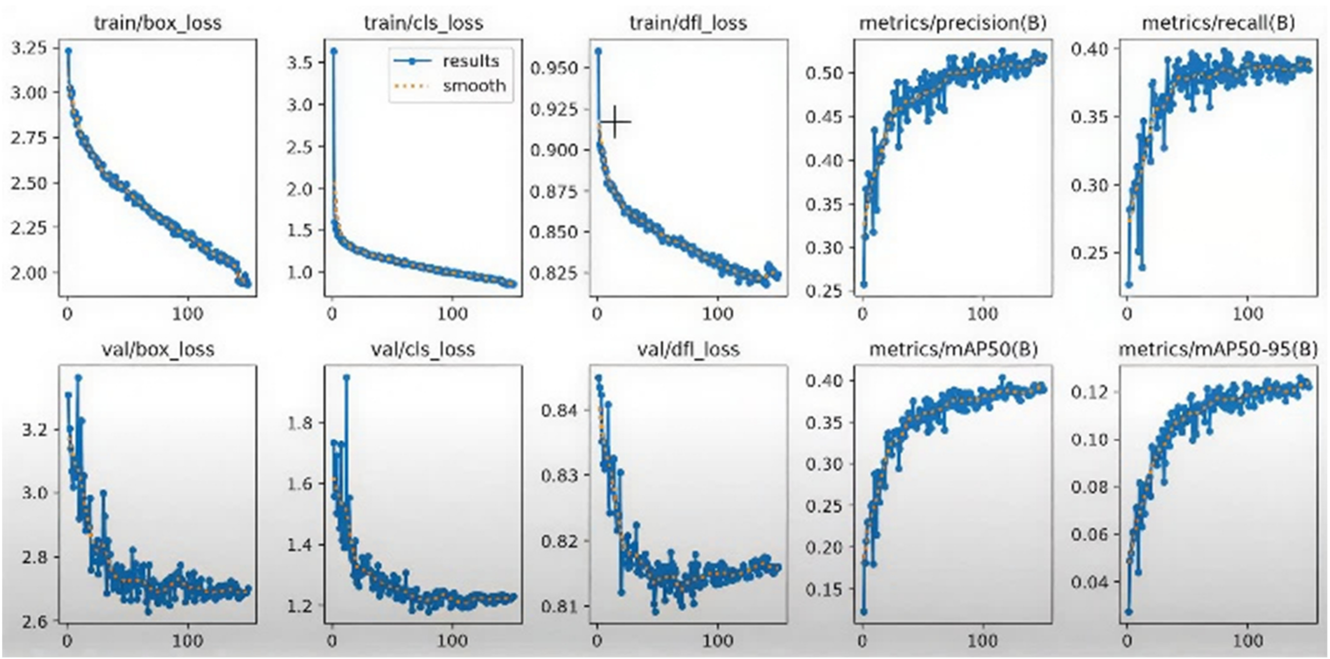

Fig. 5. Initial test results.

Accurate person localization in beach environments is important for early risk identification, as drowning incidents can develop rapidly and are often difficult to detect visually at a distance. Compared to infrared-based systems, which are sensitive to ambient temperature variations and may degrade under strong sunlight or reflective water surfaces, vision-based deep learning approaches can operate effectively in daylight conditions and provide precise spatial information relative to predefined safety zones. This capability enables more reliable risk assessment and timely intervention in dynamic beach environments.

IV. CONCLUSION

This study successfully developed and evaluated a real-time human object detection system aimed at improving visitor safety in beach environments. The application of the YOLOv11 deep learning model effectively detects human heads and shoulders and classifies their positions relative to predetermined safety zones. The YOLOv11-m model achieved a mean Average Precision @ 0.50-0.95 Intersection over Union (mAP@0.50-0.95 IoU) score of 67.6%, demonstrating reliable detection performance across different object scales. Real-time inference tests showed an average processing time of approximately 2.1 ms per image on the training GPU and 20–40 ms per frame on the edge device. The main contribution of this research lies in a proactive surveillance framework that can provide early warnings when individuals enter high-risk areas, thereby reducing potential hazards and providing significant improvements over traditional monitoring methods. Future work will explore expanding the dataset to multiple beach locations to improve model robustness and adaptability across diverse environmental conditions.

ACKNOWLEDGMENT

This work was funded by the Ministry of Higher Education, Science and Technology under contract no. 126/C3/DT.05.00/PL/2025, 0498.21/LL5-INT/AL.04/2025, and 006/KONTRAK-LPPM/AMIKOM/VI/2025. The authors gratefully acknowledge the invaluable support of all partner institutions and team members.

DATA AVAILABILITY STATEMENT

The dataset used in this study is available from the corresponding author upon reasonable request.

REFERENCES

[1] S. Borra *et al.*, "Identifying Rip Channels Along RK Beach, Visakhapatnam Using Video and Satellite Imagery Analysis," *Journal of the Indian Society of Remote Sensing*, vol. 51, no. 2, pp. 405–423, Feb. 2023, https://doi.org/10.1007/s12524-023-01666-4.

[2] A. O. Nur'aini and B. W. Mutaqin, "Morphometric and Sediment Characterization of Beach Cusps in Connection to Rip Currents Along Indonesia's Parangtritis Coast," *Journal of Earth System Science*, vol. 134, no. 1, Jan. 2025, Art. no. 28, https://doi.org/10.1007/s12040-024-02477-y.

[3] B. Castelle *et al.*, "Environmental Controls on Lifeguard-Estimated Surf-zone Hazards, Beach Crowds, and Resulting Life Risk at a High-Energy Sandy Beach in Southwest France," *Natural Hazards*, vol. 120, no. 2, pp. 1557–1576, Jan. 2024, https://doi.org/10.1007/s11069-023-06250-0.

[4] R. Irvine, W. Stephenson, and S. M. Mager, "Beach User Competencies in Rip Current Identification Using Videos and Photographs on a New Zealand Surf Beach," *Natural Hazards*, vol. 121, no. 6, pp. 7707–7726, Apr. 2025, https://doi.org/10.1007/s11069-024-07089-9.

[5] M. Z. Yusof *et al.*, "Public Understanding of Rip Current and Beach Safety at Teluk Cempedak Recreational Beach in Pahang, Malaysia," *Natural Hazards*, vol. 115, no. 1, pp. 489–506, Jan. 2023, https://doi.org/10.1007/s11069-022-05563-w.

[6] Y. Li *et al.*, "Analysis and Machine-Learning-Based Prediction of Beach Accidents on a Recreational Beach in China," *Anthropocene Coasts*, vol. 7, no. 1, Dec. 2024, Art. no. 29, https://doi.org/10.1007/s44218-024-00067-5.

[7] K.-M. Ha, "Improving Disaster Management in International Tourism," *Management Review Quarterly*, vol. 74, no. 3, pp. 1423–1437, Sept. 2024, https://doi.org/10.1007/s11301-023-00338-4.

[8] Q. Tang *et al.*, "YOLO-SS: Optimizing YOLO for Enhanced Small Object Detection in Remote Sensing Imagery," *The Journal of Supercomputing*, vol. 81, no. 1, Jan. 2025, Art. no. 303, https://doi.org/10.1007/s11227-024-06765-8.

[9] Y. Li *et al.*, "YOLO-EDGE: An Object Detection Algorithm for Traffic Scenarios," *The Journal of Supercomputing*, vol. 81, no. 7, May 2025, Art. no. 840, https://doi.org/10.1007/s11227-025-07275-x.

[10] Y. Xiao and N. Di, "SOD-YOLO: A Lightweight Small Object Detection Framework," *Scientific Reports*, vol. 14, no. 1, Oct. 2024, Art. no. 25624, https://doi.org/10.1038/s41598-024-77513-4.

[11] B. Ren, Z. Xu, J. Zhao, R. Hao, and J. Zhang, "Complex Dark Environment-Oriented Object Detection Method Based on YOLO-AS," *Scientific Reports*, vol. 15, no. 1, July 2025, Art. no. 21873, https://doi.org/10.1038/s41598-025-07348-0.

[12] W. Chen, J. Liu, T. Liu, and Y. Zhuang, "PCPE-YOLO with a Lightweight Dynamically Reconfigurable Backbone for Small Object Detection," *Scientific Reports*, vol. 15, no. 1, Aug. 2025, Art. no. 29988, https://doi.org/10.1038/s41598-025-15975-w.

[13] N. Alharbi, L. Aljohani, A. Alqasir, T. Alahmadi, R. Alhasiri, and D. Aldajan, "Improved Automatic Drowning Detection Approach with YOLOv8," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18070–18076, Dec. 2024, https://doi.org/10.48084/etasr.8834.

[14] N. Alharbi, "Exploring Advance Approaches for Drowning Detection: A Review," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 16032–16039, Aug. 2024, https://doi.org/10.48084/etasr.7804.

[15] A. K. Zhuravlyov and K. A. Grigorian, "Automatic Annotation of Training Datasets in Computer Vision Using Machine Learning Methods," *Automatic Documentation and Mathematical Linguistics*, vol. 58, no. S5, pp. S279–S282, Dec. 2024, https://doi.org/10.3103/S0005105525700347.

[16] S. Bhalla, A. Kumar, and R. Kushwaha, "Analysis of Recent Techniques in Marine Object Detection: A Review," *Multimedia Tools and Applications*, vol. 84, no. 19, pp. 20339–20437, July 2024, https://doi.org/10.1007/s11042-024-19782-9.

[17] J. Peng, K. Lv, G. Wang, W. Xiao, T. Ran, and L. Yuan, "MLSA-YOLO: A Multi-Level Feature Fusion and Scale-adaptive Framework for Small Object Detection," *The Journal of Supercomputing*, vol. 81, no. 4, Feb. 2025, Art. no. 528, https://doi.org/10.1007/s11227-025-06961-0.

[18] A. K. Salih, A. K. Faraj, M. A. Ahmed, and A. N. A. Al-Hasnawi, "The Impact of Data Splitting Strategy on Drilling Rate Prediction in the Rumaila Oil Field," *Petroleum Chemistry*, vol. 64, no. 7, pp. 781–786, July 2024, https://doi.org/10.1134/S0965544124050025.

[19] H. Babaei, M. Zamani, and S. Mohammadi, "The Impact of Data Splitting Methods on Machine Learning Models: A Case Study for Predicting Concrete Workability," *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, June 2025, Art. no. 21, https://doi.org/10.1007/s44379-025-00021-3.

[20] A. Sutikno, Sugiharto, and R. Kusumaningrum, "Enhanced Automatic License Plate Detection and Recognition using CLAHE and YOLOv11 for Seat Belt Compliance Detection," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 20271–20278, Feb. 2025, https://doi.org/10.48084/etasr.9629.

[21] Q. Deng, L. Du, W. Han, W. Ren, R. Yu, and J. Luo, "CB-YOLO: Composite Dual Backbone Network for High-frequency Transformer Coding Defect Detection," *Signal, Image and Video Processing*, vol. 18, no. 6–7, pp. 5535–5548, Aug. 2024, https://doi.org/10.1007/s11760-024-03253-7.

[22] P. Gupta and U. Singh, "Evaluation of Several YOLO Architecture Versions for Person Detection and Counting," *Multimedia Tools and Applications*, vol. 84, no. 30, pp. 37353–37376, Feb. 2025, https://doi.org/10.1007/s11042-025-20662-z.

[23] F. Huang, J. Yao, C. Su, and S. Xu, "An Online Multi-Camera Multi-Vehicle Tracking Using Lightweight YOLO11 and Improved Association Strategies," *Journal of King Saud University Computer and Information Sciences*, vol. 37, no. 7, Sept. 2025, Art. no. 180, https://doi.org/10.1007/s44443-025-00190-4.

[24] X. Zhang and B. Li, "Tennis Ball Detection Based on YOLOv5 with TensorRT," *Scientific Reports*, vol. 15, no. 1, July 2025, Art. no. 21011, https://doi.org/10.1038/s41598-025-06365-3.

[25] M.-H. Bae, S.-W. Park, J. Park, S.-H. Jung, and C.-B. Sim, "YOLO-RACE: Reassembly and Convolutional Block Attention for Enhanced Dense Object Detection," *Pattern Analysis and Applications*, vol. 28, no. 2, June 2025, Art. no. 90, https://doi.org/10.1007/s10044-025-01471-4.

[26] R. A. Elsheikh, M. A. Mohamed, A. M. Abou-Taleb, and M. M. Ata, "Accuracy is not enough: A Heterogeneous Ensemble Model Versus FGSM Attack," *Complex & Intelligent Systems*, vol. 10, no. 6, pp. 8355–8382, Dec. 2024, https://doi.org/10.1007/s40747-024-01603-z.

[27] J. Terven, D.-M. Cordova-Esparza, J.-A. Romero-González, A. Ramírez-Pedraza, and E. A. Chávez-Urbiola, "A Comprehensive Survey of Loss Functions and Metrics in Deep Learning," *Artificial Intelligence Review*, vol. 58, no. 7, Apr. 2025, Art. no. 195, https://doi.org/10.1007/s10462-025-11198-7.