



An improved YOLOv8 for foreign object debris detection with optimized architecture for small objects

Javaria Farooq¹ · Muhammad Muaz^{1,2} · Khurram Khan Jadoon³ · Nayyer Aafaq¹ · Muhammad Khizer Ali Khan⁴

Received: 31 May 2023 / Revised: 16 November 2023 / Accepted: 7 December 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Automated Foreign Object Debris (FOD) detection offers significant benefit to the aviation industry by reducing human error and enabling continuous surveillance. This paper focuses on addressing the intricacies of FOD detection, with a specific emphasis on treating FODs as “small” objects, a facet which has received limited attention in prior research. This study provides a pioneering evaluation of state-of-the-art object detectors, including both anchor-based models including SSD, YOLOv5m, Scaled YOLOv4 and anchorless models CenterNet and YOLOv8m, applied to a multiclass FOD dataset, “FOD in Airports (FOD-A)”, as well as meticulously curated subset of FOD-A featuring small FODs. The findings reveal that the anchorless object detector YOLOv8m gives the best time accuracy trade off outperforming all compared anchor-based and anchorless approaches. To address the challenge of detecting small FODs, this study optimizes YOLOv8m model by making architectural modifications and incorporating a dedicated, shallow detection head that is purpose-built for the precise identification of small objects. The proposed model, termed as “Improved YOLOv8”, outperforms YOLOv8m by a margin of 1.02 in Average Precision for small objects (APs), achieving a mean average precision (mAP) of 93.8%. Notably, Improved YOLOv8 also has better mAP than all the considered anchor-based and anchorless object detectors examined, as well as those featured in prior FOD-A dataset research.

Keywords Foreign object debris · FOD detection · Deep learning

✉ Javaria Farooq
jfarooq.ms14avecae@student.nust.edu.pk

¹ College of Aeronautical Engineering, National University of Sciences and Technology, Islamabad, Pakistan

² Hong Kong Industrial Artificial Intelligence and Robotics Centre Limited, Hong Kong, China

³ Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Pakistan

⁴ Khalifa University of Science and Technology, Abu Dhabi, UAE

1 Introduction

In aviation terminology, foreign object debris (FOD) are defined as undesirable objects on the airport runways / taxiways, that acts as a hazard to the aircraft operating on those surfaces. Stones, small hardware (nuts, bolts, washers etc.), maintenance tools, key chains, garment accessories, and vehicles on the runway are some of the most common sources of FOD. These objects, no matter how small, pose a serious risk to aircraft and flight safety, as they are either drawn inside rotating engines [1], or collide with aircraft outer body, resulting in severe damage [2]. One such example is the incident of Air France Flight 4590 where a small metal strip on the runway ruptured the concorde aircraft tyres, that eventually led to its crash, killing all 113 passengers onboard [3]. Hence, the risk that FOD poses to aircraft and flight safety cannot be overstated.

Apart from the short and long term disruption of flight schedules for the airlines, FOD related damages account for a global economic loss of approximately 12 billion USD per year [4, 5]. To mitigate the occurrence of FOD-related damages, many airports rely on manual preventive measures such as magnetic sweepers and visual runway inspections, however, these methods are prone to human errors and consume significant time and resource. In this context, automated FOD detection is a more viable alternative. Conventional automated FOD detection methods mainly encompass millimeter-wave radars [6, 7], multiple-sensor hybrid target detection technology [8] and optical image sensors technology [9–11]. Radar-based method uses two to three units of milli-meter-wave radar for each runway. Although, these technologies can accurately locate FOD; radar based technologies are expensive and result in high false-positive rate for detecting small sized FOD items [12]. Optical Image sensor technology leverages advanced image processing algorithms to detect FODs under varying lighting and surface conditions. In general, five to ten units of high-resolution cameras installed along the specified airport pavement which visually inspect the runway surface. If FOD is detected, the optical system readily captures image/ video, which is relayed to the airport authorities for subsequent removal from the indicated location. FOD differs in features from traditional objects because FOD items are generally small in size with a lot of variation in appearance [13]. Different definitions have been proposed for various datasets to determine whether the objects are “small” or not [14]. One definition calls them as the objects having smaller physical sizes in the real world [15]. Another definition refers to objects occupying area less than and equal to 32×32 pixels in an image [15]. Third definition refers to objects with sizes filling 20% of an image [14, 16]. Considering these definitions, FOD detection can be treated as a small object detection problem. With the recent advancements in deep learning, sophisticated computer vision techniques have been researched for object detection. Although there has been a significant improvement in the accuracy of object detection techniques including You Only Look Once (YOLO) [17], Single Shot Multibox Detector (SSD) [18], Fast Region-based Convolutional Neural Network (Fast-RCNN) [19], Faster-RCNN [20], and Feature Pyramid Network (FPN) [21] for large scale standard datasets such as MS-COCO [22] and PASCAL VOC [23], their effectiveness in identifying small objects with fewer pixels remains constrained [24]. One possible explanation is that standard datasets are not focused on small objects causing the detection models to be biased towards medium to large objects [14, 25].

Small object detection is difficult due to low-resolution of images, intricate backgrounds [26], view point variations, and diverse object geometry. Additionally, due to fewer pixels, there are less information representatives and key features of small objects get gradually reduced while going through different layers of convolutional neural networks (CNN) [14].

Many of the methods optimized for detecting small objects have been assessed using single-category datasets such as traffic signs [16], vehicles [27, 28], and pedestrian detection [29, 30]. In contrast, FODs encompass multiple categories with diverse appearances and sizes.

Over the past few years, researchers have explored various deep learning techniques for FOD detection by employing Convolutional Neural Networks. The techniques evaluated for various FOD datasets are predominantly founded on anchor based object detectors including two stage object detectors such as Convolutional Neural Networks with Faster RCNN [31], Convolutional Neural Networks with RPN [9], and Convolutional Neural Network with improved RPN [11], as well as one stage anchor-based object detectors such as Single Shot Detector (SSD) [32], Lighter Network SSD [10], YOLOv3 [32, 33], and YOLOv4-csp w/Augmentation [34]. Additionally, Visual transformer (ViT) has also been explored for the FOD detection problem [35]. This study reveals that a significant portion of earlier research in deep learning-based FOD detection has centered around training and assessing techniques on proprietary datasets (not publicly accessible) [9, 10, 31, 33, 36] typically limited to 2-5 FOD categories. One plausible explanation for this trend is the non-availability of publicly available multi-class FOD detection datasets. However, a recent development in this domain is the introduction of a comprehensive FOD dataset referred to as 'FOD in Airport (FOD-A)' by Munyer et al. [32], which encompasses 31 categories across more than 33,000 images. Munyer et al have also evaluated their FOD-A dataset with YOLOv3, SSD, and ViT [32, 35]. Nevertheless, these approaches have not been optimized to excel in small object detection

This work makes a threefold contribution. Firstly, an in-depth examination is offered for both traditional and deep learning computer vision techniques applied to FOD detection, along with a detailed overview of all available FOD datasets. Subsequently, a comprehensive analysis of the characteristics of the publicly accessible FOD-A dataset is conducted, revealing that FOD objects within this dataset occupy less than 20% of the image area in more than 85% of cases. As a result, FOD-A is categorized as a dataset primarily focused on small objects. Building on this insight, the small object detection algorithms are examined from the literature and two anchor-based approaches: Scaled YOLOv4 (P5 and P6) [37, 38], YOLOv5 (medium, large) [39] and two anchor-free approaches CenterNet [40, 41] and YOLOv8 (medium) [42] for performance analysis, both in terms of accuracy and inference time, for FOD detection problem are selected. These models are chosen based on their better detection accuracy, particularly their average precision for small objects (with an area smaller than 32^2), as established within the MS-COCO dataset [22] along with their real-time inference speed as reported in [43, 44]. SSD performs better than YOLOv3 for FOD-A dataset [32], therefore it is also re-evaluated in the similar experimentation settings with the selected algorithms. The performance of selected algorithms is analyzed in terms of speed and accuracy for FOD-A dataset as well as a manually selected subset of FOD-A dataset comprising of only small objects. Based upon the performance analysis of distinct state-of-the-art anchor based and anchor free object detectors specifically suited to the characteristics of small sized FOD detection problem, an architecture level improvement is proposed to optimize the best performing model i.e YOLOv8m for detecting small FODs. The contributions of this work are summarized as follows. This work:

- provides first comprehensive analysis of computer vision based FOD detection techniques and datasets, with a focus on FOD-dataset
- conducts performance evaluation of meticulously selected state-of-the-art (SOTA) anchor free object detection techniques, including CenterNet, and YOLOv8, as well as anchor based object detection techniques such as SSD, Scaled YOLOv4, and YOLOv5 to establish benchmark for recently introduced FOD-A dataset comprising 31 classes

- curates a more challenging subset from FOD-A dataset constituting solely small objects (i.e., objects occupying an area less than 20% in an image) and examines the selected SOTA techniques with a specific focus on their performance concerning small object detection.
- optimizes the best performing model i.e YOLOv8m performance for accurately detecting small target objects by introducing a specialized small target detection layer in the original architecture.
- qualitatively evaluates generalization of the selected techniques for out-of-distribution images from three classes of FOD-A dataset i.e. wrench, metal part and bolt, on actual runway under varying lighting conditions.

The rest of the paper is organized as follows. Section 2 provides an extensive review of object detection algorithms from existing literature, categorizing small object detection algorithms with respect to baseline approaches. It also provides a comprehensive overview of both conventional and deep learning approaches for FOD detection. Section 3 offers a concise analysis of available datasets for the FOD detection problem and explains the classification of the FOD-A dataset as a small object dataset. Section 4 presents the experimental details of the object detection algorithms and offers a brief overview of the architectural characteristics of the selected anchor-based and anchor-free approaches. Section 5 explains the architecture of the proposed “Improved YOLOv8” model. Section 6 delivers an exhaustive comparative analysis, evaluating the performance of selected anchor based and anchor free approaches with the Improved YOLOv8 in terms of speed, accuracy, training time, and generalization for both the FOD-A dataset and a manually curated subset of the FOD-A test dataset. Section 7 concludes this study and suggests directions for future research.”

2 Object detection algorithms

Object detection is comprised of two sub tasks: object localization and categorization [45]. Conventional object detection approaches [46, 47] employ shallow training architectures coupled with rule based methodologies. These approaches are typically dependent on manually engineered features such as colour, shape, and texture which can be effective for identifying a particular object type. However, due to their vulnerability to variations in lighting, scale, and viewpoints, conventional object detection approaches encounter difficulties in achieving effective generalization, particularly when dealing with objects that exhibit diverse appearances and backgrounds.

In contrast, modern deep learning techniques, particularly convolutional neural networks (CNNs), have emerged as powerful tools capable of assimilating semantic, higher-level, and deeper features from the data, making them more robust and capable of generalizing better. A DCNN-based object detector is typically composed of two integral components: a backbone network responsible for extracting image features and a head component for predicting bounding boxes and classifications. In contemporary developments, object detectors often incorporate additional layers between the backbone network and the head. These layers, termed the “neck”, primarily facilitate the fusion of features across different layers. Typically, the neck incorporates several bottom-up and top-down pathways to achieve this feature fusion [21].

Based on the approach used to ascertain the geometry and class of an object, CNN based object detection approaches can be broadly classified into anchor-based and anchor-free approaches [48]. Anchor-based approaches first position several anchors having distinct scale

and shape across the image plane as regional proposals, and then train a dedicated classifier to ascertain both the likelihood of an object's presence and its corresponding class. The fundamental quality of regional features obtained from the proposal can be further enhanced by leveraging more potent network backbones [49] and achieving a more precise alignment between the proposed elements and the extracted features [20, 36].

Anchor-based detectors can be further classified into one stage and two stage methods. One stage methods such as YOLO [17], SSD [18], Retina Net [50] predict location and label of object concurrently without RPN whereas two stage detectors such as Faster RCNN [20], Mask RCNN [51], and FPN [21] divide detection in two stages: region proposal stage and detection stage [26]. Two-stage detectors, in comparison to one-stage detectors, tend to offer higher accuracy at the cost of slower processing speeds [52].

Anchor-free approaches do not presume that objects are associated with uniformly distributed anchors. These approaches do not require complicated anchor related calculations and generalize better for objects with varying shapes [53]. The foundation of anchor-free methodologies often revolves around a single or a few keypoints while the manner in which these keypoints are utilized to represent objects varies depending on the specific approach [52].

Anchor free approaches can be sub-divided into key-point based and center-based methods. Spatial extent of objects is bounded by key-point based methods [40, 54] using self learned / predefined key-points. Center-based methods [55, 56] first points the center of objects and then regress to the object boundaries [48]. Figure 1 illustrates the object detection sub categories along with state-of-the-art anchor-based and anchor-free methods for object detection.

2.1 Small object detection algorithms

Small object detection is a specific scenario within the broader context of object detection, addressing the unique difficulties associated with detecting objects that are comparatively diminutive in size. Small object detection is challenging owing to less representation (number of pixels) of small objects in an image, the background, and variations in view point. Since last few years, one stage detection algorithms such as the one in [17] and SSD [18], and two stage detection algorithms such as Faster RCNN [20] and FPN [21] have been modified / improved for better detection performance for small objects. A summary of recently proposed state-of-the-art object detection methods is tabulated in Table 1. In the literature, small object

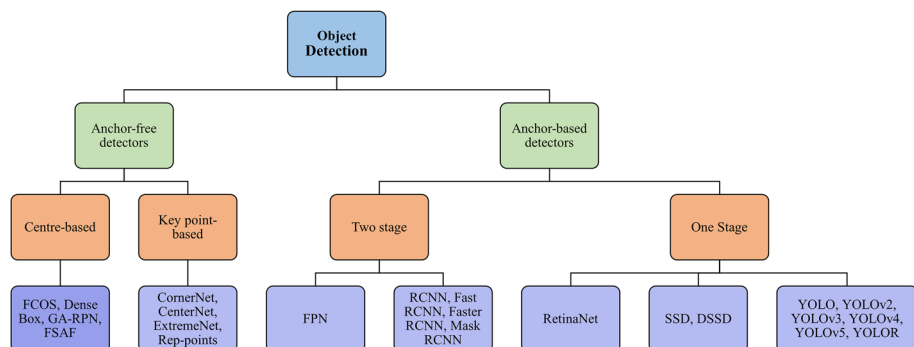


Fig. 1 State-of-the-art approaches for anchor-based and anchor-free object detection

Table 1 Small object detection algorithms with evaluated datasets

Algorithm type	Algorithm	Brief overview	Dataset evaluated
YOLO-based algorithms	YOLO-Fine [57]	Introduces a refined object search grid in YOLOv3 to detect 'small' and 'very small' objects.	VEDAI, Munich and xView
	YOLO-MXANet [58]	Utilizes Complete-Intersection over Union (CIoU) to optimize loss function for improving positioning accuracy of small objects.	KITTY and CCTSDB
	Improved YOLOv3 algorithm [59]	Enhances feature map acquisition network and adds a size recognition module to input image for better recognition of small objects.	VEDAI, DLR 3K Munich vehicle
	UAV-YOLO [60]	Based on YOLOv3, UAV-YOLO optimizes the Resblock in Darknet by concatenating two ResNet units having identical width and height along with improvement in darknet structure.	UAV123 and UAV-viewed
	YOLOv4 /improved YOLOv4 [61]	YOLOv4 modifies backbone and utilizes SPP block integrated with CSPDarknet53 backbone along with mosaic augmentation for better small object detection. YOLOv4 is further improved/optimized for particular small object detection task such as oil well detection [62], bridge crack detection [63].	MS COCO
	YOLOv5/ Improved YOLOv5 [39]	YOLOv5 employs CSPDarknet53 backbone with PANet based on FPN for small object detection. YOLOv5 is further improved/ optimized for small object problem such as UAV images [64], autonomous vehicles [65], planetary images [66] etc.	MS COCO, VOC 2007
	YOLOv8/ Improved YOLOv8 [42]	YOLOv8 employs a dynamic head network with default mosaic augmentation that enhances the accuracy of object object detection. YOLOv8 is further improved/ optimized for small object problem such as fire detection [67], helmet detection [68] etc.	MS COCO, VOC 2007
SSD-based algorithms	Mask-guided SSD [69]	Employs a mask of segmentation branch to identify focused regions of the captured image, which assists detection branch to locate target objects in these regions.	TT100K, Caltech
	Multi-block SSD [70]	Slices the input image to four overlapped sub images to enhance local contextual information.	ILSVRC CLS-LOC

Table 1 continued

Algorithm type	Algorithm	Brief overview	Dataset evaluated	
Faster RCCN based algorithms	Feature-Fused SSD [71]	Adds contextual information in base-line SSD by designing multiple modules at different levels for feature fusion.	VOC 2007	
	Context and Attention algorithms [72]	Enhance SSD with context aware information from multi scale feature fusion and attention.	VOC 2007	
	Improved Faster R-CNN [73]	Optimizes the loss function based upon intersection over Union (IoU) to achieve bounding box regression, with bilinear interpolation to improve RoI pooling operation which improves positioning of small objects.	TT100K	
	Enhanced faster RCNN [74]	Improves feature maps resolution for enhancing small object detection accuracy.	Flicker Dataset	
FPN based algorithms	Deep feature pyramid network [75]	Builds feature pyramids during Region proposal stage and introduces special anchors to detect small objects.	TT100K	
	Enhanced Fusion SSD [76]	Merges selected multiscale feature layer with scale-invariant convolutional layer across FPN to generate a set of enhanced feature maps.	VOC 2007	
	Extended feature pyramid (EFPN) [24]	Uses a novel module called feature texture transfer (FTT) for super resolution of features along with extraction of credible regional details concurrently.	TT100K, MS COCO	
	Effective Fusion Factor in FPN [77]	Uses a statistical method for determining fusion factor for various datasets which assists shallower layers to detect smaller objects with a better accuracy.	VOC 2007, MS COCO and City Persons	
GAN based algorithms	Image Pyramid Guidance Network [78]	Introduces the image pyramid information into the backbone stream to balance between spatial and the semantic information.	VOC 2007, MS COCO	
	End-to-End Edge-Enhanced GAN [79]	Enhances super-resolution architecture to detect small objects by integrating GAN based on residual blocks into a cycle model.	ISPRS xView	Potsdam,
	Perceptual GAN [80]	Boosts detection performance by creating super resolved representations for small objects similar to large objects by competitive optimization of generator and discriminator.	TT100K, Caltech	

Table 1 continued

Algorithm type	Algorithm	Brief overview	Dataset evaluated
	End-to-End multi-task GAN (MTGAN) [81]	Uses generator to super resolve low resolution images whereas the discriminator grades each upscaled image from the generator with a real/fake score and concurrently provides scores for bounding box regression and object categorization.	MS COCO
Combination approaches	Feature based super resolution [82]	Boosts small RoI features and shows its appreciable impact on small object detection.	TT100K, VOC 2007 and MS COCO
	Image Tiling [83]	Uses PeleeNet and improves small object detection for high resolution images from a micro aerial vehicle (MAV) by tiling images in both training and inference phase to prevent loss of detail.	ImageNet, VOC 2007 and MS COCO
	CenterNet [40]	Detects object using only center point of object bounding box as key point without requiring post-processing. CenterNet has been optimized to achieve better accuracy for a particular small object problem [84–86].	MS COCO, VOC 2007

detection algorithms can be broadly classified into the following five subcategories according to the baseline approaches:

- YOLO,
- SSD,
- Faster RCNN,
- FPN, and
- GAN.

From Table 1, it can be observed that most of small object detection algorithms are the extended / modified forms of such baseline architectures evaluated commonly for Tsinghua-Tencent 100K [16], MS-COCO [22], and VOC 2007 [23] datasets.

2.2 FOD detection algorithms

With the rapid improvement in object detection techniques, CV algorithms are employed for FOD detection. CV systems employ high-resolution optical sensors to scan the runway surface and generate alarm upon FOD detection for subsequent removal by airport staff [87]. Recent advancements in CV can equip computers to automatically detect and localize FODs round the clock with high degree of accuracy and fidelity. CV approaches for FOD detection can be broadly categorized into

- Conventional CV approaches, and
- Deep Learning CV approaches.

2.2.1 Conventional CV approaches for FOD detection

Before the wide adaptation of CNN architectures, research on FOD detection was directed towards manually engineered feature-based approaches. Although manual methods are effective, yet the effort and prior knowledge to build manually engineered feature extractors decrease the efficiency and generalization of such techniques. Li et al. [88] proposed an FOD detection algorithm using edge detection to identify runway region and background subtraction to detect FOD objects. Khan et al. [89] combined image sensing with ultrasonic proximity sensors for runway inspection. The prototype system is semi-automated and involved active participation of air traffic controller (ATC). ATC switches on the camera upon receiving FOD detection information from sensors and subsequently applies image processing algorithm on MATLAB image processing tool using MSER algorithms to get a better image of FOD for prompt removal by airport staff. Since airport runway images are observed to have obvious texture features, Liang et al. [90] proposed a method based on texture segmentation to detect foreign objects on the runway using Gabor Filter. At first, Gabor filter is applied to suppress background texture information and enhance image contrast. The enhanced image is further de-noised using bilateral filtering and subsequently threshold detection is employed to segment FOD on the runway.

2.2.2 Deep learning CV approaches for FOD detection

CNN achieves better detection accuracy than conventional feature-based methods using traditional features [11] such as scale invariant feature transform (SIFT) [94], histograms of oriented gradients (HOG) [95], and local binary patterns (LBP) [96]. New ideas for automatic FOD detection have been proposed aimed at improving object detection accuracy using CNN [13]. Since Region Proposal Network (RPN) performs better than sliding windows for target detection, Cao et al. [9] proposed a two-stage method for FOD detection using CNN with RPN. In the first stage, RPN is employed for generating detection proposals. RPN detection rate is enhanced by reducing the number of proposals by adopting a candidate selection method. In the second stage, Spatial Transform Network [97] is employed with CNN classifier, which allows spatial manipulation of input data and transforms feature maps within the network without any additional supervised training.

Caffe toolbox [98] is used for FOD detection using pre-trained models of VGG [99] on ILSVRC2012 dataset [100]. Upon evaluation on Airfield Pavement Dataset [9] comprising of 12231 images belonging to two classes (screws, stones), detection accuracy as well as detection rate of algorithm is found to be better as compared to Selective Search method [101]. The main drawback of the algorithm is its increased computational cost as for 512×512 input, it gives 2 Frames Per Second (FPS) on an Nvidia 1080, which is far from real time object detection [10]. To increase detection speed and accuracy on Air Field Pavement dataset, Cao et al. [10] proposed a one stage detector based on SSD [18]. The algorithm makes detection network lighter by using less number of CNN layers and employs dilated convolution [102] to increase the number of points linked with FOD in feature maps for better detection accuracy. To cater for the size variations of FOD objects, SPP-net [103] is employed, which resizes all input images and then fuses output of each image. The airfield pavement dataset is again evaluated by the author for one-stage lighter dilated network. This network is found to have better recall and false alarm rates than the classical SSD [18]. In terms of speed, [10] achieves about 10 FPS on an Nvidia GTX 1080, which is 5 times more than the speed achieved by

[9]. Cao et al. [11] further optimized RPN with STN two stage detector [9] using optical image sensor technology for better FOD detection accuracy. The algorithm employs an improved RPN at the first stage and CNN classifier with STN at the second stage for FOD detection. Improved RPN of [9] is achieved by setting additional rules for region proposal, which enhances the quality and reduces the number of region proposals to a great extent (approx. 60%). The improved two stage RPN with STN achieves better results compared to faster-RCNN, SSD [18], and Selective Search [101] for FOD detection on Airfield Pavement Dataset. It achieves 26 FPS on NVIDIA GTX 1080 with a higher accuracy as well as better speed (at least five times faster) for Aircraft Pavement dataset than [9] and [10].

Liu et al. [31] proposed a FOD detection approach based on CNN for optical image sensors and evaluated Airport Runway FOD Image dataset comprising four classes of objects namely small steel balls, metal nuts, large screws, and small screws. The algorithm uses Dense Net instead of conventional VGG16Net to extract features for Faster RCNN, which improves the computational speed. It modifies the RPN classification loss function and optimizes weights using Focal Loss to achieve better classification accuracy. The algorithm performs better than Faster RCNN in terms of detection accuracy with almost twice the speed. Mayiu et al. [92] proposed an approach termed as Bi-directional YOLO which introduces a weighted bidirectional operation similar to BiFPN into the PANet, and evaluate their approach on a customized dataset termed as FODInSues as well as FOD-A dataset. Ying et al. [93] proposed an unsupervised anomaly detection approach called Multi-Scale Feature Inpainting (MSFI) in which a deep learning feature inpainting module is designed for runway FOD detection. Jing et al. [93] evaluated MSFI for a customized FOD detection dataset.

Unmanned Aerial Vehicle (UAV) has also been used for FOD detection. Papadopoulos et al. [33] proposes FOD detection system containing RGB camera mounted on a UAV. FOD detection is achieved using AI algorithm trained on a locally collected dataset comprised of five classes namely paper, metal, bolts, plastic, and plastic bottles. Papadopoulos and Gonzalez [33] employs YOLOv3 [104] as well as Microsoft Azure Custom Vision [105] to compare FOD detection accuracy. The results reveal that Microsoft Azure Custom Vision performs better than the YOLOv3 for FOD detection on locally collected dataset. Noroozi et al. [34] proposed a framework based on YOLOv4-csp encompassing an Unmanned Aircraft System (UAS) designed for surveying and gathering data within an airport setting along incorporating data augmentation methods. They evaluated their proposed approach for a customized dataset collected on runway as well as for FOD-A dataset.

Semantic Segmentation classifies every pixel as per its semantic content in an image and has also been used for FOD detection. DeepLabv3+ is a framework employed for semantic segmentation. Gao et al. [91] collected Airport Runway Foreign Object Dataset from Guangzhou Baiyun International Airport containing 3000 images of common FOD items such as screws, wires and tools, and evaluated the dataset on DeepLabv3+. Upon experimentation, although the model is effective for segmenting large objects, its performance deteriorates for segmenting small FOD items.

Munyer et al. [32] proposed FOD in airports (FOD-A) dataset and upon evaluated, found SSD to be better than YOLOv3 in terms of detection accuracy. Munyer et al. [35] also proposed a Vision transformer based self supervised FOD detection method and found better results than SSD upon evaluating FOD-A dataset.

A summary for deep learning FOD detection approaches along with the datasets evaluated is given in Table 2.

Table 2 Deep learning computer vision approaches for FOD detection with evaluated datasets

S No	Deep Learning CV Approaches	Dataset	Availability/Number of Classes
1	CNN with RPN [9]	Airfield pavement dataset	Not available/ 02
2	Lighter Network SSD [10]	Airfield pavement dataset	Not available/ 02
3	CNN with improved RPN [11]	Airfield pavement dataset	Not available/ 02
4	CNN with Faster RCNN [31]	Airport runway FOD dataset	Not available/ 04
5	Microsoft Azure Custom Vision and YOLOv3 [33]	1500 locally collected images	Not available/ 05
6	Semantic Segmentation [91]	3000 locally collected images	Not available/ 01
7	YOLOv3 and SSD [32]	FOD-A dataset	Publicly available/ 31
8	Outer ViT [35]	FOD-A dataset	Publicly available/ 31
9	YOLOv4-csp w/Augmentation [34]	FOD-A dataset	Publicly available/ 31
10	Bidirectional YOLO [92]	FOD-A dataset and FODInSues	FOD-A (Publicly available)/ 31 FODInSues (Not available)/ 10
11	MSFI [93]	9042 locally collected images	Not available/ 15

3 Datasets for FOD detection

Although [9–11] evaluate detection accuracy for same dataset (Airport Pavement) and calculate computational efficiency on the same GPU (Nvidia GTX-1080), it is noteworthy that the dataset has only two classes and is not readily available. Table 2 shows that [31, 33, 91] also evaluate FOD detection accuracy on locally collected datasets with limited number of classes (four to five). According to Federal Aviation Authority (FAA), FOD generally includes engine fasteners, aircraft parts, tools, flight line items, catering material, apron items, construction material, construction debris, plastic, natural materials and weather contamination [106].

A dataset [107] for FOD classification is publicly available but it comprises of only 3000 zoomed in images and is focused on FOD material recognition. It classifies FOD into only three object categories namely metal, plastic, and concrete, which does not cover all types of FODs mentioned by FAA [106]. Moreover, algorithms trained on only zoomed-in images may not perform well for small FOD detection. FOD in Airports (FOD-A) dataset has been recently proposed by [32]. The objects covered in FOD-A dataset include common FOD items such as battery, bolts, nails, nuts, metal sheets, screws, tools, rocks, plastic parts, pliers, paint chip, clamps, wood, tape, wires, and fuel caps which are captured using static camera as well as from UAV mounted camera. FOD-A is a publicly available multi-class dataset with

- 31 categories comprising common FOD items,
- data imbalance depicted in Fig. 2,
- zoomed-out FOD images,
- images captured under two weather conditions: wet/dry,
- images captured under three lighting conditions: bright, dim, dark
- high intra-class variance, and
- low inter-class variance.

In this paper, FOD-A dataset with image resolution 300×300 has been analyzed as a small object detection dataset, where a small object [14–16] is mentioned in Section 1. Figure 3 depicts bounding box area in pixels with respect to the number of images in FOD-A. It can be observed that the FOD objects in 30,406 images occupy less than 20% area in its image, hence more than 85% images are treated as the small objects. It can also be observed that FOD objects in 19,820 images (more than 55% of FOD-A dataset) occupy area less than even 5% of image. MS COCO criteria classifies objects as small (area less than 32^2), medium ($32^2 < \text{area} < 96^2$) and large ($\text{area} > 96^2$), by measuring area with respect to pixels in segmentation mask. Since segmentation masks are not provided for FOD-A dataset, it is found that almost 75% dataset comprises of small (30%) and medium (45%) objects whereas large objects make only 25% of FOD-A dataset according to bounding box area (Fig. 3).

Figure 2 illustrates distribution of images across the 31 classes highlighting the class imbalance in FOD-A dataset. Multiple factors including varying camera viewpoint, illumination conditions and object poses make even the same object appear significantly different, which is termed as intra-class variance. High intra-class variance degrades performance of object detection algorithms [108]. Figure 4 visualizes the inter-class and intra-class variances in FOD-A dataset. Out of 31 classes of FOD-A dataset, images from five classes namely bolt, nut, screw, bolt-nut and bolt-washer have the highest intra-class similarity.

Choosing FOD-A dataset will not only provide the benchmark for the FOD detection problem but also evaluate the compared object detection algorithms as discussed in Section 4 for a real-world application.

To evaluate the selected algorithms specifically for small FODs, a new test dataset constituting manually picked images from the FOD test dataset is compiled and termed as small-FOD test set. Small-FOD test set has

- 4825 zoomed out test images from 31 FOD categories,
- objects in all images occupy less than 15% image area, and
- objects in 99% images occupy less than 5% image area.

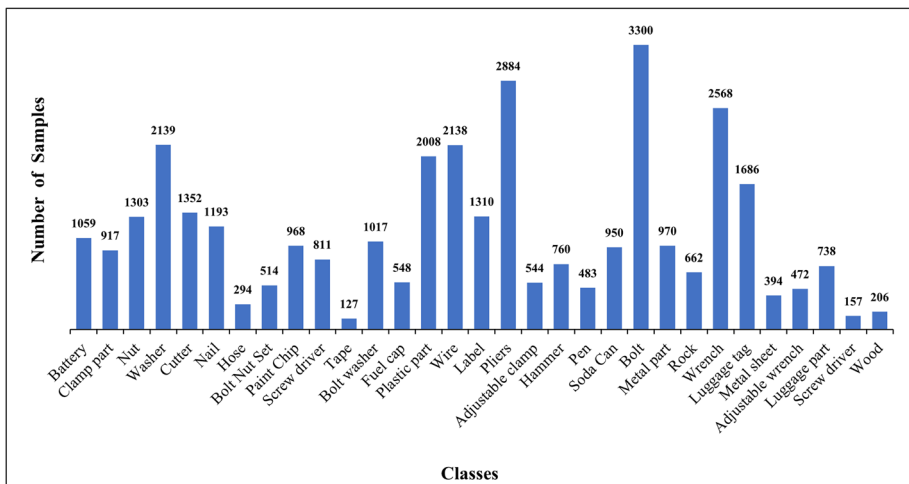


Fig. 2 The distribution of images across the 31 classes in the FOD-A dataset highlights the data imbalance

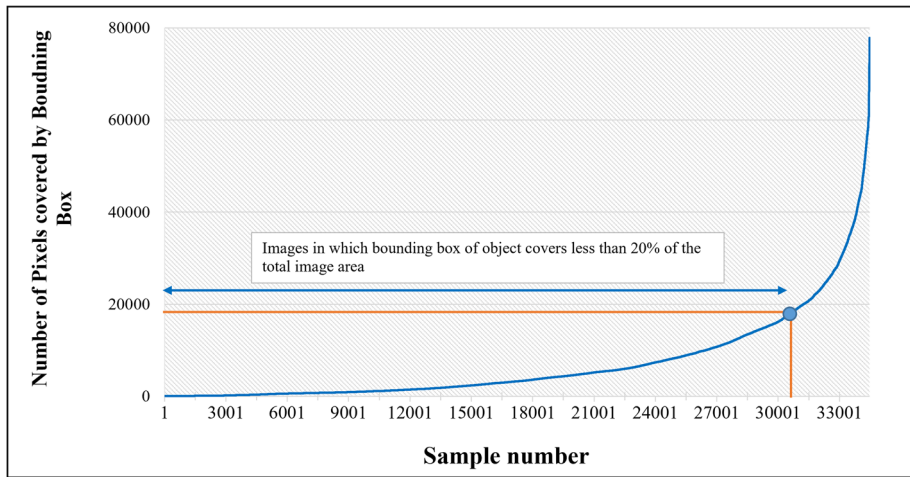
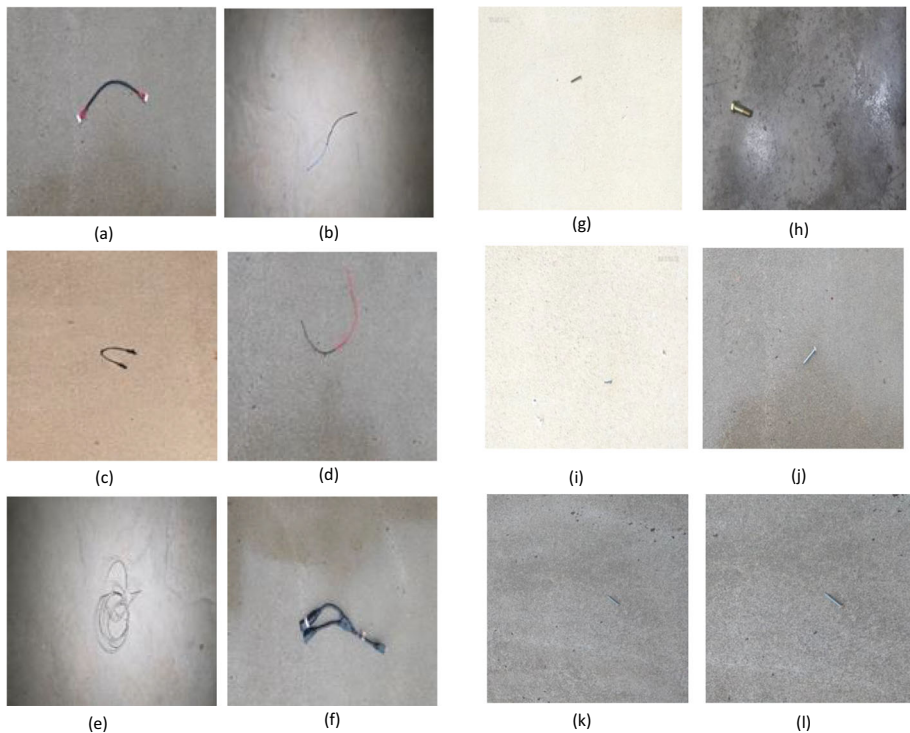


Fig. 3 Bounding box of more than 85% images in FOD-A small dataset occupy less than 20% of total area of the image



Images (a,b,c,d,e,f,) from FOD-A belong to same class: **Wire** with different backgrounds and illumination conditions

Images (g,h) belong to class: **Bolt**

Images (i,j) belong to class: **Nail**

Images (k,l) belong to class: **Screw**

Fig. 4 Background, illumination, inter-class and intra-class variations in FOD-A dataset



Fig. 5 Sample images from small FOD testset

Figure 5 illustrates samples of images taken from Small FOD test set. Small-FOD test set has also been evaluated for selected object detection algorithms and results are termed as APs(average precision small) in this paper.

4 Experiments for performance analysis of YOLOv5, Scaled YOLOv4, SSD, CenterNet and YOLOv8 for FOD-A dataset

The FOD-A dataset was subject to experimental evaluation using three anchor-based object detectors namely YOLOv5 (YOLOv5m and YOLOv5l) [39], Scaled YOLOv4 (P5 and P6) [37, 38], SSD [18], as well as two anchor-free object detectors, CenterNet [40, 41], and YOLOv8m [42]). High accuracy and fast inference are key aspects to ensure swift removal of FOD from runway surface, hence this assessment is specifically focused on detection accuracy and inference speed.

YOLOv4 is a state-of-the-art object detector with CSPDarknet53 CNN back bone. The CSP block in YOLOv4 maintains the fine grained features while SPP (Spatial Pyramid Pooling) block added between the backbone CSPDarkNet53 and modified path aggregator network (PANet) increases the receptive field. Mosaic data augmentation improves focus of network which results in better detection of small objects whereas Self-Adversarial Training (SAT) assists the network in learning new features [61]. YOLOv4 improved accuracy than previous YOLO models but with an increase in network size and computational cost. Scaled YOLO v4 [37] introduced a synergistic compound scaling method to reduce computational cost of YOLOv4. Scaled YOLOv4 CSP-ized YOLOv4 for general GPU by using original Darknet residual layer instead of first CSP stage in backbone and reduces computation by 40

percent. For evaluating detection performance for FOD-A dataset, scaled YOLOv4 large P5 and P6 models are selected.

YOLOv5 [39] was released shortly after YOLOv4. It is one of the latest YOLO object detection models using the CSPDarknet53 backbone. In YOLOv5, backbone extracts feature maps of different sizes [66], neck performs feature fusion, prevents loss of small object information [65] and aggregates maximum information extracted by backbone. YOLOv5 uses Path Aggregation Network (PANet) to transfer low-level features to high-level features using the bottom-up path based on FPN [109]. PANet improves the localization in lower layers, which enhances the localization accuracy of objects. Head in YOLOv5 is the same as previous models, which helps in detection of small to large objects by multiscale prediction. Mosaic, copy-paste, random-affine and mix-up augmentations are incorporated in YOLOv5 v6.1 [39, 110] which are also useful in small object detection. YOLOv5 models train faster and are smaller in size and hence are more feasible for real world deployments [65]. Four different models of YOLOv5 are introduced namely YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x, which refer to small, medium, large and extra large. These models just differ in scales due to application of different multipliers, which changes the size and complexity of models while maintaining the overall structure. YOLOv5m and YOLOv5l have been evaluated for FOD-A dataset in this work.

YOLO models are anchor-based methods which generate a large number of pre-set fixed bounding boxes on the convolution features for object detection. The anchor-free methods use keypoints instead of pre-set boxes for object detection. CenterNet proposed by [40] is one of the latest one-stage anchor-free object detector. CenterNet predicts key points (object-centers) and their attributes such as size, orientation, and offset to detect objects. It eliminates the requirement of anchors as well as computationally expensive post-processing techniques such as non-maximum suppression. CenterNet feeds input image in a fully convolutional network to generate heatmap. The object centers correspond to peak in heatmap and CenterNet uses the predicted center for calculating bounding box offsets. Zhou et al. [40] modify DLA-34 by adding deformable convolutions (DCN) proposed by [111] and evaluate CenterNet for four backbone architectures namely ResNet-18, ResNet-101, modified DLA-34 and Hourglass-104.

DLA refines shallow features by propagating through different aggregation layers. Unlike standard convolution where receptive field and sampling location are fixed, the receptive field and sampling location of DCN are adaptive as per the object shape and scale enhancing CNN's capability of modeling geometric transformations of objects [111, 112]. Since FOD-A has objects in varying shapes and orientations, we have selected CenterNet with modified Deep layer aggregation (DLA-34) [113] as backbone without flip augmentation for experimentation.

YOLOv8 model is a state-of-the-art object detector, introduced in 2023 by the developers of YOLOv5 [42]. The YOLOv8 model adopts an anchor-free approach, which implies that instead of using predefined anchors or bounding boxes, YOLOv8 partitions the input image into a grid of cells and each cell is expected to detect the object(s) contained within it. Within every cell, YOLOv8 makes predictions for objectness scores, probabilities of classes, and adjustments in terms of geometry to approximate the object's bounding box. This anchor-free methodology results in smaller number of predicted boxes, which in turn expedites post-inference filtering Non-Maximum Suppression (NMS) process.

YOLOv5m, YOLOv5l, scaled YOLOv4 P5, scaled YOLO v4 P6, SSD, CenterNet and YOLOv8 algorithms have been trained for 70 epochs for FOD-A dataset. 75% images from FOD-A dataset have been taken for training whereas the remaining 25% have been used for testing according to the train/ test split of FOD-A dataset [32, 114] to compare performance with [32]. All algorithms have been trained from scratch without using any pre-trained weights

Table 3 Hyperparameters for evaluated algorithms

S No	Algorithm	Learning rate	Optimizer	Batch size
1	SSD	1e-3	SGD	8
2	CenterNet	1.25e-4	ADAM	8
3	Scaled YOLOv4 P5 and Scaled YOLOv4 P6	1e-2	SGD	8
4	YOLOv5m and YOLOv5l	1e-2	SGD	8
5	YOLOv8m and Improved YOLOv8m	1e-3	AdamW	8

or transfer learning. The input image resolution is set to 416×416 for all algorithms whereas hyperparameters are depicted in Table 3. Frames per second (FPS) depicted in Table 7 are calculated using inference (netforwarding + decoding + non-maximum suppression (NMS)) time for all algorithms. In case of CenterNet, NMS is not applicable as it is an anchor-less detector. The CPU used is Intel I7 7700 K, GPU is NVIDIA GeForce GTX 1070, and the Operating System is 64-bit Windows 10 x64-based processor.

5 Improved YOLOv8 for FOD-A dataset

YOLOv8 is an anchor-free object detection model which implies that the model does not rely on predefined anchor boxes to detect objects; instead, it dynamically generates bounding boxes during the detection process offering more flexibility and accuracy in detection results. To improve YOLOv8 model for FOD detection task, we introduce a specialized small-object detection layer in the head of original YOLOv8 architecture. The original YOLOv8 network employs three feature maps with sizes 80×80 , 40×40 , and 20×20 cells for detection of small, medium and large objects respectively. To introduce a specialized prediction layer tailored for small targets, an up-sampling operation is performed on 80×80 scale feature

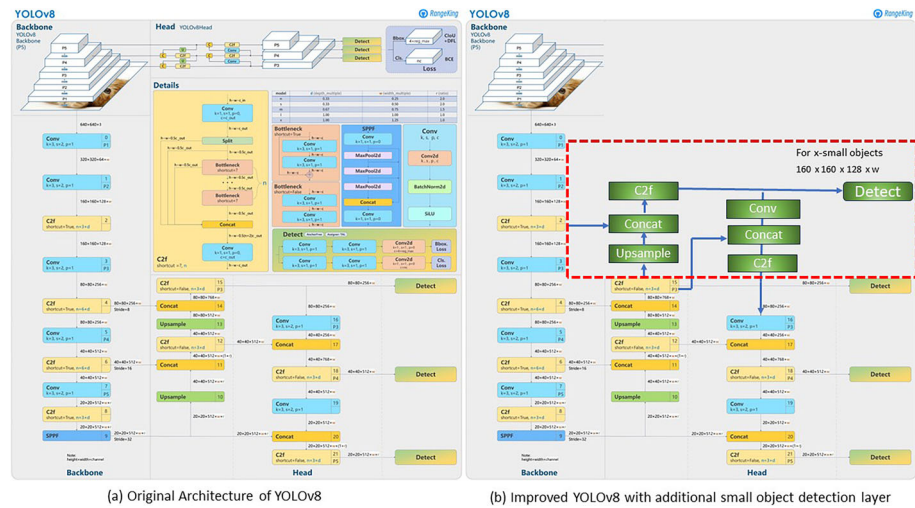


Fig. 6 YOLOv8 architecture: (a) Original architecture [115], (b) Improved YOLO v8 with additional small object detection layer

map from FPN module to get 160×160 feature map which is subsequently fused with the 160×160 feature map from the shallow layer (second layer) of the backbone module containing higher amount of inherent image information. Following this, the fused feature map is directly input to the prediction module. Figure 6 depicts the addition of specialized small object detection layer in the original architecture of YOLOv8. The specialized small object detection layer improves model output granularity, which in turn enhances its ability to make precise predictions on a smaller scale. Additionally, because smaller objects occupy a larger portion of a 160×160 grid compared to a smaller grid, the model allocates increased computational resources to these smaller objects, potentially leading to improved detection accuracy. This study also leverages transfer learning and uses pre-trained weights on COCO dataset to start training the improved YOLOv8 model.

For experimental purposes, the Improved YOLOv8 model is trained for 70 epochs using the same experimental settings as other evaluated algorithms, employing the hyper parameters outlined in Table 3. We then proceeded to compare the results with the assessed anchor-based and anchor-free techniques.

6 Results and discussion

Evaluation parameters for object detection algorithms generally include mean average precision (mAP) and precision-recall (PR) curve. Mean Average Precision (mAP) is the mean AP over all IoU thresholds or all classes. mAP at IOU (0.5 – 0.95) refers to mAP averaged for $\text{IOU} \in [0.5 : 0.05 : 0.95]$. Performance for selected object detection algorithms is compared with respect to mAP@ (0.5 – 0.95) and mAP@0.5 in this work. Table 4 gives comparative analysis of YOLOv5m, YOLOv5l, scaled YOLOv4-P6, scaled YOLOv4-P5, CenterNet, YOLOv8m and the proposed model improved YOLOv8m for FOD-A dataset. The empirical results for small FOD-A test set (manually selected small-FOD test set comprising of

Table 4 Results for evaluation on FOD-A dataset

S No	Algorithm	mAP @0.5	mAP @0.5 - 0.95
Comparison with Other Results for FOD-A Dataset			
1	SSD [32]	—	79.6
2	YOLOv3 [32]	—	69.7
3	Outer ViT [35]	—	82.7
4	YOLOv4-csp w/Augmentation [34]	—	90.12
Results for Evaluated Algorithms in this Work			
5	CenterNet	99.6	89.8
6	SSD	99.6	84.6
7	Scaled YOLOv4-P6	99.1	87.6
8	Scaled YOLOv4-P5	99.1	86.7
9	YOLOv5m	99.3	91.1
10	YOLOv5l	99.3	91.2
11	YOLOv8m [42]	99.4	93.2
12	Improved YOLOv8	99.4	93.8

Table 5 Results of evaluated algorithms for small FOD testset

S No	Results for Small FOD Testset	
	Algorithm	APs for mAP (0.5-0.95)
1	CenterNet	85.9
2	SSD	81.4
3	Scaled YOLOv4-P6	83.1
4	Scaled YOLOv4-P5	82.5
5	YOLOv5m	87.3
6	YOLOv5l	87.9
7	YOLOv8	90
8	Improved YOLOv8	91.8

all small objects) are depicted in as APs in Table 5. Upon analysis, Average Precision for small objects (APs) is found to be lower in values than Average precision (AP) for evaluated anchor-based and anchor-free approaches, confirming that performance of SOTA algorithms degrades for small object detection tasks.

6.1 Accuracy comparison

Precision-recall curve (PR curve) determines detection accuracy of the model. Large area under PR curve indicates high AP for the particular class. We compare accuracy of evaluated anchor based and anchor free models as well as the proposed approach Improved YOLOv8 with other works on FOD-A [32, 34] and [35] in Table 4. From empirical evaluation, we conclude that the Improved YOLOv8 model effectively increases the detection accuracy for both FOD-A dataset as well as small-FOD-A dataset with a minimal increase in parameter count. In the Improved YOLOv8 model, the addition of a dedicated shallow detection head in YOLOv8 architecture focuses the network towards small objects, hence improving their detection accuracy. Additionally, it is deduced from the experimentation that anchor free object detector CenterNet with DLA-34 backbone performs better in respect of AP and APs than scaled YOLOv4-P5, scaled YOLOv4-P6 and SSD as per the set hyperparameters. In anchor-based approaches, YOLOv5l has better performance than SSD, scaled YOLOv4 and CenterNet. In YOLOv5 models, PANet plays its part in retaining small-object information as it prevents information from being lost by adopting a feature pyramid network composed of many bottom up and top down layers. This assists in the propagation of low level features which in turn improves localization accuracy in the model [116].

The table presented in Table 6 illustrates the class-wise Average Precision (AP) scores for the FOD-A test set, serving as an indicator of the models' classification accuracy under the specified hyperparameters. Additionally, we perform a performance comparison among the evaluated models for classes that exhibit objects with similar appearances. As demonstrated in Fig. 7, the proposed anchor-free model, Improved YOLOv8, outperforms all other models in detecting objects with lower inter-class variance. Furthermore, the anchor-free model, CenterNet, also exhibits superior accuracy compared to YOLOv5l for categories such as nails, bolt-nut, and screws, as illustrated in Fig. 7. It's worth noting that CenterNet displays only minimal differences in classification accuracy compared to YOLOv5l for the bolt-washer and bolt classes.

Table 6 Class wise results of evaluated algorithms for FOD-A dataset

SNo	Class	mAP (0.5-0.95) for Classes of FOD-A Dataset						
		CenterNet	Scaled YOLOv4 P5	Scaled YOLOv4 P6	YOLOv5m	YOLOv5l	YOLOv8m	Improved YOLOv8m
1	Battery	0.874	0.809	0.843	0.901	0.886	0.914	0.92
2	BoltWasher	0.839	0.818	0.821	0.861	0.869	0.883	0.888
3	Bolt	0.828	0.794	0.787	0.826	0.837	0.872	0.888
4	ClampPart	0.933	0.9	0.899	0.942	0.947	0.962	0.971
5	FuelCap	0.932	0.896	0.901	0.927	0.943	0.956	0.963
6	MetalPart	0.908	0.879	0.882	0.929	0.93	0.955	0.967
7	Nut	0.784	0.735	0.73	0.797	0.819	0.832	0.852
8	PlasticPart	0.966	0.952	0.966	0.982	0.982	0.989	0.989

Table 6 continued

SNo	Class	mAP (0.5-0.95) for Classes of FOD-A Dataset						
		CenterNet	Scaled YOLOv4 P5	Scaled YOLOv4 P6	YOLOv5m	YOLOv5l	YOLOv8m	Improved YOLOv8m
9	Rock	0.898	0.857	0.866	0.922	0.913	0.936	0.942
10	Washer	0.759	0.73	0.733	0.786	0.8	0.808	0.83
11	Wire	0.952	0.939	0.952	0.973	0.967	0.988	0.988
12	Wrench	0.784	0.753	0.763	0.8	0.799	0.844	0.852
13	Cutter	0.955	0.948	0.951	0.973	0.973	0.98	0.979
14	Label	0.901	0.89	0.896	0.925	0.925	0.943	0.95
15	LuggageTag	0.885	0.843	0.855	0.899	0.879	0.928	0.929
16	Nail	0.706	0.639	0.641	0.693	0.703	0.721	0.743
17	Pliers	0.92	0.911	0.91	0.942	0.944	0.958	0.96
18	MetalSheet	0.997	0.978	0.991	0.993	0.995	0.995	0.995
19	Hose	0.947	0.941	0.944	0.982	0.966	0.987	0.986
20	AdjustableClamp	0.912	0.882	0.895	0.906	0.904	0.944	0.952
21	AdjustableWrench	0.898	0.876	0.897	0.93	0.933	0.95	0.95
22	BoltNutSet	0.728	0.617	0.631	0.67	0.681	0.733	0.753
23	Hammer	0.954	0.944	0.946	0.974	0.973	0.98	0.98
24	LuggagePart	0.952	0.937	0.961	0.981	0.973	0.989	0.98
25	PaintChip	0.984	0.982	0.973	0.988	0.988	0.992	0.992
26	Pen	0.896	0.842	0.888	0.915	0.913	0.96	0.969
27	Screw	0.895	0.816	0.801	0.897	0.891	0.927	0.945
28	Screwdriver	0.948	0.93	0.931	0.972	0.976	0.987	0.986
29	SodaCan	0.988	0.958	0.971	0.991	0.99	0.992	0.993
30	Wood	0.965	0.95	0.965	0.97	0.98	0.995	0.99
31	Tape	0.933	0.929	0.958	0.989	0.995	0.993	0.993

These results highlight better performance of both anchor-free object detectors, CenterNet and Improved YOLOv8, underscoring the potential of anchor-free approaches in effectively detecting objects that share similar physical appearances

6.2 Inference time comparison

For a safe airfield, prompt detection and removal of FOD is essential. Table 7 depicts that YOLOv5m model takes 24.5 hrs in training which is much lower in value as compared to YOLOv5l and infers at 95 FPS making it the fastest among the evaluated detectors. CenterNet with DLA-34 backbone despite having better AP and APs compared to scaled YOLOv4 (P5 and P6), infers at speed of 27 FPS and takes 50 hrs to train for 70 epochs. FPS vs mAPs (mAP for small-FOD test set) is plotted in Fig. 8 to compare speed/ accuracy of evaluated algorithms for small object detection. YOLOv5m is found comparable in terms of APs with YOLOv5l with an advantage of 36 FPS in inference time. Figure 8 depicts that the proposed Improved YOLOv8 model infers at 90 FPS, with better detection accuracy than YOLOv8 for small FODs, as reflected in Table 5, hence making it the most suitable algorithm for real time detection of FODs on runway.

6.3 Performance evaluation for out-of-distribution data

To assess the generalization of anchor based and anchor free models trained on the FOD-A dataset, we conducted practical testing using images captured in real runway environment. These images were acquired using a Machine Vision Camera GTX 2750 with a 6.1MP resolution, simulating various conditions, including bright sunlight, low-light settings, and scenarios where multiple objects appeared in a single image. We captured images having three common categories with FOD-A: wrenches, metal parts, and bolts. The images were

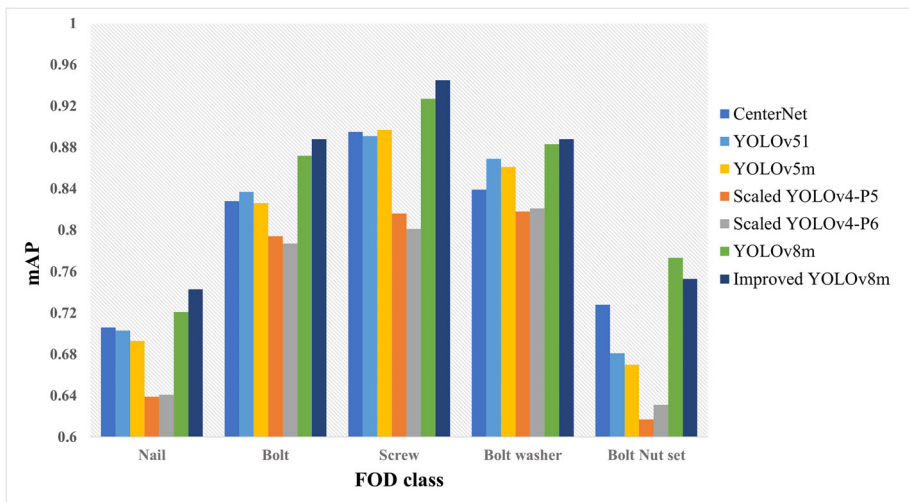


Fig. 7 mAP comparison for classes with low inter-class variance

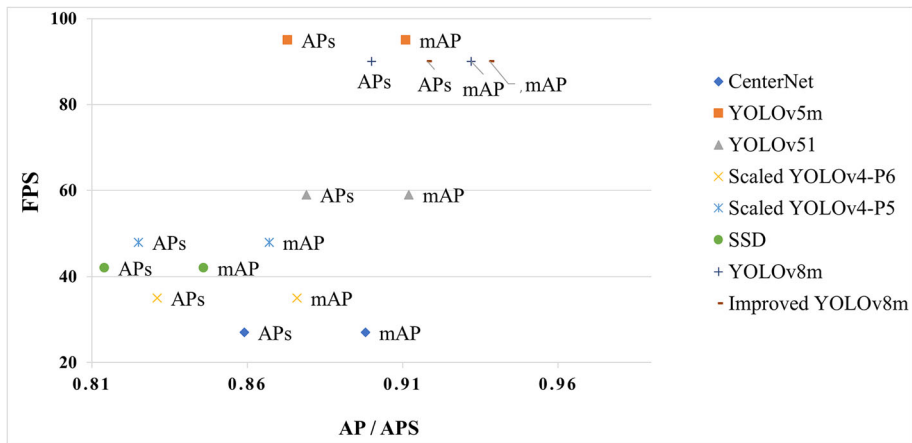


Fig. 8 mAP/APs vs FPS for evaluated algorithms

inferred from trained models and a qualitative comparison was conducted, as illustrated in Fig. 9. Notably, certain limitations were observed with the generalization of YOLOv5 model, such as its inability to detect wrenches in bright sunlight and bolts positioned on yellow lines on the runway. Among the evaluated models, CenterNet exhibited superior generalization ability when compared to anchor-based models like YOLOv5m and Scaled YOLOv4P6. Furthermore, the enhanced YOLOv8 demonstrated performance on par with CenterNet when detecting out-of-distribution images in various lighting conditions. This observation provides evidence of the superior generalization capabilities of anchor-free object detectors compared to their anchor-based counterparts.

Despite the improved performance of anchor-free detectors, all the models we assessed experienced a decrease in performance when dealing with out-of-distribution images. This underscores the necessity for ongoing efforts to enhance model generalization in such scenarios. Additionally, we identified certain limitations in the FOD-A dataset, which can be attributed to limiting generalization ability. Although the FOD-A dataset provides a substantial number of images encompassing diverse lighting conditions and wet/dry scenarios, it lacks representation of realistic runway backgrounds, including features such as cracks and taxiway lines. Furthermore, each training image within the dataset features only a single FOD object, neglecting scenarios in which multiple FODs from various categories might

Table 7 Training and inference time for evaluated algorithms for FOD-A dataset

S No	Method	Training Time (hrs)	Inference (FPS)
1	CenterNet	50	27
2	SSD	22.5	42
3	Scaled YOLOv4-P6	41.6	35
4	Scaled YOLOv4-P5	34	48
5	YOLOv5m	24.5	95
6	YOLOv5l	41	59
7	YOLOv8m	30	90
8	Improved YOLOv8m	34	90

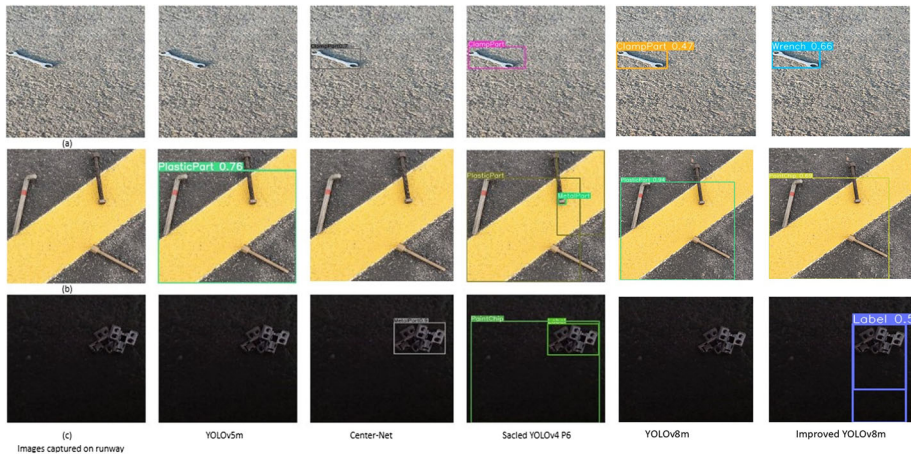


Fig. 9 (a) Wrench under sunlight (b) Multiple bolts on runway way yellow lines (c) Metal parts in dark. From left: Results of YOLOv5m, Center-Net, Scaled YOLOv4 P6, YOLOv8m and Improved YOLOv8m

appear within a single frame. Therefore, exploring data-driven approaches may present a promising avenue for enhancing the FOD-A dataset, ultimately leading to improved model generalization.

7 Conclusion

In this research endeavor, a comprehensive evaluation of state-of-the-art anchor-based object detectors, namely YOLOv5, SSD, Scaled YOLOv4 and anchor-free object detectors Center-Net and YOLOv8, is undertaken for the multiclass FOD-A dataset. YOLOv8m emerged as the leading performer, striking an optimal balance between speed and accuracy for the FOD-A dataset. To enhance the mean average precision (mAP) specifically for small FODs, the architecture of YOLOv8 is further refined by introducing a dedicated small object detection layer. This refinement led to the development of an “Improved YOLOv8” model, surpassing YOLOv8 by a factor of 1.02 in small FOD detection. Notably, our “Improved YOLOv8” model outperformed all anchor-based and anchor-free detectors evaluated in this study, as well as those from prior research involving the FOD-A dataset. While anchor-free object detectors such as CenterNet and Improved YOLOv8 demonstrated superior detection accuracy, it was observed that these models exhibited limited generalization capabilities when handling out-of-distribution images. Given the diverse potential forms of foreign object debris (FOD) on airport runways, it is neither economically feasible nor practically attainable to create a comprehensive dataset that encompasses every possible instance of FOD. Such an endeavor would incur substantial costs and might still not adequately anticipate the emergence of new and novel types of FOD. Therefore, in the realm of future research endeavors, emphasis should be directed towards improving generalization capability by exploration and implementation of generative deep learning models. Enrichment of FOD datasets by incorporating realistic synthetic images [117, 118] illustrating an array of runway scenarios, while considering factors such as pavement conditions, the positioning of runway lights, and vari-

ations in weather and lighting, stands as a promising avenue to improve both the detection accuracy and the generalization capability of deep learning FOD detection models.

Data Availability The FOD-A dataset analysed during the current study is available in the [FOD-UNOmaha] repository, [<https://github.com/FOD-UNOmaha/FOD-data>]

Declarations

Conflicts of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Mishra R, Srivastav D, Srinivasan K, Nandi V, Bhat RR (2015) Impact of foreign object damage on an aero gas turbine engine. *J Fail Anal Prev* 15:25–32
2. Chauhan T, Goyal C, Kumari D, Thakur AK (2020) A review on foreign object debris/damage (fod) and its effects on aviation industry. *Mater Today: Proc* 33:4336–4339
3. Cramoisi G (2010) Air Crash Investigations: The End of the Concorde Era, the Crash of Air France Flight 4590 Lulu. com
4. Öztürk S, Kuzucuoglu AE (2016) A multi-robot coordination approach for autonomous runway foreign object debris (fod) clearance. *Robot Auton Syst* 75:244–259
5. Rafiq HA, Manarvi IA, Iqbal A (2013) Identification of major fod contributors in aviation industry. In: Business strategies and approaches for effective engineering management, pp 237–250, IGI Global
6. Zhong J, Gou X, Shu Q, Liu X, Zeng Q (2021) A fod detection approach on millimeter-wave radar sensors based on optimal vmd and svdd. *Sensors* 21(3):997
7. Yonemoto N, Kohmura A, FUTATSUMORI S, Morioka K, Kanada N (2018) Two dimensional radar imaging algorithm of bistatic millimeter wave radar for fod detection on runways. In 2018 International topical meeting on microwave photonics (MWP), pp 1–4, IEEE
8. Hong J-B, Kang M-S, Kim Y-S, Kim M-S, Hong G-Y (2018) Experiment on automatic detection of airport debris (fod) using eo/ir cameras and radar. *J Adv Navig Technol* 22(6):522–529
9. Cao X, Gong G, Liu M, Qi J (2016) Foreign object debris detection on airfield pavement using region based convolution neural network. In: 2016 International conference on digital image computing: Techniques and applications (DICTA) pp 1–6, IEEE
10. Cao X, Gu Y, Bai X (2017) Detecting of foreign object debris on airfield pavement using convolution neural network. In: LIDAR imaging detection and target recognition 2017, vol 10605, pp 840–846, SPIE
11. Cao X, Wang P, Meng C, Bai X, Gong G, Liu M, Qi J (2018) Region based cnn for foreign object debris detection on airfield pavement. *Sensors* 18(3):737
12. Ni P, Miao C, Tang H, Jiang M, Wu W (2020) Small foreign object debris detection for millimeter-wave radar based on power spectrum features. *Sensors* 20(8):2316
13. Yuan Z-D, Li J-Q, Qiu Z-N, Zhang Y (2020) Research on fod detection system of airport runway based on artificial intelligence. In: Journal of Physics: Conference Series, vol 1635, p 012065, IOP Publishing
14. Nguyen N-D, Do T, Ngo TD, Le D-D (2020) An evaluation of deep learning methods for small object detection. *J Electr Comput Eng* 2020:1–18
15. Tong K, Wu Y, Zhou F (2020) Recent advances in small object detection based on deep learning: A review. *Image Vis Comput* 97:103910
16. Zhu Z, Liang D, Zhang S, Huang X, Li B, Hu S (2016) Traffic-sign detection and classification in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2110–2118
17. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
18. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC (2016) Ssd: Single shot multibox detector. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, pp 21–37, Springer
19. Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448

20. Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv in Neural Inf Process Syst*, 28
21. Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 2117–2125
22. Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL, Microsoft coco: Common objects in context. In: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*, pp 740–755, Springer
23. Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A (2010) The pascal visual object classes (voc) challenge. *Int J Comput Vis* 88:303–338
24. Deng C, Wang M, Liu L, Liu Y, Jiang Y (2021) Extended feature pyramid network for small object detection. *IEEE Trans Multimed* 24:1968–1979
25. Kisantal M, Wojna Z, Murawski J, Naruniec J, Cho K (2019) Augmentation for small object detection [arXiv:1902.07296](https://arxiv.org/abs/1902.07296)
26. Liu Y, Sun P, Wergeles N, Shang Y (2021) A survey and performance evaluation of deep learning methods for small object detection. *Expert Syst Appl* 172:114602
27. Liu K, Mattyus G (2015) Fast multiclass vehicle detection on aerial images. *IEEE Geosci Remote Sens* 12(9):1938–1942
28. Geiger A, Lenz P, Urtasun R (2011) Are we ready for autonomous driving? the kitti vision benchmark suite. In: *2012 IEEE conference on computer vision and pattern recognition*, pp 3354–3361, IEEE
29. Alahi A, Goel K, Ramanathan V, Robicquet A, Fei-Fei L, Savarese S (2016) Social lstm: Human trajectory prediction in crowded spaces. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 961–971
30. Zhang S, Benenson R, Schiele B (2017) Citypersons: A diverse dataset for pedestrian detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 3213–3221
31. Liu Y, Li Y, Liu J, Peng X, Zhou Y, Murphey YL (2018) Fod detection using densenet with focal loss of object samples for airport runway. In: *2018 IEEE symposium series on computational intelligence (SSCI)*, pp 547–554, IEEE
32. Munyer T, Huang P-C, Huang C, Zhong X (2021) Fod-a: A dataset for foreign object debris in airports. [arXiv:2110.03072](https://arxiv.org/abs/2110.03072)
33. Papadopoulos E, Gonzalez F (2021) Uav and ai application for runway foreign object debris (fod) detection. In: *2021 IEEE aerospace conference (50100)*, pp 1–8, IEEE
34. Noroozi M, Shah A (2023) Towards optimal foreign object debris detection in an airport environment. *Expert Syst Appl* 213:118829
35. Munyer T, Brinkman D, Zhong X, Huang C, Konstantzos I (2022) Foreign object debris detection for airport pavement images based on self-supervised localization and vision transformer. [arXiv:2210.16901](https://arxiv.org/abs/2210.16901)
36. Cai Z, Vasconcelos N (2018) Cascade r-cnn: Delving into high quality object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 6154–6162
37. Wang C-Y, Bochkovskiy A, Liao H-YM (2021) Scaled-yolov4: Scaling cross stage partial network. In: *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition*, pp 13029–13038
38. Wang C-Y, Bochkovskiy A, Liao H-YM (2020) Scaled-yolov4: Scaling cross stage partial network. pp. 2011
39. Glenn-Jocher (2023) ultralytics/yolov5.” GitHub Repository, Accessed 2023. <https://github.com/ultralytics/yolov5>
40. Zhou X, Wang D, Krähenbühl P (2019) Objects as points. [arXiv:1904.07850](https://arxiv.org/abs/1904.07850)
41. Zhou X (2023) Centrenet. GitHub Repository, Accessed 2023. <https://github.com/xingyizhou/CenterNet>
42. Jocher G, Chaurasia A, Qiu J (2023) YOLO by Ultralytics
43. Meta (2022) Real time object detection on COCO. GitHub Repository, Accessed June 2022. <https://paperswithcode.com/sota/real-time-object-detection-on-coco?metric=FPS>
44. Meta AI (2022) APs object detection on COCO GitHub Repository, Accessed June 2022. https://paperswithcode.com/sota/object-detection-on-coco?metric=APS&tag_filter=14%2C15%2C13%2C17
45. Qi G, Zhang Y, Wang K, Mazur N, Liu Y, Malaviya D (2022) Small object detection method based on adaptive spatial parallel convolution and fast multi-scale fusion. *Remote Sensing* 14(2):420
46. Harris C, Stephens M, et al (1988) A combined corner and edge detector. In *Alvey vision conference*, vol 15, pp 10–5244, Citeseer
47. Lin Z, Davis LS (2010) Shape-based human detection and segmentation via hierarchical part-template matching. *IEEE Trans Pattern Anal Mach Intell* 32(4):604–618
48. Zhang S, Chi C, Yao Y, Lei Z, Li SZ (2020) Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 9759–9768

49. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708
50. Lin T-Y, Goyal P, Girshick R, He K, Dollár P (2017) Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision, pp 2980–2988
51. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 2961–2969
52. Duan K, Xie L, Qi H, Bai S, Huang Q, Tian Q (2020) Corner proposal network for anchor-free, two-stage object detection. In European Conference on Computer Vision, pp 399–416, Springer
53. Wang T, Zhu X, Pang J, Lin D (2021) Fcos3d: Fully convolutional one-stage monocular 3d object detection. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 913–922
54. Zhou X, Zhuo J, Krahenbuhl P (2019) Bottom-up object detection by grouping extreme and center points. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 850–859
55. Zhu C, He Y, Savvides M (2019) Feature selective anchor-free module for single-shot object detection In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 840–849
56. Wang J, Chen K, Yang S, Loy CC, Lin D (2019) Region proposal by guided anchoring. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 2965–2974
57. Pham M-T, Courtrai L, Friguet C, Lefèvre S, Baussard A (2020) Yolo-fine: One-stage detector of small objects under various backgrounds in remote sensing images. *Remote Sensing* 12(15):2501
58. He X, Cheng R, Zheng Z, Wang Z (2021) Small object detection in traffic scenes based on yolo-mxnet. *Sensors* 21(21):7422
59. Xianbao C, Guihua Q, Yu J, Zhaomin Z (2021) An improved small object detection method based on yolo v3. *Pattern Anal Appl* 24:1347–1355
60. Liu M, Wang X, Zhou A, Fu X, Ma Y, Piao C (2020) Uav-yolo: Small object detection on unmanned aerial vehicle perspective. *Sensors* 20(8):2238
61. Bochkovskiy A, Wang C-Y, Liao H-YM (2020) Yolov4: Optimal speed and accuracy of object detection. [arXiv:2004.10934](https://arxiv.org/abs/2004.10934)
62. Shi P, Jiang Q, Shi C, Xi J, Tao G, Zhang S, Zhang Z, Liu B, Gao X, Wu Q (2021) Oil well detection via large-scale and high-resolution remote sensing images based on improved yolo v4. *Remote Sensing* 13(16):3243
63. Yu Z, Shen Y, Shen C (2021) A real-time detection approach for bridge cracks based on yolov4-fpm. *Autom Constr* 122:103514
64. Zhan W, Sun C, Wang M, She J, Zhang Y, Zhang Z, Sun Y (2022) An improved yolov5 real-time detection method for small objects captured by uav. *Soft Comput* 26:361–373
65. Benjumea A, Teeti I, Cuzzolin F, Bradley A (2021) Yolo-z: Improving small object detection in yolov5 for autonomous vehicles. [arXiv:2112.11798](https://arxiv.org/abs/2112.11798)
66. Zhu L, Geng X, Li Z, Liu C (2021) Improving yolov5 with attention mechanism for detecting boulders from planetary images. *Remote Sensing* 13(18):3776
67. Talaat FM, ZainEldin H (2023) An improved fire detection approach based on yolo-v8 for smart cities. *Neural Comput Appl*, 1–16,
68. Aboah A, Wang B, Bagci U, Adu-Gyamfi Y (2023) Real-time multi-class helmet violation detection using few-shot data sampling technique and yolov8 In: Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition, pp 5349–5357,
69. Sun C, Ai Y, Wang S, Zhang W (2021) Mask-guided ssd for small-object detection. *Appl Intell* 51:3311–3322
70. Yundong L, Han D, Hongguang L, Zhang X, Zhang B, Zhifeng X (2020) Multi-block ssd based on small object detection for uav railway scene surveillance. *Chinese J Aeronaut* 33(6):1747–1755
71. Cao G, Xie X, Yang W, Liao Q, Shi G, Wu J (2018) Feature-fused ssd: Fast detection for small objects. In: Ninth International Conference on Graphic and Image Processing (ICGIP 2017), vol 10615, pp 381–388, SPIE
72. Lim J-S, Astrid M, Yoon H-J, Lee S-I (2021) Small object detection using context and attention In: 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), pp 181–186, IEEE
73. Cao C, Wang B, Zhang W, Zeng X, Yan X, Feng Z, Liu Y, Wu Z (2019) An improved faster r-cnn for small object detection. *IEEE Access* 7:106838–106846
74. Eggert C, Brehm S, Winschel A, Zecha D, Lienhart R (2017) A closer look: Small object detection in faster r-cnn. In: 2017 IEEE international conference on multimedia and expo (ICME), pp 421–426, IEEE
75. Liang Z, Shao J, Zhang D, Gao L (2018) Small object detection using deep feature pyramid networks. In: Advances in Multimedia Information Processing–PCM 2018: 19th Pacific-Rim Conference on Multimedia, Hefei, China, September 21–22, 2018, Proceedings, Part III 19, pp 554–564, Springer

76. Li H, Lin K, Bai J, Li A, Yu J (2019) Small object detection algorithm based on feature pyramid-enhanced fusion ssd. *Complexity* 2019:1–13
77. Gong Y, Yu X, Ding Y, Peng X, Zhao J, Han Z (2021) Effective fusion factor in fpn for tiny object detection. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp 1160–1168
78. Liu Z, Gao G, Sun L, Fang L (2020) Ipg-net: Image pyramid guidance network for small object detection. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp 1026–1027
79. Rabbi J, Ray N, Schubert M, Chowdhury S, Chao D (2020) Small-object detection in remote sensing images with end-to-end edge-enhanced gan and object detector network. *Remote Sensing* 12(9):1432
80. Li J, Liang X, Wei Y, Xu T, Feng J, Yan S (2017) Perceptual generative adversarial networks for small object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1222–1230
81. Bai Y, Zhang Y, Ding M, Ghanem B (2018) Sod-mtgan: Small object detection via multi-task generative adversarial network. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp 206–221
82. Noh J, Bae W, Lee W, Seo J, Kim G (2019) Better to follow, follow to be better: Towards precise supervision of feature super-resolution for small object detection. In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp 9725–9734
83. Ozge Unel F, Ozkalayci BO, Cigla C (2019) The power of tiling for small object detection. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp 0–0
84. He Z, Huang L, Zeng W, Zhang X, Jiang Y, Zou Q (2021) Elongated small object detection from remote sensing images using hierarchical scale-sensitive networks. *Remote Sensing* 13(16):3182
85. Li G, Xie H, Yan W, Chang Y, Qu X (2020) Detection of road objects with small appearance in images for autonomous driving in various traffic situations using a deep learning based approach. *IEEE Access* 8:211164–211172
86. Xu J, Ye Y, Liu Z, Meng F, Zhang D, Sun M (2021) Small object detection with improved centernet. In: *2021 IEEE 2nd international conference on information technology, big data and artificial intelligence (ICIBA)*, vol 2, pp 956–960, IEEE
87. Xiao-jing G, Xue-you Y, Zhi-jing Y (2013) Application of wavelet analysis in detecting runway foreign object debris. *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 11(4):759–766
88. Li Y, Xiao G (2011) A new fod recognition algorithm based on multi-source information fusion and experiment analysis. In: *International symposium on photoelectronic detection and imaging 2011: Advances in infrared imaging and applications*, vol 8193, pp 769–778, SPIE
89. Khan T, Alam M, Kadir K, Shahid Z, Mazliham M, Khan S, Miqdad M (2017) Foreign objects debris (fod) identification: A cost effective investigation of fod with less false alarm rate. In: *2017 IEEE 4th international conference on smart instrumentation, measurement and application (ICSIMA)*, pp 1–4, IEEE
90. Liang W, Zhou Z, Chen X, Sheng X, Ye X (2020) Research on airport runway fod detection algorithm based on texture segmentation. In: *2020 IEEE 4th information technology, networking, electronic and automation control conference (ITNEC)*, vol 1, pp 2103–2106, IEEE
91. Gao Q, Hong R, Chen Y, Lei J (2021) Research on foreign object debris detection in airport runway based on semantic segmentation. In: *The 2nd International Conference on Computing and Data Science*, pp 1–3
92. Ren M, Wan W, Yu Z, Zhao Y (2022) Bidirectional yolo: improved yolo for foreign object debris detection on airport runways. *J Electron Imaging* 31(6):063047–063047
93. Jing Y, Zheng H, Zheng W, Dong K (2022) A pixel-wise foreign object debris detection method based on multi-scale feature inpainting. *Aerospace* 9(9):480
94. Lindeberg T (2012) Scale invariant feature transform
95. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol 1, pp 886–893, IEEE
96. Pietikäinen M, Hadid A, Zhao G, Ahonen T (2011) *Computer vision using local binary patterns*, vol 40. Springer Science & Business Media
97. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 3431–3440
98. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: Convolutional architecture for fast feature embedding. In: *Proceedings of the 22nd ACM international conference on multimedia*, pp 675–678
99. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)

100. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. *Int J Comput Vis* 115:211–252
101. Uijlings JR, Van De Sande KE, Gevers T, Smeulders AW (2013) Selective search for object recognition. *Int J Comput Vis* 104:154–171
102. Yu F, Koltun V (2015) Multi-scale context aggregation by dilated convolutions. [arXiv:1511.07122](https://arxiv.org/abs/1511.07122)
103. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell* 37(9):1904–1916
104. Redmon J, Farhadi A (2018) Yolov3: An incremental improvement. [arXiv:1804.02767](https://arxiv.org/abs/1804.02767)
105. Copeland M, Soh J, Puca A, Manning M, Gollob D, Copeland M, Soh J, Puca A, Manning M, Gollob D (2015) Microsoft azure and cloud computing. Planning, Deploying, and Managing Your Data center in the Cloud, Microsoft Azure, pp 3–26
106. Administration FA (2023) Ac 150/5220-24 - foreign object debris detection equipment. Advisory Circular, Year. https://www.faa.gov/regulations_policies/advisory_circulars
107. Xu H, Han Z, Feng S, Zhou H, Fang Y (2018) Foreign object debris material recognition based on convolutional neural networks. *EURASIP J Image Vid Process* 2018:1–10
108. Zhang Z, Luo C, Wu H, Chen Y, Wang N, Song C (2022) From individual to whole: reducing intra-class variance by feature aggregation. *Int J Comput Vis* 130(3):800–819
109. Jing Y, Ren Y, Liu Y, Wang D, Yu L (2022) Automatic extraction of damaged houses by earthquake based on improved yolov5: a case study in yangbi. *Remote Sensing* 14(2):382
110. Glenn-Jocher (2023) ultralytics/yolov5: Issue 6998. GitHub Issue, Accessed 2023. <https://github.com/ultralytics/yolov5/issues/6998>
111. Dai J, Qi H, Xiong V, Li Y, Zhang G, Hu H, Wei Y (2017) Deformable convolutional networks. In: Proceedings of the IEEE international conference on computer vision, pp 764–773
112. Li D-Y, Wang G-F, Zhang Y, Wang S (2022) Coal gangue detection and recognition algorithm based on deformable convolution yolov3. *IET Image Process* 16(1):134–144
113. Yu F, Wang D, Shelhamer E, Darrell T (2018) Deep layer aggregation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2403–2412
114. FOD-UNomaha (2023) FOD-data. GitHub Repository, Accessed 2023. <https://github.com/FOD-UNomaha/FOD-data>
115. Glenn-Jocher (2023) ultralytics/yolov5: Issue 189. GitHub Issue, Accessed August 2023. <https://github.com/ultralytics/yolov5/issues/6998>
116. Roy AM, Bose R, Bhaduri J (2022) A fast accurate fine-grain object detection model based on yolov4 deep neural network. *Neural Comput Appl*, 1–27
117. Farooq J, Fatima S, Aafaq N (2023) Synthetic randomized image augmentation (SRIA) to address class imbalance problem. 2023 3rd International conference on computing and information technology (ICCIT), 308–313. <https://doi.org/10.1109/ICCIT58132.2023.10273972>
118. Farooq J, Aafaq N, Khan M, Saleem, Siddiqui MI (2023) Randomize to generalize: domain randomization for runway FOD detection. *ArXiv Preprint* [arXiv:2309.13264](https://arxiv.org/abs/2309.13264)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.