



Applications, databases and open computer vision research from drone videos and images: a survey

Younes Akbari¹ · Noor Almaadeed¹ · Somaya Al-maadeed¹ · Omar Elharrouss¹

Published online: 22 February 2021

© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract

Analyzing videos and images captured by unmanned aerial vehicles or aerial drones is an emerging application attracting significant attention from researchers in various areas of computer vision. Currently, the major challenge is the development of autonomous operations to complete missions and replace human operators. In this paper, based on the type of analyzing videos and images captured by drones in computer vision, we have reviewed these applications by categorizing them into three groups. The first group is related to remote sensing with challenges such as camera calibration, image matching, and aerial triangulation. The second group is related to drone-autonomous navigation, in which computer vision methods are designed to explore challenges such as flight control, visual localization and mapping, and target tracking and obstacle detection. The third group is dedicated to using images and videos captured by drones in various applications, such as surveillance, agriculture and forestry, animal detection, disaster detection, and face recognition. Since most of the computer vision methods related to the three categories have been designed for real-world conditions, providing real conditions based on drones is impossible. We aim to explore papers that provide a database for these purposes. In the first two groups, some survey papers presented are current. However, the surveys have not been aimed at exploring any databases. This paper presents a complete review of databases in the first two groups and works that used the databases to apply their methods. Vision-based intelligent applications and their databases are explored in the third group, and we discuss open problems and avenues for future research.

Keywords Drones · Survey article · Computer vision · Remote sensing · Navigation · Applications · Database · Open research

✉ Younes Akbari
akbari_younes@semnan.ac.ir

Noor Almaadeed
n.alali@qu.edu.qa

Somaya Al-maadeed
s_alali@qu.edu.qa

Omar Elharrouss
elharrouss.omar@gmail.com

¹ Department of Computer Science and Engineering, Qatar University, Doha, Qatar

1 Introduction

The advantages of using drones, compared with enduring platforms (manned aerial vehicles and satellites), are the lower-altitude flights, images with high spatial resolution and low cost use and maintenance for monitoring and sensing environments. In recent years, the powerful features of drones have been improving, resulting in drones becoming a major field instrument for researchers. Thus, an increasing number of companies are being persuaded by the simple mechanics of drones for surveillance and infrastructure inspection applications. Generally, drones have the ability to fly at various speeds indoors or outdoors and control their position around targets and obstacles using various sensors to detect their environment. All of these advantages and features make them increasingly suitable to replace human operations in situations in which experts cannot participate, especially in dangerous, difficult, expensive or exhausting conditions Kanellakis and Nikolakopoulos (2017). Drones can be controlled remotely from a ground control station (GCS) by the pilot (remotely piloted aerial system (RPAS)) or can be automated by the onboard, programmable sensors mounted on it. As a vehicle, drones refer to the supporting hardware such as sensors, microcontrollers, ground stations, and software including communication protocols, and user interfaces. To perform most of the unmanned aerial vehicle (UAV) applications, the computer vision method has a vital role. Computer vision aims to interpret the 3D world into metric data by processing 2D images of planes in different applications. Each computer vision simulation should consider four tasks, namely, acquiring, processing, analyzing and understanding digital videos and images Elharrouss et al. (2020). The image deciphering, assists in automating the real-world problems, especially those that are difficult to the average human to perceive. The computer vision methods in drone applications ranges from basic and simple aerial imagery to super complex tasks such as aerial refueling or rescue operations. The methods for performing the application accurately require reliable decision-making and precise maneuvering tasks Al-Kaff et al. (2018).

In this paper, based on videos and images captured by drones in computer vision, we present a survey of works that have introduced a database for the various applications of the videos and images and works that have used these databases. We have categorized applications into three groups. The first group of applications is related to remote sensing with challenges such as camera calibration, image matching, and aerial triangulation. The second group of applications use their own drone navigation in which computer vision methods are designed to explore challenges such as flight control, visual localization and mapping, and target tracking and obstacle detection. The third group of applications is dedicated to using images and videos captured by drones in applications such as surveillance, agriculture and forest, animal detection, disaster detection, and face recognition. This survey summarizes the knowledge generated by 228 articles and provides insights based on many additional articles and supporting literature. A statistical report on the surveyed literature from 2005 until the present (October 2020) is shown in Fig. 1. As it can be verified in the figure that growing academic interest based on papers that provide databases on the topic in terms three categories, from 2017 until the present (October 2020). All works have been categorized into 116 journals, 75 conferences, 28 preprints, 6 reports and 3 books/thesis classes.

For researchers, our survey is an introduction to open research. Additionally, we provide an overview of the existing literature and present databases for remote sensing and navigation based on computer vision and applications related to images captured by drones.

Because of the breadth of the research area, many other related topics considering our literature survey are considered in this article:

- An accurate, sharp boundary cannot be found to separate military security and civil applications. We try to include articles with civil and commercial applications of drones, which generally can be used in both contexts.
- We try to concentrate on outdoor applications, which can also be used in indoor environments.
- Since the infrastructure of each application in computer vision is a database, we have focused on databases in various applications and works related to these databases.
- Papers were explored that explicitly introduce or use databases based on drones in their title, abstract, or keywords or used databases in the experimental results section using any relevant term or description. Our keywords for searching were “vehicle, UAV, drone, unmanned aircraft, unmanned aerial system (UAS), remotely piloted aerial system (RPAS), and remotely piloted vehicle”, but they were not limited to these words. Then, we choose only papers related to computer vision and those that created a database. It should be noted that topics related to drones such as operations planning of mobile robots (including ground-based drones), mobile sensors, vehicle routing and machine scheduling are not part of this survey.
- Papers were considered if they met the following publishing criteria: peer-reviewed English journals, peer-reviewed conference proceedings, or recent manuscripts from open-source archives. Additionally, we have tracked all the studies from authors who are distinguished and experts in the field. Due to a large number of publications, we are unable to include all the publications. However, we try to include important articles on the topic.
- All papers of the sections (text, tables, figures, and plots) are sorted in terms of categories and publication year.

This paper presents comprehensive insights into the evaluation and benchmarking of videos and images captured by drones based on the three categories, as shown in Fig. 2.

We provide a background on drones and their developed applications based on computer vision in Sect. 2. In Sect. 3, we summarize databases related to remote sensing and navigation groups. Based on our survey categories in the literature, we then describe applications that can be applied to images and videos captured from drones in Sect. 4. Section 5 is devoted to open challenges and research that can be done in the future. Section 6 outlines future directions and concludes our article.

2 Background

In this section, we first present a history of and developments in drone technology and then a brief description of the types of drones and cameras used are presented. While a brief description of surveys presented up until recently is described, finally, a general description of drone-based computer vision is presented.

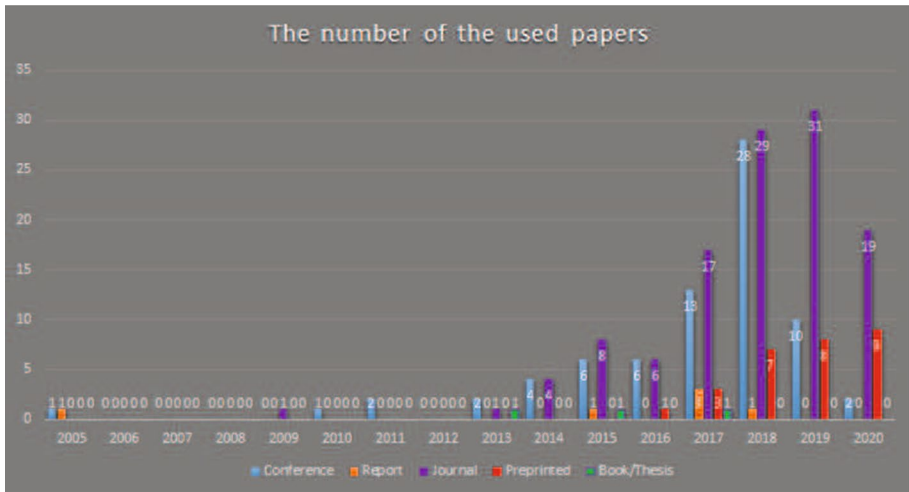


Fig. 1 Number of papers used in this survey by publication year (from 2005 to October 2020)

2.1 History and developments

The history of using drones dates back to the First Italian War of Independence (1849), when, in response to dropping bombs on Venice, a system of unmanned hot air balloons were designed by the Austrian Empire. This development led to the use of hot air balloons and kites for communication during the American Civil War and the Spanish–American War, and it has endured and been developed for military use until the twenty first century. Advances were observed when the tensions between the U.S. and the Soviet Union during the Cold War were increasing, during which the U.S. government started a UAS research program under the code name “Red Wagon”. In parallel with the advances, the first version of the Global Positioning System (GPS) based on the global satellite navigation system was introduced by the Defense Advanced Research Projects Agency (DARPA). The genesis of commercial application of drones was in 2006, as shown in Fig. 3. The figure summarizes the commercial aspects of drones until the present. Dajiang (DJI), as a leader in the commercial and civilian drone industry, created the first commercial drone in 2006. DJI has steadily developed drones for various applications around the world. From 2012, the Federal Aviation Administration (FAA), according to U.S. law, has managed to integrate small drones into the airspace and reports each year details about the drones, including the number and distance limitations. In 2013, Amazon announced plans to deliver products by drones. For more information regarding history and development of drones, readers can refer to Rakha and Gorodetsky (2018).

As shown in Fig. 3, from 2018, researchers in computer vision based on drone systems have produced and developed databases. To the best of our knowledge, no one field or industry has presented a comprehensive review of studies using databases of videos and images captured from drones. In addition, since each study based on computer vision methods has specific database needs in various applications, a summary of all of them is useful for continuing research. This survey is dedicated to gathering, comparing, contrasting, and assessing current and emerging research in drone fields based on the databases created.

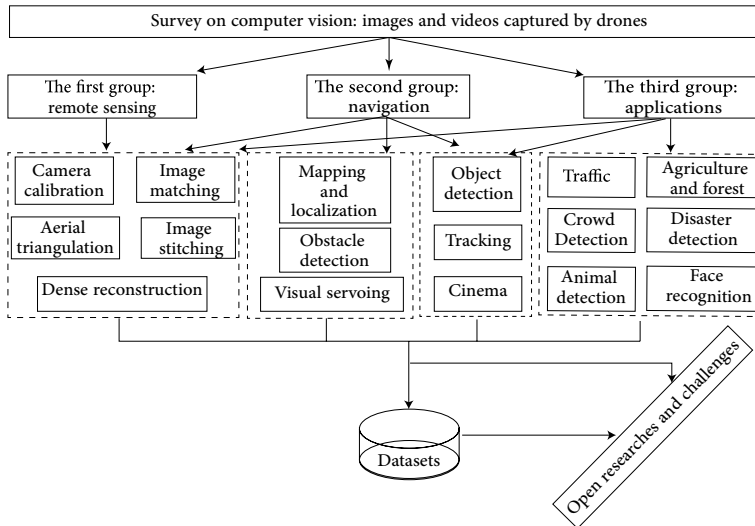


Fig. 2 Overview of our survey structure

2.2 UAVs and cameras types

Drones fly without needing roads, and thus, they can reach difficult locations for various aims. Many companies have produced types of drone models for different missions to reduce labor costs. In the production process, issues such as the weight of the aircraft and thus its energy consumption, thermal control, and cabin pressurization are important. Figure 4 illustrates several models of drones.

To sense different situations, we need to use a variety of sensors. For example, to sense the environment and estimate their position and orientation in space, exteroceptive, and proprioceptive sensors such as the global positioning system (GPS) are mounted on drones. In addition, drones can be equipped and embedded with different types of sensors to extract useful data and information. Ultrasonic sensors and visual stereo or monocular camera systems can be directly used to detect and avoid obstacles and map 3D environments. This can be integrated with laser range finders and inertial measurement units (IMUs) to provide more accurate results and visual-inertial ego-motion estimation. Some examples of modular vision systems are depicted in Fig. 5. In this survey, we explore images and videos, and accordingly, we consider studies that include a camera as a primary or secondary sensor.

2.3 Related to previous surveys

A number of representative surveys concerning drone-based computer vision have been presented, as summarized in Table 1. The research reviewed in Colomina and Molina (2014), Pádua et al. (2017) was dedicated to presenting 3D reconstruction and geometric correction methods. Reference Xiang et al. (2018) focused on surveying the issues of specific aerial remote sensing data processing, such as image matching and dense image matching. As mentioned in Xiang et al. (2018), some other drone data processing technologies and their recent advances were presented with a focus on deep learning and related

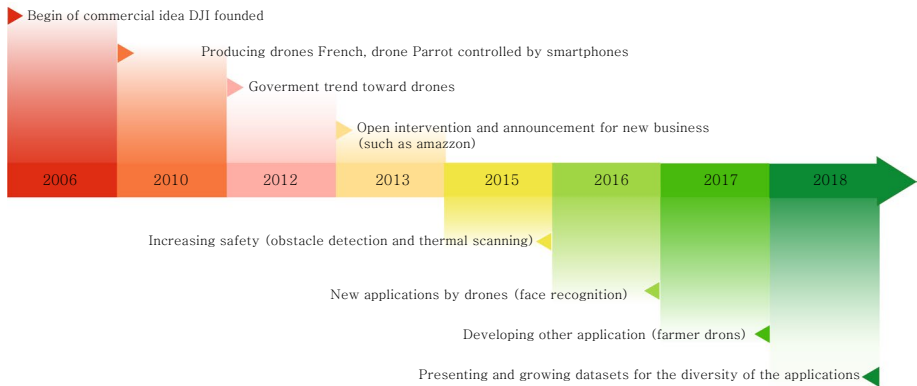


Fig. 3 A historical timeline of UAS technology developments based on commercial aspects

methods on drone data geometric processing. References Kanellakis and Nikolakopoulos (2017) and Al-Kaff et al. (2018) provide a comprehensive review of navigation systems, which also include advances in computer vision. However, recent developments in current procedures and methodologies of drone-based thermal imaging practices were detailed in Rakha and Gorodetsky (2018). In addition, some surveys reviewed specific applications of UAVs in remote sensing fields, such as agriculture Gago et al. (2015), forestry Yuan et al. (2015), disaster Adams and Friedland (2011), Giordan et al. (2018) and surveillance Puri (2005), Kanistras et al. (2015). Extensive work on other hot issues, such as optimization approaches for civil applications Otto et al. (2018) and machine learning approaches Choi and Cha (2019), was explored separately. Considering the problems discussed above, it is imperative to provide a comprehensive survey of drones, centering on drone-based computer vision methods based on databases, recent applications, and future directions. A thorough review and summarization of existing work is essential for further progress in drone computer vision, particularly for researchers wishing to enter the field. The objectives of this paper are the following:

- a systematic survey of computer vision methods based on databases are categorized into three different themes; (in each section, we provide a critical overview of databases and the methods applied to them);
- a detailed overview of recent potential applications of drones in computer vision tasks;
- a discussion of the future directions and challenges of drones from the point of view of databases.

2.4 General description of UAV-based computer vision

Today, computer vision methods are applied in most drone applications. By developing computer vision algorithms and decreasing their errors and embedding them into sensors, drones can not only be used for simple applications such as photography and filming but also in more complex applications. Cameras obtain images and videos. After obtaining images and videos by drones, tasks related to them (e.g., image processing and analysis



Fig. 4 Different models of drones

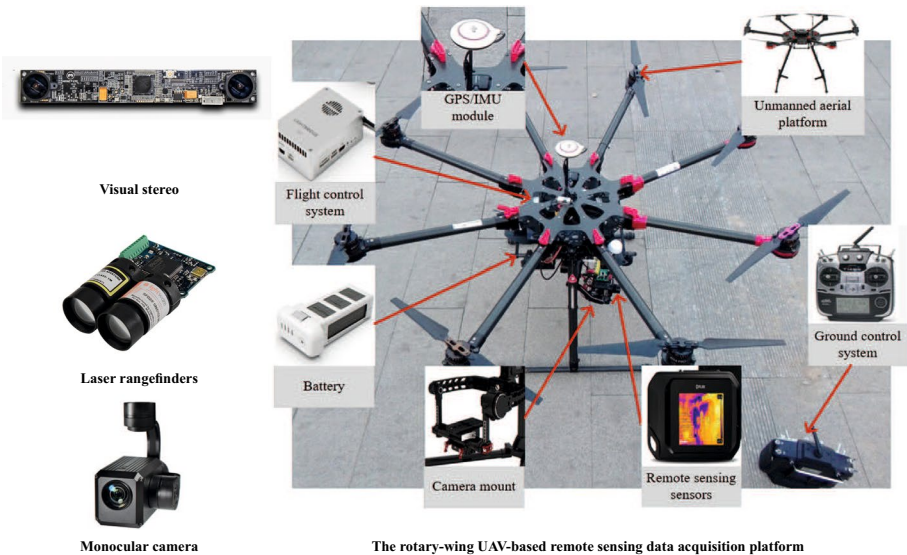


Fig. 5 Different models of cameras used in computer vision applications

Table 1 List of a number of related surveys on UAVs in recent years

Survey	Category	Year	Content
Colomina and Molina (2014)	Remote sensing	2014	A survey of recent technologies (until 2014) in drones and its applications in photogrammetry and remote sensing
Pádua et al. (2017)	Remote sensing	2017	A survey of sensors for collecting and processing data, applications of drones in agroforestry, and some open researches
Xiang et al. (2018)	Remote sensing	2018	A comprehensive survey of mini-UAV-based remote sensing, focusing on techniques, applications and future development
Kanellakis and Nikolakopoulos (2017)	Navigation	2017	A survey on computer vision for controlling UAVs based in current developments and trends
Al-Kaff et al. (2018)	Navigation	2018	A survey of computer vision methods and applications for drones
Rakha and Gorodetsky (2018)	Navigation	2018	A survey on drones applications in the built and application environment for automating building inspection procedures
Puri (2005)	Application	2005	A survey of drones for traffic surveillance (until 2005)
Kanistras et al. (2015)	Application	2015	A survey of drones for traffic surveillance (until 2015)
Gago et al. (2015)	Application:agriculture	2015	A survey on drones challenges for sustainable agriculture
Giordan et al. (2018)	Application	2018	A survey of the use of drones for natural hazards monitoring and management
Otto et al. (2018)	Applications	2018	A literature review on optimization approaches for civil applications of drones
Choi and Cha (2019)	Application	2019	A survey on current machine learning methods for autonomous flight
Kerle et al. (2020)	Application:disaster	2020	A review on structural damage mapping
Cazzato et al. (2020)	Application:object detection	2020	A survey on computer vision methods for 2D object detection from drones
Ilyas et al. (2020)	Application:crowd counting	2020	A survey on convolutional-neural network-based image crowd counting
Adams and Friedland (2011)	Remote sensing and application	2011	A survey of drones usage for collecting data in disaster environments
Yuan et al. (2015)	Remote sensing and application	2015	A survey for exploring automatic forest fire monitoring, detection, and fighting using drones in terms of techniques

to collect scene information, including drone attitude and position) can be considered. Additionally, the distance of drones from the building should be considered. The distance depends on the laws of the specific country (for example, in the U.S., the FAA controls and manages the rules). However, the distance for commercial purposes is almost 5 m. All the distance variations over the years can be found in Rakha and Gorodetsky (2018). The term computer vision includes characteristics and analyses of the real 3D world from 2D image planes. To implement a computer vision system, three fields are involved, namely, image processing, pattern recognition and machine learning. In the first step, it needs to use image processing methods to provide images and videos to execute processes, such as noise removal and morphology tools. Then, in terms of each application, several methods are applied on the processed images to extract features and patterns. Finally, machine learning methods are used to learn the various patterns to automate the processes. In new machine learning methods, such as deep learning, all or two steps are integrated. Computer vision, in general, focuses on interactions with the environment as well as the basic applications of machine inspection, navigation, 3D model building, and surveillance. One of the other contexts related to drones is imaging, which includes the process of producing images and involves image processing and computer vision. Consequently, the development of drones and their corresponding capabilities in computer vision can be used in object recognition, object tracking, pose estimation, ego-motion estimation, optical flow, and scene reconstruction Kanellakis and Nikolakopoulos (2017).

In the following section, we present information about events related to computer vision and drones, as shown in Table 2. It should be noted that in conjunction with ICCV 2017¹, ECCV 2018² and CVPR 2019³, three workshops based on computer vision problems for drones have been presented. Each year starting in 2013, the International Conference on Unmanned Aircraft Systems presents new issues in the field. Additionally, many competitions are organized that use images and videos captured by UAVs, such as Kristan et al. (2017). As shown in our references, journals attracting the most attention in the field were Sensors, Remote Sensing, and the IEEE Transactions journals. It should be noted that most of the detection methods in computer vision applications are real-time methods and since most of researchers cannot provide a real condition for their methods based on drones, in next sections, we consider papers that provide a database for these purposes. In the survey, we explored all databases, including RGB, thermal and multispectral images and videos. Additionally, we consider the databases in terms of the type of availability: public, private and upon request.

3 Remote sensing and navigation databases

With the increase in the number of applications by drones in recent decades, advances in photogrammetry and remote sensing have turned into a commercial competition. In remote sensing, it is important to know the quality of the information and the acquisition obtained

¹ Workshop in conjunction with International Conference on Computer Vision: <https://sites.google.com/site/uavision2017/>.

² Workshop in conjunction with European Conference on computer vision: <https://sites.google.com/site/uavision2018/>.

³ Workshop in conjunction with Conference on Computer Vision and Pattern Recognition: <https://sites.google.com/site/uavision2019/>.

Table 2 The journals and conferences attracting the most attention in the UAV based on computer vision field

Journal or conference	Category	No. of papers	References
IEEE Transaction journals	Navigation, obstacle detection, traffic, crowd and object detection	16	Avola et al. (2018), Bharati et al. (2018), Fan and Ling (2019), Ke et al. (2017), Zhang et al. (2020), Minaeian et al. (2018), Rozantsev et al. (2017), Sommer et al. (2018), Stahl et al. (2019), Tzelepi and Tefas (2019), Zhang et al. (2018), Zimmermann et al. (2009), Ke et al. (2018), Li et al. (2019), Kellenberger et al. (2019), Chen et al. (2020)
Remote sensing (journal)	Obstacle detection, animal detection, agriculture, disaster, object detection and navigation	12	Adão et al. (2017), Bejiga et al. (2017), Dandois et al. (2015), Duarte et al. (2018), Nex et al. (2019), Rahmoumfar et al. (2019), Turner et al. (2014), Xu et al. (2016), Xue et al. (2018), Zhu et al. (2019), Zhang et al. (2020), Zhang et al. (2020)
Sensors (journal)	Obstacle detection, animal detection, agriculture, disaster, object detection and navigation	11	Al-Kaiff et al. (2017), Gonzalez et al. (2016), Kragh et al. (2017), Li et al. (2018), Rivas et al. (2018), Tian et al. (2016), Xu et al. (2016), Xue et al. (2018), Barbedo et al. (2019, 2020), Ilyas et al. (2020)
ECCV (conference)	Navigation, object detection, animal detection, face recognition, crowd detection and camera calibration	11	Ballan et al. (2016), Du et al. (2018), van Gemert (2014), Marcu et al. (2018), Mueller et al. (2016), Muller et al. (2018), Robicquet et al. (2016), Layne et al. (2014), Majid Azimi (2018), Yin et al. (2018), Zhu et al. (2018)
ICCV (conference)	Obstacle detection, object detection and navigation	9	Berker Logoglu et al. (2017), Cehovin Zajc et al. (2017), Hsieh et al. (2017), Wang et al. (2019), Zhang et al. (2019) Kristan et al. (2017), Tijtgat et al. (2017), Du et al. (2019), Wang et al. (2019)
CVPR (conference)	Object detection, face recognition and surveillance	3	Barekatain et al. (2017), Oreifej et al. (2010), Oh et al. (2011)

by the sensors. Remote sensing based on drones provides high-resolution images and videos a low photographic altitude and other data at spatial, spectral and temporal scales compared with satellite and manned aerial remote sensing. Camera calibration, image matching, aerial triangulation, dense reconstruction, image stitching, and multisensor registration are computer vision problems in remote sensing, and these problems have recently been explored in a survey Xiang et al. (2018). The important role of large databases is not only for evaluating traditional methods but also can be useful for applying a new approach, such as deep learning models Elharrouss et al. (2019). However, in recent years, a few works have provided publicly available databases, which requires more effort. To prepare a standard database, we need to follow a series of the rules. Reference Long et al. (2020) discussed the rules to create a standard database for remote sensing applications. Remote sensing databases are as follows.

The International Society for Photogrammetry and Remote Sensing (ISPRS) and EuroSDR presented a database Nex et al. (2015) for image orientation and dense matching. The database provided oblique airborne, UAV-based and terrestrial images captured from Dortmund, Germany, and Zurich, Switzerland. Additionally, terrestrial laser scanning, aerial laser scanning, topographic networks, and GNSS points accompany it as ground truth data. In addition, 3D coordinates on checkpoints (CPs) and cross-sections and residuals on generated point cloud surfaces were presented.

To mosaic images captured by drones, Xu et al. (2016) presented a large database that can also be used for image matching and camera calibration. Images with a resolution of 3680 by 2456 pixels and from flying heights of 558 m, 405 m, and 988 m were captured over Yongzhou, Hechi and HeJiangdong of Hunan Province, China. One of the drones was used Pix4D⁴ with a Panasonic DMC-GF1 camera with a 20 mm focal length lens mounted on it.

In Al Kaff (2017), three database groups were introduced, and some of the state-of-the-art methods for image matching were applied to them. The images were captured by quadcopter drone with a 1270 by 720 resolution from flying heights of 61.1 m, 78.6, and 153.6 m. The images were obtained from both outdoor and indoor scenarios.

The Image Fisheye database Yin et al. (2018), with the aim of camera calibration, evaluating distortion parameter settings, and rectification of images, was created. Additionally, a deep learning method based on an end-to-end multi-contextual collaborative network was presented that estimated the distortion parameter and subsequently removed them from captured images. As recommended in Xiang et al. (2018), the database can be used for evaluating the camera calibration of drones.

To estimate 3D pose, a drone-assistant database synthesis was introduced in Albanis et al. (2020). In the study, the DJI Mavic Enterprise drone was equipped with a HoloLens 2.0 external color camera. The database has both the egocentric view of a cooperative drone and the exocentric view of the user.

Some sample images from the databases are shown in Fig. 6.

Stabilizing and automating flight accurately are the targets of modern drones, which leads to the design of navigation systems at higher levels than previous systems in terms of speed, accuracy, and autonomy with accurate flight stabilization. The main part of an autonomous UAV is the navigation system and its supporting subsystems. The navigation supporting subsystems (pose estimation, obstacle detection, and visual servoing) use

⁴ <https://www.pix4d.com/>.

data captured by various sensors and integrate the data for the navigation system. One of the important tasks in the system is estimating the pose of the drone in terms of positions (x, y, z) and orientations (u, v, w) , and the rest of the tasks, such as detecting obstacles and tracking targets statically and dynamically, are handled by other subsystems, which are finally integrated. Today, due to the increase in sensors based on vision and the improvement of computer vision methods, companies tend to design and produce drone navigation systems using cameras and analyze their data Al-Kaff et al. (2018). Additionally, in the system, three subsystems of pose estimation, obstacle detection, and visual servoing should be redesigned based on computer vision methods. In the navigation systems group, two survey papers presented in 2017 Kanellakis and Nikolakopoulos (2017) and 2018 Al-Kaff et al. (2018) did not explore any databases, and therefore, we introduce the databases of the group in this section.

The Video Verification of Identity (VIVID) database Collins et al. (2005) includes images captured on a runway with changing drone flight heights in both visible and thermal IR imagery for the aim of tracking a vehicle. In addition, they provided ground truth images for tracking aim and a website for interacting with the researcher for testing new methods. Original videos are in AVI format, and their frames are presented separately.

The database presented in Zimmermann et al. (2009) is not based on images captured by a drone; however, many researchers use this database for pose estimation and tracking based on drones Luna (2013). The database includes three objects (MOUSEPAD (MP), TOWEL, and PHONE) and position them in each frame and were labeled as their ground truth.

Reference Pestana et al. (2013) presents a database that follows navigation purposes, and many researchers have used it. To collect the database, AR Drone 2.0 is used in an unstructured condition at flying heights ranging from 1 to 2 m and from 10 to 15 m. The database is useful for training state-of-the-art and deep learning network methods.

Reference Tian et al. (2016) presented a database for adjusting the brightness of two matched images and can also be used for other image processing steps. The study area was the northwestern part of the Sichuan Basin, China, which was captured by a drone at a height of 400 m and a speed of 50 km/s equipped with a nonmeasurement array charge-coupled device (CCD) camera with a resolution of 0.3 m.

Reference Robicquet et al. (2016) introduced a database for navigation aims, such as multitarget tracking and predicting the trajectories. The Stanford Drone Dataset (SDD) includes images and videos recorded by a quadcopter drone (a 3DR solo) equipped with a 4K camera at a flying height of 8 m over intersections of the Stanford University campus with a resolution of 1400 by 1904 pixels. Additionally, due to providing comprehensive ground truth, the database is sufficient to test deep learning methods Wang et al. (2018).

In Rozantsev et al. (2017) to evaluate navigation problems such as obstacle detection, two databases were created, one of which was based on drones. A drone equipped with a camera flew in various weather conditions and flying heights recorded environment at a resolution of 752 by 480 pixels. They evaluated their approach, a convolutional neural network (CNN), on the databases. Additionally, because the paper used a CNN, they provided patches of images with the original sizes.

The UAV mosaicking and change detection (UMCD) database Avola et al. (2018) includes images and videos at low altitude for mosaicking and detecting altitude changes.

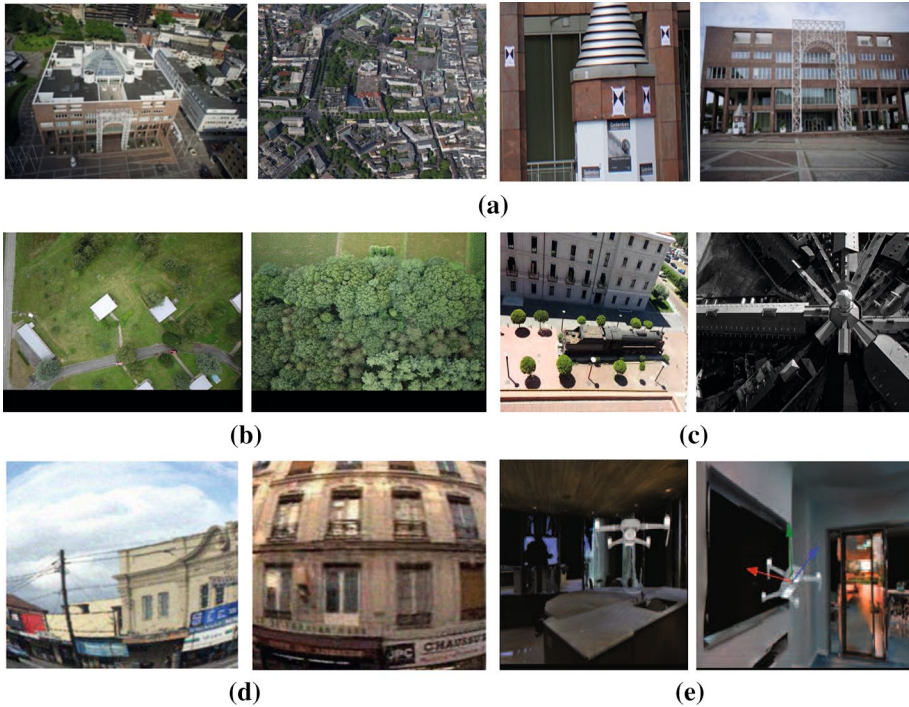


Fig. 6 Samples of the database images related to remote sensing methods: **a** Nex et al. (2015), **b** Xu et al. (2016), **c** Al Kaff (2017), **d** Yin et al. (2018), and **e** Albanis et al. (2020)

Compared with other aerial databases^{5,6} that have many goals, this database focused on these two goals. Images recorded by drone of the National Marine Electronics Association (NMEA⁷) are from flying heights from 6 m to 15 m with speeds from 2 m/s to 12 m/s with spatial resolutions ranging from 720 by 540 (4:3, standard definition) up to 1920 by 1080 (16:9, high definition) pixels per frame.

In Bharati et al. (2018), to detect obstacles and track moving objects by a drone with a forward-looking camera, a database and a method based on a kernelized correlation filter (KCF) framework tested on it were presented with variations in the scale, axial and planar rotation, partial occlusion, illumination variation, and camera instability.

Reference Chen and Lee (2018) presented the National Campus Taiwan University (NCTU) campus database to detect obstacles such as pedestrians, cars, trees, leaves, trunks, trucks, poles and buses by autonomous drone flight. The drone used in the research was a small quadrotor equipped with a Pixhawk⁸ controlling system and an Nvidia TX2 embedded system suitable for applying deep learning methods. Additionally, they applied a deep learning network (UAVNet) on patches extracted from the database.

⁵ <http://eros.usgs.gov/aerial-photography>.

⁶ <http://sipi.usc.edu/database/database.php?volume=aerials>.

⁷ <http://www.gpsinformation.org/dale/nmea.htm>.

⁸ <https://pixhawk.org>.

In Loquercio et al. (2018), a database (original images and patch-based images) and a deep learning method (DroNet) were presented to an autonomous flight system. DroNet is a deep learning method based on CNNs with the aim of flying drones over the streets of a city. A forward-looking camera was mounted on Parrot Bebop 2.0 drone⁹ with flying heights ranging from 5 to 30 m. Additionally, in Palossi et al. (2019), the database was developed by images captured by a COTS Crazyflie 2.0 nano quadrotor¹⁰.

Reference Müller et al. (2018) presented a simulator (Sim4CV) along with a related database to cover many applications in the computer vision community that is suitable for autonomous drone flights and moving objects. They collected images and videos from two drones with speeds of 4 m/s, 6 m/s, and 8 m/s equipped with stabilized cameras. Additionally, the Sim4CV project presented a deep learning method for the aims above and displays comprehensive information on their website.

Reference Mantegazza et al. (2018) introduced a database for autonomous flight and moving object detection. The quadrotor drone captured images at an altitude of 1–2 m for 45 min. They applied some state-of-the-art machine learning methods and deep learning networks and compared them with the database.

A synthetic 3D database obtained by flying a drone in suburban and urban areas at high speeds was presented in Marcu et al. (2018). The database can be used to estimate depth and safe landings and to test deep learning methods in an environment with obstacles. They provided additional data such as RGB, depth and safe-landing information from Google Earth.

In Kang et al. (2019), a database based on images captured by a Crazyflie 2.0 nano drone equipped with a 3.4-g monocular camera at an altitude of 40 cm and a speed of 30 cm/s to meet autonomous flight challenges was presented. The images were collected at Cory Hall at UC Berkeley and to evaluate them, a deep reinforcement learning method was tested on it. However, the research was designed for indoor scenarios, but the database can be used for outdoor scenarios.

The benchmarking database Backes et al. (2019) was designed for flood mapping and modeling for images captured by drones. A Pix4D drone was used at flying heights of 50 m and 60 m to take high-resolution images for creating a 3D mapping. Designing accurate models can help people affected by floods, especially in urban areas.

The database presented in Karaduman et al. (2019) can be useful in patrolling and tracking challenges by detecting the drone route. The drone speed and flight altitude were 50 km/s and 100 m, respectively. To see the achieved results of the method presented in the paper, readers can refer to the supplementary material¹¹ of the paper.

Some image samples from the databases are shown in Fig. 7. Also, the databases presented in the navigation and remote sensing groups are summarized in Table 3.

⁹ <https://www.parrot.com/us/drones/parrot-bebop-2>.

¹⁰ <https://www.bitcraze.io/crazyflie-2/>.

¹¹ <https://link.springer.com/article/10.1007%2Fs10846-018-0954-x>.

4 Applications of images and videos captured by drones

The section is dedicated to using images and videos captured by drones in various applications, such as surveillance, agriculture and forestry, animal detection, disaster detection, and face recognition as shown in categories of Fig. 2. In the group, we present new methods based on databases in each subgroup.

4.1 Surveillance

One of the important applications of drones is surveillance. We divided this application into traffic, crowd, and object detection.

4.1.1 Traffic and crowd detection

The significant increase in the number of vehicles in urban areas and on roadways has led transportation managers to propose new capabilities and systems for the surveillance of traffic and the issues related to it. One of the systems for this aim is use of drones and devices mounted on them, in contrast to the traditional technologies such as inductive loop detectors. The use of drones not only can increase the mobility and coverage domain but also the cost of its operation is significantly lower than that of manned aerial vehicles (MAVs). In Kanistras et al. (2015), Puri (2005) two surveys of drone-based systems for traffic monitoring and management are presented. In the following section, we explore databases created for traffic issues.

The VIRAT video database Oh et al. (2011) includes videos of both humans and vehicles in single-object and two-object categories of with annotated details. The database was collected by a camera used on a drone with an aerial video resolution of 640 by 480 pixels in natural scenes with people performing normal actions in standard contexts, with uncontrolled, cluttered backgrounds. Therefore, the database can be used in continuous visual event recognition (CVER) in which an event can be recognized.

Reference Liu and Mattyus (2015) presented aerial images based on a drone over Munich, Germany, equipped with a German Aerospace Center (DLR) 3K camera system with a resolution of 5616 by 3744 pixels at a flying height of 1000 m. The database is suitable for detecting vehicles in multiclass and multidirectional scenarios. They applied a method based on a fast binary detector using integral channel features in a soft cascade structure.

A video database based on a car parking with the aim of privacy inspection and in three categories (normal, suspicious, and illicit behaviors) was presented in Bonetto et al. (2015). A DJI Phantom 2 Vision+ mini-drone was used to collect videos with a mounted camera with full HD resolution. The database was manually annotated for persons and vehicles in each scene using the ViPER-GT tool¹² in XML format. Additionally, a method using privacy filters was applied to evaluate the database goal.

To collect images in Xu et al. (2016), a quadcopter (Phantom 2) with a GoPro Hero Black Edition 3 camera (resolution of 1920 by 1080) was used. Some scenarios were considered for different weather conditions, locations, time and flight altitudes (refer to Table 1, page 12 in Xu et al. (2016)). Additionally, for traffic monitoring, a method based

¹² <http://vipер-toolkit.sourceforge.net/>.

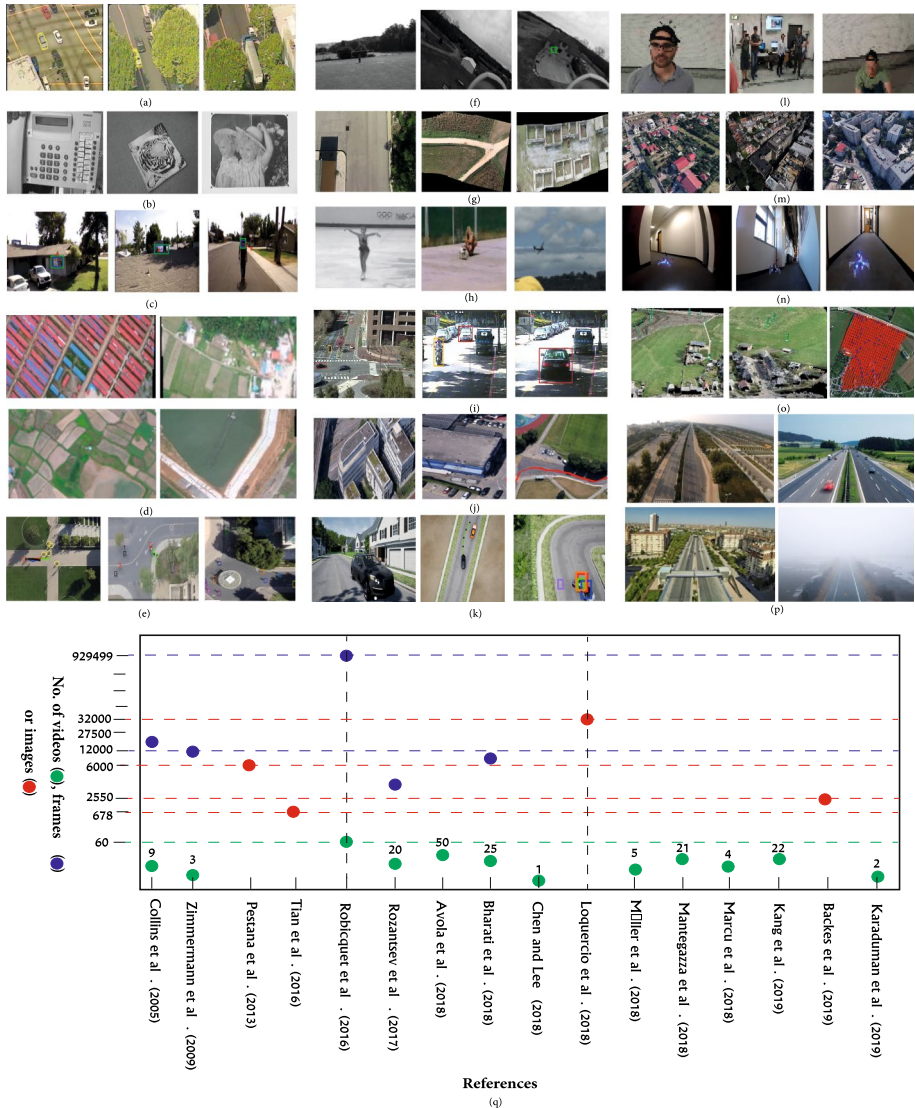


Fig. 7 Samples of the databases images related to navigation methods: **a** Collins et al. (2005), **b** Zimmermann et al. (2009), **c** Pestana et al. (2013), **d** Tian et al. (2016), **e** Robicquet et al. (2016), **f** Rozantsev et al. (2017), **g** Avola et al. (2018), **h** Bharati et al. (2018), **i** Chen and Lee (2018), **j** Loquerio et al. (2018), **k** Müller et al. (2018), **l** Mantegazza et al. (2018), **m** Marcu et al. (2018), **n** Kang et al. (2019), **o** Backes et al. (2019), **p** Karaduman et al. (2019), and **q** a chart of the number of images, frames, and videos in terms of the databases (the order is corresponding to Table 3)

on the Viola-Jones (V-J) and linear support vector machine (SVM) classifier with HOG features (HOG + SVM) was proposed.

In Najiya and Archana (2018) a method to detect vehicles along with the amount, speed, and densities of bidirectional flow based on enhanced videos, a Kanade–Lucas–Tomasi (KLT) tracker, a SVM, and connected graphs for traffic surveillance was presented. The

Table 3 List of the databases used by navigation and remote sensing groups

Database	Category	Description	Year	Availability	Technique used
Nex et al. (2015)	Remote sensing (Image matching)	oblique images: 1260 images, terrestrial images: 4310 images	2015	Public ^a	Song et al. (2019), Alidoost and Arefi (2015)
Xu et al. (2016)	Remote sensing (Image matching)	591 images	2016	Upon request ^b	Tian et al. (2018)
Al Kaff (2017)	Remote sensing (Image matching)	27,500 high resolution images	2017	Private	Hussein (2018), Al-Kaff et al. (2017)
Yin et al. (2018)	Remote sensing (Camera calibration)	2550 source images, 10 samples with various distortion parameter settings	2018	upon request ^c	–
Albanis et al. (2020)	Remote sensing (Pose estimation)	Number of trajectories: 90, environments: 90	2020	Public ^d	–
Collins et al. (2005)	Navigation (tracking)	9 videos, 16284 frames	2005	Public ^e	Abughaliieh et al. (2018), Askar et al. (2017)
Zimmermann et al. (2009)	Navigation (Pose estimation)	3 videos, 12000 frame	2009	Public ^f	Luna (2013), Pestana et al. (2013)
Pestana et al. (2013)	Navigation (object tracking)	6k images	2013	Public ^g	Luna (2013), Pestana et al. (2014), Pestana Puerta (2017), Carrio et al. (2018)
Tian et al. (2016)	Navigation	678 images	2016	Private: Email for request ^h	–
Robicquet et al. (2016)	Navigation	60 videos, 929499 frame	2016	Public ⁱ	Wang et al. (2018, 2019), Hu et al. (2019), Robicquet et al. (2016), Ballan et al. (2016)
Rozantsev et al. (2017)	Navigation	20 videos, 4000 frames	2017	Public ^j	Rozantsev (2017)
Avola et al. (2018)	Navigation	50 challenging aerial videos with and without the presence of vehicles, persons, and objects, plus metadata and telemetry	2018	Public ^k	Avola et al. (2018)
Bharati et al. (2018)	Navigation (Obstacle detection)	25 challenging videos, 6584 frames	2018	Public ^l	–
Chen and Lee (2018)	Navigation (Obstacle detection)	One video, 8 different kinds of obstacles	2018	Public ^m	–
Loquerio et al. (2018)	Navigation (obstacle detection)	32,000 images	2018	Public ⁿ	Palossi et al. (2019)
Müller et al. (2018)	Navigation	5 videos	2018	Public ^o	Müller et al. (2018), Müller et al. (2017, 2019)
Mantegazza et al. (2018)	Navigation	21 different videos	2018	Public ^p	Mantegazza et al. (2019)
Marcu et al. (2018)	Navigation	4 videos, 11907 samples	2018	Public ^q	–
Kang et al. (2019)	Navigation	22 videos	2019	Public ^r	–

Table 3 (continued)

Database	Category	Description	Year	Availability	Technique used
Backes et al. (2019)	Navigation	2200 images	2019	Upon request ⁸	–
Karaduman et al. (2019)	Navigation (Patrolling)	2 videos	2019	Upon request ⁴	–

^a Available at http://www2.isprs.org/commissions/comm1/cwgl15/benchmark_main.html

^b Email: oujl@hnu.edu.cn

^c Email: yinxiaqing89@gmail.com

^d Available at <https://vc13d.github.io/DronePose>

^e Available at <http://vision.cse.psu.edu/data/vividEval/main.html>

^f Available at <http://cmp.felk.cvut.cz/cmp/demos/Tracking/linTrack/>

^g Available at <http://vision4uav.com/>

^h duanfuzhou@cnu.edu.cn

ⁱ Available at http://cvgl.stanford.edu/projects/uav_data/

^j Available at <https://drive.switch.ch/index.php/s/3b3bdb6f8fb61e05d8b0560667ea992>

^k Available at <http://www.umcd-database.net/>

^l Available at <http://www.itc.ku.edu/cviiu/tracking.html>

^m Available at <https://www.csie.ntu.edu.tw/~r01944012/newcroptool.html>

ⁿ Available at <http://rpg.ifi.uzh.ch/dronet.html>

^o Available at <http://www.sim4cv.org>

^p Available at <https://github.com/idsia-robotics/proximity-quadrotor-learning>

^q Available at <https://sites.google.com/site/aerialimageunderstanding/safeuav-learning-to-estimate-depth-and-safe-landing-areas-for-uavs>

^r Available at github.com/gkahl13/GtS

⁸ Email: dietmar.backes@uni.lu

⁴ Email: he.edu.tr@gmail.com

method was applied to the presented database by using drones to collect the videos with a resolution of 336 by 596.

Reference Kyrkou et al. (2018) presented different models using deep learning methods for traffic monitoring on a database created by drones in different illumination, viewpoint, and occlusion conditions. Since the speed of transmitting and processing data from drone to GCS is vital, they designed a light CNN for the aim that is compared with other deep networks.

Reference Ke et al. (2018) introduced a database for traffic surveillance by drones over different roadway segments with an orthographic camera, and images at a resolution of 60 by 40 were captured. Additionally, they proposed a method based on deep learning to address irregular ego-motion, low estimation accuracy in a dense traffic situation, and high computational complexity. It should be noted that the database was an updated version from that developed in Ke et al. (2017).

The database created by original images and patches presented in Zhu et al. (2018), is suitable for detecting, counting, and tracking vehicles and location and type (car, bus, or truck) recognition. The database used a Zenmuse X3 camera at a 3840 by 2178 resolution mounted on an Inspire 1 Pro quadcopter in sunny and cloudy weather. Since the authors provided images based on (512 by 512) patches, deep learning methods can be applied to the database.

Crowd detection is one of the challenging problems in surveillance and behavioral analysis that attracts researchers in drone fields to address it. Upright views, detecting boundaries of crowds in places such as sport stadiums, drone locations and flight altitudes, and moving objects have been explored in images based on drones Minaeian et al. (2015). In the following section, we explore databases created for crowd issues.

In Tzelepi and Tefas (2017), a video and image drone database was created based on videos collected from the YouTube¹³, senseFly-Example-drone¹⁴, as well as the UAV123¹⁵ databases. The database is for detecting human crowds for applications in which crowd and non-crowd scenes should be classified. To solve this problem, the authors proposed a deep learning method. Additionally, patch-based images are publicly prepared for studies that use deep learning approaches.

The database presented in Al-Sheary and Almagbile (2017) includes three subgroups of images. The first group¹⁶ of images was collected via a low-altitude Pix4D drone with a Canon camera¹⁷ over Leftous. The second group was images download from the internet, while the third group was images captured over Mecca. To evaluate the database, the authors tested a segmentation method to extract the crowd.

In Almagbile (2019), images of different orientations and positions with resolutions of 691 by 1359, 683 by 471, and 689 by 1366 pixels were captured to detect and count people. The authors tested a method that uses features from accelerated segment test (FAST) and filters for extracting crowd features.

A drone-based vehicle re-identification (ReID) database was presented in Wang et al. (2019). Two DJI Phantom4 drones captured vehicles in different locations, with diverse

¹³ Available at <http://www.youtube.com/>.

¹⁴ Available at <https://www.sensefly.com/drones/example-datasets.html>.

¹⁵ Available at <https://ivul.kaust.edu.sa/Pages/Dataset-UAV123.aspx>.

¹⁶ <https://www.sensefly.com/drones/ebee.html>.

¹⁷ CanonDIGITALIXUS120IS_5.0_3000x4000.

view-angles and flight-altitudes. In addition, a deep learning method was tested for vehicle ReID.

To explore the congested urban environment in traffic monitoring, in Barmounakis and Geroliminis (2020), a new database (pNEUMA) was presented. The images were captured by 10 consumer quadcopter DJI drones equipped with a camera with a resolution of 4096×2160 pixels. The study place was included a total of 10 Km area, 10 km road network, low, medium, and high-volume arterial, more than 100 intersections, and more than 30 bus stops.

Reference Chen et al. (2020) extracted vehicle trajectory based on images recorded by a DJI Mavic professional drone. Images were collected with a resolution of 3840×2160 pixels at an altitude of 223 m and 281 m. Both free flow and congested scenarios were considered in the database. Three procedures of the region of interest (ROI), the kernelized correlation filter (KCF), and transforming positions from the Cartesian coordinate in the video to the Frenet coordinate applied on the database.

DroneVehicle database Zhu et al. (2020a) was created by the Lab of Machine Learning and Data Mining, Tianjin University, China. The database was recorded by both RGB and infrared cameras mounted on a drone. The database covers scenarios such as urban roads, residential areas, parking lots, highways, objects such as car, bus, truck, van, and density such as sparse and crowded scene.

To detect and segment vehicles based on images captured by drone in Zhang et al. (2020), a database was presented. A DJI matrice 200 quadcopter equipped with a zenmuse X5S gimbal and camera collected the images with resolution from a range of 960×540 pixels to 5280×2970 pixels. A Multi-Scale and Occlusion Aware Network (MSOA-Net) with two parts of Multi-Scale Feature Adaptive Fusion Network (MSFAF-Net) and Regional Attention-based Triple Head Network (RATH-Net) was tested on the database.

Several deep learning methods were conducted on a database based on drone images presented in Lyu et al. (2020). The database is suitable for semantic segmentation in complex urban scenes for applications such as robotics and autonomous driving. The image resolution was considered in 4096×2160 pixels and 3840×2160 pixels.

Some image samples from the databases are shown in Fig. 8. Also, the databases presented in the traffic and crowd tasks are summarized in Table 4.

4.1.2 Object detection

Object detection (segmenting scenes to certain classes such as humans, buildings, or cars) is a basic step in computer vision that covers different areas in the field, such as image retrieval and video surveillance. In the following section, we explore drone-based object detection databases.

In Saif et al. (2014), a dynamic motion model (DMM) was applied to the UAV video database¹⁸ (actions1.mpg and actions2.mpg) from the Center for Research in Computer Vision (CRCV) at the University of Central Florida, while in Maria et al. (2016), a database based on YouTube videos was collected to detect cars in a scene.

The UAV123 database Mueller et al. (2016) introduced 123 videos captured by drones at low altitudes for tracking issues and a simulator to evaluate moving targets in a real-time state. Attributes such as the aspect ratio change, full and partial occlusion, low resolution, illumination variation, fast motion, and camera motion were provided for researchers.

¹⁸ Available at https://www.crcv.ucf.edu/data/UCF_Aerial_Action.php.

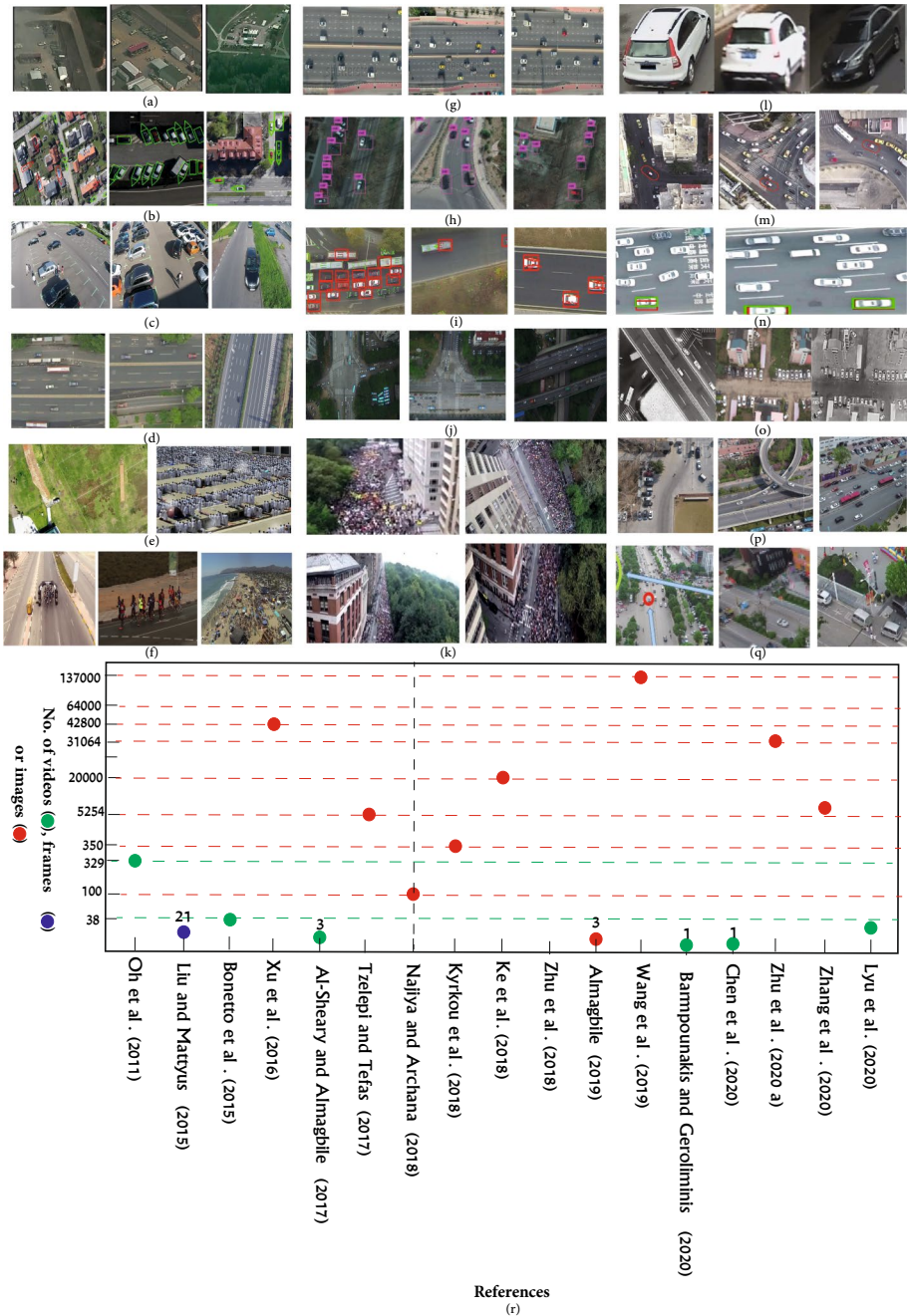


Fig. 8 Samples of the database images related to traffic and crowd methods: **a** Oh et al. (2011), **b** Liu and Matyus (2015), **c** Bonetto et al. (2015), **d** Xu et al. (2016), **e** Al-Sheary and Almagbile (2017), **f** Tzelepi and Tefas (2017), **g** Najiya and Archana (2018), **h** Kyrkou et al. (2018), **i** Ke et al. (2018), **j** Zhu et al. (2018), **k** Almagbile (2019), **l** Wang et al. (2019), **m** Barpounakis and Geroliminis (2020), **n** Chen et al. (2020), **o** Zhu et al. (2020a), **p** Zhang et al. (2020), **q** Lyu et al. (2020), and **r** chart of the number of images, frames, and videos of the databases (the order is corresponding to Table 4)

Table 4 List of the databases used in traffic and crowd tasks

Database	Year	Availability	Description	Technique used
Oh et al. (2011)	2011	Public ^a	Surveillance: 23 event types distributed throughout 29 hours of video (329 videos)	Chen et al. (2017), Vega et al. (2015), Perera et al. (2018), Lee (2016)
Liu and Mattyus (2015)	2015	Public ^b	Traffic: 21 MPixel original frame images	Tayara et al. (2018), Azimi et al. (2018), Yang et al. (2019), Majid Azimi (2018), Sommer et al. (2018)
Bonetto et al. (2015)	2015	Public ^c	Privacy protection and anomaly detection 38 different contents (videos) captured in full HD resolution	Ruchaud (2015), Henrio and Nakashima (2018)
Xu et al. (2016)	2016	Private	Vehicle detection: over 42800 images	–
Al-Sheary and Almagbile (2017)	2017	Private	Crowd monitoring system: collected from 3 databases	–
Tzelepi and Tefas (2017)	2017	Public ^d	Crowd detection: 5254 images, labeled for deep learning methods	Tzelepi and Tefas (2019)
Najjya and Archana (2018)	2018	Private	Traffic surveillance: 100 images above a freeway segment	–
Kyrkou et al. (2018)	2018	Private	Vehicle detection: 350 images	Plastiras et al. (2018a, 2018b)
Ke et al. (2018)	2018	Public ^e	Traffic: 20,000 images	–
Zhu et al. (2018)	2018	upon request ^f	Urban traffic density estimation: 64,000 images	–
Almagbile (2019)	2019	Private	Estimation of crowd density: 3 images	–
Wang et al. (2019)	2019	Private	Vehicle Re-identification: 1,37,000 images	–
Barpounakis and Geroliminis (2020)	2020	Public ^g	Urban traffic monitoring: 1 video (59 h)	–
Chen et al. (2020)	2020	Public ^h	Vehicle trajectory: 1 video	–
Zhu et al. (2020a)	2020	Public ⁱ	Vehicle detection: 31,064 images	–
Zhang et al. (2020)	2020	Public ^j	Vehicle detection: 5874 images	–
Lyu et al. (2020)	2020	Public ^k	Semantic segmentation in traffic: 30 videos	–

Table 4 (continued)

- ^a Available at www.viratdata.org
- ^b Available at https://www.dlr.de/eoc/desktopdefault.aspx/tabid-5431/9230_read-42467/
- ^c Available at <http://mmsg.epfl.ch/mini-drone>
- ^d Available at: <https://github.com/mitzelepi/GraphEmbeddedCNN>
- ^e Available at <http://www.uwstarlab.org/research.html>
- ^f Email: zhujiason@gmail.com
- ^g Available at open-traffic.epfl.ch
- ^h Available at <https://seutrafic.com>
- ⁱ Available at <https://github.com/VisDrone/DroneVehicle>
- ^j Available at <https://github.com/liuchunsense/UVSD>
- ^k Available at <https://uavid.nl/>

The Okutama-Action database and its annotations Barekatin et al. (2017) for concurrent human action detection present some challenging issues, such as a non-static camera with abrupt motion, the dynamic transition of actions, multiple concurrent actions and multi-labeled actors. The database was recorded by two drones with cameras with a resolution of 4 K and 45 or 90 degrees at flying heights of 10–45 m.

In Cehovin Zajc et al. (2017), the presented 360-degree videos can be used in active-camera robotics applications such as circling over a target object. The videos captured by a drone with a Ricoh Theta 360-degree camera over objects of different sizes have also annotated frames.

A car parking database (CARPK) was presented in Hsieh et al. (2017) in which challenges for object counting in parking lots were considered. A Phantom 3 professional drone at a flying height of 40 m recorded high-resolution videos. Additionally, the authors tested an object-counting method on the database based on layout proposal networks (LPNs) and spatial kernels.

The UAVDT benchmark Du et al. (2018) is a database for object detection and tracking that includes high-density, small-object camera motion, and real-time challenges with attributes such as different weather conditions (daylight, night and fog), flying altitudes (10–30 m), and different camera views (front view, side view and bird's-eye view). The authors claimed that state-of-the-arts methods used in the database achieved a disappointing result because of new challenges presented in the database. Also, a developed version of the database was presented in Yu et al. (2020).

UG² Vidal et al. (2018) includes uncontrolled videos at resolutions of 600 by 400 to 3840 by 2026 recorded by the drone that collected videos from the YouTube website. The database provided challenges related to glare, lens flare, low image quality, camera shaking, and images converted to patches for testing deep learning methods.

Mivia, a research laboratory of the University of Salerno, presented the Mivia database Carletti et al. (2018) for multiobject tracking. The database was a collection of DJI F-450 drone videos that the mounted Nilox F60 camera recorded with variable altitudes, speeds, and angles (yaw and pitch). Additionally, they proposed a method based on local data association with a backward chain for multiobject tracking.

Reference Zhu et al. (2018) was a report of the Vision Meets Drone 2018 challenge workshop in conjunction with the 15th European Conference on Computer Vision (ECCV 2018) that also presented a database for the Vision Meets Drone Video Detection and Tracking (VisDrone-VDT2018) challenge. Additionally, the database was developed in Zhu et al. (2018), Du et al. (2019), Zhu et al. (2020b, 2020c).

Reference Xu et al. (2018) describes the database downloaded from the DJI website¹⁹, which is videos captured by various types of drones and cameras and suitable for a lower power object detection challenge (LPODC). Additionally, the paper reported the results of the System Design Contest (SDC) in conjunction with the 55th Design Automation Conference (DAC) in 2018.

The Urban Drone dataset (UDD) Chen et al. (2018) includes images over Beijing, Huludao, Zhengzhou, and Cangzhou (China) collected from images of a DJI Phantom 4 drone at flying heights of 60–100 m with resolutions of 4 K (4096 by 2160 pixels) and 12 M (4000 by 3000 pixels). Additionally, the images can be fed into deep learning networks.

¹⁹ <https://www.dji.com/>.

Finally, the UAVP100 database Wang et al. (2019) was considered for tracking people (online single-person tracking (OSPT)) by DJI Phantom 4, Inspire 2 and Spark drones with flying heights of 5–30 m and a cameras resolution of 1920 by 1080 pixels. The challenges explored when collecting the database are similar to those for UAV123.

Reference Qi et al. (2019) presented a database based on other database images and their own drone. The aim of the database was to detect and track objects. Several challenges such as parking lots, street views, social parties, traveling were explored in the study.

Virtual Aerial Image Dataset (VALID) Chen et al. (2020) is a virtual database, which can be considered as images captured from drones. The authors presented a comprehensive ground truth that is suitable for image segmentation on 30 categories in 6 different virtual scenes and 5 various ambient conditions (sunny, dusk, night, snow, and fog).

ERA (Event Recognition in Aerial videos) database was presented in Mou et al. (2020). The database was collected for recognizing events from images recorded by drones on YouTube. Several deep learning methods were tested on the database. The database covered 25 classes related to different events such as traffic congestion, harvesting, ploughing, constructing, police chase, conflict, baseball, basketball, boating.

Reference Mandal et al. (2020) introduced a moving object recognition (MOR) database based on videos recorded by drones. The videos were captured on highways, flyovers, traffic intersections, urban areas, and agricultural regions. The range of image resolutions was from 1280×720 pixels to 1920×1080 pixels. In addition, a deep learning method was tested on the database.

EyeTrackUAV2 database Perrin et al. (2020) is useful to explore saliency researches related to drones. The EyeLink 1000 Plus eye-tracking system²⁰ was used to conduct the experiment and create gaze information. Image resolution was considered from 1280×720 pixels or 720×480 pixels. Additionally, the database is suitable to test a deep learning approach.

Some image samples from the databases are shown in Fig. 9. Also, the databases presented in the object detection are summarized in Table 5.

4.2 Agriculture and forestry

Today, compared with satellite imagery, there is a growing interest in using drones to present effective solutions in autonomous applications, such as inspections of the state of farming. From the viewpoint of farmers, drones can provide a bird's-eye view over their fields to assess and lead to a precise monitoring system for crop and water statuses and biomass estimation Adão et al. (2017).

Reference Zarco-Tejada et al. (2014) used the database presented by the Institute for Sustainable Agriculture (IAS) of the Spanish Council for Scientific Research (CSIC). The database was obtained by consumer-grade cameras at a resolution of 4000 by 3000 and at a flying height of 200 m for tree height estimation.

In Turner et al. (2014), a collection of ultrahigh-resolution visible, multispectral and thermal images were captured by three sensors. A Canon 550D digital single-lens reflex (DSLR) camera (resolution of 5184 by 3456 pixels), a FLIR Photon 320²¹ uncooled

²⁰ Available at <https://www.sr-research.com/eyelink-1000-plus/>.

²¹ <https://www.flir.com>.

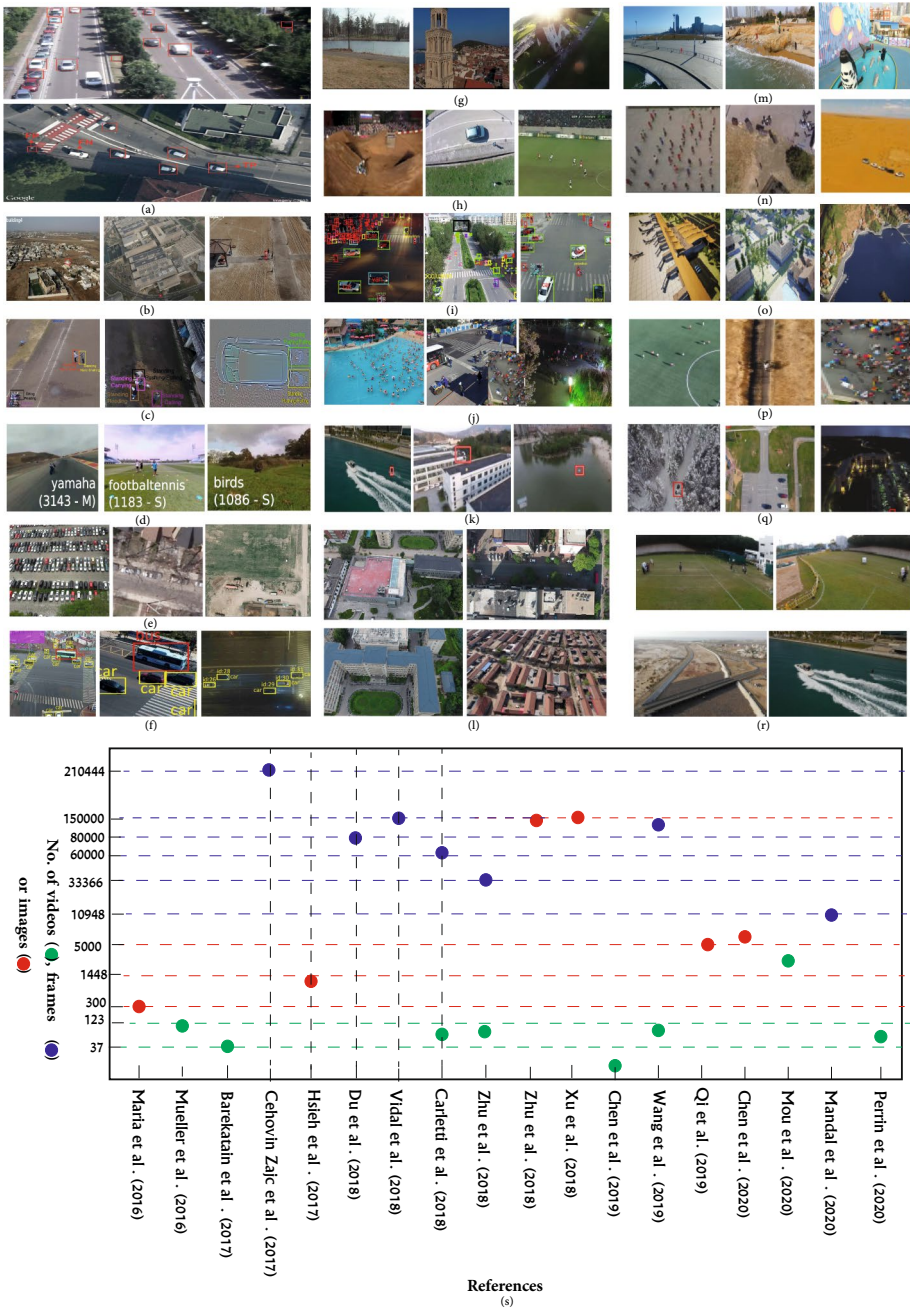


Fig. 9 Samples of the database images related to object detection methods: **a** Maria et al. (2016), **b** Mueller et al. (2016), **c** Barekatin et al. (2017), **d** Cehovin Zajc et al. (2017), **e** Hsieh et al. (2017), **f** Du et al. (2018), **g** Vidal et al. (2018), **h** Carletti et al. (2018), **i** Zhu et al. (2018), **j** Zhu et al. (2018), **k** Xu et al. (2018), **l** Chen et al. (2018), **m** Wang et al. (2019), **n** Qi et al. (2019), **o** Chen et al. (2020), **p** Mou et al. (2020), **q** Mandal et al. (2020), **r** Perrin et al. (2020), and **s** a chart of the number of images, frames, and videos in terms of the databases

Table 5 List of databases used in object detection

Database	Year	Availability	Description	Technique used
Maria et al. (2016)	2016	Private	Car detection: 300 images	–
Mueller et al. (2016)	2016	Public ^a	Object tracking: 14 state-of-the-art trackers 123 new and fully annotated HD video sequences	Mueller et al. (2016), Tijgat et al. (2017), Xue et al. (2018), Carletti et al. (2018), Berker Logoglu et al. (2017), Fan and Ling (2019), Cavaliere et al. (2019), Micheal and Vani (2019), Xue et al. (2018), Dinh et al. (2019), Li et al. (2018), Kuai et al. (2018), Wang et al. (2018), Touil et al. (2019), Xiaoyuan et al. (2019)
Barekatin et al. (2017)	2017	Public ^b	Human action detection: 37 videos, 43 minute-long fully-annotated sequences with 12 action classes	Murray (2017), Wang et al. (2018, 2019), Soleimani et al. (2018)
Cehovin Zajc et al. (2017)	2017	Upon request ^c	Visual object tracking: 17 tracker, 210,444 frame	Lukežič et al. (2019)
Hsieh et al. (2017)	2017	Public ^d	Object counting: 1448 images, 90,000 cars captured from different parking lots	Majid Azimi (2018), Li et al. (2019), Stahl et al. (2019)
Du et al. (2018)	2018	Public ^e	Object detection and tracking: 10 h raw videos, 80000 representative frames	Long et al. (2019), Perreault et al. (2019), Zhang et al. (2020)
Vidal et al. (2018)	2018	Public ^f	Object detection: over 150,000 annotated frames	VidalMata et al. (2019)
Carletti et al. (2018)	2018	Public ^g	Multi-object tracking: 53 videos and more than 60,000 frames	Carletti et al. (2019)
Zhu et al. (2018)	2018	Public ^h	Object detection: 79 video clips with about 1.5 million annotated bounding boxes in 33,366 frames	Zhang et al. (2019), Wei and Duan (2020), Bochinski et al. (2018), Tang et al. (2020), Liu et al. (2020), Wang et al. (2019)
Zhu et al. (2018)	2018	Public ⁱ	Object detection: 2.5 million annotated instances in 179,264 images/video frames	Zhang et al. (2018), Li et al. (2019), Vaddi et al. (2019)
Xu et al. (2018)	2018	Upon request ^j	Object detection: 95 categories and 150 k images	Wang et al. (2019), Hao (2019)
Chen et al. (2018)	2018	Public ^k	Object detection: 10 video sequences taken in 4 different cities in China	–
Wang et al. (2019)	2019	Upon request ^l	100 fully annotated aerial videos with nearly 130 K frames and 11 challenging factors	–
Qi et al. (2019)	2019	Upon request ^m	Object detection: 5000 images	–
Chen et al. (2020)	2020	Public ⁿ	Object detection: 6690 images	–
Mou et al. (2020)	2020	Public ^o	Event recognition: 2864 videos	–

Table 5 (continued)

Database	Year	Availability	Description	Technique used
Mandal et al. (2020)	2020	Upon request ^p	Object recognition: 10,948 frames	–
Perrin et al. (2020)	2020	Public ^q	Eye tracking: 43 videos	–

^a<https://ivul.kaust.edu.sa/Pages/pub-benchmark-simulator-uav.aspx>

^b<https://okutama-action.org/>

^cluka.cehovin@fri.uni-lj.si

^d<https://lafi.github.io/LPN/>

^e<https://sites.google.com/site/daviddo0323/>

^f<http://www.ug2challenge.org/>

^g<https://mivia.unisa.it/>

^h<http://www.aiskyeye.com/>

ⁱ<http://aiskyeye.com/>

^jshu@nd.edu

^k<https://github.com/MarcWong/UDD>

^lwdice@dlut.edu.cn

^myunlongqi@my.unt.edu

ⁿ Available at <https://sites.google.com/view/valid-dataset/>

^o Available at <https://lcmou.github.io/ERA-Dataset/>

^pmurarimandal.cv@gmail.com

^q Available at <ftp://dissoctie@ftp.ivic.polytech.univ-nantes.fr/EyeTrackUAV2/>



Fig. 10 Samples of the database images related to agriculture and forest methods: **a** Tripicchio et al. (2015), **b** Escalante et al. (2019), **c** Oppenheim et al. (2017), **d** Kragh et al. (2017), **e** Murugan et al. (2017), **f** Zarco-Tejada et al. (2014), **g** Dandois et al. (2015), **h** Turner et al. (2014), and **i** Wang and Luo (2019)

thermal sensor (resolution of 324 by 256 pixels) and a Tetracam mini-MCA sensor with six channels (resolution of 1280 by 1024), respectively, mounted on an oktokooper drone. It has been demonstrated that drones carrying multiple sensors can be considered to accurately map vegetation canopies.

Reference Tripicchio et al. (2015) describes a collection method for drone videos that can be used for analyzing soil characteristics. The videos captured by an Asus Xtion Pro Sensor collected RGB and depth data. In addition, a new approach to classify plow field by the sensor is studied. Finally, two different metric, re-orientation method based on Principal component analysis (PCA) and Delaunay triangulation method, have been developed for this purpose.

In Dandois et al. (2015), images captured by a Canon ELPH 520 HS digital camera on board a hobbyist, commercial multirotor and ArduCopter drone over a temperate deciduous forest in Maryland, USA. The database is suitable for producing 3D multispectral point clouds at different flying heights.

Reference Oppenheim et al. (2017) presented a database for the detection and counting of yellow tomato flowers in a greenhouse. The images captured by a smartphone's LG-G4 camera and a Canon PowerShot 590IS at resolutions of 5312 by 2988 and 3264 by 1832 mounted on drone with top and front views.

In Kragh et al. (2017), a multimodal database for obstacle detection in agriculture based on a DJI Phantom 4 drone equipped with three sensors. Web, thermal and stereo cameras at resolutions of 1920 by 1080, 640 by 512, and 1024 by 544 pixels, respectively, and at altitudes of 1.5–50 m is studied. The database comprises approximately 2 h of data in a grass-mowing scenario in Denmark.

In Murugan et al. (2017), a multispectral image database for agriculture monitoring in a large farm in Roorkee, Uttarakhand, India, was presented. The drone used was a DJI Phantom at an altitude of 100 m and mounted on it was a high-definition 4K-resolution RGB camera. Additionally, the database can be used for image segmentation based on a multichannel imaging process.

The authors in Escalante et al. (2019), designed and produced a hexacopter drone equipped with six 700-KVA brushless motors and four 40A electronic speed controllers for monitoring barley fields in the state of Nuevo Leon, Mexico. They used a Parrot Sequoia multispectral sensor to capture multispectral images in the five channels of red, green, red-edge, and near-infrared at a resolution of 1.2 Mpx and at a height flying of 24.4 m. Additionally, a deep learning method was applied to the database.

To recognize the bayberry tree, a database of the images was collected by drone in Wang and Luo (2019). The database can be useful to extract the position and crown information of the tree and to estimate yield. The drone used in the study was a DJI Phantom 4 to collect to take the aerial photography in Dayangshan Forest Park, Yongjia county, Zhejiang province with a resolution of 5472×3648 pixels from January 23 to 24, 2019. A deep learning method based on Mask RCNN (Mask Region Convolutional Neural Networks) was tested on the database.

Some image samples from the databases are shown in Fig. 10. Also, the databases presented in agriculture and forest methods are summarized in Table 6.

4.3 Animal detection

One of the other applications based on drones that has recently been increasing is monitoring animals in large areas with the aim of detecting, counting, and tracking. In the following section, we explore drone-based animal detection databases.

A conservation animal database was collected in van Gemert (2014) for the localization and counting animals such as rhinos or elephants. Specially, the database is suitable for

Table 6 List of the databases used in agriculture and forestry

database	Year	Availability	Description	Technique used
Zarco-Tejada et al. (2014)	2014	Email for request ^a	Agriculture: 750 original images acquired for study area 1, and 679 images acquired for study area 2	–
Turner et al. (2014)	2014	private Email for request ^b	696 RGB images, 474 multi-spectral images, and 1503 thermal Infrared images	–
Tripicchio et al. (2015)	2015	Private	Agriculture: a short RGB-D video has been captured for each field storing the rgb and depth data	–
Dandois et al. (2015)	2015	Public ^c	Forest: 1219 images	–
Oppenheim et al. (2017)	2017	Private	Agriculture: detecting tomato flowers	–
Kragh et al. (2017)	2017	Public ^d	Agriculture: obstacle detection in agriculture 2 h of raw sensor data	Korthals et al. (2018)
Murugan et al. (2017)	2017	Upon request ^e	Agriculture: mosaicked, orthorectified, and georegistered images for the two locations	Maurya et al. (2018)
Escalante et al. (2019)	2019	Private	Agriculture: a total of 72 rectangular plots	–
Wang and Luo (2019)	2019	Public ^f	Agriculture: 3690 images	–

^apablo.zarco@csic.es^bDarren.Turner@utas.edu.au^c<http://ecosynth.org/video>^d<https://vision.eng.au.dk/fieldsafe/>^emdeepakitr@gmail.com^f Available at <http://www.geodoi.ac.cn>

animal detection and counting. An Ascending Technologies Pelican quadcopter drone with a GoPro Hero 3 (Black Edition) action camera (resolution of 1920 by 1080 pixels) recorded the videos. In addition, a object recognition method based on the three light-weights was evaluated in terms of the database.

Reference Chamoso et al. (2014) presented a database to detect cattle in an area with a very large number of animals captured by a multirotor drone equipped with a GoPro Hero 5 in full HD auxiliary camera with a resolution of 1080 pixels per inch. The database was evaluated by a CNN architecture for animal detection and counting. Therefore, the database can be used for applying deep learning methods.

A monitoring wildlife database (koala tracking and detection above the canopy) Gonzalez et al. (2016) was created by S800 EVO Hexacopter²² drone over the Sunshine Coast, 57 km north of Brisbane, Queensland, Australia. RGB and thermal images and videos of were obtained by a Mobius RGB camera (resolution 1080) and FLIR thermal camera (resolution 640 by 510), respectively.

A process of data augmentation on the database provided in Okafor et al. (2017) was applied to develop the database for use in animal detection and deep learning approaches. The image was taken by a DJI Phantom 3 drone. To obtain promising and accurate results in deep learning approaches, they applied a data augmentation method on the database. It should be noted that data augmentation is an important step in deep learning methods to increase training data. Additionally, several deep learning methods were evaluated in terms of the database.

For detecting and enumerating marine wildlife over breeding colonies in eastern Canada, a database was collected by a senseFly eBee drone equipped with two sensors: an RGB camera (Canon S110 with image resolution of 12 megapixel) and a thermal infrared camera (senseFly LLC, Thermomapper, with an image resolution of 640 by 512) Seymour et al. (2017). Moreover, a counting animal method based on polygon/convex hull proportion and high-pass filter combination was applied on the database.

Reference Kellenberger et al. (2018) introduced a database for detecting animals over the Kuzikus²³ Wildlife Reserve in eastern Namibia. A Canon PowerShot S110 RGB camera, a multispectral and a thermal sensor with a resolution of 3000 by 4000 was mounted on a single wing of a senseFly 3 eBee. The database was suitable for exploring the challenge of monitoring and covering large areas and for applying deep learning methods.

To detect and count sheep over the Pirinoa region of New Zealand, a database was presented in Sarwar et al. (2018) that was evaluated with deep learning methods. The deep learning method was based on Region-based convolutional neural networks (R-CNNs). The results showed that the R-CNNs obtained great promise for sheep detection and counting compared with CNNs. The database captured by a drone with an image resolution of 20,148 by 1080 pixels at an altitude of 80 m.

Since analyzing the population and migration of marine animals such as stingrays and dolphins is important for biologists, Saqib et al. (2018) presented drone videos with a resolution of 3840 by 2160 pixels for stingrays and a resolution of 4096 by 2160 pixels for dolphins over beaches in Queensland, Australia. A deep learning method based on Faster-R-CNNs was tested on the database. The approach obtained better results than CNNs and R-CNNs.

For the counting, assistance, and management of cattle, a DJI Phantom 4 drone with a flight time of 28 min and image resolution of 4000 by 3000 pixels flew over Kumamoto, Japan Shao et al. (2019). The database captured four sets of normal, truncated, blurred, and occluded images in different weather conditions and areas. In addition, a CNN method was

²² DJI—The World Leader in Camera Drones/Quadcopters for Aerial Photography.

²³ http://kuzikus-namibia.de/xe_index.html.

applied on the database. To achieve more accurate results a three-dimensional model was considered on the images. Therefore, the database is suitable to test deep learning methods.

A database was introduced in Sykora-Bodie et al. (2017) and developed in Gray et al. (2019)^{24, 25} for sea turtle detection during a mass nesting event on the coast of Ostional, Costa Rica. The database was obtained by flying a Canon PowerShot S110 near-infrared (NIR) camera at a flying height of 90 m. Moreover, to increase the quality of the images, a post-processing step based on a threshold function was applied on the database.

Reference Rahnemoonfar et al. (2019) presented a collection of images captured by a fixed-wing drone from the Measurement Analytics Lab (MANTIS) at Texas A&M University-Corpus Christi under a blanket Certificate of Authorization (COA) and equipped with Canon IXUS 127 HS 16.1 MP RGB and resolution of 3456 by 4608 pixels over the Welder Wildlife Foundation in Sinton. The database aimed to cover animal detection and counting and was evaluated by a deep learning method.

A DJI Phantom 4 Pro drone with 20-MPixel (resolution of 4864×3648 pixels) the camera was used to collect the Cattle database introduced in Barbedo et al. (2019). The place of collecting the database was at the Canchim farm, São Carlos, Brazil at 11 dates over the year of 2018. One of the study aims was to determine the ideal ground sample distance (GSD). In addition, a deep learning method was applied to the database.

The first method applied to the database presented in Xu et al. (2020) was Mask RCNN to detect cattle and sheep. The places of collecting the database are the Tullimba Research Feedlot (AEC18-038) owned by the University of New England, New South Wales, Australia and surrounding farmlands (AEC19-009) across seasons from Summer to Spring (February to October). The images were captured by an integrated PTZ camera which was mounted on a MAVIC PRO drone. The resolution of the images are 4096×2160 pixels. Additionally, a preprocessing on the images provided them to use in deep learning methods.

Some image samples from the databases are shown in Fig. 11. Also, the databases presented in animal detection are summarized in Table 7.

4.4 Disaster detection

Imagery based on drones has been opening up a growing, interesting, important role in disaster analysis due to its ability in real-time tasks, its high spatial resolution images, its oblique imagery, etc. These results lead to effective results in detecting cracks and damage and to help transportation planners make the right decisions. In the following section, we introduce drone-based databases related to disaster analysis.

The drone used in Jeon et al. (2013) was equipped with several sensors such as a mirrorless camera, a GPS, an IMU, and a sensor integration and synchronization module. The authors designed and produced a micro drone Sony NEX-55 camera with a resolution of 4912 by 3264 pixels that captured images at an altitude of 100m. The database can be used for disaster detection and monitoring.

²⁴ <https://sites.nicholas.duke.edu/uas/>.

²⁵ <https://sites.nicholas.duke.edu/uas/>.

Table 7 List of the databases used in animal detection

Database	Year	Availability	Description	Technique used
van Gemert (2014)	2014	Public ^a	Localization and counting of animals: 6 videos	–
Chamoso et al. (2014)	2014	Private	Cattle detection: 13,520 images, 443 frames; 70 frames included cattle	Rivas et al. (2018)
Gonzalez et al. (2016)	2016	Private:Email for request; ^b	Wildlife monitoring and conservation: thermal images	–
Okafor et al. (2017)	2017	Private: Email for request; ^c	Animal detection: 3981 samples	Okafor et al. (2018)
Seymour et al. (2017)	2017	Public ^d	Detection and enumeration of marine wildlife thermal images	–
Kellenberger et al. (2018)	2018	Public ^e	Animal detection: 469 images with 969 animals	Kellenberger et al. (2018, 2019)
Sarwar et al. (2018)	2018	Private: Email for request ^f	Detecting and counting sheep: 4 images (2048 * 1080 * 3)	–
Sajib et al. (2018)	2018	Upon request ^g	Surveillance and population estimation of marine 1970 frames	–
Shao et al. (2019)	2019	Public ^h	Cattle detection and counting:two databases of pasture aerial images, 670 images	–
Gray et al. (2019)	2019	public ⁱ	Detecting sea turtles: 1,059 UAS images	–
Rahmooonfar et al. (2019)	2019	Private: Email for request ^j	Counting and localization of sparse animal: images of 11 regions	–
Barbedo et al. (2019)	2019	Private: Email for request ^k	Cattle detection: 1853 images	Barbedo et al. (2020)
Xu et al. (2020)	2020	Private: Email for request ^l	Cattle counting: 2250 images	Xu et al. (2020)

^a<http://fisic-data.science.uva.nl/jvgemert/conservationDronesECCV14w/conservationDronesECCVwData.tar.gz>

^bfelipe.gonzalez@qut.edu.au

^ce.okafor@rug.nl

^d<http://seamap.env.duke.edu/dataset/1462>

^ehttps://kuzikus-namibia.de/xc_index.html

^ffarah.sarwar@aut.ac.nz

^gMuhammad.Sajib@uts.edu.au

^h<http://bird.nae-lab.org/cattle/>

ⁱ<https://datadryad.org/resource/doi:10.5061/dryad.5h06vv2>

^jmaryam.rahmooonfar@tamucc.edu

^kjayme.barbedo@embrapa.br

^lwangwensheng@caas.cn

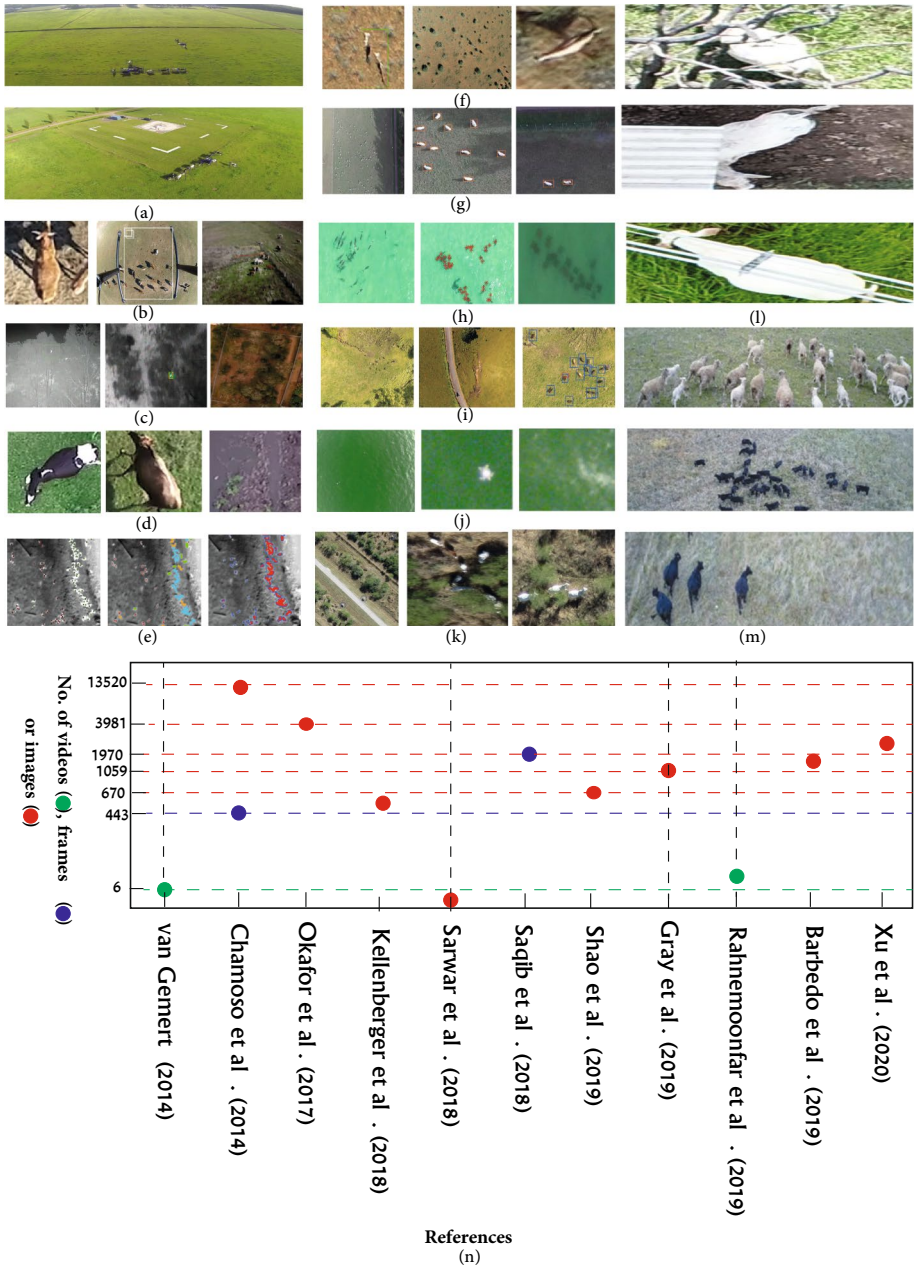


Fig. 11 Samples of the database images related to animal detection methods: **a** van Gemert (2014), **b** Chamoso et al. (2014), **c** Gonzalez et al. (2016), **d** Okafor et al. (2017), **e** Seymour et al. (2017), **f** Kellenberger et al. (2018), **g** Sarwar et al. (2018), **h** Saqib et al. (2018), **i** Shao et al. (2019), **j** Gray et al. (2019), **k** Rahnemooifar et al. (2019), **l** Barbedo et al. (2019), **m** Xu et al. (2020), and **n** a chart of the number of images (there are not information for the number of images for references Gonzalez et al. (2016) and Seymour et al. (2017)), frames, and videos in terms of the databases

The purpose of Offi et al. (2016) was to provide a database for disaster response and wildlife protection and anti-poaching efforts. The SAVMAP project research was a collaboration of Drone Adventures²⁶ and the EPFL Cooperation & Development Center²⁷. The authors also presented a solution based on machine learning approaches. Features extracted were based on the histogram of oriented gradient (HOG). Several machine learning methods such as SVM and logistic regression were selected accordingly.

The UAV Mosaicking and Change Detection (UMCD) database Avola et al. (2017) was used to support five tasks: object detection, people search and rescue, people and vehicle classification, military camp monitoring as well as urban area monitoring. The tasks are suitable for mosaicking and changing detection methods at low altitude. The database includes two sets of 30 and 20 videos, respectively.

In Kakooei and Baleghi (2017), oblique images were collected for disaster assessment including the Haitian earthquake of 2010²⁸, Hurricane Irene of 2011²⁹, Hurricane Sandy of 2012³⁰, the Illinois tornadoes of 2015³¹, and abc7chicago³². The database is suitable for earthquakes and hurricanes assessment. Moreover, a segmentation algorithm was considered for estimating facade and building damage in the areas.

In Bejiga et al. (2017), two databases were introduced for search and rescue (SAR) operations by drone images. The first database was a collection of different videos of a ski area from the website with a resolution of 1280 by 720 pixels, and the second database was captured by a CyberFed “Pinocchio” hexacopter equipped with a GoPro camera over a mountain close to the city of Trento at a flying height of 2–4 m for low flights and 20–40 m for high flights. Additionally, the database is appropriate for applying deep learning methods.

The database used in Attari et al. (2017) was provided by the World Bank in collaboration with the Humanitarian UAV Network (UAViators) during Cyclone Pam in Vanuatu in 2015. The database was targeted for monitoring damage and object detection in affected environments. In addition, a deep learning method (Nazr-CNN³³) was proposed for this goal. Therefore, the database can be used to compare deep learning methods.

L’Aquila database Duarte et al. (2017) was collected from the damage left by the earthquake in L’Aquila, Italy, in 2009. The database was obtained by flying an Aibot X6 hexacopter equipped with a Sony ILCE-6000 camera at an altitude of 100m. The database is appropriate for damage detection. A segmentation method based on CNNs was tested on the database. Additionally, they presented a solution using a sparse point cloud.

In Li et al. (2018) a database of 5 different scenes (urban, suburban, rural, wilderness and green land) from an airborne drone was collected and can be used for scene recognition and damage detection. Additionally, the authors used superpixel-based features for this purpose. The features were used to segment and detect the damages. In addition, an SVM classifier was considered for classifying the scenes.

²⁶ <http://droneadventures.org/>.

²⁷ <http://cooperation.epfl.ch/>.

²⁸ Available at <http://www.reuters.com/news/picture/ruins-of-haitis-national-palace?articleId=USRTR370GT>.

²⁹ Available at <http://environmentalheadlines.com/ct/2011/09/01/new-england-feels-hurricane-irene%E2%80%99s-impacts/hurricane-irene-damage-ct-nat-guard-east-haven>.

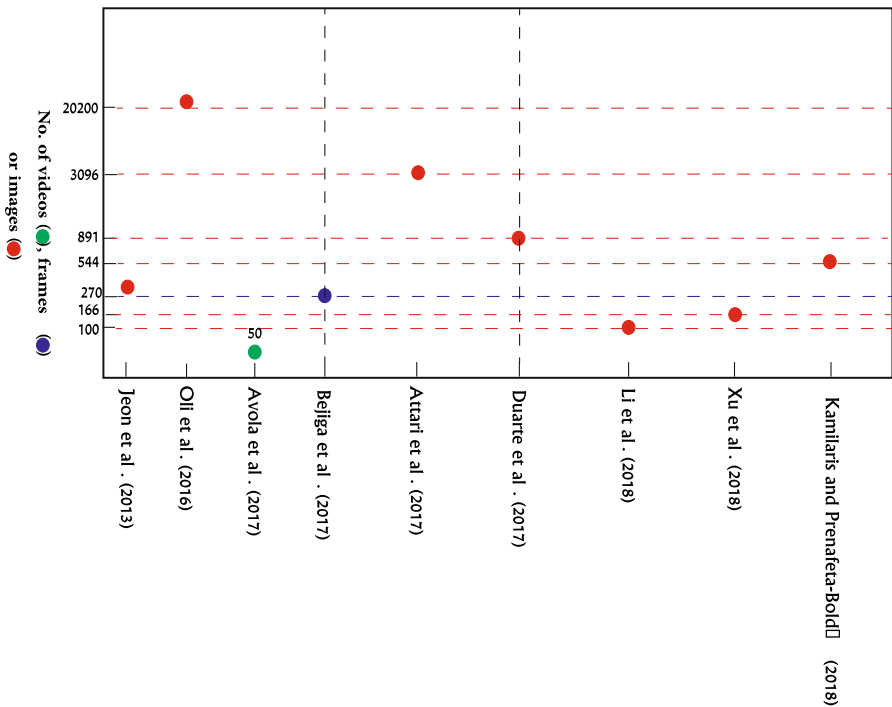
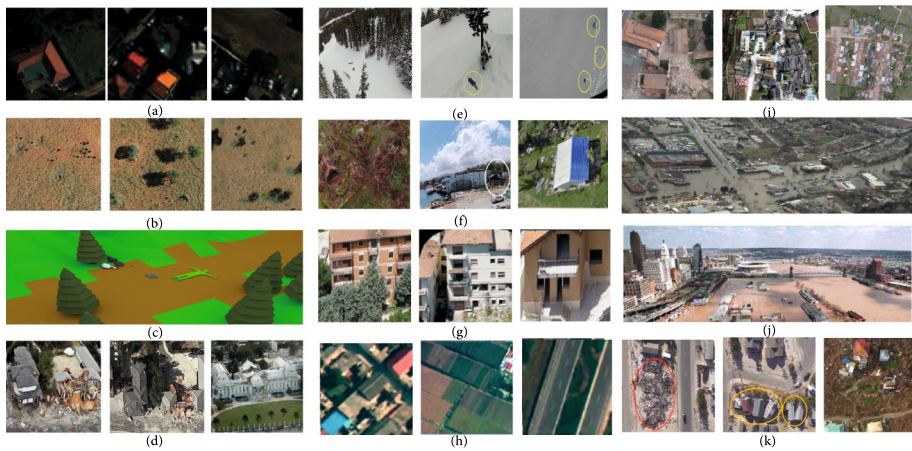
³⁰ Available at <http://www.defense.gov/Media/Photo-Gallery?igphoto=2001185999>.

³¹ Available at <http://www.chicagotribune.com/news/nationworld/83269837-132.html>.

³² <http://abc7chicago.com/news/illinois-tornado-victims-how-to-help/648502/>.

³³ Nazr means “sight” in Arabic.

Reference Xu et al. (2018) introduced three databases from drone earthquake images over three locations of Mirabello, Italy, 2012; Lushan County in Sichuan Province, China,



References

(l)

Fig. 12 Samples of the database images related to disaster methods: **a** Jeon et al. (2013), **b** Offi et al. (2016), **c** Avola et al. (2017), **d** Kakooei and Baleghi (2017), **e** Bejiga et al. (2017), **f** Attari et al. (2017), **g** Duarte et al. (2017), **h** Li et al. (2018), **i** Xu et al. (2018), **j** Kamilaris and Prenafeta-Boldo (2018), **k** Li et al. (2019), and **l** a chart of the number of images, frames, and videos in terms of the databases (there are not information for the number of images for references Kakooei and Baleghi (2017) and Li et al. (2019))

Table 8 List of the databases used in disaster detection

Database	Year	Availability	Description	Technique used
Jeon et al. (2013)	2013	Private Email for request ^a	Damage assessment: 300 images	–
Ofri et al. (2016)	2016	Public ^b	Disaster analysis and animal detection: 15,000 images and 5200 microimages	Rey et al. (2017), Kellenberger et al. (2017)
Avola et al. (2017)	2017	Public ^c	People search and rescue (50 videos)	Avola et al. (2017)
Kakooei and Baloghi (2017)	2017	Private: used images are available	Disaster damage assessment	–
Bejiga et al. (2017)	2017	Private:Email for request, ^d	Search and rescue operations 270 frames	–
Aitan et al. (2017)	2017	Private: used from images: the Humanitarian UAV Network (UAViators) ^e and Artificial Intelligence for disaster response (AIDR) ^f	Damage assessment: 3096 images	–
Duarte et al. (2017)	2017	Private Email for request ^g	Earthquake damage assessment 891 images in total	Duarte et al. (2018), Nex et al. (2019)
Li et al. (2018)	2018	Private Email for request ^h	Scene recognition 100 images	–
Xu et al. (2018)	2018	Private Email for request ⁱ second set: 95 images third set: 215 images	Earthquake damage mapping first set: 166 images	–
Kamilaris and Prenafeta-Boldú (2018)	2018	Private Email for request ^j	Disaster monitoring 544 images	Kamilaris et al. (2019)
Li et al. (2019)	2019	Public	Building damage detection Hurricane Sandy and Hurricane Irma ^k	Li et al. (2020)

^a dmlr0323@uos.ac.kr^b Available at <http://lasig.epfl.ch/savmap>^c <http://www.umcd-dataset.net/>^d mesaybelete.bejiga@unim.it^e <http://uaviators.org/>^f <http://aidr.qcri.org/>^g d.duarte@utwente.nl^h newsey@buaa.edu.cnⁱ awulixin@263.net^j andreas.kamilaris@irta.cat^k <https://storms.ngs.noaa.gov/storms/irma/index.html#6/26.657/-78.783>

2013; and Hanwang County in Sichuan Province, China, 2008. The databases were captured by multirotors, rotors, and fix-wing drones. To segment the damaged areas, feature extraction and classification methods were based on geometrical features and K-k-nearest neighbors (KNN), respectively. Additionally, the database images were generated in a 3D point cloud format.

Reference Kamilaris and Prenafeta-Boldú (2018) presented a small database for disaster detection and monitoring that was captured by a drone. The database is based on images of fires, earthquakes, collapsed buildings, tsunami and flooding, as well as “non-disaster” related scenes. Deep learning methods were evaluated in terms of the database.

Finally, reference Li et al. (2019) introduced a damaged building assessment database based on images from Hurricane Sandy in 2012 and Hurricane Irma at a resolution of 1920 by 1080 pixels which were collected by Drexel University. The database has classes with labels of undamaged buildings, damaged buildings, and ruins. Additionally, the deep learning method can be applied to the database.

Some image samples from the databases are shown in Fig. 12. Also, the databases presented in disaster detection are summarized in Table 8.

4.5 Face recognition

Since drone videos are more often captured from the top view, face and action recognition are challenging problems that should be solved when inspection and security are important for such videos. In the following, we introduce drone-based databases related to face recognition problems.

A very challenging database for human identity recognition was presented in Oreifej et al. (2010). The database is appropriate for detecting, segmenting, aligning, and recognition of humans viewed from aerial cameras with low resolution and adverse conditions. The images were captured by a drone and tested by the weighted region matching (WRM) method as the feature extraction and SVM as classification steps.

In Davis et al. (2013), a database was created to support low-cost facial detection and recognition tasks using an AR.Drone 1.0 that captured images with a resolution of 640 by 480 pixels. Feature extraction method applied on the database was based on local binary pattern (LBP) and the classifier used to train the features was KNN.

The mobile reidentification platforms (MRPs) database Layne et al. (2014) is a collection of images captured at a resolution of 640 by 360 pixels by a quadcopter drone. The database was the first platform for mobile reidentification to be used for face recognition. In addition, the authors used several feature extraction and classifier methods to evaluate the database.

DroneFace Hsu and Chen (2017) is a database that simulated a drone at an altitude of 1–5 m by a GoPro Hero3 camera. The aim of the database is a face recognition task with frontal and side portrait images. Additionally, they evaluated the database with several methods such as wavelet transform and LBP features and an SVM classifier.

The IARPA Janus Surveillance (IJB-S) database Kalka et al. (2018) was presented for face recognition. A small fixed-wing drone collected images captured by a Panasonic WV-SW3955 and Speco O4P30X6 dome cameras with resolutions of 1280 by 960 and 2592 by 1520 pixels, respectively.

The DroneSURF database for exploring the challenges of motion, variations in poses, illumination, and background in face recognition was introduced in Kalra et al. (2019). The images were captured by DJI Phantom 4 in a variety of altitudes and regulations for active and passive scenarios.

Table 9 List of the databases used in face recognition

Database	Year	Availability	Description	Technique used
Oreifej et al. (2010)	2010	Public ^a	Face recognition: 12,000 images	Perera et al. (2018), Walha et al. (2015), Yeh et al. (2016), Minaeian et al. (2018), Mliki et al. (2020)
Davis et al. (2013)	2013	Private: Email for request ^b	Facial recognition	–
Layne et al. (2014)	2014	Public ^c	Person re-identification: 16,000 frames	–
Hsu and Chen (2017)	2015	Public ^d	Face recognition: 620 pictures	Hsu and Chen (2015), Daryanavard and Harifi (2018), Deeb et al. (2020)
Kalka et al. (2018)	2018	Public ^e	Person identification: 10 videos, 1487 images	–
Kalra et al. (2019)	2019	Public ^f	Face recognition: 200 videos: 411,451 frames	–
Perera et al. (2019)	2019	Public ^g	Action recognition: 240 videos: 66,919 frames	–
Zhang et al. (2020)	2020	Public ^h	Person Re-identification: 39,461 images	–
Grigorev et al. (2020)	2020	Email for request ⁱ	Person Re-identification: 46,359 images	–

^a Available at http://csrcv.ucf.edu/data/UCF_Aerial_Action.php

^b Nicholas.Davis@tufts.edu

^c <http://homepages.inf.ed.ac.uk/thospeda/>

^d <https://hjsu.github.io/DroneFace/>

^e <https://www.nist.gov/programs-projects/face-challenges>

^f <http://www.iab-rubric.org/resources/dronesurf.html>

^g Available at <https://asankagp.github.io/droneaction>

^h Available at <https://github.com/stormyoung/PRAI-1581>

ⁱ igrigorev@hit.edu.cn



Fig. 13 Samples of the database images related to face recognition methods: **a** Oreifej et al. (2010), **b** Davis et al. (2013), **c** Layne et al. (2014), **d** Hsu and Chen (2017), **e** Kalka et al. (2018), **f** Kalra et al. (2019), **g** Perera et al. (2019), and **h** Zhang et al. (2020) (there are not image sample for reference Grigorev et al. (2020))

A new database (Drone-Action) was presented in Perera et al. (2019) for action recognition based on person images captured by a drone equipped with a GoPro Hero 4 Black camera. The images have HD (1920×1080 pixels) format. The actions were classified into three categories of the following, side-view, and front-view actions. Deep learning methods were tested on the database that showed the database is suitable for such methods.

PRAI-1581 database Zhang et al. (2020) was introduced to re-identify persons based on images captured by two DJI consumer drones. The images of the database were collected with flying heights ranging from 20 to 60 m. Several state-of-the-art methods such as deep learning methods were tested on the database. Therefore, the database is suitable to test deep learning methods.

Reference Grigorev et al. (2020) presented a database for person re-identification purposes. A remote-operated quadcopter mounted by HD camera collected images with a resolution of 1920×1080 pixels at a height of 25 m. The ground truth of the database was included with 18 attributes such as gender (male and female) and type of lower-body clothing (pants and overcoat). Additionally, a deep learning method was applied to the database.

Some image samples from the databases are shown in Fig. 13. Also, the databases presented in face recognition are summarized in Table 9.

5 Open research

As the FAA predicts, the number of drones will exceed 4 million units Boroujerdian et al. (2018); therefore, the design and implementation of an accurate system for different applications will play an important role. More research needs to be done in the community, and this research will not happen unless researchers provide more databases for different purposes. Additionally, the domain of applications will increase, and new applications and problems will be introduced without introducing databases for others to use.

One of the new applications is cinematography (for movies and sports) by drones. Although the application is currently manually operated, autonomous approaches based on machine learning and computer vision are being developed. However, several challenges exist, such as tracking fast and unpredictably moving targets. For handling some of these challenges, researchers can use videos from certain websites³⁴Huang et al. (2019).

Another application in the area is archeology, which can use computer vision to document archeological sites, including 3D maps, orthophotos, and thermal images Xiang et al. (2018). To the best of our knowledge, there is no database for this application.

Recently, indoor approaches for drones such as in Kaufmann et al. (2018) have been introduced, and public databases need to be introduced to facilitate more research. It should be noted that methods used outdoors can also be used in indoor approaches.

One of the new applications related to surveillance and traffic is paying attention to and tracking pedestrian movement for future cities, especially detecting pedestrians, vehicles, and cyclists at traffic intersections for determining transit times Zhu et al. (2019).

Growing websites for datacenters and repositories, similar to Dronestagram³⁵ are required for researchers to share their achievements in the field. As mentioned in Hochmair and Zielstra (2015), the Dronestagram project provides a space for sharing photos captured by drones. Information such as drone models, camera models and the upload dates are shown on the website. The first photo was uploaded in July 2013. Reference Johnson et al. (2017) also introduced some other hosting services^{36, 37, 38} and provided a website³⁹ for consulting companies or volunteer groups that do not have any space to share their data (especially images and videos).

By exploring the tables presented in this survey, researchers can decide to define new projects and present new databases for different applications. For example, providing the databases related to disaster detection, such as fire detection, can be very useful in issues related to assistance and rescue. As shown in Table 3 for the remote sensing and navigation databases, 8 databases have not yet been used, and researchers can use the database for these topics.

³⁴ <http://gettyimages.com>.

³⁵ Available at <http://www.dronestagr.am/>.

³⁶ <https://openaerialmap.org/>.

³⁷ <https://github.com/openimagerynetwork>.

³⁸ <http://coastalresilience.org/project-areas/california/el-nino-california/>.

³⁹ <http://droneadventures.org/>.

6 Conclusion

Today, drones play an important role in automating processes that are too hard for humans. In this paper, we have surveyed applications related to drones and computer vision methods. We categorized images and videos captured by drones into three groups: remote sensing (camera calibration, image matching and aerial triangulation), navigation (flight control, visual localization and mapping, and target tracking and obstacle detection), and applications related to the sensed environment (surveillance, agriculture and forest, animal detection, disaster detection, face recognition). In this paper, we focused on databases for the three categories. Finally, we presented open research based on information obtained in the survey. As mentioned in the open research section, researchers in the field still need to present databases for existing applications and develop databases for new applications. Additionally, because the number of drones based on new hardware is growing exponentially and the rapid advancement of drones is unstoppable, the increasing power and accuracy of software with embedded computer vision methods is an essential reality. This aim seems as if it will not be achieved unless databases are presented in various applications for applying and testing the methods proposed by other researchers.

Acknowledgements This publication was made possible by NPRP Grant # NPRP8-140-2-065 from Qatar National Research Fund (a member of Qatar Foundation). The statement made herein are solely the responsibility of the authors.

References

- Abughalieh KM, Sababha BH, Rawashdeh NA (2018) A video-based object detection and tracking system for weight sensitive uavs. *Multimed Tools Appl* 78:9149–9167
- Adams SM, Friedland CJ (2011) A survey of unmanned aerial vehicle (uav) usage for imagery collection in disaster research and management. In: 9th international workshop on remote sensing for disaster response, vol 8
- Adão T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, Sousa J (2017) Hyperspectral imaging: a review on uav-based sensors, data processing and applications for agriculture and forestry. *Remote Sens* 9(11):1110
- Al-Kaff A, García F, Martín D, De La Escalera A, Armingol J (2017) Obstacle detection and avoidance system based on monocular camera and size expansion algorithm for uavs. *Sensors* 17(5):1061
- Al-Kaff A, Martín D, García F, de la Escalera A, Armingol JM (2018) Survey of computer vision algorithms and applications for unmanned aerial vehicles. *Expert Syst Appl* 92:447–463
- Al Kaff AHA (2017) Vision-based navigation system for unmanned aerial vehicles. Ph.D. dissertation, Universidad Carlos III de Madrid, 2017. <https://e-archivo.uc3m.es/handle/10016/26603>
- Al-Sheary A, Almagbile A (2017) Crowd monitoring system using unmanned aerial vehicle (uav). *J Civ Eng Archit* 11:1014–1024
- Albanis G, Zioulis N, Dimou A, Zarpalas D, Daras P (2020) Dronepose: photorealistic uav-assistant dataset synthesis for 3d pose estimation via a smooth silhouette loss. [arXiv:2008.08823](https://arxiv.org/abs/2008.08823)
- Alidoost F, Arefi H (2015) An image-based technique for 3d building reconstruction using multi-view uav images. *Int Arch Photogram Remote Sens Spatial Inf Sci* 40(1):43
- Almagbile A (2019) Estimation of crowd density from uavs images based on corner detection procedures and clustering analysis. *Geo-spatial Inf Sci* 22(1):23–34
- Askar W, Elmowafy O, Youssif A, Elnashar G (2017) Optimized uav object tracking framework based on integrated particle filter with ego-motion transformation matrix. In: MATEC web of conferences, vol 125. EDP Sciences, p 04027
- Attari N, Offi F, Awad M, Lucas J, Chawla S (2017) Nazr-cnn: fine-grained classification of uav imagery for damage assessment. In: 2017 IEEE international conference on data science and advanced analytics (DSAA). IEEE, pp 50–59
- Avola D, Cinque L, Foresti GL, Martinel N, Pannone D, Piciarelli C (2018) A uav video dataset for mosaicking and change detection from low-altitude flights. *IEEE Trans Syst Man Cybern Syst* 99:1–11

- Avola D, Cinque L, Foresti GL, Pannone D (2018) Visual cryptography for detecting hidden targets by small-scale robots. In: International conference on pattern recognition applications and methods. Springer, pp 186–201
- Avola D, Foresti GL, Martinel N, Micheloni C, Pannone D, Piciarelli C (2017) Aerial video surveillance system for small-scale uav environment monitoring. In: 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS). IEEE, pp 1–6
- Avola D, Foresti GL, Martinel N, Pannone D, Piciarelli C (2017) The umcd dataset. [arXiv:1704.01426](https://arxiv.org/abs/1704.01426)
- Azimi SM, Fischer P, Körner M, Reinartz P (2018) Aerial lanenet: lane marking semantic segmentation in aerial imagery using wavelet-enhanced cost-sensitive symmetric fully convolutional neural networks. [arXiv:1803.06904](https://arxiv.org/abs/1803.06904)
- Backes D, Schumann G, Teferele F, Boehm J (2019) Towards a high-resolution drone-based 3d mapping dataset to optimise flood hazard modelling. *Int Arch Photogramm Remote Sens Spatial Inf Sci* 42(W13):181–187
- Ballan L, Castaldo F, Alahi A, Palmieri F, Savarese S (2016) Knowledge transfer for scene-specific motion prediction. In: European conference on computer vision. Springer, pp 697–713
- Barbedo JGA, Koenigkan LV, Santos PM, Ribeiro ARB (2020) Counting cattle in uav images—dealing with clustered animals and animal/background contrast changes. *Sensors* 20(7):2126
- Barbedo JGA, Koenigkan LV, Santos TT, Santos PM (2019) A study on the detection of cattle in uav images using deep learning. *Sensors* 19(24):5436
- Barekatin M, Marti M, Shih HF, Murray S, Nakayama K, Matsuo Y, Prendinger H (2017) Okutama-action: an aerial view video dataset for concurrent human action detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 28–35
- Barmounakis E, Geroliminis N (2020) On the new era of urban traffic monitoring with massive drone data: the pneuma large-scale field experiment. *Transp Res Part C Emerg Technol* 111:50–71
- Bejiga M, Zeggada A, Nouffidj A, Melgani F (2017) A convolutional neural network approach for assisting avalanche search and rescue operations with uav imagery. *Remote Sens* 9(2):100
- Berker Logoglu K, Lezki H, Kerim Yucel M, Ozturk A, Kucukmurler A, Karagoz B, Erdem E, Erdem A (2017) Feature-based efficient moving object detection for low-altitude aerial platforms. In: Proceedings of the IEEE international conference on computer vision, pp 2119–2128
- Bharati SP, Wu Y, Sui Y, Padgett C, Wang G (2018) Real-time obstacle detection and tracking for sense-and-avoid mechanism in uavs. *IEEE Trans Intell Veh* 3(2):185–197
- Bochinski E, Senst T, Sikora T (2018) Extending iou based multi-object tracking by visual information. In: 2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS). IEEE, pp 1–6
- Bonetto M, Korshunov P, Ramponi G, Ebrahimi T (2015) Privacy in mini-drone based video surveillance. In: 2015 11th IEEE international conference and workshops on automatic face and gesture recognition (FG), vol. 4, pp. 1–6. IEEE
- Boroujerdian B, Genc H, Krishnan S, Cui W, Faust A, Reddi V (2018) Mavbench: micro aerial vehicle benchmarking. In: 2018 51st annual IEEE/ACM international symposium on microarchitecture (MICRO). IEEE, pp 894–907
- Carletti V, Greco A, Saggese A, Vento M (2018) Multi-object tracking by flying cameras based on a forward-backward interaction. *IEEE Access* 6:43905–43919
- Carletti V, Greco A, Saggese A, Vento M (2019) An intelligent flying system for automatic detection of faults in photovoltaic plants. *J Ambient Intell Hum Comput* 11:2027–2040
- Carrío A, Vemprala S, Ripoll A, Saripalli S, Campoy P (2018) Drone detection using depth maps. In: 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, pp 1034–1037
- Cavaliere D, Loia V, Saggese A, Senatore S, Vento M (2019) A human-like description of scene events for a proper uav-based video content analysis. *Knowl-Based Syst* 178:163–175
- Cazzato D, Cimarelli C, Sanchez-Lopez JL, Voos H, Leo M (2020) A survey of computer vision methods for 2d object detection from unmanned aerial vehicles. *J Imag* 6(8):78
- Cehovin Zajc L, Lukezic A, Leonardis A, Kristan M (2017) Beyond standard benchmarks: parameterizing performance evaluation in visual object tracking. In: Proceedings of the IEEE international conference on computer vision, pp 3323–3331
- Chamoso P, Raveane W, Parra V, González A (2014) Uavs applied to the counting and monitoring of animals. In: Ambient intelligence—software and applications. Springer, pp 71–80
- Chen L, Liu F, Zhao Y, Wang W, Yuan X, Zhu J (2020) Valid: a comprehensive virtual aerial image dataset. In: 2020 IEEE international conference on robotics and automation (ICRA). IEEE, pp 2009–2016. <https://doi.org/10.1109/ICRA40945.2020.9197186>
- Chen PH, Lee CY (2018) Uavnet: an efficient obstacle detection model for uav with autonomous flight. In: 2018 international conference on intelligent autonomous systems (ICoIAS). IEEE, pp 217–220

- Chen X, Li Z, Yang Y, Qi L, Ke R (2020) High-resolution vehicle trajectory extraction and denoising from aerial videos. *IEEE Trans Intell Transp Syst*
- Chen Y, Liu L, Gong Z, Zhong P (2017) Learning cnn to pair uav video image patches. *IEEE J Sel Topics Appl Earth Obs Remote Sens* 10(12):5752–5768
- Chen Y, Wang Y, Lu P, Chen Y, Wang G (2018) Large-scale structure from motion with semantic constraints of aerial images. In: Chinese conference on pattern recognition and computer vision (PRCV). Springer, pp 347–359
- Choi SY, Cha D (2019) Unmanned aerial vehicles using machine learning for autonomous flight; state-of-the-art. *Adv Robot* 33:265–277
- Collins R, Zhou X, Teh SK (2005) An open source tracking testbed and evaluation web site. In: IEEE international workshop on performance evaluation of tracking and surveillance, vol 2, p 35
- Colomina I, Molina P (2014) Unmanned aerial systems for photogrammetry and remote sensing: a review. *ISPRS J Photogramm Remote Sens* 92:79–97
- Dandois J, Olano M, Ellis E (2015) Optimal altitude, overlap, and weather conditions for computer vision uav estimates of forest structure. *Remote Sens* 7(10):13895–13920
- Daryanavard H, Harifi A (2018) Implementing face detection system on uav using raspberry pi platform. In: Iranian conference on electrical engineering (ICEE). IEEE, pp 1720–1723
- Davis N, Pittaluga F, Panetta K (2013) Facial recognition using human visual system algorithms for robotic and uav platforms. In: 2013 IEEE conference on technologies for practical robot applications (TePRA). IEEE, pp 1–5
- Deeb A, Roy K, Edoh KD (2020) Drone-based face recognition using deep learning. In: International conference on advanced machine learning technologies and applications. Springer, pp 197–206
- Dinh M, Morris B, Kim Y (2019) Uas-based object tracking via deep learning. In: 2019 IEEE 9th annual computing and communication workshop and conference (CCWC). IEEE, pp 0217–0275
- Du D, Qi Y, Yu H, Yang Y, Duan K, Li G, Zhang W, Huang Q, Tian Q (2018) The unmanned aerial vehicle benchmark: Object detection and tracking. In: Proceedings of the European conference on computer vision (ECCV), pp 370–386
- Du D, Zhu P, Wen L, Bian X, Ling H, Hu Q, Zheng J, Peng T, Wang X, Zhang Y, et al. (2019) Visdrone-sot2019: the vision meets drone single object tracking challenge results. In: Proceedings of the IEEE international conference on computer vision workshops
- Duarte D, Nex F, Kerle N, Vosselman G (2017) Towards a more efficient detection of earthquake induced facade damages using oblique uav imagery. *Int Arch Photogramm Remote Sens Spatial Inf Sci* 42:93
- Duarte D, Nex F, Kerle N, Vosselman G (2018) Multi-resolution feature fusion for image classification of building damages with convolutional neural networks. *Remote Sens* 10(10):1636
- Elharrouss O, Almaadeed N, Al-Maadeed S, Akbari Y (2019) Image inpainting: a review. *Neural Process Lett* 51:2007–2028. <https://doi.org/10.1007/s11063-019-10163-0>
- Elharrouss O, Almaadeed N, Al-Maadeed S, Bouridane A, Beghdadi A (2020) A combined multiple action recognition and summarization for surveillance video sequences. *Appl Intell*. <https://doi.org/10.1007/s10489-020-01823-z>
- Escalante H, Rodríguez-Sánchez S, Jiménez-Lizárraga M, Morales-Reyes A, De La Calleja J, Vazquez R (2019) Barley yield and fertilization analysis from uav imagery: a deep learning approach. *Int J Remote Sens* 40(7):2493–2516
- Fan H, Ling H (2019) Parallel tracking and verifying. *IEEE Trans Image Process* 28(8):4130–4144
- Gago J, Douthe C, Coopman R, Gallego P, Ribas-Carbo M, Flexas J, Escalona J, Medrano H (2015) Uavs challenge to assess water stress for sustainable agriculture. *Agric Water Manag* 153:9–19
- Giordan D, Hayakawa Y, Nex F, Remondino F, Tarolli P (2018) The use of remotely piloted aircraft systems (rpas) for natural hazards monitoring and management. *Nat Hazards Earth Syst Sci* 18(4):1079–1096
- Gonzalez L, Montes G, Puig E, Johnson S, Mengersen K, Gaston K (2016) Unmanned aerial vehicles (uavs) and artificial intelligence revolutionizing wildlife monitoring and conservation. *Sensors* 16(1):97
- Gray PC, Fleishman AB, Klein DJ, McKown MW, Bézy VS, Lohmann KJ, Johnston DW (2019) A convolutional neural network for detecting sea turtles in drone imagery. *Methods Ecol Evol* 10(3):345–355
- Grigorev A, Liu S, Tian Z, Xiong J, Rho S, Feng J (2020) Delving deeper in drone-based person re-id by employing deep decision forest and attributes fusion. *ACM Trans Multimед Comput Commun Appl (TOMM)* 16(1):1–15
- Hao C, Zhang X, Li Y, Huang S, Xiong J, Rupnow K, Hwu Wm, Chen D (2019) Fpga/dnn co-design: an efficient design methodology for iot intelligence on the edge. [arXiv:1904.04421](https://arxiv.org/abs/1904.04421)
- Henrio J, Nakashima T (2018) Anomaly detection in videos recorded by drones in a surveillance context. In: 2018 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, pp 2503–2508
- Hochmair HH, Zielstra D (2015) Analysing user contribution patterns of drone pictures to the dronestagram photo sharing portal. *J Spatial Sci* 60(1):79–98

- Hsieh MR, Lin YL, Hsu WH (2017) Drone-based object counting by spatially regularized regional proposal network. In: Proceedings of the IEEE international conference on computer vision, pp 4145–4153
- Hsu HJ, Chen KT (2015) Face recognition on drones: issues and limitations. In: Proceedings of the first workshop on micro aerial vehicle networks, systems, and applications for civilian use. ACM, pp 39–44
- Hsu HJ, Chen KT (2017) Droneface: an open dataset for drone research. In: Proceedings of the 8th ACM on multimedia systems conference. ACM, pp 187–192
- Hu B, Yang H, Wang L, Chen S (2019) A trajectory prediction based intelligent handover control method in uav cellular networks. *China Commun* 16(1):1–14
- Huang C, Yang Z, Kong Y, Chen P, Yang X, Cheng KTT (2019) Learning to capture a film-look video with a camera drone. In: 2019 International conference on robotics and automation (ICRA). IEEE, pp 1871–1877
- Hussein AAM (2018) Control and communication systems for automated vehicles cooperation and coordination. PhD thesis, Universidad Carlos III de Madrid. <https://e-archivo.uc3m.es/handle/10016/27674>
- Ilyas N, Shahzad A, Kim K (2020) Convolutional-neural network-based image crowd counting: review, categorization, analysis, and performance evaluation. *Sensors* 20(1):43
- Jeon E, Choi K, Lee I, Kim H (2013) A multi-sensor micro uav based automatic rapid mapping system for damage assessment in disaster areas. *ISPRS-Int Arch Photogramm Remote Sens Spatial Inf Sci* 1(2):217–221
- Johnson P, Ricker B, Harrison S (2017) Volunteered drone imagery: challenges and constraints to the development of an open shared image repository. In: Proceedings of the 50th Hawaii International Conference on System Sciences. Available from: <http://scholarspace.manoa.hawaii.edu/handle/10125/41396>. Accessed 23 Feb 2017
- Kakooei M, Baleghi Y (2017) Fusion of satellite, aircraft, and uav data for automatic disaster damage assessment. *Int J Remote Sens* 38(8–10):2511–2534
- Kalka ND, Maze B, Duncan JA, O'Connor K, Elliott S, Hebert K, Bryan J, Jain AK (2018) Ijb-s: Iarpa janus surveillance video benchmark. In: 2018 IEEE 9th international conference on biometrics theory, applications and systems (BTAS). IEEE, pp 1–9
- Kalra I, Singh M, Nagpal S, Singh R, Vatsa M, Sujit P (2019) Dronesurf: benchmark dataset for drone-based face recognition
- Kamilaris A, van den Brink C, Karatsiolis S (2019) Training deep learning models via synthetic data: application in unmanned aerial vehicles. In: International conference on computer analysis of images and patterns. Springer, pp 81–90
- Kamilaris A, Prenafeta-Boldú FX (2018) Disaster monitoring using unmanned aerial vehicles and deep learning. [arXiv:1807.11805](https://arxiv.org/abs/1807.11805)
- Kanellakis C, Nikolakopoulos G (2017) Survey on computer vision for uavs: current developments and trends. *J Intell Robot Syst* 87(1):141–168
- Kang K, Belkhal S, Kahn G, Abbeel P, Levine S (2019) Generalization through simulation: integrating simulated and real data into deep reinforcement learning for vision-based autonomous flight. [arXiv:1902.03701](https://arxiv.org/abs/1902.03701)
- Kanistras K, Martins G, Rutherford MJ, Valavanis KP (2015) A survey of unmanned aerial vehicles (UAVs) for traffic monitoring. In: 2013 international conference on unmanned aircraft systems (ICUAS), Atlanta, GA, 2013, pp 221–234. <https://doi.org/10.1109/ICUAS.2013.6564694>
- Karaduman M, Çınar A, Eren H (2019) Uav traffic patrolling via road detection and tracking in anonymous aerial video frames. *J Intell Robot Syst*, pp 1–16
- Kaufmann E, Loquercio A, Ranftl R, Dosovitskiy A, Koltun V, Scaramuzza D (2018) Deep drone racing: learning agile flight in dynamic environments. [arXiv:1806.08548](https://arxiv.org/abs/1806.08548)
- Ke R, Li Z, Kim S, Ash J, Cui Z, Wang Y (2017) Real-time bidirectional traffic flow parameter estimation from aerial videos. *IEEE Trans Intell Transp Syst* 18(4):890–901
- Ke R, Li Z, Tang J, Pan Z, Wang Y (2018) Real-time traffic flow parameter estimation from uav video based on ensemble classifier and optical flow. *IEEE Trans Intell Transp Syst* 99:1–11
- Kellenberger B, Marcos D, Lobry S, Tuia D (2019) Half a percent of labels is enough: efficient animal detection in uav imagery using deep cnns and active learning. *IEEE Trans Geosci Remote Sens* 57(12):9524–9533
- Kellenberger B, Marcos D, Tuia D (2018) Best practices to train deep models on imbalanced datasets—a case study on animal detection in aerial imagery. In: Joint European conference on machine learning and knowledge discovery in databases. Springer, pp 630–634
- Kellenberger B, Marcos D, Tuia D (2018) Detecting mammals in uav images: best practices to address a substantially imbalanced dataset with deep learning. *Remote Sens Environ* 216:139–153

- Kellenberger B, Volpi M, Tuia D (2017) Fast animal detection in uav images using convolutional neural networks. In: 2017 IEEE international geoscience and remote sensing symposium (IGARSS). IEEE, pp 866–869
- Kerle N, Nex F, Gerke M, Duarte D, Vetrivel A (2020) Uav-based structural damage mapping: a review. *ISPRS Int J Geo-inf* 9(1):14
- Korthals T, Kragh M, Christiansen P, Karstoft H, Jørgensen RN, Rückert U (2018) Multi-modal detection and mapping of static and dynamic obstacles in agriculture for process evaluation. *Front Robot AI* 5:28
- Kragh M, Christiansen P, Laursen M, Larsen M, Steen K, Green O, Karstoft H, Jørgensen R (2017) Field-safe: dataset for obstacle detection in agriculture. *Sensors* 17(11):2579
- Kristan M, Leonardis A, Matas J, Felsberg M, Pflugfelder R, Cehovin Zajc L, Vojir T, Hager G, Lukežič A, Eldesokey A, et al (2017) The visual object tracking vot2017 challenge results. In: Proceedings of the IEEE international conference on computer vision, pp 1949–1972
- Kuai Y, Wen G, Li D (2018) Multi-task hierarchical feature learning for real-time visual tracking. *IEEE Sens J* 19(5):1961–1968
- Kyrkou C, Plastiras G, Theocharides T, Venieris SI, Bouganis CS (2018) Dronet: efficient convolutional neural network detector for real-time uav applications. In: 2018 design, automation & test in Europe conference & exhibition (DATE). IEEE, pp 967–972
- Layne R, Hospedales TM, Gong S (2014) Investigating open-world person re-identification using a drone. In: European conference on computer vision. Springer, pp 225–240
- Lee SC (2016) A trajectory based event classification from uav videos and its evaluation framework. In: 2016 IEEE applied imagery pattern recognition workshop (AIPR). IEEE, pp 1–4
- Li D, Wen G, Kuai Y, Porikli F (2018) End-to-end feature integration for correlation filter tracking with channel attention. *IEEE Signal Process Lett* 25(12):1815–1819
- Li H, Shi Y, Zhang B, Wang Y (2018) Superpixel-based feature for aerial image scene recognition. *Sensors* 18(1):156
- Li W, Li H, Wu Q, Chen X, Ngan KN (2019) Simultaneously detecting and counting dense vehicles from drone images. *IEEE Trans Ind Electron* 66(12):9651–9662. <https://doi.org/10.1109/TIE.2019.2899548>
- Li Y, Hu W, Dong H, Zhang X (2019) Building damage detection from post-event aerial imagery using single shot multibox detector. *Appl Sci* 9(6):1128
- Li Y, Lin C, Li H, Hu W, Dong H, Liu Y (2020) Unsupervised domain adaptation with self-attention for post-disaster building damage detection. *Neurocomputing* 415:27–39
- Liu K, Mattyus G (2015) Fast multiclass vehicle detection on aerial images. *IEEE Geosci Remote Sens Lett* 12(9):1938–1942
- Liu Y, Yang F, Hu P (2020) Small-object detection in uav-captured images via multi-branch parallel feature pyramid networks. *IEEE Access* 8:145,740–145,750
- Long H, Chung Y, Liu Z, Bu S (2019) Object detection in aerial images using feature fusion deep networks. *IEEE Access* 7:30980–30990
- Long Y, Xia GS, Li S, Yang W, Yang MY, Zhu XX, Zhang L, Li, D (2020) Dirs: on creating benchmark datasets for remote sensing image interpretation. [arXiv:2006.12485](https://arxiv.org/abs/2006.12485)
- Loquercio A, Maqueda AI, del Blanco CR, Scaramuzza D (2018) Dronet: learning to fly by driving. *IEEE Robot Autom Lett* 3(2):1088–1095
- Lukežič A, Zajc LČ, Vojř T, Matas J, Kristan M (2019) Performance evaluation methodology for long-term visual object tracking. [arXiv:1906.08675](https://arxiv.org/abs/1906.08675)
- Luna CVM (2013) Visual tracking, pose estimation, and control for aerial vehicles. Ph.D. thesis, Universidad Politécnica de Madrid
- Lyu Y, Vosselman G, Xia GS, Yilmaz A, Yang MY (2020) Uavid: a semantic segmentation dataset for uav imagery. *ISPRS J Photogramm Remote Sens* 165:108–119
- Majid Azimi S (2018) Shuffledet: real-time vehicle detection network in on-board embedded uav imagery. In: Proceedings of the European conference on computer vision (ECCV)
- Mandal M, Kumar LK, Vipparthi SK (2020) Mor-uav: a benchmark dataset and baselines for moving object recognition in uav videos. [arXiv:2008.01699](https://arxiv.org/abs/2008.01699)
- Mantegazza D, Guzzi J, Gambardella LM, Giusti A (2018) Vision-based control of a quadrotor in user proximity: mediated vs end-to-end learning approaches. [arXiv:1809.08881](https://arxiv.org/abs/1809.08881)
- Mantegazza D, Guzzi J, Gambardella LM, Giusti A (2019) Learning vision-based quadrotor control in user proximity. In: 2019 14th ACM/IEEE international conference on human-robot interaction (HRI). IEEE, pp 369–369

- Marcu A, Costea D, Licaret V, Pirvu M, Slusanschi E, Leordeanu M (2018) Safeuav: learning to estimate depth and safe landing areas for uavs from synthetic data. In: Proceedings of the European conference on computer vision (ECCV)
- Maria G, Baccaglioni E, Brevi D, Gavelli M, Scopigno R (2016) A drone-based image processing system for car detection in a smart transport infrastructure. In: 2016 18th mediterranean electrotechnical conference (MELECON). IEEE, pp 1–5
- Maurya AK, Singh D, Singh K (2018) Development of fusion approach for estimation of vegetation fraction cover with drone and sentinel-2 data. In: IGARSS 2018–2018 IEEE international geoscience and remote sensing symposium. IEEE, pp 7448–7451
- Micheal AA, Vani K (2019) Automatic object tracking in optimized uav video. *J Supercomput* 75(8):4986–4999
- Minaeian S, Liu J, Son YJ (2015) Crowd detection and localization using a team of cooperative uav/ugvs. In: IIE annual conference. Proceedings, p. 595. Institute of industrial and systems engineers (IISE)
- Minaeian S, Liu J, Son YJ (2018) Effective and efficient detection of moving targets from a uav's camera. *IEEE Trans Intell Transp Syst* 19(2):497–506
- Mliki H, Bouhlef F, Hammami M (2020) Human activity recognition from uav-captured video sequences. *Pattern Recogn* 100:107,140
- Mou L, Hua Y, Jin P, Zhu XX (2020) Era: a dataset and deep learning benchmark for event recognition in aerial videos. [arXiv:2001.11394](https://arxiv.org/abs/2001.11394)
- Mueller M., Sharma G, Smith N, Ghanem B (2016) Persistent aerial tracking system for uavs. In: 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, pp 1562–1569
- Mueller M, Smith N, Ghanem B (2016) A benchmark and simulator for uav tracking. In: European conference on computer vision. Springer, pp 445–461
- Müller M, Casser V, Lahoud J, Smith N, Ghanem B (2018) Sim4cv: a photo-realistic simulator for computer vision applications. *Int J Comput Vis* 126(9):902–919
- Müller M, Casser V, Smith N, Michels DL, Ghanem B (2017) Teaching uavs to race using sim4cv. [arXiv:1708.05884](https://arxiv.org/abs/1708.05884)
- Muller M, Casser V, Smith N, Michels DL, Ghanem B (2018) Teaching uavs to race: end-to-end regression of agile controls in simulation. In: Proceedings of the European conference on computer vision (ECCV). https://doi.org/10.1007/978-3-030-11012-3_2
- Müller M, Li G, Casser V, Smith N, Michels DL, Ghanem B (2019) Learning a controller fusion network by online trajectory filtering for vision-based uav racing. [arXiv:1904.08801](https://arxiv.org/abs/1904.08801)
- Murray S (2017) Real-time multiple object tracking—a study on the importance of speed. [arXiv:1709.03572](https://arxiv.org/abs/1709.03572)
- Murugan D, Garg A, Singh D (2017) Development of an adaptive approach for precision agriculture monitoring with drone and satellite data. *IEEE J Sel Topics Appl Earth Obs Remote Sens* 10(12):5322–5328
- Najiya K, Archana M (2018) Uav video processing for traffic surveillance with enhanced vehicle detection. In: 2018 second international conference on inventive communication and computational technologies (ICICCT). IEEE, pp 662–668
- Nex F, Duarte D, Steenbeek A, Kerle N (2019) Towards real-time building damage mapping with low-cost uav solutions. *Remote Sens* 11(3):287
- Nex F, Remondino F, Gerke M, Przybilla HJ, Bäumker M, Zurhorst A (2015) Isprs benchmark for multi-platform photogrammetry. *ISPRS Ann Photogramm Remote Sens Spatial Inf Sci* 2
- Ofli F, Meier P, Imran M, Castillo C, Tuia D, Rey N, Briant J, Millet P, Reinhard F, Parkan M et al (2016) Combining human computing and machine learning to make sense of big (aerial) data for disaster response. *Big Data* 4(1):47–59
- Oh S, Hoogs A, Perera A, Cuntoor N, Chen CC, Lee JT, Mukherjee S, Aggarwal J, Lee H, Davis L, et al (2011) A large-scale benchmark dataset for event recognition in surveillance video. In: CVPR 2011. IEEE, pp 3153–3160
- Okafor E, Schomaker L, Wiering MA (2018) An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals. *J Inf Telecommun* 2(4):465–491
- Okafor E, Smit R, Schomaker L, Wiering M (2017) Operational data augmentation in classifying single aerial images of animals. In: 2017 IEEE international conference on innovations in intelligent systems and applications (INISTA). IEEE, pp 354–360
- Oppenheim D, Edan Y, Shani G (2017) Detecting tomato flowers in greenhouses using computer vision. *World Acad Sci Eng Technol Int J Comput Electr Autom Control Inf Eng* 11(1):104–109
- Oreifej O, Mehran R, Shah M (2010) Human identity recognition in aerial images. In: 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, pp 709–716
- Otto A, Agatz N, Campbell J, Golden B, Pesch E (2018) Optimization approaches for civil applications of unmanned aerial vehicles (uavs) or aerial drones: a survey. *Networks* 72(4):411–458

- Pádua L, Vanko J, Hruška J, Adão T, Sousa JJ, Peres E, Morais R (2017) Uas, sensors, and data processing in agroforestry: a review towards practical applications. *Int J Remote Sens* 38(8–10):2349–2391
- Palossi D, Loquercio A, Conti F, Flamand E, Scaramuzza D, Benini L (2019) A 64mw dnn-based visual navigation engine for autonomous nano-drones. *IEEE Internet Things J* 6(5):8357–8371
- Perera AG, Al-Naji A, Law YW, Chahl J (2018) Human detection and motion analysis from a quadrotor uav. In: IOP conference series: materials science and engineering, vol 405. IOP Publishing, p 012003
- Perera AG, Law YW, Chahl J (2019) Drone-action: an outdoor recorded drone video dataset for action recognition. *Drones* 3(4):82
- Perreault H, Bilodeau GA, Saunier N, Gravel P (2019) Road user detection in videos. [arXiv:1903.12049](https://arxiv.org/abs/1903.12049)
- Perrin AF, Krassanakis V, Zhang L, Ricordel V, Perreira Da Silva M, Le Meur O (2020) Eyetrackuv2: a large-scale binocular eye-tracking dataset for uav videos. *Drones* 4(1):2
- Pestana J, Sanchez-Lopez JL, Campoy P, Saripalli S (2013) Vision based gps-denied object tracking and following for unmanned aerial vehicles. In: 2013 IEEE international symposium on safety, security, and rescue robotics (SSRR). IEEE, pp 1–6
- Pestana J, Sanchez-Lopez JL, Saripalli S, Campoy P (2014) Computer vision based general object following for gps-denied multirotor unmanned vehicles. In: 2014 American control conference. IEEE, pp 1886–1891
- Pestana Puerta J (2017) Vision-based autonomous navigation of multirotor micro aerial vehicles. Ph.D. thesis, Industriales
- Plastiras G, Kyrkou C, Theocharides T (2018) Efficient convnet-based object detection for unmanned aerial vehicles by selective tile processing. In: Proceedings of the 12th international conference on distributed smart cameras. ACM, p 3
- Plastiras G, Terzi M, Kyrkou C, Theocharides T (2018) Edge intelligence: challenges and opportunities of near-sensor machine learning applications. In: 2018 IEEE 29th international conference on application-specific systems, architectures and processors (ASAP). IEEE, pp 1–7
- Puri A (2005) A survey of unmanned aerial vehicles (uav) for traffic surveillance. Department of Computer Science and Engineering, University of South Florida, Florida, pp 1–29
- Qi Y, Wang D, Xie J, Lu K, Wan Y, Fu S (2019) Birdseyeview: aerial view dataset for object classification and detection. In: 2019 IEEE Globecom workshops (GC Wkshps). IEEE, pp 1–6
- Rahmehoonfar M, Dobbs D, Yari M et al (2019) Discountnet: discriminating and counting network for real-time counting and localization of sparse objects in high-resolution uav imagery. *Remote Sens* 11(9):1128
- Rakha T, Gorodetsky A (2018) Review of unmanned aerial system (uas) applications in the built environment: towards automated building inspection procedures using drones. *Autom Constr* 93:252–264
- Rey N, Volpi M, Joost S, Tuia D (2017) Detecting animals in african savanna with uavs and the crowds. *Remote Sens Environ* 200:341–351
- Rivas A, Chamoso P, González-Briones A, Corchado J (2018) Detection of cattle using drones and convolutional neural networks. *Sensors* 18(7):2048
- Robicquet A, Alahi A, Sadeghian A, Anenberg B, Doherty J, Wu E, Savarese S (2016) Forecasting social navigation in crowded complex scenes. [arXiv:1601.00998](https://arxiv.org/abs/1601.00998)
- Robicquet A, Sadeghian A, Alahi A, Savarese S (2016) Learning social etiquette: human trajectory understanding in crowded scenes. In: European conference on computer vision. Springer, pp 549–565
- Rozantsev A (2017) Vision-based detection of aircrafts and uavs. Tech. rep, EPFL
- Rozantsev A, Lepetit V, Fua P (2017) Detecting flying objects using a single moving camera. *IEEE Trans Pattern Anal Mach Intell* 39(5):879–892
- Ruchaud N (2015) Privacy protection filter using stegoscambling in video surveillance. In: MediaEval 2015 Workshop, Wurzen, Germany
- Saif A, Prabuwo AS, Mahayuddin ZR (2014) Moving object detection using dynamic motion modelling from uav aerial images. *Sci World J* 2014. <https://doi.org/10.1155/2014/890619>
- Saqib M, Khan SD, Sharma N, Scully-Power P, Butcher P, Colefax A, Blumenstein M (2018) Real-time drone surveillance and population estimation of marine animals from aerial imagery. In: 2018 international conference on image and vision computing New Zealand (IVCNZ). IEEE, pp 1–6
- Sarwar F, Griffin A, Periasamy P, Portas K, Law J (2018) Detecting and counting sheep with a convolutional neural network. In: 2018 15th IEEE International conference on advanced video and signal based surveillance (AVSS). IEEE, pp 1–6
- Seymour A, Dale J, Hammill M, Halpin P, Johnston D (2017) Automated detection and enumeration of marine wildlife using unmanned aircraft systems (uas) and thermal imagery. *Sci Rep* 7:45,127
- Shao W, Kawakami R, Yoshihashi R, You S, Kawase H, Naemura T (2019) Cattle detection and counting in uav images based on convolutional neural networks. *Int J Remote Sens* 41(1):31–52

- Soleimani A, Nasrabadi NM (2018) Convolutional neural networks for aerial multi-label pedestrian detection. In: 2018 21st International conference on information fusion (FUSION). IEEE, pp 1005–1010
- Sommer L, Schuchert T, Beyerer J (2018) Comprehensive analysis of deep learning based vehicle detection in aerial images. *IEEE Trans Circuits Syst Video Technol* 29(9):2733
- Song WH, Jung HG, Gwak IY, Lee SW (2019) Oblique aerial image matching based on iterative simulation and homography evaluation. *Pattern Recogn* 87:317–331
- Stahl T, Pintea SL, van Gemert JC (2019) Divide and count: generic object counting by image divisions. *IEEE Trans Image Process* 28(2):1035–1044
- Sykora-Bodie ST, Bezy V, Johnston DW, Newton E, Lohmann KJ (2017) Quantifying nearshore sea turtle densities: applications of unmanned aerial systems for population assessments. *Sci Rep* 7(1):17,690,690
- Tang Z, Liu X, Shen G, Yang B (2020) Penet: object detection using points estimation in aerial images. [arXiv:2001.08247](https://arxiv.org/abs/2001.08247)
- Tayara H, Soo KG, Chong KT (2018) Vehicle detection and counting in high-resolution aerial images using convolutional regression neural network. *IEEE Access* 6:2220–2230
- Tian J, Li X, Duan F, Wang J, Ou Y (2016) An efficient seam elimination method for uav images based on wallis dodging and Gaussian distance weight enhancement. *Sensors* 16(5):662
- Tian Y, Sun A, Wang D (2018) Seam-line determination via minimal connected area searching and minimum spanning tree for uav image mosaicking. *Int J Remote Sens* 39(15–16):4980–4994
- Tijtgat N, Van Ranst W, Goedeme T, Volckaert B, De Turck F (2017) Embedded real-time object detection for a uav warning system. In: Proceedings of the IEEE international conference on computer vision, pp 2110–2118
- Touil DE, Terki N, Medouakh S (2019) Hierarchical convolutional features for visual tracking via two combined color spaces with svm classifier. *SIViP* 13(2):359–368
- Tropicchio P, Satler M, Dabisiias G, Ruffaldi E, Avizzano CA (2015) Towards smart farming and sustainable agriculture with drones. In: 2015 International conference on intelligent environments. IEEE, pp 140–143
- Turner D, Lucieer A, Malenovsky Z, King D, Robinson S (2014) Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-uav over antarctic moss beds. *Remote Sens* 6(5):4003–4024
- Tzelepi M, Tefas A (2017) Human crowd detection for drone flight safety using convolutional neural networks. In: 2017 25th European signal processing conference (EUSIPCO). IEEE, pp 743–747
- Tzelepi M, Tefas A (2019) Graph embedded convolutional neural networks in human crowd detection for drone flight safety. *IEEE Trans Emerg Topics Comput Intell*
- Vaddi S, Kumar C, Jannesari A (2019) Efficient object detection model for real-time uav applications. [arXiv:1906.00786](https://arxiv.org/abs/1906.00786)
- van Gemert JC, Verschoor CR, Mettes P, Epema K, Koh LP, Wich S (2014) Nature conservation drones for automatic localization and counting of animals. In: European conference on computer vision. Springer, pp 255–270
- Vega A, Lin CC, Swaminathan K, Buyuktosunoglu A, Pankanti S, Bose P (2015) Resilient, uav-embedded real-time computing. In: 2015 33rd IEEE International conference on computer design (ICCD). IEEE, pp 736–739
- Vidal RG, Banerjee S, Grm K, Struc V, Scheirer WJ (2018) Ug²: A video benchmark for assessing the impact of image restoration and enhancement on automatic visual recognition. In: 2018 IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1597–1606
- VidalMata RG, Banerjee S, RichardWebster B, Albright M, Davalos P, McCloskey S, Miller B, Tambo A, Ghosh S, Nagesh S, et al (2019) Bridging the gap between computational photography and visual recognition. [arXiv:1901.09482](https://arxiv.org/abs/1901.09482)
- Walha A, Wali A, Alimi AM (2015) Video stabilization with moving object detecting and tracking for aerial video surveillance. *Multimed Tools Appl* 74(17):6745–6767
- Wang D, Luo W (2019) Bayberry tree recognition dataset based on the aerial photos and deep learning model. *J Global Change Data Discover* 3(3):290–296
- Wang J, Feng Z, Chen Z, George S, Bala M, Pillai P, Yang SW, Satyanarayanan M (2018) Bandwidth-efficient live video analytics for drones via edge computing. In: 2018 IEEE/ACM symposium on edge computing (SEC). IEEE, pp 159–173
- Wang J, Feng Z, Chen Z, George S, Bala M, Pillai P, Yang SW, Satyanarayanan M (2019) Edge-based live video analytics for drones. *IEEE Internet Comput* 23(4):27–34
- Wang P, Jiao B, Yang L, Yang Y, Zhang S, Wei W, Zhang Y (2019) Vehicle re-identification in aerial imagery: dataset and approach. In: Proceedings of the IEEE international conference on computer vision, pp 460–469

- Wang T, Xiong J, Xu X, Shi Y (2019) Senn: a general distribution based statistical convolutional neural network with application to video object detection. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol 33. pp 5321–5328. <https://doi.org/10.1609/aaai.v33i01.33015321>
- Wang X, Cheng P, Liu X, Uzochukwu B (2018) Fast and accurate, convolutional neural network based approach for object detection from uav. In: IECON 2018-44th annual conference of the IEEE industrial electronics society. IEEE, pp 3171–3175
- Wang Y, Ding L, Laganieri R (2019) Real-time uav tracking based on psr stability. In: Proceedings of the IEEE international conference on computer vision workshops Seoul, Korea (South), 2019, pp 144–152. <https://doi.org/10.1109/ICCVW.2019.00023>
- Wang Y, Luo X, Ding L, Fu S, Hu S (2018) Collaborative model based uav tracking via local kernel feature. *Appl Soft Comput* 72:90–107
- Wang Z, Liu Z, Wang D, Wang S, Qi Y, Lu H (2019) Online single person tracking for unmanned aerial vehicles: benchmark and new baseline. In: ICASSP 2019–2019 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 1927–1931
- Wei Z, Duan C (2020) Amrnet: chips augmentation in aerial images object detection. [arXiv:2009.07168](https://arxiv.org/abs/2009.07168)
- Xiang TZ, Xia GS, Zhang L (2018) Mini-uav-based remote sensing: techniques, applications and perspectives. [arXiv:1812.07770](https://arxiv.org/abs/1812.07770)
- Xiaoyuan Y, Ridong Z, Jingkai W, Zhengze L (2019) Real-time object tracking via least squares transformation in spatial and fourier domains for unmanned aerial vehicles. *Chin J Aeronaut* 32(7):1716–1726
- Xu B, Wang W, Falzon G, Kwan P, Guo L, Chen G, Tait A, Schneider D (2020) Automated cattle counting using mask r-cnn in quadcopter vision system. *Comput Electron Agric* 171:105,300
- Xu B, Wang W, Falzon G, Kwan P, Guo L, Sun Z, Li C (2020) Livestock classification and counting in quadcopter aerial images using mask r-cnn. *Int J Remote Sens*, pp 1–22
- Xu X, Zhang X, Yu B, Hu XS, Rowen C, Hu J, Shi Y (2018) Dac-sdc low power object detection challenge for uav applications. [arXiv:1809.00110](https://arxiv.org/abs/1809.00110)
- Xu Y, Ou J, He H, Zhang X, Mills J (2016) Mosaicking of unmanned aerial vehicle imagery in the absence of camera poses. *Remote Sens* 8(3):204
- Xu Y, Yu G, Wang Y, Wu X, Ma Y (2016) A hybrid vehicle detection method based on viola-jones and hog+ svm from uav images. *Sensors* 16(8):1325
- Xu Z, Wu L, Zhang Z (2018) Use of active learning for earthquake damage mapping from uav photogrammetric point clouds. *Int J Remote Sens* 39(15–16):5568–5595
- Xue X, Li Y, Dong H, Shen Q (2018) Robust correlation tracking for uav videos via feature fusion and saliency proposals. *Remote Sens* 10(10):1644
- Xue X, Li Y, Shen Q (2018) Unmanned aerial vehicle object tracking by correlation filter with adaptive appearance model. *Sensors* 18(9):2751
- Yang MY, Liao W, Li X, Cao Y, Rosenhahn B (2019) Vehicle detection in aerial images. *Photogramm Eng Remote Sens* 85(4):297–304
- Yeh MC, Chiu HK, Wang JS (2016) Fast medium-scale multiperson identification in aerial videos. *Multimed Tools Appl* 75(23):16117–16133
- Yin X, Wang X, Yu J, Zhang M, Fua P, Tao D (2018) Fisheyerecnet: a multi-context collaborative deep network for fisheye image rectification. In: Proceedings of the European conference on computer vision (ECCV), pp 469–484
- Yu H, Li G, Zhang W, Huang Q, Du D, Tian Q, Sebe N (2020) The unmanned aerial vehicle benchmark: object detection, tracking and baseline. *Int J Comput Vis* 128(5):1141–1159
- Yuan C, Zhang Y, Liu Z (2015) A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Can J For Res* 45(7):783–792
- Zarco-Tejada PJ, Diaz-Varela R, Angileri V, Loudjani P (2014) Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (uav) and automatic 3d photo-reconstruction methods. *Eur J Agron* 55:89–99
- Zhang P, Zhong Y, Li X (2019) Slimyolov3: narrower, faster and better for real-time uav applications. In: Proceedings of the IEEE international conference on computer vision workshops
- Zhang R, Shao Z, Huang X, Wang J, Li D (2020) Object detection in uav images via global density fused convolutional network. *Remote Sens* 12(19):3140
- Zhang S, Zhang Q, Yang Y, Wei X, Wang P, Jiao B, Zhang Y (2020) Person re-identification in aerial imagery. *IEEE Trans Multimed* 23:281–291. <https://doi.org/10.1109/TMM.2020.2977528>
- Zhang W, Liu C, Chang F, Song Y (2020) Multi-scale and occlusion aware network for vehicle detection and segmentation on uav aerial images. *Remote Sens* 12(11):1760
- Zhang W, Song K, Rong X, Li Y (2018) Coarse-to-fine uav target tracking with deep reinforcement learning. *IEEE Trans Autom Sci and Eng* 16(4):1522–1530

- Zhu J, Chen S, Tu W, Sun K (2019) Tracking and simulating pedestrian movements at intersections using unmanned aerial vehicles. *Remote Sens* 11(8):925
- Zhu J, Sun K, Jia S, Li Q, Hou X, Lin W, Liu B, Qiu G (2018) Urban traffic density estimation based on ultrahigh-resolution uav video and deep neural network. *IEEE J Sel Topics Appl Earth Obs Remote Sens* 11(12):4968–4981
- Zhu P, Sun Y, Wen L, Feng Y, Hu Q (2020) Drone based rgbt vehicle detection and counting: a challenge. [arXiv:2003.02437](https://arxiv.org/abs/2003.02437)
- Zhu P, Wen L, Bian X, Ling H, Hu Q (2018) Vision meets drones: a challenge. [arXiv:1804.07437](https://arxiv.org/abs/1804.07437)
- Zhu, P., Wen, L., Du, D., Bian, X., Hu, Q., Ling, H (2020) Vision meets drones: past, present and future. [arXiv:2001.06303](https://arxiv.org/abs/2001.06303)
- Zhu P, Wen L, Du D, Bian X, Ling H, Hu Q, Wu H, Nie Q, Cheng H, Liu C, et al (2018) Visdrone-vdt2018: the vision meets drone video detection and tracking challenge results. In: Proceedings of the European conference on computer vision (ECCV)
- Zhu P, Zheng J, Du D, Wen L, Sun Y, Hu Q (2020) Multi-drone based single object tracking with agent sharing network. [arXiv:2003.06994](https://arxiv.org/abs/2003.06994)
- Zimmermann K, Matas J, Svoboda T (2009) Tracking by an optimal sequence of linear predictors. *IEEE Trans Pattern Anal Mach Intell* 31(4):677–692

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.