Array 23 (2024) 100361

Contents lists available at ScienceDirect

Array

journal homepage: www.sciencedirect.com/journal/array

Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review

Aditya Vardhan Reddy Katkuri [a], Hakka Madan [b], Narendra Khatri [b,*], Antar Shaddad Hamed Abdul-Qawy [c], K. Sridhar Patnaik [d]

[a] Department of Aeronautical and Automobile Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
[b] Department of Mechatronics, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
[c] Department of Mathematics and Computer Science, Faculty of Science, Abdulrahman Al-Sumait University, Zanzibar, 1933, Tanzania
[d] Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, 835215, Ranchi, Jharkhand, India

ARTICLE INFO

Keywords:
Autonomous UAV
Deep learning
Computer vision
Systematic literature review
UAV applications
You only look once (YOLO) framework

ABSTRACT

The increasing use of unmanned aerial vehicles (UAVs) in both military and civilian applications, such as infrastructure inspection, package delivery, and recreational activities, underscores the importance of enhancing their autonomous functionalities. Artificial intelligence (AI), particularly deep learning-based computer vision (DL-based CV), plays a crucial role in this enhancement. This paper aims to provide a systematic literature review (SLR) of Scopus-indexed research studies published from 2019 to 2024, focusing on DL-based CV approaches for autonomous UAV applications. By analyzing 173 studies, we categorize the research into four domains: sensing and inspection, landing, surveillance and tracking, and search and rescue. Our review reveals a significant increase in research utilizing computer vision for UAV applications, with over 39.5 % of studies employing the You Only Look Once (YOLO) framework. We discuss the key findings, including the dominant trends, challenges, and opportunities in the field, and highlight emerging technologies such as in-sensor computing. This review provides valuable insights into the current state and future directions of DL-based CV for autonomous UAVs, emphasizing its growing significance as legislative frameworks evolve to support these technologies.

1. Introduction

Owing to rapid development in technology and key changes to public policy, autonomous Unmanned Aerial Vehicles have presented as one of the most challenging and high-potential solutions in both the military and civil domains. UAVs, commonly known as drones, have applications ranging from remote sensing [1] and surveillance to cargo delivery and precision agriculture [2,3]. The term UAV represents a broad category of aerial vehicles, including fixed-wing, multirotor, and Vertical Take-off and Landing Aircraft (VTOL) aircraft, that are further segregated into different classes based on their range, size, payload capacity, and altitude [4].

UAVs have an inherent advantage over Unmanned Ground Vehicles (UGV) because of an added degree of freedom pertaining to the possibility of altering their altitude [5]. UAVs are known for their cost-effectiveness, fast mobility, and easy deployment. The recent development of VTOL has enhanced their use cases for stable hover over the target [6]. They are distinguished for their ability to fly at a range of speeds, operate in proximity to an object, carry required sensors, and fly indoors and outdoors. These functions allow them to replace humans in dangerous, inaccessible, complicated, or expensive environments [7].

Typically, UAVs are equipped with vision sensors, such as cameras, among other sensors, in order to collect information and perceive the environment, enabling it to navigate autonomously in that environment. Autonomous UAV Navigation refers to navigation and accomplishing tasks without human intervention. Autonomy in this area has been proven to improve efficiency and reduce risks and costs incurred due to human factors. The role of AI in increasing UAV autonomy is of great significance in this AI era, where many applications of such have been proposed in the past five years alone.

Numerous industries, including precision agriculture, search and rescue, environmental monitoring, and infrastructure inspection, have been transformed by employing technologies like UAVs. For instance, UAVs with cameras and sensors can be used in precision agriculture to track crop health, detect crop stress, and improve irrigation and fertilization [8,9]. UAVs can be employed in search and rescue missions to

* Corresponding author.
E-mail addresses: narkhatri@gmail.com, narendra.khatri@manipal.edu (N. Khatri).

https://doi.org/10.1016/j.array.2024.100361
Received 22 March 2024; Received in revised form 26 August 2024; Accepted 27 August 2024
Available online 29 August 2024

swiftly find missing people or inspect disaster zones [10]. UAVs can be used in environmental monitoring to gather information on wildlife populations, water quality, and air quality. UAVs can be used to inspect infrastructure and evaluate the state of buildings, roads, and bridges.

Various Computer Vision (CV) based autonomous navigation frameworks have been increasingly researched and reviewed in the past few years. Computer vision is a method to extract meaning from visual input; it is done in real-time on UAVs to perceive their operating environment. The AI revolution has added to this development, allowing for employing state-of-the-art CV approaches in domains never thought possible, such as achieving autonomy in aerial refueling [11].

In general, CV navigation models are trained using a transfer learning approach where the training is conducted in a virtual 3D environment that simulates the UAV's operating conditions, and then the learnings are deployed onto a physical machine. This method reduces the crash cost associated with real-world testing [12]. Furthermore, UAVs house limited recourses on-board to stay in the air, so they cannot sustain long-duration robust training and testing.

In the vision-based navigation approach, the information provided by cameras is used as input and processed through the trained CV pipeline. Finally, the CV pipeline outputs an action that determines the direction of UAV movement, the thrust of the propelling unit, etc., to accomplish the task. In this study, the CV applications in UAVs are organized into four categories, as shown in Fig. 1.

The scientific community reviewed several AI-based approaches for autonomous UAV applications, including computational intelligence-based path planning, DL in drone autonomy, DL in forestry, navigation using the RL approach, and civil infrastructure health monitoring (see Fig. 2). A few excellent complementary surveys are presented in Table 1.

This review aims to encompass research articles from Scopus-indexed journals that specifically examine the utilization of Deep Learning in Computer Vision methods for enhancing the autonomous capabilities of Unmanned Aerial Vehicles (UAVs). Moreover, analyzes the difficulties and potential advantages in this field. For this purpose, a systematic literature review approach (SLR) was utilized. To our knowledge, no thorough literature reviews of studies published after 2017 have been done on this subject. In detail, this review addresses the following research questions.

RQ1: *How many quality papers have been published on the DL-based CV approach for UAV autonomy since 2019?*
RQ2: *How can we systematically identify relationships among them, categorize them, and present them with an eye on reproducibility?*
RQ3: *What is the state-of-the-art in employing CV to enhance UAV autonomy for navigation?*
RQ4: *What are the key research tracks and their significant contribution to CV in autonomous UAV navigation?*
RQ5: *What are the open challenges and opportunities in this domain?*

The remainder of this paper is organized as follows: Section 2 Presents the systematic literature review process – RQ1, RQ2. Section 3 introduces computer vision and the state-of-the-art technology it can offer to UAV applications. – RQ 3. Section 4: Gives a detailed review of computer vision approaches aiding UAV autonomy –RQ3, RQ4. Section 5: Discusses the challenges and opportunities in this domain – RQ 5. Section 6: Provides the conclusion to this review.

2. Review process

This section describes the inclusion criteria, paper identification process, and threats to validity.

2.1. Inclusion criteria and identification of papers

The objective of the conducted SLR was to thoroughly identify and analyze studies related to deep learning-based computer vision methods for autonomous UAV navigation. The process commenced with a methodical examination of relevant literature carried out in July 2024, utilizing the Scopus database. The inclusion criteria stipulated that only journal publications from 2019 onwards were taken into account, in accordance with the emphasis on recent progress in the field.

At first, the search produced 731 articles, which then went through additional stages of filtering. A manual filtering process was utilized to eliminate articles that were not directly pertinent to the research topic, resulting in a total of 173 articles that were considered appropriate for further analysis. The retained articles were subsequently subjected to certainty filtering, which involved excluding articles with a certainty level below 24 %. Ultimately, this step verified the inclusion of all 173 articles that were initially kept. After selecting the articles, the next step was to categorize the chosen literature into specific review categories based on the application domains associated with autonomous UAV navigation. The identified categories were Sensing and Inspection, Landing, Surveillance and Tracking, and Search and Rescue. These categories included 44, 33, 44, and 52 articles, respectively.

The methodology provided additional information on how to create search strings that are customized for the specific keywords related to autonomous unmanned aerial vehicles (UAVs) and computer vision. The terms encompassed a variety of concepts, including unmanned aerial vehicles, vision, image processing, and artificial intelligence, among others. The search strings were formulated to efficiently limit the range of investigation while guaranteeing thorough inclusion of pertinent scholarly works. In addition, the methodology established specific criteria for article selection, including the time frame (from 2019 to July 2024), document type (journal articles), field of study (engineering and computer science), and language (English). These criteria were designed to provide clear parameters and maintain consistency throughout the review process.

The systematic approach utilized in this systematic literature review methodology enabled a thorough analysis of existing literature that specifically investigates deep learning-based computer vision techniques for autonomous UAV navigation. Through the implementation of

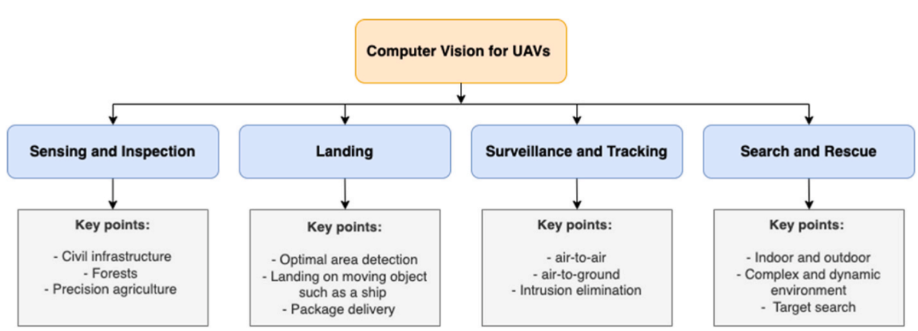

Fig. 1. Taxonomy of Computer Vision for UAV systems.

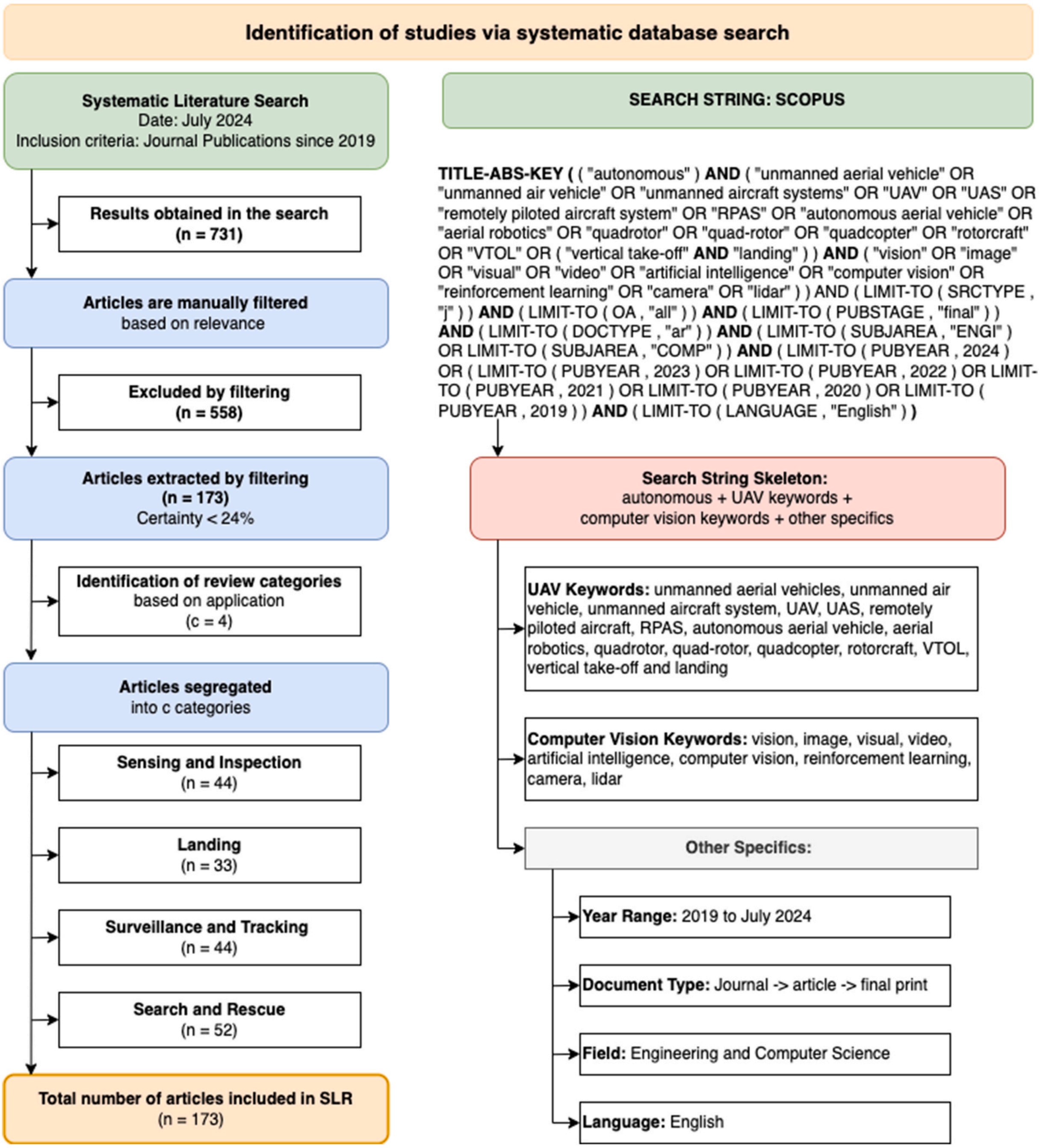

Fig. 2. Systematic search consort diagram.

systematic search and filtering methods, the review ensured a meticulous examination and classification of pertinent studies, thereby enhancing our comprehensive comprehension of progressions in this specialized field of research. Fig. 3 clearly demonstrates a growing research focus on the utilization of computer vision in Unmanned Aerial Vehicles (UAVs).

2.2. *Threats to validity*

Despite the authors' effort to review all the relevant studies without any bias, our work might be subjected to the following main threats to the validity.

* Time bias: The search query is only limited to procuring papers published between 2019 and July 2024, which results in excluding relevant studies published before 2019.
* Language bias: Only English language papers are included.
* Selector bias: The research literature on Computer Vision techniques is identified using authors' knowledge of variant approaches and related work in the recognized approaches. It is not practical to pinpoint all the approaches utilizing a search query, which could result in missing some CV techniques for autonomous UAV navigation.
* Retrieval bias: The search query is as mentioned above, so relevant studies that have used terms not in the search query might have been excluded. This was done to make the process systematic and to enhance reproducibility.

Table 1
Summary of excellent complementary survey contributions on artificial intelligence applications in UAVs.

Paper	Title	Publisher	Year	Scope
[13]	A survey on vision-based UAV navigation	Geo-spatial Information Science Journal by Taylor & Francis	Jan-18	Vision-based methods for UAV navigation, no emphasis on AI involvement, comprehensive review, not an SLR
[14]	Survey on Computational-Intelligence-based UAV Path Planning	Knowledge-Based Systems Journal by Elsevier	Oct-18	Specific to CI-based UAV path planning, no emphasis on computer vision, semi-SLR
[15]	Computer Vision in Autonomous Unmanned Aerial Vehicles – A Systematic Mapping Study	MDPI applied sciences	Aug-19	Computer Vision based approaches for autonomous UAV, not specific to navigation, SLR, from 1999 to 2017.
[16]	Flying free: A research overview of deep learning in drone navigation autonomy	MDPI drones	Jun-21	DL for UAV navigation autonomy, specific to autonomy, no emphasis on computer vision, comprehensive review, not an SLR
[1]	Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review	MDPI remote sensing	Jul-21	CV in Forestry for UAV data, specific to forestry, comprehensive review, not an SLR
[17]	Autonomous Unmanned Aerial Vehicle navigation using Reinforcement Learning: A systematic review	Engineering Applications of Artificial Intelligence Journal by Elsevier	Feb-22	Specific to autonomous UAV navigation using RL techniques, with no emphasis on computer vision, SLR
[18]	Artificial Intelligence Approaches for UAV Navigation: Recent Advances and Future Challenges	IEEE Access	Mar-22	AI approaches in autonomous UAV navigation, not specific to vision, comprehensive review, not an SLR
[19]	Artificial intelligence in civil infrastructure health monitoring—Historical perspectives, current trends, and future visions	Frontiers	Sep-22	AI is applied to civil infrastructure health monitoring, and it is specific to civil infrastructure, not an SLR.
[20]	A review of artificial intelligence applied to path planning In UAV swarms	Springer Neural Computing and Applications	Oct-22	AI applied to path planning in UAV swarms, no emphasis on computer vision or other aspects of navigation, not a comprehensive review, semi-SLR
[4]	Vision-Based Navigation Techniques for Unmanned Aerial Vehicles: Review and Challenges	MDPI drones	Jan-23	Vision-base navigation approaches for UAV, no emphasis on deep learning-based CV, comprehensive review, not an SLR
[21]	A Systematic Literature Review (SLR) on Autonomous Path Planning of Unmanned Aerial Vehicles	MDPI drones	Feb-23	Specific to autonomous path planning, no emphasis computer vision and AI involvement, SLR, from 2003 to 2022
[22]	A Review of Navigation Algorithms for Unmanned Aerial Vehicles Based on Computer Vision Systems	Springer Gyroscope and Navigation	Mar-23	Computer Vision based approaches for outdoor UAV navigation, not a comprehensive review or SLR
[23]	Evaluation of Machine Learning Approaches for Precision Farming in Smart Agriculture System: A Comprehensive Review	IEEE Access	Apr-24	AI approaches in precision farming, no emphasis on computer vision or other aspects of navigation, comprehensive review, not an SLR
-	Our Review	–	–	A comprehensive SLR focused on Computer Vision based approaches for autonomous UAV operation from 2019 to July 2024

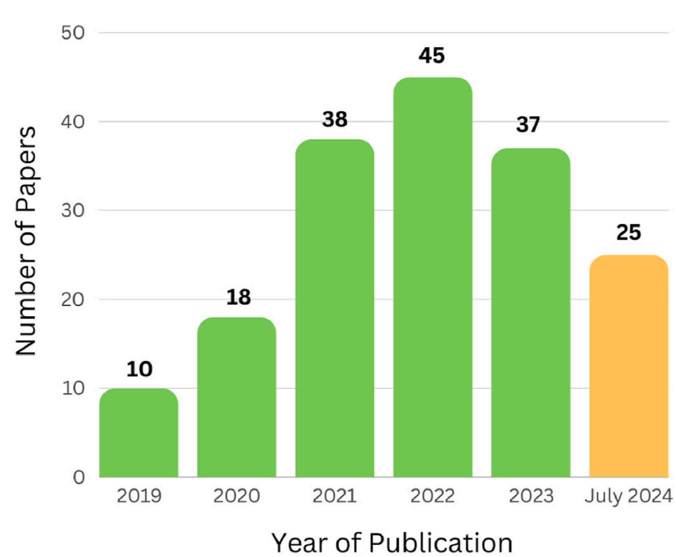

Fig. 3. Number of papers published per year.

However, the authors have tried their best to counter these biases to the possible extent.

3. Computer vision for UAV autonomy

Computer vision has great potential for applications involving

Unmanned Aerial Vehicles (UAVs). UAVs can achieve real-time scene understanding, object detection, and semantic segmentation from visual data captured by onboard cameras by utilizing advanced image processing algorithms, deep learning models, and convolutional neural networks [24]. The images taken by the camera unit can be of various types, including greyscale, RGB, RGB-D, thermal, infrared, etc. This fusion of computer vision and UAV technology paves the way for a new era of intelligent and adaptive aerial systems, transforming various industries and revolutionizing how we perceive and interact with the world.

Convolutional Neural Network (CNN) is a deep learning architecture specifically designed for processing and analyzing visual data, such as images and videos. It consists of multiple layers, including convolutional, pooling, and fully connected layers, that are adept at automatically learning and extracting hierarchical features from input images. The core operation of a CNN is convolution, where small filters slide across the input image to detect patterns and features. Pooling layers then reduce the spatial dimensions of the feature maps, reducing the computation and promoting spatial invariance [25]. The learned features are then fed into fully connected layers for classification or regression tasks.

Transfer learning has emerged as a game-changer in the domain of UAVs. It involves leveraging knowledge from pre-trained models on a source domain and applying it to a related target domain with limited data, thereby enhancing the learning process by benefiting from shared features and representations [26]. The source domain can be a simulation or a real environment. The applicability of transfer learning in UAVs is rapidly gaining momentum due to its potential to address data scarcity and domain shift challenges faced during aerial missions. By fine-tuning pre-trained models on new tasks or domains encountered in real time, UAVs can quickly adapt to novel environments, optimize resource usage, and improve their performance and decision-making capabilities [27].

Integrating transfer learning in UAVs promises to revolutionize aerial operations by enabling sophisticated functionalities like object detection, target recognition, and terrain classification, even in previously unexplored or dynamic scenarios. As transfer learning advances, we can foresee UAVs becoming more agile, intelligent, and versatile, facilitating seamless collaboration in complex missions, ranging from precision agriculture and wildlife monitoring to disaster response and urban surveillance, ultimately contributing to safer, more efficient, and sustainable aerial ecosystems [28]. Moreover, EMDL enhances reliability in multiview deep learning by dynamically fusing view-specific evidence, crucial for high-risk applications like autonomous UAV navigation [29]. Similarly, TABLE method's approach to balancing multi-view data for stock ranking aligns with our focus on optimizing multi-modal inputs in UAV navigation using deep learning [30].

Furthermore, we observed that the You Only Look Once (YOLO) model had a significant impact on CV research for the UAV domain, as indicated in Fig. 4. YOLO is a one-stage object detection algorithm that simultaneously does regressive prediction for target placement and classification using a unified convolutional neural network structure. Its ability to detect multiple objects in a single image gives it an advantage over other models in the case of highly complex images with many objects or intricate patterns. YOLO uses a grid-based approach to detect objects within the image at different scales and locations. This makes YOLO more efficient and accurate for object detection tasks compared to other models [31].

Various YOLO versions have been used in UAV applications, including (in no specific order) YOLOX, YOLOv2, YOLOv2-Tiny, YOLOv3, YOLOv3-Tiny, YOLOv4, YOLOv4-Tiny, YOLOv5s, YOLOv5x, YOLOv5I, YOLOv5m, YOLOv7, YOLOvR etc. [32].

Fig. 5 depicts the usage and range of applications of YOLO (You Only Look Once) in the domain of computer vision between 2019 and 2024. The rising prevalence of YOLO usage indicates its expanding significance in both academic research and practical applications. This increase highlights the effectiveness and versatility of the technology in tackling different obstacles in object detection and recognition. Fig. 6 provides a comprehensive overview of the dataset types that are frequently used in conjunction with YOLO. It highlights a significant dependence on well-known benchmarks such as COCO (Common Objects in Context) and VOC (Visual Object Classes). The existence of custom datasets further emphasizes the adaptability of YOLO to meet specific application requirements, demonstrating its flexibility beyond standardized datasets. Fig. 7 demonstrates that the use of sensor modalities with YOLO is primarily focused on RGB cameras. This preference is due to the wide availability and suitability of RGB cameras for visual-based tasks. Efforts are being made to improve YOLO's capabilities in various environmental conditions and sensing scenarios by incorporating additional modalities like LiDAR and multispectral sensors. Fig. 8 analyzes the various environments where YOLO has been evaluated, demonstrating a comprehensive investigation of indoor, outdoor, urban, and rural/natural settings. The wide range of testing scenarios highlights the strength and versatility of YOLO, making it suitable for different operational settings, including controlled indoor environments and challenging outdoor and natural landscapes.

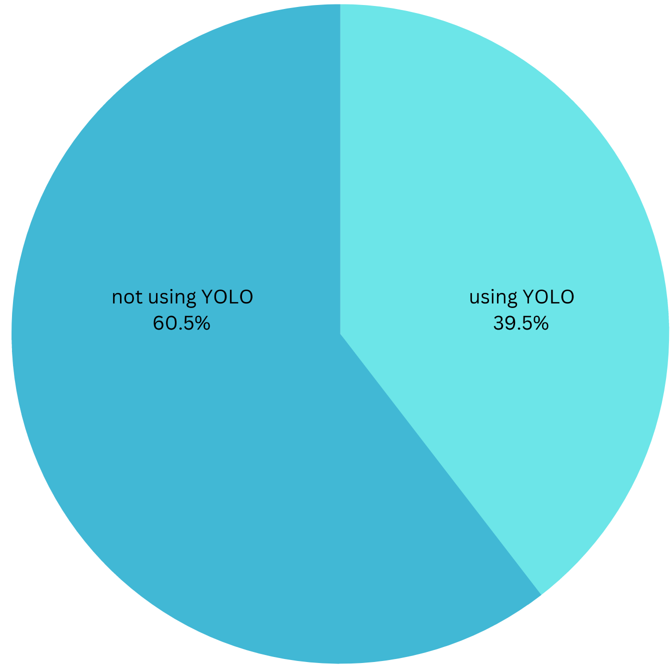

Fig. 4. Extent of YOLO usage in Computer Vision for UAV research.

Evaluation metrics are essential for evaluating the performance of YOLO (You Only Look Once), a commonly employed object detection framework. Two commonly used metrics are Average Precision (AP) and mean Average Precision (mAP) [33]. These metrics assess the trade-off between precision and recall, which is crucial for measuring the accuracy of detection at different thresholds. Recall, also known as Sensitivity, measures the ability of a model to correctly identify all positive instances in a dataset. Precision, on the other hand, evaluates the accuracy of positive predictions made by the model. IoU, short for Intersection over Union, quantifies the degree of overlap between predicted and ground truth bounding boxes, offering valuable information about the accuracy of localization [34]. Moreover, the F1 Score combines precision and recall, providing a well-balanced evaluation metric. These metrics are essential benchmarks that guide the evaluation and improvement of YOLO models [35]. They are crucial for advancing object detection capabilities in various applications and scenarios in computer vision research and development.

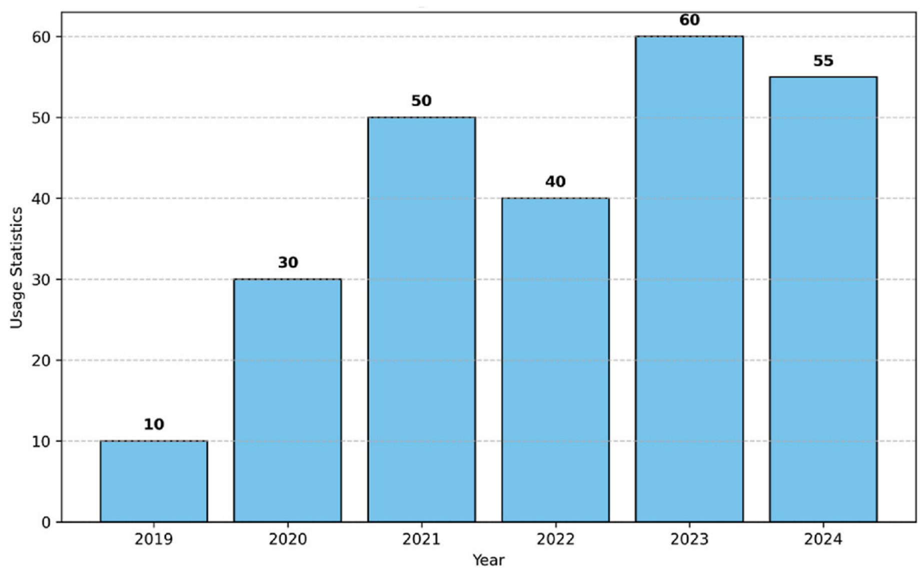

Fig. 5. YOLO usage Statistics (2019–2024).

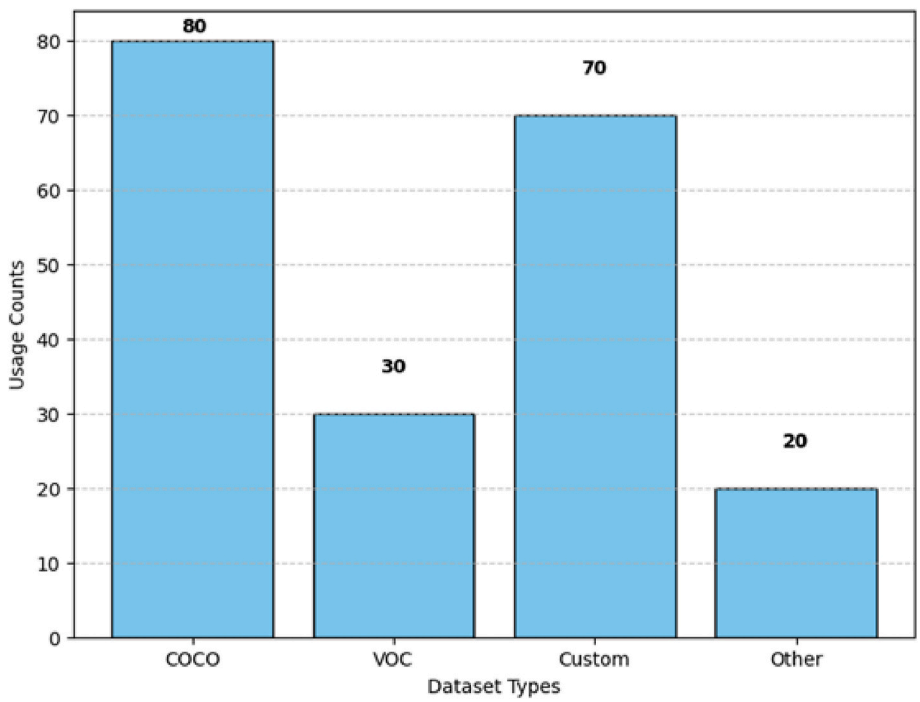

Fig. 6. Dataset types used with YOLO.

3.1. Computer vision-based sensing and inspection for UAVs

The inspection involves the detection of anomalies such as cracks, incorrect placement, and pest infestation in critical infrastructure, including bridges, roads, crops, transmission towers, and lines. Individuals or swarms of UAVs are used to collect and analyze 2D or 3D aerial RGB, thermal, or infrared images of the structure or concerned area. This eliminates the risk of a person physically getting involved and collecting the data from the structure already in question for damage. All the collected studies are compared and presented in Table 2.

Roads are one of the primary means of transportation. In the fiscal year 2019, movement through road transport amounted to around 22.6 trillion passengers per kilometer in India alone [36]. For inspecting such critical infrastructure spanning millions of kilometers, it is in the best interest to use technologies like UAVs to detect and analyze the state of the roads. One such approach is proposed by Ref. [37], where a deep CNN-based autonomous road inspection framework is developed to analyze aerial RGB images and detect potholes with a precision of 98.26 %. Other classes include road cracks and the yellow line in real-time. The yellow line is used as a marker along which the UAV navigates while simultaneously detecting road cracks and potholes. An improved highway center marking detection model using YOLOv3 proposed by Ref. [38] can be employed to further enhance UAV following, achieving an AP of 82.79 %.

For UAV-assisted bridge inspection [39], presented an automatic crack segmentation approach using Mask R-CNN with an accuracy of 90 %. First, the UAV is used to collect 2D RGB images in a predetermined trajectory, followed by using Mask R-CNN in the damage identification

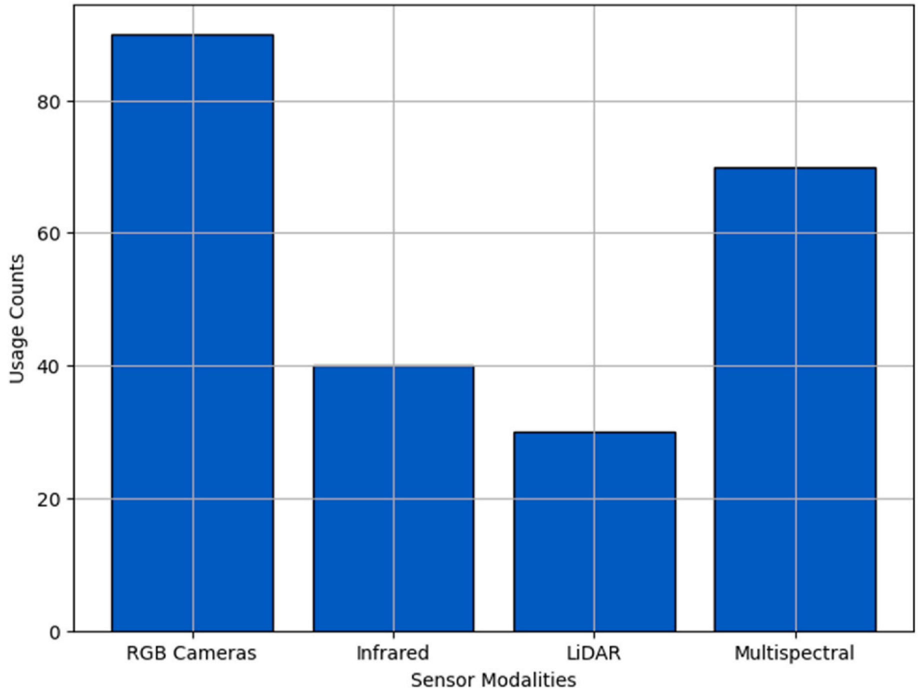

Fig. 7. Sensor modalities utilized with YOLO.

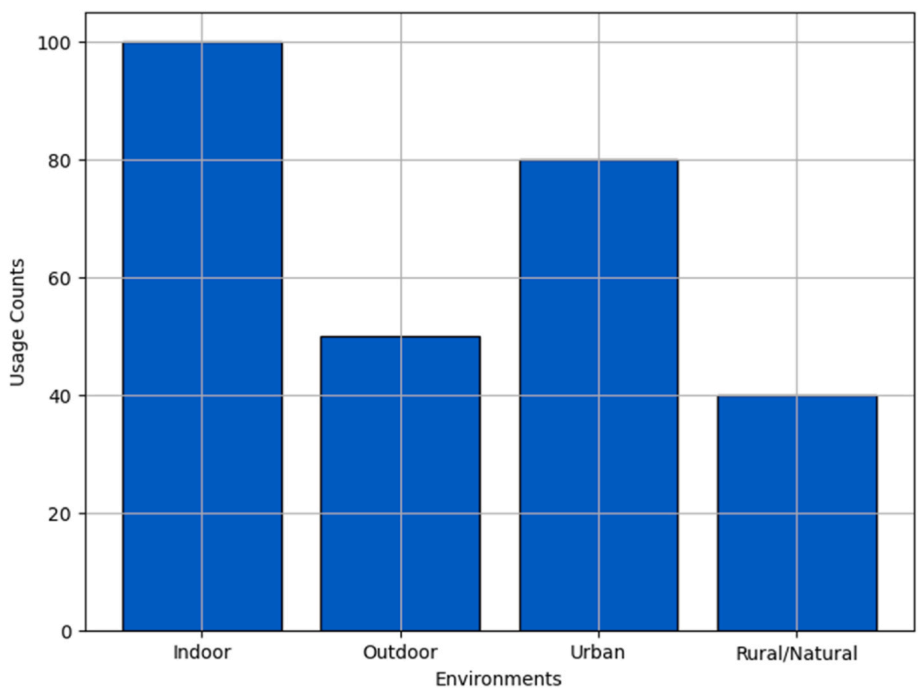

Fig. 8. Environment tested with YOLO.

stage for per-pixel segmentation of cracks. They have also used post-processing elements such as photogrammetry to maintain a record of the geometry of each bridge asset, which could be used for navigation purposes by producing 3D imagery of the bridge from the collected 2D images. Making this process semi-autonomous [40], proposed using digital models and a hybrid approach based on DL and photogrammetry. They employed Inceptionv3 and U-Net architectures for crack detection and segmentation. The output of this CNN is used to generate a 3D model that enabled crack localization and quantification on this 3D model, achieving an impressive accuracy of 98.71 %.

Researchers have not only worked on the anomaly detection stage but also explored aiding navigation for inspection. For example, a frame-wise RGB image processing approach using UNET3+ is employed by Ref. [10] for detecting and localizing critical structural components, such as columns of a railway bridge post-earthquake, for generating waypoints in an online manner. This approach reported an mPR of 96 % [41]. used YOLOv4 to identify unburied offshore pipelines, followed by an image processing technique such as Canny edge detection and Hough transform, which is used to detect the pipeline reference with a mAP of 72 %. Similarly [42,43], used GAN, SOLOv2, and AdaLSN to extract the

Table 2
Summary of studies classified under Computer Vision based Sensing and Inspection for UAVs.

Reference	Year	Optical sensor	Dataset	Method	Performance
[46]	2019	RGB	Self Collected	R-CNN	AP = 89.9 %
[47]		RGB	Self Collected	R-CNN	aPR = 89.9 %
[48]	2020	RGB	Molley Caren Research Farm 2017	CNN	Acc = 89.8 %
[39]		RGB	Self Collected	CNN	Acc = 90 %
[49]		RGB	Self + "Earthquake Datasets from Purdue University"	Faster RCNN Inception-ResNet-v2	mAP = 60.8 %
[50]		RGB-D	Self + "NYUv2″ "KITTI depth completion dataset"	CNN	RMSE = 0.510
[51]		RGB	Synthetic Self Generated	LS-Net-VGG-16	aPR = 80.04 % F1 Score = 59.40 %
[52]		RGB	"WeedMap by Inkyu Sa et al."	SegNet, HoughCNet	MPR = 90.1 % MR = 70.72 % mF1 Score = 77.94 %
[53]	2021	RGB	Self + Open Datasets	YOLO v4 Tiny	Inference Time (ms) = 29.4
[54]		RGB	Self Collected	PADENet	mAP = 87.34 %
[55]		RGB, Depth, Thermal	Synthetic Self Generated	YOLOv2	A = 90 % RMSE = 0.4
[37]		RGB	Self Collected	CNN	Potholes class: Precision = 98.26 % F1 Score = 94.04 %
[56]		RGB	Self Collected	YOLOv4-tiny	Acc = 84.4 %
[57]		multi-spectral	Self Collected	CNN	Acc = 88 %
[58]		RGB	Self Collected	YOLOv4	F1 score = 77 %–81 %
[38]		RGB	Self + "UAV123" + "CULane"	YOLOv3	AP = 82.79 %
[59]		RGB	Self Collected	CNN	Acc = 92 %
[41]	2022	RGB	Synthetic Self Generated	YOLOv4	mAP = 72 %
[42]		RGB	Self-Collected ("iRailway")	RT-GAN with CSPDarknet53 backbone, FC Patch-GAN	IoU = 88.07 %
[60]		RGB and thermal	Self Collected	NN	RMSE = 0.016 (m)
[61]		RGB-D	Synthetic Self Generated	DCNN	IoU = 93 % mAcc = 96.5 %
[45]		RGB	Self Collected	Fast R-CNN	FP = 2.1 %–2.6 %
[62]		RGB	Self Collected	YOLOv5-s	mAP = 92.1 %
[7]		infrared	Synthetic Self Generated	DQN	SR = 80 %
[10]		RGB	Self + "Tokaido Dataset"	UNET3+	mPR = 96 %
[9]		RGB	Self Collected	ResNet50	Acc = 98.77 %
[63]		RGB	Self Collected	YOLOv3	Acc = 99.23 %
[64]		RGB	Self Collected	YOLOv2, v3, v4	mAP = 95 %
[65]		RGB, infrared	Self Collected	YOLOv5x, v5s	Acc = 75 % PR = 85 %
[43]	2023	RGB	Self Collected	SOLOv2, YOLOv7 AdaLSN	mAP = 93.98 %
[66]		RGB-D	Self Collected	YOLOv4	Acc = 99 %
[67]		RGB	Self Collected ("Flower Detection Dataset")	YOLOv5, v7, vR	PR = 84 % Recall = 80 % F1 score = 82 %
[44]		RGB and thermal	Self Collected	FCN, RetinaNet	mAcc = 90.5 %
[8]		RGB	Self Collected	NN	Acc = 97.3 %
[40]		RGB and Depth	Self Collected	Inceptionv3 and U-Net	Acc = 98.71 %
[68]		RGB and thermal	Self Collected	D-LinkNet	Acc = 99 % Precision = 80 %
[69]		Depth	Synthetic Self Generated	CNN-AD and LSTM-AD	Acc = 80 % and 90 %
[70]	2024	RGB	FLAME Dataset	CNN	Acc = 99.46 % mAP = 99.64 %
[71]		RGB and Depth	Self Collected	RoMP Transformer	mSKEWIoU + DIoU = 92.4 %
[72]		RGB	Self Collected	YOLOv5-v1	mAP = 86.8 %
[73]		RGB	Self Collected	YOLOv5	Precision = 88 % Recall = 90 %
[74]		RGB	Synthetic Self Generated and Hybrid	attention-based U-Net	Precision = 99.5 % IoU = 98.6 %
[75]		RGB	Self Collected	U-Net	Acc = 97.42 %
[76]		RGB	Self Collected	YOLOv5-Large	Precision = 95 % mAP = 96 %

segment and extract the structure of railway lines, achieving an IoU of 88.07 % and mAP of 93.98 %, respectively. These extracted structures are paired with navigation approaches such as line-following algorithms to guide the UAV along the structure to collect images of those components at a desired level of detail in the later stage.

Concrete buildings require regular maintenance checks, especially as the building gets older, to ensure residents' safety. Unlike roads, buildings don't experience constant heavy movement. In Ref. [44], the authors suggested using RGB and thermal images as input for detecting and labeling visual and thermal anomalies, respectively. The proposed approach achieved a mean accuracy of 90.5 %. They employed a strategy of a comprehensive analysis of RGB images using RetinaNet in the macro and FCN in the microanalysis stages. Other CV methods detect thermal anomalies in the macro analysis stage. Taking a step further [45], proposed using Fast R-CNN to not only detect wall cracks in pre-processed RGB images but also classify them as hazardous or non-hazardous, achieving a false positive rate of around 2.1 %.

Electricity transmission systems have been one of the hotspots for UAV inspection research owing to their high-risk involvement because of both the structure's height and lethal amounts of electricity being transmitted. Transmission power would be turned off to allow inspection. However, in the case of a UAV, the risk of life can be eliminated, and the power need not be switched off as the UAV inspects the system from a safe distance. Researchers have taken a segmented approach in dealing with the transmission systems where the transmission tower, lines, insulators, and other anomalies are researched individually.

For close proximity power transmission tower inspection [46], proposed an R-CNN-based approach to detect the tower and reported an average precision of 89.9 %. After this, the UAV uses the predetermined waypoints to navigate itself around the tower. A different approach is presented by Ref. [74] employing an attention-based U-Net model to precisely segment images taken by a camera mounted on a UAV in order to allow a motion module to generate collision-free and inspection-relevant maneuvers of the UAV along different types of towers, achieving a precision of 99.5 % for precise movement of visually guided UAVs. As a next step [47], suggested an end-to-end approach for the detection of towers and power lines for inspection whilst maintaining an aPR of 89.9 %. The tower detection and localization in real-time are performed using Fast R-CNN, and LSD is used to detect the transmission lines. Additionally [53], proposed using YOLOv4 Tiny for power line insulator detection with the approach having an inference time of 29.4 ms, after which a cleaning tool can be used to perform the insulator's maintenance work. The fault detection part hasn't been incorporated into the approach; they have only presented a way to safely and effectively navigate the system.

In extension, works by Refs. [56,58] provided approaches to classifying the detected components as faulty or normal. These methods using YOLOv4 and YOLOv4-tiny, respectively, could detect and differentiate multiple components, including insulators, vibration dampers, and antennas of a transmission tower, and then classify them as normal and faulty, using RGB images as input with an F1 score of 81 % and accuracy of 84.4 % respectively. Taking it a step further [66], proposed using a swarm of drones to do the same and achieved an impressive accuracy of 99 %. Furthermore, an approach using YOLOv5-s and RGB images for real-time detection of anomalies in transmission systems such as birds' nests is presented by Ref. [62]. This method is reported to carry out the task with a mean average precision of 92.1 %. Taking advantage of the latest models [71], proposed an autonomous flight strategy for the transmission tower approach that integrates multimodal information from novel sensors, including optical camera and 3D light detection and ranging. They introduced a novel deep neural network architecture titled the rotational bounding box with a multi-level feature pyramid transformer for accurate object detection, achieving mSKEWIoU + DIoU of 92.4 %.

A significant amount of research has also been observed in precision agriculture, where UAVs are utilized for sensing and inspecting crop and forest health indicators. For example [72], presented a fruit detection and counting method that involves simultaneous capture and synchronization of video frames from multiple UAV cameras, converting them into a cohesive data structure and, ultimately, a continuous image. Employing YOLOv5-v1, this approach could achieve an mAP of 86.6 %. Taking this further on the diagnosis direction [48], proposed a fully autonomous aerial scouting method, emphasizing application over large crop fields, employing CNN, with RGB images as input, for sensing and predicting crop health with an accuracy of 89.8 %. Similarly [65], proposed using YOLOv5s for real-time detection of forest health indicators (FHIs) detection and achieved an accuracy of 75 %. A unique approach has been taken by pairing a semi-autonomous quadruped robot with a hexacopter UAV, enabling effective detection of FHIs from both the top and ground levels. Additionally, much effort is seen in weed and infestation detection, where a model is trained to identify specific plant/insect species/symptoms [8,9,63,64].

3.2. Computer vision-based landing for UAVs

Landing involves the detection of an optimal region, avoiding obstacles, and dealing with challenging situations like landing on a moving platform such as a ship. Depending on the application, the optimal landing region varies. All the collected studies are compared and summarized in Table 3.

Applications involving package delivery make use of landing identification markers such as AprilTags (see Table 4). For example [77,78], worked on using CNNs and YOLOv3, respectively, for autonomous landing marker detection in a real environment. The former method achieved an accuracy of 97 %. The approach involving CNNs was capable of detecting the marker from as high as 20m. The dataset for training the models so far was real images or a mix of real and simulator images of the landing marker [79]. proposed a sim-to-real approach that generated a success rate of 91 % by training an SDQN model composed of DQNs with CNN layers solely in simulation using simple uniform textures and later tested in complex simulation and real-world environments.

In order to enhance the robustness of the landing marker detection in challenging situations such as extremely low illumination, researchers have suggested employing a two-phase framework of image enhancement and object detection. For instance Ref. [80],proposed employing CNNs for outdoor low-illumination landing. The model reported an F1 score of 96.1 %, consisting of a model-based scheme for low-illumination image enhancement and a hierarchical-based method including a decision tree associated with lightweight CNN for coarse-to-fine landing marker detection. Also [81], presented a method using R-CNN for the night landing of fixed-wing UAVs that achieved an average precision of 84.37 %; they adopted a strategy of visual and infrared image fusion and then used Faster R-CNN for runway detection on top of the enhanced image. In addition to low light, blur images captured during the fast motion of the UAV also pose a problem for precise marker detection. The work employed the SlimDeblurGan model for deblurring the input grayscale images and then used YOLOv2 for landing marker detection with a 64.5 % precision.

QR detection is the most common landing approach, but it is not versatile or flexible. If the drone faces a failure mid-flight, it would be impractical to find a landing spot with a QR code. Solving this [82], suggested using a human as a marker for the landing sport, with the model achieving a recall of 76 %; human detection is performed using SSDLite-MobileNet-V2. This approach would be especially beneficial for drone delivery, where a drone could detect the receiver and land near him. However, this approach assumes that the person is standing still during the detection phase and in a safe area for the UAV to land.

A more convenient way would be to detect any safe flat landing area. As a step closer to this, the works by Refs. [83,84] have employed CNNs and YOLOv5l models, achieving a mean IoU of 68.88 % and F1 Score of 65.5 %, respectively, to determine a flat landing area. However, the approach cannot differentiate between land and water, both of which can be equally flat. As a solution [85], used CNNs for ground surface classification and the optical flow method for flat surface detection. The results are presented on a heat map. The output of both models is a heatmap, which will later be fused to generate a single heatmap, which will, in turn, determine the safe location to land. This approach is unique as it does not just detect a flat surface to land but distinguishes between land and water along with nine other surfaces to determine which surface is safe to perform landing. However, the landing zone doesn't have to be completely flat for the UAV to land. It can tolerate a certain degree

Table 3
Summary of studies classified under Computer Vision-based Landing for UAVs.

Paper	Year	Optical sensor	Dataset	Method	Performance
[77]	2019	RGB	Natural and Synthetic Self Generated	CNN	Acc = 97 % F1 Score = 95.6 %
[87]	2020	Grayscale	Synthetic Self Generated ("SMBD-DB1″ and "RMBD-DB1″)	SlimDeblurGAN YOLOv2	PR = 64.5 % Recall = 64.1 % F1 Score = 64.3 %
[79]		RGB	Synthetic Self Generated	SDQN, CNN	SR = 91 %
[82]		RGB	Self Collected	YOLO SSDLite-MobileNet-V2	Recall = 76 %
[88]		RGB	"VisDrone"	FCN	Acc = 87.95 % PR = 88.37 %
[80]	2021	RGB	Self Collected	CNN	Recall = 94.9 % F1 Score = 96.1 %
		RGB	Self Collected	CNN-LSTM	RMSE = 0.1715
[89]		RGB	Self + "Venice" + "VisDrone"	Pruned BL CNN	IoU = 61.1 %
[90]		RGB	"Aerial Semantic Segmentation Drone Dataset"	DeepLab–ResNet34, PSPNet–ResNet34	IoU = 64.4 %
[91]		RGB	Synthetic Self Generated	image-based RL	Lateral Error = 0.04–0.5 m
[92]		RGB	Natural and Synthetic Self Generated	BboxLocate-Net PointRefine-Net	mAP = 96.3 %
[85]		RGB	Self Collected	CNN	Heat Map
[93]	2022	BW	Self Collected	YOLOv4-tiny	Landing Precision (x) = 0.06 m Landing Precision (y) = 0.05 m
[94]		RGB	Self Collected	YOLOv3	FPS = 8.83
[95]		RGB	Self Collected	YOLOX	mAP = 77.3 %
[96]		RGB	Synthetic Self Generated	YOLOv5	mAP = 99.5 % mIoU = 95.4 %
[81]		RGB, Infrared	Synthetic Self Generated	Faster R-CNN	AP = 84.37 %
[97]		RGB	Self Collected	YOLOv5	TP = 0.99
[83]		RGB, 3D LiDAR	"Aerospace Dataset"	CNN	mIoU = 68.88 %
[78]		RGB	Synthetic Self Generated	YOLOv3	_
[98]		RGB	"VisDrone" and "Stanford Drone Dataset"	YOLOv5	_
[99]		RGB	"UAV123"	VitP-RCNN	AP = 92 %
[84]		RGB	"DOTA"	YOLOv3, v4,v5I	Recall = 0.611 F1 Score = 0.655
[100]		RGB	Self Collected	YOLOv4	RMSE (x) = 0.0335 m
[101]	2023	RGB	Self Collected	YOLOv3	Detection Range = 250 m
[102]		RGB	Synthetic Self Generated	YOLO	RMSE = 0.63 m
[103]		RGB	Self Collected	YOLOv5 m	mAP_50 = 97.6 %
[104]		Grayscale	Synthetic Self Generated	CNN	Acc_x = 0.23 m Acc_y = 0.02
[105]		RGB	Self Collected	YOLOv5	AP = 99.5 % Acc = 0.066 m
[106]	2024	RGB	"HelipadCAT"	ORB	RMSE = 10.51 m
[86]		RGB	Self Collected	VGG16 and Xception	Precision = 0.91 Recall = 0.82
[107]		LiDAR	"Paradise", "Holyrood" and "Semantic3D"	ConvPoint	Acc = 92.1 %
[108]		RGB	Synthetic and Self Collected	DeepLabV3Plus and U-Net	F1_People = 99.4 %

of slope. To estimate this [86], introduced an approach based on VGG16 and Xception models with a precision and recall of 91 % and 82 %, respectively.

An ideal landing approach should not only include detecting a safe landmass but also be capable of landing safely in a populated area. In this light [88], implemented a crowd detection algorithm using FCN, achieving an accuracy of 87.95 % and outputs a heat map. Also [89], presented a Pruned BL CNN model that generates a density map with 61.1 % IoU. Taking this a step further [108], proposed a much more robust, fast, and lightweight approach that uses a DL-accelerated image processing pipeline, employing DeepLabV3Plus and U-Net models, for accurate detection and relative pose estimation of the UAV with respect to the landing pad with error estimation for each. Adding to the robustness of this approach, it also calculates human presence probability via segmentation with an F1 score of 99.4 %.

Methods so far include stationary landing areas. Landing on mobile platforms such as ships and moving cars is more challenging. Researchers [101] employed a vision system utilizing YOLOv3 for long-range ship tracking with a detection range of 250 m and classical computer vision for estimating aircraft relative position and orientation using the horizon bar during the final approach for landing. Additionally, Authors [93,100] proposed using YOLOv4-tiny and YOLOv4 reporting landing precision of 0.06 m and RMSE of 0.0335 m concerning the x-axis, respectively, to detect each ArUco marker on a moving platform.

3.3. Computer vision-based surveillance and tracking for UAVs

Surveillance is to detect, track, and eliminate intrusive objects; this is especially significant in places like airport obstacle-free zones and no-fly zones. Air-to-air and air-to-ground surveillance are two key applications of unmanned aerial systems. All the collected studies are compared and presented in Table 3.

Counter-drone systems are gaining popularity owing to the increasing number of drones taking to the skies [109]. have employed YOLOv3 and implemented a real-time drone detection system that can achieve an accuracy of 91 % and F1 score of 94 % [110]. presented an approach using RetinaNet with a ResNet50 backbone for vision-based following of multirotor. Their model was trained using synthetic data generated by a commercial game engine and was able to generalize to a wide range of real-world multirotors and environments, reporting an average precision of 75 %. Adding another dimension to images [111],

Table 4
Summary of studies classified under Computer Vision based Surveillance and Tracking for UAVs.

Reference	Year	Optical sensor	Dataset	Method	Performance
[122]	2019	RGB	Self Collected	YOLOv3	TP = 0.98
[110]		RGB	Synthetic Self Generated + "COCO"	RetinaNet-ResNet50	AP = 75 %
[123]		RGB	Self Collected	VGG-M YOLO v2, CNN	mAP = 96 %
[124]		RGB	Self Collected	SSD	Acc = 98 %
[121]	2020	RGB	Self Collected	YOLO	Acc = 94.52 %
[125]		RGB	Self Collected + "COCO"	MobileNet Faster-RCNN	Acc = 72.2 % MSE = 0.113
[114]		RGB	Self Collected	YOLOv3	HIT Rate = 94.44 %
[109]	2021	RGB	Synthetic Self Generated	YOLO EfficientNet	Acc = 91 % F1 Score = 94 %
[126]		RGB	Synthetic Self Generated	CNN	mIoU = 0.2488
[127]		RGB	Self Collected	YOLOv3, CNN	RMSE = 0.93
[118]		RGB	"VisDrone"	R-CNN	AP = 34.57 %
[128]		RGB	Synthetic Self Generated	CNN	Acc = 95.67 % F1 Score = 84.68 %
[115]		RGB, Depth	Self Collected	Faster R-CNN	AP = 88.33 %
[129]		RGB	Self Collected	OpenPose	SR = 10/10
[130]		RGB-D	Self Collected	YOLO v4 Tiny	mAP = 79.36 %
[131]		RGB	Self Collected	Yolo-v3	MTE (x) = 0.46 m MTE (y) = 0.02 m SSE = 0.41 m
[132]		RGB	Open Datasets (16)	YOLOv3	Q = 79.2 %
[116]		Grayscale	Self Collected	YOLOv3	Acc = 92 % PR = 93 %
[133]		Grayscale	Self Collected	PULP-Frontnet CNN	MSE (x) = 0.066 MSE (y) = 0.078 MSE (z) = 0.02
[134]	2022	RGB	"UC Merced Land Use Dataset"	ResNet	Acc = 99.15 %
[135]		RGB	Self Collected	YOLOv3-tiny	IoU = 0.67–0.94
[12]		RGB	Synthetic Self Generated	RetinaNet	Acc = 88 %
[136]		RGB	Self Collected	CNN and CNN + GAN - ResNet-18	_
[119]		RGB	Self Collected	CNN	Error (u) = 24–50 Error (v) = 10-20
[137]		RGB	Synthetic Self Generated	Airpose	MPE = 0.15 m MPJPE = 0.072
[138]		RGB	pre trained	CNN	_
[139]		RGB	Self Collected	YOLOv5	Acc = 99.5 %
[120]		RGB	"COCO" and "SEAGULL"	YOLOv5	mAP_0.5 = 57 %
[117]		RGB	"UCM"	OSQN-DNN	Acc = 98.99 % F Score = 98.54 % Recall = 98.94 %
[113]	2023	RGB	"UCF-ARG"	CNN	SR = 96.15 %
[112]		RGB	Self Collected ("visioDECT" and "attached object recognition dataset"	YOLOv5	F1 Score = 99.8 %
[140]		Depth	Synthetic Self Generated	CNN YOLOv4-tiny, DQN	_
[141]		RGB, Depth	Self Collected	YOLO	RR = 99.7 %
[142]		RGB	"Traffic Img … Management" and "Roundabout Ae … Detection"	YOLOv5 and v8, EfficientDet and DETR	mAP50 = 80 %
[143]		RGB	Synthetic Self Generated	MobileNetV3	mAP = 98.23 %
[111]		RGB	Self Collected	YOLOv5s	mAP = 86.87 %
[144]		RGB	"MOT20"	YOLOv5	Precision = 61 %
[145]	2024	RGB	Self Collected	CNN	V_error = 0.07 %
[146]		RGB-D	Self Collected	YOLO-MAD	V_error = 0.23 m/s X_error = 0.11 m
[147]		RGB-D	Self Collected	YOLOv3	_
[148]		RGB	Self Collected	YOLOv5-SE	mAP50 = 96.7
[149]		RGB	"UAVDT" and "VisDrone-2021"	CSPRegNet-800M	mAP50 = 62.5 %
[150]		RGB and Greyscale	Self Collected	MobileNetV2-based CNN and PULP-Frontnet-based CNN	MAE improvement = 24 %
[151]		RGB	Synthetic Self Generated	YOLOv5	_

introduced an improved UAV detection in camera videos through optical flow and spatiotemporal information. Here, the authors have employed the YOLOv5s model and proposed a novel detection method that extends the input of the model to the continuous sequence of images and inter-frame optical flow. This integration of spatiotemporal data helps detect small and weak targets better, adding to the method's robustness and achieving an average precision of 86.87 %.

Taking air-to-air surveillance a step further [112], proposed a multimodal approach for determining the malicious status of a drone by detecting and classifying the kind of object carried by the drone, such as guns, missiles, etc. They achieved this using the backbone layer comprising a cross-partial network (CSP) and a focus on feature extraction. The neck layer consists of a path aggregation network for feature aggregation, while the head layer has the YOLOv5 for actual prediction. Their approach can achieve an impressive F1 score of 99.8 %. Additionally [113],presented a real-time-based lightweight CNN named HarNet, which proposed to classify human actions from UAV sequence videos with a 96.15 % success rate.

However, the approaches so far only detect drones but cannot eliminate them. A solution to this is seen in the work of [114], where they presented a counter object system to detect and eliminate airborne balloon objects and small aircraft for airport obstacle-free zone monitoring. Detection was performed using YOLOv3, and the object was eliminated using a gel ball blaster with a hit rate of 94.44 %.

Table 5
Summary of studies classified under Computer Vision based Search and Rescue for UAVs.

Paper	Year	Optical sensor	Dataset	Method	Performance
[159]	2019	RGB	Self Collected	Tiny-YOLOv3	mAP = 67 % Recall = 70 %
[158]		RGB	Self Collected	CNN	PR = 97.8 % F1 Score = 98.7 % Recall = 99.6 %
[160]		RGB	Self Collected	CNN	_
[161]	2020	RGB	"VOC2012"	YOLOv3	_
[162]		RGB	Self Collected	CNN	Acc = 84 %
[163]		RGB	Synthetic Self Generated	CNN	RMSE = 0.199
[154]		RGB	Self Collected	YOLOv2	Acc = 77.4 %
[155]		RGB-D	"COCO"	Tiny-YOLOv2	ATE = 0.225
[164]	2021	RGB	Self Collected	CNN	Acc = 90 %
[165]		RGB	Self Collected	YOLOv5	mAP = 95 %
[166]		RGB, depth	Synthetic Self Generated	DQN	SR = 91 %
[167]		RGB	Synthetic Self Generated	D3RQN	SR = 96.7 %
[168]		RGB	"Udacity" and "Collision"	CNN	Acc = 95.6 % F1 Score = 90 % EMSE = 0.114
[169]		RGB	Synthetic Self Generated	YOLOv3	_
[170]		RGB	"COCO"	Mask R-CNN	_
[171]		RGB	Synthetic Self Generated	CNN	Avg_rewards = 0.2515 Num_success = 40.5
[172]		RGB	Synthetic Self Generated + "VisDrone-2018"	Decomm-Net Enchance-Net	Acc = 72.7 %
[173]		RGB	Synthetic Self Generated	DQN, CNN	SR = 94 %
[174]	2022	RGB	Self + Open Datasets	CNN	_
[175]		RGB	"Udacity" and "Collision-Sequence"	Drone-STM-RENet	Acc = 96.26 % Recall = 95.47 %
[176]		RGB	Self Collected	CNN	Error = 3 %
[177]		RGB-D	Synthetic Self Generated	CNN	SR = 98 %
[178]		3D SCDM, WDM	"WDM"	CNN	Acc = 99.6 %
[156]		RGB-D	Self Collected	BiSeNetV2	Acc = 89 %
[179]		RGB	Synthetic Self Generated	FCNN, R-CNN	AP = 97.35 %
[157]		RGB	Self Collected	pre-trained VGG16	Acc = 91 %
[180]		RGB	Self Collected	DeepLabV3	mIoU = 95.6 %
[181]		RGB	_	YOLOv3	_
[182]		RGB-D	"TartanAir"	R-MSFM, MiDaS, GLPDepth, MiDaS	RMSE = 0.023 RMSElog = 0.205
[183]	2023	RGB	"VR-EyeTracking"	Lite-TA-MSNet-EfficientNet, GNN	CC = 0.2 SIM = 0.25 KL = 1.22
[184]		RGB, depth	Synthetic Self Generated	DCCN, DQN, SAC	_
[152]		RGB	"OpenImages", "DoorDetect" and "Stairs Detection"	CNN	AP = 88.79 %
[185]		RGB	Synthetic Self Generated	CNN	SR = 90 %
[186]		RGB, thermal	"COCO"	lightweight CNN	mIoU = 68 %
[187]		RGB	Self Collected	SSN	mIoU = 63 %
[188]		2D range image	"KITTI odometry"	Darknet from YOLOv3	translation error = 0.80
[189]		RGB-D	Synthetic Self Generated	ETHP	Success Rate = 91 %
[190]		Depth	Synthetic Self Generated	SAC_CNN, SAC_ViT and SAC_FAE	Flight Quality SR, S1 = 0.82
[191]		RGB	Self Collected	lightweight YOLOv8-pose	avg relative error = 4.14 %
[192]		RGB	Self Collected	YOLOX and PIDNet	AP50 = 90.3 %
[193]		RGB	Self Collected	ResNet	Loc error = 12.6 m
[194]		RGB	Self Collected	2chADCNN	Acc_summer = 94 %
[195]		RGB	default	tiny-YOLO	_
[196]		RGB	"Cityscapes" and PASCAL VOC	DPNet	mIoU = 78.69 % and 79.51 %
[197]		RGB	Synthetic Self Generated	DeepLabv3 ResNet101	mIoU = 61.7 %
[198]	2024	RGB	"VisDrone"	CSP-Darknet53, SPP and PANet	mAP = 34.52 %
[199]		RGB	Self Collected	YOLOv4 with MobileNetV3Small backbone	Precision = 90.48 %
[200]		RGB	Self Collected ("HID") and "VisDrone"	YOLO-IHD	mAP50 = 77.71 %
[153]		RGB	Self Collected	YOLOv5s-MTL	mAP = 64.8 %
[201]		RGB	Synthetic Self Generated	CNN	Acc = 72 %
[202]		RGB	Synthetic Self Generated	Faster-RCNN	Arrival rate = 95.2 %
[203]		RGB	Synthetic Self Generated	GoogleNet	_

We observed that one of the major research areas for UAV surveillance is air-to-ground, with the advantage of UAV allowing surveillance over a larger area. For example, in Ref. [115], researchers have trained Faster R-CNN for on-board detection of vehicles and pedestrians in aerial RGB images with an average precision of 88.33 %. This ability is employed in applications like intrusion monitoring utilizing YOLOv3 [116], achieving an accuracy of 92 % and building detection employing RetinaNet [12], reaching an accuracy of 88 % and multi-class classification of urban landscapes using optimal Squeezenet with a deep neural network (OSQN-DNN) [117] achieving an impressive accuracy of 98.99 %. Enhancing this ability [118], proposed using a method for robust detection of small and dense objects in aerial images. They have used a model based on cascade R-CNN and adopted the recursive feature pyramid and switchable atrous convolution for robust detection.

Researchers have mentioned applications of UAV surveillance in an urban setting so far. Expanding this into maritime surveillance [119], proposed extracting shoreline using image segmentation based on CNN, and the part of the segmented image that includes the detected shoreline is then fed into a CNN real-time optical flow estimator; this can then be used for obtaining tracking info. For object detection in maritime environments [120], employed YOLOv5 for identifying that there is an object in the frame with a 57 % mean average precision (for 0.5 thresholds), and [121] proposed Sharkeye, employing YOLO-like object detector for shark detection with a 94.52 % accuracy.

Another notable research direction is media production. For example [138,141], presented methods using YOLO and CNNs, respectively, for object-following applications that can aid in stable and autonomous cinematography. A better approach was developed by Ref. [125], with an aerial cinematography pipeline involving the UAV visually detecting the actor's motion using a Faster-RCNN + MobileNet-based model, amping the environment with the on-board LiDAR sensor, determining a safe, artistic, and low-occlusion trajectory. This approach reported an accuracy of 72.2 % and an MSE of 0.113. A similar approach is presented by Ref. [147]. However, the proposed method is for single UAV cinematography. For increased dimensions, the work of [136] suggested a multiple-UAV architecture for autonomous media production. They have employed CNN and GAN to process video streams from cameras onboard the UAVs.

3.4. Computer vision-based search and rescue for UAVs

Search and rescue (SAR) operations through UAVs involve autonomous navigation and detection of humans in crisis scenes, which may be in GNSS-denied, complex, natural, outdoor, and indoor environments. Computer Vision frameworks that are capable of handling the unprecedented and dynamic environments of a crisis scene are reviewed under this category. All the collected studies are compared and summarized in Table 5.

Indoor environments are challenging to navigate, especially during crises like house fires, because of the tight spaces, proximity to objects, and the risk of collapse. UAVs were employed to navigate the environment autonomously using various logic. One of the preliminary approaches for indoor navigation is to detect objects like doors and windows to both escape/enter and localize the UAV's position. The authors [152] have presented escape route detection, where a UAV uses CNN to identify windows, doors, and stairs in the building, and these objects were classified as valid or invalid with an average precision of 88.79 % for the path planning algorithm. Additionally [153], proposed an approach for the outdoor-indoor transition of UAVs by detecting damaged building openings using their YOLOv5s-MTL model with mAP of 64.8 %.

In recent times, the availability of a precise CAD model of a building has become common; this can be used during indoor SAR operations. Leveraging these readily available 3D CAD models provides valuable information about the environment of the UAV. Taking advantage of this, researchers [154] presented a method to localize a UAV using macro-feature detection and matching. The main contribution of this work is the real-time creation of the macro-feature description vector from the UAV captured images, which are simultaneously matched with an offline pre-existing vector from a CAD mode. This study employed YOLOv2 for macro-feature detection, achieving an accuracy of 77.4 %.

Utilizing the abundant presence of high-level semantic information in indoor environments [155,156],offered a method to better understand the environment for UAVs to improve the uncertainties in their pose estimation. In the former, the authors were able to estimate the 6DoF pose of the aerial robot while simultaneously creating a sparse semantic map of the environment. Semantic object detection was performed using Tiny-YOLOv2 with an ATE of 0.225. In the latter, the authors used a direct method to estimate the pose of the camera while simultaneously generating a 3D map of the environment using their 3D reconstruction system, and semantic segmentation was performed using BiSeNetV2, achieving 89 % accuracy.

In 2020, researchers [155] presented an approach to detect possible directions from pictures using a pre-trained VGG16 framework with 84 % accuracy. Later, in 2022, they improved their work through a method for exploring unknown indoor environments [157]. The newly presented multi-task learning method is capable of direction and low-level position prediction based only on the information from monocular images, achieving an accuracy of 91 %. Additionally, one of the cited works by Ref. [158] proposed a fully autonomous aerial robotics solution for executing complex SAR missions in an unstructured environment. They employed CNNs for target/background classification, achieving 97.8 % precision and an F1 score of 98.7 %.

Outdoor navigation has its challenges, such as search spanning through large areas, which will require efficient path planning and mapping in order to completely cover the area under diverse illumination settings and deal with complex natural environments.

To deal with obstacles in an outdoor environment, one of the cited works by Ref. [169] offered a real-time 3D path planning approach employing the YOLO framework to detect objects, including pedestrians, windows, electric poles, tunnels, trees, and barely visible nets; that it encounters during a disaster monitoring mission [182]. proposed an NN approach employing GLPDepth and Boosting Monocular Depth for monocular depth estimation networks, achieving an RMSE of 0.023, with the main idea of using depth maps for obstacle detection for a UAV system in complex natural environments with low altitude navigation.

The outdoor environment is large and dynamic, thus making it challenging for safe UAV navigation. To address this problem [171], proposed a solution using DLR with the CNN and RNN in the network structure with 40.5 % success. Another DLR approach using DQN with CNN for path planning was presented by Ref. [173], and this method reported a success rate of 94 %. However, this approach suggests employing a hybrid path planning that uses an anytime graph-based path planning algorithm for global planning and DLR for local planning. Additionally [175], proposed using the Drone-STM-RENet model based on the split transform merge concept. This model predicts collision probability and steering angle with a 96.26 % accuracy. It is essential to cover the whole unknown search area efficiently. To solve this [179], suggested employing a feature learning-enabled FCNN model that is developed to identify the edges of the workspace to be explored from the satellite images, achieving an average precision of 97.35 %, and then a CCPP path is generated.

A challenge unique to the outdoor environment is the change in season. We know that the environment's color, texture, etc., changes significantly between seasons (summer-winter). So, this calls for a model that is generalized to be season-invariant. Researchers [174] proposed a localization solution relying on matching UAV camera images to georeferenced orthophotos with a trained CNN model that is invariant to significant seasonal appearance differences between the camera and map. A similar template-matching approach for both UAV aerial images and satellite images is proposed by Ref. [194] here, the authors have used a two-channel deep convolutional neural network to achieve an accuracy of 94 % in the summer case. On the flip side, change in seasons takes months, but illumination conditions change drastically throughout the day; to tackle this [172], offered a method that employs a low-light image enhancement technique based on the Retinex theory. Their proposed FCNN architecture consists of two subnetworks, Decomposition-Net and Enhancement-Net, achieving an accuracy of 72.7 %.

Studies reviewed so far have proposed CV methods to navigate various challenging environments safely and effectively. However, another dimension of a SAR mission is target search. One of the cited works by Ref. [159] proposed an unsupervised approach for real-time human detection employing the Tiny-YOLOv3 model, with the model generating a mean average precision of 67 % and a recall of 70 %. A similar improved approach is presented by Ref. [200], using their YOLO-IHD model for indoor real-time human detection with a mAP50 of 77.71 %.

A similar problem is detecting humans in water bodies like lakes and seas. Work by Ref. [199] introduces a solution employing the YOLOv4 model with MobileNetV3Small backbone trained on the custom dataset, allowing the model to detect persons in water with 90.48 % precision. A more comprehensive method would be to make use of both Unmanned Surface Vehicles (USVs) and UAVs in a cooperative manner. This approach enhances USV's visual perception ability. This helps detect a person with cooperation and provides help to the person using the USV. For example, in the work done by Ref. [190], this approach employs the YOLOX model, PIDNet model, and distance measurement using a monocular camera, which is capable of detecting and classifying multiple objects surrounding USV. It can also differentiate between navigable and non-navigable regions for USV, along with the distance between the object and the USV, achieving an AP50 of 90.3 %.

4. Challenges and opportunities

The utilization of DL in computer vision for UAV autonomy has experienced a notable increase in recent years, indicating substantial advancements and prompting significant inquiries for future investigation. Since 2019, there have been many well-researched papers on deep learning-based computer vision approaches for UAV autonomy. However, it is important to conduct a systematic review to understand the connections between these papers, classify them into categories, and assess their ability to be reproduced. This review provides a comprehensive analysis of the current advancements in utilizing computer vision to improve the autonomy of UAVs, with a specific focus on navigation. It explores various research areas and highlights their notable contributions. Furthermore, it will emphasize the current difficulties and potential advantages in this field. To tackle these challenges, it is necessary to examine the accessibility of data, implement sophisticated AI techniques, overcome limitations in computational power, address concerns related to energy consumption, and optimize hardware. This review aims to offer clear guidance for future research, with the goal of enhancing the autonomy and efficiency of UAVs.

4.1. Data availability

The large quantity and quality of data collected from the visual sensors can be used to extract various details. Processes such as mapping the data to a multidimensional space and analyzing the visual images in various spectrums of light are now possible through deep learning like CNNs, DQNs, and DRL approaches. Better data means better performance of data-centric DL models. Work by Ref. [139] is a unique example of this increase in data availability, where authors used a self-collected dataset to develop a multi-scale video target detection algorithm based on YOLOv5 with regression method to recognize Shenzhou re-entry capsules and obtain positioning data. This method aids in the real-time search of re-entry capsules and guarantees astronauts' safety.

A major challenge is to derive a comparison among prior research, owing to variability in the datasets used. Most of the reviewed studies in these papers use self-collected datasets to evaluate their methods. This is not a favorable practice for improving collective research direction. Therefore, the authors recommend using open-published datasets along with their self-collected datasets. Just to name a few: VEPL dataset for power line inspection [204], FLAME dataset for aerial fire detection [205], VisDrone dataset for general object detection from aerial view [206], HelipadCat dataset for helipad detection [207], SEAGULL dataset for sea surveillance [208], IHD dataset for human detection [200], DoorDetect dataset for door detection [209] etc.

4.2. AI approaches

It has been observed the recent research is moving away from the usage of out-of-the-box YOLO models and towards custom models. However, SOTA approaches like transformers are underutilized in UAV computer vision, possibly due to the higher computational requirement of these models. Additionally, rather than only relying on image data for detection, classification, and localization tasks, researchers can utilize the increasingly available multimodal data using SOTA models like vision transformers [210,211]. One such notable example of multimodal information usage is [71], where the authors have integrated multimodal information from novel sensors, including optical cameras and 3D light detection and ranging.

AdaViT presents an adaptive computational framework that improves the efficiency of vision transformers by selectively utilizing patches, self-attention heads, and transformer blocks. This approach demonstrates a more than twofold enhancement in efficiency while simultaneously preserving a minimal decrease in accuracy of 0.8 % [212]. CrossViT introduces a dual-branch transformer architecture that integrates image patches of varying sizes to enhance the representation of features. The utilization of a cross-attention module enables the effective integration of multi-scale tokens, resulting in superior performance compared to previous approaches like DeiT, with a 2 % improvement on the ImageNet1K dataset [213].

Additionally, Federated Learning, an innovative decentralized machine learning paradigm, has recently garnered significant attention due to its potential to revolutionize UAV applications. With this innovative method, model training is carried out cooperatively across numerous edge devices, such as UAVs, without the need for raw data exchange or simply transmitting model updates. This reduces privacy and security concerns while allowing UAVs to actively participate in the learning process because of their onboard computing capabilities. FL may be combined with any AI system for autonomous UAV navigation, and leveraging central learning lowers the complexity in both space and time. However, according to our search, FL for UAVs hasn't been researched well. In 2022 [134,214], proposed employing FL in the image classification stage. This indicates a new research direction, which has the potential to improve real-time decision-making, optimize resource-constrained operations, and enable adaptive and autonomous missions. Additionally, this technique promotes effective information sharing among UAVs operating in various situations, promoting resilience and generalization.

4.3. Computational power

Deep Learning algorithms demand extensive computational power to operate. Considering that UAVs have limited payload capacity, this calls for efficiency improvement on both the hardware and software side. On the hardware side, companies have made significant contributions to improving the on-board UAV computation unit size and efficiency [215]. On the software side, models such as MobileNet were developed that require significantly less computational power to run while only compromising very little accuracy [216].

With the onset of strong satellite communication, there is even little need for UAVs to carry out extensive computing on-board [217]. It is possible to compute on remote servers and transfer the results to the UAV within a certain timeframe, which is critical to the mission [218]. This way, UAVs can carry more useful payloads and reduce huge sensors that occupy space and weight.

4.4. Energy consumption

Batteries are the sole energy source used to power the hardware and, thus, the software onboard the UAV. Therefore, it is essential to efficiently manage the energy availability throughout the mission, especially for long-range missions such as SAR over uninhabited areas. However, no significant studies have been conducted in this domain in the past five years. By leveraging DL algorithms, UAVs can efficiently monitor and optimize battery performance, capacity, and health, thereby enhancing their overall flight endurance, mission range, and operational reliability [219]. AI battery management involves real-time data analysis from various sensors, including camera, voltage, current, temperature, and impedance, to predict battery degradation, assess state-of-charge, and estimate remaining flight time accurately. Using a visual sensor such as a camera, a UAV can perceive its surroundings and determine the most efficient flight path for full area coverage, considering the battery capacity.

This data-driven approach allows UAVs to dynamically adjust their flight parameters, such as speed, altitude, and payload, to maximize battery utilization and extend mission duration while ensuring safe and sustainable battery operation. As AI battery management continues to advance, we can anticipate UAVs becoming more self-aware and adaptive, enabling longer and more complex missions, such as extended surveillance, remote inspections, and aerial data collection, with reduced downtime and minimal human intervention [220].

4.5. Hardware

UAVs now have robust yet reasonably priced camera modules and Inertial Measurement Units (IMUs) at their disposal, owing to developments in embedded systems technology. These modules can retrieve essential data from the device and deliver it back to the user, along with inertial sensor measurements.

Deploying computer vision systems on UAV hardware platforms poses various practical difficulties, especially when taking into account limited computational resources, power limitations, and system resilience. The speed at which data is processed in CV systems is crucial, especially during onboard inference. These systems often have strict latency requirements that can put a strain on their limited computational capabilities. The effectiveness of these systems depends on their capacity to strike a balance between processing speed and accuracy, a task that is frequently hindered by the limitations of UAV hardware. Power efficiency is a major concern when it comes to UAVs [221]. These unmanned aerial vehicles usually have a limited battery life, so it is crucial to optimize the energy consumption of computer vision algorithms. Optimizing algorithms to minimize computational load while maintaining performance, along with advancements in energy-efficient hardware, is necessary for this task [222–225]. Furthermore, the resilience of CV systems is assessed in practical situations where different environmental conditions and operational pressures may result in unforeseen malfunctions. These factors encompass hardware malfunctions, software errors, or performance degradation in highly demanding circumstances [226,227]. Therefore, it is essential to assess and address these potential sources of failure through thorough testing and validation. To tackle these challenges, it is necessary to adopt a comprehensive strategy that combines efficient algorithms with strong hardware solutions. This approach guarantees dependable performance while effectively handling the limitations that are inherent to UAV platforms [228–231]. This ultimately enhances the operational effectiveness and reliability of UAVs in various applications.

5. Conclusion

In this SLR, studies are identified and reviewed, focusing on Deep Learning Computer Vision approaches for autonomous UAV navigation. A systematic search string is utilized to collect all relevant literature indexed on Scopus. Subsequently, the collected studies are further filtered and organized into the identified categories. Sensing and inspection, landing, surveillance and tracking, and search and rescue are identified as four primary UAV applications that extensively extract visual information from images. These four categories are thoroughly reviewed and analyzed, with mention of the type of optical sensor onboard, the dataset used for model development, the methodology employed, and the performance achieved for each cited study. It is observed that previous studies have made significant use of transfer learning and efficient models, such as YOLO, with YOLO alone accounting for over 39.5 % of usage. Additionally, the majority of studies employ self-collected datasets, which presents a bottleneck for computer vision in UAV research. This challenge can only be addressed by promoting open data sharing and establishing more benchmark datasets. Few such benchmark datasets are mentioned in this study. Furthermore, a suggestion is made to consider SOTA approaches like vision transformers to capitalize on multimodal data, decentralized training approaches like federated learning, efficient models like MobileNet, and battery management using AI and hardware, including DL model optimization for UAVs. All five research questions are comprehensively addressed through this SLR approach.

CRediT authorship contribution statement

Aditya Vardhan Reddy Katkuri: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Hakka Madan:** Writing – original draft, Validation, Investigation, Conceptualization. **Narendra Khatri:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization. **Antar Shaddad Hamed Abdul-Qawy:** Writing – review & editing, Conceptualization. **K. Sridhar Patnaik:** Writing – review & editing, Conceptualization.

Declaration of competing 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.

Data availability

No data was used for the research described in the article.

References

[1] Diez Y, Kentsch S, Fukuda M, Caceres MLL, Moritake K, Cabezas M. Deep learning in forestry using UAV-acquired RGB data: a practical review. Rem Sens 2021;13:2837. https://doi.org/10.3390/RS13142837. 2021;13:2837.
[2] Mogili UR, Deepak BBVL. Review on application of drone systems in precision agriculture. Procedia Comput Sci 2018;133:502–9. https://doi.org/10.1016/J.PROCS.2018.07.063.
[3] Tsouros DC, Bibi S, Sarigiannidis PG. A review on UAV-based applications for precision agriculture. Information 2019;10:349. https://doi.org/10.3390/INFO10110349. 2019;10:349.
[4] Arafat MY, Alam MM, Moh S. Vision-based navigation techniques for unmanned aerial vehicles: review and challenges. Drones 2023;7:89. https://doi.org/10.3390/DRONES7020089. 2023;7:89.
[5] Minaeian S, Liu J, Son YJ. Vision-based target detection and localization via a team of cooperative UAV and UGVs. IEEE Trans Syst Man Cybern Syst 2016;46:1005–16. https://doi.org/10.1109/TSMC.2015.2491878.
[6] Herisse B, Russotto FX, Hamel T, Mahony R. Hovering flight and vertical landing control of a VTOL Unmanned Aerial Vehicle using optical flow. In: IEEE/RSJ international conference on intelligent robots and systems, IROS; 2008. p. 801–6. https://doi.org/10.1109/IROS.2008.4650731. 2008.
[7] Byun HJ, Nam H. Autonomous control of unmanned aerial vehicle for chemical detection using deep reinforcement learning. Electron Lett 2022;58:423–5. https://doi.org/10.1049/ELL2.12485.

[8] Fragassa C, Vitali G, Emmi L, Arru M. A new procedure for combining UAV-based imagery and machine learning in precision agriculture. Sustainability 2023;15: 998. https://doi.org/10.3390/SU15020998. 2023;15:998.

[9] Feng J, Sun Y, Zhang K, Zhao Y, Ren Y, Chen Y, et al. Autonomous detection of spodoptera frugiperda by feeding symptoms directly from UAV RGB imagery. Appl Sci 2022;12:2592. https://doi.org/10.3390/APP12052592. 2022;12:2592.

[10] Narazaki Y, Hoskere V, Chowdhary G, Spencer BF. Vision-based navigation planning for autonomous post-earthquake inspection of reinforced concrete railway viaducts using unmanned aerial vehicles. Autom ConStruct 2022;137: 104214. https://doi.org/10.1016/J.AUTCON.2022.104214.

[11] Liu Y, Li H, Wang L, Ai J. Deep Learning Approach to Drogue Detection for Fixed-Wing UAV Autonomous Aerial Refueling with Visual Camera 2023:827–34. https://doi.org/10.1109/ICUAS57906.2023.10156472.

[12] Boone J, Goodin C, Dabbiru L, Hudson C, Cagle L, Carruth D. Training artificial intelligence algorithms with automatically labelled UAV data from physics-based simulation software. Appl Sci 2023;13:131. https://doi.org/10.3390/APP13010131. 2022;13:131.

[13] Lu Y, Xue Z, Xia GS, Zhang L. A survey on vision-based UAV navigation 2018;21: 21–32. https://doi.org/10.1080/10095020.2017.1420509.

[14] Zhao Y, Zheng Z, Liu Y. Survey on computational-intelligence-based UAV path planning. Knowl Base Syst 2018;158:54–64. https://doi.org/10.1016/J.KNOSYS.2018.05.033.

[15] Belmonte LM, Morales R, Fernández-Caballero A. Computer vision in autonomous unmanned aerial vehicles—a systematic mapping study. Appl Sci 2019;9:3196. https://doi.org/10.3390/APP9153196. 2019;9:3196.

[16] Lee T, McKeever S, Courtney J. Flying free: a research overview of deep learning in drone navigation autonomy. Drones 2021;5:52. https://doi.org/10.3390/DRONES5020052/S1.

[17] AlMahamid F, Grolinger K. Autonomous unmanned aerial vehicle navigation using reinforcement learning: a systematic review. Eng Appl Artif Intell 2022; 115:105321. https://doi.org/10.1016/J.ENGAPPAI.2022.105321.

[18] Rezwan S, Choi W. Artificial intelligence approaches for UAV navigation: recent advances and future challenges. IEEE Access 2022;10:26320–39. https://doi.org/10.1109/ACCESS.2022.3157626.

[19] Mondal TG, Chen G. Artificial intelligence in civil infrastructure health monitoring—historical perspectives, current trends, and future visions. Front Built Environ 2022;8:1007886. https://doi.org/10.3389/FBUIL.2022.1007886/BIBTEX.

[20] Puente-Castro A, Rivero D, Pazos A, Fernandez-Blanco E. A review of artificial intelligence applied to path planning in UAV swarms. Neural Comput Appl 2022; 34:153–70. https://doi.org/10.1007/S00521-021-06569-4/FIGURES/7.

[21] Ul Husnain A, Mokhtar N, Shah NM, Dahari M, Iwahashi M. A systematic literature review (SLR) on autonomous path planning of unmanned aerial vehicles. Drones 2023;7:118. https://doi.org/10.3390/DRONES7020118. 2023; 7:118.

[22] Ali B, Sadekov RN, Tsodokova VV. A review of navigation algorithms for unmanned aerial vehicles based on computer vision systems. Gyroscopy and Navigation 2022;13:241–52. https://doi.org/10.1134/S2075108722040022/TABLES/3.

[23] Mohyuddin G, Khan MA, Haseeb A, Mahpara S, Waseem M, Saleh AM. Evaluation of machine learning approaches for precision farming in smart agriculture system: a comprehensive review. IEEE Access 2024;12:60155–84. https://doi.org/10.1109/ACCESS.2024.3390581.

[24] Xu C, Zhao W, Zhao J, Guan Z, Song X, Li J. Uncertainty-aware multiview deep learning for internet of things applications. IEEE Trans Ind Inf 2023;19:1456–66. https://doi.org/10.1109/TII.2022.3206343.

[25] O'Shea K, Nash R. An introduction to convolutional neural networks. Int J Res Appl Sci Eng Technol 2015;10:943–7. https://doi.org/10.22214/ijraset.2022.47789.

[26] Mantau AJ, Widayat IW, Leu JS, Köppen M. A human-detection method based on YOLOv5 and transfer learning using thermal image data from UAV perspective for surveillance system. Drones 2022;6:290. https://doi.org/10.3390/DRONES6100290. 2022;6:290.

[27] Al-Qubaydhi N, Alenezi A, Alanazi T, Senyor A, Alanezi N, Alotaibi B, et al. Detection of unauthorized unmanned aerial vehicles using YOLOv5 and transfer learning. Electronics 2022;11:2669. https://doi.org/10.3390/ELECTRONICS11172669. 2022;11:2669.

[28] Xu C, Zhao W, Zhao J, Guan Z, Song X, Li J. Uncertainty-aware multiview deep learning for internet of things applications. IEEE Trans Ind Inf 2023;19:1456–66. https://doi.org/10.1109/TII.2022.3206343.

[29] Liu Y, Xu C, Chen L, Yan M, Zhao W, Guan Z. TABLE: time-aware balanced multi-view learning for stock ranking. Knowl Base Syst 2024:112424. https://doi.org/10.1016/J.KNOSYS.2024.112424.

[30] Aliyari M, Droguett EL, Ayele YZ. UAV-based bridge inspection via transfer learning. Sustainability 2021;13:11359. https://doi.org/10.3390/SU132011359. 2021;13:11359.

[31] Liu K, Tang H, He S, Yu Q, Xiong Y, Wang N. Performance validation of yolo variants for object detection. In: Proceedings of the 2021 international conference on bioinformatics and intelligent computing. BIC; 2021. p. 239–43. https://doi.org/10.1145/3448748.3448786. 2021.

[32] Jiang P, Ergu D, Liu F, Cai Y, Ma B. A review of yolo algorithm developments. Procedia Comput Sci 2022;199:1066–73. https://doi.org/10.1016/J.PROCS.2022.01.135.

[33] Nepal U, Eslamiat H. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. Sensors 2022;22:464. https://doi.org/10.3390/S22020464/S1.

[34] Gongal A, Amatya S, Karkee M, Zhang Q, Lewis K. Sensors and systems for fruit detection and localization: a review. Comput Electron Agric 2015;116:8–19. https://doi.org/10.1016/J.COMPAG.2015.05.021.

[35] Wang Q, Cheng M, Huang S, Cai Z, Zhang J, Yuan H. A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed Solanum rostratum Dunal seedlings. Comput Electron Agric 2022;199:107194. https://doi.org/10.1016/j.compag.2022.107194.

[36] India: Road transport passengers | Statista n.d.

[37] Hassan SA, Rahim T, Shin SY. An improved deep convolutional neural network-based autonomous road inspection scheme using unmanned aerial vehicles. Electronics 2021;10:2764. https://doi.org/10.3390/ELECTRONICS10222764. 2021;10:2764.

[38] Zhao Z, Han J, Song L. YOLO-highway: an improved highway center marking detection model for unmanned aerial vehicle autonomous flight. Math Probl Eng 2021;2021. https://doi.org/10.1155/2021/1205153.

[39] Ayele YZ, Aliyari M, Griffths D, Droguett EL. Automatic crack segmentation for UAV-assisted bridge inspection. Energies 2020;13:6250. https://doi.org/10.3390/EN13236250. 2020;13:6250.

[40] Mirzazade A, Popescu C, Gonzalez-Libreros J, Blanksvärd T, Täljsten B, Sas G. Semi-autonomous inspection for concrete structures using digital models and a hybrid approach based on deep learning and photogrammetry. J Civ Struct Health Monit 2023;13:1633–52. https://doi.org/10.1007/S13349-023-00680-X/FIGURES/23.

[41] da Silva YMR, Andrade FAA, Sousa L, de Castro GGR, Dias JT, Berger G, et al. Computer vision based path following for autonomous unmanned aerial systems in unburied pipeline onshore inspection. Drones 2022;6:410. https://doi.org/10.3390/DRONES6120410. 2022;6:410.

[42] Yang H, Li X, Guo Y, Jia L. RT-GAN: GAN based architecture for precise segmentation of railway tracks. Appl Sci 2022;12:12044. https://doi.org/10.3390/APP122312044. 2022;12:12044.

[43] Li X, Guo Y, Yang H, Ye Q, Jia L. DWPIS: dynamic-weight parallel instance and skeleton network for railway centerline detection. Appl Sci 2023;13:6133. https://doi.org/10.3390/APP13106133. 2023;13:6133.

[44] Filippo M De, Asadiabadi S, Kuang JS, Mishra DK, Sun H. Ai-powered inspections of facades in reinforced concrete buildings. HKIE Transactions Hong Kong Institution of Engineers 2023;30:1–14. https://doi.org/10.33430/V30N1THIE-2020-0023.

[45] Mo DH, Wu YC, Lin CS. The dynamic image analysis of retaining wall crack detection and gap hazard evaluation method with deep learning. Appl Sci 2022; 12:9289. https://doi.org/10.3390/APP12189289. 2022;12:9289.

[46] Bian J, Hui X, Zhao X, Tan M. A monocular vision–based perception approach for unmanned aerial vehicle close proximity transmission tower inspection. Int J Adv Rob Syst 2019;16. https://doi.org/10.1177/1729881418820227/ASSET/IMAGES/LARGE/10.1177_1729881418820227-FIG16.JPEG.

[47] Hui X, Bian J, Zhao X, Tan M. A monocular-based navigation approach for unmanned aerial vehicle safe and autonomous transmission-line inspection. Int J Adv Rob Syst 2019;16. https://doi.org/10.1177/1729881419829941/ASSET/IMAGES/LARGE/10.1177_1729881419829941-FIG18.JPEG.

[48] Zhang Z, Boubin J, Stewart C, Khanal S. Whole-field reinforcement learning: a fully autonomous aerial scouting method for precision agriculture. Sensors 2020; 20:6585. https://doi.org/10.3390/S20226585. 2020;20:6585.

[49] Mondal TG, Jahanshahi MR, Wu RT, Wu ZY. Deep learning-based multi-class damage detection for autonomous post-disaster reconnaissance. Struct Control Health Monit 2020;27:e2507. https://doi.org/10.1002/STC.2507.

[50] Teixeira L, Oswald MR, Pollefeys M, Chli M. Aerial single-view depth completion with image-guided uncertainty estimation. IEEE Rob Autom Lett 2020;5: 1055–62. https://doi.org/10.1109/LRA.2020.2967296.

[51] Nguyen VN, Jenssen R, Roverso D. LS-Net: fast single-shot line-segment detector. Mach Vis Appl 2021;32:1–16. https://doi.org/10.1007/S00138-020-01138-6/TABLES/6.

[52] Bah MD, Hafiane A, Canals R. CRowNet: deep network for crop row detection in UAV images. IEEE Access 2020;8:5189–200. https://doi.org/10.1109/ACCESS.2019.2960873.

[53] Lopez RL, Sanchez MJB, Jimenez MP, Arrue BC, Ollero A. Autonomous uav system for cleaning insulators in power line inspection and maintenance. Sensors 2021;21:8488. https://doi.org/10.3390/S21248488/S1.

[54] Luo C, Yu L, Yan J, Li Z, Ren P, Bai X, et al. Autonomous detection of damage to multiple steel surfaces from 360° panoramas using deep neural networks. Comput Aided Civ Infrastruct Eng 2021;36:1585–99. https://doi.org/10.1111/MICE.12686.

[55] Khelifi A, Ciccone G, Altaweel M, Basmaji T, Ghazal M. Autonomous service drones for multimodal detection and monitoring of archaeological sites. Appl Sci 2021;11:10424. https://doi.org/10.3390/APP112110424. 2021;11:10424.

[56] Ayoub N, Schneider-Kamp P. Real-time on-board deep learning fault detection for autonomous UAV inspections. Electronics 2021;10:1091. https://doi.org/10.3390/ELECTRONICS10091091. 2021;10:1091.

[57] Blekos K, Tsakas A, Xouris C, Evdokidis I, Alexandropoulos D, Alexakos C, et al. Analysis, modeling and multi-spectral sensing for the predictive management of verticillium wilt in olive groves. J Sens Actuator Netw 2021;10:15. https://doi.org/10.3390/JSAN10010015. 2021;10:15.

[58] Rahman EU, Zhang Y, Ahmad S, Ahmad HI, Jobaer S. Autonomous vision-based primary distribution systems porcelain insulators inspection using UAVs. Sensors 2021;21:974. https://doi.org/10.3390/S21030974. 2021;21:974.

[59] Rosende SB, Sánchez-Soriano J, Muñoz CQG, Andrés JF. Remote management architecture of UAV fleets for maintenance, surveillance, and security tasks in

solar power plants. Energies 2020;13:5712. https://doi.org/10.3390/EN13215712. 2020;13:5712.
[60] Morando L, Recchiuto CT, Calla J, Scuteri P, Sgorbissa A. Thermal and visual tracking of photovoltaic plants for autonomous UAV inspection. Drones 2022;6:347. https://doi.org/10.3390/DRONES6110347. 2022;6:347.
[61] de Figueiredo RP, Sejersen JLF, Hansen JG, Brandão M. Integrated design-sense-plan architecture for autonomous geometric-semantic mapping with UAVs. Front Robot AI 2022;9:911974. https://doi.org/10.3389/FROBT.2022.911974/BIBTEX.
[62] Li H, Dong Y, Liu Y, Ai J. Design and implementation of UAVs for bird's nest inspection on transmission lines based on deep learning. Drones 2022;6. https://doi.org/10.3390/DRONES6090252. Page 252 2022;6:252.
[63] Dutta AK, Albagory Y, Sait ARW, Keshta IM. Autonomous unmanned aerial vehicles based decision support system for weed management. Comput Mater Continua (CMC) 2022;73:899–915. https://doi.org/10.32604/CMC.2022.026783.
[64] Safonova A, Hamad Y, Alekhina A, Kaplun D. Detection of Norway spruce trees (picea abies) infested by bark beetle in UAV images using YOLOs architectures. IEEE Access 2022;10:10384–92. https://doi.org/10.1109/ACCESS.2022.3144433.
[65] Idrissi M, Hussain A, Barua B, Osman A, Abozariba R, Aneiba A, et al. Evaluating the forest ecosystem through a semi-autonomous quadruped robot and a hexacopter UAV. Sensors 2022;22:5497. https://doi.org/10.3390/S22155497. 2022;22:5497.
[66] Jacobsen RH, Matlekovic L, Shi L, Malle N, Ayoub N, Hageman K, et al. Design of an autonomous cooperative drone swarm for inspections of safety critical infrastructure. Appl Sci 2023;13:1256. https://doi.org/10.3390/APP13031256/S1.
[67] Pinheiro I, Aguiar A, Figueiredo A, Pinho T, Valente A, Santos F. Nano aerial vehicles for tree pollination. Appl Sci 2023;13:4265. https://doi.org/10.3390/APP13074265. 2023;13:4265.
[68] Tsellou A, Livanos G, Ramnalis D, Polychronos V, Plokamakis G, Zervakis M, et al. A UAV intelligent system for Greek power lines monitoring. Sensors 2023;23:8441. https://doi.org/10.3390/S23208441. 2023;23:8441.
[69] Hickling T, Aouf N, Spencer P. Robust adversarial attacks detection based on explainable deep reinforcement learning for UAV guidance and planning. IEEE Transactions on Intelligent Vehicles 2023;8:4381–94. https://doi.org/10.1109/TIV.2023.3296227.
[70] Ramadan MNA, Basmaji T, Gad A, Hamdan H, Akgün BT, Ali MAH, et al. Towards early forest fire detection and prevention using AI-powered drones and the IoT. Internet of Things 2024:101248. https://doi.org/10.1016/J.IOT.2024.101248.
[71] Jeon M, Moon J, Jeong S, Oh KY. Autonomous flight strategy of an unmanned aerial vehicle with multimodal information for autonomous inspection of overhead transmission facilities. Comput Aided Civ Infrastruct Eng 2024;39:2159–86. https://doi.org/10.1111/MICE.13188.
[72] Melnychenko O, Scislo L, Savenko O, Sachenko A, Radiuk P. Intelligent integrated system for fruit detection using multi-UAV imaging and deep learning. Sensors 2024;24:1913. https://doi.org/10.3390/S24061913. 2024;24:1913.
[73] Chen Y, Song X. Unmanned Aerial Vehicle Fire Detection Platform Based on Semantic Yolov5 and Autonomous Recognition. Https://ServicesIgi-GlobalCom/Resolvedoi/ResolveAspx?Doi=104018/IJSWIS344026 1AD;20:1–26 https://doi.org/10.4018/IJSWIS.344026.
[74] Maitre G, Martinot D, Tuci E. On the design of deep learning-based control algorithms for visually guided UAVs engaged in power tower inspection tasks. Front Robot AI 2024;11:1378149. https://doi.org/10.3389/FROBT.2024.1378149/BIBTEX.
[75] Song F, Sun Y, Yuan G. Autonomous identification of bridge concrete cracks using unmanned aircraft images and improved lightweight deep convolutional networks. Struct Control Health Monit 2024;2024:7857012. https://doi.org/10.1155/2024/7857012.
[76] Mathur P, Sharma C, Azeemuddin S. Autonomous inspection of high-rise buildings for façade detection and 3D modeling using UAVs. IEEE Access 2024;12:18251–8. https://doi.org/10.1109/ACCESS.2024.3360209.
[77] Jin R, Owais HM, Lin D, Song T, Yuan Y. Ellipse proposal and convolutional neural network discriminant for autonomous landing marker detection. J Field Robot 2019;36:6–16. https://doi.org/10.1002/ROB.21814.
[78] Lee S, Jo D, Kwon Y. Camera-based automatic landing of drones using artificial intelligence image recognition. International Journal of Mechanical Engineering and Robotics Research 2022;11:357–64. https://doi.org/10.18178/IJMERR.11.5.357-364.
[79] Polvara R, Patacchiola M, Hanheide M, Neumann G. Sim-to-Real quadrotor landing via sequential deep Q-networks and domain randomization. Robotics 2020;9. https://doi.org/10.3390/ROBOTICS9010008. Page 8 2020;9:8.
[80] Lin S, Jin L, Chen Z. Real-time monocular vision system for UAV autonomous landing in outdoor low-illumination environments. Sensors 2021;21:6226. https://doi.org/10.3390/S21186226. 2021;21:6226.
[81] Wang Z, Zhao D, Cao Y, Wang Z, Zhao D, Cao Y. Visual navigation algorithm for night landing of fixed-wing unmanned aerial vehicle. Aerospace 2022;9:615. https://doi.org/10.3390/AEROSPACE9100615. 2022;9:615.
[82] Safadinho D, Ramos J, Ribeiro R, Filipe V, Barroso J, Pereira A. UAV landing using computer vision techniques for human detection. Sensors 2020;20:613. https://doi.org/10.3390/S20030613. 2020;20:613.
[83] Liu F, Shan J, Xiong B, Fang Z. A real-time and multi-sensor-based landing area recognition system for UAVs. Drones 2022;6:118. https://doi.org/10.3390/DRONES6050118. 2022;6:118.
[84] Nepal U, Eslamiat H. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. Sensors 2022;22:464. https://doi.org/10.3390/S22020464/S1.
[85] Kikumoto C, Harimoto Y, Isogaya K, Yoshida T, Urakubo T. Landing site detection for UAVs based on CNNs classification and optical flow from monocular camera images. J Robot Mechatron 2021;33:292–300. https://doi.org/10.20965/JRM.2021.P0292.
[86] Alorf A. A robust deep learning model for terrain slope estimation. Int J Adv Comput Sci Appl 2024;15:1231–45. https://doi.org/10.14569/IJACSA.2024.01501121.
[87] Truong NQ, Lee YW, Owais M, Nguyen DT, Batchuluun G, Pham TD, et al. SlimDeblurGAN-based motion deblurring and marker detection for autonomous drone landing. Sensors 2020;20:3918. https://doi.org/10.3390/S20143918. 2020;20:3918.
[88] Castellano G, Castiello C, Mencar C, Vessio G. Crowd detection in aerial images using spatial graphs and fully-convolutional neural networks. IEEE Access 2020;8:64534–44. https://doi.org/10.1109/ACCESS.2020.2984768.
[89] Gonzalez-Trejo J, Mercado-Ravell D, Becerra I, Murrieta-Cid R. On the visual-based safe landing of UAVs in populated areas: a crucial aspect for urban deployment. IEEE Rob Autom Lett 2021;6:7902–9. https://doi.org/10.1109/LRA.2021.3101861.
[90] Igonin DM, Kolganov PA, Tiumentsev YV. Situational awareness and problems of its formation in the tasks of uav behavior control. Appl Sci 2021;11. https://doi.org/10.3390/app112411611.
[91] Zhao J, Sun J, Cai Z, Wang L, Wang Y. End-to-End deep reinforcement learning for image-based UAV autonomous control. Appl Sci 2021;11:8419. https://doi.org/10.3390/APP11188419. 2021;11:8419.
[92] Li M, Hu T. Deep learning enabled localization for UAV autolanding. Chin J Aeronaut 2021;34:585–600. https://doi.org/10.1016/J.CJA.2020.11.011.
[93] Dong X, Gao Y, Guo J, Zuo S, Xiang J, Li D, et al. An integrated UWB-IMU-Vision framework for autonomous approaching and landing of UAVs. Aerospace 2022;9:797. https://doi.org/10.3390/AEROSPACE9120797. 2022;9:797.
[94] Wang L, Jiang X, Wang D, Tu Z, Ai J, Huang Y, et al. Research on aerial autonomous docking and landing technology of dual multi-rotor UAV. Sensors 2022;22:9066. https://doi.org/10.3390/S22239066. 2022;22:9066.
[95] Xu Y, Zhong D, Zhou J, Jiang Z, Zhai Y, Ying Z. A novel UAV visual positioning algorithm based on A-YOLOX. Drones 2022;6:362. https://doi.org/10.3390/DRONES6110362. 2022;6:362.
[96] Ma N, Weng X, Cao Y, Wu L. Monocular-vision-based precise runway detection applied to state estimation for carrier-based UAV landing. Sensors 2022;22:8385. https://doi.org/10.3390/S22218385. 2022;22:8385.
[97] Arrahmah AI, Rahmania R, Saputra DE. Comparison between convolutional neural network and K-nearest neighbours object detection for autonomous drone. Bulletin of Electrical Engineering and Informatics 2022;11:2303–12. https://doi.org/10.11591/EEI.V11I4.3784.
[98] Shadakshri VHC, Veena MB, Dev VKRG. OpenCV implementation of grid-based vertical safe landing for UAV using YOLOv5. Int J Adv Comput Sci Appl 2022;13:500–6. https://doi.org/10.14569/IJACSA.2022.0130957.
[99] Yuan B, Ma W, Wang F. High speed safe autonomous landing marker tracking of fixed wing drone based on deep learning. IEEE Access 2022;10:80415–36. https://doi.org/10.1109/ACCESS.2022.3195286.
[100] Chang CW, Lo LY, Cheung HC, Feng Y, Yang AS, Wen CY, et al. Proactive guidance for accurate UAV landing on a dynamic platform: a visual–inertial approach. Sensors 2022;22:404. https://doi.org/10.3390/S22010404. 2022;22:404.
[101] Lee B, Saj V, Kalathil D, Benedict M. Intelligent vision-based autonomous ship landing of VTOL UAVs. ArXiv 2022;68. https://doi.org/10.4050/JAHS.68.022010.
[102] Santos NP, Lobo V, Bernardino A. Fixed-wing unmanned aerial vehicle 3D-model-based tracking for autonomous landing. Drones 2023;7:243. https://doi.org/10.3390/DRONES7040243. 2023;7:243.
[103] Cui Q, Liu M, Huang X, Gao M. Coarse-to-fine visual autonomous unmanned aerial vehicle landing on a moving platform. Biomimetic Intelligence and Robotics 2023;3:100088. https://doi.org/10.1016/J.BIROB.2023.100088.
[104] Wang B, Ma R, Zhu H, Sha Y, Yang T. An autonomous tracking and landing method for unmanned aerial vehicles based on visual navigation. Drones 2023;7:703. https://doi.org/10.3390/DRONES7120703. 2023;7:703.
[105] Mu L, Li Q, Wang B, Zhang Y, Feng N, Xue X, et al. A vision-based autonomous landing guidance strategy for a micro-UAV by the modified camera view. Drones 2023;7:400. https://doi.org/10.3390/DRONES7060400. 2023;7:400.
[106] Subramanian JA, Asirvadam VS, Zulkifli SABM, Singh NSS, Shanthi N, Lagisetty RK, et al. Integrating computer vision and photogrammetry for autonomous aerial vehicle landing in static environment. IEEE Access 2024;12:4532–43. https://doi.org/10.1109/ACCESS.2024.3349419.
[107] Balasooriya NM, de Silva O, Jayasiri A, Mann GKI. AI-based landing zone detection for vertical takeoff and land LiDAR localization and mapping pipelines. Drone Systems and Applications 2024;12:1–19. https://doi.org/10.1139/DSA-2022-0038/ASSET/IMAGES/DSA-2022-0038_F28.JPG.
[108] Pieczyński D, Ptak B, Kraft M, Piechocki M, Aszkowski P. A fast, lightweight deep learning vision pipeline for autonomous UAV landing support with added robustness. Eng Appl Artif Intell 2024;131:107864. https://doi.org/10.1016/J.ENGAPPAI.2024.107864.
[109] Çetin E, Barrado C, Pastor E. Improving real-time drone detection for counter-drone systems. Aeronaut J 2021;125:1871–96. https://doi.org/10.1017/AER.2021.43.

[110] Rodriguez-Ramos A, Alvarez-Fernandez A, Bavle H, Campoy P, How JP. Vision-based multirotor following using synthetic learning techniques. Sensors 2019;19:4794. https://doi.org/10.3390/S19214794. 2019;19:4794.
[111] Sun Y, Zhi X, Han H, Jiang S, Shi T, Gong J, et al. Enhancing UAV detection in surveillance camera videos through spatiotemporal information and optical flow. Sensors 2023;23:6037. https://doi.org/10.3390/S23136037. 2023;23:6037.
[112] Ajakwe SO, Ihekoronye VU, Kim DS, Lee JM. ALIEN: assisted learning invasive encroachment neutralization for secured drone transportation system. Sensors 2023;23:1233. https://doi.org/10.3390/S23031233. 2023;23:1233.
[113] Othman NA, Aydin I. Development of a novel lightweight CNN model for classification of human actions in UAV-captured videos. Drones 2023;7:148. https://doi.org/10.3390/DRONES7030148. 2023;7:148.
[114] Wang L, Ai J, Zhang L, Xing Z. Design of airport obstacle-free zone monitoring UAV system based on computer vision. Sensors 2020;20:2475. https://doi.org/10.3390/S20092475. 2020;20:2475.
[115] Kim J, Cho J. RGDiNet: efficient onboard object detection with faster R-CNN for air-to-ground surveillance. Sensors 2021;21:1677. https://doi.org/10.3390/S21051677. 2021;21:1677.
[116] Rodriguez-Gomez JP, Eguiluz AG, Dios JRM-D, Ollero A. Auto-tuned event-based perception scheme for intrusion monitoring with UAS. IEEE Access 2021;9:44840–54. https://doi.org/10.1109/ACCESS.2021.3066529.
[117] Minu MS, Canessane RA, Ramesh SSS. Optimal squeeze net with deep neural network-based arial image classification model in unmanned aerial vehicles. Trait Du Signal 2022;39:275–81. https://doi.org/10.18280/TS.390128.
[118] Lee JC, Yoo JY, Kim Y, Moon ST, Ko JH. Robust detection of small and dense objects in images from autonomous aerial vehicles. Electron Lett 2021;57:611–3. https://doi.org/10.1049/ELL2.12245.
[119] Aspragkathos SN, Karras GC, Kyriakopoulos KJ. A hybrid model and data-driven vision-based framework for the detection, tracking and surveillance of dynamic coastlines using a multirotor UAV. Drones 2022;6:146. https://doi.org/10.3390/DRONES6060146. 2022;6:146.
[120] Vasilopoulos E, Vosinakis G, Krommyda M, Karagiannidis L, Ouzounoglou E, Amditis A. A comparative study of autonomous object detection algorithms in the maritime environment using a UAV platform. Computation 2022;10:42. https://doi.org/10.3390/COMPUTATION10030042. 2022;10:42.
[121] Gorkin R, Adams K, Berryman MJ, Aubin S, Li W, Davis AR, et al. Sharkeye: real-time autonomous personal shark alerting via aerial surveillance. Drones 2020;4:18. https://doi.org/10.3390/DRONES4020018. 2020;4:18.
[122] Unlu E, Zenou E, Riviere N, Dupouy PE. Deep learning-based strategies for the detection and tracking of drones using several cameras. IPSJ Transactions on Computer Vision and Applications 2019;11:1–13. https://doi.org/10.1186/S41074-019-0059-X/FIGURES/15.
[123] Akhloufi MA, Arola S, Bonnet A. Drones chasing drones: reinforcement learning and deep search area proposal. Drones 2019;3. https://doi.org/10.3390/DRONES3030058. Page 58 2019;3:58.
[124] Rohan A, Rabah M, Kim SH. Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2. IEEE Access 2019;7:69575–84. https://doi.org/10.1109/ACCESS.2019.2919332.
[125] Bonatti R, Wang W, Ho C, Ahuja A, Gschwindt M, Camci E, et al. Autonomous aerial cinematography in unstructured environments with learned artistic decision-making. J Field Robot 2019;37:606–41. https://doi.org/10.1002/rob.21931.
[126] Kakaletsis E, Mademlis I, Nikolaidis N, Pitas I. Multiview vision-based human crowd localization for UAV fleet flight safety. Signal Process Image Commun 2021;99:116484. https://doi.org/10.1016/J.IMAGE.2021.116484.
[127] Pandya A, Jha A, Cenkeramaddi LR. A velocity estimation technique for a monocular camera using mmWave FMCW radars. Electronics 2021;10:2397. https://doi.org/10.3390/ELECTRONICS10192397. 2021;10:2397.
[128] Martínez A, Belmonte LM, García AS, Fernández-Caballero A, Morales R. Facial emotion recognition from an unmanned flying social robot for home care of dependent people. Electronics 2021;10:868. https://doi.org/10.3390/ELECTRONICS10070868. 2021;10:868.
[129] Pan N, Zhang R, Yang T, Cui C, Xu C, Gao F. Fast-Tracker 2.0: improving autonomy of aerial tracking with active vision and human location regression. IET Cyber-Systems and Robotics 2021;3:292–301. https://doi.org/10.1049/CSY2.12033.
[130] Lo LY, Yiu CH, Tang Y, Yang AS, Li B, Wen CY. Dynamic object tracking on autonomous UAV system for surveillance applications. Sensors 2021;21:7888. https://doi.org/10.3390/S21237888. 2021;21:7888.
[131] Li M, Cai Z, Zhao J, Wang Y, Wang Y, Lu K. MNNMs integrated control for UAV autonomous tracking randomly moving target based on learning method. Sensors 2021;21:7307. https://doi.org/10.3390/S21217307. 2021;21:7307.
[132] Maltezos E, Douklias A, Dadoukis A, Misichroni F, Karagiannidis L, Antonopoulos M, et al. The INUS platform: a modular solution for object detection and tracking from UAVs and terrestrial surveillance assets. Computation 2021;9:12. https://doi.org/10.3390/COMPUTATION9020012. 2021;9:12.
[133] Palossi D, Zimmerman N, Burrello A, Conti F, Muller H, Gambardella LM, et al. Fully onboard AI-powered human-drone pose estimation on ultralow-power autonomous flying nano-UAVs. IEEE Internet Things J 2022;9:1913–29. https://doi.org/10.1109/JIOT.2021.3091643.
[134] Abunadi I, Althobaiti MM, Al-Wesabi FN, Hilal AM, Medani M, Hamza MA, et al. Federated learning with blockchain assisted image classification for clustered UAV networks. Comput Mater Continua (CMC) 2022;72:1195–212. https://doi.org/10.32604/CMC.2022.025473.
[135] Bożko A, Ambroziak L. Influence of insufficient dataset augmentation on IoU and detection threshold in CNN training for object detection on aerial images. Sensors 2022;22:9080. https://doi.org/10.3390/S22239080. 2022;22:9080.
[136] Mademlis I, Torres-González A, Capitán J, Montagnuolo M, Messina A, Negro F, et al. A multiple-UAV architecture for autonomous media production. Multimed Tool Appl 2023;82:1905–34. https://doi.org/10.1007/S11042-022-13319-8/FIGURES/10.
[137] Saini N, Bonetto E, Price E, Ahmad A, Black MJ. AirPose: multi-view fusion network for aerial 3D human pose and shape estimation. IEEE Rob Autom Lett 2022;7:4805–12. https://doi.org/10.1109/LRA.2022.3145494.
[138] Adekola OD, Udekwu OK, Saliu OT, Dada DW, Maitanmi SO, Odumuyiwa V, et al. Object tracking-based "follow-me" unmanned aerial vehicle (UAV) system. Comput Syst Sci Eng 2021;41:875–90. https://doi.org/10.32604/CSSE.2022.021029.
[139] Li B, Liu B, Han D, Wang Z. Autonomous tracking of ShenZhou reentry capsules based on heterogeneous UAV swarms. Drones 2023;7. https://doi.org/10.3390/DRONES7010020. Page 20 2022;7:20.
[140] Nguyen H, Thudumu S, Du H, Mouzakis K, Vasa R. UAV dynamic object tracking with lightweight deep vision reinforcement learning. Algorithms 2023;16:227. https://doi.org/10.3390/A16050227. 2023;16:227.
[141] Chen WC, Lin CL, Chen YY, Cheng HH. Quadcopter drone for vision-based autonomous target following. Aerospace 2023;10. https://doi.org/10.3390/AEROSPACE10010082. Page 82 2023;10:82.
[142] Bemposta Rosende S, Ghisler S, Fernández-Andrés J, Sánchez-Soriano J. Implementation of an edge-computing vision system on reduced-board computers embedded in UAVs for intelligent traffic management. Drones 2023;7:682. https://doi.org/10.3390/DRONES7110682. 2023;7:682.
[143] Mohsan H, Lorenz P, Rabie K, Khan MA, Shafiq M, Gong K, et al. Research of an unmanned aerial vehicle autonomous aerial refueling docking method based on binocular vision. Drones 2023;7:433. https://doi.org/10.3390/DRONES7070433. 2023;7:433.
[144] Thakur N, Nagrath P, Jain R, Saini D, Sharma N, Hemanth DJ. Autonomous pedestrian detection for crowd surveillance using deep learning framework. Soft Comput 2023;27:9383–99. https://doi.org/10.1007/S00500-023-08289-4.
[145] Lv J, Tu L. Deep learning-based visual navigation control method for autonomous trajectory of UAVs. Applied Mathematics and Nonlinear Sciences 2024;9. https://doi.org/10.2478/AMNS.2023.2.01249.
[146] Xu Z, Zhan X, Xiu Y, Suzuki C, Shimada K. Onboard dynamic-object detection and tracking for autonomous robot navigation with RGB-D camera. IEEE Rob Autom Lett 2023;9:651–8. https://doi.org/10.1109/LRA.2023.3334683.
[147] Pueyo P, Dendarieta J, Montijano E, Murillo AC, Schwager M. CineMPC: a fully autonomous drone cinematography system incorporating zoom, focus, pose, and scene composition. IEEE Trans Robot 2024;40:1740–57. https://doi.org/10.1109/TRO.2024.3353550.
[148] Xu Y, Fu Z, Peng W, Ding Z, Lu G, Liu Q. A recurrent approach for uninterrupted tracking of rotor blades using kalman filter. IEEE Access 2023;12:5321–32. https://doi.org/10.1109/ACCESS.2023.3344805.
[149] Ou K, Dong C, Liu X, Zhai Y, Li Y, Huang W, et al. Drone-TOOD: a lightweight task-aligned object detection algorithm for vehicle detection in UAV images. IEEE Access 2024;12:41999–2016. https://doi.org/10.1109/ACCESS.2024.3378248.
[150] Cereda E, Bonato S, Nava M, Giusti A, Palossi D. Vision-state fusion: improving deep neural networks for autonomous robotics. J Intell Rob Syst: Theory and Applications 2024;110:1–13. https://doi.org/10.1007/S10846-024-02091-6/METRICS.
[151] Worth D, Choate J, Lynch J, Scott Nykl, Clark Taylor. Relative vectoring using dual object detection for autonomous aerial refueling. Neural Comput Appl 2024;36:17. https://doi.org/10.1007/S00521-024-09589-Y. 2024;36:10143–10163.
[152] Alparslan O, Cetin O. Real-time indoor path planning using object detection for autonomous flying robots. Intelligent Automation & Soft Computing 2023;36:3355–70. https://doi.org/10.32604/IASC.2023.035689.
[153] Surojaya A, Zhang N, Bergado JR, Nex F. Towards fully autonomous UAV: damaged building-opening detection for outdoor-indoor transition in urban search and rescue. Electronics 2024;13:558. https://doi.org/10.3390/ELECTRONICS13030558. 2024;13:558.
[154] Haque A, Elsaharti A, Elderini T, Elsaharty MA, Neubert J. UAV autonomous localization using macro-features matching with a CAD model. Sensors 2020;20:743. https://doi.org/10.3390/S20030743. 2020;20:743.
[155] Bavle H, Puente PD La, How JP, Campoy P. VPS-SLAM: visual planar semantic SLAM for aerial robotic systems. IEEE Access 2020;8:60704–18. https://doi.org/10.1109/ACCESS.2020.2983121.
[156] Li Z, Zhao J, Zhou X, Wei S, Li P, Shuang F. RTSDM: a real-time semantic dense mapping system for UAVs. Machines 2022;10:285. https://doi.org/10.3390/MACHINES10040285. 2022;10:285.
[157] Duc V, Bui T, Shirakawa HS. A UAV exploration method by detecting multiple directions with deep learning. https://doi.org/10.18178/ijmerr.9.10.1419-1426; 2020.
[158] Bui VD, Shirakawa T, Sato H. Autonomous unmanned aerial vehicle flight control using multi-task deep neural network for exploring indoor environments 2022;15:130–44. https://doi.org/10.1080/18824889.2022.2087413.
[159] Sampedro C, Rodriguez-Ramos A, Bavle H, Carrio A, de la Puente P, Campoy P. A fully-autonomous aerial robot for search and rescue applications in indoor environments using learning-based techniques. J Intell Rob Syst: Theory and Applications 2019;95:601–27. https://doi.org/10.1007/S10846-018-0898-1/METRICS.
[160] Lygouras E, Santavas N, Taitzoglou A, Tarchanidis K, Mitropoulos A, Gasteratos A. Unsupervised human detection with an embedded vision system on

a fully autonomous UAV for search and rescue operations. Sensors 2019;19:3542. https://doi.org/10.3390/S19163542. 2019;19:3542.
[161] Wu C, Ju B, Wu Y, Lin X, Xiong N, Xu G, et al. UAV autonomous target search based on deep reinforcement learning in complex disaster scene. IEEE Access 2019;7:117227–45. https://doi.org/10.1109/ACCESS.2019.2933002.
[162] Chatziparaschis D, Lagoudakis MG, Partsinevelos P. Aerial and ground robot collaboration for autonomous mapping in search and rescue missions. Drones 2020;4. https://doi.org/10.3390/DRONES4040079. Page 79 2020;4:79.
[163] Loquercio A, Kaufmann E, Ranftl R, Dosovitskiy A, Koltun V, Scaramuzza D. Deep drone racing: from simulation to reality with domain randomization. IEEE Trans Robot 2019;36:1–14. https://doi.org/10.1109/TRO.2019.2942989.
[164] Niculescu V, Lamberti L, Conti F, Benini L, Palossi D. Improving autonomous nano-drones performance via automated end-to-end optimization and deployment of DNNs. IEEE J Emerg Sel Top Circuits Syst 2021;11:548–62. https://doi.org/10.1109/JETCAS.2021.3126259.
[165] Fan J, Lu R, Yang X, Gao F, Li Q, Zeng J. Design and implementation of intelligent EOD system based on six-rotor UAV. Drones 2021;5:146. https://doi.org/10.3390/DRONES5040146. 2021;5:146.
[166] Krishnan S, Boroujerdian B, Fu W, Faust A, Reddi VJ. Air learning: a deep reinforcement learning gym for autonomous aerial robot visual navigation. Mach Learn 2019;110:2501–40. https://doi.org/10.1007/s10994-021-06006-6.
[167] Ou J, Guo X, Zhu M, Lou W. Autonomous quadrotor obstacle avoidance based on dueling double deep recurrent Q-learning with monocular vision. Neurocomputing 2021;441:300–10. https://doi.org/10.1016/J.NEUCOM.2021.02.017.
[168] He W, Yang D, Peng H, Liang S, Lin Y. An efficient ensemble binarized deep neural network on chip with perception-control integrated. Sensors 2021;21:3407. https://doi.org/10.3390/S21103407. 2021;21:3407.
[169] Tullu A, Endale B, Wondosen A, Hwang HY. Machine learning approach to real-time 3d path planning for autonomous navigation of unmanned aerial vehicle. Appl Sci 2021;11:4706. https://doi.org/10.3390/APP11104706/S1.
[170] Morad SD, Mecca R, Poudel RPK, Liwicki S, Cipolla R. Embodied visual navigation with automatic curriculum learning in real environments. IEEE Rob Autom Lett 2020;6:683–90. https://doi.org/10.1109/LRA.2020.3048662.
[171] Xie R, Meng Z, Wang L, Li H, Wang K, Wu Z. Unmanned aerial vehicle path planning algorithm based on deep reinforcement learning in large-scale and dynamic environments. IEEE Access 2021;9:24884–900. https://doi.org/10.1109/ACCESS.2021.3057485.
[172] Hai J, Hao Y, Zou F, Lin F, Han S. A Visual Navigation System for UAV under Diverse Illumination Conditions 2021;35:1529–49. https://doi.org/10.1080/08839514.2021.1985799.
[173] Maw AA, Tyan M, Nguyen TA, Lee JW. iADA*-RL: anytime graph-based path planning with deep reinforcement learning for an autonomous UAV. Appl Sci 2021;11:3948. https://doi.org/10.3390/APP11093948. 2021;11:3948.
[174] Kinnari J, Verdoja F, Kyrki V. Season-invariant GNSS-denied visual localization for UAVs. IEEE Rob Autom Lett 2021;7:10232–9. https://doi.org/10.1109/LRA.2022.3191038.
[175] Arshad MA, Khan SH, Qamar S, Khan MW, Murtza I, Gwak J, et al. Drone navigation using region and edge exploitation-based deep CNN. IEEE Access 2022;10:95441–50. https://doi.org/10.1109/ACCESS.2022.3204876.
[176] Liu L, Wu Y, Fu G, Zhou C. An improved four-rotor UAV autonomous navigation multisensor fusion depth learning. Wireless Commun Mobile Comput 2022;2022. https://doi.org/10.1155/2022/2701359.
[177] Dooraki AR, Lee DJ. A multi-objective reinforcement learning based controller for autonomous navigation in challenging environments. Machines 2022;10:500. https://doi.org/10.3390/MACHINES10070500. 2022;10:500.
[178] Almalki FA, Angelides MC. Autonomous flying IoT: a synergy of machine learning, digital elevation, and 3D structure change detection. Comput Commun 2022;190:154–65. https://doi.org/10.1016/J.COMCOM.2022.03.022.
[179] Lei T, Luo C, Jan GE, Bi Z. Deep learning-based complete coverage path planning with Re-joint and obstacle fusion paradigm. Front Robot AI 2022;9:843816. https://doi.org/10.3389/FROBT.2022.843816/BIBTEX.
[180] Kazim M, Zaidi A, Ali S, Raza MT, Abbas G, Ullah N, et al. Perception action aware-based autonomous drone race in a photorealistic environment. IEEE Access 2022;10:42566–76. https://doi.org/10.1109/ACCESS.2022.3168710.
[181] Chen C, Wang Z, Gong Z, Cai P, Zhang C, Li Y. Autonomous navigation and obstacle avoidance for small VTOL UAV in unknown environments. Symmetry 2022;14:2608. https://doi.org/10.3390/SYM14122608. 2022;14:2608.
[182] Romero-Lugo A, Magadan-Salazar A, Fuentes-Pacheco J, Pinto-Elías R. A comparison of deep neural networks for monocular depth map estimation in natural environments flying at low altitude. Sensors 2022;22:9830. https://doi.org/10.3390/S22249830. 2022;22:9830.
[183] Xie MR, Jung SY, Chen KW. Three-dimensional drone exploration with saliency prediction in real unknown environments. Aerospace 2023;10:488. https://doi.org/10.3390/AEROSPACE10050488. 2023;10:488.
[184] Kalidas AP, Joshua CJ, Md AQ, Basheer S, Mohan S, Sakri S. Deep reinforcement learning for vision-based navigation of UAVs in avoiding stationary and mobile obstacles. Drones 2023;7:245. https://doi.org/10.3390/DRONES7040245. 2023;7:245.
[185] Xu Y, Wei Y, Jiang K, Wang D, Deng H. Multiple UAVs path planning based on deep reinforcement learning in communication denial environment. Mathematics 2023;11:405. https://doi.org/10.3390/MATH11020405. 2023;11:405.
[186] Bultmann S, Quenzel J, Behnke S. Real-time multi-modal semantic fusion on unmanned aerial vehicles with label propagation for cross-domain adaptation. Robot Autonom Syst 2023;159:104286. https://doi.org/10.1016/J.ROBOT.2022.104286.
[187] Naso D, Pohudina O, Yashin S, Pohudin A, Bartolo R. Autonomous flight insurance method of unmanned aerial vehicles parot mambo using semantic segmentation data. Radioelectronic and Computer Systems 2023:147–54. https://doi.org/10.32620/REKS.2023.1.12.
[188] Cao Q, Liao Y, Fu Z, Peng H, Ding Z, Huang Z, et al. An iterative closest point method for lidar odometry with fused semantic features. Appl Sci 2023;13:12741. https://doi.org/10.3390/APP132312741. Page 12741 2023;13.
[189] Chen C, Song B, Fu Q, Xue D, He L. Event-triggered hierarchical planner for autonomous navigation in unknown environment. Drones 2023;7:690. https://doi.org/10.3390/DRONES7120690. 2023;7:690.
[190] Hu K, Li H, Zhuang J, Hao Z, Fan Z. Efficient focus autoencoders for fast autonomous flight in intricate wild scenarios. Drones 2023;7:609. https://doi.org/10.3390/DRONES7100609. 2023;7:609.
[191] Si X, Xu G, Ke M, Zhang H, Tong K, Qi F. Relative localization within a quadcopter unmanned aerial vehicle swarm based on airborne monocular vision. Drones 2023;7:612. https://doi.org/10.3390/DRONES7100612. 2023;7:612.
[192] Cheng C, Liu D, Du JH, Li YZ. Research on visual perception for coordinated air–sea through a cooperative USV-UAV system. J Mar Sci Eng 2023;11:1978. https://doi.org/10.3390/JMSE11101978. 2023;11:1978.
[193] Kinnari J, Renzulli R, Verdoja F, Kyrki V. Lsvl. Robot Autonom Syst 2023;168. https://doi.org/10.1016/J.ROBOT.2023.104497.
[194] Ren Y, Liu Y, Huang Z, Liu W, Wang W. 2chADCNN: a template matching network for season-changing UAV aerial images and satellite imagery. Drones 2023;7:558. https://doi.org/10.3390/DRONES7090558. 2023;7:558.
[195] Dsouza JM, Rafikh RM, Nair VG. Autonomous navigation system for multi-quadrotor coordination and human detection in search and rescue. J Robot Mechatron 2023;35:1084–91. https://doi.org/10.20965/JRM.2023.P1084.
[196] Wang J, Zhang X, Yan T, Tan A. DPNet: dual-pyramid semantic segmentation network based on improved Deeplabv3 plus. Electronics 2023;12:3161. https://doi.org/10.3390/ELECTRONICS12143161. 2023;12:3161.
[197] Hinniger C, Rüter J. Synthetic training data for semantic segmentation of the environment from UAV perspective. Aerospace 2023;10:604. https://doi.org/10.3390/AEROSPACE10070604. 2023;10:604.
[198] Horikawa Y, Yoshida R, Mochizuki S, Matsumura T. [Paper] deep learning based hierarchical object detection system adopting a depth correction scheme for high-resolution aerial images towards realization of autonomous UAV flight. ITE Transactions on Media Technology and Applications 2024;12:85–92. https://doi.org/10.3169/MTA.12.85.
[199] Taipalmaa J, Raitoharju J, Queralta JP, Westerlund T, Gabbouj M. On automatic person-in-water detection for marine search and rescue operations. IEEE Access 2024;12:52428–38. https://doi.org/10.1109/ACCESS.2024.3386640.
[200] Kucukayan G, Karacan H. YOLO-IHD: improved real-time human detection system for indoor drones. Sensors 2024;24:922. https://doi.org/10.3390/S24030922. 2024;24:922.
[201] Lamberti L, Cereda E, Abbate G, Bellone L, Morinigo VJK, Barcis M, et al. A sim-to-real deep learning-based framework for autonomous nano-drone racing. IEEE Rob Autom Lett 2023;9:1899–906. https://doi.org/10.1109/LRA.2024.3349814.
[202] Wang F, Zhu X, Zhou Z, Tang Y. Deep-reinforcement-learning-based UAV autonomous navigation and collision avoidance in unknown environments. Chin J Aeronaut 2024;37:237–57. https://doi.org/10.1016/J.CJA.2023.09.033.
[203] Wang H, Shen Q, Deng Z, Cao X, Wang X. Absolute pose estimation of UAV based on large-scale satellite image. Chin J Aeronaut 2024;37:219–31. https://doi.org/10.1016/J.CJA.2023.12.028.
[204] Cano-Solis M, Ballesteros JR, Branch-Bedoya JW. VEPL dataset: a vegetation encroachment in power line corridors dataset for semantic segmentation of drone aerial orthomosaics. Data 2023;8:128. https://doi.org/10.3390/DATA8080128. 2023;8:128.
[205] FLAME Dataset | Papers With Code n.d https://paperswithcode.com/dataset/flame (accessed July 17, 2024).
[206] VisDrone - Ultralytics YOLO Docs n.d. https://docs.ultralytics.com/datasets/detect/visdrone/(accessed July 17, 2024).
[207] Bitoun J, Winkler S. HelipadCat: categorised helipad image dataset and detection method. In: IEEE region 10 annual international conference, proceedings/TENCON; 2020. p. 685–9. https://doi.org/10.1109/TENCON50793.2020.9293911. 2020.
[208] Seagull Dataset – VisLab – Computer and Robot Vision Laboratory n.d. https://vislab.isr.tecnico.ulisboa.pt/seagull-dataset/(accessed July 17, 2024).
[209] GitHub - MiguelARD/DoorDetect-Dataset: Labelled image dataset for door and handle detection. n.d https://github.com/MiguelARD/DoorDetect-Dataset (accessed July 17, 2024).
[210] Tummala S, Kadry S, Bukhari SAC, Rauf HT. Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling. Curr Oncol 2022;29:7498–511. https://doi.org/10.3390/CURRONCOL29100590. 2022;29:7498–511.
[211] Zheng Y, Jiang W. Evaluation of vision transformers for traffic sign classification. Wireless Commun Mobile Comput 2022;2022:3041117. https://doi.org/10.1155/2022/3041117.
[212] Meng L, Li H, Chen B-C, Lan S, Wu Z, Jiang Y-G, et al. AdaViT: adaptive vision transformers for efficient image recognition. 2022. p. 12309–18.
[213] Chen C-F, Richard, Fan Q, Panda R. CrossViT: cross-attention multi-scale vision transformer for image classification. 2021. p. 357–66.
[214] Xu C., Guan Z., Zhao W., Wu H., Niu Y., Ling B. Adversarial Incomplete Multi-view Clustering. Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19), proceedings; 2019.p. 3933-3939.

[215] McEnroe P, Wang S, Liyanage M. A survey on the convergence of edge computing and AI for UAVs: opportunities and challenges. IEEE Internet Things J 2022;9: 15435–59. https://doi.org/10.1109/JIOT.2022.3176400.
[216] Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018), proceedings;2018.p. 4510-4520.
[217] Ampatzidis Y, Partel V, Costa L. Agroview: cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. Comput Electron Agric 2020;174:105457. https://doi.org/10.1016/J.COMPAG.2020.105457.
[218] Yazid Y, Ez-Zazi I, Guerrero-González A, Oualkadi A El, Arioua M, Lakas A, et al. UAV-enabled mobile edge-computing for IoT based on AI: a comprehensive review. Drones 2021;5:148. https://doi.org/10.3390/DRONES5040148. 2021;5: 148.
[219] Ure NK, Chowdhary G, Toksoz T, How JP, Vavrina MA, Vian J. An automated battery management system to enable persistent missions with multiple aerial vehicles. IEEE ASME Trans Mechatron 2015;20:275–86. https://doi.org/10.1109/TMECH.2013.2294805.
[220] Zhou Z, Feng J, Gu B, Ai B, Mumtaz S, Rodriguez J, et al. When mobile crowd sensing meets UAV: energy-efficient task assignment and route planning. IEEE Trans Commun 2018;66:5526–38. https://doi.org/10.1109/TCOMM.2018.2857461.
[221] Atzori L, Iera A, Morabito G. The internet of things: a survey. Comput Network 2010;54:2787–805. https://doi.org/10.1016/J.COMNET.2010.05.010.
[222] Abeywickrama HV, Jayawickrama BA, He Y, Dutkiewicz E. Comprehensive energy consumption model for unmanned aerial vehicles, based on empirical studies of battery performance. IEEE Access 2018;6:58383–94. https://doi.org/10.1109/ACCESS.2018.2875040.
[223] Ozer S, Ilhan HE, Ozkanoglu MA, Cirpan HA. Offloading deep learning powered vision tasks from UAV to 5G edge server with denoising. IEEE Trans Veh Technol 2023;72:8035–48. https://doi.org/10.1109/TVT.2023.3243529.
[224] Choi K, Oh BS, Yu S. Memory access minimization for mean-shift tracking in mobile devices. Multimed Tool Appl 2021;80:34173–87. https://doi.org/10.1007/S11042-020-09364-W/FIGURES/13.
[225] Lee J, Wang P, Xu R, Jain S, Dasari V, Weston N, et al. Virtuoso: energy- and latency-aware streamlining of streaming videos on systems-on-chips. ACM Transact Des Autom Electron Syst 2023;28:31. https://doi.org/10.1145/3564289/ASSET/D6F06553-717E-47EF-871F-85052B461AA3/ASSETS/GRAPHIC/TODAES-2022-P-2205-F21.JPG.
[226] Cheng E, Mitra S. Cross-layer resilience. Cross-Layer Reliability of Computing Systems 2020:113–54. https://doi.org/10.1049/PBCS057E_ch5.
[227] Tong Q, Yang M, Zinetullina A. A dynamic bayesian network-based approach to resilience assessment of engineered systems. J Loss Prev Process Ind 2020;65: 104152. https://doi.org/10.1016/J.JLP.2020.104152.
[228] Bemposta Rosende S, Ghisler S, Fernández-Andrés J, Sánchez-Soriano J. Implementation of an edge-computing vision system on reduced-board computers embedded in UAVs for intelligent traffic management. Drones 2023;7:682. https://doi.org/10.3390/DRONES7110682. 2023;7:682.
[229] Lu Y, Xue Z, Xia GS, Zhang L. A survey on vision-based UAV navigation. Geo Spatial Inf Sci 2018;21:21–32. https://doi.org/10.1080/10095020.2017.1420509.
[230] Low-Cost E, Ceccarelli M, Becerra HM, Ramirez-Paredes J-P, Flores G, Brandão A, et al. Low-cost computer-vision-based embedded systems for UAVs. Robotics 2023;12:145. https://doi.org/10.3390/ROBOTICS12060145. 2023;12:145.
[231] Aksenov AY, Kuleshov SV, Zaytseva AA. An application of computer vision systems to solve the problem of unmanned aerial vehicle control. Transport and Telecommunication 2014;15:209–14. https://doi.org/10.2478/TTJ-2014-0018.