Data in Brief 65 (2026) 112448

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

VisioDECT: A robust dataset for aerial and scenario based multi-drone detection, identification, and neutralization

Simeon Okechukwu Ajakwe[a,*], Vivian Ukamaka Ihekoronye[b], Golam Mohtasin[a], Rubina Akter[c], Jae Min Lee[d], Dong Seong Kim[d,*]

[a] *Information and Communication Technology Convergence Research Centre, Kumoh National Institute of Technology, Gumi, South Korea*
[b] *Department of Computer and Information Sciences, Towson University, Towson, MD, USA, 2152*
[c] *Schulich School of Engineering, Department of Electrical and Software Engineering, University of Calgary, Calgary, AB, Canada*
[d] *Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea*

ARTICLE INFO

Article history:
Received 2 October 2025
Revised 28 November 2025
Accepted 2 January 2026
Available online 8 January 2026

Keywords:
Anti-drone
Deep Learning
Detection
Drone
Security
Surveillance
UAV

ABSTRACT

The rapid proliferation of unmanned aerial vehicles (UAVs) for logistics, surveillance, and civilian applications continues to pose significant challenges to airspace security, particularly through unauthorized or malicious deployments. Existing UAV datasets are limited in scope, often focusing on single-drone scenarios, synthetic imagery, or restricted environmental conditions, thereby constraining the development of robust counter-UAV systems. To bridge these gaps, we present vision-based drone detection dataset named as **VisioDECT**, a comprehensive and scenario-rich dataset for multi-drone detection, identification, and neutralization. The dataset comprises 20,924 annotated images and labels from six UAV models (Anafi-Extended, DJI FPV, DJI Phantom, EFT-E410S, Mavic Air 2, and Mavic 2 Enterprise), captured across three distinct scenarios (sunny, cloudy, and evening) at varying altitudes (30–100 m) and distances. Importantly, all UAVs included in this dataset are rotary-wing (multirotor) platforms, which dominate low-altitude airspace and are the most commonly encountered in real-world surveillance and

---

* Corresponding authors.
E-mail addresses: simeon.ajakwe@kumoh.ac.kr (S.O. Ajakwe), dskim@kumoh.ac.kr (D.S. Kim).
Social media: @AjakweSimeon (S.O. Ajakwe)

https://doi.org/10.1016/j.dib.2026.112448

counter-UAV scenarios. Data were collected over 20 months from more than 12 locations in South Korea, ensuring diversity in illumination, weather, and background complexity. Each sample is provided in three standard formats (.txt, .xml, .csv), with detailed metadata and quality-verified annotations for detection and classification tasks. Illustrative benchmark evaluations using state-of-the-art detection models (e.g., DRONET, YOLO variants) are included solely to validate the quality and practical usability of the dataset for real-time drone defense research. VisioDECT provides a standardized, reproducible, and scalable resource that enables benchmarking, model training, and evaluation for airspace surveillance, UAV traffic management, and national security applications.

Specifications Table

Subject	Engineering & Materials science
Specific subject area	Computer Vision, Drone Security, Autonomous Systems, Multi-drone detection and identification (supporting pre-neutralization decision-making in counter-UAV workflows).
Data Format	Raw and Processed Data – JPEG image frames (raw) and annotated labels in TXT, XML, and CSV formats (processed).
Data collection	Data were collected over 20 months at 12 outdoor sites in South Korea using six UAV models (Anafi-Extended, DJI FPV, DJI Phantom, EFT-E410S, Mavic Air 2, Mavic 2 Enterprise) flown at 30–100 m under sunny, cloudy, and evening scenarios. Videos were captured with a high-definition digital camera (HD 1080p) on tripod, converted to JPEG frames (852 × 480) via Video to JPG Converter. Frames without drones were excluded. Annotation was performed with MakeSense.ai, exporting labels in TXT, XML, and CSV formats.
Data source location	The dataset was collected at more than 12 outdoor test sites across South Korea, such as Gumi, Daegu, Pohang, and surrounding regions, under varied atmospheric and is stored and illumination conditions (sunny, cloudy, and evening).
Data accessibility	(1) IEEE Dataport: VisioDECT Dataset: An Aerial Dataset for Scenario-Based Multi-Drone Detection and Identification Data Identification Number: 10.21227/n27q-7e06 Direct URL to data: https://dx.doi.org/10.21227/n27q-7e06 (2) Kaggle: VisioDECT for Scenario-Based Multi-Drone Detection URL: https://doi.org/10.34740/kaggle/dsv/13155551
Related research article	DRONET [1]

Note: The dataset does not include any direct neutralization mechanisms, its detection and identification capabilities form essential inputs for decision-making modules within counter-UAV and threat-mitigation systems.

1. Value of the Data

- **Benchmark contribution**: VisioDECT provides 20,924 manually annotated images of six UAV models captured under diverse illumination conditions (sunny, cloudy, and evening) and varied outdoor environments. This large-scale, scenario-rich imagery offers a standardized benchmark for evaluating and developing robust UAV detection and identification models. The dataset supports reliable performance assessment across easy, moderate, and challenging visual conditions, making it a valuable resource for researchers working on vision-based counter-UAV systems.

- **Diversity and realism:** VisioDECT captures UAV imagery across three illumination scenarios (sunny, cloudy, evening), multiple outdoor environments, and a wide range of altitudes (30–100 m). These conditions reflect real-world surveillance challenges and operational variability, providing realistic samples for developing algorithms that must perform reliably in practical counter-UAV settings such as airport monitoring, critical-infrastructure protection, and airspace security. Also, it ensures distance-based analysis, as altitude annotations enable research into detection performance degradation as a function of distance, critical for establishing operational ranges of defence systems. The dataset's realistic scenes, therefore, directly support downstream practical applications without requiring a separate application-focused bullet. Unlike existing benchmarks such as VisDrone, DroneRF, etc., that capture a single environmental condition, VisioDECT explicitly emphasizes scenario diversity (sunny, cloudy, and evening scenarios), enabling evaluation of environmental robustness. Finally, with six distinct drone models systematically captured, the dataset enables fine-grained identification research beyond simple detection to multi-drone identification.

2. Background

The rapid growth of unmanned aerial vehicles (UAVs) for commercial, recreational, and industrial purposes has created new challenges for airspace safety and surveillance [2,3]. UAV misuse, including unauthorized flights near airports and sensitive facilities, has highlighted the need for reliable detection and neutralization systems [3–6]. Existing datasets for UAV research [7–14] are often limited in scope, focusing on single-drone operations, synthetic imagery, or constrained environmental conditions, which reduces their applicability for real-world counter-UAV system development.

The VisioDECT dataset was compiled to address these gaps by providing a comprehensive collection of annotated images of multiple drone models under diverse operational scenarios. The data were generated through systematic flight tests of six (6) UAVs under varying weather and illumination conditions at multiple outdoor locations in South Korea. All UAVs included in these tests are rotary-wing (multirotor) platforms, selected because they dominate low-altitude civilian, industrial, and unauthorized-airspace operations and represent the primary targets of vision-based counter-UAV research. High-definition cameras were used to capture video sequences, later converted into images and manually annotated with bounding boxes using standardized labelling tools.

This dataset builds on prior research such as the DRONET framework [1], ALIEN [15], and iBANDA [16], which explored real-time UAV detection, by providing a larger and more scenario-rich source of training and benchmarking data. This Data in Brief article supplements that work by documenting the dataset design, methodology, and structure, thereby ensuring accessibility and reproducibility for the broader research community.

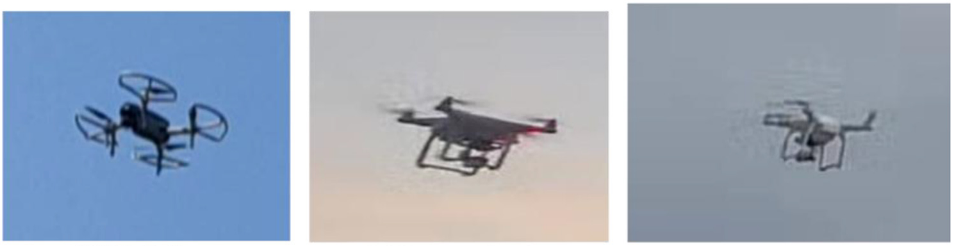
3. Data Description

The VisioDECT dataset is a comprehensive collection of aerial imagery specifically developed to support cutting-edge research in multi-drone detection, identification, and neutralization. This section describes the development process of the dataset, beginning with the initial system setup and culminating in the meticulous organization of the data. Additionally, it offers comprehensive specifications detailing the dataset's structure and contents.

3.1. Dataset specifications, structure, and development methodology

The VisioDECT dataset comprises 20,924 annotated image samples extracted from multi-UAV surveillance operations, using six (6) distinct unmanned aerial vehicles (UAVs): Anafi-Extended,

Anafi-Extended-Sunny **DJI-FPV-Cloudy** **EFT-E410S-Sunny**



MAVIC-Air-Sunny **DJI-Phantom-Evening** **DJI-Phantom-Cloudy**

Fig. 1. Samples of Captured Drone Models at different locations, heights, and weather conditions.

DJI FPV, DJI Phantom, EFTE410S, Mavic 2 Air, and Mavic 2 Enterprise [21] (see Fig. 1). These six UAV models were selected to cover a diverse range of commercially relevant platforms, including hobby-grade, professional, enterprise, and industrial drones. Their differing physical sizes, propulsion characteristics, and visual signatures provide a representative variety of UAV appearances needed for robust model development. The Parrot Anafi-Extended is a lightweight (320 g) foldable quadrotor with a 21 MP 4K/30 frame-per-second (fps) camera, 55 km/h top speed, and 25 min endurance; the DJI FPV racing-style quadcopter (795 g) offers Manual mode speeds up to 140 km/h, 4K/60 fps video, and 20 min flight time; the DJI Phantom 4 Pro V2.0 (1 388 g) features a 1 20 MP sensor, 72 km/h maximum speed, and 30 min endurance; the EFT E410S agricultural quadcopter (6 kg empty, 26 kg MTOW) supports 10 L payloads, 15–35 min missions, and IP67-rated components; the DJI Mavic Air 2 (570 g) delivers a 48 MP 4K/60 fps camera, 68 km/h top speed, and 34 min flight time; and the DJI Mavic 2 Enterprise (900 g–1 100 g MTOW) offers 12 MP imaging, optional thermal modules, 72 km/h maximum speed, and 31 min endurance. These platforms represent diverse aerodynamic configurations, propulsion systems, and electromagnetic signatures, providing comprehensive coverage of contemporary UAV technologies employed in both civilian and potentially hostile applications.

The dataset was gathered over 20 months at twelve outdoor test sites using six commercial drones, their respective controllers, a smart mobile device equipped with the UAV controller, high-definition digital cameras, and a tripod for mounting a smartphone to capture videos. The dataset encompasses three primary environmental scenarios designed to simulate realistic operational conditions: sunny, cloudy, and evening at altitudes ranging from 30 m to 100 m and varying standoff distances [1]. The altitude range of 30–100 m was chosen to reflect realistic UAV operating conditions. At lower altitudes, drones exhibit larger and clearer visual profiles, while higher altitudes introduce reduced pixel occupancy and greater detection difficulty. This range ensures that the dataset includes both easy and challenging visual conditions, enabling researchers to design detection algorithms that remain effective across varying operational distances. This multi-scenario approach ensures robust model training capabilities across various lighting conditions and atmospheric variations, thereby preventing class imbalance. The data cleaning phase involved rigorous quality assessment procedures, including frame extraction from video sequences, noise reduction, and standardization of image resolutions to 852 × 480 pixels in JPEG format. Subsequently, comprehensive annotation processes were implemented to generate accurate bounding box labels and metadata for each captured drone instance. Expert annotators performed manual quality control on each image repository, removing any frames lacking visible UAVs. Subsequently, trained labelers drew bounding boxes around each drone instance and exported the annotations in three standard formats (TXT, XML, and CSV) [22] (See Fig. 6). The VisioDECT dataset is organized into six primary directories, each corresponding to one of the six drone models. Within each model directory, there are two subdirectories: one for images and one for labels. The image subdirectory contains three scenario-specific folders, sunny, cloudy, and evening, each housing JPEG-formatted image files. The label subdirectory mirrors this structure, providing annotations in TXT, CSV, and XML formats for each scenario. Each annotation includes precise bounding box coordinates, drone model identification, scenario classification, and neutralization outcome indicators where applicable. The metadata structure incorporates flight trajectory information, altitude specifications, ground speed measurements, and environmental detection conditions. This hierarchical arrangement facilitates the integration of the dataset into classification, object detection, and broader computer vision workflows, leveraging various state-of-the-art AI frameworks.

3.2. Sampling framework and site selection

Data was collected through a randomized sampling approach within Gumi City, covering densely populated, urban, and remote areas to ensure environmental diversity and geographic representativeness. This framework supports the dataset's generalizability and relevance to real-world UAV detection scenarios. A total of 20,924 images were captured at altitude intervals of 10 meters between 30 m and 100 m, with 78% of samples obtained at lower altitudes (30–60 m), reflecting realistic operational threats rather than sampling convenience. Temporal sampling spanned November 2019 to June 2021, encompassing 20 months of data collection. Seasonal variations were transparently reported, with 18% fewer samples in winter and 32% in summer, attributed to regional weather constraints. Geographic distribution ensures that detection models trained on VisioDECT demonstrate generalizability across diverse airspace environments while complying with South Korean airspace restrictions and safety regulations. Annotations were conducted by three independent annotators with years of experience in computer vision and image processing. For bounding box conflicts, consensus was achieved through discussion-based resolution, with final validation by a senior domain expert. Annotation inclusion criteria require at least 60% of the drone area to be visible, ensuring sufficient visual information for the detection model's training. This threshold was selected based on a preliminary analysis showing a degradation in detection accuracy below 35% visibility. Table 1: Distribution of images across altitude ranges with operational justification. Lower altitudes prioritized to reflect realistic unauthorized drone threat zones for infrastructure security applications.

Table 1
Altitude distribution of image samples.

Altitude Range (m)	Image Count	Percentage	Operational Justification
30-40	6,689	32%	Primary threat zone
40-50	5,856	28%	Primary threat zone
50-60	3,769	18%	Secondary threat zone
60-70	2,511	12%	Extended monitoring range
70-80	1,047	5%	Surveillance altitude
80-90	628	3%	High altitude operations
90-100	420	2%	Maximum collection range
Total	20,924	100%	Overall operational purpose

Fig. 2. Data capturing system highlighting different equipment used for data gathering.

4. Experimental Design, Materials, and Methods

4.1. Data collection and scenario system setup

The dataset is generated by flying 6 different drone models under different climatic conditions (sunny, cloudy, and evening) at varying altitudes. The data-capturing operation was carried out by a team of scientists and volunteers at more than 12 different locations, at different time intervals of the day, and over a period of 20 months. The data capturing equipment includes drone models, drone controllers, a mobile phone with a controller application, high-definition digital cameras, and a tripod stand, as shown in Fig. 2. During data acquisition, comprehensive telemetry was recorded including GPS trajectory, altitude, ground speed, heading, and camera gimbal angle from onboard flight control systems. The publicly available annotation files (Table 2) provide bounding box coordinates and class labels in three formats (txt, .xml, .csv) optimized for detection model training. Raw telemetry data is preserved and available to researchers requesting detailed trajectory or altitude analysis by contacting the corresponding author. For example, DJI Phantom: estimated 22 × 18-pixel footprint at 100m, identifiable by

Table 2
Captured drone model specification.

Drone Types and Configurations			
Sno.	Drone Name	Model	Configuration
1	Anafi-Extended	Parrot	Max. Ascent speed:4m/s; Flight Ceiling: 4345m; Wind Resistance:50 km/h
2	DJI FPV	V2	Max.Ascent speed:15 m/s; lens:FOV:150/14.66 mm; Wind Resistance:39-49km/h
3	DJI-Phantom	Phantom 4	Max.Ascent speed:6m/s; Range:7km; Sensors:5vision sensors
4	EFT-E410S	E410s 10L	Flight speed:0-12m/s; Max.weight:25 kg; Max.Thrust:15kg
5	Mavic2-Air	Mavic2-Air	Max.Ascent speed:4m/s; weight:570g; Wind Resistance:29-38km/h
6	Mavic2-Enterprise	Zoom	Max.Ascent speed: 72km/h; weight:1100g; Wind Resistance:29-38km/h

symmetric quad-rotor configuration, characteristic landing gear visibility, and boxy fuselage profile.

The captured drone models are Anafi Extended, DJI FPV, DJI Phantom, EFT-E410S, Mavic2-Air, and Mavic2-Enterprise Zoom. The choice of these drone models is due to their wide-range usage for varied purposes, especially in carrying out nefarious activities in the airspace. Also, the need to consider both hobbyist and industrial-based drones of varied sizes justifies the inclusiveness of these drones in the sample space. Table 2 provides a succinct detail of the drone model's specification.

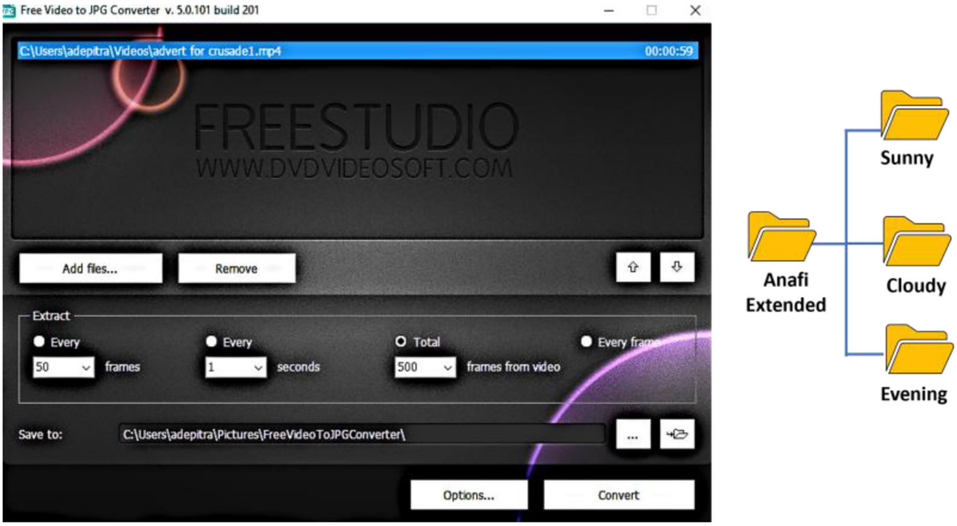
All flights were conducted following predetermined grid-based patterns optimized for diverse viewing angle acquisition. Each flight mission maintained consistent altitude in 10-meter increments (30m, 50m, 70m, 90m, or 100m). At each altitude, the imaging system executed systematic heading changes at 45-degree intervals (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°), with 60-second hold at each heading to accumulate temporal samples. Horizontal velocity during transitions was maintained at constant 5 m/s to ensure consistent motion characteristics beneficial for motion-based detection model training. Complete flight coverage including all eight headings at a given altitude required approximately 12 minutes, enabling approximately 1,440 frames (72 frames per heading at 2 fps extraction rate) per altitude level.

For the scenario design, each of the drone models is flown into the airspace at altitudes ranging between 30 m and 100 m at various degrees and speeds with the aid of the controller and a mobile phone. While the flight operation is ongoing, the video-capturing process is conducted simultaneously with the aid of a high-definition digital camera mounted on a tripod stand. The time span for each captured video frame is between 5-10 minutes, covering the various altitudes, heights, rotational angles, and speeds of the drone in the airspace, and at various locations. This ensured sample variability and data representativeness for testing model robustness. To enhance video capturing quality, the deployed high-definition digital camera is mounted on a tripod stand and operated accordingly. The video capturing was conducted in sunny (bright), cloudy (hazy), and dark (evening) climatic conditions. Thereafter, the captured video frame was edited and transformed into a usable format using video editing software for data preprocessing. Although the drones were flown at varying standoff distances relative to the camera, the exact camera-to-drone distance was not recorded as metadata in this current dataset release. The variations in drone size within the image frames, therefore, implicitly reflect changes in distance, altitude, viewpoint, and camera zoom. Nonetheless, future updates of the VisioDECT dataset will incorporate explicit distance information extracted from synchronized flight telemetry and fixed camera geolocation to support research in distance detection and long-range UAV tracking.

4.2. Data preprocessing: Conversion, annotation, and cleaning

Transforming the captured video data into a usable format is performed in three (3) phases, namely: video frame to raw data frame (image) conversion, data frame cleaning and labeling, and dataset organization. First, the edited video frames are fed into Video to JPG Converter [17] application, where the video sequences are converted into a series of data frames in the form of raw Joint Photographic Expert Group (JPEG) images. To achieve this, JPEG images of the same size (aspect ratio) are extracted in a specified number of frames and seconds from each captured video. This forms the total number of raw images per scenario of each edited captured video file fed into the application. Then, these image files are stored in repositories for further preprocessing as shown in Fig. 3.

Video footage was captured at 1920 × 1080 resolution, 30 fps native capture rate. Frames were extracted at a 2-fps sampling rate (one frame every 15 video frames) to provide temporal diversity while maintaining a manageable dataset size. This sampling strategy yielded approximately 10,800 candidate frames per 90-minute flight, with subset selection for annotation as detailed in Table 3. Image resolution standardization: Original 1920 × 1080 frames were downsampled to 852 × 480 pixels. Drones included in the annotation set met the following criteria: (1) minimum 70% of the bounding area visible within the image frame; (2) drone clearly distin-



Free Video to JPG Converter v. 5.0.101 build 201

C:\Users\adepitra\Videos\advert for crusade1.mp4 00:00:59

Extract

● Every 50 frames

● Every 1 seconds

○ Total 500 frames from video

● Every frame

Save to: C:\Users\adepitra\Pictures\FreeVideoToJPGConverter\

Fig. 3. Video sequence transformation to drone images, highlighting the file structure of the Anafi – Extended drone with sunny, cloudy, and evening scenario drone images.

Table 3
Technical specifications for image processing and annotation.

Parameter	Value
Source resolution	1920 × 1080 px, 30 fps
Frame extraction rate	2 fps (every 15 frames)
Output resolution	852 × 480 px
Downsampling method	Bicubic interpolation (Lanczos kernel)
Partial drone threshold	≥40% visible area
Annotation formats	.txt, .xml, .csv
Class labels	Drone model names + generic "drone"

guishable from background or other objects; (3) no severe motion blur exceeding visual distinctiveness. Frames with drones partially occluded by clouds but identifiable (e.g., rotor blur visible) were included. This threshold ensures training images contain sufficient visual information for model learning while avoiding extremely degraded samples.

Thereafter, the drone images for each scenario are manually cleaned to elicit image frames without drones in the background, as well as frames with partial drone images (defective sample images). To minimize the number of image frames without drones in the background and defective samples, the digital camera handler maintains communication with the drone controller during video capture. This is to ensure that the flown drones are within a specified boundary range of the video frame (up and down, left and right). This process helps to speed up the data labeling phase, thereby minimizing the likelihood of errors in the dataset. After cleaning the images, per scenario, the data labeling phase commences.

Labeling the captured data entails annotating or assigning labels or tags (in this context, ground truth values) to raw images using annotation tools. The tags represent the unique class of objects each drone image or data belongs to, and its corresponding values. This helps the AI model to distinctly identify and classify each class of object during pattern discovery and inference. Make Sense AI, an image annotation tool, is used [18] to annotate the captured drone images. To avert the problem of inconsistent labeling of data, a consensus label file or annotation dictionary was created, which contains the specifics (nomenclature, color codes, etc.) of each label. Each scenario's repository (for instance, Sunny for Anafi-extended drone) containing

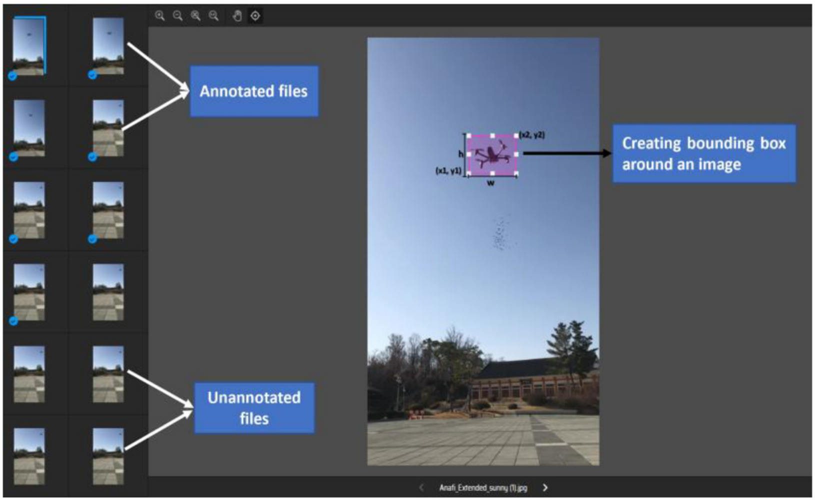


Fig. 4. Labelling of Anafi – Extended Sunny scenario drone images, highlighting the annotated files, a bounding box around a drone in the airspace, and unannotated files of Anafi-Extended sunny scenario.

the image files is fed into the Make Sense environment. Then, the annotation class file, data.txt contains the class list, which is imported into the same application environment. Image annotation was performed manually by drawing bounding boxes around each image. A bounding box is a rectangular box that is used to define the location of a target object on a background. The location of the target object is determined by either two coordinates of the x and y axis (x_1,y_1) (x_2,y_2), or by one coordinate (x_1,y_1) of the upper left-corner and lower- right of the box as well as the height (h) and width (w) of the bounding box (see Fig. 4). Each annotated file is saved automatically in cache memory with its corresponding ground truth values. This process continues until all image files have been annotated, as shown in Fig. 4. On completion of the annotation, the annotated files are exported and saved in different repositories corresponding to the various file formats: text, extensible markup language, and comma-separated values corresponding to txt, xml, and csv, respectively. Subsequently, other scenario files are annotated, adopting the same steps.

4.3. Data Validation & Quality Assurance

To ensure that the eventual output is a quality and reliable dataset, quality assurance and control measures were carried out on each annotated and image repository. This is done by carefully scanning the entire annotated files with their corresponding raw image files to ensure that each annotation file has a corresponding image file. This scanning and cleaning process helps in error minimization during model training, testing, and validation using the final dataset. Hence, the resultant dataset contains an equal number of image files (images) and corresponding annotation files (labels) per drone scenario, as shown in Fig. 5. Furthermore, the final dataset has undergone extensive review and has been validated by UAV experts, computer vision researchers, and engineers for accuracy and robustness [19,1,15].

4.4. VisioDECT dataset characterisation and strucutre

The final output, VisioDECT, is a vision-based dataset that contains all the information needed by various object detection models in conducting feature extraction to identify, classify, and differentiate drones in the airspace from other aerial objects based on distinct characteristics. In

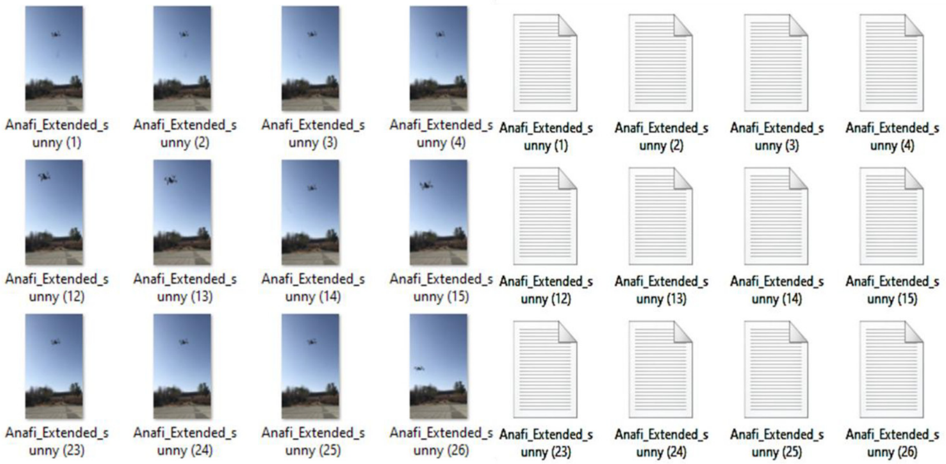


Fig. 5. Sample of final dataset highlighting some image files and their corresponding annotation files of Anafi – extended Sunny Scenario.

Table 4
VisioDECT dataset distribution, highlighting the sample variability, representatives, and size.

Drone Models	Scenario			Samples Totals
	Cloudy	Sunny	Evening	
Anafi-Extended	1200	1071	1200	**3471**
DJI-FPV	1200	1200	1200	**3600**
DJI-Phantom	900	1200	1200	**3300**
ETF-E410S	1200	1200	1200	**3600**
Mavic2-Air	1165	1191	1165	**3521**
Mavic2 Enterprise Zoom	1188	1142	1102	**3432**
Totals	**6853**	**7004**	**7067**	**20924**

most deep learning object detection algorithms, input files used for training and testing of models are usually in.txt.xml, or .csv format. Hence, the dataset is available in three formats to serve all object detectors, including single-shot, dual-shot, zero-shot, and other hybrid models that use text, Pascal Visual Object Class (VOC), and comma-separated values files as input. The VisioDECT contains 20,924 (see Table 2) samples of drone images and annotation files, which are drawn from 6 different drone models as mentioned in Section 3.1. These 6 drone models represent 6 folders with corresponding names, for instance, Anafi–Extended. Each folder contains both the image (images) and annotation (labels) subfolders, as shown in Fig. 6.

Each folder contains the flown drone model's scenarios. Each of these scenarios corresponds to the climatic conditions under which the drones are flown, sunny, cloudy, and dark (evening) during the different time intervals of the day. Each of the scenarios contains the files that store the metrics in three (3) formats: txt, voc, csv, for model training. The values in Table 4 show the distribution of the various classes of the dataset. It reveals that all the data samples are sufficiently represented to avoid errors caused by non-representative data during model training using the dataset. This representative sample distribution covers the range of drone altitudes, ranging between 30m and 100m during the data capture. This further helps in subjecting any drone detector to robust training and validation covering all possible real-life scenarios, which are necessary during the practical implementation of the developed model.

To access the image files, the name of each file (image and label) follows the pattern, "drone model-scenario (number)", where "drone model" is the name of the drone model, "scenario" is

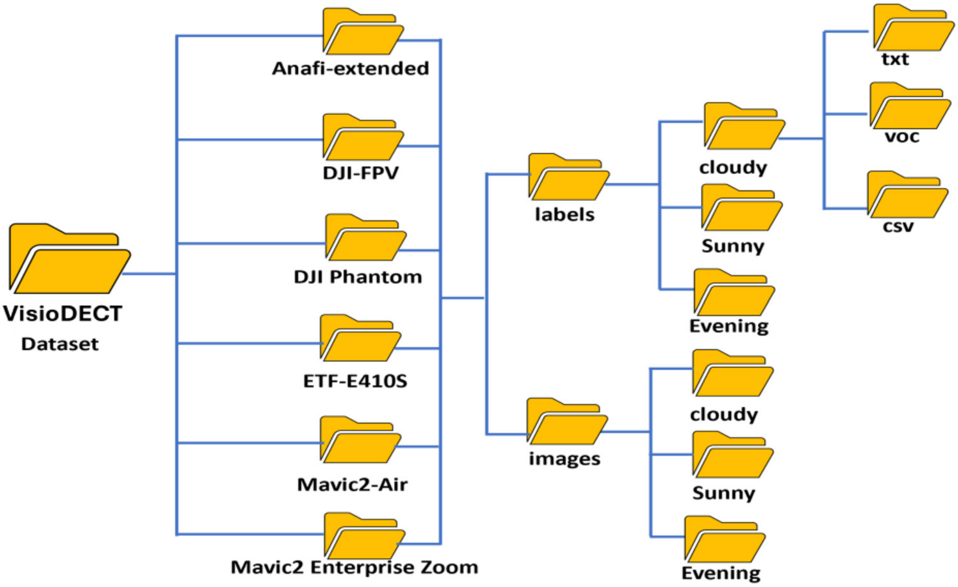


Fig. 6. Dataset structure showing the drone models, scenarios, and their corresponding images and annotations in different file formats.

Table 5
Dataset file description (CSV File).

CSV File Description	
Column Header	Meaning
A	Drone model
B	Xmin
C	Ymin
D	breath
E	length
F	filename
G	width
H	height

the climatic condition, and “(number)” is the file number. For instance, a file with the name “Mavic2-Air-Cloudy (1)” represents a Mavic2 – Air drone flown in the airspace in a cloudy condition with 1 as the unique file number. Each image file is in JPEG format with a dimension of 852 × 480 pixels.

To access the annotation file, the corresponding annotation files (for instance, Anafi-Extended-Cloudy) have contents/values that are characteristic of the image file structure. These values define the complexity of the dataset. For instance, the CSV file for an Anafi – extended Cloudy scenario has its contents arranged in rows and columns. Each row contains each drone’s scenario file record, and the total number of rows corresponds to the number of labeled objects. The columns contain the specific values of each drone scenario file record as summarized in Table 5. The values of columns B to E, G, and H in Table 4 are those of the bounding box drawn around the target object, i.e, the captured drone. These values correspond to the content of each Pascal VOC file, though in a different format with further details for the bounding box, such as “xmin-top-left”, “ymin-top-left”, “xmax-bottom-right”, and “ymax-bottom-right”, as well as the file “path” and “source”. Similarly, the “.txt” file, which is mostly used by You Only Look Once

Table 6
VisioDECT performance across scenarios.

Uav Models	*DRONET Model Classification Detection and Performance*				
	Scenario	mAP (%)	Sensitivity (%)	Specificity (%)	G-mean (%)
Anafi-Ext	sunny	99.50	75.00	49.86	61.15
	evening	99.50	100.00	46.10	67.89
	cloudy	99.50	85.00	45.20	61.98
DJI-Phantom	sunny	99.50	100.00	47.21	68.70
	evening	94.50	95.00	49.71	68.72
	cloudy	99.50	100.00	46.13	67.91
DJI-FPV	sunny	99.50	100.00	44.85	66.97
	evening	24.90	100.00	47.50	68.92
	cloudy	99.50	100.00	44.15	66.45
EFTE410S	sunny	99.50	100.00	46.12	67.91
	evening	44.60	75.00	52.25	62.59
	cloudy	90.80	100.00	49.13	70.09
Mavic-Ent	sunny	80.00	82.30	48.50	63.17
	evening	66.70	61.60	49.68	55.48
	cloudy	81.00	82.50	43.75	60.08
Mavic-Air	sunny	99.50	100.00	44.12	66.42
	evening	12.00	100.00	49.95	70.67
	cloudy	99.50	15.00	45.12	26.01

Fig. 7. (Left:) Detection performance of various object detection models on VisioDECT. (Right:) Drone Detection samples by DRONET for various UAVs at different heights and scenarios [16].

(YOLO) contains annotations that correspond to the object class (index), object coordinates, and height and width of each image file.

The result in Table 6 captures the detection performance of DRONET on 34.4% samples of VisioDECT, i.e., 7200 out of 20924 dataset, [1, 21]. Furthermore, Fig. 7 captures the performance of different object detection models on the VisioDECT based on average precision, sensitivity, and F1-score (balanced accuracy measure).

Table 7
Quantitative comparison of VisioDECT with existing UAV detection datasets.

Dataset	Total samples	Modality	Drone Models	Scenarios	Altitude (m)	Annotations	Formats	Resolution	Primary Task
VisioDECT	20,924 images	RGB visual	6 models	3 weather conditions (sunny dark, cloudy), diverse heights and distances (30m-100m), and multiple environmental locations	30-100	20,924 bounding boxes	3	852 × 480	Scenario-based detection, identification
VisDrone	271,117 images/frames	RGB visual	DJI, Parrot, others	14 Chinese cities, diverse urban/r rural	20-60	2.6M bounding boxes	1	480 × 360 to 2000 × 1500px	Detection, tracking, crowd counting
DroneRF	227 segments, 2.54B samples	RF signal s (2.4 GHz)	Parrot Bebo, Parrot AR, DJI Phantom	4 flight modes, lab-controlled	1 meter (hovering)	RF signature labels	1	N/A (signal data)	RF-based detection, Classification, identification
Anti- UAV	318 video pairs	RGB Ther mal	DJI, Parrot	Day/night	Not specified	580k+ bounding boxes	1	1920 × 10 80 (Full HD)	Multi-modal tracking, detection

These results attest to the quality, veracity, validity, and authenticity of the VisioDECT as a robust dataset for drone detection, identification, localization, and neutralization for safe and secure UAV navigation, surveillance, and logistics. The VisioDECT is a robust vision-based drone detection dataset that spans popular drone models. It took into consideration the drone's real-time operation scenarios (sunny, cloudy, and dark), and other characteristics needed for comprehensive and scalable model training and validation. This is necessary to interface with emerging drone technologies. The detailed dataset generation process and methodology provide a scientific, verifiable, and transparent approach needed to reproduce similar work. Though the VisioDECT contains 20,924 annotated images of only six (6) different UAV models, they are captured under three diverse scenarios at varying altitudes and distances and distributed across multiple file formats for broader accessibility. This unique characteristic is rare in related datasets. Unlike prior datasets, VisioDECT emphasizes both diversity and realism, incorporating cluttered backgrounds, dynamic environmental conditions, and multiple perspectives to simulate real-world surveillance contexts. This design ensures that algorithms trained on VisioDECT are better equipped to handle variations encountered in operational deployments. In the future, we hope to periodically update the dataset to include more drone models and other aerial objects, especially drones with payloads.

Summarily, the VisioDECT addresses specific research gaps in multi-drone detection. (1) Scenario diversity - existing benchmarks (such as VisDrone, DroneRF, etc.) typically capture single environmental conditions, while VisioDECT explicitly documents performance across sunny, cloudy, and evening scenarios, enabling evaluation of environmental robustness. (2) Distance-based analysis - altitude annotations enable research into detection performance degradation as a function of distance, critical for establishing operational ranges of defense systems. (3) Multi-drone identification - with six distinct drone models systematically captured, the dataset enables fine-grained identification research beyond simple detection. Table 7 captures the quantitative difference between VisioDECT and related datasets.

Limitations

VisioDECT provides a robust, vision-based dataset for drone detection, identification, and neutralization, comprising 20,924 annotated images from six UAV models under three environmental scenarios (sunny, cloudy, and evening). By integrating diverse altitudes, backgrounds, and atmospheric conditions, the dataset addresses major gaps in existing UAV resources and enables realistic benchmarking for counter-UAV research. Its multi-format structure (.txt, .xml, .csv) and consistent annotations make it adaptable across object detection frameworks, including single-shot, multi-shot, and transformer-based models. Nevertheless, VisioDECT has certain limitations. The current release includes only six UAV models and does not cover payload-equipped drones, thermal imaging, or radar/RF modalities. While sufficient for vision-based detection and classification tasks, its scope can be further expanded to represent broader UAV ecosystems and adversarial scenarios. Also, VisioDECT does not explicitly include camera–drone distance metadata. Hence, the absence of calibrated distance values limits its direct use in distance-aware detection research. Therefore, future VisioDECT release will integrate camera geolocation to generate distance annotations for each frame. In conclusion, VisioDECT establishes a reproducible, standardized, and high-quality foundation for developing and evaluating UAV detection and defense models. Future updates will aim to incorporate additional UAV types, payload configurations, and multi-sensor data, thereby strengthening its role as a scalable and comprehensive benchmark for secure and intelligent airspace management.

Ethics Statement

This dataset does not involve human or animal subjects. All drone operations were conducted in compliance with local aviation and safety regulations, and permissions were obtained for test flights.

Data Availability

The VisioDECT dataset is publicly available at IEEE Dataport; and can be downloaded freely on the Zenodo repository and on the Kaggle repository.

Credit Author Statement

Simeon Okechukwu Ajakwe: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft, Writing – Review & Editing, Supervision; **Vivian Ukamaka Ihekoronye**: Methodology, Data Curation, Validation, Writing – Original Draft, Writing – Review & Editing; **Golam Mohtasin**: Conceptualization, Methodology, Data Curation, Validation, Writing – Original Draft, Writing – Review & Editing; **Rubina Akter**: Conceptualization, Methodology, Writing – Original Draft, Writing – Review & Editing; **Jae Min Lee**: Conceptualization, Supervision, Funding, Writing – Review & Editing; **Dong Seong Kim**: Supervision, Funding, Writing – Review & Editing.

Acknowledgements

This research was supported by the Priority Research Centers Program through the NRF funded by the MEST (2018R1A6A1A03024003) (50%) and by the Institute of Information & Communications Technology Planning & Evaluation (IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government (MSIT)(IITP-2025-RS-2020-II201612) (50%).

Declaration of Competing Interest

The authors declare no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S.O. Ajakwe, V.U. Ihekoronye, D.-S. Kim, J.M. Lee, Dronet: Multi-tasking framework for real-time industrial facility aerial surveillance and safety, Drones 6 (2022) 46.
[2] J. Liu, L. Plotegher, E. Roura, C. de Souza Junior, S. He, Real-time detection for small uavs: Combining yolo and multi-frame motion analysis, IEEE Transac. Aerospace Electr. Syst. (2025) 1–15, doi:10.1109/TAES.2025.3577592.
[3] S.O. Ajakwe, D.-S. Kim, Facets of security and safety problems and paradigms for smart aerial mobility and intelligent logistics, IET Intellig. Transport Syst. 18 (2024) 2827–2855.
[4] The LocalUpdate: 143 flights cancelled at Frankfurt Airport due to drone sighting, 2019. Accessed: May 9, 2019. [Online]. Available: https://www.thelocal.de/20190509/disruption-after-frankfurt-airport-halts-flights-due-to-drone-sighting.
[5] BBCGatwick Airport drone attack: Police have ’no lines of inquiry’, 2019. Accessed: Sep. 27, 2019. [Online]. Available: https://www.bbc.com/news/uk-england-sussex-49846450.
[6] R. Akter, M. Golam, J.-M. Lee, D.-S. Kim, Doppler radar-based real-time drone surveillance system using convolution neural network, in: 2021 International Conference on Information and Communication Technology Convergence (ICTC), 2021, pp. 474–476, doi:10.1109/ICTC52510.2021.9620998.
[7] M.S. Allahham, M.F. Al-Sa’d, A. Al-Ali, A. Mohamed, T. Khattab, A. Erbad, Dronerf dataset: A dataset of drones for rf-based detection, classification and identification, Data Brief 26 (2019) 104313.
[8] J. Zhao, J. Zhang, D. Li, D. Wang, Vision-based anti-uav detection and tracking, IEEE Transac. Intellig. Transport. Syst. 23 (2022) 25323–25334.
[9] M.S. Awan, S.A.A. Zaidi, J. Mir, Uett4k anti-uav: A large scale 4k benchmark dataset for vision-based drone detection in high-resolution imagery, IEEE Access, 2025.
[10] A. Rouhi, S. Patel, N. McCarthy, S. Khan, H. Khorsand, K. Lefkowitz, D. K. Han, Lrddv2: Enhanced long-range drone detection dataset with range information and comprehensive real-world challenges, arXiv preprint arXiv:2508.03331 (2025).
[11] T.R. Lenhard, A. Weinmann, K. Franke, T. Koch, Syndronevision: A synthetic dataset for image-based drone detection, in: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), IEEE, 2025, pp. 7637–7647.

[12] M. Muzammul, A. Algarni, Y.Y. Ghadi, M. Assam, Enhancing uav aerial image analysis: Integrating advanced sahi techniques with real-time detection models on the visdrone dataset, IEEE Access 12 (2024) 21621–21633.
[13] J. Suo, T. Wang, X. Zhang, H. Chen, W. Zhou, W. Shi, Hit-uav: A high-altitude infrared thermal dataset for unmanned aerial vehicle-based object detection, Scientific Data 10 (2023) 227.
[14] A. Antonini, W. Guerra, V. Murali, T. Sayre-McCord, S. Karaman, The blackbird dataset: A large-scale dataset for uav perception in aggressive flight, in: International Symposium on Experimental Robotics, Springer, 2018, pp. 130–139.
[15] S.O. Ajakwe, V.U. Ihekoronye, D.-S. Kim, J.-M. Lee, Alien: Assisted learning invasive encroachment neutralization for secured drone transportation system, Sensors 23 (2023) 1233.
[16] S.O. Ajakwe, I.S. Igboanusi, J.-M. Lee, D.-S. Kim, i banda, A blockchain-assisted defense system for authentication in drone-based logistics, Drones 9 (2025) 590.
[17] Uptodown, Free Video to JPG Converter, 2022. URL: https://free-video-to-jpg-converter.en.uptodown.com/windows/download.
[18] MakeSense, Alpha Make Sense Application, 2022. URL: https://www.makesense.ai/.
[19] S.O. Ajakwe, V.U. Ihekoronye, G. Mohtasin, R. Akter, A. Aouto, D.S. Kim, J.M. Lee, Visiodect dataset: An aerial dataset for scenario-based multi-drone detection and identification, IEEE Dataport 2 (2022) 13.