A Systematic Review: Computer Vision Algorithms in Drone Surveillance

Hamzah Fawzy [1], Anas Elbrawy [2], Moustafa Amr [3], Omar Eltanekhy [4], Esraa Khatab [5], and Omar Shalash [6]

[1,2,3,4,6] College of Artificial Intelligence, Arab Academy for Science and Technology and Maritime Transport, New Alamein, Egypt.
[5] School of Mathematics and Computer Science Herriot Watt University, Dubai Knowledge Park, Dubai, UAE.

h.fawzy0048@student.aast.edu, a.elbrawy01235@student.aast.edu, m.mohsen0390@student.aast.edu, o.eltanekhy0345@student.aast.edu, e.khatab@hw.ac.uk, omar.o.shalash@aast.edu

Received on, 23 December 2024 Accepted on, 19 February 2025 Published on, 03 March 2025

ABSTRACT

Drones have emerged as advanced Cyber-Physical Systems (CPSs) with significant potential in data collection and environmental monitoring. Their ability to operate via wireless communication channels makes them integral to various IoT applications, such as surveillance, delivery services, traffic monitoring, and precision agriculture. Drones surpass traditional surveillance methods by offering better mobility and broader coverage, enabling efficient decision-making in diverse contexts; however; however, object detection in drone-captured images poses challenges due to varying spatial resolutions, a large number of objects, and their diverse sizes in aerial imagery. This paper provides a comprehensive review of drone-based surveillance techniques, focusing on object detection and tracking algorithms, relevant datasets, and exploration strategies. By analyzing current methods and identifying key trends, this study aims to highlight advancements and opportunities for improving the performance and reliability of drone-based surveillance systems.

Key-words: Drones, Surveillance, Objects Detection, Objects Tracking

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), also called drones, are considered high-end Cyber-Physical Systems (CPSs) that are used for data collection and environment monitoring [1], [2] This is because they can perform different tasks via wireless communication channels. UAVs can provide real-time data for different IoT applications and enable efficient decision-making, such as surveillance, delivery services, traffic monitoring, and precise agriculture [3]–[5]. Drones could be classified according to their speed, stability, hovering, and their targeted flying environment. Moreover, they could also be classified according to the autonomy and the type of their wings.

There is a high demand for intelligent drones for surveil- lance in contrast to fixed cameras as they provide higher mobility and a bigger surveillance scope. However, object detection is a challenging task when performed on drone-captured images; this is due to spatial sensor resolutions in addition to the large number and diverse sizes of objects in aerial images [6]–[11].

The surveillance task could be described as monitoring a specific target, for example,, environmental phenomena, buildings, people, certain behaviors, or activities [12]. Compared to traditional surveillance methods, drones provide more sustainable solutions; they can perform more complex tasks and can cover larger and harder-to-access areas in a a short time, see Figure 1. The optimization objectives in target tracking based UAV surveillance are:

- To maximize the number of targets for surveillance
- To maximize the quality of Surveillance (QoS) for surveillance
- To minimize the cost of surveillance.

Multiple review studies have addressed different aspects of drone operations, challenges, and applications:

- Surveillance application [13]–[16]
- object detection [17]
- traffic management [17]
- disaster management [15], [17]
- Most commonly, UAV platforms [16], [18]

This paper presents a comprehensive review of the different surveillance techniques performed by drones, including objects detection and tracking algorithms, datasets, and exploration algorithms.

II. DATASETS

A. *Mini Drone Video (MDV)*

The mini-drone video (MDV) dataset [19] was initially proposed in order to design privacy filtering methods. It includes scenes of three different types: normal, suspicious, and abnormal. Therefore, it is used in anomaly detection methods. Later, it was annotated using the ViPEr-GT [20] tool, describing the different objects in each frame, along with their Position; this made the dataset suitable for object detection and tracking methods. This dataset has 38 videos; each is 16-24 seconds. It is recorded by a drone flying at a low altitude. It is divided into a training set of 15 videos and a test set of 23 videos.

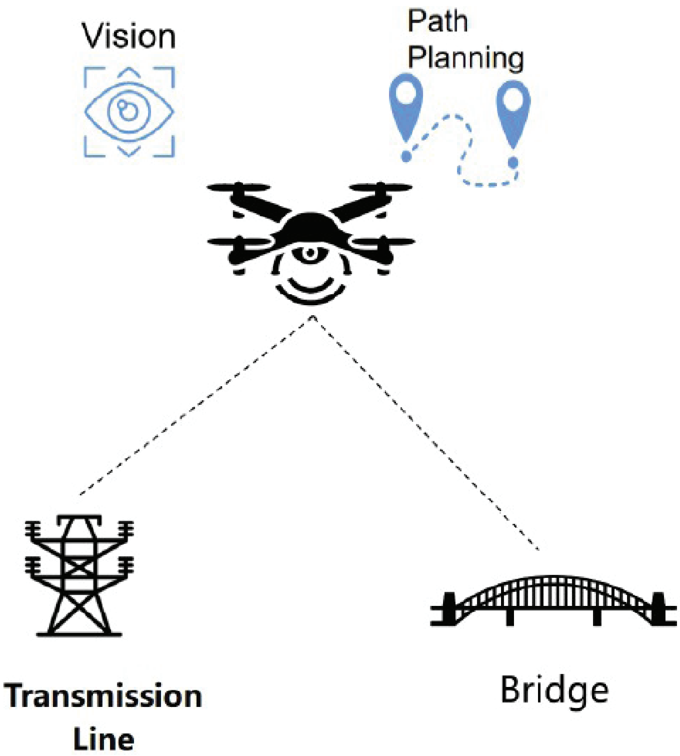

Figure 1. UAV Infrastructure Inspection Based Surveillance. [16]

TABLE I. DIFFERENT ACTIONS ANNOTATED IN THE MDV DATASET

Category	Action
Normal	walking, standing, talking, nothing, parking, parked, moving, stopping
Suspicious	loitering
Abnormal	fighting, picking up, attacking, stealing, cycling, running, falling, repairing

Scene categories are identified by the type of actions people perform in a scene, as shown in Table I. For example, normal scenes include people doing normal activities, such as walking, talking, or parking their cars. Abnormal scenes include illicit situations such as fighting. Suspicious scenes contain non-illicit scenes but would draw the surveillance staff's attention. Moreover, it also includes videos with anomalies made by objects other than human actions, such as cars parking outside parking spots. Finally, the videos are recorded at different times of the day; therefore, the videos have different luminosity levels. The MDV dataset is considered a complex dataset as it includes diverse conditions under which the videos were recorded.

B. UAVid Dataset

The UAVid dataset [21] is a high-resolution aerial video dataset designed for semantic segmentation tasks focusing on urban scenes. It comprises 8 object categories: Buildings, roads, static cars, trees, low vegetation, people, moving cars, and background clutter. The dataset addresses challenges unique to UAV imagery, such as different altitudes, perspectives, and object sizes,, making it suitable for designing computer vision solutions in real-time urban monitoring and surveillance applications. A sample of the UAVid dataset is shown in Figure 2.

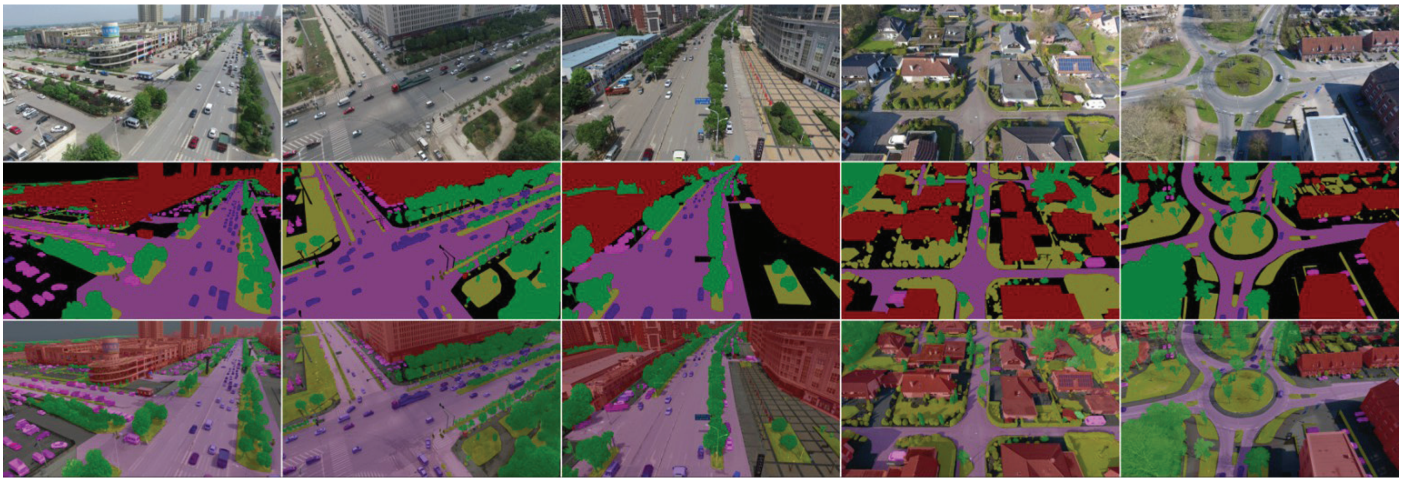

Figure 2. A sample of the UAVid dataset [21]

C. Stanford Drone Dataset

The Stanford Drone Dataset (SDD) [22] is a large-scale dataset that is designed for studying human and object interaction in public outdoor environments. This dataset was captured by drones over the Stanford University campus. It comprises scenes such as walking paths, intersections, and plazas. It consists of 20,000 annotated trajectories for various object classes, including pedestrians, bicycles, skateboards, cars, buses, and golf cars. SSD is widely used for trajectory prediction, object tracking, and activity analysis. A sample of the dataset is shown in Figure 3.

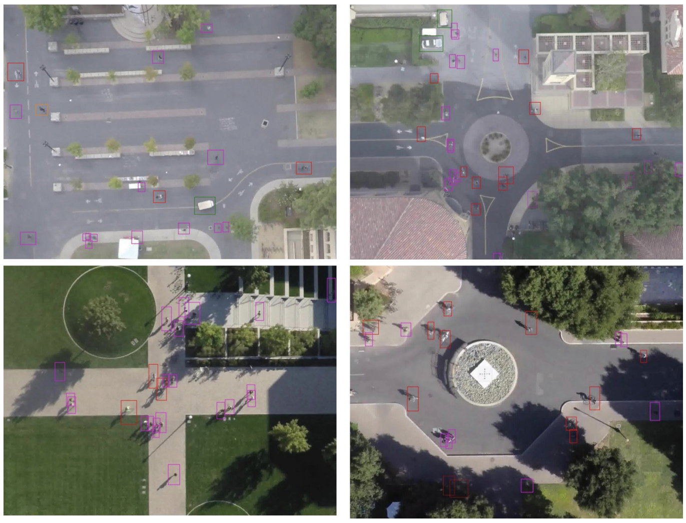

Figure 3. A sample of the Stanford drone dataset [22]

D. Drone Face Dataset

In 2019, Hwai-Jung Hsu et al.et al. [23] presented the DroneFace dataset,, which is used in testing face recognition for drones. DroneFace contains images of facial features having a combinations of different distances and heights for finding out how a face recognition technique works identifying faces from a high altitude. The dataset consists of 2,057 images, 3,680x2,760 resolution, and the altitude of the images were taken from 1.5, 3, 4, and 5 meters high.

E. Drone Surf Dataset

In 2019, Isha Kalra et. al. [24] presented Drone SURF, which is implemented for face recognition, containing 200 videos of 58 subjects from 411,0000 frames, which contains 786,000 facial expressions. The dataset shows different situations across surveillance cases:: active and passive, two different locations, and two different times. DroneSURF sums up difficult challenges due to motion, different poses, illumination, background, height, and quality (resolution).

F. VisDrone

The VisDrone [25] has been collected for visual analysis tasks involving UAVs. It comprises of 400 videos containing 265K frames captured by drones flying over different urban and suburban environments. It includes annotations for multiple object categories with precise bounding boxes (2.6M bounding boxes) and tracking data. It also addresses challenges facing aerial imagery, such as small object sizes, occlusions, and cluttered background.

G. Okutama-Action Dataset

The Okutama-Action Dataset [26] is a drone-based benchmark designed for action detection in aerial videos. It consists of 43 annotated video sequences recording frames of 3840 x 2160 pixels. It was captured from a low-altitude UAV. The dataset contains 12 action categories, such as walking, running, carrying, opening a vehicle, and talking. It addressed challenges such as small objects, motion blur, and different perspectives. A sample is shown in Figure 4.

H. UIT-Aerial Drone Dataset

In 2023, Tung Minh Tran et al. [28] proposed the UIT- ADrone dataset, which seeks to resolve the abnormality identification in urban traffic scenarios, specifically on roundabouts in Ho Chi Minh City, Vietnam. It contains 51 videos with close to 6.5 hours of traffic video material recorded over 206,000 frames and includes ten categories of abnormal activities. This broad dataset has been taken from drones to fill up the gap of resources for unusual event detection in troubled traffic dynamics. The report also inspects the performance of some advanced anomaly detection algorithms and offers preliminary experimental results and their bearing in relation to the other benchmark datasets. Table II shows the actions done by detected objects that were not supposed to happen. Figure 5 shows a sample of this dataset.

A summary of the aerial dataset for drone object detection and tracking is shown in Table III.

Figure 4. Frame captured from Okutama-Action dataset detecting objects [27]

Figure 5. Movement caused by objects captured from the UIT-ADrone dataset [28]

III. SURVEILLANCE ALGORITHMS

In 2009 Axel Bu¨rkle [29] created an architecture of a multi-agent system for the realization of team collaboration in UAVs, where software agents model entities of the UAV team. Thus, the agents mimic the properties and behaviors of their physical forms while their autonomous and context--

TABLE II. COUNT OF ABNORMAL EVENTS IN THE UIT-ADRONE DATASET

Types of abnormal events	Number of actions
Crossing the road at the wrong line	80
Walking under the street	344
Driving the wrong roundabout	636
Driving on the sidewalk	145
Illegal left or right turn	28
Illegal parking on the street	225
Carrying bulky goods	144
Parking on the sidewalk	68
Driving in the opposite direction	214
Falling off motorcycles	1

TABLE III. A SUMMARY OF THE AERIAL DATASET FOR DRONE OBJECT DETECTION AND TRACKING

Dataset Name	Objects	Number of Samples	Resolution
UMN	human	7738	320 x 240
Drone Face	human face	620	3680 x 2760
Drone Surf	human face	411.5k	1280 x 720
Okutama-Action	human actions	77.4k	3840 x 1260
VisDrone	Vehicles	10,209	2000 x 1500
UIT-Drone	vehicles, pedestrians, and roads	2,000	1280 x 720
Stanford Drone Dataset	Skateboard, cart, Pedestrian, Bicyclist, bus, and car	929.5k	1400 x 1904
Mini Drone Video	human actions in the parking area	23.3k	1920 x 1080

TABLE IV. AGENTS USED IN AXEL BU¨ RKLE'S ARCHITECTURE [29]

Agents	Primary/ Secondary agents	Brief Description
Teamleader Agent	primary	Manages a group of agents, coordinates tasks, assigns subtasks, and racks team members' positions and capabilities
Copter Agent	Primary	Individual copters are represented and quadcopter features are modeled, and status information is shared with the assigned team leader
Sensor Agent	Primary	Represents sensors and their properties, assigned to a copter agent. These agents can be combined in particular applications
IRCameraCopter Agent	Secondary	Could integrate a copter agent and a sensor agent to enable infrared imaging
Communication Agent	Secondary	Handles protocols and network configurations concerning inter-agent communication

Aware functionality is ensured. A multi-agent system is ideal for realizing intelligent UAV swarms because agents perceive an environment, act independently, and can achieve certain design goals. The system's architecture is based on primary/ secondary agent classes, as shown in Table IV.

In 2015, the Naval Postgraduate School's (NPS) [30] Advanced Robotic Laboratory Systems Engineering flew 50 autonomous drones all at once. This demonstration proved that autonomous drone swarm technology is evolving at a daunting control. As academia, industry, and defense sectors are continuing to miniaturize sensors and improve swarm operating systems [31], the transition from demonstrations to tactical employment will be rapid. Authors in [30] comprehensively elaborated on drone swarm use in support of a marine infantry company. Simulations showed that the drone swarms allow the fire support team to engage and take on twice as many enemy fighters compared to today's ISR drone available at the company level. For the hierarchical swarm, this amounts to up to 50% fewer casualties. Table V summarizes swarm technology related to drones.

In 2020 J Zhang et al. [32] proposed the following strategies: (1) Stage One, Generating Waypoints, this phase estimates the minimum number of needed waypoints for surveil- lance, (2) Stage Two, UAV Path Planning, the objective of path planning is to find an efficient trajectory for UAVs to visit waypoints while minimizing travel time in the presence of kinematic constraints. This work represents a variant of

TABLE V. SUMMARY TABLE DESCRIBING SWARM TECHNOLOGY AND ITS APPLICATIONS [30]

Aspects	Description
Evolution of Swarm Technology	Enabled en masse deployment, rapid advancements in sensors and operating systems
Research Methods	Agent-based simulation, advanced experimental design, and parallel computing
Scenario	Focused clearance operation involving a marine infantry company fighting a peer adversary in rugged terrain
Findings	Fire support teams engaged twice as many enemies, Hierarchical swarms reduced U.S. casualties by up to 50%, Swarm volume and sensor overlap reduced sensor requirements

The Traveling Salesman Problem by using a clustered spiral-alternating algorithm combined with Be´zier curves for smooth paths. This method shortens the path and is appropriate for low-altitude surveillance.

In 2018 Carlos Paucar et al. [33] performed the Pre- Experimental "One Group Pre-test, Post-test" design to model the algorithm for communication in a system and the detection and tracking of a Parrot Bebop 2 drone. It utilizes Visual Servoing, popularly known as Vision-Based Robot Control, to track the movement of the drone with the help of a camera that feeds spatial data to perform an action. It is based on the methodology of Image-Based Visual Servoing, which estimates the error between current and desired image features, such as visual coordinates, lines, or regional moments. For precise target tracking using the Bebop 2 quad-copter, the sequence of operations in this process involves the following:

- Image acquisition
- Image pre-processing
- Noise reduction
- Segmentation
- Feature extraction
- Recognition and interpretation

In 2022, H Gupta et. al. [34] conducted research to evaluate and optimize DL-based approaches for traffic surveillance and monitoring of aerial imagery. To handle some existing problems such as model generalizability and high-class imbalance in datasets. Traditional object detection techniques such as background subtraction, frame difference, and SVMs are fast but poorly generalized. Deep Learning (DL), particularly with Convolutional Neural Networks (CNNs), has been shown to be highly effective in the detection of complex objects. Examples include SSD for human detection, YOLOv2 for high voltage line detection, YOLOv3 for small object detection, Faster R- CNN, and YOLOv3 for vehicle detection. Although DL has shown excellent performance in UAV-based traffic monitoring, challenges persist. For example, the performance of pre-trained models is poor on aerial data, which decreases practical applications. Also, existing models focus on detecting single components (e.g., cars), and complexity escalates significantly when detecting multiple components (e.g., pedestrians, vans, trucks, bicycles).

In 2020 Dilshad et al. [17] proposed a strong framework underpinning Real-time moving object detection, tracking, and classification from UAV video streams. This model uses SqueezeNet, a lightweight and efficient DL network. The proposed model is more robust in challenging situations with higher accuracy in detection.

In 2018, Chang et. al. [35] came up with a system for drone localization and tracking using acoustic arrays. The Localization Methods include the following:

- **DOA:** signal subspace decomposition techniques that are employed are multiple signal classification (MUSIC) and the incoherent signal-subspace method. However, they are susceptible to noise.
- **Received Signal Strength (RSS):** It is not that common for drone acoustic signals because the RSS values cannot be known beforehand during passive localization and also because environmental interference greatly affects the accuracy.
- **Time Difference of Arrival (TDOA):** It is calculated by the generalized cross-correlation function; it has low computational complexity and high accuracy and robustness. However, it suffers from (1) multi-path caused by the reflection of a signal and (2) poor Signal to Noise (SNR) due to the weak strength of drone acoustic signals.

Due to the shortcomings of TDOA estimation, an enhanced TDOA estimation has been proposed to enhance TDOA estimation by mitigating challenges such as multipath effects and low SNR. Moreover, a localization technique has been developed to accurately determine the position of a drone. Also, a drone tracking algorithm has been proposed to effectively track the drone and make proper predictions of its movement using Kalman Filter. Table VI summarizes the different drone localization and tracking algorithms used and proposed by Chang et al. [35]

IV. EXPLORATION ALGORITHMS

In 2019, J. Sa´nchez-Garc´ıa et al. [36] introduced the PSO-based algorithm in a study named dPSO-U for exploring disaster scenarios by UAVs integrated with a Delay-Tolerant Network (DTN) application. It was assessed by simulations to identify the effective settings and configurations to achieve the fastest search and find victims. The authors discovered that applying the dPSO-U algorithm was able to generate scenarios with varying numbers of victims and clusters and that the overall victim discovering rate was well over 79%. The algorithm was able to successfully converge multiple UAVs to victim clusters.

In 2022, Muhammad Arif Arshad et al. [37] presented the Drone-STM-ResNet architecture to improve the flight dynamics of drones in complex surroundings. It is based on the Split Transform Merge (STM), which is seamlessly embedded in a CNN. The authors focus on two main tasks: (1) predicting the current steering angle and (2) detecting an impending collision. The system produces 96.26% accuracy, 95.47% recall, and 91.95% F-score to develop precise predictions of steering angles and collision risks.

In 2024, Bilal Yousuf et al. [38] developed a new approach to identifying the number of fixed objects by a drone, which was unknown beforehand. The method unifies a multi-target filter for target conditioning, the planner according to target exploration and target refinement, and the way to remove targets once they are identified. The system used is a quad-copter, the Parrot Mambo, equipped with an IMU, an ultrasound sensor, a height pressure sensor, and a downwards-facing camera for the estimation of optical flow, target detection, and localization. Real-time experiments showed that the drone was able to detect targets with moderately high RMSE of actual and estimated targets.

In 2024, Rupayan Das et al. [39] presented a novel Reinforcement Learning (RL) system for the autonomous drone delivery system. They used Unity's ML-Agents toolkit to model a drone that flies around objects while picking targets. The training environment was designed with parameters for incentivizing drone actions: positive rewards for picking tar-gets and negative for hitting obstacles. The authors compared the performance of the curiosity-enhanced PPO algorithm with two baseline algorithms: standard PPO and DQN. The PPO with curiosity has a higher cumulative reward and a better exploration capacity than the other algorithms.

V. CONCLUSION

Drones have proven to be transformative tools in modern surveillance systems, offering unparalleled mobility, efficiency, and adaptability in comparison to traditional methods. Through this review paper, we analyzed existing survey papers, datasets, exploration, and surveillance techniques to provide a comprehensive understanding of the current state of drone-based surveillance. The findings reveal substantial advancements in object detection and tracking algorithms, which play a critical role in enhancing the performance of drones in diverse IoT applications. However, challenges such as varying spatial resolutions, object diversity, and scale in aerial imagery requires further research.

TABLE VI. SUMMARY OF DRONE LOCALIZATION AND TRACKING SYSTEM BY CHANG ET AL. [35]

Aspect	Details
Methods for Localization	• **DOA:** Uses MUSIC; high accuracy, noise-prone. Example: $\hat{\vartheta} = \arg\max_{\vartheta} \frac{1}{\mathbf{a}^H(\vartheta)\mathbf{R}_n\mathbf{a}(\vartheta)}$, where $\mathbf{R}_n$ is the noise covariance matrix. • **RSS:** Rarely used due to environmental interference and unknown RSS values. • **TDOA:** Low complexity and robust. $\Delta t_{ij} = t_i - t_j$, where Δt_{ij} is the time difference between receivers *i* and *j*.
Enhanced TDOA Estimation	• GCC-based a algorithm to enhance TDOA estimation: $\mathrm{GCC}(\tau) = \int_{-\infty}^{\infty} X_1(f)X_2^*(f)e^{j2\pi f\tau}\,df$, where $X_1(f)$ and $X_2(f)$ are the Fourier transforms of the signals at two microphones.
Drone Localization	• Position determination using enhanced TDOA values: $\mathbf{p} = \arg\min_{\mathbf{p}} \sum_{i,j} (\lVert\mathbf{p} - \mathbf{p}_i\rVert - \lVert\mathbf{p} - \mathbf{p}_j\rVert - c\Delta t_{ij})^2$, where $\mathbf{p}_i$ and $\mathbf{p}_j$ are sensor positions, c is the speed of sound, and Δt_{ij} is the TDOA.
Drone Tracking	• Kalman filter for refining positional estimates: $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$, $\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$, where $\mathbf{x}_k$ is the state vector, $\mathbf{z}_k$ is the measurement vector, and $\mathbf{w}_k$ *and* $\mathbf{v}_k$ are noise terms.

REFERENCES

[1] E. Khatab, A. Onsy, and A. Abouelfarag, "Evaluation of 3D Vulnerable Objects' Detection Using a Multi-Sensors System for Autonomous Vehicles," *Sensors*, vol. 22, no. 4, p. 1663, Feb. 2022, doi: 10.3390/s22041663.

[2] Y. Salah, O. Shalash, and E. Khatab, "A lightweight speaker verification approach for autonomous vehicles," *Robotics : Integration, Manufacturing and Control*, vol. 1, no. 2, p. 15, Dec. 2024, doi: 10.21622/RIMC.2024.01.2.1112.

[3] A. Métwalli, M. H. Sallam, E. Khatab, and O. Shalash, "Polygraph-Based Truth Detection System: Leveraging Machine Learning Model on Physiological and Behavioral Data Using Data Fusion," 2024. doi: 10.2139/ssrn.5031332.

[4] E. Khatab, A. Onsy, M. Varley, and A. Abouelfarag, "A Lightweight Network for Real-Time Rain Streaks and Rain Accumulation Removal from Single Images Captured by AVs," *Applied Sciences*, vol. 13, no. 1, p. 219, Dec. 2022, doi: 10.3390/app13010219.

[5] S. Bouassida, N. Neji, L. Nouvelière, and J. Neji, "Evaluating the impact of drone signaling in crosswalk scenario," *Applied Sciences (Switzerland)*, vol. 11, no. 1, 2021, doi: 10.3390/app11010157.

[6] A. Gohari, A. Bin Ahmad, R. B. A. Rahim, A. S. M. Supa'at, S. A. Razak, and M. S. M. Gismalla, "Involvement of Surveillance Drones in Smart Cities: A Systematic Review," *IEEE Access*, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3177904.

[7] O. Shalash and P. Rowe, "Computer-assisted robotic system for autonomous unicompartmental knee arthroplasty," *Alexandria Engineering Journal*, vol. 70, pp. 441–451, May 2023, doi: 10.1016/j.aej.2023.03.005.

[8] Z. Hua, "From Battlefield to Border: The Evolving Use of Drones in Surveillance Operations," *ITEJ (Information Technology Engineering Journals)*, vol. 9, no. 1, pp. 44–52, Jul. 2024, doi: 10.24235/itej.v9i2.126.

[9] Z. Fang and A. V. Savkin, "Strategies for Optimized UAV Surveillance in Various Tasks and Scenarios: A Review," *Drones*, vol. 8, no. 5, p. 193, May 2024, doi: 10.3390/drones8050193.

[10] N. Dilshad, J. Y. Hwang, J. S. Song, and N. M. Sung, "Applications and Challenges in Video Surveillance via Drone: A Brief Survey," in *International Conference on ICT Convergence*, 2020. doi: 10.1109/ICTC49870.2020.9289536.

[11] J. Doornbos, K. E. Bennin, O. Babur, and J. Valente, "Drone Technologies: A Tertiary Systematic Literature Review on a Decade of Improvements," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3364676.

[12] M. Bonetto, P. Korshunov, G. Ramponi, and T. Ebrahimi, "Privacy in mini-drone based video surveillance," in *Proceedings - International Conference on Image Processing, ICIP*, 2015. doi: 10.1109/ICIP.2015.7351245.

[13] D. Doermann and D. Mihalcik, "Tools and techniques for video performance evaluation," *Proceedings - International Conference on Pattern Recognition*, vol. 15, no. 4, 2000, doi: 10.1109/icpr.2000.902888.

[14] Y. Lyu, G. Vosselman, G. S. Xia, A. Yilmaz, and M. Y. Yang, "UAVid: A semantic segmentation dataset for UAV imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 165, 2020, doi: 10.1016/j.isprsjprs.2020.05.009.

[15] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, "Learning social etiquette: Human trajectory understanding in crowded scenes," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016. doi: 10.1007/978-3-319-46484-8_33.

[16] M. Nalamati, A. Kapoor, M. Saqib, N. Sharma, and M. Blumenstein, "Drone detection in long-range surveillance videos," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2019*, 2019. doi: 10.1109/AVSS.2019.8909830.

[17] I. Kalra, M. Singh, S. Nagpal, R. Singh, M. Vatsa, and P. B. Sujit, "DroneSURF: Benchmark dataset for drone-based face recognition," in *Proceedings - 14th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2019*, 2019. doi: 10.1109/FG.2019.8756593.

[18] P. Zhu *et al.*, "Detection and Tracking Meet Drones Challenge," *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 11, 2022, doi: 10.1109/TPAMI.2021.3119563.

[19] M. Barekatain *et al.*, "Okutama-Action: An Aerial View Video Dataset for Concurrent Human Action Detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2017. doi: 10.1109/CVPRW.2017.267.

[20] S. Sambolek and M. Ivasic-Kos, "Person Detection in Drone Imagery," in *2020 5th International Conference on Smart and Sustainable Technologies, SpliTech 2020*, 2020. doi: 10.23919/SpliTech49282.2020.9243737.

[21] T. M. Tran, T. N. Vu, T. V. Nguyen, and K. Nguyen, "UIT-ADrone: A Novel Drone Dataset for Traffic Anomaly Detection," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 16, 2023, doi: 10.1109/JSTARS.2023.3285905.

[22] A. Bürkle, "Collaborating miniature drones for surveillance and reconnaissance," in *Unmanned/Unattended Sensors and Sensor Networks VI*, 2009. doi: 10.1117/12.830408.

[23] N. J. Gulosh, "Employment Of Intelligence, Surveillance, And Reconnaissance Drone Swarms To Enhance Ground Combat Operations," Thesis (Ph.D.), Naval Postgraduate School, Monterey, CA, 2018.

[24] M. Yasser, O. Shalash, and O. Ismail, "Optimized Decentralized Swarm Communication Algorithms for Efficient Task Allocation and Power Consumption in Swarm Robotics," *Robotics*, vol. 13, no. 5, p. 66, Apr. 2024, doi: 10.3390/robotics13050066.

[25] J. Zhang and Y. Zhang, "A method for UAV reconnaissance and surveillance in complex environments," in *2020 6th International Conference on Control, Automation and Robotics, ICCAR 2020*, 2020. doi: 10.1109/ICCAR49639.2020.9107972.

[26] C. Paucar *et al.*, "Use of drones for surveillance and reconnaissance of military areas," in *Smart Innovation, Systems and Technologies*, 2018. doi: 10.1007/978-3-319-78605-6_10.

[27] H. Gupta and O. P. Verma, "Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach," *Multimed Tools Appl*, vol. 81, no. 14, 2022, doi: 10.1007/s11042-021-11146-x.

[28] X. Chang, C. Yang, J. Wu, X. Shi, and Z. Shi, "A Surveillance System for Drone Localization and Tracking Using Acoustic Arrays," in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2018. doi: 10.1109/SAM.2018.8448409.

[29] J. Sánchez-García, D. G. Reina, and S. L. Toral, "A distributed PSO-based exploration algorithm for a UAV network assisting a disaster scenario," *Future Generation Computer Systems*, vol. 90, 2019, doi: 10.1016/j.future.2018.07.048.

[30] M. A. Arshad *et al.*, "Drone Navigation Using Region and Edge Exploitation-Based Deep CNN," *IEEE Access*, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3204876.

[31] S. Wei, S. Phanse, T. Koduru, M. Chen, D. W. Yeo, and M. Z. Li, "Exploration-Exploitation Guided Drone Dispatch Strategies in Uncertain Environments," in *AIAA SCITECH 2024 Forum*, Reston, Virginia: American Institute of Aeronautics and Astronautics, Jan. 2024. doi: 10.2514/6.2024-1079.c1.

[32] R. Das, A. Khan, and G. Paul, "A proximal policy optimization with curiosity algorithm for virtual drone navigation," *Engineering Research Express*, vol. 6, no. 1, 2024, doi: 10.1088/2631-8695/ad1f14.