Enhancing unmanned aerial vehicle capabilities: integrating YOLO algorithms for diverse industrial applications

Nikolai Guliutin¹*, and Oleslav Antamoshkin²,³

¹Siberian Federal University, Artificial Intelligence Systems Department, ul. Akademika Kirenskogo 26k1, Krasnoyarsk, 660074, Russia
²Siberian Federal University, Artificial Intelligence Laboratory, ul. Akademika Kirenskogo 26, Krasnoyarsk, 660074, Russia
³Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Ave, Krasnoyarsk, 660037, Russia

Abstract. The integration of UAVs with advanced deep learning algorithms, particularly the You Only Look Once models, has opened new horizons in various industries. This paper explores the transformative impact of YOLO-based systems across diverse sectors, including agriculture, forest fire detection, ecology, marine science, target detection, and UAV navigation. We delve into the specific applications of different YOLO models, ranging from YOLOv3 to the lightweight YOLOv8, highlighting their unique contributions to enhancing UAV functionalities. In agriculture, UAVs equipped with YOLO algorithms have revolutionized disease detection, crop monitoring, and weed management, contributing to sustainable farming practices. The application in forest fire management showcases the capability of these systems in real-time fire localization and analysis. In ecological and marine sciences, the use of YOLO models has significantly improved wildlife monitoring, environmental surveillance, and resource management. Target detection studies reveal the efficacy of YOLO models in processing complex UAV imagery for accurate and efficient object recognition. Moreover, advancements in UAV navigation, through YOLO-based visual landing recognition and operation in challenging environments, underscore the versatility and efficiency of these integrated systems. This comprehensive analysis demonstrates the profound impact of YOLO-based UAV technologies in various fields, underscoring their potential for future innovations and applications.

1 Introduction

In recent years, the fusion of Unmanned Aerial Vehicles (UAVs) with the You Only Look Once (YOLO) algorithm has revolutionized various industries, showcasing an impressive range of applications. The YOLO models, known for their efficiency in object detection, have

* Corresponding author: gulyutin@gmail.com
been tailored and optimized for UAV imagery, leading to significant advancements in fields such as agriculture, forest fire detection, ecology, marine science, and UAV navigation.

YOLO, short for "You Only Look Once," represents a major advancement in the field of computer vision, especially in object detection tasks. This method stands out for its speed and efficiency in detecting objects in images, making it a popular choice in numerous real-world applications such as autonomous vehicles, security systems, and various automated inspection and monitoring tasks.

YOLO was first introduced by Joseph Redmon et al. [1] in a 2015 paper, marking a departure from the traditional two-step approach of first proposing regions and then classifying them. It unified these steps into a single neural network model, significantly speeding up the process. YOLO divides the input image into a grid, with each grid cell predicting several bounding boxes and confidence scores for these boxes. These scores reflect the system's certainty that the box contains an object and the accuracy of the box. Additionally, it predicts the class probabilities of the detected object.

In agriculture [3-8], UAVs equipped with YOLO models have enhanced monitoring and disease detection. Studies have shown the algorithm's effectiveness in early disease detection in crops like rice, olives, maize, and apple trees. This advancement in precision agriculture allows for early intervention, potentially preventing widespread crop damage.

Forest fire detection [9-13] has also benefited greatly from the integration of YOLO models with UAV technology. Various studies have employed YOLO-based models for efficient and accurate fire detection using UAV aerial images. These models have been crucial in real-time geo-localization of fires, allowing for rapid assessment and emergency response. They have also proven effective in identifying forest fires through infrared UAV imagery, enabling precise discrimination of fire-affected areas.

In ecology and marine science [14-18], UAVs combined with YOLO models have made strides in wildlife monitoring, riverbank inspection, and marine surveillance. Innovations like the SE-YOLO model, which incorporates a channel self-attention mechanism into YOLOv7, have enhanced the capability to detect small targets from aerial images, crucial for wildlife conservation. UAVs have also been utilized in environmental surveillance, particularly in inaccessible areas, using image processing techniques like the scale-invariant feature transform (SIFT) algorithm.

Target detection [19-22] in UAV imagery has seen advancements with the development of more efficient and accurate neural network methods. These methods, including adaptations like Drone-YOLO and YOLO-ViT-based models, have been designed to overcome challenges like detecting small objects and ensuring computational efficiency for real-time processing on UAV platforms.

Enhanced object [23-27] detection and tracking capabilities in UAVs have been crucial in evolving their functionalities for applications ranging from environmental monitoring to search and rescue operations. Adaptations of the YOLO v8 algorithm for small object
detection and the integration of YOLO's object detection with onboard edge computing have improved UAVs' operational capabilities in various scenarios.

Overall, these developments highlight the growing importance of UAV technology and advanced image processing models like YOLO in a wide range of applications. The continued evolution of these technologies promises further advancements in these critical areas, making UAVs more versatile and effective tools in the modern technological landscape.

2 Usage of YOLO algorithm for UAVs application in various industries

2.1 Agriculture

The integration of UAVs with advanced deep learning algorithms, particularly the YOLO models, has revolutionized agricultural monitoring and disease detection. This synergy is evident in a series of recent studies that showcase the varied applications of UAV-based imagery analysis using YOLO models.

Lin et al. [4] have made strides in early disease detection in rice using UAV imagery. Their approach involves a semi-supervised model that enhances the capability to identify diseases from aerial images. This method marks a significant advancement in precision agriculture, leveraging technology to detect problems at their nascent stage, potentially saving crops from widespread damage.

Similarly, Mamalis et al. [5] have applied the YOLO algorithm for detecting Verticillium fungus in olive trees, a critical issue in olive cultivation. Their work demonstrates the algorithm's effectiveness in processing UAV imagery to identify specific plant diseases with high precision, showcasing the potential of YOLO in targeted disease management.

Pu et al.'s [6] contribution lies in developing Tassel-YOLO, a method focusing on real-time, high-precision maize tassel detection and counting. This approach highlights the versatility of YOLO in crop monitoring, offering a tool for farmers and researchers to accurately assess crop growth and development.

Jemaa et al. [7] have further expanded the application of this technology by developing a UAV-based computer vision system for apple tree health assessment in orchards. Their work underlines the comprehensive potential of UAVs combined with deep learning models in assessing plant health, which is crucial for maintaining the quality and yield of orchard crops.

Gallo et al. [8] have focused on weed management, evaluating the performance of YOLOv7 in the deep object detection of crop weeds from real-case UAV datasets. Their research is significant for weed control strategies in agriculture, demonstrating how deep learning can assist in distinguishing between crops and weeds effectively.

These studies collectively illustrate the growing trend of employing YOLO algorithms in various agricultural contexts. They range from disease detection to crop health assessment and weed control. The synergy between UAV technology and YOLO models opens new avenues for efficient, accurate, and timely agricultural monitoring, offering promising solutions for sustainable farming practices. As these technologies continue to evolve, they are set to play a pivotal role in the advancement of precision agriculture.

2.2 Forest fires detection

The integration of YOLO models with UAV technology has marked a significant advancement in forest fire detection, as evidenced by several recent studies. These studies
have utilized various versions of YOLO models, each tailored to enhance specific aspects of fire detection.

In the study by Choutri et al. [9], YOLO-based models were employed for fire detection using UAV aerial images. This approach emphasized real-time geo-localization of fires, showcasing the models' ability to accurately identify and locate fire spots from aerial perspectives. This capability is crucial for efficient firefighting and emergency response, as it allows for rapid localization and assessment of fire situations.

Niu et al. [10] developed an improved YOLOv5s-Seg model, focusing on the accurate identification of forest fires through UAV infrared images. This model innovatively combined detection and segmentation, enabling more precise discrimination of fire-affected areas. Such an approach is essential in forest fire management as it aids in understanding the extent and severity of fires, thereby facilitating effective containment strategies.

Chen et al. [11] introduced the LMDFS model based on YOLOv7, specifically designed for detecting forest fire smoke in UAV imagery. The significance of this model lies in its lightweight design, which reduces the demand for computational resources while maintaining high detection accuracy. This feature is particularly beneficial in scenarios where computational power is limited, ensuring that fire detection capabilities are not compromised.

Liu et al. [12] utilized YOLOv5 for forest flame detection in UAV imagery. Their study demonstrated the model's effectiveness in recognizing forest flames, an essential factor for initiating timely firefighting actions. The accuracy and speed of YOLOv5 in detecting flames play a vital role in early fire detection, which is critical in minimizing damage and controlling forest fires.

Bahhar et al.'s study [13] stood out for its use of a staged YOLO model combined with an ensemble CNN for wildfire and smoke detection. This approach highlighted the benefits of integrating different machine learning techniques to enhance detection accuracy. In complex scenarios where smoke and fire coexist, such a combined approach proves to be more reliable and effective in distinguishing between different elements of a fire scene.

These studies collectively underscore the potential of UAVs equipped with YOLO models in forest fire management and emergency response. The versatility and precision of these models in varied environmental conditions and fire scenarios demonstrate the growing importance of UAV and YOLO model integrations in contemporary fire detection and management strategies.

### 2.3 Ecology and marine science

Recent developments in UAV technology and image processing have led to significant strides in environmental monitoring and resource management. The integration of advanced deep learning models, particularly the YOLO architecture, has been pivotal in these advancements.

Mou et al. [14] in their work "WAID: A Large-Scale Dataset for Wildlife Detection with Drones," have made a notable contribution by developing the SE-YOLO model. This model innovatively incorporates a channel self-attention mechanism into YOLOv7, enhancing its capability to detect small targets effectively. This is particularly useful in wildlife monitoring, where the precision of detecting smaller animals from aerial images is crucial. Alongside the model, the study introduces the WAID dataset, a comprehensive collection of wildlife aerial images. This dataset is invaluable for training and validating models for UAV-based wildlife detection, marking a significant step forward in conservation efforts.

Chiang and Juang's study, "Application of UAVs and Image Processing for Riverbank Inspection," [15] showcases the utilization of UAVs coupled with image processing for environmental surveillance, especially in areas that are otherwise inaccessible. Their approach involves the use of UAVs for capturing riverbank images, which are then processed using the scale-invariant feature transform (SIFT) algorithm. This method proves effective
in identifying riverine waste, demonstrating the potential of UAVs in environmental protection and monitoring.

In the realm of maritime surveillance, Yang et al. have made strides with their paper, "A High-Precision Detection Model of Small Objects in Maritime UAV Perspective Based on Improved YOLOv5." [16] Their work centers around enhancing the YOLOv5 model for the high-precision detection of small objects in maritime environments using UAVs. This improved model addresses the critical need for accurate and efficient surveillance in marine settings, highlighting the versatility of the YOLO architecture in various applications.

Hou et al., in their research titled "Detection and Recognition Algorithm of Arbitrary-Oriented Oil Replenishment Target in Remote Sensing Image," [17] focus on the development of an algorithm for detecting and recognizing oil replenishment targets in remote sensing images. This study is vital for efficient resource management and environmental conservation, as it aids in the monitoring and management of oil resources.

Lastly, Idrissi et al.'s work "Evaluating the Forest Ecosystem through a Semi-Autonomous Quadruped Robot and a Hexacopter UAV" [18] explores the use of a semi-autonomous quadruped robot and a hexacopter UAV for forest ecosystem evaluation. This novel approach combines ground and aerial surveillance, enhancing the capabilities of environmental monitoring and data collection.

These studies collectively underscore the growing importance of UAV technology and advanced image processing models like YOLO in various fields, ranging from wildlife conservation to resource management and environmental monitoring. The continued development and integration of these technologies promise further advancements in these critical areas.

2.4 Target detection

In the realm of UAV imagery, recent advancements have been made in developing more efficient, accurate neural network methods for target detection, harnessing the capabilities of YOLO and its variations. These methods are tailored to overcome the unique challenges posed by UAV imagery, such as detecting small or densely arranged objects against complex backgrounds, and ensuring computational efficiency suitable for real-time processing on UAV platforms.

Zhengxin Zhang's "Drone-YOLO" [19] offers an insightful adaptation of the YOLOv8 model, focusing on the 'neck' component to enhance multi-scale UAV image object detection. This adaptation introduces a three-layer PAFPN structure and a detection head optimized for small objects, significantly boosting the capability to pinpoint small-sized targets. A novel aspect of Drone-YOLO is its integration of a sandwich-fusion module that combines network features with low-level features, providing a richer spatial understanding of objects at different layer detection heads. The implementation of RepVGG modules in the network backbone marks another strategic enhancement, fostering the network's ability to learn multi-scale features more effectively.

Parallel to this, Zhao [20] and colleagues' study on a YOLO-ViT-based method focuses on vehicle target detection in UAV infrared imagery. Their approach innovatively merges the YOLO framework with elements of Vision Transformer (ViT), aiming to capture global dependencies. This is particularly pertinent in infrared UAV images, where challenges like variable lighting conditions and low contrast are prevalent. The fusion of YOLO and ViT in this context underscores a strategic move to leverage the strengths of both architectures for improved detection accuracy in the nuanced realm of infrared imagery.

Yang, Zhang, Shang, and Li's contribution [21], a lightweight small target detection algorithm with multi-feature fusion, addresses the critical balance between computational efficiency and detection accuracy. Their algorithm is designed to cater to UAV applications
where real-time processing is essential. By fusing multiple features and employing a lightweight neural network architecture, their solution offers a viable route for real-time applications on UAV platforms, without compromising on the accuracy of detection.

Luo, Wu, and Wang's research [22] on enhancing YOLOv5 for UAV aerial imagery target detection represents another significant stride in this field. Their improvements include the integration of three feature-extraction modules with asymmetric convolutions (ASResNet, AEFE, ASRes2Net), and the incorporation of an Improved Efficient Channel Attention module along with Group Spatial Pyramid Pooling. These enhancements are complemented by the application of the K-Means++ algorithm for more precise anchor box determination and the introduction of a new EIOU-NMS method for improved model postprocessing. This method demonstrates notable improvements in mean average precision (mAP) on various object detection datasets, showcasing its efficacy in addressing the challenges specific to UAV aerial imagery.

Collectively, these studies represent a significant leap forward in UAV-based remote sensing. They not only address specific challenges such as small target detection and complex background analysis but also contribute to the broader narrative of advancing UAV-based imaging and analysis. The continual evolution of these methods points to a future where UAV imagery can be utilized more effectively and efficiently for a wide range of applications, bolstering the capabilities of remote sensing and aerial surveillance.

### 2.5 UAV Navigation

Recent advancements in the field of UAV technology have been marked by significant research focusing on enhanced object detection and tracking capabilities. These studies are crucial in evolving UAV functionalities for a range of applications, from environmental monitoring to search and rescue operations.

Starting with Huangfu and Li's work [23], they have made notable progress in adapting the YOLO v8 algorithm for UAV imagery, specifically targeting small object detection. This lightweight model stands out for its improved efficiency and accuracy in identifying small objects from aerial perspectives. Such enhancement is pivotal in areas where UAVs operate under constrained computational resources, ensuring that even small yet crucial details are not missed.

On a similar trajectory of innovation, Ma, Shen, and Huang [24] have directed their research towards refining UAV visual landing recognition. Integrating YOLO's object detection with onboard edge computing, their approach addresses the critical need for UAVs to autonomously identify suitable landing spots. This capability is especially vital for the safe operation of UAVs in unfamiliar or emergency scenarios. By deploying advanced algorithms on edge computing devices like the NVIDIA JetsonTM Xavier NX, the UAVs are equipped with real-time processing capabilities, a fundamental requirement for accurate and safe landings.

Cao and his team's [25] introduction of GCL-YOLO marks another milestone in UAV technology. This GhostConv-based lightweight YOLO network is designed to enhance the detection of small objects in UAV imagery. The innovative use of GhostConv architecture allows the model to be efficient in processing, which is a crucial consideration for UAVs where battery life and processing power are often limited. The GCL-YOLO represents a significant step forward in making UAV operations more efficient and prolonged, particularly in scenarios where detailed and accurate detection of small objects is essential.

Fang and colleagues [26] have tackled the challenge of UAV operations in adverse weather conditions, particularly fog. Their multi-task learning model is proficient in image dehazing and object detection, which are critical for UAV operations in foggy environments. This research highlights the necessity for robust and adaptive algorithms in UAV technology,
especially for applications like environmental monitoring and search and rescue missions, where weather conditions can be unpredictable and challenging.

Lastly, Hong and colleagues [27] have focused on developing a real-time multi-target tracking algorithm. This deep learning-based approach is tailored for real-time performance and is crucial in scenarios like monitoring extensive areas or conducting search operations. The ability to track multiple targets simultaneously in real-time demonstrates the potential of deep learning in enhancing UAV functionalities for complex and dynamic tasks.

In summary, these studies collectively represent a leap forward in the UAV domain, showcasing how the integration of sophisticated algorithms and edge computing can significantly enhance UAV capabilities in object detection, tracking, and autonomous operations. These advancements are not just technical achievements; they open up new possibilities for UAV applications across various sectors, including environmental monitoring, security, and emergency response, making UAVs more versatile and effective tools in the modern technological landscape.

3 Basic technical approach

The advancements in UAV technology have been significantly influenced by the evolution of the YOLO object detection algorithms. Each version of YOLO offers unique features suitable for UAV applications, emphasizing different aspects of detection efficiency, accuracy, and computational demands.

YOLOv3 and YOLOv3-tiny: These versions marked a significant improvement in detection accuracy, especially for smaller objects, which are common in UAV imagery. The "tiny" variant, designed for devices with limited computational power, is particularly suited for smaller UAVs, offering a balance between detection performance and resource usage.

YOLOv4 and YOLOv4-tiny: YOLOv4 brought enhancements in speed and accuracy, making it ideal for real-time applications in UAVs. Its architecture, CSPDarknet53, improved feature extraction, beneficial for detailed aerial views. The "tiny" version further optimized this balance, providing efficiency for edge computing scenarios in UAVs.

Lightweight YOLOv8: An upgraded version designed for small object identification in UAV images. It caters to the unique challenge of detecting small objects from aerial perspectives, which is crucial for detailed analysis and decision-making in UAV operations.

GCL-YOLO (GhostConv-based YOLO): This variant focuses on reducing the computational load without significantly compromising detection accuracy. By integrating GhostNet's linear transformations, GCL-YOLO reduces parameter count and computational needs, making it suitable for UAVs with stringent payload and power limitations.

Multi-task learning models for foggy conditions: These models, tailored for UAVs, address the challenge of object detection in adverse weather conditions like fog. By simultaneously performing image dehazing and object detection, they enhance the UAV's operational capabilities in less-than-ideal visual conditions.

Each YOLO variant contributes uniquely to UAV technology, catering to the diverse needs of UAV applications ranging from surveillance and mapping to emergency response and agricultural monitoring. The choice of a specific YOLO model depends on the UAV's intended use, computational resources, and environmental conditions it will encounter.

4 Conclusions

The advancements in UAV technologies, particularly through the integration of YOLO object detection algorithms, have significantly transformed various industries, offering innovative and efficient solutions. This fusion of cutting-edge aerial imaging with deep
learning models has revolutionized fields such as agriculture, forest fire management, ecological monitoring, target detection, and UAV navigation.

In agriculture, the application of YOLO models for disease detection and crop monitoring has greatly enhanced precision farming practices. Early detection of plant diseases, accurate counting of crop elements, and efficient weed management using UAVs equipped with YOLO algorithms demonstrate a major step forward in sustainable agriculture.

For forest fire detection, YOLO-based UAV systems have proven to be invaluable. Their ability to accurately identify and localize fire spots in real-time has greatly improved emergency response efforts. The development of models tailored for the detection of specific elements like smoke and flames in varied environmental conditions has further solidified the role of UAVs in forest fire management.

In the realm of ecology and marine science, UAVs with YOLO algorithms have facilitated groundbreaking research in wildlife monitoring, riverbank inspection, maritime surveillance, and environmental conservation. The enhanced capability to detect and analyze small and distant objects in complex backgrounds is pivotal for effective resource management and ecosystem assessment.

The role of UAVs with YOLO in target detection has been particularly noteworthy. Their ability to detect small or densely arranged objects and perform real-time processing on UAV platforms has opened new possibilities for remote sensing and aerial surveillance. The innovations in network architectures like lightweight models and multi-feature fusion algorithms have tailored UAV capabilities for a range of complex detection tasks.

Moreover, the advancements in UAV navigation, including the development of models for visual landing recognition, small object detection, and operation in adverse weather conditions, have greatly expanded the operational scope of UAVs. These technologies not only enhance the UAVs' functional capabilities but also ensure their safe and efficient deployment in various scenarios.

In conclusion, the synergy between UAV technology and YOLO models marks a transformative era in technological advancements. These developments not only address the immediate needs of specific industries but also pave the way for future innovations. As UAV and YOLO technologies continue to evolve, they promise to play an even more integral role in a wide range of applications, making UAVs more versatile, efficient, and indispensable in the technological landscape.

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Grant No.075-15-2022-1121).

References

