



OPEN ACCESS

EDITED BY
Salil Bharany,
Chitkara University, India

REVIEWED BY
Ali Sarrafi Nik,
Islamic Azad University, Iran
Shahid Khan,
Gdansk University of Technology, Poland

*CORRESPONDENCE
Nasir Saeed
✉ mr.nasir.saeed@ieee.org

RECEIVED 15 January 2025
ACCEPTED 26 February 2025
PUBLISHED 14 March 2025

CITATION
Javaid S, Khan MA, Fahim H, He B and
Saeed N (2025) Explainable AI and monocular
vision for enhanced UAV navigation in smart
cities: prospects and challenges.
Front. Sustain. Cities 7:1561404.
doi: 10.3389/frsc.2025.1561404

COPYRIGHT
© 2025 Javaid, Khan, Fahim, He and Saeed.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Explainable AI and monocular vision for enhanced UAV navigation in smart cities: prospects and challenges

Shumaila Javaid^{1,2,3}, Muhammad Asghar Khan⁴, Hamza Fahim⁵, Bin He^{1,2,3} and Nasir Saeed^{6*}

¹Frontiers Science Center for Intelligent Autonomous Systems, Shanghai, China, ²Department of Control Science and Engineering, College of Electronics and Information Engineering, Tongji University, Shanghai, China, ³National Key Laboratory of Autonomous Intelligent Unmanned Systems, Tongji University, Shanghai, China, ⁴Department of Electrical Engineering, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia, ⁵Department of Computer Science, Faculty of Computer Sciences, Ilma University, Karachi, Pakistan, ⁶Department of Electrical and Communication Engineering, College of Engineering, United Arab Emirates University, Al Ain, United Arab Emirates

Explainable Artificial Intelligence (XAI) is increasingly pivotal in Unmanned Aerial Vehicle (UAV) operations within smart cities, enhancing trust and transparency in AI-driven systems by addressing the 'black-box' limitations of traditional Machine Learning (ML) models. This paper provides a comprehensive overview of the evolution of UAV navigation and control systems, tracing the transition from conventional methods such as GPS and inertial navigation to advanced AI- and ML-driven approaches. It investigates the transformative role of XAI in UAV systems, particularly in safety-critical applications where interpretability is essential. A key focus of this study is the integration of XAI into monocular vision-based navigation frameworks, which, despite their cost-effectiveness and lightweight design, face challenges such as depth perception ambiguities and limited fields of view. Embedding XAI techniques enhances the reliability and interpretability of these systems, providing clearer insights into navigation paths, obstacle detection, and avoidance strategies. This advancement is crucial for UAV adaptability in dynamic urban environments, including infrastructure changes, traffic congestion, and environmental monitoring. Furthermore, this work examines how XAI frameworks foster transparency and trust in UAV decision-making for high-stakes applications such as urban planning and disaster response. It explores critical challenges, including scalability, adaptability to evolving conditions, balancing explainability with performance, and ensuring robustness in adverse environments. Additionally, it highlights the emerging potential of integrating vision models with Large Language Models (LLMs) to further enhance UAV situational awareness and autonomous decision-making. Accordingly, this study provides actionable insights to advance next-generation UAV technologies, ensuring reliability and transparency. The findings underscore XAI's role in bridging existing research gaps and accelerating the deployment of intelligent, explainable UAV systems for future smart cities.

KEYWORDS

consumer electronics, explainable AI, monocular vision, unmanned aerial vehicles, UAV navigation

1 Introduction

Unmanned Aerial Vehicles (UAVs) have evolved from early military reconnaissance tools into essential assets across various industries (Mozaffari et al., 2019; Sharma et al., 2020). Initially developed for surveillance during World War I and later refined with radio-controlled advancements in World War II (Fahlstrom et al., 2022; Sullivan, 2006), UAVs remained largely confined to military applications until the late 20th century (Keane and Carr, 2013). Advances in consumer electronics, including lightweight batteries, high-resolution cameras, and compact processors, alongside regulatory changes, drove their expansion into civilian and commercial domains, enabling applications in photography, disaster response, environmental monitoring, and humanitarian efforts (Xing et al., 2024; Battsengel et al., 2020; Fotouhi et al., 2019; Nex et al., 2022).

Between 2000 and 2010, research focused on improving UAV design, navigation, control systems, and communication technologies, enhancing their ability to execute sophisticated autonomous operations (Hassanalian and Abdelkefi, 2017; Gyagenda et al., 2022; Gu et al., 2020; Ebeid et al., 2018). These advancements facilitated AI-driven navigation, adaptive flight control, and real-time decision-making, allowing UAVs to support critical applications such as disaster relief and climate monitoring (Hernández et al., 2021; Adade et al., 2021). With increasing adoption, modern UAVs now face pressing challenges, including scalability, efficient autonomy, and enhanced situational awareness. Addressing these challenges requires novel approaches such as Explainable AI (XAI) to improve trust, transparency, and decision-making in dynamic environments.

As UAVs have evolved, AI has emerged as a transformative force in navigation and control, enabling unprecedented levels of autonomy and precision. AI-driven systems process large datasets in real-time, allowing UAVs to adapt to dynamic environments and perform complex tasks such as obstacle avoidance, route optimization, and accurate target tracking (Lahmeri et al., 2021; Rovira-Sugranes et al., 2022; Su et al., 2023). These advancements have significantly expanded UAV operational scope, reducing human intervention and improving efficiency in applications like surveillance, environmental monitoring, and disaster response. A key component of AI-enabled navigation is monocular vision, which relies on a single camera to perceive the environment, offering a cost-effective and lightweight alternative to multi-sensor systems. By leveraging AI and Machine Learning (ML) algorithms, monocular vision enables UAVs to estimate depth, detect obstacles, and construct environmental maps in real time, facilitating precise navigation and responsive control even in dynamic and unpredictable environments (Padhy et al., 2018b). This integration allows UAVs to operate efficiently without the reliance on expensive or bulky multi-sensor setups, making monocular vision an essential technology for autonomous UAV operations.

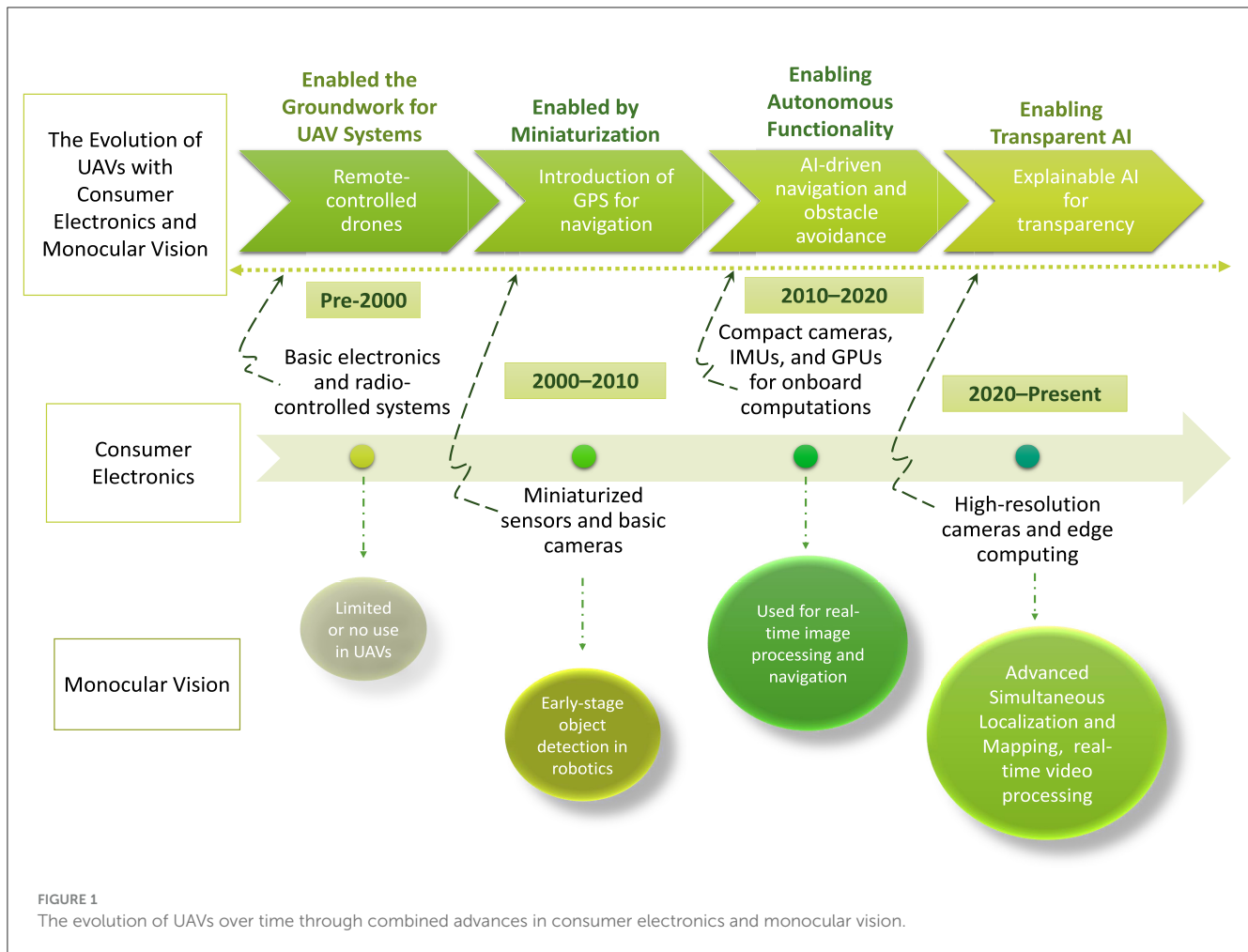
However, as UAVs become increasingly autonomous, the need for transparency and trust in their decision-making processes has brought Explainable Artificial Intelligence (XAI) to the forefront. XAI frameworks address the "black-box" nature of AI algorithms by providing interpretable and human-understandable explanations for UAV navigation decisions (Minh et al., 2022; Došilović et al.,

2018). For instance, XAI techniques can explain how a UAV interprets sensor data, including data from consumer electronics-based components such as high-resolution cameras and lightweight processors, to avoid obstacles or prioritize routes under specific constraints. XAI is also significantly important for enhancing monocular vision by explaining the UAV's interpretations of visual data, such as how depth was inferred or why a specific obstacle avoidance maneuver was chosen. This is especially critical in high-stakes applications such as environmental monitoring and urban planning, where stakeholders require assurance about the reliability and rationale of UAV operations (Kurunathan et al., 2023; Wilson et al., 2021). Therefore, by combining AI-driven automation with XAI interpretability and advances in consumer electronics technologies, UAVs can achieve both operational efficiency and the transparency needed for greater acceptance and trust in their capabilities (Shakhatreh et al., 2019; Sharma et al., 2020). Figure 1 illustrates how the advancements in consumer electronics and monocular vision have collectively driven the evolution of UAVs, enabling innovations such as GPS navigation, autonomous functionality, and XAI.

1.1 Related surveys

State-of-the-art survey papers have significantly advanced UAV navigation efforts by addressing various challenges and proposing innovative solutions. For instance, Lu et al. (2018) reviewed vision-based methods for UAV navigation, focusing on key components like visual localization, mapping, obstacle avoidance, and path planning. Similarly, Rezwan and Choi (2022) provided a comprehensive survey of AI approaches for autonomous UAV navigation, categorizing optimization-based and learning-based techniques while exploring their fundamentals, working principles, and applications. Gyagenda et al. (2022) analyzed independent navigation solutions for GNSS-denied environments, emphasizing perception, localization, and motion planning, whereas Balamurugan et al. (2016) reviewed vision-based navigation techniques, including Visual Odometry (VO) and Simultaneous Localization and Mapping (SLAM), alongside sensor fusion methods for enhanced performance. Alam and Oluoch (2021) examined adaptive navigation techniques for safe flight and landing, while Gu et al. (2020) focused on ANN-based flight controllers for optimization. Ebeid et al. (2018) highlighted the role of open-source platforms in accelerating UAV research and development.

In the context of monocular vision, existing reviews such as Kakaletsis et al. (2021); Yang et al. (2018), and Tong et al. (2023) have explored its applications in UAV navigation. Belmonte et al. (2019) examined vision-based autonomous UAVs for tasks like navigation, control, tracking or guidance, and sense-and-avoid, emphasizing advancements in vision systems, algorithms, UAV platforms, and validation methods. This study highlighted the growing potential of UAVs as personal aerial assistants enabled by computer vision technologies. Similarly, Xin et al. (2022) investigated vision-based autonomous landing for UAVs, showcasing advantages like strong autonomy, low cost, and high anti-interference capability, while categorizing landing scenarios



into static, dynamic, and complex environments. [Bijjahalli et al. \(2020\)](#) emphasized the use of AI-based methods, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and ensemble approaches, to enhance localization, obstacle detection, and data fusion for small UAVs. Collectively, these studies demonstrate the transformative impact of advanced navigation and control strategies, pushing UAV capabilities in increasingly complex environments. [Table 1](#) presents a comparative analysis of existing survey papers and this review. The selected surveys were chosen based on their significant contributions to UAV navigation, obstacle avoidance, path planning, and vision-based methodologies. These works represent foundational studies in AI-driven UAV operations, encompassing techniques such as sensor fusion and monocular vision. However, despite their relevance, none of these surveys comprehensively integrate XAI into their frameworks, highlighting the unique contribution of this review.

1.2 Motivation and contribution

Existing survey papers have primarily focused on UAV navigation technologies, monocular vision, or the challenges of AI integration in isolation. However, the critical issue

of interpretability in AI-driven UAV systems remains largely unaddressed. This review bridges these domains by examining how XAI can enhance trust, reliability, and decision-making in UAV navigation. The study provides a comprehensive analysis of UAV navigation’s evolution—from traditional methods to modern AI and ML approaches while critically assessing the limitations of monocular vision and the transformative potential of XAI. Unlike prior works, it adopts a holistic perspective, emphasizing the synergy between XAI, monocular vision, and consumer electronics. By leveraging advancements such as high-resolution cameras and compact processors, it demonstrates how consumer-grade hardware facilitates the integration of XAI and monocular vision, making UAV navigation more accessible and interpretable. Furthermore, this review identifies key research gaps and offers actionable insights to enhance the reliability and performance of AI-driven UAV systems, particularly in high-stakes applications. By addressing these critical challenges, it represents a significant step toward bridging the gap between cutting-edge innovation and real-world deployment in UAV operations. The key contributions of this work are as follows:

- First, this paper provides a comprehensive overview of the evolution of UAV navigation and control systems, highlighting the transition from traditional methods, such

TABLE 1 A Comparison between existing reviews and presented work.

References	Localization and mapping	Obstacle avoidance	Path planning	Vision-based methods	Monocular vision	AI & ML	Sensor fusion	XAI
Lu et al. (2018)	✓	✓	✓	✓	×	×	×	×
Rezwan and Choi (2022)	×	×	✓	×	×	✓	×	×
Gyagenda et al. (2022)	✓	×	×	×	×	×	×	×
Balamurugan et al. (2016)	✓	✓	×	✓	×	×	✓	×
Alam and Oluoch (2021)	×	×	✓	×	×	×	×	×
Gu et al. (2020)	×	×	×	×	×	✓	×	×
Belmonte et al. (2019)	✓	✓	✓	✓	×	×	×	×
Xin et al. (2022)	×	×	✓	✓	×	×	×	×
Bijahalli et al. (2020)	✓	✓	✓	×	×	✓	✓	×
This work	✓	✓	✓	✓	✓	✓	✓	✓

as GPS and inertial navigation, to advanced AI- and ML-driven approaches. It investigates the role of XAI in UAV systems, addressing the "black-box" nature of traditional ML models. This study explores how XAI enhances autonomous navigation and decision-making in safety-critical applications, emphasizing the integration of technologies, including monocular vision and modern consumer electronics, to expand UAV capabilities across diverse use cases.

- Then, this work underscores the transformative potential of XAI in monocular vision frameworks, which, despite being lightweight and cost-effective, face challenges such as depth perception ambiguities and limited fields of view. This study demonstrates how the incorporation of XAI, significantly improves the reliability and interpretability of monocular vision systems by providing clearer insights into navigation paths, obstacle detection, and avoidance strategies.
- The work then examines how XAI frameworks enhance transparency and trust in UAV decision-making, particularly in high-stakes applications like urban planning and environmental monitoring. It identifies key challenges, including scalability, adaptability to dynamic environments, balancing explainability with performance, ensuring robustness in adverse conditions, and fostering collaboration between vision models and Large Language Models (LLMs) to improve UAV operations. By exploring opportunities for integrating advancements in consumer electronics, monocular vision, and XAI, the paper provides actionable insights that pave the way for future research and the development of next-generation UAV technologies that are both reliable and transparent.

The rest of the paper is organized as follows: Section methodology presents the methodology, Section navig-contr

discusses UAV navigation and control, Section moncol explores monocular vision for UAVs, highlighting challenges and advances, Section sec:XAI focuses on XAI integration for UAV navigation and control, Section challenges addresses the challenges and future research directions, and Section conclusion concludes the paper.

2 Methodology

This study employs a structured methodological approach to investigate the role of XAI in monocular vision-based UAV navigation for smart cities. The methodology includes a systematic literature review and a methodology-driven analysis of research gaps and future directions, focusing on the integration of XAI into UAV systems to enhance interpretability, trust, and overall performance in dynamic urban environments.

2.1 Systematic literature review approach

The hypothesis driving this study asserts that the integration of XAI with monocular vision-based UAV navigation systems substantially improves interpretability, trust, and overall performance, especially within the dynamic and complex environments of smart cities. By addressing the "black-box" nature of AI models, XAI offers greater transparency into the decision-making processes of UAVs. This increased transparency is crucial for improving the safety, reliability, and adaptability of UAV operations in rapidly evolving urban settings. To establish the current state of UAV navigation, monocular vision techniques, and XAI applications, a comprehensive and systematic literature review was conducted. The review process was structured as follows:

- **Selection Criteria:** Articles and survey papers were selected from high-impact journals, conference proceedings, and technical reports to ensure credibility. The selection was based on relevance to UAV navigation, AI-based path planning, monocular vision methodologies, and explainability in AI-driven UAV applications. Only articles with more than 20 citations were initially considered to ensure the inclusion of high-quality, well-established research. Additionally, newer papers with fewer citations but emerging relevance were selected to capture the latest advancements in the field. Papers addressing advances in AI-driven UAV systems and their practical implementations in smart cities were prioritized.
- **Classification of Studies:** The studies were classified into key areas that reflect the core aspects of UAV navigation, monocular vision, and the integration of XAI. The first category, UAV Navigation Evolution, examines the transition from traditional navigation systems such as GPS and Inertial Measurement Units (IMUs) to AI-driven approaches, which leverage ML and data fusion techniques to enhance the autonomy, precision, and adaptability of UAV systems. The second category, Monocular Vision Limitations, addresses challenges inherent in monocular vision systems, such as depth perception issues, occlusions, and environmental sensitivities, which can limit the effectiveness of vision-based navigation in UAVs. Finally, the Role of XAI in UAV Decision-Making explores how various XAI techniques, including *post-hoc* explainability methods and interpretable model architectures, have been applied to improve transparency, trust, and decision-making in UAV operations. This classification emphasizes the evolution of these fields and underscores their interconnectedness in advancing UAV technologies, particularly for applications in smart cities.
- **Comparative Analysis:** A detailed comparative analysis was performed to assess the strengths and limitations of existing studies on UAV navigation, monocular vision, and XAI integration. While much of the existing literature addresses individual aspects such as AI-driven navigation or monocular vision techniques, there is a significant gap in studies that integrate XAI into monocular vision-based UAV systems. This gap highlights the need for further research in this area, particularly to enhance the interpretability and decision-making capabilities of UAV systems in dynamic environments. The current study aims to address this gap by focusing on the synergy between XAI and monocular vision, contributing to more reliable and transparent UAV operations.

2.2 Methodology-driven research gaps and future directions

A critical evaluation of existing UAV navigation studies identified several challenges and research gaps. These include scalability challenges, particularly the integration of XAI in UAV systems that operate at large scales in both urban and rural environments. There is also a need to explore computational efficiency improvements to deploy XAI without compromising UAV battery life and processing capacity. Furthermore, adaptability

to environmental conditions is a key challenge, as weather variations, lighting conditions, and terrain differences impact the performance of XAI-based UAV navigation systems. Techniques for adaptive learning and real-time model adjustments must be examined to enhance performance in dynamic environments. The balancing of explainability with performance remains an ongoing challenge, requiring the development of hybrid approaches where simpler, interpretable models can assist more complex deep-learning models in critical decision-making processes. Ensuring robustness in adverse conditions is another key area of focus, particularly the resilience of XAI-enhanced UAV navigation in challenging environments, such as disaster zones and low-visibility conditions. Redundancy mechanisms should be investigated to ensure system stability in unpredictable scenarios.

Looking toward the future, synergy between XAI, vision models, and LLMs presents promising opportunities. This study explores the potential of multimodal AI models for UAVs, where LLMs are combined with computer vision models to improve UAV situational awareness and autonomous decision-making. The integration of LLMs with computer vision in UAV-based perception systems is examined through real-world applications. Human-UAV interaction via natural language explanations is also explored, leveraging LLMs to provide human-readable explanations of UAV decision-making processes, thereby enhancing human trust in autonomous UAV operations. Finally, enhancing UAV decision-making with LLMs involves evaluating how LLMs can be integrated into UAV mission planning, coordination, and adaptive response strategies, alongside examining novel architectures that blend vision, language understanding, and real-time UAV decision logic.

Based on these findings, several actionable insights and recommendations for advancing XAI-integrated UAV technologies are provided. These include guidelines for the effective deployment of XAI in UAV applications, strategies for selecting the most appropriate XAI models based on operational requirements, and the development of hybrid AI models that optimize the trade-offs between transparency and efficiency. The study also proposes the use of lightweight explainable models suited for real-time UAV applications and explores the feasibility of leveraging consumer-grade hardware, such as compact processors and high-resolution cameras, combined with edge AI techniques, to support cost-effective UAV navigation solutions. These insights inform the subsequent sections, which focus deeper into the application of these recommendations and explore real-world use cases.

3 UAV's navigation and control

Navigation and control have been primary areas of research since the inception of UAVs, undergoing significant evolution over the years. The focus has gradually shifted from basic navigation techniques to advanced systems that enable UAVs to operate autonomously in complex environments. Rapid advancements in consumer electronics have further bridged the gap between high-end military UAVs and affordable civilian models by introducing compact, high-performance components originally designed for consumer markets. This section provides a detailed discussion of the evolution of UAV navigation systems, the advancements in consumer electronics for UAV applications, the integration of AI

and ML technologies, and the challenges in achieving robust UAV navigation and control.

3.1 Evolution and current state of UAV navigation systems

Initially, the navigation and control techniques focused on remotely piloted or semi-autonomous systems that attracted a growing interest in the sophisticated control mechanisms used in UAVs today, such as Ground Positioning System (GPS) guidance (Javaid et al., 2023b; Kwak and Sung, 2018), Inertial Navigation Systems (INS) (Zhang and Hsu, 2018), and visual navigation through cameras and computer vision (Elkaim et al., 2015). Over time, advancements in consumer electronics introduced compact, high-resolution cameras and lightweight processors, enabling more accessible and efficient visual navigation systems. These developments, alongside more sophisticated control mechanisms, led to the creation of algorithms such as Proportional-Integral-Derivative (PID), Linear Quadratic Regulator (LQR), Model Predictive Control (MPC), adaptive control, robust control, sliding mode control, and fuzzy logic controllers (Kim et al., 2006). These control techniques aim to enhance UAV stability, maneuverability, and the ability to respond to changing environmental conditions. With the rapid growth in UAV applications, new control strategies have been introduced to address the increasing complexity of tasks and environments. State-of-the-art schemes (Balamurugan et al., 2016; Gu et al., 2020; Ebeid et al., 2018) have also focused on addressing challenges in complex and GPS-denied environments through autonomous solutions that enhance perception, localization, and motion planning. Vision-based techniques have also been developed to ensure safe flight in dynamic conditions. Neural network-based controllers optimize performance, and open-source platforms have accelerated innovation, making advanced tools more accessible. These advancements have made UAVs highly versatile and capable of operating autonomously in a wide range of situations, from agriculture and logistics to disaster response.

3.2 AI and ML-based navigation and challenges

The integration of AI and ML has transformed UAV navigation systems, enabling unprecedented autonomy and precision in operational performance. Techniques such as Deep Reinforcement Learning (DRL) allow UAVs to optimize navigation strategies by learning from dynamic environments through trial-and-error frameworks (Wang C. et al., 2019; Wang et al., 2017). Similarly, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are employed for real-time feature extraction and sequence analysis in tasks such as obstacle detection and trajectory prediction (Maravall et al., 2015; Padhy et al., 2018a). Advanced sensor fusion techniques integrate data from multiple sources, including monocular vision, LiDAR, and Radar, to provide UAVs with comprehensive situational awareness (Alam and Oluoch, 2021). These AI-driven methodologies allows UAVs

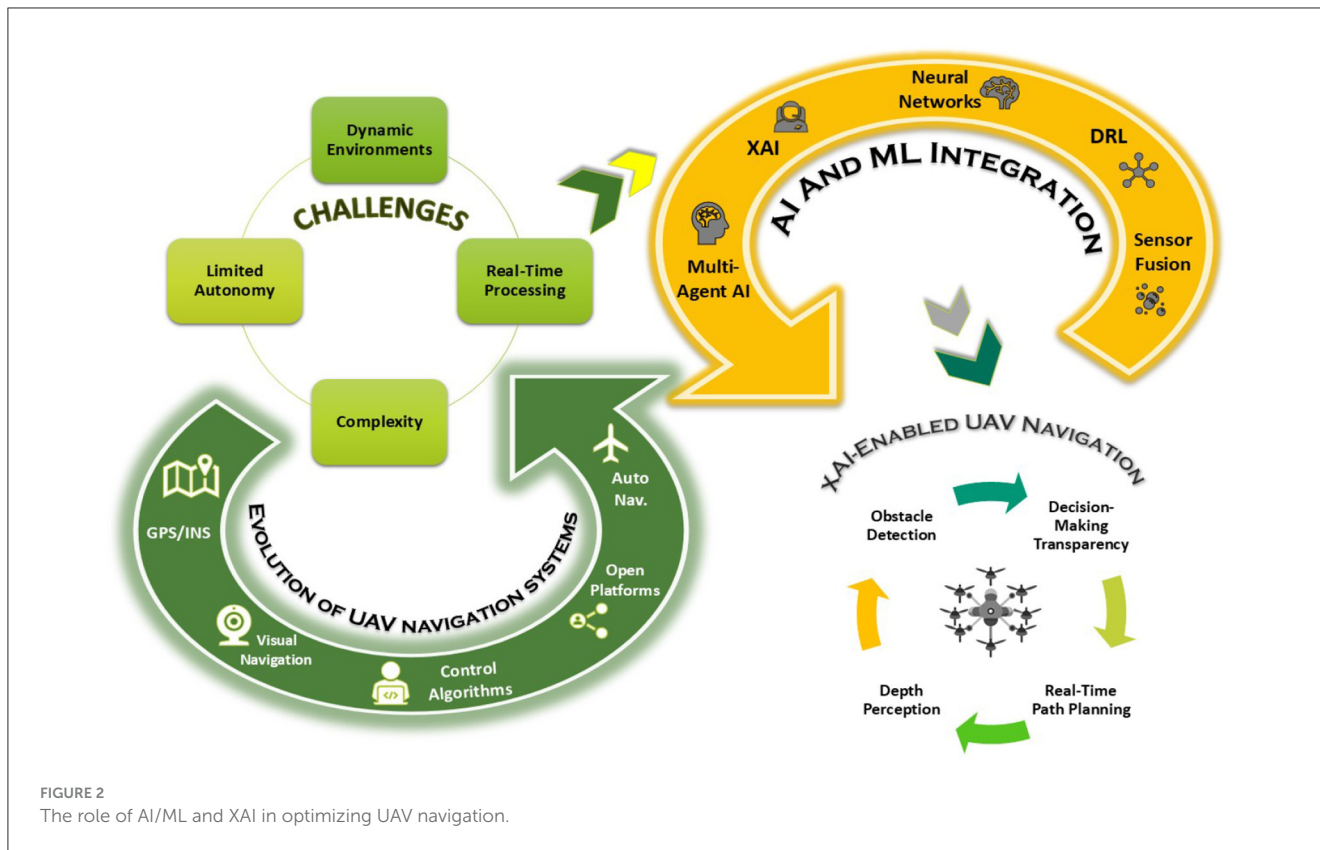
to execute complex tasks, such as multi-agent coordination, real-time path planning, and adaptive flight control, in challenging and GPS-denied environments.

Despite these advancements, the integration of AI and ML in UAVs poses several challenges. Computational limitations remain a significant barrier, as real-time AI processing often exceeds the onboard capabilities of lightweight UAV platforms, even with optimized processors (Hashesh et al., 2022). Furthermore, the generalizability of AI models is hindered by the variability in operational environments, such as diverse terrains, dynamic obstacles, and unpredictable weather conditions. The reliance on training data introduces risks of bias and underperformance in novel scenarios, requiring robust domain adaptation techniques (Puente-Castro et al., 2022; Kurunathan et al., 2023). Additionally, the black-box nature of AI algorithms raises concerns about interpretability and trust, particularly in high-stakes applications. Integrating diverse sensors, including consumer electronics-based monocular vision and advanced sensing technologies such as LiDAR, introduces synchronization and computational challenges, further complicating real-time data fusion and decision-making. Addressing these challenges requires leveraging advancements in consumer electronics, which have introduced innovative hardware solutions to enhance UAV navigation and control capabilities. These contributions are explored in detail in the following section (Sun et al., 2024; McEnroe et al., 2022). Figure 2 further illustrates the progression from traditional UAV navigation to AI/ML integration, emphasizing XAI's role in enhancing capabilities such as obstacle detection.

3.3 Consumer electronics in UAV navigation

Advancements in consumer electronics have profoundly transformed UAV systems, driving remarkable improvements in navigation accuracy, operational efficiency, and system versatility. High-resolution cameras, lightweight and powerful processors optimized for edge computing, and energy-efficient batteries have enabled smaller UAVs to perform computationally intensive tasks under stringent resource constraints (Bebortta et al., 2024). Previously, UAVs depended on bulky multi-sensor arrays and centralized processing units, which limited their agility and deployment flexibility. The shift to monocular vision systems, utilizing compact and consumer-grade cameras, has revolutionized depth estimation and environmental mapping by leveraging cutting-edge advancements in image sensor technology and optical design. These systems efficiently capture high-fidelity, real-time visual data crucial for tasks such as obstacle detection, terrain mapping, and target tracking, offering a cost-effective alternative to traditional setups like stereo vision or LiDAR.

Furthermore, the integration of consumer-grade Graphics Processing Units (GPUs) and Tensor Processing Units (TPUs) has significantly enhanced UAV computational capabilities by enabling the deployment of on-device AI models. Modern GPUs deliver accelerated processing power tailored for ML inference, reducing latency in decision-making processes. This advancement facilitates real-time navigation and rapid response maneuvers, critical for



operations in dynamic and unpredictable environments (Ahmed and Jenihhin, 2022). Edge computing frameworks, enabled by these processors, allow UAVs to process data locally, reducing reliance on ground-based control systems and increasing the reliability of autonomous operations. Energy-efficient battery technologies, such as Lithium-Polymer (LiPo) batteries, developed through consumer electronics innovations, have further bolstered UAV performance by extending flight durations and increasing payload capacities. These batteries, characterized by high energy density and improved discharge rates, support longer missions and enable UAVs to carry additional sensors or communication equipment without sacrificing efficiency. Complementary advancements in power management systems and lightweight electronic components have also contributed to UAV miniaturization, enabling the development of agile and versatile designs suitable for a wide range of applications (Sarkar, 2024; Mahmood et al., 2023).

The collaboration between consumer electronics and UAV technology has additionally fostered the creation of modular and customizable UAV systems. Wireless communication technologies, including Wi-Fi and 4G/5G modules, enhance UAV capabilities by enabling seamless data transmission and real-time remote control, essential for applications that require continuous monitoring and rapid intervention (Ullah et al., 2020; Huang et al., 2024). These synergies have propelled UAV systems into a new era of operational excellence, making them indispensable tools across various domains. Notably, monocular vision systems, driven by compact and consumer-grade cameras, exemplify

how advancements in consumer electronics have reshaped UAV navigation. These systems have become a transformative solution for depth estimation and environmental mapping, offering cost-effective and efficient alternatives to traditional approaches, as explored in the next section (Yang et al., 2018; Padhy et al., 2019).

4 Monocular vision for UAVs: challenges and advances

Monocular vision systems have emerged as a pivotal innovation in UAV navigation, offering lightweight and cost-effective solutions for visual sensing. This section provides the overview, examines the key challenges, and discusses the existing methodologies that can enhance their performance and reliability in diverse applications.

4.1 Overview and challenges

Monocular vision enables UAVs to perceive depth and three-dimensional objects using a single camera, unlike binocular vision, which relies on two cameras for stereopsis. While monocular vision offers a cost-effective and lightweight alternative to stereoscopic systems, its primary drawback is limited depth perception, making accurate distance estimation inherently challenging. This limitation directly impacts essential UAV

functions such as obstacle detection, navigation, and scene understanding (Miclea and Nedevschi, 2021; Madhuanand et al., 2021), increasing the risk of navigation errors and collisions.

Despite these challenges, monocular vision remains fundamental to UAV perception, thanks to recent technological advancements. High-resolution cameras, lightweight processors, and efficient image-processing algorithms have enhanced UAVs' ability to estimate distances, detect objects, and construct real-time environmental maps. These innovations provide a streamlined alternative to multi-sensor setups, reducing the size, weight, and complexity of onboard equipment while maintaining robust performance. However, while these advancements improve reliability, they do not eliminate the fundamental limitations of monocular vision, particularly in complex and dynamic environments.

In addition to depth estimation issues, monocular vision systems face several challenges that hinder UAV navigation and control (Lin et al., 2017; Choi and Kim, 2014; Zhao et al., 2019b). Size estimation remains problematic, as a single camera struggles to distinguish objects at varying distances, leading to ambiguities in detection (Kendall et al., 2014). A limited field of view further restricts situational awareness, particularly in complex urban environments where UAVs must track multiple moving objects such as vehicles and pedestrians. The constrained visual range reduces the UAV's ability to anticipate obstacles, often requiring frequent reorientation to compensate for blind spots. This leads to increased computational demands, energy consumption, and potential delays in real-time decision-making. In dense areas such as intersections or pedestrian zones, UAVs may fail to detect fast-moving objects in time, increasing the risk of navigation errors or inefficient path planning. Additionally, lighting variations, environmental textures, and sudden scene changes can degrade the reliability of object detection and tracking (Al-Kaff et al., 2016). The high mobility of UAVs introduces motion blur, reducing visual data quality and complicating feature extraction. Occlusions, where objects block one another within the camera's field of view, further hinder effective tracking (Lin et al., 2017; Bharati et al., 2018). Adverse weather conditions such as rain, snow, fog, and low light significantly affect monocular vision performance, limiting UAVs' ability to operate reliably. Ensuring precise camera calibration and alignment is also critical, as misalignment can cause navigation errors (Mori and Scherer, 2013).

To mitigate these limitations, ongoing research explores AI-driven image processing, sensor fusion, and adaptive enhancement techniques. Deep learning models refine depth estimation from single-camera inputs, improving object recognition and spatial awareness. Integrating monocular vision with other sensing modalities, such as LiDAR and IMUs, enhances depth perception and environmental mapping. Self-powered sensors (Javaid et al., 2023a) further improve UAV perception by enabling continuous environmental monitoring while reducing reliance on external power sources, contributing to more efficient and sustainable navigation. These advancements continue to refine monocular vision, ensuring it remains a viable solution for autonomous UAV operations.

4.2 Current state of monocular vision

State-of-the-art literature aims to address the aforementioned challenges and enhance monocular vision by developing robust algorithms and techniques (Qin and Shen, 2017). Improved depth estimation algorithms, ML approaches, and sensor fusion techniques enhance UAVs' understanding of their surroundings, enabling more accurate obstacle avoidance, scene reconstruction, and overall autonomous capabilities. For instance, Ross et al. (2013) introduced a novel method for achieving high-speed, autonomous flight of micro aerial vehicles in densely forested environments using monocular vision. The proposed algorithm learns from human expertise that allows micro aerial vehicles to predict and avoid collisions with trees and foliage using affordable visual sensors. In another work (Schmuck and Chli, 2017), the authors introduced a novel centralized architecture for collaborative monocular SLAM using multiple small UAVs. Each UAV independently explores the environment, performing limited-memory SLAM, and sends its mapping data to a central server with greater computational resources, leading to enhanced mapping accuracy and efficiency. In another work (Kendall et al., 2014), the authors developed a cost-effective, self-contained quadcopter controller using monocular vision for object tracking without GPS or external sensors. The system employs a low-frequency algorithm and parallel PID controllers to track colored objects, enabling reliable and efficient UAV applications. Minaeian et al. (2018) introduced an effective, efficient, and robust method for the accurate detection and segmentation of multiple independently moving foreground targets within video sequences captured by UAVs. The presented scheme efficiently estimates camera motion by tracking background key points and employs a combination of local motion history and spatiotemporal differencing over a sliding window to robustly detect multiple moving targets. In another work (Zhao et al., 2019b), the authors proposed a framework combining a monocular camera, GPS, and inertial sensors to detect, track, and geolocate moving vehicles. Using correlation filters for visual tracking and a passive geolocation method, the system calculates vehicle GPS coordinates and guides the UAV to follow the target based on image processing results.

Zheng et al. (2016) developed a robust monocular vision-based control system for quadrotor UAVs, utilizing an onboard monocular camera and IMU sensors to enable precise navigation and positioning. Their approach defines image features on a virtual image plane, simplifying control and enhancing the UAV's ability to regulate its position, making it suitable for applications such as monitoring, landing, and other scenarios requiring accurate and efficient positioning. Similarly, Fu et al. (2017) introduced an improved pose estimation method for multirotor UAVs using an off-board camera. The scheme employed a novel nonlinear constant-velocity process model tailored to multirotor UAV characteristics, enhancing observability and robustness against noise and occlusion for improved vision-based pose estimation.

For collision avoidance, Saha et al. (2014) developed a real-time obstacle detection and avoidance system for low-cost UAVs operating in unstructured, GPS-denied environments. Their mathematical model estimates the relative distance from the UAV's camera to obstacles, enabling effective collision avoidance

while ensuring safety and applicability in complex scenarios. In another study, Al-Kaff et al. (2016) proposed a real-time obstacle detection and avoidance method for micro and small UAVs with limited onboard sensors due to payload and battery constraints. This method mimics human-like behavior by tracking changes in obstacle size within the image sequence, identifying feature points, and extracting obstacles with a high likelihood of approaching the UAV. By comparing area ratio changes in detected obstacles and estimating their 2D positions in the image, the UAV determines the presence of obstacles and takes appropriate avoidance actions. Additionally, Wu et al. (2018) combined obstacle detection with waypoint tracking to further refine avoidance maneuvers, enhancing the overall safety and operability of lightweight UAVs. These advancements in monocular vision have significantly improved UAV capabilities in autonomous navigation, optimized localization, pose estimation, and collision avoidance. However, as UAVs become increasingly autonomous, the need for transparency and trust in their decision-making processes has brought XAI to the forefront. The following section explores the role of XAI in enhancing UAV navigation and control, offering insights into how interpretable AI frameworks can address the “black-box” nature of traditional algorithms and improve operational reliability. Table 2 further summarizes the main contributions of these schemes.

5 Explainable Artificial Intelligence (XAI) integration for UAV navigation and control

This section examines how XAI enhances UAV navigation and monocular vision systems by improving the interpretability and reliability of navigation decisions. It also highlights the role of XAI in seamlessly integrating with monocular vision techniques to optimize depth perception, obstacle detection, and overall efficiency, enabling transparent and effective UAV operations in complex environments.

5.1 Role of XAI in UAV navigation

XAI plays a pivotal role in enhancing the navigation capabilities of UAVs by providing transparency, interpretability, and trustworthiness to the autonomous decision-making processes. Traditional AI models, particularly deep learning algorithms, often function as “black boxes”, making it challenging to understand why UAVs make certain navigation decisions. XAI addresses this limitation by clarifying the underlying mechanisms and factors influencing UAV navigation, ensuring more reliable and accountable operations.

One of the primary contributions of XAI in UAV navigation is improving obstacle detection and avoidance. Monocular vision systems on UAVs capture real-time visual data, which XAI-enhanced algorithms analyze to identify potential obstacles and determine safe flight paths. State-of-the-art techniques such as saliency maps, Local Interpretable Model-agnostic Explanations (LIME), and SHAP highlight the most critical regions in the

TABLE 2 Summary of Key Contributions and Limitations in UAV monocular vision.

References	Key contributions	Limitation of the work
Ross et al. (2013)	Learning-based collision avoidance, Human-demonstration input	Limited adaptability to varying lighting/weather
Schmuck and Chli (2017)	Centralized SLAM, Multi-UAV mapping	Scalability challenges with large UAV fleets
Kendall et al. (2014)	Monocular tracking algorithm, PID-based control	Reduced robustness under severe motion blur
Minaeian et al. (2018)	Multi-Target Detection and Segmentation	Performance degradation in low-light environments
Zhao et al. (2019b)	Vehicle Detection and Geolocation	Limited testing across varied terrains/objects
Zheng et al. (2016)	Virtual image-plane features, IMU integration	Less tested under rapidly changing conditions
Fu et al. (2017)	Nonlinear CV process model, Noise/occlusion resilience	Increased computational demands in complex scenarios
Saha et al. (2014)	Real-time distance estimation, Low-cost solution, GPS-denied environment	Limited robustness at higher speeds and complex obstacle types
Al-Kaff et al. (2016)	Human-inspired detection logic, Minimal sensors, Area-based approach	Less tested under diverse obstacle shapes and severe motion blur
Wu et al. (2018)	Integration with path planning, Dynamic obstacle handling, Real-time navigation adjustments	Reduced evaluation under extreme weather or lighting conditions

visual input that influence navigation decisions, making the decision-making process more interpretable. By providing visual justifications for avoidance maneuvers, XAI allows developers and operators to verify and refine navigation models, ultimately improving UAV safety and adaptability in complex environments (Çiçek et al., 2021; Gundermann, 2020).

Moreover, XAI enhances the adaptability of UAVs in diverse operational scenarios. In various applications, including search and rescue, environmental monitoring, and urban traffic management, UAVs must dynamically adjust their navigation strategies based on changing conditions and mission objectives. XAI provides valuable insights into how environmental factors such as weather variations, moving obstacles, and terrain shifts influence UAV navigation. Understanding these influences allows AI models to be fine-tuned, thereby increasing the robustness and flexibility of UAV navigation systems across different operational landscapes.

As XAI becomes more embedded in UAV navigation, it is critical to establish standardized evaluation metrics to assess its effectiveness and reliability consistently. Current evaluation methods often prioritize accuracy and performance but lack

metrics for explainability and user trust. Future research should focus on developing benchmark datasets and evaluation protocols that incorporate interpretability scores, explanation fidelity, uncertainty quantification, and real-time obstacle avoidance accuracy. Establishing universally accepted evaluation frameworks will allow for a comparative analysis of XAI-driven navigation models, ultimately ensuring the scalability and robustness of these systems in real-world UAV applications.

In addition, trust and accountability are significantly strengthened through the integration of XAI in UAV navigation. Operators, regulatory bodies, and end-users are more likely to trust autonomous UAV systems when they can understand and verify the decision-making processes. XAI fosters this trust by providing clear explanations for route selection, speed adjustments, and obstacle avoidance maneuvers, which can be audited and validated against regulatory safety standards and operational protocols. This transparency is essential for regulatory compliance and for ensuring the broader acceptance of autonomous UAV technologies in industries such as logistics, surveillance, and disaster management (He et al., 2021; Zhu et al., 2024).

Furthermore, XAI supports human-in-the-loop decision-making, where human operators collaborate with UAVs to oversee navigation in complex or high-risk environments. In scenarios that require real-time intervention, such as search-and-rescue missions or disaster response, XAI-powered UAVs can explain their navigational choices, allowing human operators to assess the reasoning behind a maneuver and intervene if necessary. For example, if a UAV initiates an emergency rerouting due to an unexpected obstacle, XAI can clarify whether this action was based on obstacle proximity, terrain constraints, or learned behavioral patterns. This transparent and interpretable decision-making process ensures that UAVs operate with both autonomy and human oversight, reinforcing trust and reliability in critical missions.

5.2 Integration of XAI with monocular vision

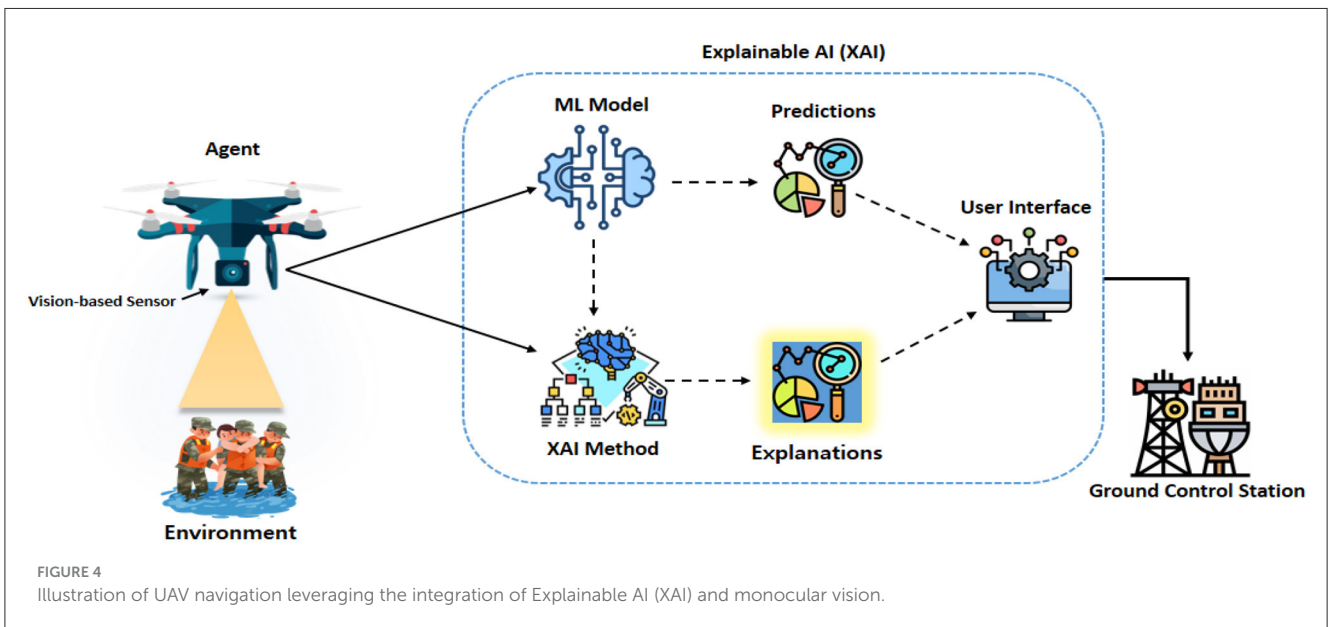
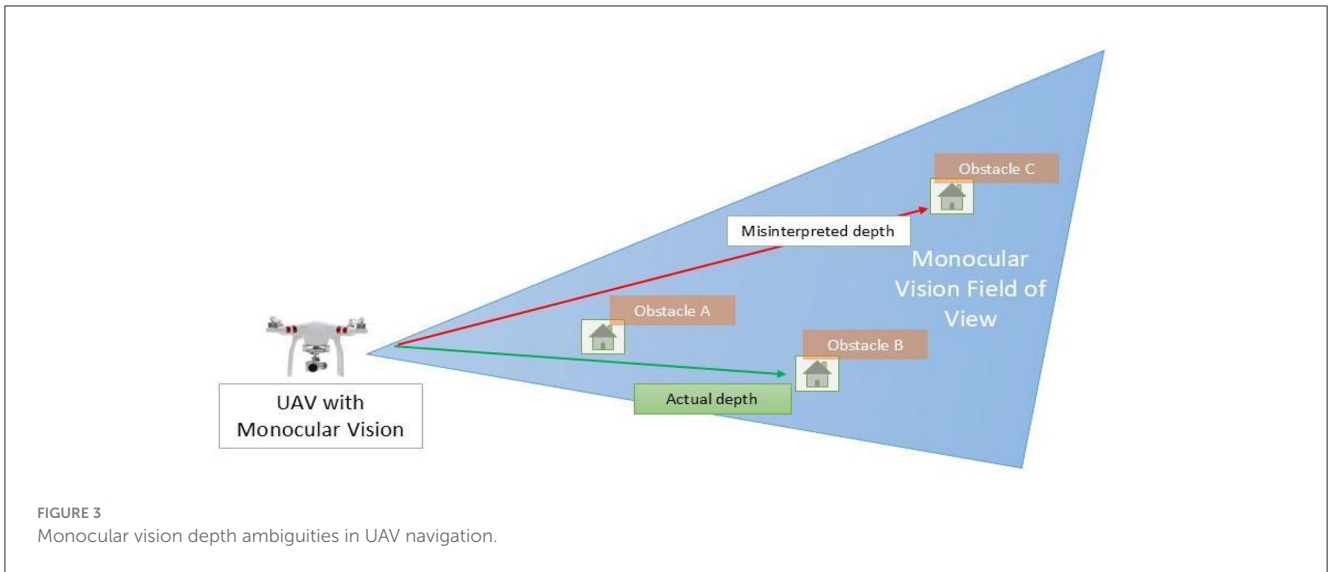
The integration of XAI with monocular vision systems significantly enhances the autonomy, reliability, and transparency of UAV operations. XAI provides interpretable insights into the visual perception and decision-making processes, fostering trust among users and stakeholders while facilitating effective human-AI collaboration. This, in turn, strengthens the adaptability and robustness of UAVs in dynamic and unpredictable environments. Despite their cost and weight advantages, monocular vision systems face inherent limitations such as depth perception ambiguities and sensitivity to lighting variations. Figure 3 illustrates the challenges of depth perception in monocular vision-based UAV navigation. The UAV perceives obstacles within its field of view, but due to the lack of stereo depth cues, it misinterprets the actual distance of objects. Obstacle C is misjudged as being closer than it actually is (red arrow), while Obstacle B represents the correct depth perception (green arrow). Such depth estimation errors can lead to inaccurate obstacle avoidance and inefficient path planning in UAV operations.

Embedding XAI within these AI-driven monocular vision systems has the potential to improve interpretability and reliability and bridge the gap between autonomous UAV functionality and human oversight, ensuring more transparent and dependable operations (Ghasemieh and Kashef, 2024). Figure 4 illustrates UAV navigation leveraging the integration of XAI and monocular vision. However, to fully utilize the benefits of XAI in monocular vision-based UAV navigation, it is crucial to address inherent challenges such as depth estimation ambiguities and environmental sensitivity. This necessitates the use of AI-driven adaptive techniques, which will be explored in the following section.

5.2.1 AI-driven adaptive techniques for robust vision-based UAV navigation

To mitigate the impact of lighting variations and adverse weather conditions on vision-based UAV navigation, several adaptive techniques have been developed. AI-driven image enhancement methods, such as deep learning-based dehazing, contrast enhancement, and super-resolution techniques, improve visibility in challenging conditions such as fog, rain, and low-light environments (Cai et al., 2016; Ren et al., 2018). Dehazing networks leverage CNNs to restore clear images from foggy environments. Additionally, Contrast-Limited Adaptive Histogram Equalization (CLAHE) enhances image contrast in low-light conditions, improving overall visibility. Super-resolution algorithms, such as the Enhanced Deep Super-Resolution Network (EDSR), reconstruct finer details in degraded images, improving the UAV's perception of obstacles and terrain (Lim et al., 2017). Moreover, the integration of multi-spectral and infrared imaging further enhances situational awareness, allowing UAVs to detect thermal signatures and obstacles beyond the visible spectrum (Maes and Steppe, 2019). Infrared (IR) sensors help UAVs function during nighttime operations, heavy fog, or extreme glare conditions, where traditional RGB cameras fail. AI-based sensor fusion techniques combine thermal and visible-light images to improve real-time detection of objects, pedestrians, and vehicles (Zhao et al., 2019a). Additionally, polarization-based vision correction methods reduce glare from reflective surfaces such as water bodies, wet roads, or glass structures, making UAV navigation more stable in urban and coastal environments (Liu et al., 2021). ML models play a crucial role in weather condition classification and real-time adaptation. AI-powered frameworks, such as deep weather recognition networks, analyze environmental conditions and adjust UAV camera settings dynamically (Tahir et al., 2024). RL-based models further enable UAVs to adapt to varying illumination and weather conditions, ensuring reliable navigation in rapidly changing scenarios (Chronis et al., 2023). These systems improve UAV decision-making by adjusting exposure settings, sensor sensitivity, and image filtering techniques based on real-time weather data.

Moreover, sensor fusion approaches, such as combining monocular vision with LiDAR or radar, provide robust depth estimation and obstacle detection, reducing dependency on lighting conditions (Aung et al., 2024). LiDAR-assisted UAV navigation enhances depth perception in low-visibility environments, while radar-based object detection ensures obstacle avoidance even



in extreme weather such as heavy rain or snow (Tang et al., 2021). By integrating multiple sensing modalities, UAVs can create a more comprehensive environmental model, ensuring greater accuracy in navigation, object tracking, and collision avoidance. These AI-driven advancements significantly enhance UAV autonomy, enabling safer and more efficient navigation in complex, dynamic environments. By leveraging adaptive image enhancement, multi-sensor fusion, and real-time weather-aware AI models, UAVs can maintain reliable operations across diverse terrains and environmental conditions, ensuring greater mission success in applications such as disaster response, surveillance, and environmental monitoring (Adão et al., 2017; Singh et al., 2020).

However, while these sensor fusion approaches enhance depth perception and environmental modeling, they do not inherently improve the interpretability of UAV decision-making. In safety-critical applications, it is equally important to understand

the reasoning behind UAV navigation choices to ensure trust and accountability. This necessitates the integration of XAI techniques, which provide insights into UAV perception and decision-making processes.

5.2.2 XAI-driven interpretability and decision transparency in UAV navigation

One of the primary benefits of integrating XAI with monocular vision is the enhancement of perceptual understanding. Monocular vision systems capture real-time visual data, which DRLs process to perform tasks such as object detection, segmentation, and scene understanding. XAI facilitates the generation of interpretable insights into how visual features are utilized in decision-making processes. Saliency maps and Gradient-weighted Class Activation Mapping (Grad-CAM) can highlight regions of interest within the captured images that significantly influence the UAV's perception

TABLE 3 Comparing traditional approaches with XAI-enhanced strategies for UAV navigation and monocular vision.

Focus area	Aspect	Traditional (Non-XAI) approach and limitations	XAI-enhanced approach	Improvements/outcomes
UAV navigation	Transparency and interpretability	Black-box decisions, no clear rationale	Saliency maps, LIME, SHAP highlight decision factors	Enhanced trust, easier debugging, compliance with safety standards
	Adaptability to environmental changes	Hidden logic hinders handling of varying weather/obstacles	Interpretable insights on how conditions affect decisions	Greater robustness, flexibility in dynamic scenarios
	Human-in-the-loop collaboration	Operators struggle to validate or intervene effectively	Real-time explanations guide operator input	Improved teamwork, timely corrective actions, better mission success
Monocular vision	Perceptual understanding and scene interpretation	Limited insight into which visual features influence object recognition/detection	Grad-CAM, saliency maps reveal critical visual features	More accurate scene understanding, confident verification of object focus
	Obstacle detection and avoidance	Relies on raw outputs, limited understanding of key features	Feature attribution highlights critical regions/objects	Safer navigation, refined avoidance strategies, targeted error analysis
	Computational and real-time performance	Limited/no efficient explanation methods may slow decisions	Lightweight, edge-based XAI for quick interpretability	Real-time explanations without major performance impact

and subsequent actions. This enhanced perceptual understanding enables operators to verify that the UAV correctly identifies and prioritizes relevant objects, such as obstacles or points of interest, thereby improving the accuracy and reliability of autonomous operations (Zablocki et al., 2022).

Improving decision-making transparency is another critical aspect of integrating XAI with monocular vision. Autonomous UAVs must make critical decisions based on visual inputs, such as navigating through complex environments or selecting optimal paths for mission objectives. XAI provides transparency in these decision-making processes by elucidating the rationale behind each choice. For instance, when a UAV opts to alter its flight path to avoid an obstacle, XAI algorithms can generate explanations detailing the factors considered, such as the size, speed, and proximity of the obstacle. This transparency is crucial for debugging AI models, ensuring compliance with safety standards, and fostering trust among users and stakeholders. By making the decision-making process transparent, XAI facilitates a better understanding of UAV behavior, which is essential for both developers and end-users. The integration of XAI also facilitates adaptive learning and robustness in UAV operations. The dynamic nature of UAV missions, especially in environments with unpredictable elements, demands adaptive learning capabilities to maintain robust performance. XAI-enabled monocular vision systems allow UAVs to adapt their learning strategies based on explainable feedback. For example, if a UAV consistently misidentifies a particular type of obstacle under certain lighting conditions, XAI can provide insights into the underlying reasons, such as feature misrepresentation or model bias. This feedback allows developers to fine-tune AI models, enhancing their ability to generalize across diverse scenarios and improving the overall robustness of UAV operations (Hwu et al., 2021).

Furthermore, a real-world implementation of XAI with monocular vision can be seen in the Lifeseeker system, which integrates XAI-driven monocular vision for search and rescue operations. UAVs utilize monocular depth estimation and AI-powered scene segmentation to locate missing individuals in

smoke-covered or remote areas. By leveraging XAI-powered saliency maps, the system highlights temperature variations and movement patterns, allowing human operators to validate and refine UAV decisions, reducing false detections, and improving mission success rates. This highlights a broader advantage of XAI-integrated monocular vision that supports human-in-the-loop systems, which is particularly crucial in real-time, high-stakes applications such as search and rescue or urban surveillance. XAI-enabled explanations provide critical insights that empower operators to make informed decisions and corrective interventions. For instance, when a UAV encounters an ambiguous visual scenario, XAI can articulate the reasoning behind its tentative classifications or actions, enabling the operator to validate, override, or adjust the UAV's behavior accordingly (He et al., 2021). Table 3 further highlights the role of XAI in UAV navigation and monocular vision.

5.3 Use cases

This section highlights the use cases for integrating XAI into UAV systems to showcase how it enhances their performance by improving decision-making transparency, enabling real-time adaptability, and optimizing navigation and operational efficiency in complex environments. Figure 5 illustrates the integration of monocular vision and XAI in key use cases, enabling enhanced predictive analytics for actionable insights.

5.3.1 Environmental monitoring and wildlife conservation

Consumer-grade drones equipped with single-camera systems can be deployed to monitor large and remote natural areas, such as forests, wetlands, and wildlife reserves, providing continuous aerial surveillance without the high costs associated with specialized equipment. Monocular vision enables these drones to capture detailed visual data, which can be used to track animal movements,

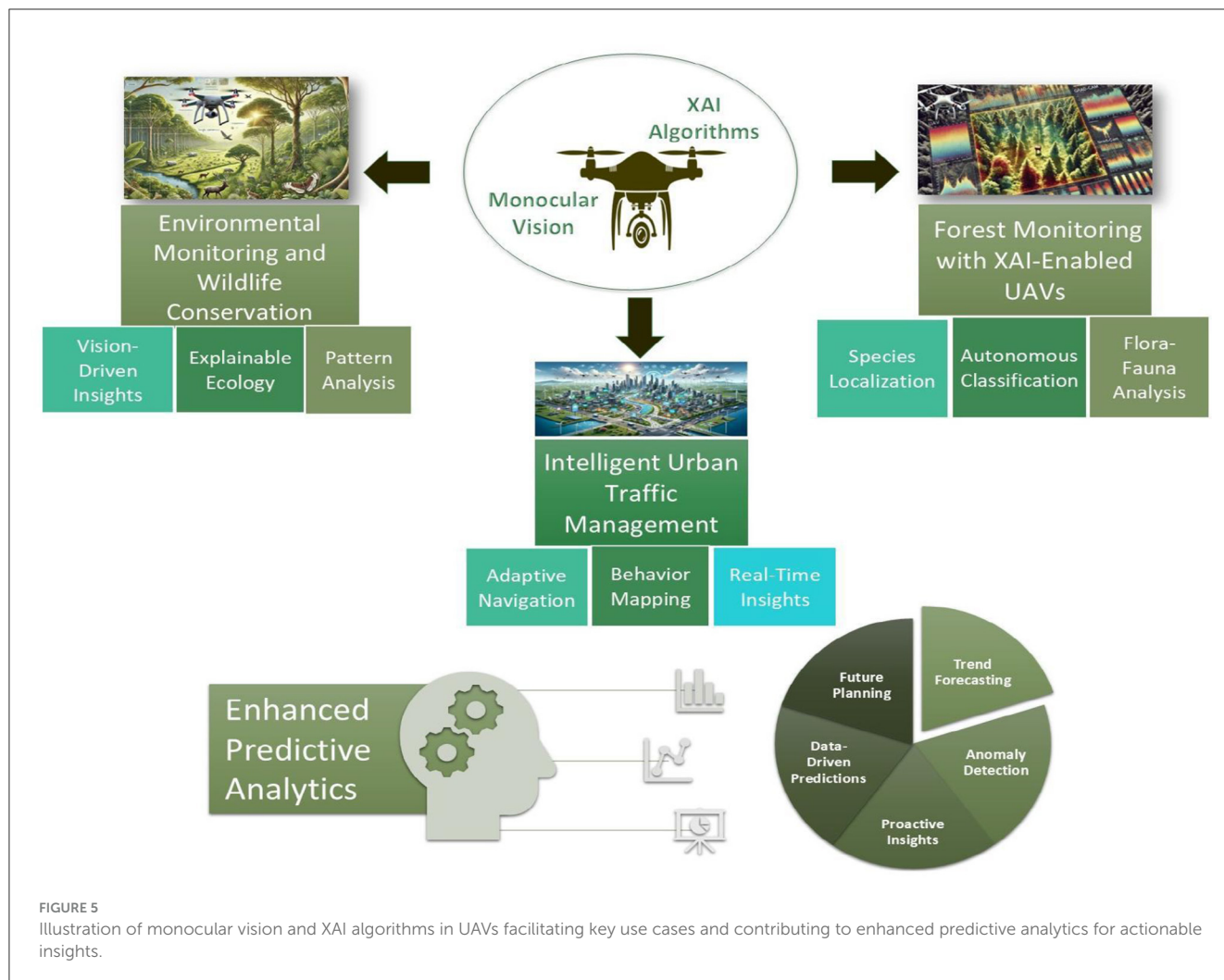


FIGURE 5 Illustration of monocular vision and XAI algorithms in UAVs facilitating key use cases and contributing to enhanced predictive analytics for actionable insights.

detect changes in vegetation, and identify signs of environmental stress or illegal activities like poaching and deforestation. XAI processes this visual information to identify and classify different species or environmental conditions and to provide clear, understandable reasons for its assessments. For example, if the AI detects a decline in a particular plant species, it can highlight the affected areas and explain the potential causes, such as disease signs or adverse weather conditions. This transparency allows conservationists and environmental scientists to trust the AI's findings and make informed decisions based on accurate, actionable insights. Additionally, XAI can assist in predicting future environmental changes by analyzing trends and patterns in the data, offering explanations for its predictions that help stakeholders understand the underlying factors. The affordability and accessibility of consumer electronics make it feasible to deploy these drone systems widely, enabling comprehensive monitoring efforts even in budget-constrained projects. By leveraging monocular vision and XAI, UAVs can play a crucial role in preserving biodiversity, managing natural resources sustainably, and responding swiftly to environmental threats, thereby supporting global conservation initiatives in an efficient and transparent manner.

5.3.2 Intelligent urban traffic management

In densely populated cities, consumer-grade drones equipped with single-camera systems can be deployed to monitor traffic flow, detect congestion, and identify traffic violations in real-time. Monocular vision allows these drones to capture continuous video streams of streets, intersections, and public spaces without the complexity and cost of multi-camera setups. XAI processes the visual data to analyze patterns such as vehicle speeds, pedestrian movements, and the occurrence of traffic incidents. For example, if the AI detects a sudden slowdown or an accident, it can provide clear, understandable explanations by highlighting the specific vehicles involved, the location of the incident, and the factors contributing to the disruption, such as a traffic light malfunction or a jaywalking pedestrian.

This transparency is crucial for city planners and law enforcement agencies, as it builds trust in the system's assessments and decisions. Urban authorities can use these insights to make informed decisions about traffic light timings, road maintenance, and emergency response strategies. Additionally, XAI can assist in optimizing traffic flow by suggesting adjustments based on observed behaviors, such as rerouting traffic to alleviate

congestion hotspots or implementing new pedestrian safety measures. The use of consumer-grade electronics ensures that these drone systems are affordable and scalable, allowing widespread deployment across various urban areas without significant investment. Moreover, the user-friendly nature of consumer drones enables local communities to participate in monitoring and improving their neighborhoods, fostering a collaborative approach to urban management.

5.3.3 Forest monitoring with XAI-enabled UAVs for flora and fauna classification

The integration of XAI into UAV systems significantly enhances their functionality by enabling the autonomous identification and classification of various flora and fauna while providing transparent explanations for these classifications. A fleet of consumer-grade UAVs equipped with high-resolution monocular cameras can be deployed across a sprawling forest ecosystem. These UAVs, integrated with advanced XAI algorithms, process real-time visual data to identify and classify different plant and animal species. The monocular vision system captures continuous video feeds, which are analyzed by DRLs trained on extensive datasets of local flora and fauna. XAI frameworks enable these models to generate interpretable insights into their decision-making processes. Techniques such as Grad-CAM and LIME highlight specific regions within captured images that influenced the UAVs' classifications. For instance, when a UAV identifies a particular tree species, Grad-CAM highlights distinctive features such as leaf patterns or bark textures, while LIME provides textual explanations, such as "The tree was identified as Oak due to its lobed leaf structure and rough bark texture."

The deployment of XAI-enabled UAVs provides substantial benefits for environmental monitoring and research. Researchers can collect comprehensive data on species distribution and abundance across large areas with greater accuracy and efficiency compared to traditional ground surveys. The explainable outputs generated by XAI algorithms facilitate data validation by allowing researchers to cross-reference UAV classifications with field observations. For example, during a survey of the forest canopy, UAVs successfully identified several rare plant species that were previously under-documented. The explanations provided by XAI enable researchers to verify these findings by inspecting highlighted features and understanding the rationale behind each classification. This level of transparency increases trust in UAV-collected data and empowers researchers to refine models by addressing potential misclassifications or biases in the AI algorithms.

Furthermore, the ability of XAI-enabled UAVs to autonomously classify fauna, such as bird species and small mammals, offers valuable insights into ecosystem dynamics and species interactions. Real-time explanations facilitate immediate validation and adaptation of monitoring strategies, improving the overall effectiveness of conservation efforts. For instance, when a UAV detects an unusual concentration of a particular bird species, XAI explanations help researchers understand behavioral patterns or environmental factors contributing to the observation, informing targeted conservation actions.

6 Challenges and future research directions

This section examines key challenges that limit the application of XAI in UAV navigation and monocular vision systems and proposes potential future research directions. It emphasizes critical areas for development, including scalability, interpretability, adaptability to dynamic environments, and robustness in real-world applications.

6.1 Scalability and interpretability in XAI for monocular vision systems

One of the foremost challenges in integrating XAI with monocular vision systems is scalability across different UAV models and environments. UAV platforms vary significantly in sensor configurations, computational capacities, and environmental operating conditions, requiring extensive model fine-tuning for each system. Ensuring that XAI methods can generalize effectively across diverse UAV architectures without excessive retraining is critical for large-scale deployment. Additionally, monocular vision systems generate high-resolution visual data that must be processed in real-time for navigation and decision-making. Balancing computational efficiency with explainability is particularly challenging, as XAI models must provide interpretable insights while maintaining low latency in dynamic and unpredictable environments (Das and Rad, 2020; Ding et al., 2022).

To address these scalability concerns, future research should focus on developing lightweight, adaptive XAI models that can efficiently process high-resolution data across various UAV platforms. Techniques such as meta-learning, transfer learning, and model compression can enhance the generalizability of XAI methods, reducing the need for extensive system-specific retraining. Additionally, edge-based explainability and hybrid cloud-edge frameworks should be explored to optimize computational workloads and enable real-time decision-making. Future studies should also prioritize the development of contextual and hierarchical explainability techniques, breaking down UAV decisions into intuitive, multi-layered insights to improve interpretability across different operational contexts. Establishing benchmark datasets and standardized evaluation protocols tailored for multi-platform XAI scalability will further facilitate the development of widely adoptable, efficient, and robust explainable UAV systems.

6.2 Adapting XAI for dynamic environmental changes

Monocular vision systems in UAVs operate within highly dynamic environments where lighting conditions, terrain variations, and the presence of moving or static obstacles can change rapidly. These fluctuations pose significant challenges for XAI models, which may struggle to maintain the accuracy and relevance of their explanations in real-time. Traditional XAI frameworks often produce static or generalized explanations

that do not capture the nuanced decision-making required in such unpredictable scenarios, thereby reducing user trust and operational reliability (Shah and Konda, 2021; Confalonieri et al., 2021).

To address these challenges, it is essential to develop adaptive XAI frameworks that incorporate real-time environmental context. Integrating DRL with XAI allows UAVs to update their explainability models based on evolving environmental data continuously. This integration ensures that explanations remain accurate and contextually relevant, even in highly dynamic settings. Additionally, context-aware XAI techniques can improve the adaptability of UAVs by utilizing temporal and spatial data to inform explanations. This approach ensures that explanations evolve with changing conditions, offering more precise and situation-specific insights. Implementing transfer learning and domain adaptation strategies further enhances the generalizability of XAI models, enabling UAVs to perform reliably across diverse and unforeseen environments.

6.3 Balancing explainability and performance

One of the critical challenges in building trust in UAV navigation using XAI and monocular vision lies in balancing the trade-off between explainability and system performance. High levels of explainability often require additional computational resources, leading to increased inference time, higher energy consumption, and reduced real-time responsiveness (Silva et al., 2024). Many explainability techniques, such as SHAP, and Grad-CAM, involve complex backpropagation computations and additional layers of processing, significantly increasing computational overhead. This additional processing load can be particularly problematic for UAVs operating in resource-constrained environments, where battery life, processing power, and response time are critical constraints. This trade-off is especially evident in autonomous UAV surveillance and disaster response missions, where UAVs must quickly analyze complex environments, detect anomalies, and make real-time navigational adjustments while simultaneously providing interpretable decisions to human operators. In high-risk applications, delays caused by computationally expensive XAI techniques can negatively impact situational awareness, collision avoidance, and time-sensitive decision-making. Additionally, excessive computational demands can drain battery life more rapidly, limiting the UAV's flight duration and mission effectiveness. As UAVs become increasingly autonomous, addressing the trade-off between explainability and computational efficiency is essential to ensure that these systems remain both transparent and operationally viable.

Future research should focus on developing lightweight XAI models optimized for real-time UAV deployment. Several approaches can help mitigate computational overhead while maintaining interpretability. Techniques such as model compression (e.g., knowledge distillation and weight pruning) reduce the complexity of deep learning models while preserving essential explainability features (Hinton et al., 2015). Edge

computing architectures allow UAVs to offload explainability computations to low-power onboard processors or external edge nodes, reducing the burden on real-time decision-making (Xu et al., 2023). Additionally, selective explainability mechanisms can be implemented, where explanations are provided only for anomalies, high-risk decisions, or human-in-the-loop scenarios, rather than for every prediction (Wang D. et al., 2019). This adaptive XAI approach ensures that UAVs dynamically adjust the level of explainability based on mission requirements, prioritizing transparency when necessary while conserving computational resources in routine operations. Furthermore, hardware-accelerated explainability solutions, such as dedicated GPUs, TPUs, and neuromorphic processors, can significantly enhance computational efficiency by parallelizing explainability computations (Tahir et al., 2024). These specialized hardware components enable UAVs to perform real-time, interpretable decision-making without compromising flight endurance or mission effectiveness. Thus, explainability and computational efficiency remain an open challenge in deploying AI-driven UAVs. Future advancements should focus on energy-aware explainability techniques, mission-adaptive XAI frameworks, and real-time optimization strategies to ensure that UAVs remain both transparent and operationally effective across diverse applications. Addressing this trade-off will be crucial for the next generation of autonomous UAVs, particularly in time-sensitive, resource-constrained environments where both real-time decision-making and interpretability are equally vital.

6.4 Ensuring robustness of XAI in adverse conditions

XAI models integrated with monocular vision systems face difficulties in maintaining reliability under adverse conditions, such as low light, fog, rain, or high-motion scenarios. These conditions can degrade visual data quality, leading to less accurate explanations and undermining user trust in UAV decisions. Future efforts should focus on developing XAI models that incorporate uncertainty quantification to account for degraded input quality. By integrating techniques such as Bayesian neural networks or ensemble learning, XAI systems can provide confidence scores for their explanations, helping users assess the reliability of UAV decisions in adverse conditions. Furthermore, future research should leverage multimodal data fusion by integrating monocular vision with complementary sensors to enhance robustness and provide trustworthy explanations, even in challenging environments (Bai et al., 2021; Ali et al., 2023).

Future research should focus on developing adaptive explainability mechanisms that dynamically adjust explanation complexity based on operational context. In high-visibility environments with reliable sensor data, UAVs should be able to generate detailed, feature-rich explanations, while in low-confidence scenarios such as heavy fog or fast motion, XAI models should be designed to simplify explanations while incorporating uncertainty estimates to prevent misleading interpretations. Additionally, future work should explore the integration of physics-informed AI models to enhance the robustness of explanations by

leveraging physical constraints, such as scene geometry and optical flow, to improve depth perception and motion estimation in degraded visual conditions. Future research should also investigate how context-aware explainability frameworks can be developed to ensure that interpretability dynamically adapts to environmental and operational uncertainties. Furthermore, advancements in self-supervised learning and reinforcement learning-based XAI models should be investigated to enhance UAV autonomy, enabling models to learn optimal explainability strategies from real-world interactions and ensuring a balance between transparency, accuracy, and operational feasibility.

6.5 Collaborative decision-making between vision models and LLMs with XAI

Monocular vision systems, while proficient in object detection and scene analysis, face challenges in contextualizing their outputs in real-world scenarios. Traditional AI models, such as CNNs, excel at recognizing patterns and detecting objects but lack interpretability, providing only binary or probability-based classifications without insight into the reasoning behind their decisions. For example, a CNN-based vision model can detect an obstacle but cannot determine its significance whether it is a temporary inconvenience or a critical mission threat. Similarly, existing XAI methods, though designed to enhance transparency, often struggle to provide intuitive and contextually relevant justifications, particularly in dynamic and high-stakes applications such as UAV navigation (Berwo et al., 2023). Furthermore, most current XAI implementations rely on static explanations, failing to adapt dynamically to real-time environmental changes, reducing user trust, and limiting deployment in critical missions.

Future work should focus on integrating monocular vision systems with contextual reasoning algorithms, such as LLMs, to bridge the gap between perception and contextual awareness (Javaid et al., 2024a). Unlike traditional AI models, LLMs can generate accurate, natural language explanations tailored to specific decision contexts, improving trust and usability. However, LLM-based vision models present significant challenges in handling real-world navigation complexities, particularly in dynamic environments where real-time decision-making is crucial. One of the key limitations is contextual misinterpretation, as LLMs rely on pre-trained datasets and lack direct interaction with real-world scenes. This limitation can lead to misclassification of transient obstacles, such as pedestrians or moving vehicles, which UAVs may treat as permanent barriers, resulting in inefficient route planning or unnecessary evasive maneuvers.

In addition, LLM-vision models struggle with spatiotemporal reasoning, making it difficult for UAVs to anticipate object movement patterns and differentiate between static and dynamic elements in complex environments. Unlike human pilots who instinctively recognize environmental cues and adapt, LLMs lack continuous learning mechanisms that allow them to refine their perception of transient vs. permanent objects in real-time. Future research should explore hybrid approaches,

integrating LLM-based reasoning with sensor fusion techniques such as LiDAR, radar, and event-based cameras to enhance environmental understanding. Furthermore, adaptive and interactive explainability techniques should be developed to ensure that UAVs can dynamically adjust the level of interpretability based on mission constraints and operational uncertainties. Another important research direction is the development of lightweight LLM architectures or hardware-accelerated LLM inference, ensuring that real-time UAV applications maintain low latency and high computational efficiency while benefiting from improved contextual awareness. Addressing these limitations will be critical for deploying LLM-enhanced vision models in UAVs for real-world applications, ensuring both operational safety and mission success (Javaid et al., 2024b).

7 Conclusion

This paper provides a comprehensive exploration of the advancements and challenges in integrating Explainable Artificial Intelligence (XAI) into UAV navigation and monocular vision systems. It outlines the progression of UAV navigation from traditional methods, such as GPS and inertial navigation, to modern AI and ML-driven approaches, highlighting the transformative role of monocular vision and consumer electronics in enhancing UAV capabilities across various applications. This study comprehensively focuses on the role of XAI in addressing the "black-box" nature of AI algorithms, which has long hindered trust and transparency in UAV systems. By providing interpretable insights into navigation and control decisions, XAI fosters greater reliability and usability, particularly in high-stakes scenarios such as urban planning and environmental monitoring. The integration of XAI with monocular vision further demonstrates its potential to overcome limitations in depth perception and obstacle detection, enhancing the overall performance and adaptability of UAVs. By examining the interaction between advancements in consumer electronics, monocular vision, and XAI, this paper offers actionable insights to address these challenges and pave the way for future research and development. This work underscores the critical importance of combining technological innovation with explainability to achieve more reliable, efficient, and scalable UAV operations. The findings serve as a foundation for further exploration into the intersection of XAI, UAV navigation, and monocular vision, promoting advancements in next-generation autonomous UAV systems.

Author contributions

SJ: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft. MK: Investigation, Project administration, Writing – review & editing. HF: Data curation, Formal analysis, Writing – original draft. BH: Supervision, Validation, Writing – review & editing. NS: Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the National Natural Science Foundation of China under Grant No. 62088101.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Ad ao, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., et al. (2017). Hyperspectral imaging: a review on uav-based sensors, data processing and applications for agriculture and forestry. *Remote Sens.* 9:1110. doi: 10.3390/rs9111110
- Adade, R., Aibinu, A. M., Ekumah, B., and Asaana, J. (2021). Unmanned aerial vehicle (UAV) applications in coastal zone management—a review. *Environ. Monit. Assess.* 193, 1–12. doi: 10.1007/s10661-021-08949-8
- Ahmed, F., and Jenihhin, M. (2022). A survey on uav computing platforms: a hardware reliability perspective. *Sensors* 22:6286. doi: 10.3390/s22166286
- Alam, M. S., and Oluoch, J. (2021). A survey of safe landing zone detection techniques for autonomous unmanned aerial vehicles (UAVS). *Expert Syst. Appl.* 179:115091. doi: 10.1016/j.eswa.2021.115091
- Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., et al. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Inform. Fusion* 99:101805. doi: 10.1016/j.inffus.2023.101805
- Al-Kaff, A., Meng, Q., Martín, D., de la Escalera, A., and Armingol, J. M. (2016). “Monocular vision-based obstacle detection/avoidance for unmanned aerial vehicles,” in *2016 IEEE Intelligent Vehicles Symposium (IV)* (Gothenburg: IEEE), 92–97.
- Aung, N. H. H., Sangwongngam, P., Jintamethasawat, R., Shah, S., and Wuttisititkulki, L. (2024). A review of LiDAR-based 3d object detection via deep learning approaches towards robust connected and autonomous vehicles. *IEEE Trans. Intellig. Vehicl.* 2024, 1–23. doi: 10.1109/TIV.2024.3415771
- Bai, X., Wang, X., Liu, X., Liu, Q., Song, J., Sebe, N., et al. (2021). Explainable deep learning for efficient and robust pattern recognition: a survey of recent developments. *Pattern Recognit.* 120:108102. doi: 10.1016/j.patcog.2021.108102
- Balamurugan, G., Valarmathi, J., and Naidu, V. P. S. (2016). “Survey on uav navigation in gps denied environments,” in *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)* (Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE)), 198–204.
- Battsengel, G., Geetha, S., and Jeon, J. (2020). Analysis of technological trends and technological portfolio of unmanned aerial vehicle. *J. Open Innovat.: Technol. Mark. Complex.* 6:48. doi: 10.3390/joitmc6030048
- Bebortta, S., Tripathy, S. S., Khan, S. B., Al Dabel, M. M., Almusharraf, A., and Bashir, A. K. (2024). Tinydeepuav: A tiny deep reinforcement learning framework for uav task offloading in edge-based consumer electronics. *IEEE Trans. Consum. Elect.* 70, 7357–7364. doi: 10.1109/TCE.2024.3445290
- Belmonte, L. M., Morales, R., and Fernández-Caballero, A. (2019). Computer vision in autonomous unmanned aerial vehicles—a systematic mapping study. *Appl. Sci.* 9:3196. doi: 10.3390/app9153196
- Berwo, M. A., Khan, A., Fang, Y., Fahim, H., Javaid, S., Mahmood, J., et al. (2023). Deep learning techniques for vehicle detection and classification from images/videos: a survey. *Sensors* 23:4832. doi: 10.3390/s23104832
- Bharati, S. P., Wu, Y., Sui, Y., Padgett, C., and Wang, G. (2018). Real-time obstacle detection and tracking for sense-and-avoid mechanism in uavs. *IEEE Trans. Intellig. Vehicl.* 3, 185–197. doi: 10.1109/TIV.2018.2804166
- Bijjahalli, S., Sabatini, R., and Gardi, A. (2020). Advances in intelligent and autonomous navigation systems for small uas. *Prog. Aerospace Sci.* 115:100617. doi: 10.1016/j.paerosci.2020.100617
- Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D. (2016). DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. Image Proc.* 25, 5187–5198. doi: 10.1109/TIP.2016.2598681
- Choi, H., and Kim, Y. (2014). Uav guidance using a monocular-vision sensor for aerial target tracking. *Control Eng. Pract.* 22, 10–19. doi: 10.1016/j.conengprac.2013.09.006
- Chronis, C., Anagnostopoulos, G., Politi, E., Dimitrakopoulos, G., and Varlamis, I. (2023). Dynamic navigation in unconstrained environments using reinforcement learning algorithms. *IEEE Access* 11, 117984–118001. doi: 10.1109/ACCESS.2023.3326435
- Çiçek, İ. B., Küçükakçalı, Z., and Yağın, F. H. (2021). Detection of risk factors of pcos patients with local interpretable model-agnostic explanations (lime) method that an explainable artificial intelligence model. *J. Cognitive Syst.* 6, 59–63. doi: 10.52876/jcs.1004847
- Confalonieri, R., Coba, L., Wagner, B., and Besold, T. R. (2021). A historical perspective of explainable artificial intelligence. *Wiley Interdiscipl. Rev.: Data Mining Knowl. Discovery* 11:e1391. doi: 10.1002/widm.1391
- Das, A., and Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv preprint arXiv:2006.11371*.
- Ding, W., Abdel-Basset, M., Hawash, H., and Ali, A. M. (2022). Explainability of artificial intelligence methods, applications and challenges: a comprehensive survey. *Inf. Sci.* 615:238–292. doi: 10.1016/j.ins.2022.10.013
- Došilović, F. K., Brčić, M., and Hlupić, N. (2018). “Explainable artificial intelligence: A survey,” in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (Opatija: IEEE), 0210–0215.
- Ebeid, E., Skriver, M., Terkildsen, K. H., Jensen, K., and Schultz, U. P. (2018). A survey of open-source uav flight controllers and flight simulators. *Microprocess. Microsyst.* 61, 11–20. doi: 10.1016/j.micpro.2018.05.002
- Elkaim, G. H., Lie, F. A. P., and Gebre-Egziabher, D. (2015). “Principles of guidance, navigation, and control of UAVS,” in *Handbook of Unmanned Aerial Vehicles* (Dordrecht: Springer Science+Business Media Dordrecht), 347–380.
- Fahlstrom, P. G., Gleason, T. J., and Sadraey, M. H. (2022). *Introduction to UAV Systems*. Hoboken, NJ: John Wiley & Sons.
- Fotouhi, A., Qiang, H., Ding, M., Hassan, M., Giordano, L. G., Garcia-Rodriguez, A., et al. (2019). Survey on uav cellular communications: practical aspects, standardization advancements, regulation, and security challenges. *IEEE Commun. Surv. Tutor.* 21, 3417–3442. doi: 10.1109/COMST.2019.2906228
- Fu, Q., Quan, Q., and Cai, K.-Y. (2017). Robust pose estimation for multirotor uavs using off-board monocular vision. *IEEE Trans. Indust. Electron.* 64, 7942–7951. doi: 10.1109/TIE.2017.2696482
- Ghasemieh, A., and Kashef, R. (2024). Towards explainable artificial intelligence in deep vision-based odometry. *Comp. Electr. Eng.* 115:109127. doi: 10.1016/j.compeleceng.2024.109127
- Gu, W., Valavanis, K. P., Rutherford, M. J., and Rizzo, A. (2020). Uav model-based flight control with artificial neural networks: a survey. *J. Intellig. Robotic Syst.* 100, 1469–1491. doi: 10.1007/s10846-020-01227-8
- Gundermann, T. (2020). *Explainable Artificial Intelligence: Shapley Additive exPlanations (SHAP) zur Erklärung der KI-basierten Entscheidungsfindung* (PhD thesis). Technische Hochschule Aschaffenburg, Fakultät Ingenieurwissenschaften, Aschaffenburg, Germany.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Gyagenda, N., Hatilima, J. V., Roth, H., and Zhmud, V. (2022). A review of gnss-independent uav navigation techniques. *Rob. Auton. Syst.* 152:104069. doi: 10.1016/j.robot.2022.104069
- Hashesh, A. O., Hashima, S., Zaki, R. M., Fouda, M. M., Hatano, K., and Eldien, A. S. T. (2022). Ai-enabled uav communications: challenges and future directions. *IEEE Access* 10, 92048–92066. doi: 10.1109/ACCESS.2022.3202956
- Hassanalain, M., and Abdelkefi, A. (2017). Classifications, applications, and design challenges of drones: a review. *Prog. Aerospace Sci.* 91, 99–131. doi: 10.1016/j.paerosci.2017.04.003
- He, L., Aouf, N., and Song, B. (2021). Explainable deep reinforcement learning for uav autonomous path planning. *Aerosp. Sci. Technol.* 118:107052. doi: 10.1016/j.ast.2021.107052
- Hernández, D., Cano, J.-C., Silla, F., Calafate, C. T., and Cecilia, J. M. (2021). Ai-enabled autonomous drones for fast climate change crisis assessment. *IEEE Intern. Things J.* 9, 7286–7297. doi: 10.1109/JIOT.2021.3098379
- Hinton, G., Vinyals, O., and Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv [preprint] arXiv:1503.02531*. doi: 10.48550/arXiv.1503.02531
- Huang, X., Zhang, Y., Qi, Y., Huang, C., and Hossain, M. S. (2024). Energy-efficient uav scheduling and probabilistic task offloading for digital twin-empowered consumer electronics industry. *IEEE Trans. Consumer Electr.* 70, 2145–2154. doi: 10.1109/TCE.2024.3372785
- Hwu, T., Levy, M., Skorheim, S., and Huber, D. (2021). Matching representations of explainable artificial intelligence and eye gaze for human-machine interaction. *arXiv [preprint] arXiv:2102.00179*. doi: 10.48550/arXiv.2102.00179
- Javaid, S., Fahim, H., He, B., and Saeed, N. (2024a). Large language models for uavs: Current state and pathways to the future. *IEEE Open J. Vehicul. Technol.* 5, 1166–1192. doi: 10.1109/OJVT.2024.3446799
- Javaid, S., Fahim, H., Zeadally, S., and He, B. (2023a). Self-powered sensors: applications, challenges, and solutions. *IEEE Sens. J.* 23, 20483–20509. doi: 10.1109/JSEN.2023.3241947
- Javaid, S., Khalil, R. A., Saeed, N., He, B., and Alouini, M.-S. (2024b). Leveraging large language models for integrated satellite-aerial-terrestrial networks: recent advances and future directions. *IEEE Open J. Communications Soc.* 6, 399–432. doi: 10.1109/OJCOMS.2024.3522103
- Javaid, S., Saeed, N., Qadir, Z., Fahim, H., He, B., Song, H., et al. (2023b). Communication and control in collaborative uavs: Recent advances and future trends. *IEEE Trans. Intellig. Transp. Syst.* 24, 5719–5739. doi: 10.1109/TITS.2023.3248841
- Kakaletsis, E., Symeonidis, C., Tzelepi, M., Mademlis, I., Tefas, A., Nikolaidis, N., et al. (2021). Computer vision for autonomous uav flight safety: an overview and a vision-based safe landing pipeline example. *ACM Comp. Surveys (Csur)* 54, 1–37. doi: 10.1145/3472288
- Keane, J. F., and Carr, S. S. (2013). A brief history of early unmanned aircraft. *Johns Hopkins APL Tech. Dig.* 32, 558–571.
- Kendall, A. G., Salvapantula, N. N., and Stol, K. A. (2014). “On-board object tracking control of a quadcopter with monocular vision,” in *2014 international conference on unmanned aircraft systems (ICUAS)* (Orlando, FL: IEEE), 404–411.
- Kim, J.-H., Sukkarieh, S., and Wishart, S. (2006). “Real-time navigation, guidance, and control of a uav using low-cost sensors,” in *Field and Service Robotics: Recent Advances in Reserch and Applications* (Cham: Springer), 299–309.
- Kurunathan, H., Huang, H., Li, K., Ni, W., and Hossain, E. (2023). Machine learning-aided operations and communications of unmanned aerial vehicles: A contemporary survey. *IEEE Commun. Surv. Tutor.* 26, 496–533. doi: 10.1109/COMST.2023.3312221
- Kwak, J., and Sung, Y. (2018). Autonomous UAV flight control for gps-based navigation. *IEEE Access* 6, 37947–37955. doi: 10.1109/ACCESS.2018.2854712
- Lahmeri, M.-A., Kishk, M. A., and Alouini, M.-S. (2021). Artificial intelligence for uav-enabled wireless networks: A survey. *IEEE Open J. Commun. Soc.* 2, 1015–1040. doi: 10.1109/OJCOMS.2021.3075201
- Lim, B., Son, S., Kim, H., Nah, S., and Mu Lee, K. (2017). *Enhanced Deep Residual Networks for Single Image Super-Resolution* (New York, NY: Association for Computing Machinery (ACM)), 136–144.
- Lin, S., Garratt, M. A., and Lambert, A. J. (2017). Monocular vision-based real-time target recognition and tracking for autonomously landing an uav in a cluttered shipboard environment. *Auton. Robots* 41, 881–901. doi: 10.1007/s10514-016-9564-2
- Liu, H., Yu, T., Hu, B., Hou, X., Zhang, Z., Liu, X., et al. (2021). Uav-borne hyperspectral imaging remote sensing system based on acousto-optic tunable filter for water quality monitoring. *Remote Sens.* 13:4069. doi: 10.3390/rs13204069
- Lu, Y., Xue, Z., Xia, G.-S., and Zhang, L. (2018). A survey on vision-based UAV navigation. *Geo-spatial Inform. Sci.* 21, 21–32. doi: 10.1080/10095020.2017.1420509
- Madhuanand, L., Nex, F., and Yang, M. Y. (2021). Self-supervised monocular depth estimation from oblique uav videos. *ISPRS J. Photogramm. Remote Sens.* 176, 1–14. doi: 10.1016/j.isprs.2021.03.024
- Maes, W. H., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. doi: 10.1016/j.tplants.2018.11.007
- Mahmood, K., Tariq, T., Sangaiah, A. K., Ghaffar, Z., Saleem, M. A., and Shamshad, S. (2023). A neural computing-based access control protocol for ai-driven intelligent flying vehicles in industry 5.0-assisted consumer electronics. *IEEE Trans. Consum. Electron.* 70, 3573–3581. doi: 10.1109/TCE.2023.3276066
- Maravall, D., de Lope, J., and Fuentes, J. P. (2015). Vision-based anticipatory controller for the autonomous navigation of an uav using artificial neural networks. *Neurocomputing* 151, 101–107. doi: 10.1016/j.neucom.2014.09.077
- McEnroe, P., Wang, S., and Liyanage, M. (2022). A survey on the convergence of edge computing and AI for UAVS: Opportunities and challenges. *IEEE Intern. Things J.* 9, 15435–15459. doi: 10.1109/JIOT.2022.3176400
- Miclea, V.-C., and Nedeveschi, S. (2021). Monocular depth estimation with improved long-range accuracy for uav environment perception. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15. doi: 10.1109/TGRS.2021.3060513
- Minaeian, S., Liu, J., and Son, Y.-J. (2018). Effective and efficient detection of moving targets from a uav’s camera. *IEEE Trans. Intellig. Transport. Syst.* 19, 497–506. doi: 10.1109/TITS.2017.2782790
- Minh, D., Wang, H. X., Li, Y. F., and Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artif. Intellig. Rev.* 2022, 1–66. doi: 10.1007/s10462-021-10088-y
- Mori, T., and Scherer, S. (2013). “First results in detecting and avoiding frontal obstacles from a monocular camera for micro unmanned aerial vehicles,” in *2013 IEEE International Conference on Robotics and Automation* (Karlsruhe: IEEE).
- Mozaffari, M., Saad, W., Bennis, M., Nam, Y.-H., and Debbah, M. (2019). A tutorial on uavs for wireless networks: applications, challenges, and open problems. *IEEE Commun. Surv. Tutor.* 21, 2334–2360. doi: 10.1109/COMST.2019.2902862
- Nex, F., Armenakis, C., Cramer, M., Cucci, D. A., Gerke, M., Honkavaara, E., et al. (2022). Uav in the advent of the twenties: Where we stand and what is next. *ISPRS J. Photogramm. Remote Sens.* 184, 215–242. doi: 10.1016/j.isprsjrs.2021.12.006
- Padhy, R. P., Choudhury, S. K., Sa, P. K., and Bakshi, S. (2019). Obstacle avoidance for unmanned aerial vehicles: using visual features in unknown environments. *IEEE Consum. Electron. Magaz.* 8, 74–80. doi: 10.1109/MCE.2019.2892280
- Padhy, R. P., Verma, S., Ahmad, S., Choudhury, S. K., and Sa, P. K. (2018a). Deep neural network for autonomous uav navigation in indoor corridor environments. *Procedia Comput. Sci.* 133, 643–650. doi: 10.1016/j.procs.2018.07.099
- Padhy, R. P., Xia, F., Choudhury, S. K., Sa, P. K., and Bakshi, S. (2018b). Monocular vision aided autonomous uav navigation in indoor corridor environments. *IEEE Trans. Sustain. Comp.* 4, 96–108. doi: 10.1109/TSUSC.2018.2810952
- Puente-Castro, A., Rivero, D., Pazos, A., and Fernandez-Blanco, E. (2022). A review of artificial intelligence applied to path planning in uav swarms. *Neural Comp. Appl.* 34, 153–170. doi: 10.1007/s00521-021-06569-4
- Qin, T., and Shen, S. (2017). “Robust initialization of monocular visual-inertial estimation on aerial robots,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Vancouver, BC: IEEE), 4225–4232.
- Ren, W., Ma, L., Gong, M., Zhang, J., Cao, X., and Wei, X. (2018). Gated fusion network for single image dehazing. *IEEE Trans. Image Proc.* 28, 2538–2551. doi: 10.1109/TIP.2018.2883802
- Rezwani, S., and Choi, W. (2022). Artificial intelligence approaches for uav navigation: Recent advances and future challenges. *IEEE Access* 10, 26320–26339. doi: 10.1109/ACCESS.2022.3157626
- Ross, S., Melik-Barkhudarov, N., Shankar, K. S., Wendel, A., Dey, D., Bagnell, J. A., et al. (2013). “Learning monocular reactive uav control in cluttered natural environments,” in *2013 IEEE International Conference on Robotics and Automation (Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE))*, 1765–1772.
- Rovira-Sugranes, A., Razi, A., Afghah, F., and Chakareski, J. (2022). A review of ai-enabled routing protocols for uav networks: Trends, challenges, and future outlook. *Ad Hoc Networks* 130:102790. doi: 10.1016/j.adhoc.2022.102790
- Saha, S., Natraj, A., and Waharte, S. (2014). “A real-time monocular vision-based frontal obstacle detection and avoidance for low cost uavs in gps denied environment,” in *2014 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology* (Yogyakarta: IEEE), 189–195.
- Sarkar, S. (2024). “Unleashing the power of machine learning for enhanced capabilities in consumer electronics drones,” in *2024 IEEE International Conference on Consumer Electronics (ICCE)*, 1–6.
- Schmuck, P., and Chli, M. (2017). “Multi-uav collaborative monocular slam,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 3863–3870.
- Shah, V., and Konda, S. R. (2021). Neural networks and explainable AI: bridging the gap between models and interpretability. *Int J Comp Sci Technol.* 5, 163–176. doi: 10.5281/zenodo.10779335
- Shakhatreh, H., Sawalmeh, A. H., Al-Fuqaha, A., Dou, Z., Almaita, E., Khalil, I., et al. (2019). Unmanned aerial vehicles (UAVS): a survey on civil applications

- and key research challenges. *IEEE Access* 7, 48572–48634. doi: 10.1109/ACCESS.2019.2909530
- Sharma, A., Vanjani, P., Paliwal, N., Basnayaka, C. M. W., Jayakody, D. N. K., Wang, H.-C., et al. (2020). Communication and networking technologies for uavs: a survey. *J. Network Comp. Appl.* 168:102739. doi: 10.1016/j.jnca.2020.102739
- Silva, A., Tambwekar, P., Schrum, M., and Gombolay, M. (2024). “Towards balancing preference and performance through adaptive personalized explainability,” in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY; Piscataway, NJ: Association for Computing Machinery (ACM); Institute of Electrical and Electronics Engineers (IEEE)), 658–668.
- Singh, P., Pandey, P. C., Petropoulos, G. P., Pavlides, A., Srivastava, P. K., Koutsias, N., et al. (2020). “Hyperspectral remote sensing in precision agriculture: Present status, challenges, and future trends,” in *Hyperspectral Remote Sensing* (London: Elsevier), 121–146.
- Su, J., Zhu, X., Li, S., and Chen, W.-H. (2023). Ai meets UAVS: a survey on AI empowered uav perception systems for precision agriculture. *Neurocomputing* 518, 242–270. doi: 10.1016/j.neucom.2022.11.020
- Sullivan, J. M. (2006). Evolution or revolution? The rise of UAVS. *IEEE Technol. Soc. Magazine* 25, 43–49. doi: 10.1109/MTAS.2006.1700021
- Sun, C., Fontanesi, G., Canberk, B., Mohajerzadeh, A., Chatzinotas, S., Grace, D., et al. (2024). Advancing UAV communications: a comprehensive survey of cutting-edge machine learning techniques. *IEEE Open J. Vehicular Technol.* 2024, 1–14. doi: 10.1109/OJVT.2024.3401024
- Tahir, N. U. A., Zhang, Z., Asim, M., Chen, J., and ELAffendi, M. (2024). Object detection in autonomous vehicles under adverse weather: a review of traditional and deep learning approaches. *Algorithms* 17:103. doi: 10.3390/a17030103
- Tang, X., Zhang, Z., and Qin, Y. (2021). On-road object detection and tracking based on radar and vision fusion: a review. *IEEE Intellig. Transp. Syst. Magaz.* 14, 103–128. doi: 10.1109/MTS.2021.3093379
- Tong, P., Yang, X., Yang, Y., Liu, W., and Wu, P. (2023). Multi-uav collaborative absolute vision positioning and navigation: a survey and discussion. *Drones* 7:261. doi: 10.3390/drones7040261
- Ullah, Z., Al-Turjman, F., and Mostarda, L. (2020). Cognition in uav-aided 5g and beyond communications: a survey. *IEEE Trans. Cognit. Commun. Network.* 6, 872–891. doi: 10.1109/TCCN.2020.2968311
- Wang, C., Wang, J., Shen, Y., and Zhang, X. (2019). Autonomous navigation of uavs in large-scale complex environments: A deep reinforcement learning approach. *IEEE Trans. Vehicular Technol.* 68, 2124–2136. doi: 10.1109/TVT.2018.2890773
- Wang, C., Wang, J., Zhang, X., and Zhang, X. (2017). “Autonomous navigation of UAV in large-scale unknown complex environment with deep reinforcement learning,” in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (Montreal, QC: IEEE), 858–862.
- Wang, D., Yang, Q., Abdul, A., and Lim, B. Y. (2019). “Designing theory-driven user-centric explainable AI,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY: Association for Computing Machinery (ACM)), 1–15.
- Wilson, A., Kumar, A., Jha, A., and Cenkeramaddi, L. R. (2021). Embedded sensors, communication technologies, computing platforms and machine learning for uavs: a review. *IEEE Sens. J.* 22, 1807–1826. doi: 10.1109/JSEN.2021.3139124
- Wu, Z., Li, J., Zuo, J., and Li, S. (2018). Path planning of uavs based on collision probability and kalman filter. *IEEE Access* 6, 34237–34245. doi: 10.1109/ACCESS.2018.2817648
- Xin, L., Tang, Z., Gai, W., and Liu, H. (2022). Vision-based autonomous landing for the uav: A review. *Aerospace* 9:634. doi: 10.3390/aerospace9110634
- Xing, T., Zhang, J., Xia, Y., Hou, X., Zhang, Z., and Ren, Y. (2024). “Quantized fuzzy fixed-time target tracking with uncertainties for consumer UAVS,” in *IEEE Transactions on Consumer Electronics* (Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE)).
- Xu, X., Yu, A., Jonker, T. R., Todi, K., Lu, F., Qian, X., et al. (2023). “Xair: a framework of explainable ai in augmented reality,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY: Association for Computing Machinery (ACM)), 1–30.
- Yang, T., Li, P., Zhang, H., Li, J., and Li, Z. (2018). Monocular vision slam-based uav autonomous landing in emergencies and unknown environments. *Electronics* 7:73. doi: 10.3390/electronics7050073
- Zablocki, É., Ben-Younes, H., Pérez, P., and Cord, M. (2022). Explainability of deep vision-based autonomous driving systems: Review and challenges. *Int. J. Comput. Vis.* 130, 2425–2452. doi: 10.1007/s11263-022-01657-x
- Zhang, G., and Hsu, L.-T. (2018). Intelligent gnss/ins integrated navigation system for a commercial uav flight control system. *Aerospace sci. Technol.* 80, 368–380. doi: 10.1016/j.ast.2018.07.026
- Zhao, C., Huang, Y., and Qiu, S. (2019a). Infrared and visible image fusion algorithm based on saliency detection and adaptive double-channel spiking cortical model. *Infrared Phys. Technol.* 102:102976. doi: 10.1016/j.infrared.2019.102976
- Zhao, X., Pu, F., Wang, Z., Chen, H., and Xu, Z. (2019b). Detection, tracking, and geolocation of moving vehicle from UAV using monocular camera. *IEEE Access* 7, 101160–101170. doi: 10.1109/ACCESS.2019.2929760
- Zheng, D., Wang, H., Wang, J., Chen, S., Chen, W., and Liang, X. (2016). Image-based visual servoing of a quadrotor using virtual camera approach. *IEEE/ASME Trans. Mechatron.* 22, 972–982. doi: 10.1109/TMECH.2016.2639531
- Zhu, H., Demirbaga, U., Aujla, G. S., Shi, L., and Zhang, P. (2024). Explainable edge ai framework for IoD-assisted aerial surveillance in extreme scenarios. *IEEE Intern. Things J.* 12, 4570–4578. doi: 10.1109/JIOT.2024.3411528