JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

A YOLO26-Based Human Detection Framework for High-Entropy Disaster Environments

[1]Abdulrashid Sani, [2]Bashar Aliyu Yauri, [3]Muhammad Garba, [4]Anas Tukur Balarabe
[1&4]Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria,
[2&3]Department of Computer Science, Faculty of Computing, Abdullahi Fodio University of Science and Technology, Aliero, Kebbi State

ABSTRACT

The rapid and accurate identification of human presence is a critical prerequisite for the effective disaster management and security monitoring, particularly in crowded environments. While traditional object detection frameworks often suffer from high latency and reduced sensitivity in cluttered scenarios, this study evaluates the efficacy of the YOLO26 architecture an up-to-date, end-to-end, and NMS-free model released in January 2026 for the specific task of human detection in a crowded environment. Employing a C2 (Combination to Application) dataset of 10,215 images reflecting diverse catastrophic domains, including structural building collapses, floods, and dense vegetation area and fire hazards. Experimental results demonstrate that the YOLO26 model achieved a peak mean Average Precision (mAP@0.5) of 0.89 and a harmonic reliability (F1-score) of 0.87. Analysis of the F1-Confidence curves indicates a robust operational hill, maintaining an F1-score above 0.80 across a broad threshold range (0.2–0.6). A functional prototype was successfully deployed via a Gradio-based interface, providing real-time inference and automated human counting with an operational recall of 0.84 at a 0.50 confidence threshold. These findings suggest that the YOLO26 framework offers a favourable balance of localization precision and deployment efficiency, making it suitable for the demanding requirements of search and rescue operations. Future work will involve a comparative performance evaluation between the convolutional-based YOLO26 and transformer-based architectures (Detection Transformers) to further investigate global context modelling in extreme disaster zones.

ARTICLE INFO
Article History
Received: December, 2025
Received in revised form: January, 2026
Accepted: February, 2026
Published online: March, 2026

KEYWORDS
Human Detection, YOLO26, Disaster Management, Computer Vision, High-Entropy Environments

INTRODUCTION

Recent advancements in deep learning have transformed the analysis of visual data, enabling systems to automatically identify and categorize objects within both images and real-time video streams. Among its most life-critical applications is human detection, a field where model accuracy and latency directly influence the success of safety, security, and emergency response operations [1], [2], [3]. In the Nigerian context, the demand for robust, localized human detection has reached a critical threshold.

The nation frequently faces diverse natural and man-made disasters, such as the catastrophic 2022 floods in Jigawa and Adamawa states, which highlighted the significant difficulty of locating victims in partially submerged or high-occlusion environments [4]. Furthermore, recurring fire outbreaks in Internally Displaced Persons (IDP) camps and structural collapses in urban centers like Kano emphasize the need for intelligent systems capable of identifying individuals in high-risk zones where visibility is severely compromised [5].

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Despite the steady progression of the YOLO (You Only Look Once) family of models, a significant research gap persists regarding their reliability in high-entropy environments. Current frameworks often struggle with high false-negative rates in specific operational contexts, such as smoke-filled and the dense vegetation of insurgent-prone forests, where background clutter frequently mimics human features [6], [7]. While contemporary benchmarks for UAV-mounted systems have achieved a mAP50 of 0.807 [6], these models remain sensitive to the extreme occlusion and environmental stressors characteristic of local disaster zones. Consequently, there is an urgent need for an empirically validated, high-precision detection framework capable of maintaining architectural stability under these unique conditions.

To address this challenge, this study evaluates the deployment-readiness of the YOLO26 architecture, a model engineered for optimized real-time responsiveness and high-speed inference. Using the specialized C2A (Combination to Application) dataset, which comprises 10,215 high-resolution images and over 25,000 meticulously annotated human instances across five critical disaster domains, this research investigates the model's effectiveness in detecting human targets.

The findings of this investigation demonstrate that YOLO26 achieves a higher mAP@0.5 of 0.89 and a peak F1-score of 0.87 over a 100-epoch training cycle, effectively moving the accuracy ceiling for disaster-response. By establishing this empirical baseline and validating the model's generalizability through a functional Gradio deployment interface, this research provides a vital foundation for a national AI-driven emergency response framework in Nigeria. The remainder of this paper is structured to detail the C2A benchmark methodology, analyze the quantitative loss and accuracy metrics, and discuss the socio-technical implications for Nigerian emergency management.

REVIEW OF RELATED LITERATURES

The evolution of real-time object detection has been defined by the iterative optimization of the You Only Look Once (YOLO) family of algorithms. Recent studies, such as that by [8], demonstrate that while earlier versions like YOLOv10 provide efficiency for edge-device deployment, the transition to YOLO26 marks a significant advancement in handling crowded environmental noise. This architectural progression is particularly relevant in Unmanned Aerial Vehicle (UAV) assisted Search and Rescue (SAR) operations. For instance, [6] achieved a mAP50 of 0.807 in aerial person detection, yet noted persistent sensitivity to adverse weather and RGB limitations. This gap is further highlighted by [9] whose autonomous rescue system reached a mAP50 of 0.995 in simulated floods but lacked validation in the unpredictable, non-simulated terrains characteristic of rural Nigeria.

The integration of multimodal sensing has been proposed to mitigate these visibility constraints. [7] utilized thermal infrared and Bidirectional Feature Pyramid Networks (BiFPN) to improve detection in smoke, though scalability for expansive forest surveillance remains a challenge. Similarly, in the context of ground-level surveillance, [10] attained high precision (96.1% mAP) using YOLO and DeepSORT, but observed significant performance degradation in cluttered zones. Security-focused frameworks have also emerged, such as the cryptographic approach by [11], which achieved 92.7% accuracy in anomaly detection but introduced high computational overhead. Despite these global advancements, a synthesis of current work (see Table 1) reveals a critical deficit in empirical data regarding model performance within the specific disaster zones, a gap this study addresses using the C2A dataset.

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 1: Summary of Related Works in Human Detection in a Crowded Environment

Author(s)	Algorithm / Framework	Key Research Findings	Identified Research Gaps & Limitations
[6]	YOLOv8s + PB-FPN	Achieved a mAP50 of 0.807; improved detection of small objects.	Performance remains sensitive to adverse weather and lighting.
[9]	UAV-Rescue + YOLOv12	mAP50 of 0.995 in simulated flood scenarios.	Based on simulated environments; lacks real-world terrain testing.
[7]	YOLO + BiFPN (Thermal)	Improved detection in smoke-occluded environments.	Limited scalability for expansive forest and high-altitude areas.
[8],	YOLOv10 – YOLOv12	YOLO26 showed superior recall in high-entropy scenes.	Sensitivity to heavy occlusion and high-density dynamics.
[10]	YOLO + DeepSORT	Attained 96.1% mAP in urban activity monitoring.	Performance degradation in cluttered and high-entropy zones.
[11]	YOLO + Cryptographic SHA	92.7% anomaly identification accuracy with data integrity.	High computational overhead; lacks modeling for dense crowds.

METHODOLOGY

This study adopts a structured and deployment-oriented methodology for human detection in crowded environments. The overall research workflow is illustrated in Figure 1, which presents the complete architecture of the proposed system. The pipeline begins with the acquisition of the C2A dataset, followed by a series of preprocessing operations designed to enhance data quality and generalization. The processed data is then fed into the YOLO26 model, which integrates a backbone for feature extraction, a neck for multi-scale feature fusion, and a native NMS-free detection head for end-to-end prediction [12]. Subsequently, the model undergoes training and evaluation to optimize detection performance. The final stage produces the system output, consisting of human detection, localization, and automated counting. This structured pipeline ensures a seamless transition from data preparation to real-time inference in complex disaster and security environments.

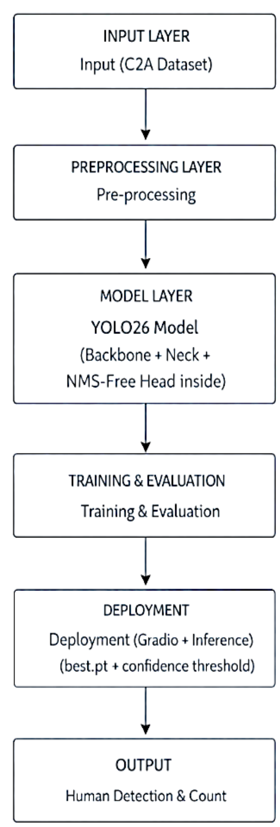

Figure 1: YOLO26-Based Human Detection Architecture

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Dataset Acquisition and Pre-Processing

The empirical foundation of this study is the C2A (Combination to Application) Dataset, a high-diversity repository specifically curated to address the limitations of standard benchmarks in disaster and security scenarios. The dataset comprises 10,215 high-resolution images featuring over 25,000 meticulously annotated human instances. Extended description of the dataset and its development is available in the following paper [13]. To ensure model robustness across the diverse Nigerian landscape, the data was stratified into four critical domains: building collapse, flood zones, forest vegetation, and fire/smoke environments.

To enhance generalizability, a series of stochastic augmentation protocols were applied, including mosaic augmentation and HSV color-space adjustments [14]. The dataset was systematically partitioned into training, validation, and testing sets using a 70/15/15 ratio. This distribution ensured that 70% of the data was dedicated to model convergence, while the remaining 30% was equally split between validation and testing to provide an unbiased evaluation of the model's performance across all four disaster domains [15]. These techniques simulate the high-entropy conditions of real-world disaster sites, forcing the model to learn invariant features of the human form rather than memorizing background patterns [16].

Model Architecture and Evaluation Strategy

This research evaluates the YOLO26 architecture, a deployment-focused evolution of the You Only Look Once framework optimized for real-time edge efficiency [12]. Moving beyond previous iterations, YOLO26 introduces a native End-to-End NMS-Free inference head, which eliminates the Non-Maximum Suppression (NMS) bottleneck and reduces total latency. The model incorporates Small-Target-Aware Label Assignment (STAL) to improve the detection of occluded subjects and distant targets in high-entropy crowds [17].

To maximize operational speed, the architecture utilizes a high-speed regression head characterized by the removal of Distribution Focal Loss (DFL). Training was stabilized using the MuSGD optimizer, a hybrid of Stochastic Gradient Descent enabling the smooth convergence of bounding box and boundary precision metrics [18]. Performance was quantified via Mean Average Precision (mAP@0.5) and F1-score over a 100-epoch training cycle on a Google Colab A100 GPU, ensuring a stable state of convergence suitable for high-stakes disaster and security monitoring [19]. Figure 2 below illustrates the internal structure of the YOLO26 model, highlighting the backbone for feature extraction, the neck for multi-scale feature fusion, and the NMS-free detection head for end-to-end human detection.

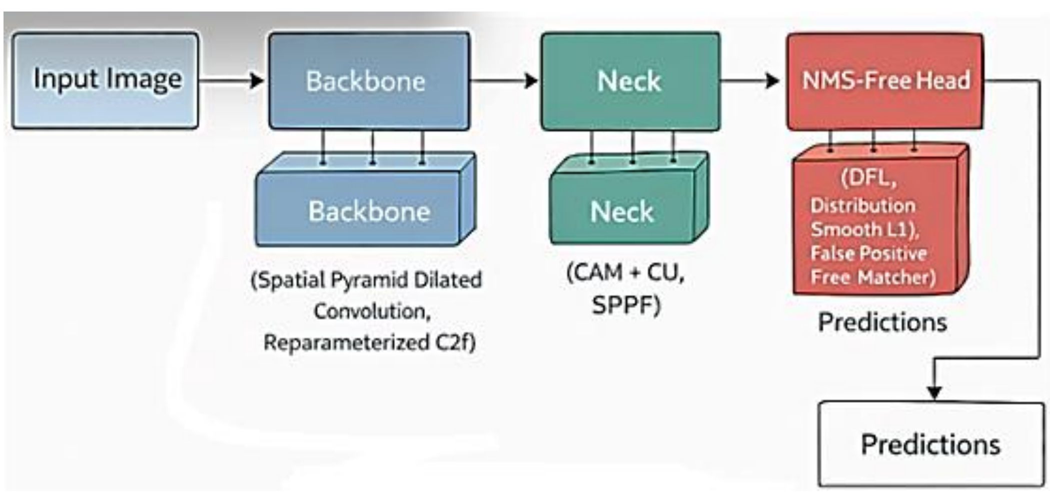

Figure 2: Internal Architecture of the YOLO26 Model

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Experimental Setup and Functional Deployment

The training utilized a structured learning rate schedule, initiating with a 5-epoch warmup phase before transitioning to linear decay to refine the model's weights [20], [21]. To ensure reproducibility and transparency, the YOLO26 model was trained using a standardized set of hyperparameters. The input images were resized to 1024 × 1024 pixels to preserve spatial details necessary for detecting partially occluded human targets in complex environments. Training was conducted using a batch size of 16 over 100 epochs [22], [23]. The MuSGD optimizer was employed with an initial learning rate of 0.01, combined with a warm-up phase to stabilize early gradient updates. The experiments were executed on a Google Colab environment equipped with an NVIDIA A100 GPU, enabling efficient training of the high-resolution dataset.

Following training, the optimal weights were validated against the test set, where the model achieved a peak mAP of 0.89 and an F1-score of 0.87. As the core contribution of this work, these weights were integrated into a functional real-time inference interface via the Gradio library [24]. To balance the requirements of high-entropy environments, the deployment was configured with a 0.5 confidence threshold; this threshold strategy was intentionally selected to maximize the system's recall in low-visibility scenarios like smoke-filled zones. This interface provides a dual-output mechanism, visual bounding box plotting and a real-time human headcount, transforming the static model into a practical, localized tool for Nigerian emergency responders.

RESULTS

This section presents the experimental results obtained from the training and evaluation of the YOLO26 model on the C2A dataset. The performance of the model is analyzed using key metrics such as training loss, mean Average Precision (mAP@0.5), precision, recall, and F1-score. In addition, qualitative results are provided to demonstrate the effectiveness of the model in detecting humans across various high-entropy disaster scenarios. The results are illustrated through Figures 3 to 11, highlighting both the learning dynamics and real-world detection capability of the proposed system.

THE SPATIAL ACCURACY OF THE MODEL

The Spatial accuracy of the model was monitored through the training bounding box and boundary precision metrics, as presented in Figure 3. The Train Box Loss (Localization) exhibited a consistent and smooth decay from an initial value of 1.62 to a stabilized minimum of approximately 1.01, indicating that the model progressively refined its ability to tightly bound human targets. Concurrently, the Train Distribution Focal Loss (DFL), which measures boundary confidence, remained exceptionally low (near zero) and stable throughout the 100-epoch cycle. This technical collaboration suggests that the YOLO26 architecture is highly effective at defining the precise physical boundaries of humans, minimizing spatial ambiguity even when detecting individuals in non-standard postures or against complex, high-entropy backgrounds.

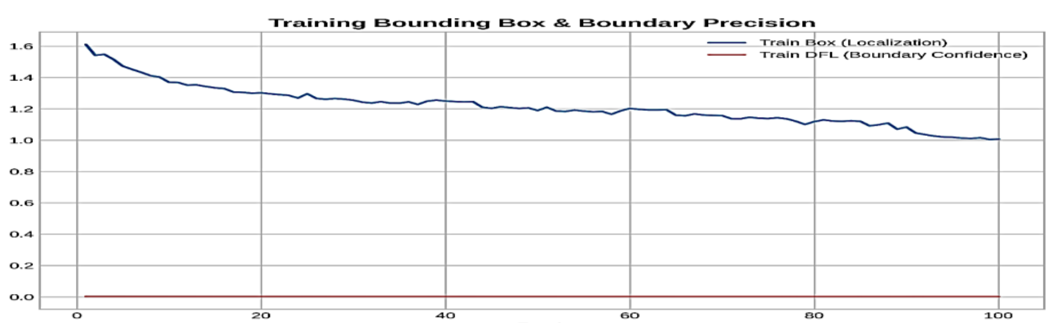

Figure 3: Training Bounding Box and Boundary Precision

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

The Training Stability

The training stability for the human detection task was facilitated by a structured learning rate schedule, as illustrated in Figure 4. The process utilized a brief warmup phase, where the learning rate was rapidly scaled from an initial value of 0.010 to a peak of approximately 0.029 within the first few epochs to stabilize initial gradient updates and prevent early divergence. Following the warmup, a linear decay strategy was implemented, gradually reducing the learning rate toward a final value of nearly 0.000 across the remaining cycle. This controlled reduction allowed the YOLO26 model to perform increasingly granular weight adjustments, ensuring that the final parameters were precisely tuned to capture the subtle visual features of human targets in cluttered disaster and security environments.

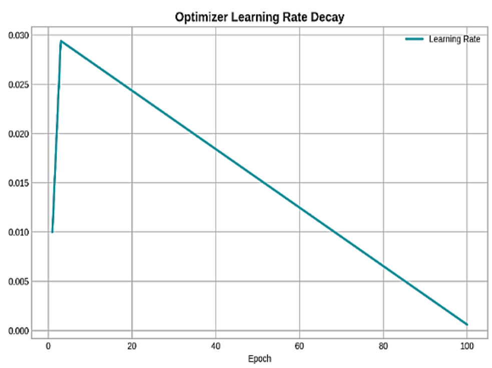

Figure 4: Optimizer Learning Rate Decay

Training Generalization and Overfitting Check

To ensure the human detection model possessed robust generalization capabilities, the training and validation box losses were monitored simultaneously, as shown in Figure 5. Both loss curves demonstrate a consistent downward trajectory without diverging, a key indicator that the YOLO26 model did not overfit the C2A training data. The validation box loss converged to a stable minimum of approximately 0.88 by the 100th epoch. This alignment between training and validation performance confirms that the model is capable of accurately localizing human targets in unseen imagery, a prerequisite for reliable operation in unpredictable disaster and security environments.

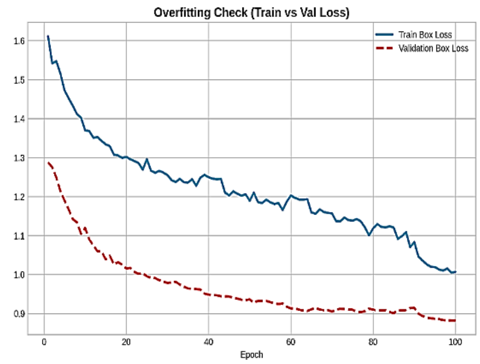

Figure 5: Overfitting Check (Train vs Val Loss)

Human Detection Accuracy (MAP@0.5)

The training progress was quantitatively assessed using the Mean Average Precision at a 0.5 Intersection over Union (mAP@0.5) metric, as shown in Figure 6. The human detection accuracy exhibited a rapid initial ascent, surpassing the 0.84 threshold within the first 15 epochs. After a period of stochastic refinement between epochs 40 and 80, the model achieved a peak mAP of approximately 0.89 by the conclusion of the 100-epoch cycle. This upward trajectory underscores the effectiveness of the YOLO26 architecture in converging toward a highly accurate representation of human features within the diverse and cluttered environments of the C2A dataset

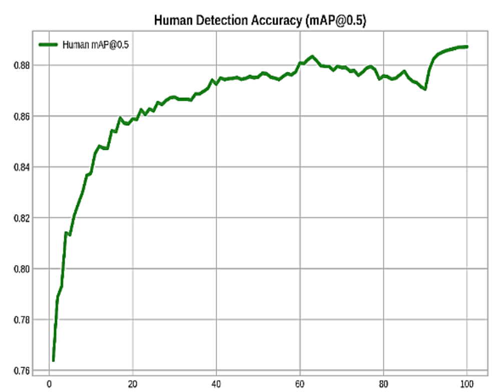

Figure 6: Human Detection Accuracy

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Human Detection Precision and Recall Convergence

The internal dynamics of the model's learning process were further examined through the simultaneous convergence of precision and recall, as illustrated in Figure 7. Human detection precision (accuracy) achieved a peak of approximately 0.89, while human detection recall (sensitivity) converged to approximately 0.84 by the 100th epoch. The steady, synchronized upward trend of both metrics indicates that the YOLO26 architecture successfully learned to identify a high percentage of human targets while simultaneously maintaining a low rate of false positives. This high sensitivity is particularly critical for the research objective, as it ensures that the model remains capable of detecting humans even when they are partially occluded or situated in visually noisy backgrounds typical of disaster environments.

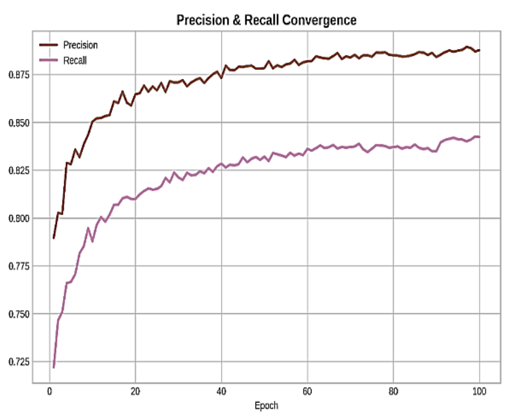

Figure 7: Precision and Recall Convergence

Human Detection Reliability Convergence (F1-Score)

The overall reliability of the system was monitored through the F1-Score convergence curve, as shown in Figure 8, throughout the training duration. Unlike the mAP, which measures general precision, the F1-score provides a combined metric of the model's ability to minimize both missed human targets and false alarms. The human detection reliability reached a stable plateau of approximately 0.865 early in the training process, specifically maintaining consistent performance from epoch 60 onward. This steady convergence indicates that the model's weights reached a state of optimal equilibrium, providing a dependable foundation for real-time deployment in high-stakes security and monitoring scenarios.

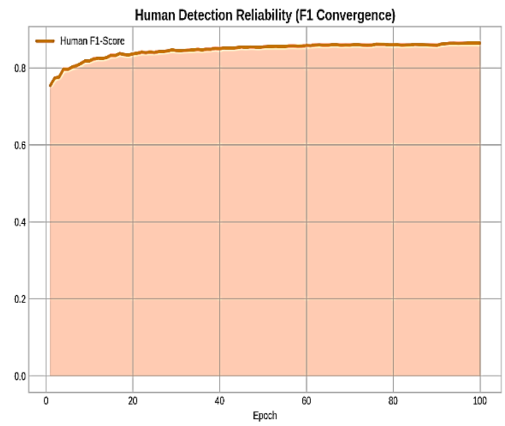

Figure 8: Human Detection Reliability (F1 Convergence)

Qualitative Human Detection Analysis

To visually validate the statistical findings, the YOLO26 model was tested on unseen imagery representing high-entropy disaster scenarios, as illustrated in Figure 9. In building collapse and rubble environments (Figures 9A and 9B), the model demonstrates an enhanced ability to identify human targets even when they are significantly occluded by structural debris or positioned in non-standard postures.

In the extreme conditions of active fire (Figure 9C) and heavy smoke (Figure 9D), the architecture maintains high detection sensitivity, successfully localizing humans amidst thermal noise and low-visibility atmospheric interference. These qualitative results confirm that the model's high recall effectively translates to reliable performance, ensuring that even partially visible individuals are accounted for in time-sensitive search and rescue operations.

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

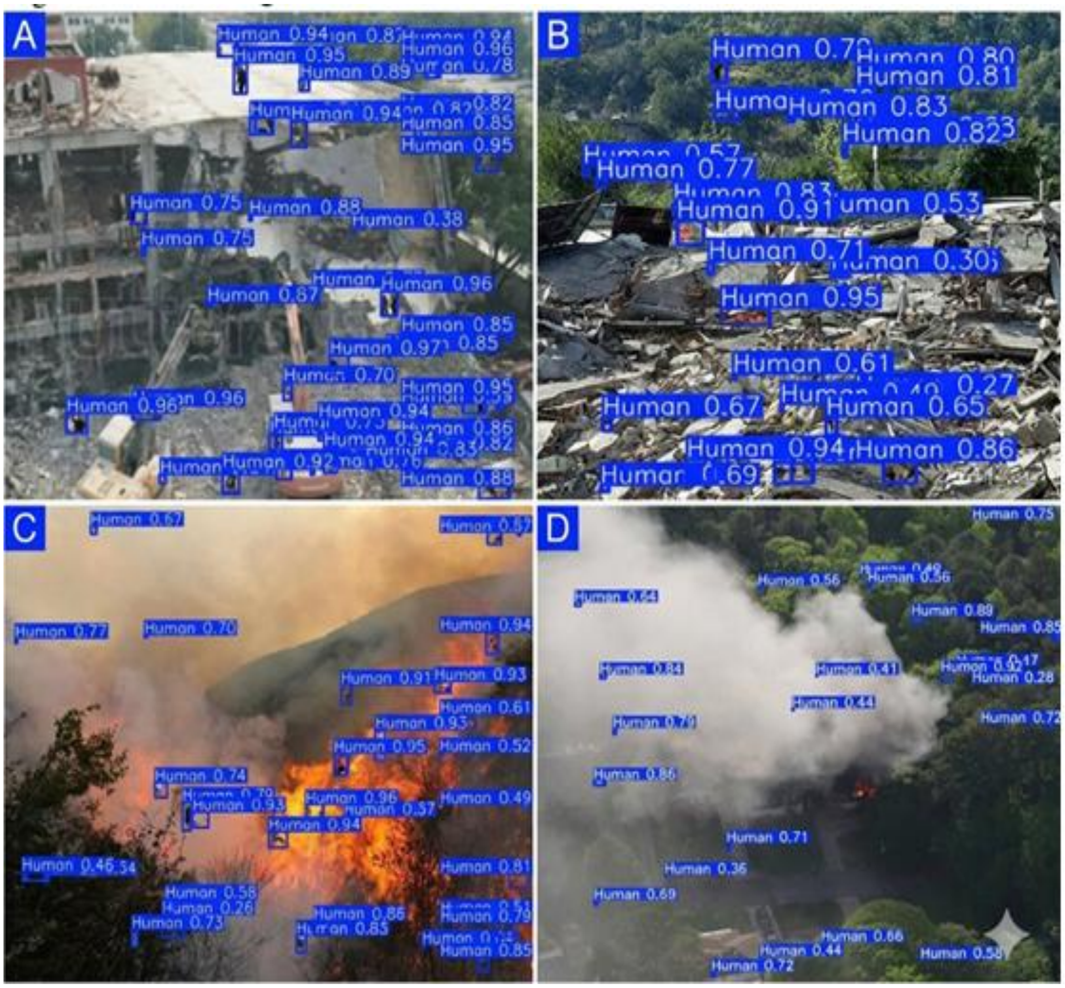

Figure 9: Qualitative Human Detection Analysis Across Critical Disaster Domains

Real-Time Deployment and Functional Prototyping

The transition from experimental training to a practical, operational tool was validated through a functional web-based deployment using the Gradio library, as shown in Figure 10. The deployment architecture utilizes the best.pt weight file, the optimized result of the 100-epoch training cycle as the primary inference engine. This allows the system to influence the specialized knowledge acquired during training to identify human features in complex environments.

As illustrated in the interface screenshots, the system provides a dual-output mechanism specifically designed for high-stakes human detection. Upon receiving an image input, the model performs real-time inference, generating localized bounding boxes with confidence scores alongside a definitive quantitative "Human Detected" count. The prototype demonstrated high efficacy across multiple disaster domains, including flooded urban streets, structural collapses, and forest fire scenarios, correctly identifying and counting humans even when they were situated far from the camera or partially obscured.

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

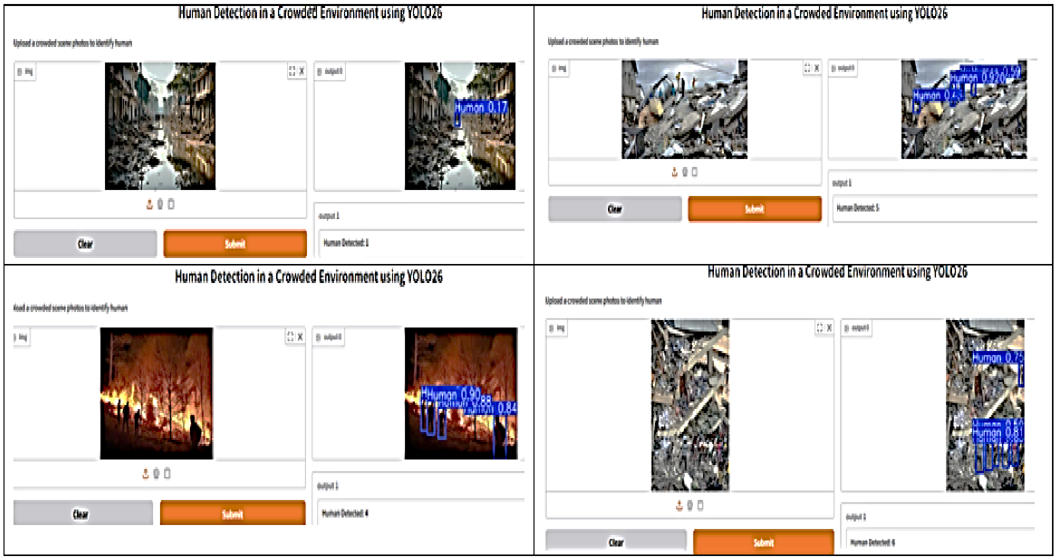

Figure 10: Deployment and Functional Prototyping

DISCUSSION OF FINDINGS

The findings of this study provide empirical evidence that the proposed YOLO26-based framework effectively addresses the challenge of accurately detecting human presence in complex, high-entropy disaster environments. As highlighted in the introduction, emergency response operations in Nigeria often occur in visually challenging conditions such as collapsed structures, flooded areas, dense vegetation, and fire-affected zones. These environments introduce significant visual noise and occlusion, reducing the reliability of traditional computer vision systems. The results demonstrate that the YOLO26 architecture maintains strong detection performance under such conditions, indicating its suitability for real-world disaster monitoring and search-and-rescue operations.

This study was motivated by key gaps identified in existing literature. Although deep learning–based object detection models have improved human detection performance, many struggle in cluttered or extreme environments. Previous studies have shown promising results using YOLO-based architectures; however, several were evaluated in controlled or simulated settings. For example, a YOLOv8-based aerial detection system achieved strong performance but remained sensitive to environmental variations [5], while high accuracy in flood-rescue simulations was limited by lack of real-world validation [6]. Additionally, while high precision has been reported using YOLO-DeepSORT frameworks, performance degradation persists in densely populated and visually complex environments [7]. These limitations highlight the need for more robust architectures. To address this, the YOLO26 model was trained and evaluated on the C2A dataset, which includes diverse disaster scenarios such as urban collapse, floods, forest environments, and fire conditions, enabling the model to learn invariant human features across varying contexts.

The experimental results demonstrate that the YOLO26 model achieved a mean Average Precision (mAP@0.5) of 0.89 and an F1-score of 0.87, indicating a strong balance between precision and recall. The relatively high recall is particularly important for disaster-response applications, where missed detections can have critical consequences. The training results further confirm the robustness of the framework, as both training and validation losses converged smoothly without divergence (Fig. 5), indicating effective generalization and minimal overfitting. The rapid improvement in detection accuracy during early

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

training stages (Fig. 6) suggests efficient feature learning, while the stable convergence of precision, recall, and F1-score (Figs. 7 and 8) demonstrates reliable model performance.

The strong performance can be attributed to key architectural improvements in YOLO26, particularly the NMS-free inference head, which reduces latency and simplifies the detection pipeline, enabling faster real-time predictions. Additionally, the Small-Target-Aware Label Assignment (STAL) mechanism enhances detection of small and partially occluded targets, which are common in disaster environments. Qualitative analysis further supports these findings, as shown in Fig. 9, where the model accurately detects human targets across challenging scenarios, including smoke-filled, collapsed, and flooded environments, consistently identifying partially obscured individuals.

Finally, the study extends beyond model evaluation by demonstrating practical deployment. The trained YOLO26 model was integrated into a real-time Gradio-based interface (Fig. 10), providing both visual detection and automated human counting. This transforms the system into a functional decision-support tool for emergency responders, improving situational awareness during rescue operations. Overall, the results confirm that YOLO26 provides a robust and deployment-ready solution for human detection in complex disaster environments, with strong potential to enhance emergency response efforts in regions such as Nigeria.

CONCLUSION

This research successfully developed and validated a human detection framework tailored for high-entropy disaster and security environments. By employing the YOLO26 architecture and the specialized C2A dataset, the model achieved a mean Average Precision (mAP@0.5) of 0.89 and a peak F1-score of 0.87. The results demonstrate that the system is capable of maintaining high sensitivity under extreme conditions, such as structural collapse and low-visibility scenarios. Furthermore, the successful deployment of a functional Gradio-based prototype proves the model's readiness for real-time application, providing a practical interface for emergency responders to obtain accurate human counts in critical situations.

FUTURE WORK

The next phase of this research will involve a systematic performance evaluation and comparative analysis between the YOLO26 architecture and the Detection Transformer (DETR) framework. The objective is to implement both algorithms on the unified C2A dataset to determine their respective technical trade-offs in terms of human detection accuracy, computational latency, and resource efficiency. Specifically, the study will investigate how the different underlying mechanisms—convolutional-based feature extraction in YOLO26 versus the self-attention mechanisms inherent in Transformers—affect model reliability across various catastrophic domains. By conducting this head-to-head comparison, we aim to identify the optimal architectural approach for real-time search and rescue operations. Additionally, we plan to expand the current dataset to include thermal and infrared imagery to assess the models' performance in night-time monitoring and low-visibility disaster scenarios.

REFERENCES

[1] M. G. Ragab *et al.*, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access*, vol. 12, no. February, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.
[2] A. Vijayakumar and S. Vairavasundaram, *YOLO-based Object Detection Models: A Review and its Applications*, vol. 83, no. 35. Springer US, 2024. doi: 10.1007/s11042-024-18872-y.
[3] M. Varatharaj and S. L. Devi, "International Journal of Research Publication and Reviews YOLO-Based Person Detection and Tracking in Dense Crowds," vol. 6, no. 1, pp. 999–1005, 2025.
[4] UNICEF, "Nigeria Flood Response Report," pp. 1–3, 2023.
[5] A. Oluwapelumi, "Overview of Displacement in Borno State, Nigeria," 2024.

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), MARCH, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

[6] F. Ciccone and A. Ceruti, "Real-Time Search and Rescue with Drones: A Deep Learning Approach for Small-Object Detection Based on YOLO," *Drones*, vol. 9, no. 8, 2025, doi: 10.3390/drones9080514.
[7] A. Basireddy, K. G. Rao, V. Vaitla, and S. R. Pedda, "Detection of Humans in Search and Rescue Operations Using Ensemble Learning and YOLOv9," *INCOFT 2025 - Int. Conf. Futur. Technol. locate*, vol. 2, no. Incoft, pp. 923–928, 2025, doi: 10.5220/0013607400004664.
[8] V. Hendriko and D. Hermanto, "Performance Comparison of YOLOv10, YOLOv11, and YOLOv12 Models on Human Detection Datasets," *Brill. Res. Artif. Intell.*, vol. 5, no. 1, pp. 440–450, 2025, doi: 10.47709/brilliance.v5i1.6447.
[9] T. Q. Thuan, V. T. N. Han, T. Quyen, D. P. Thinh, P. T. Dat, and D. A. Duy, "Development of a Drone-Based Rescue Platform with Intelligent Human Detection and Multi-Payload Delivery Mechanism," *FME Trans.*, vol. 53, no. 4, pp. 565–574, 2025, doi: 10.5937/fme2504565Q.
[10] A. B. Adege, "AI-Powered Human Activity Detection and Tracking in Dense Crowds Using YOLOv8-DeepSORT," *Inst. Eng. Technol.*, pp. 1–16, 2025, doi: 10.1049/ipr2.70227.
[11] P. Siva *et al.*, "Smart Surveillance Systems Using YOLOv8 : A Scalable Approach for Crowd and Threat Detection," *Int. J. Recent Adv. Eng. Technol.*, vol. 14, no. 1, 2025.
[12] S. Chakrabarty, "YOLO26 : A N A NALYSIS OF NMS-F REE E ND TO E ND," 2026, [Online]. Available: https://docs.ultralytics.com/models/yolo26/
[13] R. A. Nihal, B. Yen, K. Itoyama, and K. Nakadai, "UAV-Enhanced Combination to Application: Comprehensive Analysis and Benchmarking of a Human Detection Dataset for Disaster Scenarios," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* , vol. 15314 LNCS, pp. 145–162, 2025, doi: 10.1007/978-3-031-78341-8_10.
[14] B. Ç. Yavuz, "Scale-Dependent Performance Analysis of YOLO26 and YOLOv11 for PPE Detection," 2026.
[15] A. G. Tobias and J. Kittur, "Strategic innovations and future directions in deep learning for engineering applications : a systematic literature review".
[16] Z. Zhao *et al.*, "Enhancing Human Detection in Occlusion-Heavy Disaster Scenarios : A Visibility-Enhanced DINO (VE-DINO) Model with Reassembled Occlusion Dataset," 2025.
[17] V. Sharma, "Human Body Detection in Disaster Environments : A Review of Framework Design and Development Strategies," vol. 1, no. 9, pp. 1–6, 2026.
[18] R. Sapkota and M. Karkee, "Ultralytics YOLO Evolution: An Overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 Object Detectors for Computer Vision and Pattern Recognition," vol. 26, no. 2025, pp. 1–16, 2026.
[19] Z. Paszko and H. Padzik, "Estimation of high affinity estradiol binding sites in human breast cancer," *Arch. Geschwulstforsch.*, vol. 45, no. 5, pp. 430–443, 2024.
[20] D. S. Kalra, "Why Warmup the Learning Rate ? Underlying Mechanisms and Improvements," no. NeurIPS, 2024.
[21] Y. Liu, Y. Ge, R. Pan, A. Kang, and T. Zhang, "Theoretical Analysis on how Learning Rate Warmup Accelerates Convergence," 2025.
[22] K. Liu, Z. Fu, S. Jin, Z. Chen, F. Zhou, and R. Jiang, "ESOD : Efficient Small Object Detection on High-Resolution Images," pp. 1–13, 2024.
[23] E. Edozie, A. Nuhu, S. Ukagwu, K. John, and B. Olaniyi, "Comprehensive review of recent developments in visual object detection based on deep learning," 2025.
[24] P. S. Dwivedi, S. Bhosale, M. Shaikh, S. Doiphode, and S. Bansode, "Face-to-Face Language Translation," vol. 9, no. 3, pp. 256–259, 2024.

Corresponding author: Abdulrashid Sani
abdulrashid.sani@ssu.edu.ng
Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto State, Nigeria.