

Article

Localized Debris Detection in Post-Disaster Aerial Imagery Using YOLO-SDD

Hassan Al-Derham ¹, Mahitha Veeramachaneni ², Lu Gao ^{2,*} , Yunpeng Zhang ³ , Jingran Sun ⁴ ,
Ahmed Senouci ²  and Kevin Fu ⁵

¹ Department of Civil and Environmental Engineering, Qatar University, Doha 2713, Qatar; alderham@qu.edu.qa

² Department of Civil and Environmental Engineering, University of Houston, Houston, TX 77004, USA; mveerama@cougarnet.uh.edu (M.V.); asenouci@central.uh.edu (A.S.)

³ Department of Information Science Technology, University of Houston, Houston, TX 77004, USA; yzhan226@central.uh.edu

⁴ Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL 33620, USA; jingransun@usf.edu

⁵ Independent Researcher, Houston, TX 77401, USA; elfufu623@gmail.com

* Correspondence: lgao5@central.uh.edu; Tel.: +1-713-743-4436

Abstract

Post-disaster debris detection is important for rapid damage assessment, emergency response, and recovery planning. However, debris objects in aerial imagery are often fragmented, irregularly shaped, partially occluded, and visually confused with shadows, vegetation, roofs, vehicles, and damaged structures. This study proposes YOLO-SDD, a YOLO-based Shape-Guided Debris Detector built on YOLOv8 for localized debris identification in high-resolution post-disaster aerial imagery. YOLO-SDD combines a high-resolution P2 detection pathway with a shape-guided feature refinement module that uses box-supervised pseudo-mask and pseudo-boundary cues to refine P2-level features before final debris detection. A multi-event aerial imagery dataset was constructed from NOAA Emergency Response Imagery using images collected after hurricanes and a tornado in the United States. The model was evaluated using an image-level split, an event-level holdout test, component-level ablation studies, COCO-style scale-specific evaluation, and multi-seed stability analysis. On the image-level test set, YOLO-SDD achieved a precision of 0.959, recall of 0.933, mAP@50 of 0.970, and mAP@50:95 of 0.755, remaining competitive with larger YOLO-family models at lower computational complexity. In the event-level holdout test, YOLO-SDD achieved an AP@50 of 0.80 and an F1 score of 0.79, outperforming the YOLOv8s baseline and the selected large YOLO-family comparison model. The scale-specific evaluation showed improved AP@50 and recall for small and medium debris groups, while failure cases remained associated with shadows, vegetation, low contrast, and highly fragmented debris. The results indicate that shape-guided P2 refinement can improve localized debris screening under the tested conditions, although broader datasets, workflow integration, and human-in-the-loop validation are still needed before operational deployment.

Keywords: YOLO-SDD; object detection; debris detection; post-disaster aerial imagery; aerial photographs; natural disasters; consequences of natural disasters; emergency response



Academic Editor: Di Zhang

Received: 15 May 2026

Revised: 4 July 2026

Accepted: 8 July 2026

Published: 10 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Post-disaster damage assessment plays an important role in emergency response and infrastructure recovery because it supports situational awareness, resource allocation, and cleanup prioritization. Post-disaster damage assessment depends on timely imagery, reliable interpretation, and scalable analytical workflows. Although structural damage assessment has received substantial attention, localized debris detection also represents an important part of post-disaster remote sensing. Localized debris detection is important because debris can indicate damaged infrastructure, obstruct transportation access, create environmental and public health concerns, and affect the sequence of recovery operations [1–4].

Localized debris detection requires object-level information about where debris clusters appear in post-disaster imagery. A useful detection workflow should include not only visual identification, but also spatial localization, repeatable processing, robustness to background variation, and computational feasibility. In practice, debris appearance can become difficult to distinguish when objects are small, partially occluded, mixed with vegetation, affected by shadows, located near damaged buildings, or distributed across heterogeneous terrain. Therefore, detecting debris across post-disaster aerial imagery is a practical requirement for scalable disaster assessment.

Existing methods for post-disaster damage and debris detection can be organized into five main categories. Field inspection can provide direct evidence of local damage, but it requires substantial time and personnel [5]. Aerial image interpretation can cover large areas; however, manual interpretation remains sensitive to analyst subjectivity and image quality [6]. UAV-based inspection can provide fine spatial detail, although it may be limited by flight access, coverage, and deployment constraints [7]. Handcrafted image-processing and machine-learning approaches can extract texture, edge, and spectral patterns, but their performance may degrade under illumination, sensor, and regional variation [8,9]. Deep learning detectors can learn visual representations directly from imagery, yet their performance is often constrained by limited annotated debris data and small-object localization challenges [10,11].

The main limitation of these approaches is that debris detection in aerial imagery still lacks a model and dataset setting that jointly address event diversity, object-level localization, and small-object sensitivity. For emergency response and recovery planning, detection outputs should identify localized debris instances across different disaster events and geographic contexts. YOLO-based object detection can provide computationally efficient localization capability for offline or near-real-time image screening, but the standard feature pyramid may not preserve enough high-resolution spatial information for small debris objects. The remaining problem is therefore how to adapt a computationally efficient detector to localized debris detection in multi-event post-disaster aerial imagery without relying on unsupported assumptions about a single disaster type or region.

To address this gap, this study develops YOLO-SDD, a YOLO-based Shape-Guided Debris Detector, for localized debris detection using high-resolution aerial imagery collected after multiple disaster events. The framework combines a high-resolution P2 detection pathway with shape-guided P2 feature refinement using box-supervised pseudo-mask and pseudo-boundary cues. Unlike conventional P2-based YOLO variants that primarily improve scale representation, YOLO-SDD refines high-resolution features specifically for small, fragmented, and irregular debris objects in post-disaster aerial imagery. Evaluation on a multi-event dataset of hurricanes and a tornado in the United States emphasizes detection accuracy, localization quality, computational efficiency, and preliminary debris-screening value.

2. Literature Review

Post-disaster debris detection carries direct engineering significance because debris can obstruct roads, delay emergency access, influence cleanup priorities, and signal damaged infrastructure. Localized debris information therefore supports routing, staging, damage screening, and recovery sequencing for transportation agencies, emergency managers, and debris-removal planners after hurricanes, tornadoes, and other extreme events [1,6]. Recent work has begun to connect aerial image analysis with these recovery needs. Cheng et al. [12] combined AI with aerial photogrammetry to improve disaster-debris estimation, demonstrating that aerial imagery can enable more systematic debris quantification when linked to recovery planning. Amini et al. [13] fine-tuned foundational vision models for post-hurricane debris segmentation, showing that debris extraction benefits from modern vision architectures but also highlighting the need for task-specific post-disaster labels. Prakash et al. [14] employed UAV footage for automatic debris identification and damage assessment, further illustrating that debris detection is an applied engineering problem rather than a generic image-recognition task. Despite these advances, debris in aerial imagery is not a uniform object category; it can include roof fragments, construction materials, vegetation piles, household contents, vehicles, and mixed damaged structures, leading to substantial visual variation across events and locations.

Post-disaster image analysis has evolved from manual visual inspection and handcrafted features to deep learning models using UAV, satellite, and aerial imagery. Manual interpretation provides interpretable evidence but is time-consuming and difficult to scale across large areas [15,16]. Handcrafted methods rely on edge, texture, spectral, change-detection, or elevation cues, yet their performance is often sensitive to illumination, sensor differences, pre-event image availability, and site-specific calibration [8,17,18]. More recent learning-based approaches have advanced the field. Li et al. [19] applied a single-shot detector to building damage detection from post-event aerial imagery, demonstrating the value of object detection for disaster scenes but focusing on building damage rather than localized debris. Pi et al. [10] evaluated convolutional neural networks for aerial object detection in disaster response, showing that such methods can support operational workflows while also revealing that scene clutter remains a challenge. Jozi et al. [7] used UAV imagery for rapid assessment of residential buildings after disasters, highlighting the growing role of UAVs but concentrating on building-level evaluation. Al Shafian and Hu [20] reviewed the integration of machine learning and remote sensing for post-disaster building damage assessment, confirming the rapid growth of learning-based methods while noting that most studies still target buildings, roads, floods, or general damage categories rather than individual debris instances. Consequently, debris-specific object localization in high-resolution post-disaster aerial imagery remains less explored than broader building-damage or scene-level assessment.

Recent advances in small-object detection and YOLO-family models have introduced a range of strategies to improve detection in aerial and cluttered scenes, such as adding a high-resolution P2 detection head, extending Feature Pyramid Network (FPN)/Path Aggregation Network (PANet) feature fusion, integrating attention modules, refining IoU or localization losses, adopting lightweight backbones, and applying multi-scale training [21,22]. Recent aerial-image and UAV small-object detectors have similarly used high-resolution prediction heads, enhanced FPN/PANet feature fusion, attention modules, and improved localization losses to improve small-target detection in cluttered remote-sensing scenes [23–27]. Hussain [21] traced the evolution from YOLOv1 to YOLOv8 and highlighted how multi-scale feature fusion became central to modern one-stage detectors, confirming that high-resolution prediction and feature aggregation are established design directions rather than entirely new ideas. Hua and Chen [22] reviewed deep-learning-based

small-object detection in aerial images and summarized the continuing challenges caused by small target size, viewpoint variation, background complexity, scale diversity, and orientation changes. Recent YOLO-based aerial-image detectors also follow this direction. Yao et al. [23] proposed HP-YOLOv8 for remote-sensing small-object detection by strengthening spatial feature extraction and precision-oriented detection. Yuan et al. [24] proposed IA-YOLOv8 for UAV remote-sensing images by combining intra-group multi-scale fusion attention with adaptive weighted feature fusion. Li et al. [25] proposed SOD-YOLO for UAV images by integrating receptive-field attention and a balanced spatial-semantic feature pyramid. Zhang et al. [26] proposed PARE-YOLO for aerial small-target detection by restructuring multi-scale feature extraction and fusion and adding a lightweight detection head. Qu et al. [27] proposed SMA-YOLO for UAV small-object detection by combining parameter-free attention, multi-cross-scale feature fusion, and an additional P2-oriented detection design. These studies confirm the value of preserving fine spatial features, but most of them remain general small-object detectors and do not explicitly model the irregular shape and boundary ambiguity of post-disaster debris. Related work has also explored debris-like targets in other imaging contexts. Qiao et al. [11] developed a deep learning method for floating-debris detection using ground-level images, confirming the applicability of detection models to debris but in a setting distinct from high-resolution post-disaster aerial imagery. Han and Ye [28] improved YOLOv5 for water-surface garbage detection, demonstrating that YOLO variants can be adapted to debris-like objects, although the modifications target a different scene type. Bansal and Tripathi [29] reviewed automated floating-debris monitoring with optical satellite imagery and AI, noting progress while underscoring persistent challenges of scale, domain shift, and visual ambiguity. Thus, while a P2 head and a high-resolution neck are widely used small-object strategies in YOLO-style models, most of these methods primarily enhance scale representation and feature aggregation. The debris-specific difficulties, including irregular shapes, fragmented boundaries, and confusion with shadows, vegetation, roofs, and damaged structures, remain insufficiently addressed.

The remaining gap is how to make high-resolution detection features more responsive to the irregular shapes and ambiguous boundaries of post-disaster debris, rather than simply adding another small-object detection head. Prior post-disaster debris studies have demonstrated the value of aerial and satellite imagery for debris-zone extraction and recovery assessment, but they have generally focused on zone-level mapping or broader damage assessment rather than instance-level bounding-box detection with a modern one-stage detector. Existing small-object YOLO variants improve scale representation through P2 heads, enhanced feature fusion, attention modules, or lightweight detection heads, but they do not explicitly use debris-shape cues to refine high-resolution features before detection. YOLO-SDD addresses this gap by combining a high-resolution P2 pathway with shape-guided P2 feature refinement using box-supervised pseudo-mask and pseudo-boundary cues. The final inference output remains standard debris bounding boxes, allowing the model to remain compatible with object-detection evaluation and preliminary debris-screening workflows.

3. Methodology

3.1. Overview

The objective of our methodology is to develop an object detection framework for localized debris identification in high-resolution post-disaster aerial imagery. The proposed framework is designed for a setting in which debris objects may occupy only a small number of pixels, appear in cluttered urban or rural backgrounds, and vary across disaster events.

The methodology consists of four main parts: YOLOv8 baseline representation, high-resolution feature fusion, P2 detection, and shape-guided P2 feature refinement. The YOLOv8 baseline representation defines the backbone, neck, and standard prediction heads used as the starting point. The high-resolution feature fusion part extends the neck so that shallow spatial information can be combined with deeper semantic features. The P2 detection pathway uses the added high-resolution feature map to generate object-level debris predictions, while the shape-guided refinement module uses pseudo-mask and pseudo-boundary cues to refine P2 features before final detection. Figure 1 illustrates the overall structure of the proposed methodology and the relationship among the baseline detector, the enhanced fusion path, the shape-guided refinement module, and the additional detection head.

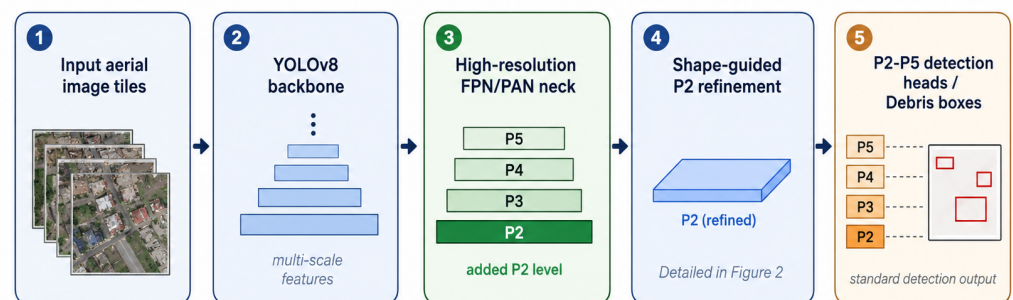


Figure 1. Overview of the YOLO-SDD architecture with shape-guided P2 feature refinement. The shape-guided P2 refinement module shown in the fourth box is detailed in Figure 2.

The methodology produces bounding boxes and confidence scores for debris objects in aerial image tiles. These outputs are evaluated in the case study using held-out images from multiple disaster events and are compared with alternative detection models.

3.2. YOLOv8 Baseline Representation

YOLOv8 is used as the baseline representation because it provides a single-stage, end-to-end detection structure that maps an input image directly to bounding box and class predictions. The architecture is organized into a backbone for hierarchical feature extraction, a neck for multi-scale feature fusion, and detection heads for prediction at multiple spatial resolutions.

The backbone progressively downsamples the input image while increasing channel depth. This structure allows the network to learn low-level spatial features in shallow layers and more abstract semantic features in deeper layers. Repeated C2f blocks preserve gradient flow while maintaining computational efficiency, and the SPPF module aggregates multi-scale context near the deepest stage of the backbone.

The standard neck combines high-level semantic features with lower-level spatial features through feature pyramid and path aggregation operations. In this structure, upsampling, concatenation, convolution, and C2f refinement are used to propagate information across scales. The standard YOLOv8 detection heads then produce predictions at P3, P4, and P5 feature levels, which are suitable for small, medium, and large objects but may still be limited for very small debris instances in aerial imagery.

3.3. High-Resolution Feature Fusion

The high-resolution feature fusion part extends the YOLOv8 neck while preserving the original backbone. In YOLO-SDD, an additional feature-fusion stage is introduced by adding one more upsampling, concatenation, and C2f refinement block in the Feature Pyramid Network (FPN) path, followed by an additional convolution-based downsampling, concatenation, and C2f refinement block in the Path Aggregation Network (PANet) path.

These are standard feature-fusion operations commonly used in YOLO-style detectors. They are used here to connect shallow spatial features with deeper semantic features so that small and fragmented debris objects can be represented at a higher spatial resolution. Because these operations are standard deep-learning components, the following section focuses on the P2 feature construction and the proposed shape-guided refinement module.

3.4. P2 Detection Head and Training Objective

The P2 detection head uses the high-resolution feature map created by the neck extension to improve the localization of small debris objects. The added head complements the original P3, P4, and P5 heads instead of replacing them, so the detector can retain multi-scale prediction while adding a denser spatial grid for small objects.

For an input size of 640×640 , the added P2 feature map operates at stride 4 and has a spatial size of 160×160 . This creates 25,600 candidate prediction locations, which is four times the number of locations in the standard P3 head. The P2 feature map is constructed as

$$F_2 = \text{C2f}(\text{Concat}(\text{Upsample}(F_3), F_{\text{shallow}})) \quad (1)$$

Here, F_3 denotes the feature map from the original small-object detection level, and F_{shallow} denotes the C2-level backbone feature map with stride 4. This shallow feature map is selected because it preserves fine spatial details that are important for detecting small, fragmented, and irregular debris objects in aerial imagery. The resulting F_2 is used by the added detection head to predict debris bounding boxes and confidence scores.

The final YOLO-SDD model extends the detection architecture with a shape-guided refinement module that operates on the high-resolution P2 feature map. Two lightweight auxiliary branches predict a debris pseudo-mask and a debris boundary map from the P2 features; these predictions are concatenated and transformed into a shape attention map, which refines the P2 representation through a residual connection before the P2 detection head. Figure 2 details this refinement pathway and separates the training-only auxiliary supervision from the box-only inference path. Importantly, the mask and boundary branches are training-only auxiliary signals and are not part of the inference outputs; the model still produces only bounding boxes and confidence scores. Pseudo masks were generated from the central region of each ground-truth debris bounding box within each already split subset, and the corresponding pseudo boundaries were obtained from the masks using a morphological gradient.

The pseudo-mask target is restricted to the central 70% of each bounding box in width and height, reducing supervision noise near uncertain boundaries. The pseudo-boundary target is obtained by applying a morphological gradient, implemented as dilation minus erosion, with a 3×3 kernel to the pseudo-mask. The mask and boundary branches each consist of a 3×3 Conv-BN-SiLU block followed by a 1×1 convolution and sigmoid activation, producing a single-channel output. The shape-attention block concatenates the mask and boundary predictions as a two-channel input, then applies a 3×3 convolution with eight output channels, batch normalization, SiLU activation, a 1×1 convolution, and sigmoid activation. The refinement branch applies a 3×3 Conv-BN-SiLU block to the attention-weighted P2 feature map, followed by a 1×1 projection back to the original P2 channel dimension. The mask loss is the sum of binary cross-entropy and Dice loss, and the boundary loss is binary cross-entropy.

$$M_2 = \sigma(f_m(F_2)) \quad E_2 = \sigma(f_e(F_2)) \quad (2)$$

$$A_{\text{shape}} = \sigma(g(\text{Concat}(M_2, E_2))) \quad (3)$$

$$F_2^r = F_2 + \gamma h(F_2 \odot A_{\text{shape}}) \quad (4)$$

In Equations (2)–(4), $f_m(\cdot)$ denotes the pseudo-mask prediction branch, and $f_e(\cdot)$ denotes the pseudo-boundary prediction branch. These two branches are lightweight auxiliary prediction branches. Each branch consists of a 3×3 Conv-BN-SiLU block followed by a 1×1 convolution and a sigmoid activation, producing a single-channel pseudo-mask map or pseudo-boundary map. The function $g(\cdot)$ denotes the shape-attention subnetwork. It takes the concatenated pseudo-mask and pseudo-boundary maps as input and applies a 3×3 convolution with eight output channels, batch normalization, SiLU activation, a 1×1 convolution, and sigmoid activation to generate the shape-attention map A_{shape} . The function $h(\cdot)$ denotes the residual refinement subnetwork. It applies a 3×3 Conv-BN-SiLU block to the attention-weighted P2 feature map, followed by a 1×1 projection back to the original P2 channel dimension.

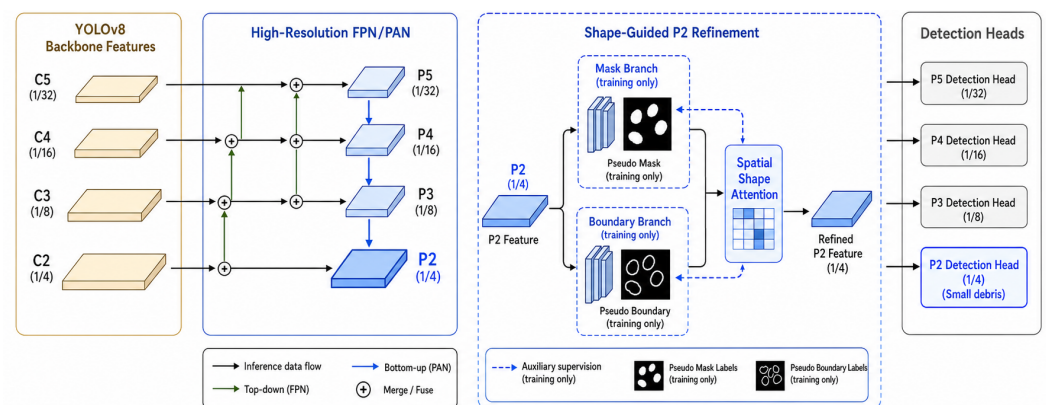


Figure 2. Local YOLO-SDD structure showing the high-resolution P2 neck and the shape-guided refinement pathway.

The detection loss (L_{det}) follows the standard Ultralytics YOLOv8 detection objective. The P2, P3, P4, and P5 detection heads are optimized jointly through the same Detect layer using the standard box loss, classification loss, and distribution focal loss. No separate manual loss weights are assigned to individual detection heads. The global detection loss gains are kept consistent with the YOLOv8 setting. To strengthen P2-level feature learning for small and irregular debris objects, the proposed shape-guided refinement module is supervised by auxiliary mask and boundary losses.

The total training loss is therefore defined as

$$L_{\text{total}} = L_{\text{det}} + \lambda_{\text{mask}} L_{\text{mask}} + \lambda_{\text{boundary}} L_{\text{boundary}} \quad (5)$$

In the final M6 configuration, the auxiliary weights were set to $\lambda_{\text{mask}} = 0.001$ and $\lambda_{\text{boundary}} = 0.0005$, and the residual scaling parameter γ was initialized as 0.0002. The value of $\lambda_{\text{boundary}} = 0.0005$ was set as a small auxiliary-loss weight to provide boundary guidance without dominating the standard YOLO detection loss, while $\gamma = 0.0002$ was initialized as a small residual scaling parameter to keep the shape-guided refinement stable during training. For negative images with no debris, all-zero pseudo masks and boundaries are supplied, and their auxiliary loss contribution is down-weighted by a factor of 0.1. Training used a batch size of 8 with gradient accumulation over two steps, yielding an effective batch size of 16.

The next section applies this methodology to the multi-event aerial imagery dataset and evaluates whether the added high-resolution prediction level and shape-guided refinement improve debris detection performance.

4. Case Study

The case study is designed to evaluate YOLO-SDD under a mixed-event image-level split using imagery collected from multiple disaster events. This setting evaluates held-out images from the multi-event dataset, rather than generalization to a completely unseen disaster event. A multi-event case is appropriate because debris in aerial imagery may appear along roads, near damaged structures, in open fields, and within vegetation-covered areas.

The case is also suitable for demonstrating the methodology because the dataset contains both debris and non-debris images, which allows the model to be evaluated with respect to false detections, missed detections, localization accuracy, and computational efficiency. The case study therefore connects the high-resolution feature design in the methodology to a practical post-disaster image-screening task.

4.1. Case Study Data

High-resolution aerial imagery was collected from post-disaster regions across the United States. The dataset was sourced from the NOAA Emergency Response Imagery portal [30]. The dataset includes imagery from eight disaster-event groups: Hurricane Harvey in Texas in 2017, Hurricane Michael in Florida in October 2018, the Nashville Tornado in Tennessee in March 2020, Hurricane Laura in Louisiana in August 2020, Hurricane Sally in September 2020, Hurricane Delta in Louisiana in October 2020, Hurricane Ida in Louisiana in August and September 2021, and Hurricane Ian in Florida in September 2022. After fixed-grid tiling and positive/negative sample selection, the final object-detection dataset included 2940 image tiles drawn from these eight disaster-event groups. These counts refer to model-input image tiles rather than original NOAA source images. Figure 3 shows the geographic distribution of the case study imagery.

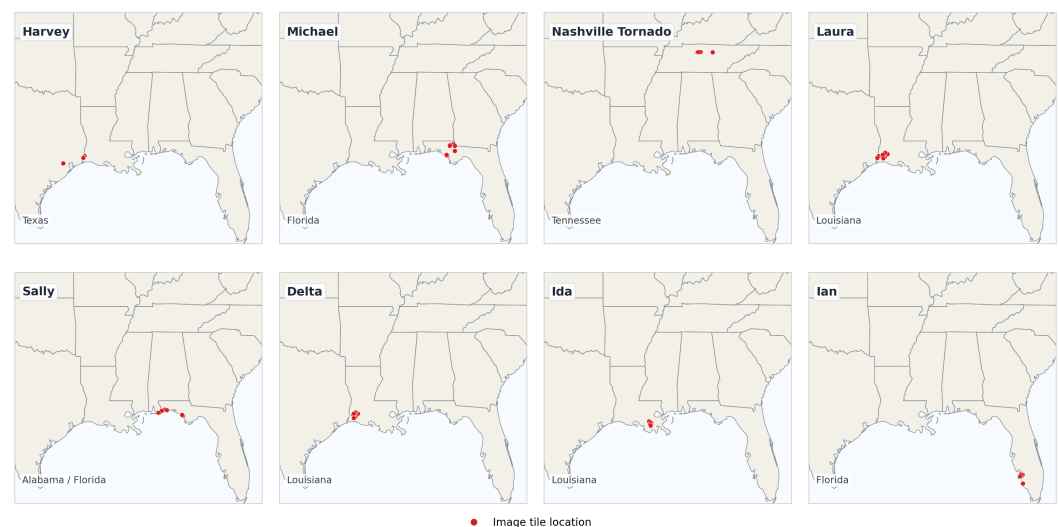


Figure 3. Geographic distribution of image-tile locations across the eight disaster-event groups.

Each original aerial image covered a relatively large geographic area; therefore, a fixed grid-based tiling procedure was used to prepare the imagery for object detection while retaining localized debris objects at the tile scale. Visible debris objects were manually annotated with bounding boxes using Labellmg. The annotations were stored in Pascal VOC XML format and were then converted to YOLO format using normalized bounding box coordinates.

The selected image tiles were drawn from the eight disaster-event groups described above. The mixed-event image-level experiment and the Sally-period event-wise experiment were constructed as two separate evaluation settings. In the mixed-event experiment,

image tiles from all eight event groups were split at the image-tile level. In the event-wise experiment, all Sally-period image tiles were reserved for testing and were excluded from the corresponding training and validation subsets. The Sally-period hold-out set contained 124 image tiles and was used only as a preliminary cross-event generalization check.

4.2. Implementation Setup

The preprocessing and training setup was designed to provide a consistent input dataset for YOLO-SDD and the comparison models.

All image tiles were resized to 640×640 pixels to match the model input requirement. For the main mixed-event image-level experiment, the selected image tiles from the eight disaster-event groups were divided into training, validation, and test subsets using a 70:20:10 ratio. As shown in Table 1, the table reports the number of debris-positive tiles, debris boxes, negative tiles, and total image tiles used in each subset. Negative tiles were selected from the same event groups within each subset. The counts in Table 1 refer to selected and manually annotated model-input image tiles used in the revised experiments and do not include offline augmented copies. Separately, the Sally-period event-wise hold-out experiment used a different split: all Sally-period image tiles were excluded from the corresponding training and validation subsets and were used only for event-level testing.

Table 1. Dataset split used for debris detection experiments.

Split	Positive Tiles	Debris Boxes	Negative Tiles	Total Tiles
Training	1176	2703	1176	2352
Validation	147	343	147	294
Test	147	330	147	294
Total	1470	3376	1470	2940

To ensure a consistent baseline comparison, all compared models were trained and evaluated using the same experimental protocol whenever the implementation supported direct configuration. The same train/validation/test split was used for all models. All input image tiles were resized to 640×640 pixels. The same held-out test set was used for final evaluation.

The YOLO-family models, including YOLO-SDD, YOLOv8s, YOLOv8l, YOLOv9, YOLOv11n, YOLOv11l, and YOLOv12l, were trained and evaluated using Ultralytics 8.4.66 under the same data split, input size, training schedule, optimizer, learning-rate settings, and evaluation protocol. YOLO-SDD was implemented by modifying the YOLOv8s configuration to add the high-resolution P2 pathway and the shape-guided P2 refinement module. No task-specific architectural changes were made to the baseline YOLO models. These models were initialized with publicly available COCO pre-trained weights when available and were fine-tuned on the debris dataset. The main training configuration used 150 epochs, a batch size of 8, gradient accumulation of 2, an effective batch size of 16, the SGD optimizer, an initial learning rate of 0.0001, a final learning-rate factor of 0.01, deterministic training, and a fixed random seed of 42. The held-out test set was used only for final evaluation and not for model selection.

The non-YOLO comparison models, including Faster R-CNN, CenterNet, Deformable DETR, and CNXResNet, were evaluated using the same dataset split, input resolution, held-out test set, and evaluation metrics. Because these architectures rely on implementation-specific training recipes and are not direct YOLO configuration variants, their results are reported as additional reference comparisons rather than one-to-one architectural ablations. The exact configuration files and training/evaluation scripts used in this study are available from the corresponding author upon reasonable request.

All models were evaluated using the same object-detection metrics: precision, recall, mAP@50, and mAP@50:95. AP and mAP values were computed from the precision–recall curves after non-maximum suppression. Fixed confidence thresholds were used only for threshold-dependent metrics such as precision, recall, and F1. Inference latency was measured under the same hardware and software environment: Ultralytics 8.4.66, Python 3.12.13, PyTorch 2.11.0+cu128, CUDA 12.8, NVIDIA Tesla T4 GPU with 14,913 MiB memory, input size 640×640 , and batch size 8. TensorRT acceleration and model compilation were not used. The reported inference time represents per-image model inference time; preprocessing, postprocessing, and non-maximum suppression were logged separately.

Following standard object detection evaluation practice [21], four standard object detection metrics were used to evaluate performance: precision, recall, mAP@0.5, and mAP@0.5:0.95. Precision measures the proportion of predicted debris instances that are correct:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

where TP denotes true positives and FP denotes false positives. Recall measures the proportion of actual debris instances that are detected:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

where FN denotes missed debris objects. The mAP@0.5 metric evaluates detections at an IoU threshold of 0.5:

$$\text{IoU} \geq 0.5 \quad (8)$$

whereas mAP@0.5:0.95 averages performance over stricter IoU thresholds:

$$\text{IoU} \in \{0.50, 0.55, 0.60, \dots, 0.95\} \quad (9)$$

Together, these metrics evaluate false positives, missed detections, standard detection accuracy, and localization quality.

4.3. Results

The training curves indicate that the evaluated detection models converged in a stable manner over 150 epochs. As shown in Figure 4, precision, recall, mAP@50, and mAP@50:95 improved rapidly during the early epochs, especially within the first 20–40 epochs. The curves then became more stable after approximately 100–120 epochs, which suggests that the models learned the main debris-related visual patterns without strong oscillation during training.

The metric curves also show that coarse detection accuracy stabilizes earlier than strict localization accuracy. In Figure 4, mAP@50 rises faster and reaches a plateau earlier than mAP@50:95. This pattern indicates that the models can learn to identify debris regions relatively quickly, whereas precise bounding-box alignment remains more difficult for small and irregular debris objects.

The held-out test results indicate that YOLO-SDD remains one of the strongest overall detectors while maintaining moderate computational cost. As shown in Table 2, YOLO-SDD achieves a precision of 0.959, recall of 0.933, mAP@50 of 0.970, and mAP@50:95 of 0.755. YOLO-SDD obtains the highest precision, recall, and mAP@50 among the compared models, while YOLOv12l gives a marginally higher mAP@50:95. These values indicate that the proposed model improves standard detection accuracy and false-detection control while maintaining localization quality comparable to the strongest large YOLO-family model.

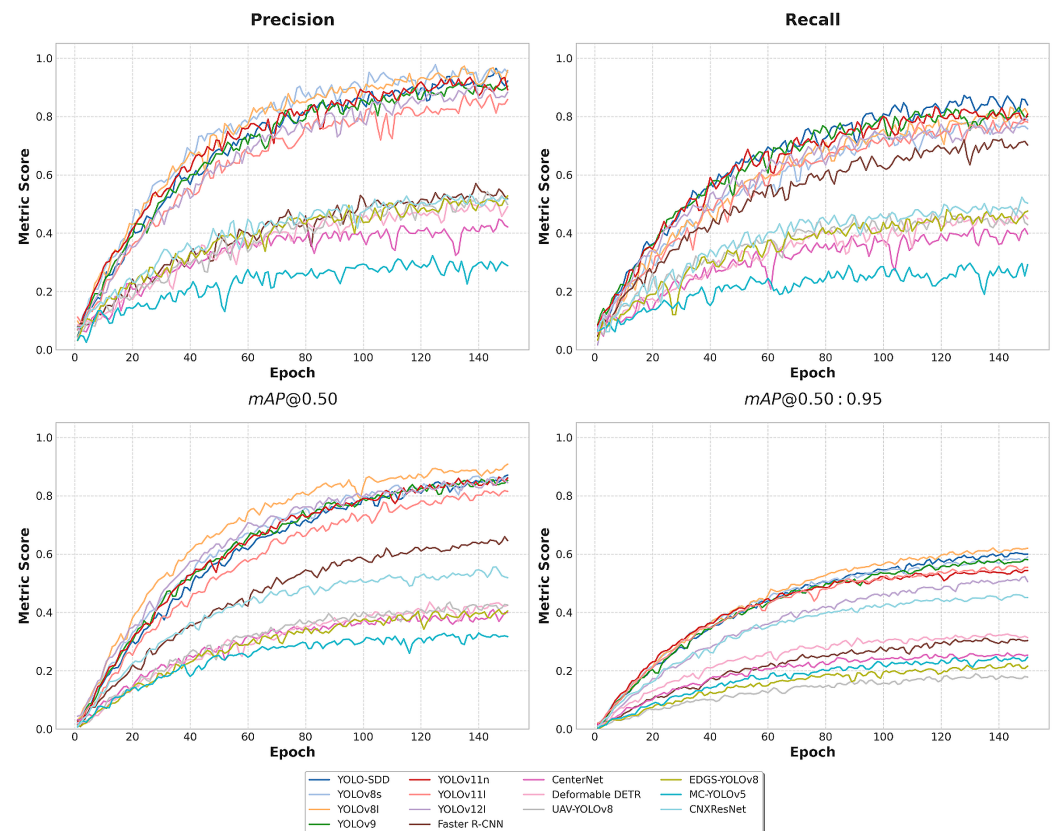


Figure 4. Validation metrics for the evaluated detection models over 150 training epochs.

Table 2. Performance and computational efficiency of debris detection models on the held-out test set using the common dataset and evaluation protocol, with inference time measured under the hardware and software settings described in Section 4.2.

Model	Precision	Recall	mAP@50	mAP@50:95	GFLOPs	Time (ms)
YOLO-SDD	0.959	0.933	0.970	0.755	36.9	24.0
YOLOv8s	0.925	0.895	0.941	0.670	28.6	17.6
YOLOv8l	0.923	0.870	0.933	0.683	165.2	50.5
YOLOv9	0.891	0.854	0.929	0.648	27.4	37.0
YOLOv11n	0.846	0.736	0.844	0.500	6.4	4.3
YOLOv11l	0.921	0.889	0.939	0.698	87.3	35.8
YOLOv12l	0.933	0.921	0.962	0.756	89.4	45.1
Faster R-CNN	0.570	0.740	0.692	0.322	180.0	95.0
CenterNet	0.427	0.412	0.397	0.259	55.0	28.0
Deformable DETR	0.512	0.483	0.451	0.326	173.0	42.0
UAV-YOLOv8	0.300	0.276	0.226	0.075	32.1	12.5
EDGS-YOLOv8	0.331	0.305	0.246	0.086	34.5	9.6
MC-YOLOv5	0.881	0.809	0.903	0.568	16.8	8.8
CNXResNet	0.528	0.494	0.546	0.467	62.0	18.5

The ablation study evaluates the contribution of each architectural component. The ablation variants were trained with a batch size of 8 and gradient accumulation of 2, resulting in an effective batch size of 16. As shown in Table 3, the YOLOv8s baseline (M1) achieved an mAP@50 of 0.9410. The partial configurations, including P2 alone (M2), the additional FPN/PAN path without P2 (M3), the combination of P2 and the extra FPN/PAN path (M4), and P2 + shape refinement (M5), yielded mAP@50 values of 0.9670, 0.9650, 0.9690, and 0.9675, respectively. In contrast, the full YOLO-SDD model (M6) achieved a precision of

0.9592, recall of 0.9333, mAP@50 of 0.9705, and mAP@50:95 of 0.7551. These results suggest that the full configuration provides the strongest overall balance among the tested ablation variants, although the improvement over the P2 + FPN/PAN configuration is incremental.

Table 3. Ablation results.

Model	Precision	Recall	mAP@50	mAP@50:95	GFLOPs	Time (ms)
M1: YOLOv8s baseline	0.9250	0.8950	0.9410	0.6700	28.6	17.6
M2: YOLOv8 + P2 only	0.9520	0.9320	0.9670	0.7420	32.8	20.0
M3: YOLOv8 + FPN/PAN only, no P2	0.9540	0.9270	0.9650	0.7380	31.9	19.3
M4: YOLO-SDD with P2 and FPN/PAN	0.9570	0.9340	0.9690	0.7480	35.6	22.8
M5: P2 + shape refinement	0.9580	0.9280	0.9675	0.7460	33.8	20.8
M6: Full YOLO-SDD	0.9592	0.9333	0.9705	0.7551	36.9	24.0

The qualitative results further indicate that YOLO-SDD can localize debris in held-out images sampled from multiple disaster events. Figure 5 shows original aerial images and the corresponding YOLO-SDD detections for Hurricane Ian, Hurricane Ida, Hurricane Michael, Hurricane Delta, Hurricane Harvey, the Nashville Tornado, and Hurricane Laura. The examples include residential neighborhoods, open-field scenes, roadside debris, vegetation-covered areas, and severe structural collapse zones, which demonstrates the range of visual conditions considered in the test set.

In residential and semi-urban scenes, YOLO-SDD identifies debris close to houses, driveways, and roadways while avoiding many visually similar background objects such as rooftops and vehicles. In open or low-contrast scenes, the model localizes debris over grassland and bare soil where visual contrast is limited. In severe structural collapse scenes, the detector captures multiple debris clusters with different scales and spatial distributions. These qualitative results support the quantitative finding that the added high-resolution detection level is useful for small and fragmented debris objects, although very small debris in heavily shadowed or chaotic scenes may still be difficult to localize precisely.

4.4. Event-Level Evaluation

The event-level evaluation used a separate Sally-period hold-out split as a preliminary cross-event generalization check. All Sally-period image tiles were excluded from training and validation and were used only for testing in this event-wise experiment. The model was trained and validated using the remaining seven disaster-event groups and tested on the Sally-period event group, which included 124 image tiles with ground-truth labels where applicable. Table 4 compares the YOLOv8s baseline, YOLOv12l, and YOLO-SDD under this setting. YOLO-SDD achieved the highest AP@50 and F1 score among the three compared models, providing preliminary evidence of better performance on this Sally-period hold-out test under the tested conditions.

Table 4. Event-level test performance on the held-out Sally-period event group.

Model	Precision	Recall	AP@50	F1
YOLOv8s baseline	0.68	0.72	0.68	0.70
YOLOv12l	0.72	0.75	0.73	0.73
YOLO-SDD	0.78	0.80	0.80	0.79



Figure 5. Representative original image tiles and YOLO-SDD detection results from multiple disaster-event groups, including hurricane and tornado scenes with residential, roadside, open-field, vegetation-covered, and structural-collapse conditions.

4.5. Comparison and Additional Analysis

The computational comparison indicates that YOLO-SDD provides a balanced accuracy–efficiency trade-off under the tested setting. Figure 6 shows that YOLO-SDD requires 36.9 GFLOPs and 24.0 ms inference time. Although it is not the fastest configuration, it remains computationally feasible for batch processing and preliminary image screening. Compared with models with similar computational complexity, such as UAV-YOLOv8 and EDGS-YOLOv8, YOLO-SDD achieves substantially higher mAP@50. The lower performance of UAV-YOLOv8 and EDGS-YOLOv8 may reflect domain-transfer limitations, because these models were originally optimized for different UAV or small-object detection settings, whereas the present dataset contains single-class, irregular, fragmented post-disaster debris that is visually confused with roofs, vehicles, vegetation, shadows, and damaged structures.

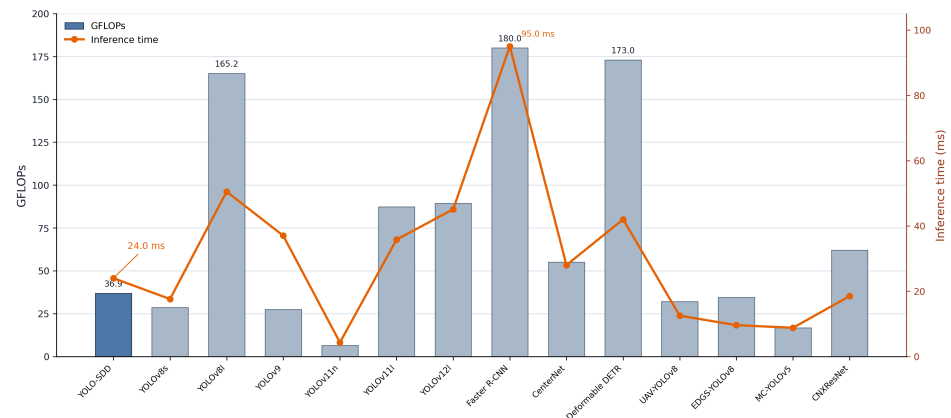


Figure 6. Comparison of computational complexity and inference latency across evaluated detection models.

The metric comparison confirms that the strongest performance is concentrated in the YOLO-family models, with YOLO-SDD, YOLOv12l, YOLOv8s, and YOLOv11l being the most competitive. As shown in Figure 7, YOLO-SDD achieves the highest mAP@50 and recall, while YOLOv12l gives a marginally higher mAP@50:95. This result suggests that YOLO-SDD improves standard detection accuracy under the tested setting, although precise bounding-box alignment for small and irregular debris objects remains a continuing challenge.

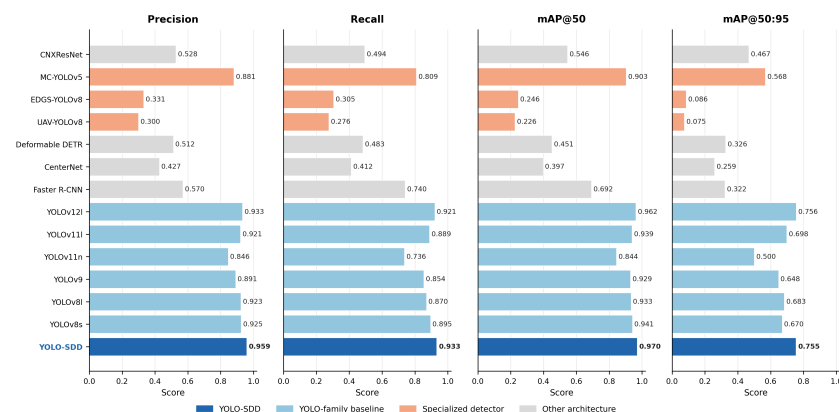


Figure 7. Comparison of evaluated detection models in terms of precision, recall, mAP@50, and mAP@50:95.

The precision–recall curves provide additional evidence of the model’s operating behavior across confidence thresholds. Figure 8 reports AP values consistent with Table 2, including 97.00% for YOLO-SDD, 94.10% for YOLOv8s, 93.30% for YOLOv8l, and 92.90% for YOLOv9. The leading YOLO-family detectors maintain relatively high precision over much of the recall range, indicating their ability to recover a substantial share of debris instances while controlling false detections.

To examine whether detection performance varied with object scale, debris boxes were grouped according to the standard COCO object-size definitions after each image was resized to 640×640 pixels. Specifically, boxes with an area smaller than 1024 pixels² were assigned to the small group, boxes with an area from 1024 to less than 9216 pixels² were assigned to the medium group, and boxes with an area of at least 9216 pixels² were assigned to the large group. AP@50 and recall were then computed separately for each size group.

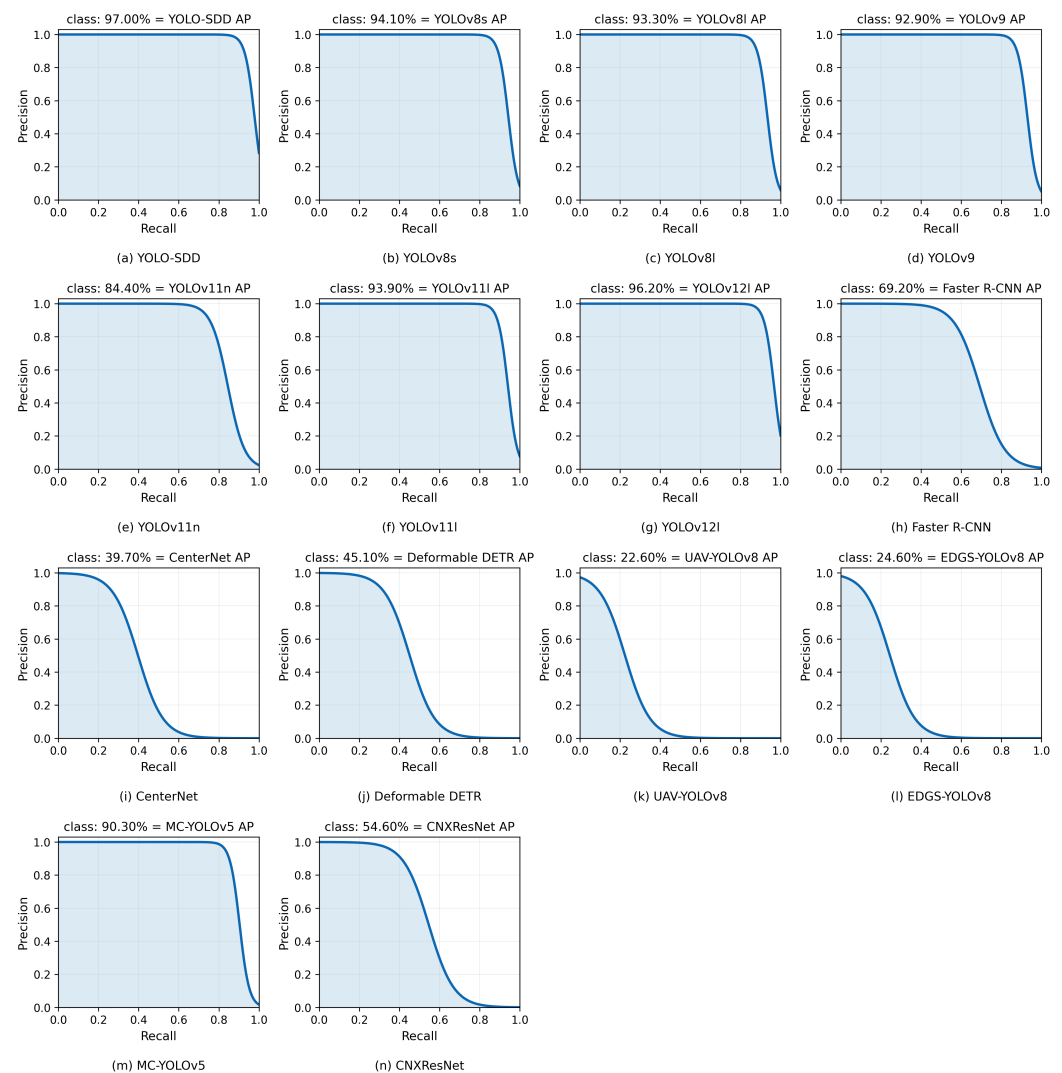


Figure 8. Precision–recall curves for the evaluated debris detection models.

As shown in Table 5, the small, medium, and large groups contain 99, 121, and 110 debris boxes, corresponding to 30.0%, 36.7%, and 33.3% of the 330 ground-truth boxes, respectively. YOLO-SDD improves AP@50 and recall for small and medium debris, indicating that the high-resolution pathway and shape-guided refinement are most beneficial for smaller and intermediate debris objects under these thresholds.

Table 5. Scale-specific AP@50 and recall on the held-out test set using COCO object-size thresholds.

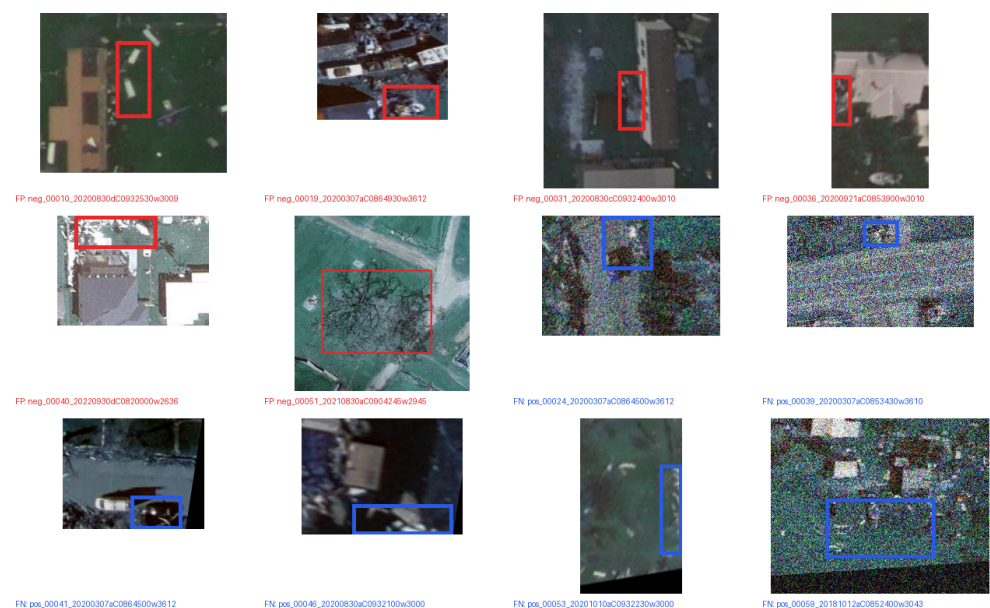
Size Group	GT Boxes	Prop.	YOLOv8s AP	YOLOv8s R	YOLO-SDD AP	YOLO-SDD R
Small (<1024 pixels ²)	99	30.0%	0.498	0.818	0.697	0.869
Medium (1024–9216 pixels ²)	121	36.7%	0.539	0.744	0.703	0.802
Large (≥9216 pixels ²)	110	33.3%	0.710	0.818	0.655	0.745

To further evaluate statistical stability, a three-seed comparison between YOLOv8s and YOLO-SDD was added. Table 6 reports the mean and standard deviation across independent training runs. The results show that YOLO-SDD maintains higher precision, recall, mAP@50, and mAP@50:95 than the YOLOv8s baseline across the three training seeds.

Table 6. Multi-seed stability comparison for the YOLOv8s baseline and YOLO-SDD.

Model	Seeds	Precision	Recall	mAP@50	mAP@50:95
YOLOv8s baseline	3	0.928 ± 0.007	0.900 ± 0.010	0.942 ± 0.007	0.672 ± 0.012
YOLO-SDD	3	0.958 ± 0.007	0.934 ± 0.009	0.969 ± 0.004	0.753 ± 0.010

To further examine the limitations of YOLO-SDD, representative false-positive and false-negative cases were reviewed. As shown in Figure 9, false positives predominantly arose from non-debris objects with debris-like visual patterns, including roof textures, vehicles, vegetation, shadows, and damaged structural elements. False negatives mainly involved very small debris, debris partially covered by vegetation, debris in heavy shadows, and low-contrast or highly fragmented debris regions. These examples indicate that while YOLO-SDD can identify many candidate debris regions, boundary ambiguity and background confusion remain significant challenges in post-disaster aerial imagery.

**Figure 9.** Representative false-positive and false-negative cases for YOLO-SDD. Red boxes indicate false positives, and blue boxes indicate false negatives.

4.6. Discussion

The case study demonstrates that YOLO-SDD can detect localized debris in high-resolution aerial imagery from multiple disaster events under the tested conditions. In the mixed-event held-out test set, YOLO-SDD achieved the highest precision, recall, and mAP@50 among the evaluated models while maintaining moderate inference time. This finding suggests that the high-resolution P2 pathway and shape-guided refinement can improve debris detection without making the model computationally impractical. However, the results should be interpreted within the scope of the available dataset, which remains limited to a single debris class. Additional labeled imagery and multi-class debris categories are needed to evaluate broader operational use.

Because the main dataset was split at the image level rather than by disaster event, the mixed-event results should not be interpreted as full evidence of generalization to a completely unseen disaster event. The separate Sally-period hold-out experiment provides a complementary event-level check. Because only one event group was used for the event-wise hold-out test, the cross-event results should be interpreted as a preliminary generalization check rather than definitive evidence of broad event-level transferability. Its

metrics are lower than the mixed-event test performance, which is expected because the held-out event introduces visual conditions not represented in the training distribution. Nevertheless, YOLO-SDD achieved the highest AP@50 and F1 score among the selected event-level comparison models under the Sally-period hold-out setting. This result provides preliminary support for improved event-level debris detection, but it should not be interpreted as definitive evidence of broad cross-event transferability.

The present study evaluates YOLO-SDD as an offline image-based debris screening model. The current framework takes post-disaster aerial image tiles as input and produces debris bounding boxes as output. It does not yet include an operational emergency-response workflow, such as automatic imagery ingestion, GIS-based routing, road-access analysis, cleanup-priority ranking, or field-team feedback. Therefore, the results should be interpreted as evidence for preliminary debris screening rather than direct operational deployment. Integration with emergency-management platforms and human-in-the-loop review remains necessary for practical use.

5. Conclusions and Future Research

This study introduced YOLO-SDD, a model for localized debris detection in high-resolution post-disaster aerial imagery. YOLO-SDD integrates a YOLOv8-based detector, a high-resolution P2 detection pathway, and a shape-guided P2 feature refinement module supervised by box-derived pseudo-mask and pseudo-boundary cues. On the held-out test set, the model achieved a precision of 0.959, recall of 0.933, mAP@50 of 0.970, and mAP@50:95 of 0.755. These results demonstrate strong detection accuracy, with the highest precision, recall, and mAP@50 among the evaluated models, while mAP@50:95 remained comparable to the strongest large YOLO-family model.

Ablation experiments show that the full model achieved the highest mAP@50 and mAP@50:95 among the tested ablation configurations, although its improvement over the P2 + FPN/PAN configuration was incremental. The Sally-period event-wise evaluation, a stricter generalization test, exhibited lower performance than the mixed-event test—an expected outcome when the held-out event group differs from the training distribution. Nevertheless, YOLO-SDD attained the highest AP@50 and F1 among the event-level models compared. These findings support the use of YOLO-SDD as a preliminary debris-screening and decision-support tool under the tested conditions.

Several limitations remain. The dataset employs a single debris class and does not distinguish debris material, source, severity, or operational priority. False positives and false negatives may arise in scenes with shadows, vegetation, roofs, vehicles, occlusion, low contrast, and highly fragmented debris. Future work should extend the dataset to additional disaster events, geographic regions, sensors, and debris categories and assess integration into emergency-response workflows, including GIS-based routing, road-access assessment, cleanup-priority ranking, and human-in-the-loop review.

Author Contributions: Conceptualization, H.A.-D. and L.G.; methodology, M.V. and L.G.; software, M.V.; validation, M.V.; formal analysis, M.V.; investigation, M.V.; data curation, M.V.; writing—original draft preparation, M.V.; writing—review and editing, H.A.-D., L.G., Y.Z., J.S., A.S. and K.F.; visualization, M.V.; supervision, H.A.-D. and L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The source imagery used in this study was obtained from the NOAA Emergency Response Imagery portal. Processed data, annotation files, model configuration files, and training/evaluation scripts are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dong, B.; Ding, S.; Wu, L.; Li, X. Short-term natural disaster impacts on transportation infrastructure: A systematic review. *Nat. Hazards* **2025**, *121*, 17321–17362. <https://doi.org/10.1007/s11069-025-07504-9>. Correction in *Nat. Hazards* **2025**, *121*, 24375. <https://doi.org/10.1007/s11069-025-07679-1>.
2. Cheng, C.; Zhang, L.; Thompson, R.G. Disaster waste clean-up system performance subject to time-dependent disaster waste accumulation. *Nat. Hazards* **2018**, *91*, 717–734.
3. Karaer, A.; Ulak, M.B.; Abichou, T.; Arghandeh, R.; Ozguven, E.E. Post-hurricane vegetative debris assessment using spectral indices derived from satellite imagery. *Transp. Res. Rec.* **2021**, *2675*, 504–523.
4. Fetter, G.; Rakes, T. Incorporating recycling into post-disaster debris disposal. *Socio-Econ. Plan. Sci.* **2012**, *46*, 14–22.
5. Bianchi, S.; Ciurlanti, J.; Pampanin, S. Comparison of traditional vs low-damage structural and non-structural building systems through a cost/performance-based evaluation. *Earthq. Spectra* **2021**, *37*, 366–385. <https://doi.org/10.1177/8755293020952445>.
6. Antwi, R.B.; Takyi, S.; Ozguven, E.E.; Volcy, K.; Alam, M.S.; Kim, K.; Horner, M.W.; Mardis, M.A. Satellite Imagery-Based Hurricane Debris Assessment: Model Development and Application in Rural Florida. *Res. Sq.* **2023**, preprint. <https://doi.org/10.21203/rs.3.rs-3652245/v1>.
7. Jozi, D.; Shirzad-Ghaleroudkhani, N.; Luhadia, G.; Abtahi, S.; Gül, M. Rapid post-disaster assessment of residential buildings using Unmanned Aerial Vehicles. *Int. J. Disaster Risk Reduct.* **2024**, *111*, 104707. <https://doi.org/10.1016/j.ijdr.2024.104707>.
8. Joshi, A.R.; Tarte, I.; Suresh, S.; Koolagudi, S.G. Damage identification and assessment using image processing on post-disaster satellite imagery. In *Proceedings of the 2017 IEEE Global Humanitarian Technology Conference (GHTC)*; IEEE: New York, NY, USA, 2017; pp. 1–7. <https://doi.org/10.1109/GHTC.2017.8239286>.
9. Gueguen, L.; Hamid, R. Large-scale damage detection using satellite imagery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 7–12 June 2015; pp. 1321–1328. <https://doi.org/10.1109/CVPR.2015.7298737>.
10. Pi, Y.; Nath, N.D.; Behzadan, A.H. Convolutional neural networks for object detection in aerial imagery for disaster response and recovery. *Adv. Eng. Inform.* **2020**, *43*, 101009. <https://doi.org/10.1016/j.aei.2019.101009>.
11. Qiao, G.; Yang, M.; Wang, H. A detection approach for floating debris using ground images based on deep learning. *Remote Sens.* **2022**, *14*, 4161. <https://doi.org/10.3390/rs14174161>.
12. Cheng, C.S.; Luo, L.; Murphy, S.; Lee, Y.C.; Leite, F. A framework to enhance disaster debris estimation with AI and aerial photogrammetry. *Int. J. Disaster Risk Reduct.* **2024**, *107*, 104468. <https://doi.org/10.1016/j.ijdr.2024.104468>.
13. Amini, K.; Liu, Y.; Padgett, J.E.; Balakrishnan, G.; Veeraraghavan, A. Post-Hurricane Debris Segmentation Using Fine-Tuned Foundational Vision Models. *arXiv* **2025**, arXiv:2504.12542. <https://doi.org/10.48550/arXiv.2504.12542>.
14. Prakash, G.; Kasthala, S.; Loya, A. Deep learning for automatic post-disaster debris identification for precise damage assessments using UAV footage. *Appl. Geomat.* **2025**, *17*, 269–279. <https://doi.org/10.1007/s12518-025-00616-8>.
15. Yamazaki, F.; Kouchi, K.; Kohiyama, M.; Muraoka, N.; Matsuo, M. Earthquake damage detection using high-resolution satellite images. In *Proceedings of the IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*; IEEE: New York, NY, USA, 2004; Volume 4, pp. 2280–2283. <https://doi.org/10.1109/IGARSS.2004.1369739>.
16. Kerle, N. Satellite-based damage mapping following the 2006 Indonesia earthquake—How accurate was it? *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 466–476. <https://doi.org/10.1016/j.jag.2010.07.004>.
17. Radhika, S.; Tamura, Y.; Matsui, M. Cyclone damage detection on building structures from pre-and post-satellite images using wavelet based pattern recognition. *J. Wind Eng. Ind. Aerodyn.* **2015**, *136*, 23–33. <https://doi.org/10.1016/j.jweia.2014.10.018>.
18. Miura, H. Fusion analysis of optical satellite images and digital elevation model for quantifying volume in debris flow disaster. *Remote Sens.* **2019**, *11*, 1096. <https://doi.org/10.3390/rs11091096>.
19. Li, Y.; Hu, W.; Dong, H.; Zhang, X. Building damage detection from post-event aerial imagery using single shot multibox detector. *Appl. Sci.* **2019**, *9*, 1128. <https://doi.org/10.3390/app9061128>.
20. Al Shafian, S.; Hu, D. Integrating machine learning and remote sensing in disaster management: A decadal review of post-disaster building damage assessment. *Buildings* **2024**, *14*, 2344. <https://doi.org/10.3390/buildings14082344>.
21. Hussain, M. YOLOv1 to v8: Unveiling each variant—a comprehensive review of yolo. *IEEE Access* **2024**, *12*, 42816–42833. <https://doi.org/10.1109/ACCESS.2024.3378568>.
22. Hua, W.; Chen, Q. A survey of small object detection based on deep learning in aerial images. *Artif. Intell. Rev.* **2025**, *58*, 162. <https://doi.org/10.1007/s10462-025-11150-9>.

23. Yao, G.; Zhu, S.; Zhang, L.; Qi, M. HP-YOLOv8: High-Precision Small Object Detection Algorithm for Remote Sensing Images. *Sensors* **2024**, *24*, 4858. <https://doi.org/10.3390/s24154858>.
24. Yuan, Z.; Gong, J.; Guo, B.; Wang, C.; Liao, N.; Song, J.; Wu, Q. Small Object Detection in UAV Remote Sensing Images Based on Intra-Group Multi-Scale Fusion Attention and Adaptive Weighted Feature Fusion. *Remote Sens.* **2024**, *16*, 4265. <https://doi.org/10.3390/rs16224265>.
25. Li, Y.; Li, Q.; Pan, J.; Zhou, Y.; Zhu, H.; Wei, H.; Liu, C. SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images. *Remote Sens.* **2024**, *16*, 3057. <https://doi.org/10.3390/rs16163057>.
26. Zhang, Y.; Zhang, W.; Wang, D. PARE-YOLO: Fusion of Multi-Scale Attention for Aerial Images Small-Target Detection. *Sci. Rep.* **2025**, *15*, 4753. <https://doi.org/10.1038/s41598-025-88857-w>.
27. Qu, S.; Dang, C.; Chen, W.; Liu, Y. SMA-YOLO: Small Object Detection in UAV Images Based on Improved YOLOv8. *Remote Sens.* **2025**, *17*, 2421. <https://doi.org/10.3390/rs17142421>.
28. Han, Z.; Ye, K. An improved YOLOV5 based water surface garbage detection algorithm. In Proceedings of the International Conference on Algorithms, Software Engineering, and Network Security, Nanchang, China, 26–28 April 2024; pp. 135–139. <https://doi.org/10.1145/3677182.3677207>.
29. Bansal, K.; Tripathi, A.K. Automated floating debris monitoring using optical satellite imagery and artificial intelligence: Recent trends, challenges and opportunities. *Remote Sens. Appl. Soc. Environ.* **2025**, *37*, 101475. <https://doi.org/10.1016/j.rsase.2025.101475>.
30. National Geodetic Survey (NGS). NOAA's Emergency Response Imagery. Available online: <https://oceanservice.noaa.gov/hazards/emergency-response-imagery.html> (accessed on 3 June 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.