Autonomous Drone Navigation Using Computer Vision

Inakshi Garg[1], Harsh Panday[2]

Department of Computer Science Chandigarh University Mohali, Punjab, India

Abstract: Autonomous drone navigation with computer vision is an innovative technology that allows drones to move through intricate surroundings without needing human control. This system uses live visual information to identify objects, steer clear of obstacles, and determine routes, improving operational safety and precision. The project's main objective is to create a vision-based navigation system that combines object detection and obstacle avoidance algorithms through deep learning methods like YOLO (You Only Look Once), alongside real-time sensor fusion. The drone uses computer vision algorithms to process aerial images and automatically changes its flight path to prevent crashes. A personalized dataset of aerial images is generated and utilized for improving object detection skills in various environments. In order to guarantee practicality in real- world situations, the system undergoes validation through simulations and field tests on different terrains, focusing on its resilience in changing environments. Improvements in navigation accuracy and obstacle detection are accomplished by implementing adaptive path-planning and integrating multiple sensors, guaranteeing the drone's efficient operation in real-life situations. The goal of this method is to enhance the drone's ability to make decisions, minimize human mistakes, and expand its range of potential uses in areas like surveillance, agriculture, and disaster response.

Keywords: 1. Autonomous Drone Navigation 2. Computer Vision 3. Object Detection 4. Obstacle Avoidance 5. YOLO Algorithm 6. Deep Learning 7. Real-Time Processing 8. Sensor Fusion 9. Path Planning 10. Aerial Image Dataset

I. INTRODUCTION

The use of computer vision technology in autonomous drone navigation is a transformative advancement that is revolutionizing multiple industries. Originally created for military and recreational use, drones have quickly become essential in fields such as agriculture, logistics, environmental monitoring, and surveillance. The increasing need for efficiency, safety, and precision is fueling the rising demand for drones that can navigate independently in complicated environments. Through the use of computer vision, autonomous drones are able to cleverly analyze their environment, steer clear of obstacles, and maneuver autonomously without the need for human input. This opening discusses the need for these technologies, the reasons for their importance, the current abilities of self-piloted drones, and the anticipated advancements in the industry.

A. Need for Independent Drone Guidance

The need for self-directed drone navigation arises from the growing utilization of drones in difficult situations where manual piloting by humans is not feasible or viable. Industries like agriculture, disaster management, and infrastructure inspection need drones to fly autonomously. Agricultural drones fly over vast, unpredictable terrains to check crop health, while disaster response drones work in hazardous, unpredictable environments like collapsed buildings or areas with disrupted GPS signals. Likewise, in the field of logistics, autonomous drones are utilized to transport items in distant or city locations where continual human supervision is not efficient.

Additionally, drones are being utilized by businesses and governments in high-risk areas where it is dangerous or costly for humans to be present. As drones play a larger role in vital tasks like forest monitoring, search-and-rescue missions, and industrial inspections, the demand for dependable, self- sufficient navigation has grown stronger. Drones are able to excel in roles that require both speed and precision by autonomously making real-time decisions and avoiding obstacles.
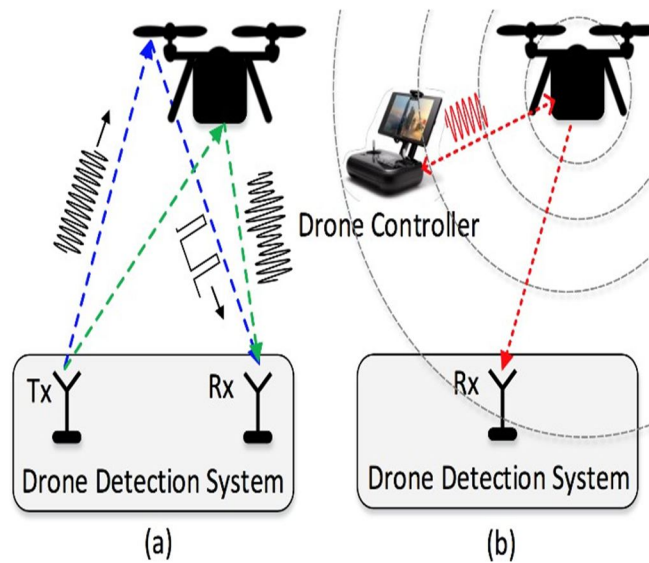
Fig 1

B. The Reason For The Necessity Of Autonomous Navigation

The origins of the demand for self-guided navigation in drones can be traced back to the constraints of manual drone operation. Previous drones needed pilots to control them manually with remote controllers, leading to susceptibility to human mistakes, tiredness, and restricted range of operation. Drones controlled by humans also encountered challenges when reacting to unforeseen obstacles immediately, resulting in accidents or less than optimal performance.

Conventional methods of navigation, like GPS systems, were sufficient for unobstructed areas lacking moving obstacles. Nevertheless, numerous drone functions need to operate in locations with poor or no GPS signals, like indoor spaces, thick forests, or regions impacted by natural calamities. In these situations, drones must depend on sensors and visual information to autonomously navigate and prevent crashes.Moreover, with the expansion in size and intricacy of drone operations, it became evident that manual manipulation was insufficient to meet the requirements for accuracy and swiftness. Having one operator managing several drones or maneuvering through intricate settings like city canyons or beneath tree canopies is not effective or dependable. Self- driving technology became essential to tackle these issues, providing quicker, more accurate decision-making, decreased need for human operators, and improved management of changing surroundings.

C. Capabilities of Self-Driving Drones in the Present

Due to technological progress in computer vision and artificial intelligence, modern autonomous drones can perform intricate functions like identifying objects, avoiding obstacles, and planning paths in real-time. The core of these capacities is found in deep learning algorithms like YOLO, SSD, and other CNNs. These drones have the capability to analyze visual information captured by cameras on board and understand their environment instantly.

Object Detection and Classification: Autonomous drones utilize deep learning algorithms to identify and categorize objects within their flight trajectory. Drones can use visual data to determine how to maneuver, whether it be identifying barriers like trees, buildings, and power lines or detecting moving objects such as vehicles or people.

Avoiding obstacles is another important ability that computer vision allows. Drones use visual data, occasionally along with other sensors such as LiDAR or ultrasonic sensors, to identify and assess the distance to objects. With this data, the drone's navigation system maps out a secure route, preventing crashes through slight alterations to its path. Numerous drones utilize real-time algorithms, enabling them to promptly respond to changing obstacles.

Modern autonomous drones are now fitted with adaptive path-planning algorithms that assist in selecting the most efficient route to reach their destination. They have the ability to alter their course during flight instantly, responding to new obstacles or shifts in weather conditions like wind or lighting.

Sensor Fusion: Drones can merge information from various sensors like visual cameras, LiDAR, infrared, and ultrasonic sensors to enhance accuracy in detecting objects and obstacles. This combination of sensors boosts the drone's ability to perceive its surroundings, particularly in situations with limited visibility like night missions or foggy weather

Fig 2. Sensor Fusion

Data Flow

1) Observe:

Ingests data from all sensor modalities and manufacturers. Available APIs for multi-sensor technology data output ingestion. Random Finite Sets analysis to determine emitter presence (Object Existence Estimation) and tracks (Object Count Estimation).

2) Orient:

Predicts track of detected objects that have disappeared from view. Single-object tracking via Non-linear Multi Hypothesis State Estimation, Gaussian Sum Filters for measurements and Motion Models for Prediction.

3) Decide:

Fuses multiple sensor "hits" if they are from the same UAS. Multiple Object Distinction and Tracking via the use of Density Clustering models and Joint Probabilistic Data Association.

4) Act:

Generates UAS tracks rather than points. Actionable intelligence with rich information, threat level and confidence level. Track Oriented - Multi Hypothesis Tracker (TO-MHT) for data association and multiple-object tracking.

II.   LITERATURE SURVEY

The use of computer vision for autonomous drone navigation has quickly advanced in the last ten years, propelled by progress in AI, ML, and sensor technologies. Incorporating computer vision into drone navigation enables drones to analyze visual information and autonomously make decisions without human involvement. This review discusses the developments in the industry, focusing on important approaches, obstacles, and improvements from previous studies.

A.   The Development Of Independent Drone Navigation

Autonomous drone navigation in its initial phases relied mainly on GPS and inertial measurement units (IMUs). Although they work well in open areas, these systems encounter major difficulties in indoor settings or locations without GPS signal, like forests, city canyons, or disaster areas. Researchers such as Achtelik et al. (2009) were among the first to introduce the concept of vision-based navigation, which involved the use of onboard cameras to gather visual information used by drones to determine their location and surroundings. Visual Simultaneous Localization and Mapping (SLAM) was among the first techniques to enable self-navigation without relying on GPS. SLAM algorithms enabled drones to create maps of their surroundings and simultaneously determine their own position within the environment, which is extremely useful for indoor navigation.

Yet, SLAM faced constraints in settings with moving obstacles, prompting the integration of object detection algorithms to improve navigation precision. For instance, Kalman and colleagues (2012) showed how drones can use live visual information to detect obstacles and adapt to avoid them. The advancement of computer vision methods such as feature extraction and optical flow has been instrumental in allowing drones to perceive depth and speed in relation to obstacles.

Deep Learning and Identifying Objects

Recent studies have greatly improved object detection algorithms, especially with the use of deep learning models like convolutional neural networks (CNNs) leading the way. The introduction of the You Only Look Once (YOLO) algorithm by Redmon et al. (2016) was a significant moment in real-time object detection. YOLO's structure enabled quick and effective detection, making it appropriate for drone navigation where speed is crucial. Research conducted by Wang et al. (2018) highlighted how YOLO can be utilized in self-piloted drone setups to identify and categorize items like cars, plants, and structures instantaneously.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue X Oct 2024- Available at www.ijraset.com

Likewise, Single Shot MultiBox Detector (SSD) and Region- based Convolutional Neural Networks (R-CNN) have been investigated in diverse research scenarios. Aslan et al. (2019) conducted experiments on SSD for the purpose of assisting drones in maneuvering through thick forests, achieving encouraging outcomes in accurately detecting obstacles. Yet, developing lightweight and energy-efficient models for drones continues to be a major issue, since deep learning models typically need a large amount of computational resources.

B. Navigation without collisions and route selection

Even though object detection is important, drones need the capability to navigate around obstacles in constantly changing scenarios. Initial research in this field, like the investigation conducted by Fiorini and Shiller (1998), first introduced the idea of Velocity Obstacle (VO), which involves drones changing their speed and direction to prevent crashes by considering anticipated object movements. This idea developed into Dynamic Window Approaches (DWA), enabling a more responsive way to avoid obstacles. Newer studies have concentrated on combining various sensors, like LiDAR and ultrasonic sensors, with computer vision to enhance accuracy and decision-making abilities (Kumar et al., 2020).

Path planning is also a vital aspect of progress. Conventional path-finding algorithms like A* and Dijkstra's algorithm offer the best navigation paths in unchanging surroundings. Nevertheless, vision-based data has been utilized to improve these algorithms in order to deal with dynamic obstacles. Scholars like Liu et al. (2020) have created mixed algorithms that blend visual data with sensor fusion for real-time autonomous navigation. On the other hand, Reinforcement Learning (RL) is becoming a hopeful option for drones to independently improve their route planning techniques by learning from ongoing interactions with their surroundings.

C. Sensor Fusion and Multi-Modal Strategies

A major advancement in self-navigating drones is the utilization of sensor fusion, which merges information from various sensors such as cameras, LiDAR, radar, and IMUs. This method that utilizes multiple modes improves the precision and dependability of navigation systems, particularly in areas with limited visibility or challenging terrain. Research conducted by Kim et al. (2021) demonstrates the effectiveness of utilizing both visual data and LiDAR to enhance drone performance in scenarios with limited visibility, such as low-light conditions or presence of visual obstacles like smoke or fog. The incorporation of multiple sensors enables enhanced depth perception and motion planning, leading to improved obstacle avoidance.

III.   METHODOLOGY

In the autonomous systems domain, especially with drones, accurate and timely recognition, classification, and reaction to changing surroundings are essential for maintaining operational efficiency. The integration of vision- based navigation with sensor fusion has become a highly dependable and effective approach for drones to autonomously navigate and evade obstacles. Drones can enhance their understanding of the environment and make educated choices instantaneously by combining information from different sensors like optical, radar, and radio frequency (RF) systems. Combining AI-driven algorithms with multiple sensor data enhances precision, durability, and flexibility, particularly in challenging or evolving settings.

We opted for a methodology that utilizes AI-based sensor fusion to gain a thorough understanding of the environment. This includes retrieving, handling, and examining information from various origins (optical, radar, and RF systems) and merging them to allow the drone to identify objects, navigate around obstacles, and form smart choices. The selected approach is thoroughly explained below.
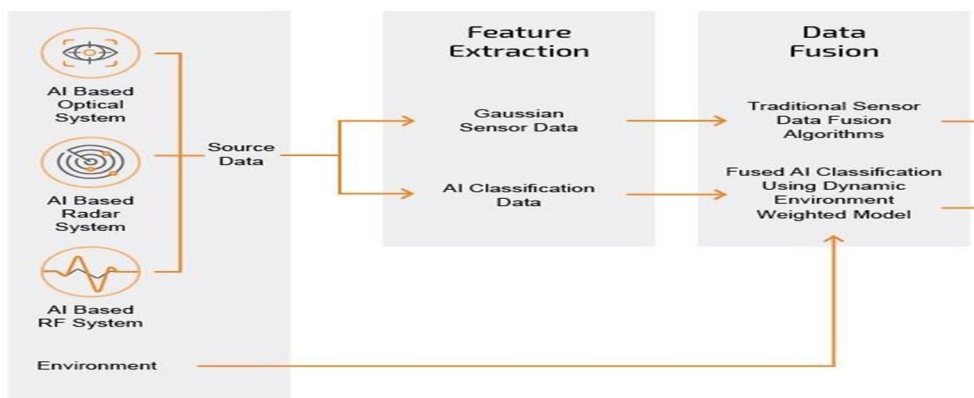

Fig 3 - initial stages of the process

After assessment, the top-performing model is put into operation to consistently observe incoming payments, identifying possibly suspicious transactions for additional

examination or intervention.

Methodology

1) AI-Driven Sensor Systems (Optical, Radar, RF)

This methodology is predicated on the integration of various sensor systems to gather environmental data:

Optical System: This system employs onboard cameras alongside computer vision methodologies to obtain real-time visual information. It plays a crucial role in the detection and identification of objects within the drone's observational range.

Radar System: Radar sensors facilitate the measurement of distances, the detection of object movement, and the tracking of velocity. This system proves particularly advantageous in scenarios where optical systems may falter, such as in low- light conditions, smoke, or fog.

RF System: This system is designed to monitor radio frequencies to identify objects that emit RF signals, including communication devices or other drones. It significantly enhances object recognition, especially when optical or radar systems yield insufficient data.

The utilization of these three distinct data sources within the methodology guarantees both redundancy and reliability in object detection across diverse environmental conditions.

2) Feature Extraction

Following the acquisition of source data from the sensors, the subsequent phase involves feature extraction. This stage encompasses the analysis and processing of raw data to discern essential attributes:

Gaussian Sensor Data: Gaussian filters are utilized to smooth and refine the sensor data, thereby minimizing noise and enhancing the visibility of features such as object edges, shapes, and movement patterns.

AI Classification Data: Artificial intelligence models are deployed to categorize the features derived from the sensor data. For instance, objects are identified as drones, vehicles, or other entities based on the distinctive characteristics of their visual, radar, and RF signatures.

Feature extraction is vital as it simplifies the complexity inherent in raw data, facilitating the identification and classification of objects within the environment.

3) Data Fusion

Following the extraction of features, the methodology employs data fusion to amalgamate information from various sensors:

Traditional Sensor Fusion Algorithms: These algorithms synthesize data from optical, radar, and RF systems to construct a cohesive representation of the environment. The objective of this fusion is to address discrepancies among the sensors, particularly in scenarios where one sensor may identify an object while another fails to do so.

AI-Based Fused Classification: Leveraging artificial intelligence, the system adaptively evaluates and processes the data according to the environmental context. For example, in conditions of reduced visibility, radar data may be prioritized over optical data, thereby facilitating a more accurate understanding of the surroundings.

By integrating data from multiple sensors, the system significantly improves the precision and dependability of object detection and classification, thereby enhancing its resilience in intricate environments.

Fig 4 – End stages of the process

4) Object Detection

Subsequent to the data fusion process, the system identifies and classifies objects within the environment based on three fundamental parameters:

Location: The object's position is ascertained, frequently utilizing GPS or visual triangulation, which provides the drone with precise information regarding the object's whereabouts.

Movement: The trajectory and velocity of the object are assessed, allowing the drone to anticipate future locations and execute evasive maneuvers if required.

Identity: The object is recognized based on its characteristics (e.g., whether it is a drone, vehicle, or human). Artificial intelligence algorithms are utilized to match the object's features with established models for accurate identification.

Object detection is crucial for facilitating real-time decision- making and obstacle avoidance, as the drone must remain cognizant of both stationary and moving objects within its environment.

5) Analytics

Upon the detection of an object, the system initiates an analytical process to further evaluate the circumstances:

Data Retrieval: The system retrieves historical information regarding the identified object, if such data exists, to establish context. For example, if the object is a previously encountered drone, its historical flight patterns may reveal whether it constitutes a potential threat.

System Performance Analysis: Continuous assessment of the system's overall performance is conducted to guarantee optimal functionality. Should any decline in performance be detected, modifications can be implemented to enhance both accuracy and response time.

This analytical layer yields profound insights into the behavior of objects and the efficacy of the system, thereby ensuring that detection remains as precise as possible.

6) Behavioral Analysis

To enhance the decision-making framework, the system scrutinizes the behavior of the identified objects:

Object Properties: This phase involves an examination of the object's dimensions, shape, and other physical characteristics, facilitating a more nuanced classification.

Proximity to Alert Zones: The system assesses whether the object is nearing or encroaching upon a restricted or critical area, which is vital for effective threat detection and defense strategies.

Payload: In instances where the object is identified as a drone or vehicle, the system evaluates whether it is transporting a payload, such as a package or potentially dangerous materials.

Previous Flight Data: If accessible, the system reviews the historical flight data of the object to analyze its behavior over time, aiding in the determination of any potential threat it may pose.

Behavioral analysis guarantees that the system not only identifies objects but also comprehends their possible risks and intentions based on their characteristics and movements.

7) Decision Support

The system ultimately offers decision support by identifying suitable responses based on the recognized objects:

Non-threatening Objects: When an object is assessed as non- threatening, passive strategies such as alert notifications may be implemented.

Threatening Objects: Conversely, if a threat is identified, active strategies including signal jamming, physical barriers, or drone interception are activated to mitigate the danger.

Alert Notifications: Notifications are dispatched to human operators or control systems to inform them of identified objects and any associated threats.

This decision support mechanism guarantees that the system executes the appropriate response according to the characteristics of the object and the potential risks it presents, whether through passive observation or active measures.

IV. CHALLENGES

Ensuring reliability and safety in real-world applications of autonomous drone navigation through computer vision poses various challenges that need to be tackled.

A key obstacle involves developing strong algorithms for detecting and classifying objects. Drones function in changing surroundings with objects that can differ significantly in size, shape, and look. Inclement weather like fog or rain can complicate visibility, so it is essential for the computer vision system to adjust instantly. Maintaining high precision and reducing false alarms is crucial for object detection, particularly when moving near obstacles or in busy environments.

Many object detectors utilizing machine learning and deep learning algorithms struggle to tackle frequently encountered difficulties, which can be grouped as the following:

1) Training at multiple scales: Many object detectors are trained to work effectively at a certain input resolution. Typically, these detectors do not work well with inputs that have varying scales or resolutions.

2) Class imbalance between foreground and background categories can significantly impact model performance.

3) Identifying smaller objects: If the model is trained on larger objects, all object detection algorithms will excel at detecting larger objects. Nonetheless, these models exhibit lower performance when it comes to smaller objects.

4) Deep learning object detection algorithms require extensive computational power, sizable datasets, and labor-intensive annotation methods to function effectively [45]. Because of the rapid growth of data being produced from different sources, annotating every object in visual content has become a time- consuming and laborious task.

5) Smaller datasets: Despite their significant superiority over traditional machine learning methods, deep learning models show weak performance when tested on datasets with limited instances.

6) Incorrect positioning in predictions: Bounding boxes serve as estimates of the actual positions. In general, the algorithm's accuracy is impacted by the inclusion of background pixels in predictions. Generally, localization errors occur because the background is mistaken for objects or when similar objects are detected .

Another important obstacle is combining computer vision with real-time decision-making algorithms. Drones need to rapidly analyze visual information in order to make prompt navigation choices, such as steering clear of barriers or changing their route. This necessitates effective data processing along with the potential utilization of machine learning models that are able to anticipate optimal actions using visual input. Moreover, creating a dependable dataset for training these models is essential. The system needs to cover different situations, lighting conditions, and surroundings in order to perform effectively in new environments. Ultimately, guaranteeing the drone's navigation system can function autonomously while also adhering to safety and regulations increases the level of difficulty, requiring extensive testing and validation in various operational scenarios.

V.   CONCLUSION

To sum up, the difficulties linked with autonomous drone navigation via computer vision emphasize the need for continuous innovation and research in this evolving field. Advancements in machine learning, sensor integration, and real-time data processing will play a crucial role in overcoming current limitations as technology continues to advance. Enhancing object detection algorithms for higher precision in various conditions is essential, as is creating effective decision-making systems that can adjust to changing environments. Overcoming these challenges successfully will improve safety and reliability, leading to increased use of drones in areas like delivery services, agriculture, and search-and-rescue missions.

In the future, there is a possibility of combining new technologies like 5G connectivity and edge computing to improve drone capabilities. These developments can help speed up data transmission and processing, enabling drones to quickly analyze visual inputs and make informed decisions in real-time. Furthermore, when advanced simulation tools are used for training and testing, it will result in the creation of comprehensive datasets, which will in turn enhance the robustness of navigation systems. As rules change to adapt to the increasing use of drones, upcoming advancements offer smart, independent aerial vehicles that can safely and effectively maneuver through challenging spaces, revolutionizing various sectors and enhancing daily experiences.

VI.   FUTURE WORK

Future research in object detection and classification for autonomous drone navigation will concentrate on improving the accuracy and durability of algorithms. Investigating sophisticated deep learning methods like transformer-based models can enhance contextual comprehension and decrease incorrect identifications. Moreover, integrating data from different sensors such as LiDAR and infrared cameras will assist drones in gaining a deeper insight into their environment, especially in difficult situations.

Using reinforcement learning in real-time decision- making can help drones adjust their navigation strategies through feedback from simulated environments. This method will improve their independence and security in changing circumstances. Additionally, delving into collaborative swarm intelligence research can boost efficiency by enabling multiple drones to communicate and coordinate, thereby improving their capacity to navigate intricate environments.

It will be essential to prioritize regulatory compliance and safety as drone technology continues to progress. Working together with regulatory agencies to set clear guidelines and implementing fail-safe mechanisms will guarantee a safe integration into shared airspace. Interacting with local residents to tackle issues regarding privacy and safety will build public approval and confidence, leading to the widespread implementation of autonomous drones in different sectors.

REFERENCES

[1] Qingwei Fu, Qianying Zheng, Fan Yu, "LMANet: A Lighter and More Accurate Multiobject Detection Network for UAV Remote Sensing Imagery", IEEE Geoscience and Remote Sensing Letters, vol.21, pp.1-5, 2024.

[2] Hotaka Oyama, Ryo Iijima, Tatsuya Mori, "DeGhost: Unmasking Phantom Intrusions in Autonomous Recognition Systems", 2024 IEEE 9th European Symposium on Security and Privacy (EuroS&P), pp.78-94, 2024.

[3] Jiabin Pei, Xiaoming Wu, Xiangzhi Liu, Longxiang Gao, Shui Yu, Xi Zheng, "SGD- YOLOv5: A Small Object Detection Model for Complex Industrial Environments", 2024 International Joint Conference on Neural Networks (IJCNN), pp.1-10, 2024.

[4] Akhil Bandarupalli, Adithya Bhat, Saurabh Bagchi, Aniket Kate, Chen-Da Liu-Zhang, Michael K. Reiter, "Delphi: Efficient Asynchronous Approximate Agreement for Distributed Oracles", 2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp.456-469, 2024.

[5] Sairam VC Rebbapragada, Pranoy Panda, Vineeth N Balasubramanian, "C2FDrone: Coarse-to-Fine Drone-to-Drone Detection using Vision Transformer Networks", 2024 IEEE International Conference on Robotics and Automation (ICRA), pp.6627-6633, 2024.

[6] Qiranul Saadiyean, S P Samprithi, Suresh Sundaram, "Learning Multi-Scale Context Mask- RCNN Network for Slant Angled Aerial Imagery in Instance Segmentation in a Sim2Realsetup", 2024 IEEE International Conference on Robotics and Automation (ICRA), pp.13573-13580, 2024.

[7] Xunkuai Zhou, Benyun Zhao, Guidong Yang, Jihan Zhang, Li Li, Ben M. Chen, "SANet: Small but Accurate Detector for Aerial Flying Object", 2024 IEEE International Conference on Robotics and Automation (ICRA), pp.17882-17888, 2024.

[8] Meijia Zhou, Yi Yang, Xuefen Wan, "Non- motorised vehicle recognition system based on drone aerial images and YOLOv8", 2024 Systems and Information Engineering Design Symposium (SIEDS), pp.313-318, 2024.

[9] Fawei Ge, Yunzhou Zhang, Li Wang, Wei Liu, Yixiu Liu, Sonya Coleman, Dermot Kerr, "Multilevel Feedback Joint Representation Learning Network Based on Adaptive Area Elimination for Cross-View Geo-Localization", IEEE Transactions on Geoscience and Remote Sensing, vol.62, pp.1-15, 2024.

[10] Yunzuo Zhang, Cunyu Wu, Tian Zhang, Yuxin Zheng, "Full-Scale Feature Aggregation and Grouping Feature Reconstruction-Based UAV Image Target Detection", IEEE Transactions on Geoscience and Remote Sensing, vol.62, pp.1-11, 2024.

[11] Ronggui Ma, Chen Liang, "Systematic Improvement and Analysis of YOLOv8 for Multiscale Targets in UAV Perspective", 2024 9th International Conference on Computer and Communication Systems (ICCCS), pp.899-905, 2024.

[12] Dong Chen, Duoqian Miao, Xuerong Zhao, "Hyneter:Hybrid Network Transformer for Multiple Computer Vision Tasks", IEEE Transactions on Industrial Informatics, vol.20, no.6, pp.8773-8785, 2024.

[13] Asma Yamani, Albandari Alyami, Hamzah Luqman, Bernard Ghanem, Silvio Giancola, "Active Learning for Single-Stage Object Detection in UAV Images", 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp.1849-1858, 2024.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY