Defence Science Journal, Vol. 74, No. 4, July 2024, pp. 535-545, DOI : 10.14429/dsj.74.19256

A Drone-Based Image Dataset Generation Methodology for Single Image Super-Resolution

Amul Batra and Meetha V. Shenoy*

Department of Electrical & Electronics Engineering, Birla Institute of Technology & Science, Pilani - 333 031, India

**E-mail: meetha.shenoy@pilani.bits-pilani.ac.in*

ABSTRACT

The advancements in drone technologies, digital imaging, computer vision techniques, and the liberalized laws related to drone flying have opened up drone-based applications such as the delivery of supplies, search and rescue, aerial surveillance, and so on. The drones, especially the nano/micro/small drones, may be mounted with only low-resolution camera(s) due to their maximum takeoff weight limitations. The low-resolution images generated by the cameras, if used for landing, can result in faulty detection unless the photos are taken from a very close distance to the point of interest. Detection and recognition of the point(s) of interest as early as possible is required to ensure sufficient response time for safe maneuvering. Hence, the images are to be captured at greater heights or distances from the point(s) of interest, and obtaining the high-resolution images from the captured low-resolution images is crucial. The High Resolution (HR) and the Low Resolution (LR) image pairs for training super-resolution models in the works presented in literature are generated using two different cameras or the HR images are captured by the camera and LR images are generated by degrading the HR images. As both methods are not appropriate for small/micro/nano category drones, we propose a novel method based on Ground Sampling Distance (GSD) to capture the LR and HR images. In this paper, we have presented the designed methodology for the creation of a dataset using drone-mounted cameras covering a broad spectrum of views of the target(s) suitable for training and testing of the Single Image Super-Resolution (SISR) models. We also present a methodology for selecting an appropriate target for imaging that enables the visual quality assessment of the developed super-resolution model.

Keywords: Drone; Deep learning; Ground sampling distance; Real-time situation awareness; Computer-vision; Super-resolution

NOMENCLATURE

I_{LR}	: Low resolution image
I	: High resolution image
σ	: Image degradation factors
D	: Image degradation function
GSD_w	: Ground sampling distance for image width
GSD_h	: Ground sampling distance for image height
h_{GSD}	: Altitude of drone at GSD value
EFL	: Effective focal length
CF	: Crop factor
Θ	: Drone camera angle
Φ	: Latitude of the selected point
λ	: Longitude of the selected point
H	: Altitude of flight above ground level
R	: Radius of earth
β	: Bearing angle of axis
n_θ	: Number of waypoints on an axis
n_β	: Number of axes
n_H	: Number of graded heights
N	: Total number of images captured

Received : 22 June 2023, Revised : 28 February 2024
Accepted : 05 February 2023, Online published : 24 July 2024

1. INTRODUCTION

The advancements in digital camera technology and computer vision have paved the way for the deployment of drones in applications such as search and rescue, aerial surveying, goods delivery, and so on[1]. The Directorate General of Civil Aviation (DGCA), India, permits the use of drones for civilian purposes in India under the following categories– nano drones (all-up weight up to 250 grams, maximum flight altitude 15m Above Ground Level (AGL)), micro (all-up weight greater than 250 grams up to 2 kg, maximum flight altitude of 60m above AGL), small (all-up weight >2 kg up to 25 kg, maximum flight altitude of 120 m above AGL), and medium drones between (all-up weight >25 kg up to 150 kg, maximum flight altitude of 9000 m above AGL) subject to registration and licensing rules[2].

A swarm of drones can be used to enable quick decision-making via aerial surveillance during floods, landslides, etc.[2-4]. Machine learning-based object identification and detection techniques can enable drones to identify critical events and make landing and takeoff decisions during search and rescue, goods delivery, etc. With the availability of efficient computing resources, these techniques can be executed on board for decision-making. The cameras used in the nano/

micro/small drones are limited in their resolution, as the high-resolution cameras are generally heavy and can account for a significant percentage of their all-up weight[4]. As the height of flight increases, the image resolution, i.e., the capability to distinguish features on the ground for a given camera configuration, reduces. It is advisable to make the vision-based decisions as early as possible, and hence, the images are to be captured at greater heights or distances from the point(s) of interest, and obtaining the high-resolution images from the captured low-resolution images is crucial[5-7].

Single Image Super-Resolution (SISR), which generates a High-Resolution (HR) image from a Low-Resolution (LR) image, is an image processing technique for generating an enhanced resolution image from a single input image[8-9]. Datasets with annotated images of the target captured at various camera orientations and heights are required so that the images provide a realistic world view to develop Super-Resolution (SR) models. Parameters such as Structural Similarity (SSIM), and Peak Signal to Noise Ratio (PSNR) are generally used as metrics for the performance assessment of the SISR applications[10-11]. Datasets consisting of labeled images captured from drones at various heights, orientations, etc., to reliably test the performance of the super-resolution are required to enable studies on super-resolution for drone-based applications. Also, unique targets are required to be designed, which, when photographed from drones, can be used to visually test the performance of SISR techniques for drone-based applications. The main contributions of this work are as follows. The paper presents

- The designed methodology for the creation of a dataset using drone-mounted cameras covering a broad spectrum of views of the target suitable for training and testing of the SISR models.
- The designed methodology for the selection of an appropriate target for imaging to visually evaluate the performance of SISR techniques for drone-based applications.
- The mission planning technique utilized for the generation of the image dataset.

The paper is organised as follows. Section 2 presents the related works and the gaps in research work. Section 3 presents the problem definition in detail. Section 4 presents the proposed methodology for dataset generation, and Section 5 covers the mission planning strategy. The performance analysis of the proposed methodology is presented in Section 6, followed by the conclusions.

2. RELATED WORKS

Super-Resolution (SR) is particularly useful in the case of nano/micro/small/medium category drones due to their limited capability to use high-resolution cameras. For higher endurance due to lower air drag, greater coverage, and the overall field of view, it is desirable to fly the drone high. The Ground Sampling Distance (GSD) is higher for higher altitudes, and the capacity of the camera to resolve small artifacts decreases with height[9]. Hence, the SR techniques are critical for object identification[9].

Supervised machine learning techniques for the implementation of SISR require apriori training of the network/model on a pair of Low-Resolution (LR) and the corresponding High-Resolution (HR) images. Let 'I′ represent an HR image and 'I_{LR}′ represent the LR image. 'I_{LR}′ can be described as

$$I_{LR} = D(I,\sigma) \quad (1)$$

where, 'D' is the degradation function that maps the HR image to the LR image, and 'σ′ refers to the image degradation factors such as camera orientation, resolution, blurring, noise, etc[8]. Unsupervised SISR techniques use an unlabeled set of images to generate new datasets from the pattern learned using deep networks[8,12-13]. Semi-supervised techniques use unlabeled datasets or a combination of unlabeled and labeled dataset (LR-HR) pairs for SR[8]. It is important to ensure that the dataset encompasses images that are captured at various orientations, heights, textures, and lighting conditions to represent the real-world scenario[8]. Thus, the generation of the dataset is a vital exercise for the training or evaluation of the deep learning based on SISR techniques/models.

Some of the publicly available image datasets utilized for the training or testing of SISR applications are described below. Berkeley Segmentation datasets, such as BSDS300 and BSDS500, and other datasets, such as Set 5, Set 14, and Urban 100, contain images of animals, buildings, food, landscapes, human beings, plants, etc.[14]. Even though the datasets mentioned above are used for SR training and evaluation, they are not specifically designed for the SISR applications and hence do not include any annotation regarding the orientation with respect to the camera, distance from the camera, etc. Agustsson, *et al.* presented a DIV2K dataset with 1000 RGB images manually crawled from the Internet to be used for SISR model training and testing[14]. The above-mentioned datasets include the HR images, and their corresponding LR images are to be created by degrading the HR images using image processing techniques (such as subsampling or introducing noise, etc.). To our knowledge, the Drone Super Resolution (DSR) dataset is the only dataset publicly available for super-resolution of images captured from drones[15]. DSR dataset provides images from various heights and orientations. In the DSR dataset, the high-resolution and low-resolution images are generated using two different cameras (a high and a low-resolution camera) placed at the same height. Also, the process of generation of dataset using drones is not elaborated in literature. From the literature, the following gaps in research work were observed.

Most of the datasets used for image SR applications generate the LR images by degrading the captured HR image using image processing techniques. Thus, the generated LR images doesn't necessarily represent the real-world scenario. Some works have utilized LR and HR cameras to generate LR-HR pairs. Both these methods are not ideal for dataset creation using small/micro/nano category drones.

It was also observed that a specific methodology or operating principle is not yet publicly available to create drone-based image datasets to train or test SISR-based deep learning models.

3. PROBLEM DEFINITION AND PROPOSED SOLUTION

Obtaining the LR and the HR image pairs captured from

drones is a major challenge for training machine learning-based super-resolution techniques. Drone Super Resolution (DSR) dataset is a dataset specifically made available for studies on super-resolution of images captured from drones. This dataset provides images of targets from various heights and orientations. In this dataset, the high-resolution and low-resolution images are generated using two different cameras (a high and a low-resolution camera) placed at the same height. The DSR dataset consists of images captured using a 4/3 inch CMOS L2D-20c Hasselblad camera (5280×3956 pixel resolution) with a focal length of 24 mm and a 1/2 inch CMOS tele-camera (4000 ×3000 pixel resolution) with a focal length of 162 mm onboard DJI Mavic 3 drone[15]. The images from the Hasselblad camera are considered as a high-resolution images and the images from the tele-camera are considered as low-resolution images. As discussed previously HR-LR pairs for training super-resolution models in the works presented in literature are generated using two different cameras (LR and HR cameras) or HR images are captured by the camera and LR images are generated by degrading the HR images. As both methods are not appropriate for small/micro/nano category drones, we propose a novel method based on Ground sampling distance (GSD) to capture the LR and HR images.

GSD depicts the effective distance on the ground captured by an aerial camera from a particular height above ground level[9]. The expression to calculate GSD is given as follows:

$$GSD_W = \frac{Sensor\ Height \times Flight\ Height}{Focal\ Length \times Image\ Width} \quad (1)$$

$$GSD_h = \frac{Sensor\ Height \times Flight\ Height}{Focal\ Length \times Image\ \ Height} \quad (2)$$

$$Focal\ Length = \frac{Effective\ Focal\ Length(EFL)}{Crop\ Factor(CF)} \quad (3)$$

$$GSD=\max(GSDw,\ GSDh) \quad (4)$$

As per Eqn. (1) and (2), GSD is directly proportional to the distance of the camera from the target and is inversely proportional to the focal length of the camera. GSD is generally expressed in centimeters per pixel. A lower value of GSD implies higher spatial resolution. As GSD increases, the capability to differentiate between finer features will reduce, and the clarity of the image will deteriorate. If the GSD value is less than the width of the smallest feature of interest in the image to be captured during photography, then the feature can be clearly identified.

The size of the smallest feature of interest in the image is decided based on the application for which the drone is being used. The height at which the GSD becomes comparable to the width of the feature of interest is considered as h_{GSD} and five (less or more number of images can be captured) images are captured at graded heights for a given target. In our work, the image captured at h_{GSD}-16 feet is considered as very high resolution, image captured at h_{GSD}-8 feet is considered as high resolution, image captured at h_{GSD} is considered as borderline, and images captured at h_{GSD}+8 feet are considered as low resolution, and images captured at h_{GSD}+16 feet are considered as very low resolution. To visually evaluate the quality of the SR model, a unique target image was required. In our work, the target image is also proposed based on the GSD and the camera parameters as mentioned in Section 4. This serves as a method of visually evaluating the SISR model without relying only on PSNR and SSIM as the evaluation metric for SISR model training and testing.

4. METHODOLOGY

In this section, the designed methodology to capture images suitable for the training and testing of deep learning models for SISR applications using a drone-mounted camera is explained.

4.1 Selection of Drone and Camera

A remotely piloted drone (DJI Mavic Air 2) belonging to the micro-category with a gimble-mounted camera was selected to capture stable, blur-free images at designated

Table 1. Specifications of the drone and camera

Parameter	Value/Range
Maximum all-up weight including, fixed payload (gimbal, camera, and sensors)	Up to 500 grams
Maximum ascent speed	4 m/s
Maximum descend speed	5m/s
Maximum sealing above mean sea level	500 m
Maximum horizontal speed	10 m/s
Maximum endurance (with the wind)	30 min
Maximum range (with safety margin)	5 km
Operating temperature	-5° to 50° centigrade
Operating Frequency	2.400-2.4835 GHz
Transmitter Power (EIRP)	Less than 26 dBm
Satellite Systems	GPS+GLONASS
Other LRUs required	Compass, IMU, internal memory to support 30 high-resolution images and 10 minutes video, Obstacle avoidance system.
Gimble mount	3-axis stable (tilt, roll, and pan)
Sensor	1/2” CMOS
Effective pixels	48 Megapixels
Field of view	At least 80°
Equivalent focal Length	24 mm
ISO	General range 100-3200 (Auto)
Maximum photo resolution	8000×6000 pixel
Photo format	JPEG

heights and orientations for the creation of the SISR dataset. The drone and camera specifications are provided in Table 1.

4.2 Design of Object for Imaging

To visually evaluate the quality of the SISR model, a unique target image was required. As indicated in Eqn. (1) - (4), the flight altitudes of the drone can be decided based on the ground details to be captured and the GSD of the camera. In our work, the target image is proposed based on the GSD and the camera parameters which can then be used for evaluation of the quality of the SISR model.

The unique features of the target created for imaging are:

- A conspicuous central point of the target (e.g., a red circle) to enable easy referencing for the calculation of the camera angle.
- An image with concentric circles of varying widths and distances between the concentric circles is proposed as a target image, as shown in the CAD view in Fig. 1.
- The target pattern can be used to visually infer the quality of any SISR model.

The width and distances between the concentric circles in the designed target are tabulated in Table 2. The rationale behind the target design and its suitability for the visual evaluation of the quality of the SR model is elaborated below. Based on Eqn. 1-4, a GSD look-up table was tabulated, as shown in Table 3, for the selected camera for various drone altitudes.

- If the value of GSD at drone altitude h (altitude of the drone above ground level) is less than the width of white and black concentric circles of the landing target, the circles are easily distinguishable and can be considered as a high-resolution image.
- As the value of GSD at drone altitude h1 ≥ h approaches the width of white and black concentric circles, the circles are not very easily distinguishable, and after reaching boundary values, i.e., GSD= width of the concentric

Table 2. Target image profile

Geometric shape	No of similar figures	Position (from the boundary of the prior figure)	Diameter (D) or Width (W)
Solid red Circle (at center)	1	At the center	D:5 cms
Alternating B&W circles	10	0.5 cm	W: 0.5 cm
Alternating B&W circles	10	1 cm	W: 1 cm
Alternating B&W circles	10	1.5 cm	W: 1.5 cm
Alternating B&W circles	10	2 cm	W: 2 cm
Alternating B&W circles	10	2.5	W: 2.5 cm
Alternating B&W circles	5	3 cm	W: 3 cm
Alternating B&W circles	5	3.5 cm	W: 3.5 cm
Alternating B&W circles	5	4 cm	W: 4 cm
Alternating B&W circles	5	4.5 cm	W: 4.5 cm
Alternating B&W circles	5	5 cm	W: 5 cm

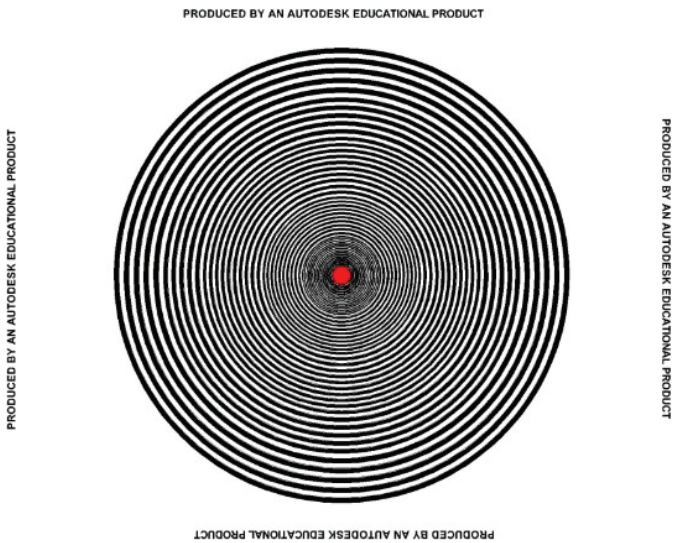

Figure 1. AutoCAD generated target object.

Table 3. GSD look-up table

Height (AGL) in meters	GSD in cm/px	Remarks
10	0.36	with a width of 0.5 cm are visible
15	0.54	with a width of 1 cm are visible
20	0.72	with a width of 1 cm are visible
25	0.90	with a width of 1 cm start to blur out
30	1.08	with a width of 1.5 cm are visible
35	1.26	with a width of 1.5 cm are visible
40	1.44	with a width of 2 cm are visible
45	1.63	with a width of 2 cm are visible
50	1.81	with a width of 2 cm start to blur out
55	1.99	with a width of 2.5 cm are visible
60	2.17	with a width of 2.5 cm are visible
65	2.35	with a width of 2.5 cm will blur out
70	2.53	with a width of 3 cm are visible
75	2.71	with a width of 3 cm are visible
80	2.89	with a width of 3 cm will blur out
85	3.07	with a width of 3.5 cm are visible
90	3.25	with a width of 3.5 cm start to blur out
95	3.43	with a width of 4 cm are visible
100	3.61	with a width of 4 cm are visible
105	3.79	with a width of 4.5 cm are visible
110	3.97	with a width of 4.5 cm are visible
115	4.15	with a width of 4.5 cm start to blur out
120	4.33	with a width of 5 cm are visible

Table 4. Axis of photography and camera profile

Bearing w.r.t N-S direction (β)	Camera angle (w.r.t. the center of the target) (θ)	No waypoints to capture images (at each camera angle)
0°- 180° at increments of 10°	10°, 20°, 30°, 40°, 50°, 60°, 70°, 80°, 90°, 100°, 110°, 120°, 130°, 140°, 150°, 160°, 170°, 180°	5

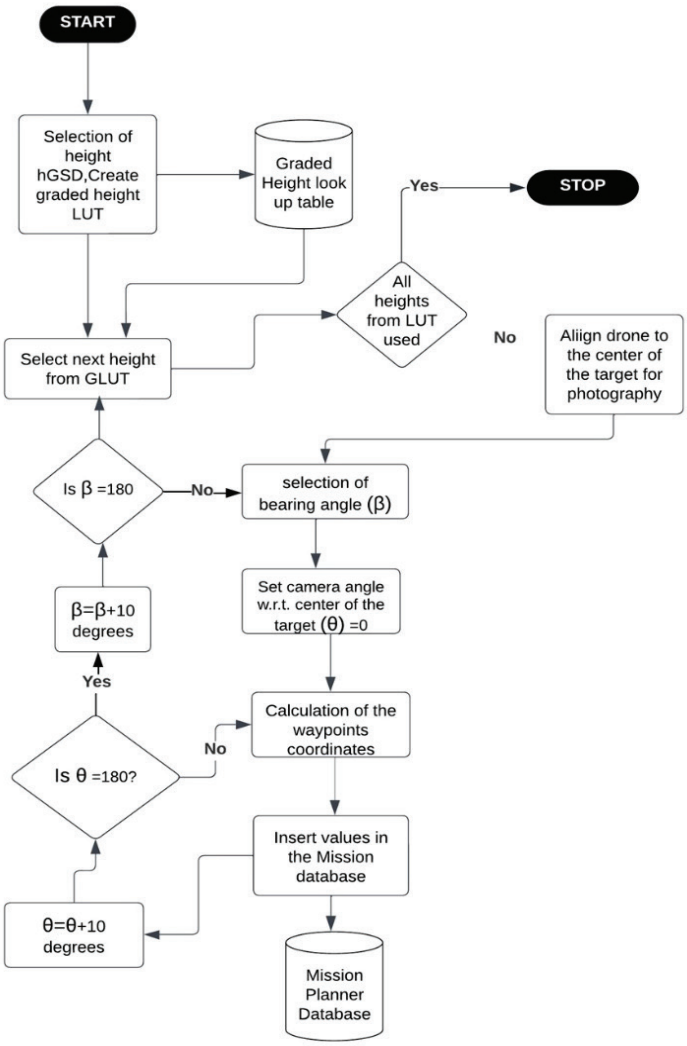

Figure 2. Methodology for imaging.

circles, the image starts blurring out, rendering poor distinguishability of white and black circles.

- As the value of GSD at drone altitude h2 ≥ h1 is more than the width of white and black concentric circles.; the circles are not distinguishable, and the captured image can be considered as an LR image.

The target image is drawn using the Autodesk AutoCAD tool and printed on a 5x5 meter glazed paper for imaging by the drone in our work. The height at which the GSD becomes comparable to the dimension (width) of the feature of interest is considered as h_{GSD}. Images captured at a height below h_{GSD} are considered as high-resolution images and images captured at a height above h_{GSD} can be considered as low-resolution images. The HR-LR image pairs created using the principle mentioned in Section 3 can be used train the SISR model (Note that images captured at h_{GSD} - 8 feet and h_{GSD}+8 feet form HR/LR pairs. Similarly, images captured at h_{GSD}-16 feet and h_{GSD}+16 feet form HR/LR pairs). To visually evaluate the SISR model, the target image presented in this section can be used. For example, to identify the concentric circles with a width of 1 cm of the target, it can be found from Table 3 that the drone should be flown at a height (i.e., h_{GSD}) of 25 mtr or below. The drone can be flown at a height of h_{GSD}+8 feet (to capture the LR image) and fed into the trained SISR model. The generated HR model by the SISR model can be visually monitored to inspect if the concentric circles of width 1 cm are clearly visible. The same principle can be used to inspect each of the concentric circles in the target image. The methodology for imaging is explained in Section 4.3

4.3 Design of Methodology for Imaging

As mentioned in Section 4.2, we intend to capture images of a target from graded heights and also from various orientations of the camera towards the target (or point of interest) to obtain low-resolution and high-resolution images. The imaging is performed in an authorized zone of flying and which provides an uninterrupted and obstacle-free conspicuous view for the line-of-sight photography of the selected target of interest. The image of the target is also required to be captured from various orientations, and hence, the axis of photography has to be suitably selected.

The overall procedure designed for photography is presented as a flowchart in Fig. 2 and further detailed here. At first, the target on the ground is placed at a location with known coordinates T(ϕ1, λ1, 0). The image capture profile is initially chosen in the form of a 180° axis with bearing 'β' radians wr.t. magnetic N-S at the location of photography as shown in Fig. 3 for a selected altitude/graded height. For a specific bearing, the photograph has to be taken from various camera angles as shown in Table 4. The profile is selected for various altitude levels (H ∈ {h_{GSD}, h_{GSD} ± 8 feet, h_{GSD} ± 16 feet}). These profiles at various altitudes form virtual arcs, as shown in

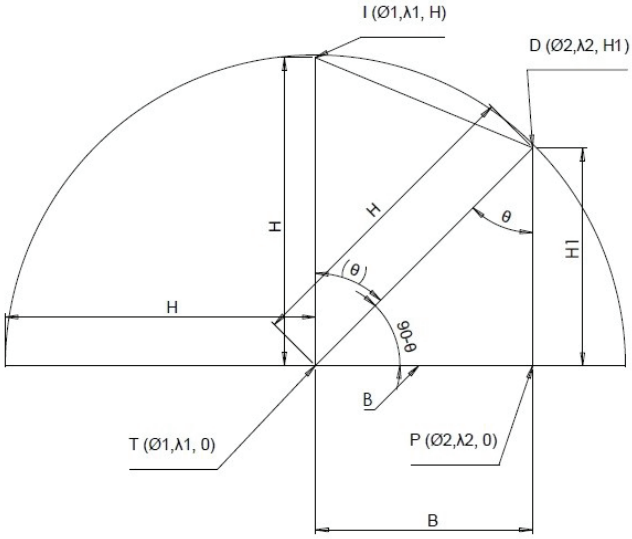

Figure 3. Side view of target photography using drone.

Table 5. Performance summary

Level	PSNR *	SSIM*	Target detection (Center) confidence score (%)
10 m	32.36	0.92	100
20 m	31.25	0.88	100
30 m	31.22	0.82	100
40 m	29.48	0.80	100
50 m	25.01	0.72	98
60 m	22.86	0.66	98
70 m	19.76	0.61	96
80 m	16.21	0.58	92
90 m	11.11	0.41	86
100 m	9.81	0.26	86
110 m	6.21	0.22	75
120 m	5.02	0.11	68

Table 6. YOLOv3 configuration

Parameter	Value
Classes	• Centre of the target (C) • Pot (P) • Solar Panel(SP) • Football Goal Post(GP)
Batch	64
Subdivisions	2
width	416
height	416
channels	3
Momentum	0.9
decay	0.0005
angle	0
saturation	1.5
exposure	1.5
hue	1
learning_rate	0.001
max_batches	8000
steps	6400,7200
CUDA used	CUDA 12.2
GPU used	MIG enabled A100-SXM4-80GB
cuDNN	8.9.3
compute_capability	800
OpenCV version	4.5.4
Pre YOLO Conv Filters	27

Fig. 4. The purpose of selecting the waypoints on the graded heights is to capture images at different clarity levels or resolution for the same camera angle from the center of the target. Waypoints are identified by latitude, longitude, and altitude.

The drone is initially flown to a height of 'H' directly above the target 'T'. Let the first waypoint for photography is chosen as I(ϕ1, λ1, H) as marked in Fig. 3. Suppose the next waypoint for photography is 'D' lying on the profile arc with a bearing 'β' radians wr.t. magnetic N-S, making a camera angle θ (as mentioned in Table 4) with the center of the target. To calculate the coordinates of D(ϕ2, λ2, H1), the following procedure is adopted. As shown in Fig. 3, let 'θ' be the camera angle.; 'ϕ' is latitude, 'λ' is longitude, 'R' is the earth's radius. Then,

$$H = H \times \mathrm{Sin}(\theta) \tag{5}$$

$$H1 = H \times \mathrm{Cos}(\theta) \tag{6}$$

We have used Haversine formula to find coordinates of unknown points lying on a spherical surface for short-distance navigation. The 'ϕ2' and 'λ2' of the waypoint 'D' can be calculated using the Haversine formula as follows[16]:

$$dbyr = \frac{B}{Radius\ of\ Earth} \tag{7}$$

$$\mathrm{Sin}\ Phi2 = (\sin(\phi 1) \times \cos(dbyr)) + (\cos(\phi 1) \times \sin(dbyr) \times \cos(\beta)) \tag{8}$$

$$t_x = \sin(\beta) \times \sin(dbyr) \times \cos(\varphi 1) \tag{9}$$

$$t_y = \mathrm{Cos}(dbyr) - (\sin(\varphi 1) \times \sin(Phi2)) \tag{10}$$

$$\varphi 2 = \sin^{-1}(\sin Phi2) \tag{11}$$

$$\lambda 2 = \lambda 1 + \tan^{-1}(t_x, t_y) \tag{12}$$

The steps, as mentioned in Eqn. (5)- Eqn. (12), are repeated to find all the waypoints lying on the arc with different values of camera angle 'θ' for a given height 'H'. After completion of axial photography along heights h_{GSD}± 16, h_{GSD} ± 8, and h_{GSD}, along the selected axis, the next axis at an angular separation or bearing offset of 10° relative to the previous bearing is chosen to take photographs. The process is repeated again and again (in Fig. 2) till bearings of 180° around the target, as mentioned in Table 4, are covered.

5. MISSION PLANNING

Mission planning includes the procedure to plan and document the drone's maneuvers and other planned actions, such as clicking the photographs. The mission planning involves

- Identification of the start point, hovering points, and stop points of the drones
- Identification of the drone profile and waypoints around the target
- Identifying the actions to be initiated at the waypoints

A mission plan can be made by using pen and paper or can also, be automated using mission planning software such as Ardupilot. We have used the ardupilot mission planner in our work. For automation of maneuvers through waypoints calculated in subsections 4.2 and 4.3, the waypoints represented through latitude, longitude, and altitude are inserted into the mission planner[18]. At each waypoint, the action points, including clicking the photograph and moving the drone to the next waypoint, are specified.

The following information is fed into the mission planner as shown in Fig. 5:

- Initially, the home position with the target's coordinates (with the object on the ground at the center of the target) is fed into the database, representing the center of the

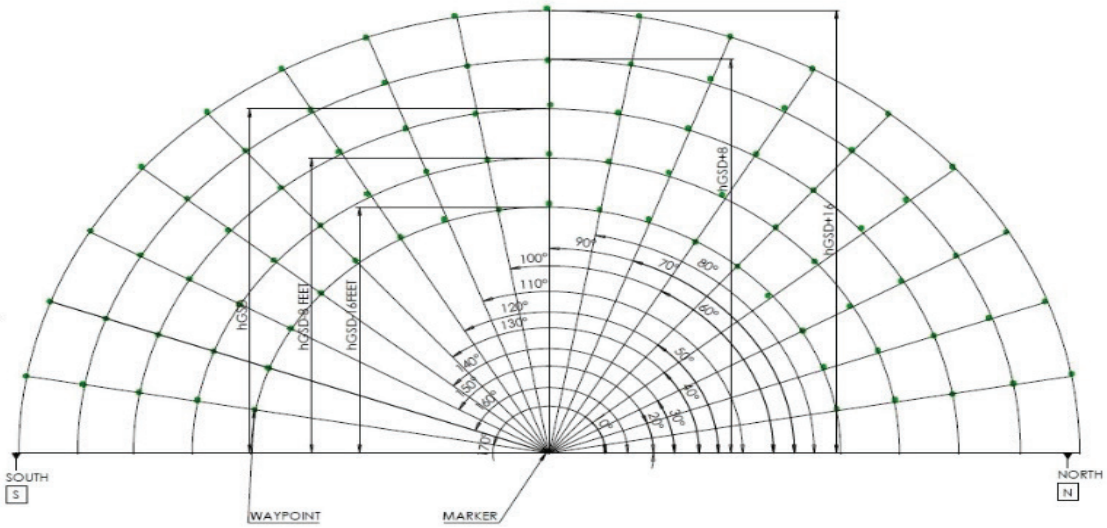

Figure 4. Axial view along N-S aligned bearing.

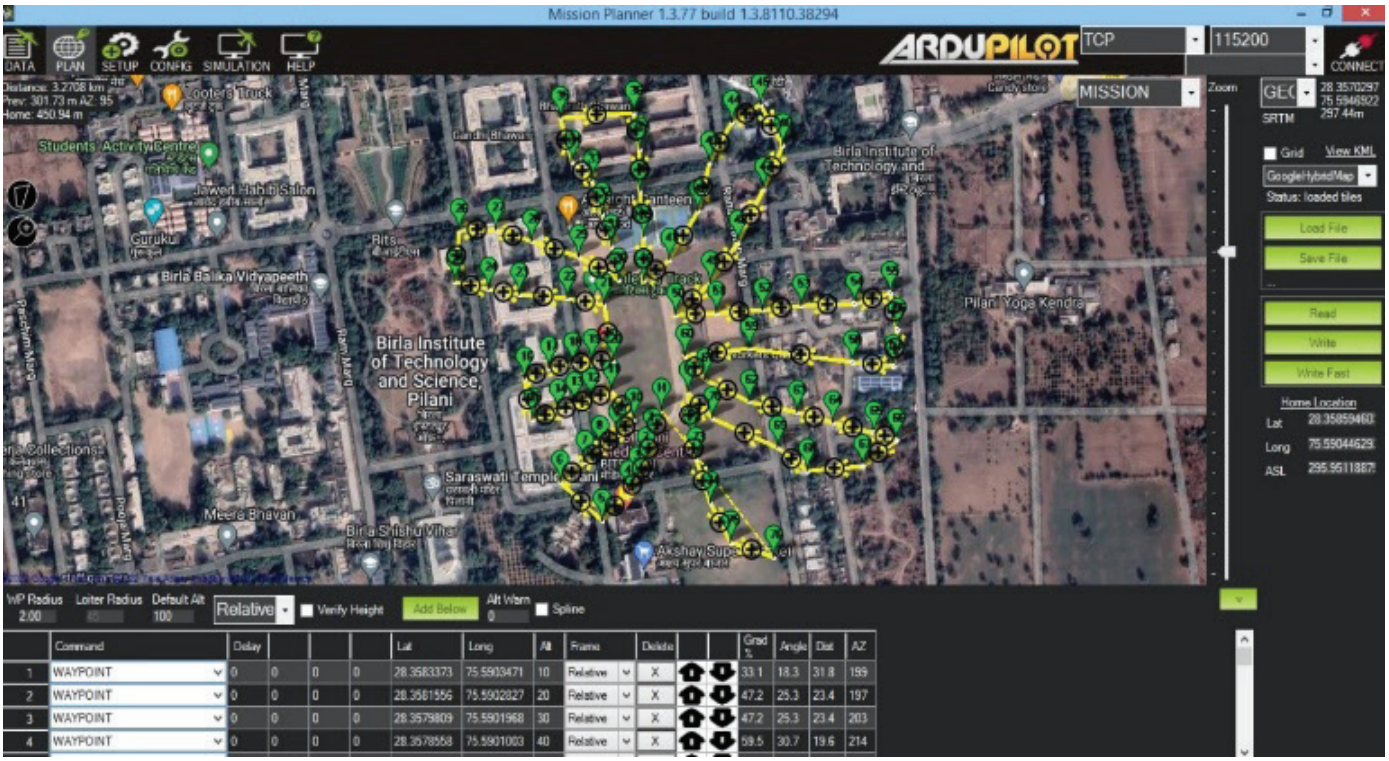

Figure 5. Sample mission planning.

target as the arming and takeoff position of the UAV. This position is also set as the Return To Home (RTH) position in the drone so that in case of communication loss, the drone can automatically navigate to this RTH position via the shortest route.

- The waypoints are then entered one by one, as calculated in Sections 4.2 and 4.3. The same is then reflected in the GIS map of the Ardupilot mission planner, as shown in Fig. 5. A similar procedure is followed to insert waypoints along all the identified axes around the target.
- The drone is placed at the center of the target.
- The drone is armed in autonomous mode to execute the planned maneuvers.
- Once all waypoints are inserted into the mission planner, the values are pushed to the mission plan to the EPROM of the UAV by clicking on the "write" button seen in Fig. 5.

Let 'nθ' be the number of the camera angles, let '$n\beta$' be the number of bearing angles or the axes, let 'nH' be the number of graded heights in consideration, and let 'nW' be the number of images to be taken for the selected bearing, and camera angle.

The bearing angles and the camera angles are listed in Table 4. The number of images (N) that can be captured for a target or point of interest are given by

$$N = nH \times n\beta \times nW \tag{13}$$

Some of the images from the dataset generated for varied bearing angles, camera angles, and altitudes using the proposed

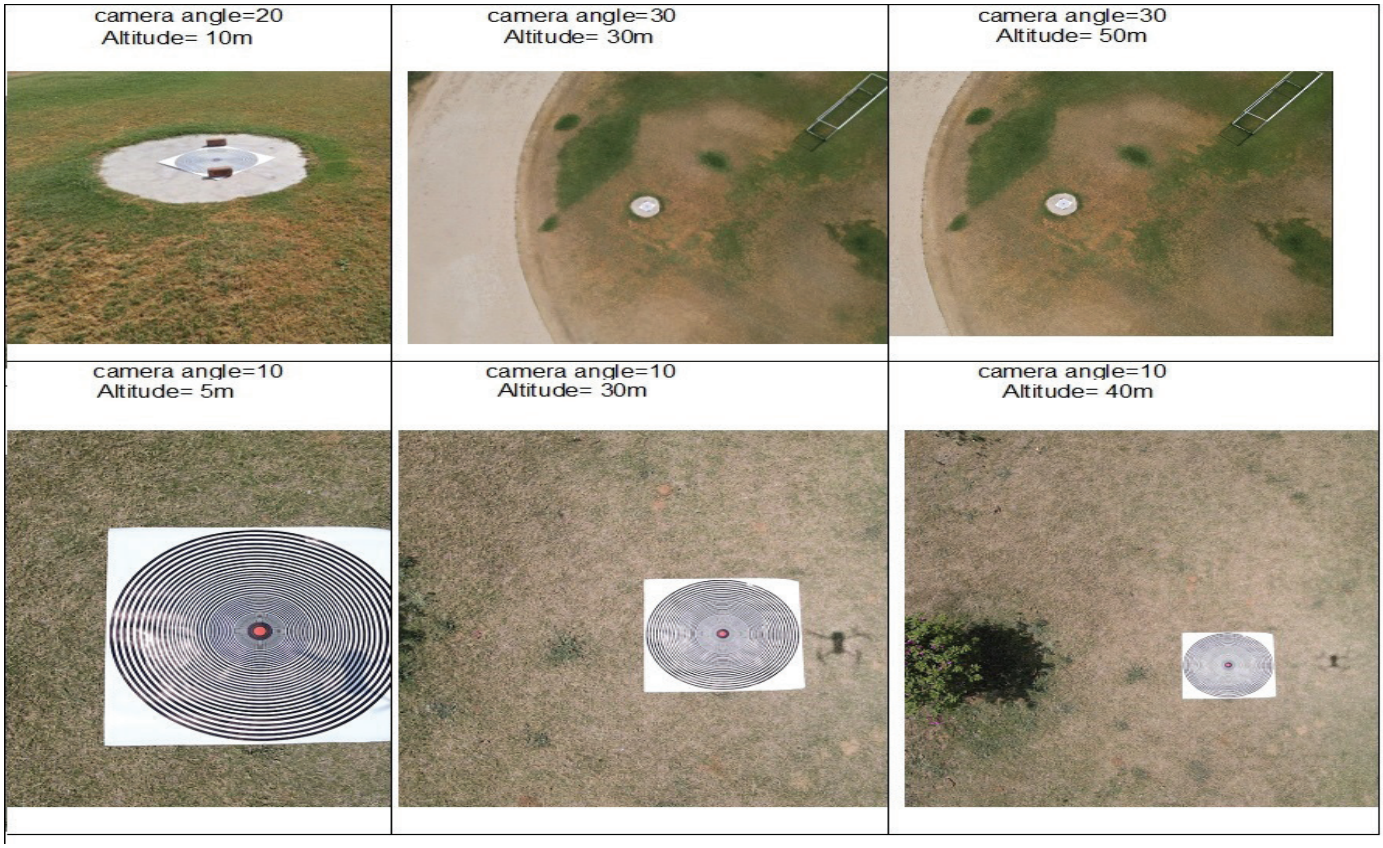

Figure 6. Sample images in the dataset.

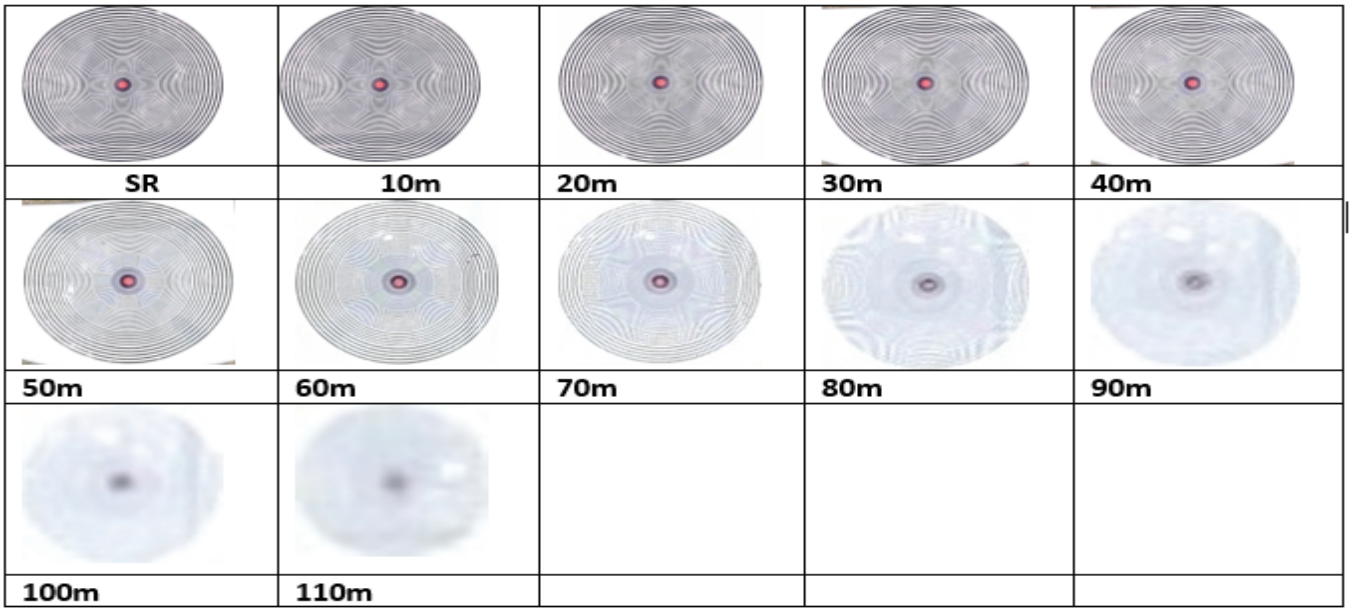

Figure 7. Images used for SSIM/PSNR evaluation.

methodology are represented in Fig. 6. The designed target image is also seen in Fig. 6. The proposed methodology can thus be used for the creation of a dataset using drone-mounted cameras covering a broad spectrum of views of the objects of interest for training and testing of the SISR models.

6. PERFORMANCE ANALYSIS

In our work, we propose to use the super-resolution technique for making real-time situation awareness decisions in applications such as such as the delivery of supplies, search and rescue, aerial surveillance, and so on. Detection and recognition of the point(s) of interest as early as possible is required to ensure sufficient response time for safe maneuvering. The super-resolution will hence target to generate the High-resolution images from the low resolution images (captured at a higher height from the objects of interest).

The high-resolution images generated by the SISR model should contain be able to depict the features captured by the camera at a lower height. The HR-LR image pairs created using the principle mentioned in Section 3 can be used train the SISR model (Note that images captured at h_{GSD}-8 feet and h_{GSD}+8 feet form the HR/LR pairs. Similarly, images captured at h_{GSD}-16 feet and h_{GSD}+16 feet form HR/LR pairs). PSNR and SSIM are generally used as metrics to compare the quality of the two images[10-11]. PSNR is given by:

$$PSNR = 10\log_{10}\frac{L^2}{MSE} \tag{14}$$

where,

$$MSE = \frac{1}{MN}\sum_{M}\sum_{N}(I_1(m,n) - I_2(m,n))^2 \tag{15}$$

where, '$I_1(m,n)$' and '$I_2(m,n)$' represents the two images (eg. high resolution image captured by the camera and the high-resolution image generated by the SR technique) being compared and M and N are row and column index of the image. SISM indicates the amount of similarity between two images based on luminance, contrast and structure.

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{16}$$

where, 'x' and 'y' represents the two images being compared, 'μ' represents the mean, 'σ' represents the variance and 'σ_{xy}' represents the covariance of the images 'x' and 'y'. The constants in Eqn. (16) are $C1=(0.01L)^2$ and $C2=(0.03L)^2$ respectively. ' L' in Eqn. (14) represent the dynamic range of the pixel values. To evaluate the techniques presented in Section 4 and to analyze the suitability of the unique target designed for visual inspection, the drone was placed at heights directly over the designed target (as mentioned in Section 4.2). The target image taken at an altitude of 5m are selected as the reference image, and other images were cropped and resized using the non-destructive INTER_AREA interpolation technique of the OpenCV library, thus preserving the pixel density and aspect ratio of the images while reducing the size of the images at higher heights to cover the same field of view as that of the reference image[21]. The image captured at a height of 5 m was then center aligned and compared with the resized and cropped images captured at heights of 10 m to 100 m and PSNR &

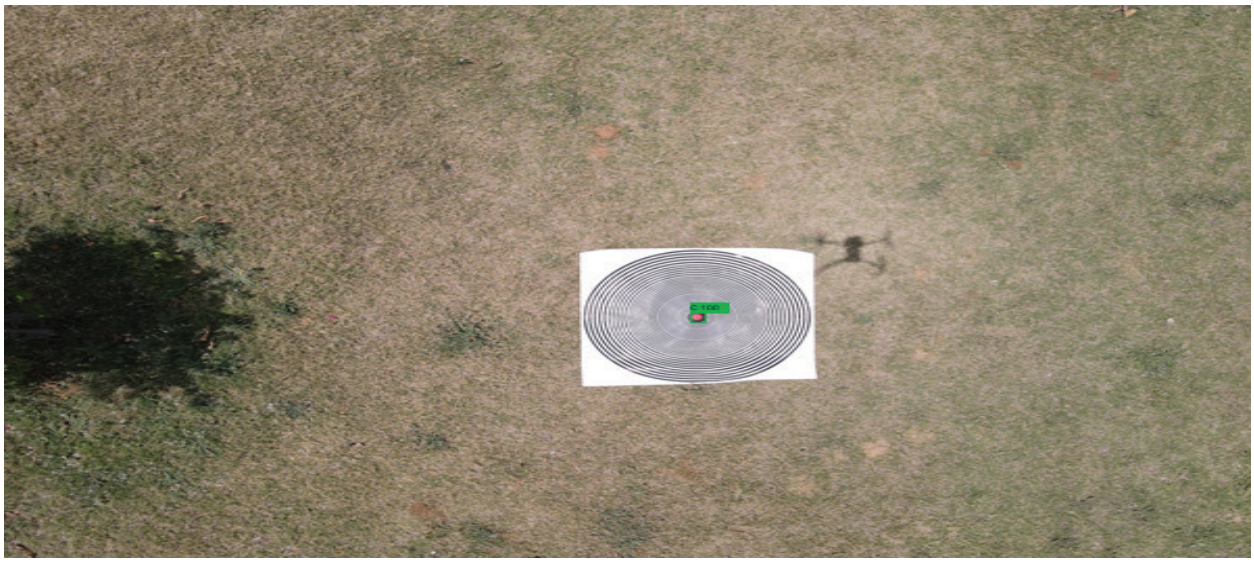

Figure 8. YOLOV3 output for image captured at 30 m altitude.

Figure 9. YOLOV3 output for image captured at 70 m altitude.

SSIM were calculated as shown tabulated in Table 5. It can be observed from Table 5 that the PSNR and SSIM decrease with height indicating that the images at lower height can be considered as HR image and the images at greater heights (based on the feature of the interest) can be considered as low-resolution image.

Even though the works presented in the literature (on drone-based super-resolution or otherwise) have demonstrated an improvement in PSNR or SSIM, whether the achieved PSNR or SSIM is enough for a specific application is not addressed yet. To evaluate usability of the dataset for decision making in real-time situation awareness applications, a popular object detection technique YOLOv3 (You only Look Once) image classifier was applied on the images captured at various heights[22]. Set of images at various heights were chosen and annotated using an open source image annotator LabelImg[23]. Four classes of objects as mentioned in Table 6 were annotated in the images using rectangular bounding boxes in YOLO text formats. YOLOv3 configuration used in our work is presented in Table 6. YOLOv3 was trained on 150 images from the dataset. Testing was done on separate set of images at various heights. Sample output with the class specific confidence scores for two images captured at the altitude of 20 m and 70 m are depicted in Fig. 8 and Fig. 9. The target is detected by YOLOv3 and the confidence score of the centre of the target can be observed in the YOLOv3 output. Table 5 also provides a summary of the confidence score for target (Centre of the target) prediction at varying heights. The higher confidence score for object detection achieved using the images captured at lower heights when compared to the images captured at higher heights indicates that those images can be considered as HR images.

7. CONCLUSIONS

In this paper, a methodology for the creation of a dataset using drone-mounted cameras covering a broad spectrum of views of the target suitable for training and testing of the SISR models is presented based on the ground sampling distance. Both LR as well as HR images are captured in the dataset. Also, a methodology for the selection of an appropriate target for imaging to visually evaluate the performance of SISR techniques for drone-based applications is proposed. The mission planning technique utilized for the generation of the image dataset is also elaborated. Further, the method and operating principles of aerial photography are not just limited to the particular target type but can be used to capture images of objects and other ground-based artifacts to create a holistic dataset for SISR model training and testing. We believe that the proposed work accelerates the use of small/micro/nano drones with LR cameras for real-time decision-making and plan to use the technique for decision making in real-time situation applications.

REFERENCES

1. Shakhatreh, H.; Sawalmeh, A.; Al-Fuqaha, Ala.; Dou, Zuochao.; Almaita, Eyad.; Khalil, Issa.; Othman, Noor.; Khreishah, A. & Guizani, M. Unmanned Aerial Vehicles (UAVs): A survey on civil applications and key research challenges, *IEEE Access*, 2019, **7**, 48572-48634. doi: 10.1109/ACCESS.2019.2909530
2. July 2021, Gazette of India, extraordinary, part II . https://www.civilaviation.gov.in /sites/default/files/Draft Drones Rules 14 Jul 2021.pdf (Accessed on 17 June 2023).
3. Ragab, A.; Isaac, M.; Luna, M. & Pena, P. Unmanned aerial vehicle swarming. *In* Proceedings of the International Conference on Engineering and Emerging Technologies (ICEET),2021. doi:10.1109/ICEET53442.2021.9659698.
4. Lahsen-Cherif, I.; Liu, H. & Lamy-Bergot, C. Real-time drone anti-collision avoidance systems: An edge artificial intelligence application. *In* Proceedings of the IEEE Radar Conference (RadarConf22), New York City, NY, USA, 2022: doi: 10.1109/RadarConf2248738.2022.9764175.
5. Nex, F.; Armenakis, C.; Cramer, M.; Cucci, D.; Gerke, M.; Honkavaara, E.; Kukko, A.; Persello. & Skaloud, J. UAV in the advent of the twenties: Where we stand and what is next. *J. Photogrammetry and Remote Sensing*, 2022, **184**, 215-242. doi:10.1016/j.isprsjprs.2021.12.006.
6. Sundar, A.; Vaithiyanathan, K.; Thangadurai, V. & Namdeo, N. Design and analysis of fusion algorithm for multi-frame super-resolution image reconstruction using framelet. *Defence Sci. J.*, 2015, **65**(4), 292-299. doi:10.14429/dsj.65.8265
7. Truong, N.; Nguyen, P.; Nam, S. & Park, K. Deep learning-based super-resolution reconstruction and marker detection for drone landing. *IEEE Access*, 2019, **7**, 61639-61655. doi: 10.1109/ACCESS.2019.2915944
8. Wang, Z.; Chen, J. & Hoi, S. Deep learning for image super resolution: A survey. *IEEE Trans. on Pattern Analysis & Machine Intell.*, 2021, **43**(10), 3365-3387. doi: 10.1109/TPAMI.2020.2982166.
9. Chen, G.; Wang, H.; Chen, K.; Li, Z.; Song, Z.; Liu, Y.; Chen, W. & Knoll, A. A survey of the four pillars for small object detection: Multiscale representation, contextual information, super-resolution, and region proposal. *In* IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, **5**(2), 936-953, 2022, doi:10.1109/TSMC.2020.3005231
10. Li, Y.; Liu, J. & Chen, Y. Learning structural coherence via generative adversarial network for single image super-resolution. *In* Proceedings of International Conference on Computer Engineering and Application (ICCEA), Kunming, China, 2021, doi: 10.1109/ICCEA53728.2021.00044.
11. Valenzuela, A. & Reyes, J. Basic Spatial Resolution Metrics for Satellite Imagers, *IEEE Sensors J.*, 2019, **19**(13), 4914-4922. doi: 10.1109/JSEN.2019.2902512.
12. Inzerillo, L.; Francesco, A.; Gaetano, D. & Mohammed, Z. Super-resolution images methodology applied to UAV datasets to road pavement monitoring. *Drones*, 2022, **171**(7). doi:10.3390/drones6070171

13. Yamanaka, J.; Kuwashima, S. & Kurita, T. Fast and accurate image super-resolution by deep CNN with skip connection and network in network. *In* Proceedings of International Conference on Neural Information Processing, 2017, *Springer, Cham.* doi: 10.48550/arXiv.1707.05425
14. Agustsson, E. & Timofte, R. NTIRE 2017 Challenge on Single image super-resolution: Dataset and study, *In* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017. doi:10.1109/CVPRW. 2017. 150.
15. Lin, X.; Ozaydin, B.; Vidit, V.; Helou, M. & Süsstrunk, S. DSR: Towards drone image super-resolution. 2023, 361-377. doi: 10.1007/978-3-031-25063-7_22
16. Suryana, A.; Reynaldi, F.; Pratama, F.; Ginanjar, G.; Indriansyah, I. & Hasman, D. Implementation of haversine formula on the limitation of e-voting radius based on android, *In* Proceedings of the International Conference on Computing, Engineering, and Design (ICCED), Bangkok, Thailand, 2018, doi: 10.1109/ICCED.2018.00050.
17. Andreou, A.; Mavromoustakis, C.; Batalla, J.; Markakis, E.; Mastorakis, G. & Mumtaz, S. UAV trajectory optimisation in smart cities using modified A* algorithm combined with haversine and vincenty formulas, *IEEE Transc. on Vehicular Tech.*, 2023, **72**(8), 9757-9769, doi: 10.1109/TVT.2023.3254604.
18. Yanming, H.; Wang, J. & Li, Bo. DSAA-YOLO: UAV remote sensing small target recognition algorithm for YOLOV7 based on dense residual super-resolution and anchor frame adaptive regression strategy, *J. King Saud University – Comp. and Info. Scs*, 2023, 1319-1578. doi: 10.1016/j.jksuci.2023.101863.
19. Reddy, K.; Molabanti, V.; Dumpala, S. & Nelakuditi, U. YOLOV3 based real time drone detection for counter drone system, *In* Proc. of the IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), Mysuru, India, 2023, doi: 10.1109/TEMSMET56707.2023.10149935.
20. Mission planner overview. Mission planner overview - Mission planner documentation. (n.d.). https://ardupilot.org/planner/docs/mission-planner-overview.html. (Accessed on 23 July 2023)
21. Bradski, G. The openCV library. *Dr. Dobbs Jl: Soft. Tools for the Profess. Program.*, 2000, **25**(11), 120-123. ISSN: 1044-789X
22. Redmon, J. & Farhadi, A. Yolov3: An incremental improvement. arXiv preprint, 2018 arXiv:1804.02767.
23. Giunchiglia, F.; Diao, X. & Bagchi, M. Incremental image labeling via iterative refinement. *In* Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Rhodes Island, Greece, 2023. doi: 10.1109/ICASSPW59220.2023.10193475.

CONTRIBUTORS

Mr Amul Batra obtained his MTech in Embedded Systems Design from Birla Institute of Technology & Science, Pilani (BITS-Pilani). He is currently a Research Scholar at the department of Electrical & Electronics Engineering department of BITS-Pilani and working in the area of design and development of real-time image resolution enhancement techniques. His research interests include: Image processing and deep learning.
His contribution to the current study includes: Design and implementation of the work proposed in the paper and the preparation of the initial draft of the paper.

Dr Meetha V. Shenoy obtained her PhD degree from the Birla Institute of Technology and Science (BITS) at Pilani, India. She is working as an Assistant Professor with the Department of Electrical and Electronics Engineering, BITS Pilani. Her research interests include: Internet of things, networked embedded systems, and autonomous systems development.
Her contribution to the current study includes: Design of the work proposed in the paper, funding, supervision of the research work, and manuscript preparation.