https://doi.org/10.1590/jatm.v17.1377

ORIGINAL PAPER

Challenges of Computer Vision for Commercial Unmanned Aerial Vehicle Detection

Valentina Grichshenko[1], Assemkhan Mukushev[2*], Andrey Kokidko[2], Nurzhan Zikiryaev[3]

1. Institute of Ionosphere – Laboratory of Reliability and Safety of Orbital and Ground Space Systems – Almaty – Republic of Kazakhstan.
2. Military Engineering Institute of Radio Electronics and Communications – Almaty – Republic of Kazakhstan.
3. Military Engineering Institute of Radio Electronics and Communications – Department of the Fundamentals of Military Radio Engineering and Electronics – Almaty – Republic of Kazakhstan.

*Correspondence author: ass.mukushev@gmail.com

ABSTRACT

The study aims to analyze the existing computer vision techniques for commercial drone detection to identify their advantages, disadvantages, and determine the best approaches in different application scenarios. The research methodology used synthesis methods to explore and propose combinations of techniques based on an analysis of the methodology and results of other works in the literature. It employed algorithms and sensor data analysis to assess the effectiveness of detection methods, and deduction to formulate hypotheses and conclusions based on data and theories. The main research results include the development of computer vision methods for detecting commercial drones, identifying their visual detectability at different altitudes, analyzing different object detection methods, and evaluating the applicability of these methods for commercial applications. In addition, the study identified the advantages and disadvantages of applying computer vision to commercial drone detection and offered recommendations for further research and practical implementation. The practical value of this study is to improve the detection systems of commercial drones, thereby enhancing the safety and efficiency of their use.

Keywords: Aircraft; Detection; Sensor systems; Performance evaluation; Advanced technology.

INTRODUCTION

Nowadays, research on the application of computer vision to the detection of commercial unmanned aerial vehicles (UAVs) is proving to be highly relevant and important due to several key factors (Mykhalevskiy *et al.* 2024; Yermolenko *et al.* 2024). With the increasing number of commercial UAVs in various industries such as transport, agriculture, surveying, and environmental monitoring, there is a need for effective detection systems (Cazzato *et al.* 2020; Chen *et al.* 2023). This is necessitated not only by the increased security and control over the use of such devices but also by the protection of data privacy, especially in areas where sensitive information is present. Research in this area is also stimulated by the active development of the drone technology market itself. The constant development of new models and types requires appropriate tools and technologies for their detection and monitoring.

The research problem includes several aspects, including ambiguity in the choice of optimal drone detection methods, limitations in the performance of existing algorithms when processing large amounts of data, difficulties in adapting algorithms to different survey conditions and customer requirements, and problems of defense against cyber-attacks and hacking of detection systems (Perry and Guo 2021). Other issues include the need to address ethical and legal considerations in the use of computer

Received: Oct. 18, 2024 | **Accepted:** Mar. 17, 2025
Peer Review History: Single Blind Peer Review.
Section editor: Alison Moraes

vision technology for UAV detection, and the need to improve the performance and accuracy of detection systems to effectively monitor the growing number of commercial drones in various industries and domains.

Kozachenko (2021) addressed the model of complex application of measures to detect small UAVs, the problems of radio-electronic suppression of the UAV navigation system, as well as the features of radio-electronic suppression of the UAV navigation system based on the reception of satellite navigation system signals, and the problems of radio-electronic suppression of UAV control and data transmission radio lines. However, it is necessary to analyze how effective the proposed methods are in real conditions and whether they apply to different types of UAVs.

According to Zhao *et al.* (2022), the field of computer vision is sufficiently developed to detect and track intruding UAVs. They propose a DUT Anti-UAV dataset that includes extensive material for training detection and tracking algorithms. However, the dataset's effectiveness and its applicability in real surveillance environments require further study. Pawełczyk and Wojtyra (2020) note the significant increase in the number of drone incidents and the need for drone detection systems running on low-performance hardware. However, the performance and reliability of such systems under different operating conditions should be investigated in more detail. Leira *et al.* (2020) study an object detection, recognition, and tracking system for UAVs applied in a maritime object tracking system. It is important to extend the research to other applications of this system and evaluate its performance in different scenarios.

Bazeltsev (2020) states that over the last 10 years, the field of UAVs has expanded rapidly. They are used in various environments such as reconnaissance, surveying, rescue operations, and mapping. UAVs are maneuverable in the air, can be operated by remote control, and can reach high altitudes and distances. Many UAVs are equipped with an inbuilt camera, such as an action camera, which allows the drone to take photos and videos from various angles. However, there are some disadvantages: drone control can be quite complicated. Even when applying the latest advances in software, the pilot must be very careful, as losing control of the drone could mean losing the UAV itself. His study did not address the aspect related to technical limitations and potential risks that may arise when using drones.

The study aims to investigate and evaluate existing computer vision techniques for commercial drone detection, with a focus on identifying the most effective approaches for UAV detection across various contexts. By considering the current limitations and challenges discussed in the literature, the study seeks to assess the performance and applicability of these techniques in real-world scenarios, providing a comprehensive understanding of the strengths and weaknesses of different methods.

METHODOLOGY

The methodology consists of nine sequential steps, each addressing a critical aspect of the research process. These steps include selecting performance indicators, conducting data analysis, synthesizing different algorithms, integrating data from various sensors, applying fusion methods, evaluating algorithm performance, analyzing trends and patterns, researching the visual detectability of drones, and formulating hypotheses and conclusions. A detailed description of each stage is provided below to ensure a clear understanding of the methodology and its application in the context of this study.

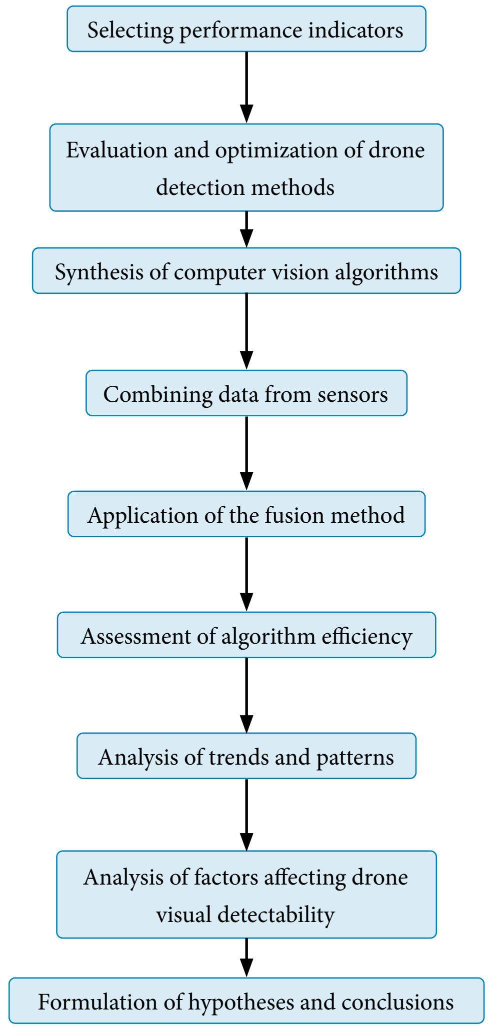
The simplified process flow (Fig. 1) summarizes the main steps in the methodology for the study of the use of computer vision to detect commercial drones.

Selecting Performance Indicators

The first step was to select appropriate performance indicators to evaluate the use of computer vision in commercial drone detection. The key metrics are accuracy, completeness, F-Score, detection rate, and the number of false positives and missed detections. These metrics enable a comprehensive quantitative evaluation of the performance of detection models, focusing on their accuracy and effectiveness under varying conditions.

Evaluation and Optimization of Drone Detection Methods

After collecting and analyzing the data, conclusions were drawn about the effectiveness of various drone detection methods. This stage allows us to evaluate which methods provide the best results in real-world conditions and helps to identify weaknesses in the applied approaches that require further optimization.



Source: Elaborated by the authors.

Figure 1. Main stages of the methodology for researching the use of computer vision to detect commercial drones.

Synthesis of Computer Vision Algorithms

In this stage of the research, a synthesis method was used to combine different computer vision algorithms to create a more efficient and reliable commercial drone detection system. By combining the strengths of several methods, better detection quality was achieved and a wider coverage of different scenarios and usage conditions was provided, which increased the system's efficiency.

Combining Data from Sensors

The fusion method was also used to combine data from different sensors, such as video cameras, radars, and Light Detection and Ranging (LIDAR). This has resulted in a more comprehensive drone detection system that takes advantage of the strengths of each sensor to improve the accuracy and reliability of the system in different environments. Combining data from different sources makes the system more adaptable to varying conditions.

Application of the Fusion Method

The fusion method involves more than just integrating various input data into a model for classification and detection tasks. It also encompasses the combination of results from different models after they have processed the data. By using outputs from multiple models together, after they have completed their respective processes, we can enhance the stability and efficiency of the drone detection system. This approach helps reduce the number of false positives and improves the accuracy of target detection, as it leverages the strengths of different models and algorithms to provide a more robust final result. Thus, the fusion method not only integrates data but also merges the outcomes of different models to achieve more reliable and precise detections.

Assessment of Algorithm Efficiency

The study used an analytical approach to evaluate the effectiveness of different drone detection algorithms based on metrics such as accuracy, completeness, and detection rate. This stage allows comparing the effectiveness of different methods and choosing the most suitable one for specific application conditions, which is important for further optimization of the system.

Analysis of Trends and Patterns

Analytical methods were used to identify the main trends and patterns in the behavior of detected objects. This allows for a better understanding of the detection process and optimization of the overall system performance, particularly by adapting to typical drone behavioral patterns. Studying such trends helps to improve detection algorithms for specific conditions.

Analysis of Factors Affecting Drone Visual Detectability

The study also examined the visual detectability of drones at different heights and distances. For this purpose, data from various sources, such as video footage, sensors, and flight simulations, were used. This research allows for the improvement of visual detection methods by taking into account various factors that affect drone visibility, such as weather conditions and flight altitude.

Formulation of Hypotheses and Conclusions

The last step is to formulate hypotheses and conclusions based on data analysis and logical thinking. Using the deductive method, researchers formulate theoretical models that explain the effectiveness or ineffectiveness of specific drone detection methods. This stage provides general principles and patterns underlying the effectiveness of detection algorithms and suggests areas for further research and improvement of detection technologies.

RESULTS

Evolution of Computer Vision Techniques in Drone Detection

The history of computer vision development in drone detection can be traced through key milestones that show the evolution from early image processing techniques to the advanced methods used today. This timeline outlines the significant advancements in the field, with each stage building upon the previous one (Cazzato *et al.* 2020).

The journey of computer vision began with the first attempts to use computers to analyze images. In the early stages, computer vision was limited to primitive image processing techniques, such as filtering and thresholding. These methods focused on basic tasks like enhancing image contrast and detecting simple patterns. During this period, researchers also began experimenting with pattern recognition techniques to detect objects in images, laying the groundwork for more advanced methods.

The foundations of computer vision were laid during the initial attempts to use computers for analyzing visual data. At the early stages, primitive image processing techniques such as filtering and thresholding were widely used. Filtering methods aimed to reduce noise, enhance image contrast, and emphasize important features in visual data. Thresholding, on the other hand, was employed to segment images into binary formats by differentiating objects from their background based on pixel intensity values. Although these methods were simple and limited in scope, they provided critical insights into the challenges and possibilities of automated visual analysis, establishing a framework for more sophisticated approaches.

These early techniques were closely tied to the development of pattern recognition methods, which marked a significant step forward in object detection. Researchers began experimenting with statistical models and feature extraction techniques to identify and classify objects within images (Borodin *et al.* 2024; Xu *et al.* 2022). This work laid the groundwork for modern methods by introducing the idea of extracting relevant information from raw visual data and using it to train algorithms. Over time, the limitations of these approaches, such as their inability to handle complex patterns or adapt to variations in lighting, texture, and object orientation, highlighted the need for more advanced methods, paving the way for the development of machine learning and deep learning-based techniques.

Between the 1970s and 1990s, computer vision saw significant development with the introduction of methods based on the extraction of characteristic features from images. Geometric analysis, pattern matching, and object classification became widely used in UAV detection tasks. These methods aimed to identify key features in images, such as edges, corners, and textures, to recognize objects. However, the computational limitations of the time and the huge variability of imaging conditions, such as lighting and scale, hindered further progress.

The period from 2000 to 2010 marked a resurgence in computer vision for drone detection, driven by advances in computing power. Machine learning algorithms started to be actively applied to computer vision tasks, and object classification and detection algorithms trained on large datasets produced significantly better results (Bay *et al.* 2006; Lowe 2004). Datasets specifically designed for UAV detection became crucial to the development of this field, as they enabled the training of models that could handle the complexities of real-world drone detection.

The true breakthrough came after 2010 with the advent of deep neural networks and deep learning techniques, which revolutionized computer vision. Convolutional Neural Networks (CNNs) emerged as the dominant method for image processing and object detection. CNNs allowed for the automatic extraction of hierarchical features from images, leading to more accurate and robust detection systems (Borges *et al.* 2024; Borodin *et al.* 2024). The application of transfer learning methods also enabled the adaptation of pre-trained models for specific tasks like drone detection. This period has seen continuous progress in the field, with CNNs powering some of the most advanced and efficient object detection models used today. The evolution of computer vision techniques for drone detection has laid a strong foundation for the development of advanced algorithms and architectures (Xu *et al.* 2022). With the advent of CNNs and their ability to extract complex hierarchical features, modern computer vision has reached new heights. Building on this progress, it is crucial to explore the key algorithms and architectures that underpin contemporary advancements in drone detection.

Key Algorithms and Architectures in Computer Vision for Drone Detection

Throughout the development of computer vision techniques for UAV detection, several key algorithms have played an important role. One such method is Harris corner detection, which identifies points in an image where the intensity changes significantly in multiple directions, ideal for detecting distinctive features in images. Another important method is the scale-invariant feature transform (SIFT), which generates local descriptors that are invariant to scale and rotation, making it particularly useful for feature description in various environments. Speeded-up robust features (SURF), a faster alternative to SIFT, improves feature extraction speed while maintaining robustness, making it suitable for real-time applications. Both SIFT and SURF are widely used in computer vision as feature description methods, typically applied in matching algorithms to find pairs of similar features (Douklias *et al.* 2022). These methods, along with others, have formed the foundational building blocks for more sophisticated and accurate object detection algorithms used in modern drone detection.

Deep learning algorithms, such as CNNs, are highly computationally demanding, especially during the training phase. These models require substantial computational resources due to the large volume of data required for training and the complexity of the models themselves, which consist of many layers and parameters (Tang *et al.* 2023). Training deep learning models typically requires powerful graphics processing units or specialized processors, as these computational resources accelerate the training process. During inference (prediction) with these models, the computational requirements are reduced but still remain high, particularly for real-time applications in complex conditions (Douklias *et al.* 2022).

Traditional computer vision methods, such as Harris, SIFT, and SURF algorithms, are generally less computationally demanding compared to deep learning-based approaches. These algorithms rely on simpler techniques to detect distinctive points in images, which reduces the strain on hardware resources, making them suitable for tasks with fewer objects or where feature recognition and matching are the primary objectives. For example, algorithms like SIFT and SURF can perform well in scenarios where object detection is not real-time or where the dataset is smaller and less complex.

While these methods require less computational power, they may struggle with large, complex datasets or tasks involving real-time object detection, where the need for higher accuracy and faster processing speeds becomes critical. In these cases, deep learning models, despite their higher computational cost, demonstrate significant advantages. Their ability to process large datasets and handle complex features allows them to outperform traditional algorithms in tasks such as detecting objects in real-time or in highly varied environments.

While traditional methods like SIFT and SURF may offer lower computational costs in specific scenarios, they cannot match the effectiveness of deep learning models when dealing with more complex, large-scale tasks. The trade-off between computational efficiency and detection performance becomes evident when considering the increasing demands of modern object detection applications.

Faster CNN (R-CNN) improves the original R-CNN by integrating the region proposal network (RPN) to streamline object detection, combining region proposals and classification into one network, significantly boosting both speed and accuracy (Kakaletsis *et al.* 2021). You Only Look Once (YOLO) takes a different approach by predicting bounding boxes and class probabilities directly from the image in a single pass, making it suitable for real-time applications. Its various versions (YOLOv1 to YOLOv5) have enhanced speed, accuracy, and handling of small objects. Single Shot Multibox Detector (SSD) eliminates region proposals, performing detection in a single pass with multiple feature maps for different object sizes, offering a balance of speed and accuracy. Mask R-CNN extends Faster R-CNN with a branch for instance segmentation, enabling pixel-level object delineation, which is crucial for detailed applications like medical imaging or autonomous driving.

The Table 1 compares different CNN architectures such as YOLO, Faster R-CNN, SSD, and Mask R-CNN on several important metrics including accuracy, precision, recall, F1-Score, frame rate, resource consumption, suitability for real-time applications, and application domains. These metrics are key to selecting the appropriate architecture depending on the specific requirements of the object detection task, such as speed, accuracy, and resource consumption.

Table 1. Comparison of CNN architectures for object detection.

Model	Accuracy	Precision	Recall	F1-Score	Speed	Resource consumption	Real-time applicability	Use case suitability
YOLO	0.75	0.74	0.77	0.75	45	Medium	Yes	Surveillance, robotics
Faster R-CNN	0.80	0.79	0.82	0.80	10	High	No	Surveillance
SSD	0.78	0.76	0.79	0.77	25	Medium	Yes	Real-time detection
Mask R-CNN	0.82	0.80	0.83	0.81	8	High	No	Segmentation

Source: Based on the study by Tang *et al.* (2023).

Following the comparison of CNN architectures for object detection presented in Table 1, it is also important to evaluate architectures designed for object classification tasks.

It is important to note that the YOLO architecture discussed here is YOLOv5, which became widely adopted due to its open-source nature, ease of use, and stable performance at the time the paper was written. However, later versions such as YOLOv6, YOLOv7, and the more recent YOLOv8 and YOLOv9 have introduced significant improvements in accuracy, speed, and resource efficiency. These newer versions have enhanced capabilities in handling small objects, improved robustness in varying environmental conditions, and optimizations for edge computing devices.

YOLOv8 and YOLOv9 represent significant advancements in the YOLO architecture, building upon the successes of earlier versions like YOLOv5 and YOLOv7. These newer iterations focus on enhancing the overall performance of the YOLO family by improving both accuracy and efficiency, addressing some of the limitations seen in previous versions.

YOLOv8, introduced with improved accuracy and speed, has brought about several innovations in object detection. One of the key features of YOLOv8 is its ability to better handle small objects, which had been a challenge for earlier versions. The model's architecture includes advanced techniques like feature fusion and multi-scale detection, which help in extracting finer details from the input images and improve the model's ability to detect objects at various scales. These advancements make YOLOv8 particularly effective in applications such as security surveillance, autonomous driving, and industrial inspections, where detecting small or partially obscured objects is crucial. Furthermore, YOLOv8 has been optimized for edge computing devices, allowing for faster inference and lower resource consumption, which is essential for real-time applications that require low latency.

YOLOv9, the latest version, continues to push the boundaries of real-time object detection by further refining the architecture for greater robustness and efficiency. It includes several optimizations, such as improved loss functions and the use of advanced backbone networks that contribute to better feature extraction. YOLOv9 introduces enhanced capabilities for handling challenging environmental conditions, including low-light scenarios and extreme weather, which often cause difficulty for traditional object detection systems. Additionally, YOLOv9 has been designed with a focus on computational efficiency, allowing it to maintain high accuracy and speed even in resource-constrained environments. This makes YOLOv9 particularly suitable for deployment in areas such as mobile robotics and embedded systems, where computational power is often limited but real-time performance is essential.

Table 2 below provides a detailed comparison of popular architectures, including AlexNet, VGG, ResNet, and Inception. These architectures are assessed based on performance metrics such as Top-1 and Top-5 accuracy, parameter count, depth, and their notable features. This comparison highlights the evolution of classification-focused CNN architectures and their distinct advantages, which play a crucial role in selecting the most suitable model for specific classification tasks.

Table 2. Comparison of CNN architectures for object classification.

Architecture	Top-1 accuracy (%)	Top-5 accuracy (%)	Parameters (millions)	Layers	Notable Features
AlexNet	57.1	80.2	60.0	8	First deep CNN for classification, ReLU activation
VGG	71.5	89.9	138.0	16	Deeper network, high parameter count
ResNet	77.0	93.3	25.6	50	Skip connections to avoid vanishing gradient
Inception	78.8	93.9	23.0	22	Multi-scale feature extraction

Source: Based on the study by Tang *et al.* (2023).

AlexNet has become one of the key architectures that has significantly advanced the field of deep learning by dramatically reducing the classification error rate compared to previous methods. However, it's worth noting that deep neural networks for classification already existed before AlexNet, in particular the LeCun LeNet5 network, which was developed in the 1990s (Bangar 2022). LeNet5 consisted of three main components: convolutional, clustering, and linear activation functions, and included seven layers. The main problem at the time was the loss of gradient along deep networks, which made training them much more difficult. With the advent of AlexNet, the use of ReLU activations and deeper architectures significantly reduced the error rate and improved training efficiency, which was an important step in the development of deep learning. Although AlexNet was not the first deep network, its innovations were key to the further development of this field.

Deep CNNs represent an evolution from shallow neural networks, characterized by their significantly larger number of layers and parameters. Unlike shallow methods, which typically involve one or two convolutional layers for simple feature extraction, deep CNNs are designed to capture hierarchical and complex patterns in data by stacking multiple convolutional and pooling layers. This depth allows for the extraction of high-level features essential for complex tasks such as object detection and segmentation.

The specialized techniques in deep CNN architectures address inherent challenges such as the vanishing gradient problem, which becomes more pronounced as the network depth increases. Innovations like batch normalization (to stabilize and accelerate training), skip connections (as introduced in ResNet, to mitigate gradient loss), and advanced loss functions (e.g., focal loss in dense object detection) enable deep networks to learn effectively even with increased complexity.

Furthermore, these architectures incorporate advanced feature extraction techniques to handle a variety of real-world challenges. For example, multi-scale feature extraction, as seen in SSD and YOLO architectures, enables robust detection of objects across different sizes and conditions. These innovations allow deep CNNs to perform effectively in diverse scenarios, such as varying lighting, background noise, and object occlusion, providing a significant advantage over earlier shallow approaches.

Deep neural networks have achieved remarkable advancements across various fields due to key innovations that address critical limitations of traditional models. Batch normalization stabilizes and accelerates training by normalizing the input distribution for each layer. Skip connections, as implemented in ResNet, ensure efficient gradient flow throughout the network, mitigating the vanishing gradient problem in deep architectures. Multi-scale feature extraction, utilized in SSD and YOLO architectures, enhances object detection accuracy across various sizes and challenging conditions. These innovations have made deep neural networks highly effective and adaptable for complex image analysis tasks.

The application of neural network architectures in computer vision opens new opportunities for solving a variety of problems related to image processing. Particularly significant are CNNs, which have become a major tool in this field. These networks consist of convolution layers that allow the automatic extraction and analysis of various features from images (Prayudi *et al.* 2020). The application of CNNs finds wide use in object detection tasks, where the network is trained to recognize and localize objects in an image. In addition, they are successfully used in image segmentation, where it is required to identify each pixel of an image and classify it to belong to a particular object or class.

The RPN is a fundamental innovation in the Faster R-CNN architecture that significantly enhances detection speed by generating region proposals directly within the network. Unlike earlier methods such as R-CNN and Fast R-CNN, which relied on external algorithms like selective search for proposal generation, the RPN integrates this process into the neural network itself, streamlining the overall workflow. The RPN operates as a fully convolutional network that scans the input image and predicts candidate bounding boxes, along with their objectness scores, which indicate the likelihood of the region containing an object. By using predefined anchor boxes of varying sizes and aspect ratios, the RPN effectively detects objects of different scales and shapes in a single forward pass. Additionally, its shared convolutional layers with the main detection network reduce computational overhead, making the approach more efficient.

YOLO, with its multiple iterations such as YOLOv1, YOLOv2, YOLOv3, YOLOv4, and YOLOv5, is renowned for its real-time object detection capabilities (Kouvaras and Petropoulos 2024; Prayudi *et al.* 2020). YOLO processes an image in a single forward pass, dividing it into a grid and predicting bounding boxes and class probabilities simultaneously. Each subsequent version of YOLO has introduced improvements, such as better feature extraction, optimized anchor boxes, and enhanced training strategies, resulting in higher speed and accuracy.

SSD achieves efficient and accurate object detection by combining multi-scale feature extraction with direct predictions of bounding boxes and class scores. Unlike Faster R-CNN, SSD eliminates the need for a separate proposal generation stage, making it more computationally efficient. By using feature maps at multiple scales, SSD excels at detecting objects of varying sizes and performs well in real-time applications.

In addition to object detection, CNN architectures are widely used for image classification tasks, where the goal is to assign an image to a specific category or class. Notable architectures such as AlexNet, VGG, ResNet, and Inception have set benchmarks in image classification by introducing innovations like deeper layers, skip connections, and improved convolutional operations. The general trend in CNN development is to create deeper and more efficient architectures capable of processing large volumes of data while maintaining high accuracy and speed. These advancements have driven the success of CNNs in both object detection and image classification tasks, solidifying their role as a cornerstone of modern computer vision (Chen *et al.* 2023).

Building on the advancements in CNN architectures and their pivotal role in object detection and classification, it is essential to consider how these methods are applied specifically to UAV detection. The unique characteristics of UAVs, their visual detectability under varying conditions, and the methods employed to identify them using computer vision form the core of modern UAV detection systems.

Characteristics, Visual Detectability, and Detection Methods for UAVs

Unmanned aerial vehicles are a variety of devices that do not require a pilot on board to perform tasks. They vary in size, shape, and characteristics, depending on their purpose (Tian *et al.* 2020a). One of the most common types of UAVs is multirotor vehicles equipped with multiple rotating rotors. This includes quadcopters with four rotors, as well as three-, six-, and eight-copters. They can range in size from small, such as the size of the palm of your hand, to large professional models with wingspans of several meters. Multi-rotor UAVs are usually highly maneuverable and capable of hovering in place in the air.

Another type of aircraft is fixed-wing vehicles, similar to those found on conventional aircraft. They provide a longer flight time than multirotor vehicles. They range in size from small radio-controlled models to large vehicles capable of carrying heavy loads and flying long distances. They are often used for monitoring, reviewing, and surveying large areas. There are also hybrid UAVs that combine the features of both multi-rotor and fixed-wing aircraft. These vehicles offer flexibility of use and can combine the advantages of both types (Ariza-Sentís *et al.* 2023). Their characteristics vary depending on their intended use and may include maximum speed, range, and duration of flight, payload capacity, types of sensors and equipment used, and degree of autonomy and protection from external factors such as weather and wind. The visual detectability of drones at different altitudes and distances depends on several factors, including their size, shape, color scheme, and environmental lighting and background.

At low altitudes and close ranges, UAVs can be easily spotted due to their distinctive sounds and movement characteristics. Multirotor vehicles, for example, often emit a characteristic noise from rotating rotors, making them visible even at low altitudes. Fixed-wings UAVs may be less conspicuous at low altitudes due to the lack of sound, but their large size and flight characteristics can attract attention (Lai and Huang 2020). However, at high altitudes and long distances, their visual detectability is reduced.

Multirotor vehicles may become less conspicuous due to their reduced size against the surrounding background and the lack of brightness of LEDs or other markers on the hull. Fixed-wings UAVs, on the other hand, may retain higher detectability due to their larger size and brighter markers. Lighting also plays an important role in the visual detectability of aircraft. In bright sunlight, the contrast between the vehicle and the environment may be low, making them less visible. However, in low-light conditions or when the angle of view changes, the vehicle may become more visible (Castellano *et al.* 2020). Thus, visual detectability at different altitudes and distances depends on many factors and can be vary depending on the viewing conditions.

Methods for detecting objects in images can be divided into two main groups: feature-based methods and methods using object detectors. Feature-based methods are designed to extract characteristic features of objects in images, such as corners, contours, or textures. The classic methods of this group include SIFT and SURF. These methods have low computational requirements and can provide fairly good accuracy in images with a small number of objects. At the same time, they have significant limitations: sensitivity to changes in lighting, scale, and viewing angle (Oyallon and Rabin 2015; Sadou and Njoya 2023).

Adapting and optimizing computer vision models for different drone detection scenarios requires the use of specialized data augmentation methodologies. While data augmentation is often employed to minimize the amount of training data, its primary objective is frequently to achieve class balancing. For instance, in cases where certain drone classes are underrepresented in the dataset, augmentation techniques such as duplicating and transforming samples of the minority class (e.g., through rotation, scaling, or mirroring) can help balance the dataset. This ensures the model does not overfit to more frequently occurring classes, improving its generalization ability across all categories.

In addition to augmentation, simpler techniques like data downsampling can also address imbalances. By reducing the number of samples in overrepresented classes, this method provides a straightforward approach to balancing datasets, particularly when computational resources or data complexity are limited. It is worth noting that in some cases, alternative strategies such as training from scratch, transfer learning, or fine-tuning on pre-trained models can eliminate the need for data augmentation entirely. These methods enable model adaptation to specific tasks by leveraging existing knowledge or highly customized training processes, thereby bypassing the need for extensive augmentation (Sivakumar and Tyj 2021).

Color transformations are an equally important tool in enhancing the robustness of drone detection models. Changing the brightness, contrast, saturation, adding noise, and simulating different color spaces help models become invariant to variations in lighting and camera sensors. This approach is critical for ensuring stable operation of drone detection systems in different weather conditions and times of day. These transformations can be applied effectively as long as spectral information is not crucial for the specific method being used. For example, if only RGB data is being used, almost any transformation can be applied to the image to increase the amount of data and make the training set as general as possible, thus representing the most diverse situations. However, if spectral information is important, such transformations should not be used. More advanced techniques include generative data augmentation methods. Generative Adversarial Networks (GANs) allow for the creation of synthetic images of drones with a high degree of realism. These artificially generated images can fill in gaps in the training data, especially for rare or complex surveillance scenarios. The result of these methodologies is the creation of more versatile and robust computer vision models that can effectively detect and classify drones in a wide range of real-world scenarios (Sonkar *et al.* 2023; Xu *et al.* 2022).

The use of color transformations and generative data augmentation techniques has significantly enhanced the robustness of drone detection models, enabling them to perform reliably under diverse environmental conditions. These advancements have laid the groundwork for more sophisticated approaches in UAV detection. The integration of deep learning techniques, particularly CNNs and GANs, has further revolutionized this field, providing unparalleled accuracy and efficiency in detecting and classifying UAVs.

Advancements in Deep Learning for UAV Detection

Image segmentation identifies each pixel in an image and classifies it as belonging to an object. This method provides accurate object boundaries and the ability to distinguish between objects with overlapping contours but requires high computational resources and is prone to errors under complex imaging conditions (Ferreira *et al.* 2020). To summarize, the choice of object detection method depends on the specific task, accuracy and speed requirements, and available computational resources (Table 3).

Table 3. Comparison of object detection methods.

Method	Pros	Cons
Feature-based	Low requirements for computing resources	Sensitivity to changes in lighting and scale
	Good accuracy on images with a small number of objects	Sensitivity to viewing angles
Object detectors	High-speed image processing	Computationally intensive
	High accuracy of object detection	The need for large datasets for training
Image segmentation	The exact boundary of the objects	High computing resource requirements
	Ability to distinguish between objects with overlapping contours	Prone to errors under difficult surveying conditions

Source: Based on Ferreira *et al.* (2020), Kozachenko (2021), and Leira *et al.* (2020).

Drone detection and classification methods using computer vision, machine learning, and deep learning have different approaches and features. An important technology is image segmentation, which allows you to identify each pixel of an image and classify it as part of a specific object. This method provides accurate object boundaries and the ability to distinguish between objects with overlapping contours, but it requires significant computing resources and is prone to errors in difficult shooting conditions (Luo *et al.* 2023; Terven *et al.* 2023).

The use of deep learning and neural networks has become one of the most effective approaches in modern computer vision technology. Deep learning allows you to create complex neural networks that can automatically extract features from images and learn from large amounts of data. The most common approach to detecting UAVs is to use CNNs. Such networks can process images efficiently and accurately, identifying the characteristic features of objects (Tian *et al.* 2020b).

Neural network architectures such as Faster R-CNN, YOLO, SSD, and Mask R-CNN are actively used to detect UAVs in images. These models are able to work in real-time and provide high accuracy of object detection even in the presence of strong background noise or changes in lighting. The use of deep learning and neural networks allows for the automation and improvement of airspace control, which is especially important in the context of the growing number and diversity of UAVs (Tang *et al.* 2023).

Although the paper primarily focuses on classification and detection, it is important to also acknowledge the significant advancements in segmentation networks, which are designed specifically for tasks requiring pixel-level understanding of images. These architectures play a critical role in many applications, including medical image analysis, autonomous driving, and scene understanding. Some of the most notable segmentation architectures include U-Net, DeepLab, SegNet, Fully Convolutional Networks (FCN), and PSPNet (Yu *et al.* 2023).

U-Net is a widely used architecture in medical image segmentation. It is characterized by its symmetric encoder-decoder structure, with skip connections between corresponding layers in the encoder and decoder. These connections help retain spatial information, making U-Net highly effective for precise pixel-level segmentation tasks. Its ability to work with relatively small datasets while achieving high performance has made it a popular choice in the field of biomedical image analysis.

DeepLab is another powerful architecture for semantic segmentation, known for its use of atrous convolutions (dilated convolutions), which allow the network to capture multi-scale context without losing resolution. DeepLab has undergone several iterations, with DeepLabv3+ being one of the most advanced versions, incorporating encoder-decoder structures and advanced atrous spatial pyramid pooling (ASPP) to improve segmentation accuracy in complex scenarios, such as urban and natural scene segmentation (Vasterling and Meyer 2013).

SegNet is another encoder-decoder network, with a focus on efficient feature extraction and segmentation. Unlike U-Net, SegNet employs max-pooling indices in the decoder to improve segmentation quality while reducing the computational load. It has been particularly useful in applications requiring real-time performance, such as autonomous vehicles and robotic systems. FCNs were among the first architectures to introduce the concept of using fully convolutional layers for semantic segmentation. By replacing the fully connected layers of traditional CNNs with convolutional layers, FCNs can handle input images of any size and generate pixel-wise predictions. This architecture has been foundational in advancing the field of deep learning-based segmentation (Mittal *et al.* 2020).

Pyramid Scene Parsing Network (PSPNet) takes a different approach by using pyramid pooling to capture context at multiple scales. This multi-scale context helps the network understand global scene information, which is essential for accurate segmentation of complex scenes. PSPNet has proven highly effective for large-scale scene parsing tasks, including street scene segmentation for autonomous driving (Minaee *et al.* 2022).

Although some of the networks discussed, such as YOLO and Faster R-CNN, can be adapted to segmentation tasks, their basic concept was developed for object detection and classification. These networks focus on defining the boundaries of objects and their categories, rather than on the exact definition of each pixel of the image, which is the main task of segmentation. Therefore, although they can be adapted for segmentation, they are not optimal for such tasks compared to networks specifically designed for segmentation, such as U-Net or DeepLab.

Adapting drone detection algorithms to specific survey conditions and customer requirements plays a key role in ensuring the effective operation of the surveillance and control system. Environmental conditions, such as lighting, weather, landscape type, and obstacles, must be taken into account. For example, when shooting in low or changing light conditions, detection algorithms must be adapted to work more efficiently. However, such adaptations must be accompanied by a rigorous evaluation of performance to ensure that the algorithms continue to deliver accurate and reliable results under varying conditions.

One of the key tools for evaluating the performance of these detection algorithms is the confusion matrix. Confusion matrices are widely used to assess classification models, particularly in object detection tasks. This tool visually demonstrates how the model classifies objects and helps identify types of errors, such as false positives and false negatives. By using confusion matrices, it becomes easier to understand whether the model is working correctly in real-world conditions and how it can be further optimized to improve results. Below is an example of a confusion matrix and its components for drone detection, which highlights its importance as an effective instrument for evaluating the accuracy of the model (Samaras *et al.* 2019; Shaharom and Tahar 2023).

Customer requirements may vary depending on the specific task and application. In the security sector, a customer may be interested in high detection accuracy and speed, while in natural resource monitoring, they may be interested in reliability and the ability to work in different climatic conditions. Therefore, algorithms must be customized and optimized to meet specific requirements (Macukow 2016; Taha and Shoufan 2019).

The types of drones are extremely diverse and are classified according to many parameters. They can differ in size – from micro to large devices, in purpose – military, commercial, or entertainment, in configuration – multi-rotor, airplane-type, or hybrid. Drones differ in flight range, payload capacity, and type of power plant – electric, gasoline, or hybrid. Thus, adapting UAV detection algorithms to specific survey conditions and customer requirements involves considering the environment, customizing them for specific tasks, taking into account the characteristics of the objects to be detected, and optimizing the use of resources. This ensures efficient and reliable operation of the detection system in various conditions and applications.

Integration of Multi-Sensor Technologies

Metrics for evaluating the performance of UAV detection algorithms are essential tools for analyzing and comparing their effectiveness under different conditions and application scenarios. Accuracy, for instance, measures how correctly an algorithm detects UAVs, calculated as the ratio of correctly detected aircraft to the total number of objects flagged as UAVs. The higher the accuracy, the fewer the false positives. Completeness, on the other hand, assesses how well the algorithm detects all real-world UAVs, defined as the ratio of correctly detected UAVs to the total number of UAVs present. A higher completeness means fewer missed detections.

Given the importance of these metrics in assessing algorithm performance, it is crucial to regularly calibrate and update UAV detection systems to adapt to changing conditions, such as varying lighting, weather, and new UAV designs. These updates ensure that detection systems remain reliable, accurate, and efficient over time.

Looking ahead, the integration of additional technologies, such as LIDAR and radar, will play a vital role in enhancing UAV detection capabilities. While computer vision algorithms are integral to real-time detection, multi-sensor approaches can significantly boost efficiency and reliability. Radar systems, for instance, are invaluable for detecting objects at long ranges, even in challenging weather conditions and low visibility. In combination with computer vision, radar data provides spatial coordinates and velocity information, improving overall system accuracy.

LIDAR technology complements these sensors by offering extremely detailed spatial information through laser scanning. LIDAR can generate highly accurate three-dimensional maps of the terrain, enabling precise detection of UAV shapes, sizes, and positions. This synergy between multiple sensors will enhance the robustness and adaptability of UAV detection systems, ensuring higher reliability and efficiency in diverse environments (Jiang *et al.* 2022).

The F-Score is the harmonic mean between accuracy and completeness. It covers both metrics in a single numerical value and provides an overall measure of the algorithm's performance. The higher the F-Score, the better the combination of accuracy and completeness. The detection rate is the time it takes for the algorithm to detect aircraft in an image or video stream. This metric is important for tasks where a fast response to object detection is required, such as in security systems or airspace monitoring. False positives are the number of objects that the algorithm mistakenly flagged as UAVs when in fact they are not. This metric is important for evaluating unwanted false positives that can affect the reliability of the detection system. Missed detections are the number of actual UAVs that the algorithm failed to detect. This metric is important for assessing detection completeness and evaluating potential gaps in the detection system.

The analysis of these metrics provides a comprehensive evaluation of the performance of the algorithms and determines their suitability for specific tasks and application conditions. When comparing different drone detection methods and algorithms, it is important to consider their features, advantages, and disadvantages (Table 1). This table provides an overview of the main methods and their characteristics, which helps make an informed decision about the most appropriate method for a particular aircraft detection task.

A confusion matrix is a valuable tool for evaluating the performance of classification models, particularly in tasks like UAV detection. It provides a clear representation of the model's predictions by categorizing them into four outcomes: true positives (correctly identified UAVs), false positives (objects incorrectly identified as UAVs), false negatives (missed UAVs), and true negatives (correctly identified non-UAVs). This framework simplifies the explanation of key concepts such as model accuracy and error types. The confusion matrix facilitates the calculation of essential metrics like sensitivity (the ability to detect all actual UAVs) and specificity (the ability to correctly identify non-UAVs), offering a comprehensive assessment of the model's effectiveness in real-world applications (Bouguettaya *et al.* 2022; Kaur *et al.* 2021).

One of the main advantages of modern computer vision algorithms is their high detection accuracy. However, accuracy alone may not fully capture the specific requirements of critical systems. In addition to accuracy, other key metrics such as recall (sensitivity), mean average precision (mAP), mAP50, and the F1-Score are also important, as these metrics provide a more comprehensive evaluation of model performance in complex scenarios. These metrics can be especially valuable in applications like UAV detection, where balancing detection quality and minimizing false positives or false negatives is crucial for system safety, defined here as the ability to accurately identify and respond to threats while minimizing errors that could lead to undetected UAVs or incorrect actions. Furthermore, in critical systems, it is important to consider uncertainty factors and ensure redundancy in the detection process. This redundancy helps maintain safety and operation in the event of a detection failure, ensuring more reliable and continuous performance.

Another significant advantage is the wide range of applications of computer vision technology. It can be used not only for UAV detection for security purposes but also in environmental monitoring, area protection, and in the aviation and transport industries to detect and track vehicles. However, besides the advantages, there are some disadvantages to applying computer vision for UAV detection (Iqbal *et al.* 2024; Kakaletsis *et al.* 2021). These may include the need for high computational resources, the dependency on environmental conditions such as lighting and weather, and the potential for false positives or negatives in complex scenarios. Addressing these challenges is crucial for the effective deployment of computer vision systems in real-world applications (Kouvaras and Petropoulos 2024).

Some computer vision methods, especially those based on deep learning, require significant computational resources for training and inference. This may require the use of powerful computing systems and infrastructure (Poplavskyi 2024). In addition, computer vision methods may be sensitive to imaging conditions such as changes in lighting, weather conditions, and other environmental factors. This can degrade the performance of algorithms under uncontrolled conditions, requiring additional efforts to adapt and optimize the methods for different imaging scenarios.

It is also worth noting that to maintain the effectiveness of the aircraft detection system, the computer vision algorithms must be calibrated and updated regularly to meet changing conditions and task requirements. This requires additional time and resources to maintain and support the system. Overall, despite some limitations, the application of computer vision for commercial UAV detection is an effective and promising approach that can significantly improve safety and efficiency in various applications.

The prospects for further research in the application of computer vision for commercial UAV detection are promising and offer significant benefits in various aspects. One of the key areas of focus in this area is to continuously improve the accuracy and reliability of detection algorithms. The development of new computer vision techniques and the improvement of existing ones will lead to more accurate results, which is essential to ensure the safety and efficiency of the detection system.

The integration of LIDAR into drone detection systems provides unique advantages that dramatically increase the efficiency of airspace monitoring. LIDAR provides extremely precise three-dimensional spatial information, allowing for the instantaneous determination of geometric parameters of drones with millimeter accuracy. Unlike traditional optical systems, LIDAR technology can clearly identify the shape, size, and spatial orientation of unmanned vehicles even in difficult environmental conditions. The key advantage of the integration approach is the ability to detect drones in conditions where traditional technologies are ineffective. This is especially relevant for ensuring the security of critical infrastructure, airports, sensitive facilities, and mass events where the most complete control of the airspace is required (Poplavskyi 2024).

Infrared sensors are vital components in UAV detection systems, as they enable effective monitoring in low visibility conditions such as at night or during poor lighting. These sensors operate by detecting thermal radiation, allowing objects to be identified based on their temperature differences from the surrounding environment. Infrared sensors are especially valuable for detecting drones in challenging weather conditions, such as fog, rain, or snow, where traditional optical cameras may struggle to provide clear images.

The integration of infrared sensors with other sensor types, such as video cameras and radars, creates more versatile and reliable drone detection systems, significantly enhancing both accuracy and detection efficiency in a variety of environmental conditions (Chen *et al.* 2023). Infrared sensors contribute additional data that complements other sensor outputs, enabling more detailed and real-time observation of objects.

The advantages of infrared sensors lie in their ability to detect objects in conditions where optical sensors are ineffective, such as low-light or in difficult weather. They also offer a non-intrusive method for detecting drones, which is crucial for maintaining security without disrupting the surrounding environment (Du *et al.* 2022). Infrared sensors are highly effective for continuous monitoring and can operate autonomously in real-time, making them crucial for applications such as critical infrastructure protection, security monitoring, and airspace control. Their importance in providing redundancy in detection systems, where reliability and safety are paramount, cannot be overstated.

Overall, further research in this area will focus on the development of more accurate, faster, and more reliable UAV detection systems to enable their effective use in various fields such as security, monitoring and control, and in the transport industry. Summarizing the research on the application of computer vision to commercial UAV detection, several important conclusions and generalizations can be drawn.

The integration of computer vision with other sensory data, such as radar and infrared sensors, plays a crucial role in enhancing detection systems. By combining these technologies, we can achieve more complete and efficient systems capable of operating effectively in a variety of environments and scenarios. This multi-sensor approach is essential not only for improving detection accuracy but also for adapting to challenging and changing conditions.

The future of real-time UAV detection systems lies in further advancing these technologies. A key focus of this development will be to optimize computer vision algorithms to ensure the fastest and most accurate detection of UAVs, even in highly dynamic and complex environments. The goal is to create intelligent systems capable of instantaneously recognizing and classifying drones. However, the development of such systems is not limited to computer vision alone. The integration of computer vision with other technologies, such as radar, acoustic, and thermal imaging systems, is vital for increasing detection reliability and minimizing recognition errors. This approach will allow for more robust and efficient UAV detection systems that can operate in a broader range of conditions.

One such technology that promises to significantly improve UAV detection is LIDAR. By combining LIDAR with other sensors like radar, LIDAR offers extremely detailed spatial information, enabling the precise detection of UAVs even in conditions where optical cameras and traditional detection methods may fail. Radar, for example, has already demonstrated its ability to detect drones at long distances, even in poor visibility conditions such as fog or rain, where optical sensors struggle. This data fusion approach, which integrates radar, LIDAR, and infrared cameras, is poised to enhance the effectiveness of UAV detection systems, ensuring they are both accurate and reliable across a variety of operational scenarios.

Practical results show that combined systems using radar, LIDAR, and infrared sensors achieve much better outcomes compared to using a single sensor. For example, one system that integrated these technologies demonstrated a 95% detection accuracy while simultaneously reducing false positives to 5%, which significantly surpasses the results of individual sensors (Du *et al.* 2022; Poplavskyi 2024). This highlights the importance of integrating data from different sensors to achieve high accuracy and reliability in the system. Data fusion compensates for the weaknesses of individual sensors: radar works well in poor visibility but does not provide detailed characteristics of objects, while LIDAR and infrared cameras offer additional capabilities for precise detection under various conditions. The use of neural networks to process fused data enables real-time drone detection, which is crucial for rapid response in complex scenarios.

Experimental and practical results confirm the effectiveness of data fusion methods from different sensors in creating reliable drone detection systems. They demonstrate significant advantages over the use of individual sensors, particularly in improving accuracy, reducing false positives, and enhancing operational efficiency in challenging conditions.

DISCUSSION

The results of this study confirm the significance of using computer vision to detect commercial drones to improve safety and efficiency in various applications. The study determined that the use of advanced computer vision techniques can provide high detection accuracy and system reliability. Similar findings are obtained in studies conducted by other researchers dealing with similar topics. For instance, the study by Akbari *et al.* (2021) on the application of computer vision to analyze videos and images captured by drones highlights the importance of using technology to improve UAV functionality and safety.

However, the results of this study are peculiar due to the focus on a specific application area – UAV detection. While the study by the researchers covers a wide range of applications of data from drones, the analysis focuses on the specific task of object detection. Thus, while both studies support the importance of using computer vision to improve safety and efficiency in the aviation industry, the results of this study add value by drawing attention to specific aspects of the application of this technology in UAV detection.

A study conducted by Kakaletsis *et al.* (2021) and this study share similar aspects in that they both address safety issues in the context of drone use and recognize the importance of integrating safety knowledge into drone algorithms and architectures. Both studies also consider the role of computer vision in improving the safety and efficiency of drone use. However, while the study by the researchers focuses on analyzing the increasing use of autonomous drones and the importance of legal regulation in this area, the study written above focuses on analyzing and optimizing computer vision techniques for detecting commercial UAVs to improve safety and control their use in various industries.

Chelluri and Manjunathachari (2019) and Mittal *et al.* (2020) reviewed state-of-the-art object detection algorithms and their applicability to low-altitude drone data. The main objective of this study is to survey and analyze algorithms such as Faster R-CNN, YOLO, SSD, and RetinaNet and their applicability to specific low-altitude data. In contrast, this study focuses on the analysis and optimization of computer vision techniques for the detection of commercial UAVs in various industries. This study seeks to develop more efficient and robust UAV detection systems, making them more applicable and practical in various fields of endeavor. The approach in the study not only considers existing algorithms but also proposes adaptation and optimization of these methods to suit specific needs and survey conditions. Thus, unlike the study by the researchers, which concentrates on algorithm review, this study aims to provide applicable solutions for UAV detection in real-world environments.

Perry and Guo (2021) present a new remote sensing approach to measure the dynamic displacement of three-dimensional structures using a sensor system on UAVs and optical and infrared cameras. This research significantly contributes by proposing the integration of different types of cameras to measure three-component displacement and the development of new data processing algorithms to extract information from video. Its performance has been validated by laboratory experiments, indicating its potential relevance in the field of measuring the dynamic structural response of three-dimensional structures. Both studies highlight the need for further research and innovation in commercial UAV detection. However, the researchers focus on the development of a new sensing method using optical and infrared cameras on UAVs, while this study offers general conclusions and recommendations for improving UAV detection systems, including computer vision algorithms, integration of data from different sensors, and system security.

A study by Ramachandran and Sangaiah (2021) and this study highlight the role of computer vision in the context of UAVs and its application to real-time object detection and tracking. Both studies emphasize the importance of these tasks for monitoring different environments and identifying gaps in existing research, which helps to identify directions for future research. The methodology of both studies includes a detailed literature review on object detection and tracking using UAVs and the development of methods to detect objects in UAV images. Both studies also enumerate specific datasets for these tasks and summarize existing research work in different UAV applications. However, the study of the researchers is more focused on the literature review and classification of object detection methods in UAV images, while this study is more specifically focused on analyzing the performance of UAV detection methods and their suitability for specific tasks and applications. Thus, although both studies address similar topics, they have different biases and approaches to analyzing the issues under investigation.

This study focuses on analyzing the findings and recommendations for improving UAV detection systems. It notes that to improve the performance and reliability of such systems, improvements in computer vision algorithms and integration of data from different sensors such as radar, LIDAR, and infrared sensors are needed. It also emphasizes the importance of real-time operation and highlights the potential threats and risks associated with the use of such systems, which requires further research in cybersecurity and data protection.

CONCLUSION

The study highlighted the significant benefits of using computer vision to detect drones, including its ability to provide high accuracy and rapid response to threats. This is an important aspect of the work, as it emphasizes the importance of the system's operational effectiveness in detecting and responding to potential threats in real-time. Computer vision for drone detection not only ensures high accuracy in UAV identification, but also guarantees a quick response, which is critical for applications that require immediate action, such as security and surveillance. In the context of the study, high accuracy is only one of the key metrics used to evaluate the effectiveness of drone detection systems. While accuracy plays a crucial role, other performance metrics such as recall, precision, and F1-Score are equally important and should be considered when assessing the overall performance of the system. These metrics have been discussed throughout the study, providing a more complete understanding of the system's capabilities beyond simple accuracy. It is important to emphasize the importance of these additional metrics as they help to provide a complete picture of the strengths and weaknesses of the detection system.

The study has determined that deep learning architectures, particularly deep CNNs, play a central role in solving UAV detection tasks. This terminology aligns with the specific usage in the paper, where deep CNNs are highlighted for their ability to automatically extract hierarchical features from visual data, significantly enhancing the accuracy and reliability of detection systems. These architectures are especially effective in challenging conditions, such as varying lighting, poor visibility, or different background obstacles, making them essential for robust and adaptive UAV detection systems.

Data fusion involves not only integrating diverse inputs into a model for classification and detection but also combining the results and methods used during the detection process. This means that fusion can occur at different stages, from pre-processing, where different types of data can be combined to create a more comprehensive view, to combining the output from multiple models after processing. In addition, fusion can include the integration of different detection methods, allowing the strengths of each method to be combined to achieve more accurate and reliable results.

Deep learning, particularly CNNs, has greatly improved the efficiency of processing large volumes of data and enabled the implementation of more complex and adaptive UAV detection methods. This confirms the importance of such technologies for achieving high results in real-time, particularly through integration with other sensors such as radar and LIDAR. At the same time, despite their numerous advantages, the use of deep networks requires significant computational resources, which must be considered when developing real-world detection systems.

The importance of adapting UAV detection algorithms to specific survey conditions and customer requirements such as lighting, weather, landscape type, and object specificity was emphasized. It is necessary to consider the different needs of customers and optimize the algorithms to meet their requirements. It was also found that to maintain the effectiveness of the drone detection system, it is important to regularly calibrate and update the computer vision algorithms according to changing conditions and task requirements. This may require additional time and resources, but it is necessary to ensure reliable system performance. The application of computer vision to the detection of commercial UAVs promises to be an effective and promising approach that can significantly improve safety and efficiency in a variety of applications. The development of new methods and the improvement of existing methods will enable more accurate results to be achieved.

The visual detectability of UAVs at different altitudes and distances is described, considering factors such as size, shape, color scheme, lighting, and environmental background. The study reveals the future of real-time UAV detection systems and suggests a direction for further research: improving the adaptability of detection systems to different imaging conditions, including lighting, climatic zones, and UAV movement scenarios. This direction will develop more efficient and robust detection systems, contributing to the safety and efficiency of UAV applications in various fields. Further research in these areas can lead to more efficient and reliable UAV detection systems, which in turn contributes to improved safety and efficiency in various UAV applications.

CONFLICT OF INTEREST

Nothing to declare.

AUTHORS' CONTRIBUTION

Conceptualization: Grichshenko V; **Methodology:** Mukushev A, Kokidko A, and Zikiryaev N; **Software:** Kokidko A and Zikiryaev N; **Validation:** Mukushev A and Kokidko A; **Formal analysis:** Grichshenko V and Zikiryaev N; **Investigation:** Mukushev A and Zikiryaev N; **Resources:** Grichshenko V and Mukushev A; **Data Curation:** Grichshenko V and Kokidko A; **Writing - Original Draft:** Grichshenko V and Mukushev A; **Writing - Review & Editing:** Grichshenko V, Mukushev A, Kokidko A, and Zikiryaev N; **Visualization:** Kokidko A and Zikiryaev N; **Supervision:** Mukushev A; **Final approval:** Mukushev A.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author.

FUNDING

Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan.
Grant No: BR109012/0221.

ACKNOWLEDGMENTS

Not applicable.

REFERENCES

Akbari Y, Almaadeed N, Al-Maadeed S, Elharrouss O (2021) Applications, databases and open computer vision research from drone videos and images: A survey. Artif Intell Rev 54:3887-3938. https://doi.org/10.1007/s10462-020-09943-1

Ariza-Sentís M, Baja H, Vélez S, Valente J (2023) Object detection and tracking on UAV RGB videos for early extraction of grape phenotypic traits. Comput Electron Agric 211:108051. https://doi.org/10.1016/j.compag.2023.108051

Bangar S (2022) LeNet 5 Architecture Explained. Medium. [accessed Feb 26 2024]. https://medium.com/@siddheshb008/lenet-5-architecture-explained-3b559cb2d52b

Bay H, Ess A, Tuytelaars T, Van Gool L (2006) SURF: speeded up robust features. Comput Vis Image Underst 110(3):346-359. https://doi.org/10.1007/11744023_32

Bazeltsev KO (2020) UAV navigation software based on openCV and tensorflow. Kyiv: National Aviation University. https://dspace.nau.edu.ua/handle/NAU/47729

Borges SF de S, Cardoso Júnior MM, Castilho DS (2024) Method for defining the automation level of an eVTOL. J Aerospace Technol Manag 16:e2224. http://doi.org/10.1590/jatm.v16.1342

Borodin V, Selin A, Kolesnichenko V, Kalyagin M (2024) Selection of routing metrics and service channel characteristics of ad hoc network for UAV swarm. J Aerospace Technol Manag 16:e2324. http://doi.org/10.1590/jatm.v16.1343

Bouguettaya A, Zarzour H, Kechida A, Taberkit AM (2022) Vehicle detection from UAV imagery with deep learning: a review. IEEE Trans Neural Netw Learn Syst 33(11):6047-6067. https://doi.org/10.1109/TNNLS.2021.3080276

Castellano G, Castiello C, Mencar C, Vessio G (2020) Crowd detection for drone safe landing through fully-convolutional neural networks. In: SOFSEM 2020: theory and practice of computer science. Cham: Springer. https://doi.org/10.1007/978-3-030-38919-2_25

Cazzato D, Cimarelli C, Sanchez-Lopez JL, Voos H, Leo M (2020) A survey of computer vision methods for 2D object detection from unmanned aerial vehicles. J Imaging 6(8):78-89. https://doi.org/10.3390/jimaging6080078

Chelluri HB, Manjunathachari K (2019) SIFT and it's variants: an overview. Paper presented 2019 Proceedings of International Conference on Sustainable Computing in Science, Technology and Management. Amity University Rajasthan; Jaipur, India.

Chen C, Zheng Z, Xu T, Guo S, Feng S, Yao W, Lan Y (2023) Yolo-based UAV technology: a review of the research and its applications. Drones 7(3):190. https://doi.org/10.3390/drones7030190

Douklias A, Karagiannidis L, Misichroni F, Amditis A (2022) Design and implementation of a UAV-based airborne computing platform for computer vision and machine learning applications. Sensors 22(5):2049. https://doi.org/10.3390/s22052049

Du K-L, Legun C-S, Mow WH, Swamy MNS (2022) Perceptron: learning, generalization, model selection, fault tolerance, and role in the deep learning era. Mathematics 10(24):4730. https://doi.org/10.3390/math10244730

Ferreira PM, de Almeida DRA, Papa DA, Minervino JBS, Veras HFP, Formighieri A, Santos CAN, Ferreira MAD, Figueiredo EO, Ferreira EJL (2020) Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. For Ecol Manage 475:118397. https://doi.org/10.1016/j.foreco.2020.118397

Iqbal U, Davies T, Perez P (2024) A review of recent hardware and software advances in GPU-accelerated edge-computing single-board computers (SBCs) for computer vision. Sensors 24(15):4830. https://doi.org/10.3390/s24154830

Jiang C, Ren H, Ye X, Zhu J, Zeng H, Nan Y, Sun M, Ren S, Huo H (2022) Object detection from UAV thermal infrared images and videos using YOLO models. Int J Appl Earth Obs Geoinf 112:102912. https://doi.org/10.1016/j.jag.2022.102912

Kakaletsis E, Symeonidis C, Tzelepi M, Mademlis I, Tefas A, Nikolaidis N, Pitas I (2021) Computer vision for autonomous UAV flight safety: an overview and a vision-based safe landing pipeline example. ACM Comput Surv 54(9):181. https://doi.org/10.1145/3472288

Kaur R, Kumar R, Gupta M (2021) Review on transfer learning for convolutional neural network. Paper presented 2021 3rd International Conference on Advances in Computing, Communication Control and Networking. Institute of Electrical and Electronics Engineers; Greater Noida, India. https://doi.org/10.1109/ICAC3N53548.2021.9725474

Kouvaras L, Petropoulos GP (2024) A novel technique based on machine learning for detecting and segmenting trees in very high-resolution digital images from unmanned aerial vehicles. Drones 8(2):43. https://doi.org/10.3390/drones8020043

Kozachenko DV (2021) Development of the concept of optical channel of secure communication with quadrocopters in the conditions of active counteraction. Vinnytsia: Vasyl Stus Donetsk National University.

Lai YC, Huang ZY (2020) Detection of a moving UAV based on deep learning-based distance estimation. Remote Sens 12(18):3035. https://doi.org/10.3390/rs12183035

Leira FS, Helgesen HH, Johansen TA, Fossen TI (2020) Object detection, recognition, and tracking from UAVs using a thermal camera. J Field Robot 38(2):242-267. https://doi.org/10.1002/rob.21985

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comp Vis 60(2):91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Luo K, Kong X, Zhang J, Hu J, Li J, Tang H (2023) Computer vision-based bridge inspection and monitoring: a review. Sensors 23(18):7863. https://doi.org/10.3390/s23187863

Macukow B (2016) Neural networks – State of art, brief history, basic models and architecture. Computer information systems and industrial management. Paper presented 2016 15th IFIP TC8 International Conference. IFIP; Vilnius, Lithuania. https://doi.org/10.1007/978-3-319-45378-1_1

Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D (2022) Image segmentation using deep learning: a survey. IEEE Trans Pattern Anal Mach Intell 44(7):3523-3542. https://doi.org/10.1109/TPAMI.2021.3059968

Mykhalevskiy D, Vasylyshyn V, Riabkov V, Myronenko R, Bryl D (2024) Method for improving the coverage efficiency of wireless sensor networks based on UAVs. Machinery & Energetics 15(2):81-94. https://doi.org/10.31548/machinery/2.2024.81

Mittal P, Singh R, Sharma A (2020) Deep learning-based object detection in low-altitude UAV datasets: a survey. Image Vis Comput 104:104046. https://doi.org/10.1016/j.imavis.2020.104046

Oyallon E, Rabin J (2015) An analysis of the SURF method. Image Process on Line 5:176-218. http://org/10.5201/ipol.2015.69

Pawełczyk MŁ, Wojtyra M (2020) Real world object detection dataset for quadcopter unmanned aerial vehicle detection. IEEE Access 8:174394-174409. http://doi.org/10.1109/ACCESS.2020.3026192

Perry BJ, Guo Y (2021) A portable three-component displacement measurement technique using an unmanned aerial vehicle (UAV) and computer vision: a proof of concept. Measurement 176:109222. https://doi.org/10.1016/j.measurement.2021.109222

Poplavskyi O (2024) Information technology for image data processing based on hybrid neural networks using geometric features. Inf Technol Comput Eng 21(2):4-16. https://doi.org/10.31649/1999-9941-2024-60-2-4-16

Prayudi A, Sulistijono IA, Risnumawan A, Darojah Z (2020) Surveillance system for illegal fishing prevention on UAV imagery using computer vision. Paper presented 2020 International Electronics Symposium. IEEE; Surabaya, Indonesia. http://doi.org/10.1109/IES50839.2020.9231539

Ramachandran A, Sangaiah AK (2021) A review on object detection in unmanned aerial vehicle surveillance. Int J Cogn Comput Eng 2:215-228. https://doi.org/10.1016/j.ijcce.2021.11.005

Sadou AM, Njoya ET (2023) Applications of artificial intelligence in the air transport industry: a bibliometric and systematic literature review. J Aerospace Technol Manag 15:e2223. http://doi.org/10.1590/jatm.v15.1312

Samaras S, Diamantidou E, Ataloglou D, Sakellariou N, Vafeiadis A, Magoulianitis V, Lalas A, Dimou A, Zarpalas D, Votis K, et al. (2019) Deep learning on multi sensor data for counter UAV applications – A systematic review. Sensors 19(22):4837. https://doi.org/10.3390/s19224837

Shaharom MF, Tahar KN (2023) Multispectral image matching using SIFT and SURF algorithm: a review. Int J Geoinformatics 19(1). https://doi.org/10.52939/ijg.v19i1.2495

Shantyr A (2024) Specifics of quality assessment models application at development and use stages of software systems. Inf Technol Comput Eng 21(1):127-138. https://doi.org/10.31649/1999-9941-2024-59-1-127-138

Sivakumar M, Tyj NM (2021) A literature survey of unmanned aerial vehicle usage for civil applications. J Aerospace Technol Manag 13:e4021. http://doi.org/10.1590/jatm.v13.1233

Sonkar S, Kumar P, Puli YT, George RC (2023) Design & implementation of an electric fixed-wing hybrid VTOL UAV for asset monitoring. J Aerospace Technol Manag 15:e0823. http://doi.org/10.1590/jatm.v15.1297

Taha B, Shoufan A (2019) Machine learning-based drone detection and classification: state-of-the-art in research. IEEE Access 7:138669-138682. http://doi.org/10.1109/ACCESS.2019.2942944

Tang G, Tang G, Ni J, Zhao Y, Gu Y, Cao W (2023) A survey of object detection for UAVs based on deep learning. Remote Sens 16(1):149. https://doi.org/10.3390/rs16010149

Terven J, Cordova-Esparza DM, Ramirez-Pedraza A, Chavez-Urbiola EA, Romero-Gonzalez JA (2023) Loss functions and metrics in deep learning. A review. arXiv:2307.02694.. https://doi.org/10.48550/arXiv.2307.02694

Tian H, Wang T, Liu Y, Qiao X, Li Y (2020a) Computer vision technology in agricultural automation – A review. Inf Process Agric 7(1):1-19. https://doi.org/10.1016/j.inpa.2019.09.006

Tian Y, Zhang C, Jiang S, Zhang J, Duan W (2020b) Noncontact cable force estimation with unmanned aerial vehicle and computer vision. Comput Aided Civ Infrastruct Eng 36(1):73-88. https://doi.org/10.1111/mice.12567

Vasterling M, Meyer U (2013) Challenges and opportunities for UAV-borne thermal imaging. In: Kuenzer C, Dech S, editors. Thermal infrared remote sensing. Remote sensing and digital image processing. Vol 17. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-6639-6_4

Xu X, Zhang L, Yang J, Cao C, Wang W, Ran Y, Tan Z, Luo M (2022) A review of multi-sensor fusion SLAM systems based on 3D LIDAR. Remote Sensing 14(12):2835. https://doi.org/10.3390/rs14122835

Yermolenko R, Klekots D, Gogota O (2024) Development of an algorithm for detecting commercial unmanned aerial vehicles using machine learning methods. Machinery & Energetics 15(2):33-45. https://doi.org/10.31548/machinery/2.2024.33

Yu Y, Wang C, Fu Q, Kou R, Huang F, Yang B, Yang T, Gao M (2023) Techniques and challenges of image segmentation: a review. Electronics 12(5):1199. https://doi.org/10.3390/electronics12051199

Zhao J, Zhang J, Li D, Wang D (2022) Vision-based anti-UAV detection and tracking. IEEE Trans Intell Transp Syst 23(12):25323-25334. http://doi.org/10.48550/arXiv.2205.10851