

A Review of Autonomous Drone Navigation Using Computer Vision and Artificial Intelligence

Rajole Prashant Baburao¹ and Dr. Surbhit Shukla²

¹Research Scholar, Department of Computer Science

²Research Guide, Department of Computer Science
Vikrant University, Gwalior, M.P.

Abstract: *Autonomous drone navigation using computer vision and artificial intelligence has emerged as an important research area in robotics, automation, and intelligent transportation systems. Unmanned aerial vehicles can now perform navigation-related tasks with reduced human intervention by analysing visual information, recognising objects, estimating motion, and making decisions in dynamic environments. Computer vision techniques enable drones to interpret images and video captured by onboard cameras, while AI algorithms support object detection, obstacle avoidance, path planning, and adaptive decision-making. Deep learning, convolutional neural networks, reinforcement learning, simultaneous localisation and mapping, and visual odometry are among the major approaches used to improve navigation performance. These technologies have applications in agriculture, disaster management, infrastructure inspection, environmental monitoring, and search-and-rescue operations.*

Keywords: Autonomous Drone Navigation, Artificial Intelligence, Computer Vision, Unmanned Aerial Vehicles, Deep Learning, Obstacle Avoidance.

I. INTRODUCTION

Autonomous drone navigation refers to the ability of an unmanned aerial vehicle to determine its position, understand its surroundings, select an appropriate route, and control its movement without continuous human intervention. Conventional drones often depend on remote pilots, GPS signals, or predefined flight paths. However, these approaches may become unreliable in indoor environments, dense forests, urban areas, and locations where satellite signals are weak or unavailable. Computer vision and artificial intelligence offer an alternative by allowing drones to extract useful information from camera images and use it to make navigation decisions.

Computer vision provides the foundation for environmental perception, including the identification of obstacles, landing areas, roads, buildings, and other objects. AI algorithms analyse this information and support decisions concerning flight direction, movement speed, route selection, and collision avoidance. Deep learning has further improved visual recognition and scene understanding by learning complex patterns from large image datasets. Katkuri et al. (2024) reviewed deep learning-based computer vision research for UAV applications and identified sensing and inspection, landing, surveillance and tracking, and search and rescue as important application areas. Their review also highlights the growing importance of vision-based intelligence in autonomous UAV development.

The navigation process generally involves four connected stages: visual perception, localisation, path planning, and flight control. The camera captures environmental information, localisation estimates the drone's position and movement, path planning identifies a suitable route, and the flight controller executes the selected movement. Each stage must operate efficiently because errors in perception or localisation can affect the entire navigation process. Therefore, integrating computer vision with AI is important for improving the flexibility, accuracy, and reliability of autonomous drones.

COMPUTER VISION TECHNIQUES FOR DRONE NAVIGATION

Computer vision enables drones to interpret visual information captured through monocular cameras, stereo cameras, or depth cameras. Image processing and feature extraction help identify important environmental structures, while object detection and semantic segmentation provide information about the location and category of objects. These techniques are useful for recognising trees, vehicles, walls, pedestrians, power lines, and potential landing surfaces. Accurate visual perception is especially important when drones operate in unfamiliar or changing environments.

Object detection models such as convolutional neural networks and You Only Look Once (YOLO) can identify multiple objects in a single image. Their outputs may be used to estimate obstacle locations and support collision avoidance. Semantic segmentation classifies individual image pixels into categories, allowing a drone to distinguish traversable regions from obstacles. Depth estimation provides additional information about the distance between the drone and nearby objects. However, camera-based systems may experience difficulties when lighting changes rapidly, images become blurred, or objects have similar visual characteristics.

Visual odometry estimates the movement of a drone by tracking visual features across consecutive images. Simultaneous localisation and mapping (SLAM) combines localisation with the construction of a map of the surrounding environment. These methods are particularly valuable when GPS is unavailable. The review by Visual SLAM for Unmanned Aerial Vehicles: Localization and Perception (2024) examines challenges associated with real-time operation, textureless environments, and dynamic scenes, while highlighting visual-inertial integration and learning-based enhancements as important research directions.

Computer vision is most effective when image interpretation is integrated with information from inertial measurement units (IMUs), GPS, depth sensors, or LiDAR where appropriate. This combination can reduce the limitations of individual sensors and improve environmental awareness.

ARTIFICIAL INTELLIGENCE AND DEEP LEARNING APPROACHES

Artificial intelligence provides the decision-making capabilities required for autonomous navigation. Machine learning models can classify environmental conditions, recognise obstacles, estimate movement patterns, and select actions based on current observations. Deep learning is particularly useful because it can learn complex visual representations that are difficult to capture through manually designed rules. Convolutional neural networks are widely used for visual feature extraction and object recognition, while recurrent neural networks can model sequential information where temporal context is useful.

Reinforcement learning offers another approach to autonomous navigation. In this method, an agent learns a navigation policy by interacting with an environment and receiving rewards or penalties for its actions. A drone can be rewarded for reaching its target efficiently and penalised for collisions, excessive energy use, or unnecessary movements. Deep reinforcement learning combines reinforcement learning with deep neural networks to handle complex observations and action spaces. AlMahamid and Grolinger (2022) systematically reviewed reinforcement learning approaches for autonomous UAV navigation, discussing navigation problems, algorithm selection, simulation environments, and research challenges.

AI-based navigation can follow either an end-to-end or a modular approach. End-to-end systems learn a mapping from visual observations to flight actions, potentially reducing the need for manually designed intermediate stages. However, such systems can be difficult to interpret and validate. Modular systems separate perception, localisation, planning, and control, making it easier to test individual components and identify failures. Hybrid architectures combine learned perception or decision-making with established geometric planning and safety constraints. For many real-world applications, this hybrid strategy offers a practical balance between adaptability, computational efficiency, and reliability.

APPLICATIONS AND COMPARATIVE ANALYSIS

AI-driven autonomous navigation is relevant to numerous civilian and industrial applications. In precision agriculture, drones can navigate crop fields, monitor plant health, and identify areas requiring intervention. During disaster management, they can explore damaged buildings and locate potential survivors. In infrastructure inspection, drones can follow inspection routes around bridges, wind turbines, communication towers, and power lines. Autonomous navigation also supports warehouse inventory management, environmental monitoring, and delivery operations.

The suitability of a navigation technique depends on the environment, onboard computing resources, sensor configuration, and safety requirements. A method that performs well in a structured indoor environment may not perform equally well in a forest or crowded urban area. Consequently, algorithms should be evaluated using measures such as navigation success rate, collision frequency, localisation error, processing latency, energy consumption, and adaptability to environmental changes.

Table 1. Comparative analysis of computer vision and AI techniques for autonomous drone navigation

Technique	Primary function	Main advantage	Key limitation
CNN-based object detection	Identifies obstacles and objects	Effective visual recognition	Sensitive to image quality and domain changes
YOLO-based detection	Real-time object detection	Fast processing potential	Accuracy depends on model, hardware, and conditions
Semantic segmentation	Classifies pixels by scene category	Detailed environmental understanding	Can require substantial computation
Visual odometry	Estimates drone movement	Supports navigation without continuous GPS	May accumulate drift
Visual SLAM	Estimates position and builds a map	Useful in unfamiliar and GPS-denied environments	Challenging in dynamic or textureless scenes
Reinforcement learning	Learns navigation actions	Adaptable decision-making	Training and real-world transfer can be difficult
Sensor fusion	Combines visual and other sensor data	Improves complementary information use	Requires calibration and reliable integration
Hybrid AI navigation	Combines learning with conventional planning or control	Balances adaptability and structured safety checks	System integration and validation are complex

Source: Author's synthesis of the reviewed literature, including AlMahamid and Grolinger (2022), Katkuri et al. (2024), and the visual SLAM review (2024). The table is a qualitative comparison, not a report of controlled experimental measurements.

The comparison demonstrates that no single technique addresses every navigation requirement. Object detection improves obstacle recognition, SLAM supports localisation and mapping, and reinforcement learning supports action selection. Sensor fusion can strengthen overall performance when individual sensing methods encounter limitations. Therefore, selecting an appropriate combination of methods is often more important than relying on a single algorithm.

CHALLENGES, RESEARCH GAPS, AND FUTURE DIRECTIONS

Despite substantial progress, autonomous drone navigation using computer vision and AI faces several technical challenges. The first is environmental variability. Changes in illumination, shadows, fog, rain, dust, and camera exposure can reduce the reliability of visual perception. Moving obstacles, including pedestrians, vehicles, and other drones, create additional difficulties because the navigation system must predict their movement and update its route continuously. Visual SLAM may also become unreliable in environments with few distinctive features or repetitive visual patterns.

The second challenge concerns computational and energy constraints. Small drones have limited onboard processing capacity, memory, battery life, and payload capability. Large neural networks may provide useful visual representations but can increase processing latency and energy consumption. Delayed obstacle detection or route updates can create safety risks during fast flight. Consequently, lightweight neural networks, model compression, quantisation, and efficient edge computing are important directions for research.

A third challenge is the transfer of AI models from simulation to real-world flight. Models trained in simulated environments may perform poorly when exposed to real camera noise, unexpected obstacles, different weather conditions, or unfamiliar landscapes. AlMahamid and Grolinger (2022) identify important research considerations involving reinforcement learning algorithms, navigation environments, and the selection of suitable approaches for different UAV tasks.

Future research should focus on the following areas:

Multimodal sensor fusion: Combining cameras with IMUs, GPS, depth sensors, or LiDAR to improve localisation and obstacle detection.

Lightweight AI models: Developing compact neural networks that operate efficiently on embedded drone hardware.

Safe reinforcement learning: Incorporating collision constraints, safety monitoring, and emergency recovery strategies into learned navigation policies.

Simulation-to-reality transfer: Improving training methods so that simulated navigation policies generalise to real environments.

Explainable AI: Making navigation decisions more interpretable for testing, auditing, and safety validation.

Collaborative drone navigation: Enabling multiple drones to share maps, observations, and task information while maintaining reliable communication.

A further research gap concerns standardised evaluation. Studies frequently use different datasets, simulation environments, hardware platforms, and performance measures, making direct comparisons difficult. Future investigations should assess navigation success, localisation accuracy, collision rates, energy use, and inference latency under consistent test conditions. They should also evaluate performance across diverse environments rather than relying solely on controlled simulations.

II. CONCLUSION

Autonomous drone navigation using computer vision and artificial intelligence represents an important advancement in intelligent aerial robotics. Computer vision enables drones to perceive objects, identify obstacles, estimate movement, and interpret environmental features, while AI supports navigation decisions, path planning, and adaptive control. Techniques such as deep learning, object detection, visual odometry, SLAM, and reinforcement learning contribute different capabilities to the overall navigation process. Their integration can support applications in agriculture, infrastructure inspection, disaster response, environmental monitoring, and delivery services.

The reviewed literature indicates that the effectiveness of autonomous navigation depends on more than the accuracy of an individual AI model. Reliable operation also requires accurate localisation, timely obstacle detection, efficient onboard computation, robust control, and appropriate responses to uncertainty. Hybrid navigation architectures that combine learning-based perception with established planning and safety mechanisms are a promising approach, although their performance must be verified experimentally. Future work should prioritise sensor fusion, lightweight models, robust sim-to-real transfer, explainable decisions, and standardised real-world evaluation. Overall, autonomous drone navigation is likely to become more capable and practical as these technical challenges are addressed. Nevertheless, safety, reliability, energy efficiency, and adaptability must remain central considerations in the design and deployment of intelligent drone systems.

REFERENCES

1. AlMahamid, F., & Grolinger, K. (2022). Autonomous unmanned aerial vehicle navigation using reinforcement learning: A systematic review. *Engineering Applications of Artificial Intelligence*, 115, Article 105321.
2. Ayala, A., Portela, L., Buarque, F., Fernandes, B. J. T., et al. (2024). UAV control in autonomous object-goal navigation: A systematic literature review. *Artificial Intelligence Review*, 57, Article 125.
3. Katkuri, A. V. R., Madan, H., Khatri, N., Abdul-Qawy, A. S. H., & Patnaik, K. S. (2024). Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review. *Array*, 23, Article 100361.
4. Tang, Y., Zhao, C., Wang, J., Zhang, C., Sun, Q., Zheng, W. X., Du, W., Qian, F., & Kurths, J. (2023). Perception and navigation in autonomous systems in the era of learning: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 34(12), 9604–9624.
5. Visual SLAM for unmanned aerial vehicles: Localization and perception. (2024). *Sensors*, 24(10), Article 2980.