Sensors and Materials, Vol. 34, No. 7 (2022) 2637–2661
MYU Tokyo

S & M 2996

Positioning Errors of Objects Measured by Convolution Neural Network in Unmanned Aerial Vehicle Images

Woosuk Kang,[1] Jisung Kim,[2*] Hongsic Yun,[2] Pooreum Lee,[1] and Heecheol Kim[1]

[1]Computer Communication and Engineering, Kangwon National University,
1 Gangwondaehakgil, Chuncheon-si, Gangwon-do 24341, Republic of Korea
[2]Department of Civil and Environmental Engineering, Sungkyunkwan University,
2066, Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do 16419, Republic of Korea

(Received April 9, 2022; accepted June 21, 2022; online published June 30, 2022)

Keywords: aerial photogrammetry, convolution neural network (CNN), object detection, positioning error, unmanned aerial vehicle

The conversion of unmanned aerial vehicles (UAVs, also called drones) and convolution neural network (CNN) facilitates the location of objects in real time using their sensors. In photogrammetry, the positional accuracy of objects is directly affected by the use of technology. It is necessary to improve the accuracy of object positioning to increase the utilization of drones and CNNs. In this study, the error factors that impede accuracy, such as the global navigation satellite system (GNSS) error of the UAV, camera distortion error, and camera posture error, were analyzed to improve the accuracy of object positioning. The effect of each error was also analyzed. The study was conducted in stages, such as establishing a method for the positioning of objects, specifying errors, and analyzing the amount and effect of error elements. The magnitude of the positioning errors was found by comparing it with accurate values measured by GNSS. Furthermore, the correlation of the errors with the factors that impeded accuracy was analyzed. Consequently, the effect of each error factor on the overall error was identified. These results can play an important role in improving positioning accuracy and developing UAV and CNN technologies employing sensors in the future.

1. Introduction

High-altitude aerial photogrammetry has been used to locate objects in wide areas. With the advent of unmanned aerial vehicles (UAVs) and the development of digital cameras, low-altitude drone photogrammetry, which can quickly and economically locate objects compared with conventional high-altitude manned aerial photogrammetry, has emerged. In contrast, artificial intelligence has facilitated the automatic identification of characteristics of objects more quickly than human eyes. In particular, convolutional neural networks (CNNs) can quickly detect and classify objects within images. When combined, UAVs and CNNs can quickly detect, classify, and locate objects in images.[1–3] Owing to these advantages, various studies have been

*Corresponding author: e-mail: jskim6687@skku.edu
https://doi.org/10.18494/SAM3939

ISSN 0914-4935
https://myukk.org/

conducted in areas such as flood mapping,[4] agriculture,[5–9] the environment,[10] traffic,[11–15] structure management,[16–20] archaeology,[21] and disaster management[22,23] using drones and CNNs.

Zhang *et al.* located weeds using UAVs and deep learning. They constructed a weeding prescription map with errors of less than 20 cm. The relative altitude of the drone from the surface was 2 m.[5] Geraldes *et al.* studied a UAV-based situational awareness system via deep learning. They detected situations through deep learning and converted them into geodetic coordinates, longitude, and latitude.[23] Biswas *et al.* estimated the speed of moving objects using a moving UAV.[12] Deng *et al.* measured the locations of dead trees using drones and deep learning. They detected trees in the images and estimated their longitude and latitude.[6] Zhi *et al.* studied distance measurements using a drone[24] and clarified the measurement errors that occurred during the positioning. Similarly, Liu and Zhang,[25] and Haamied *et al.*[26] focused on positioning. In addition, many studies have been conducted on positioning using drones and deep learning, demonstrating the importance and impact of positioning using UAVs and CNNs.

Positioning accuracy should be improved to increase the utility of drones and CNNs, as attempted in previous studies. Nassar *et al.* used machine learning to detect shapes and match figures for geolocalization. The objectives of some studies have been to improve the location of UAVs.[27] Toma *et al.* used radio frequency to improve the positioning accuracy of a drone,[28] whereas Zhang *et al.* used the time difference of arrival (TDOA)[29] and Bhoite *et al.* used CCTV.[30] However, these studies focused on the positioning accuracy of drones.

Consequently, to improve positioning accuracy, the errors of object positioning measured through UAVs and neural networks should be analyzed. However, few studies have identified the positioning accuracy of currently measured objects and analyzed the error factors. Some studies only located the objects without accuracy estimation,[6,23,25] whereas Zhang *et al.* and Biswas *et al.* demonstrated results including errors.[5,12] In contrast, Zhi *et al.* and Haamied *et al.* clarified the error factors and effects of each element.[24,26] However, further study is required to improve the accuracy due to camera distortion, the GNSS positioning of the drone, and camera posture. Therefore, in this study, the positioning accuracies of objects measured by UAVs and CNNs are analyzed, and the errors and effects of each cause are identified.

2. Objectives and Scope

2.1 Objectives

In this study, we aim to analyze the positioning accuracy of objects located through UAV aerial images and a CNN. Toward identifying the error factors, the detailed objectives are as follows:

- the analysis of the positioning accuracy of objects measured using a direct orientation method that does not use ground control points (GCPs) and
- the analysis of the effect of each cause of error.

2.2 Scope

The scope of this study includes the following:

- the accuracy of object positioning in single-camera photogrammetry using a CNN and a UAV, and
- the analysis of error and its effect.

The following are outside the scope of this study:

- the object detection accuracy of CNNs,
- the correction of error,
- indirect orientation using GCPs, and
- multi-camera photogrammetry.

We make the following assumptions:

- The locations of each object in the real world are points without length and area.
- The positions of the real objects expressed in the image are the center points of the boundary detected by the CNN or have negligible differences.
- The position of the global navigation satellite system (GNSS) receiver mounted on the UAV is the same as the principal point of the camera or has a negligible difference.

In this study, we analyze the following as factors affecting the error:

- the GNSS error of the UAV,
- the distortion error of the camera, and
- the posture error of the camera.

3. Materials and Methods

Figure 1 illustrates the detailed procedure used herein to achieve our objectives. We first established a direct orientation-based object positioning process using a UAV and a CNN. In this process, the following errors were defined for quantitative analysis: position error, camera

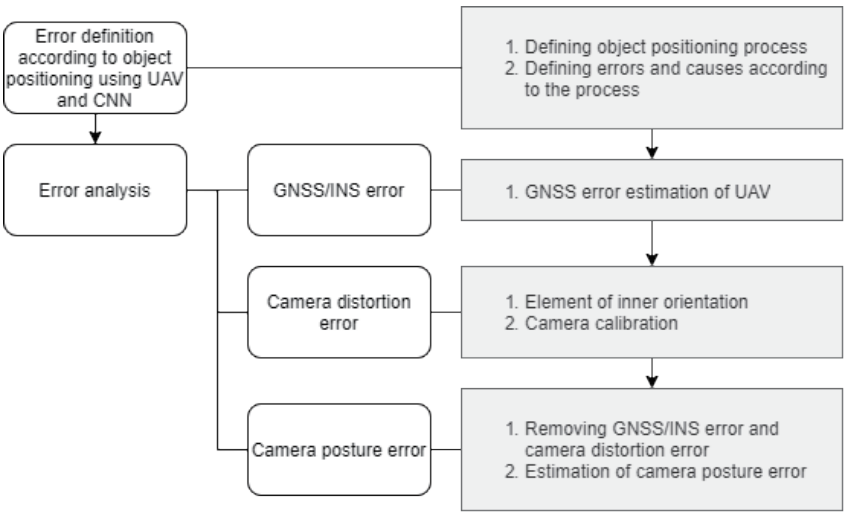


Fig. 1. Overall procedure of this study.

distortion error, and camera posture error. The UAV position error was evaluated at a control point where the location was known. The camera distortion error was analyzed in a laboratory. In the analysis of the camera posture error, the UAV position and camera distortion errors were calibrated from the field images and then compared with the actual position of the object measured by GNSS.

3.1 Object positioning

Object positioning was performed in the following order: aerial image shooting, target object boundary extraction, boundary coordinate acquisition, image coordinate transformation, and ground coordinate transformation. The camera distortion error was introduced while extracting the boundary of an object and converting it into image coordinates. The position error of the UAV and the posture error of the camera were reflected while converting the image coordinates into ground coordinates (Fig. 2).

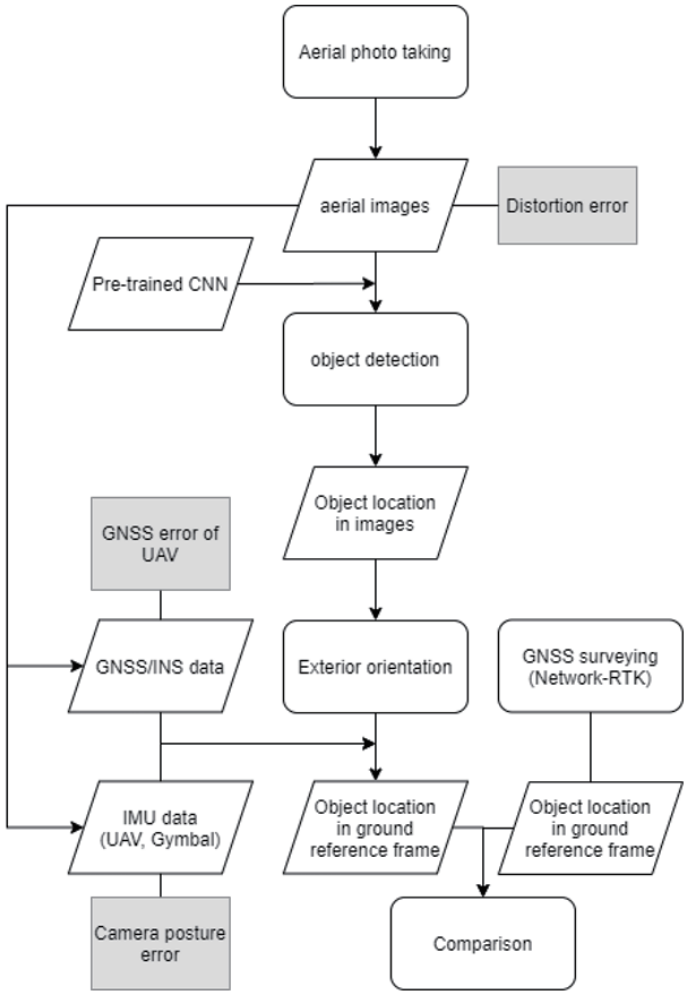


Fig. 2. Object positioning process and errors measured using UAV image and CNN in this study.

3.1.1 UAV and camera for image acquisition

In this study, a DJI Phantom 4 Pro UAV was used. Its position (x_u, y_u, z_u) and the posture of the camera (ω_c, φ_c, κ_c) were obtained through the Exchangeable image file format information of the image taken from the relevant model. Table 1 lists the specifications of the camera and UAV.

3.1.2 Object detection using CNN

The CNN model constructed in this study is shown in Fig. 3. EfficientNet-b0 was used partially as the model backbone. On the basis of MobileNet research,[32] EfficientNet consists of MBConv operation blocks (Fig. 4),[33] such as depthwise separable convolution and squeeze and excitation (SE), to reduce the number of model parameters and improve performance, respectively.

Subsequently, a region proposal network (RPN) was constructed using two convolution operations, and the results of the RPN were learned using heatmap focal loss and generalized intersection over union loss. To detect relatively small objects on the basis of the characteristics of the UAV images, the network model was configured to have a relatively high resolution in both the initial input image size and the RPN input size.

In addition, to address the problem of overfitting that occurs with limited training data, various augmentation techniques were applied during training. The accuracy was improved using test time augmentation.

Table 1
Camera and UAV specifications.[31]

Item	Value
Sensor	1” CMOS
Lens	FOV 84°, 8.8 mm/24 mm
Image size	5472 × 3648 pixels
Gimbal vibration range	±0.02°
Hovering accuracy (GNSS positioning)	±0.5 m in vertical direction ±1.5 m in horizontal direction

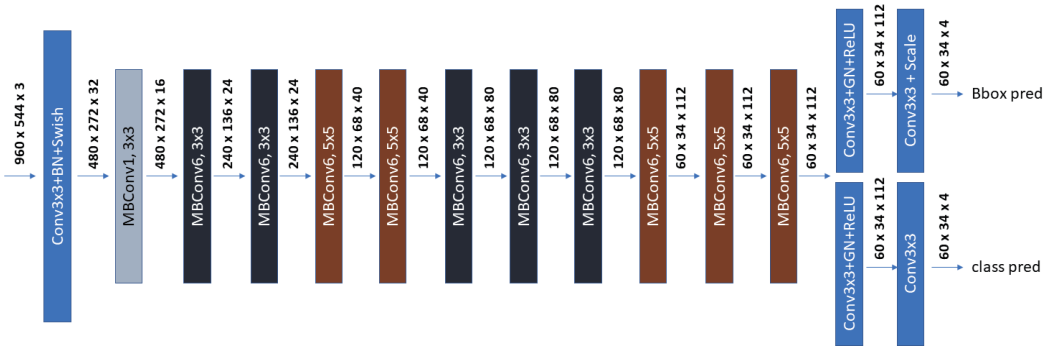


Fig. 3. (Color online) Overall network structure of CNN used in this study.

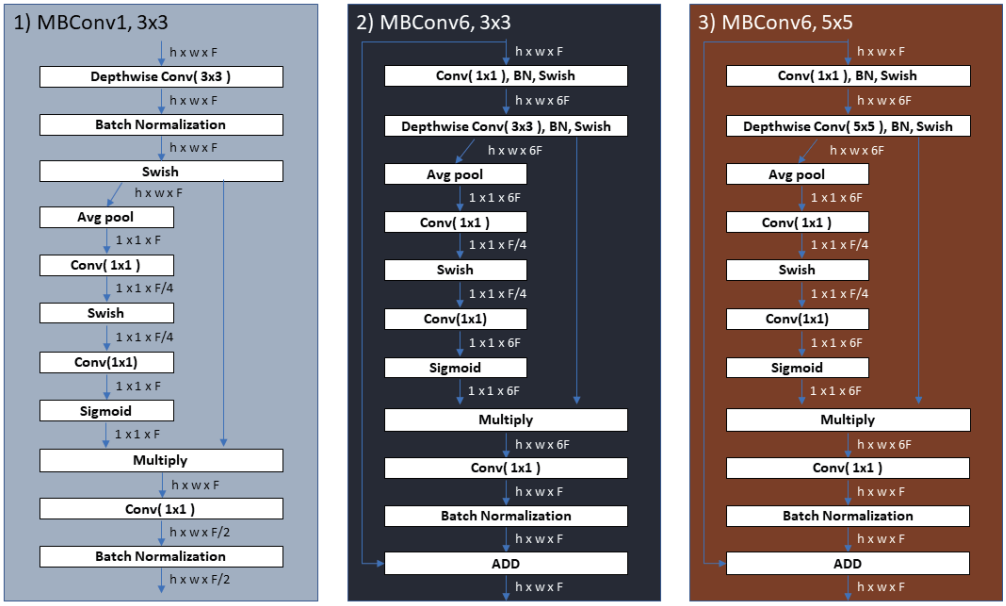


Fig. 4. (Color online) Detailed module structures of CNN used in this study.

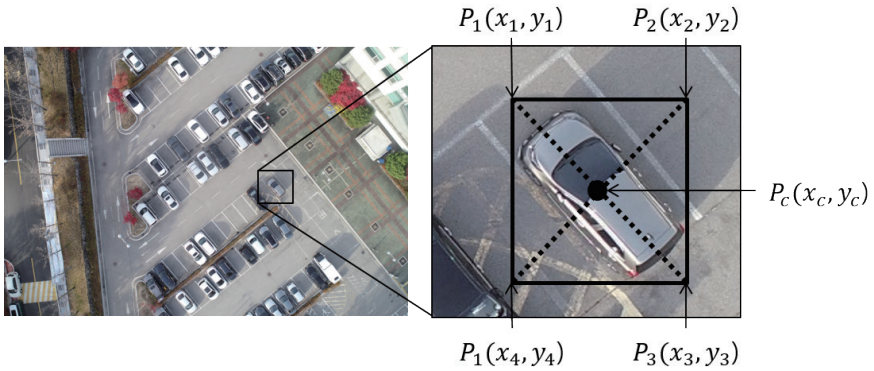


Fig. 5. (Color online) Object position in image coordinate system obtained from detected object boundary.

3.1.3 Object location in image coordinates

The coordinates of the detected object in the image were obtained. Since the boundary of the object detected through the CNN is a rectangle, the intersection of its two diagonals ($\overline{P_1P_3}$, $\overline{P_2P_4}$) was used as the object position $P_c(x_c, y_c)$ (Fig. 5).

3.1.4 Exterior orientation

The position of the object in the ground coordinate system was obtained using external orientation elements and the position of the object in the image coordinates. The external orientation elements x_u, y_u, z_u, ω_c, φ_c, and κ_c were directly obtained using GNSS and the inertial

measurement unit, which were mounted on the UAV and camera. x_u, y_u, and z_u are the positions in the x, y, and z directions based on the ground coordinate system of the UAV, and x_0, y_0, and z_0 are the positions in the x, y, and z directions based on the ground coordinate system of the object, respectively. ω_c, φ_c, and κ_c are the positions of the camera in the roll, pitch, and yaw directions, respectively (Fig. 6).

The positions of the UAV and the posture of the camera are stored in the extensible metadata platform (XMP) of the photographed image. When the focal length is f, the position of the object is a vector factor S. The real positions can be calculated as follows:

$$R = \begin{bmatrix} \cos(\kappa) & -\sin(\kappa) & 0 \\ \sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\varphi) & 0 & -\sin(\varphi) \\ 0 & 1 & 0 \\ \sin(\varphi) & 0 & \cos(\varphi) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & \sin(\omega) \\ 0 & -\sin(\omega) & \cos(\omega) \end{bmatrix},$$
$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} x_u \\ y_u \\ z_u \end{bmatrix} - RS \begin{bmatrix} f \\ y_c \\ z_c \end{bmatrix}. \tag{1}$$

Objects whose elevations are equal are assumed to be on the same plane. Their 2D coordinates are obtained on a planar coordinate system, such as Korea Geodetic Datum 2000 (European Petroleum Survey Group: 5186). The scale factor S can be determined using Eq. (2). The altitude difference H between the object and the UAV is acquired from the XMP of the captured images. After determining the scalar quantity S, the positions of the objects can be estimated using the known variables.

$$\begin{bmatrix} x_0 \\ y_0 \\ 0 \end{bmatrix} = \begin{bmatrix} x_u \\ y_u \\ H \end{bmatrix} - RS \begin{bmatrix} f \\ y_c \\ z_c \end{bmatrix} \tag{2}$$

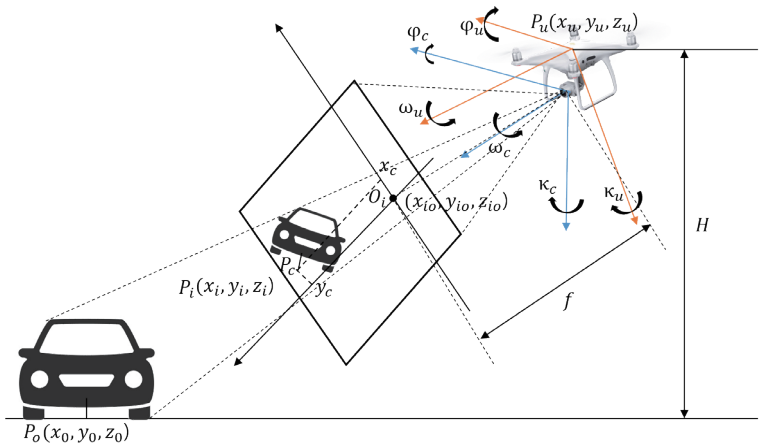


Fig. 6. (Color online) Geometric relationship among UAV, camera, and object.

3.2 True value of object position

The error is the difference between the observed and true values.[34] The observed value is the position of the object calculated from the image, and the true value is the actual position of the object. In this study, the center of an object in a parking area was assumed to be its correct position, which was measured using GNSS. The equipment used for the measurement was a Trimble R8s system using Virtual Reference System-Real Time Kinematic (VRS-RTK) (with a 2–3 cm positioning error) provided by the National Geographic Information Institute (NGII), Republic of Korea. Figure 7 indicates the true position of the object.

3.3 Error analysis

3.3.1 Error clarification

The diverse error factors in object positioning were investigated in this study. These errors were the position error of the UAV, camera distortion error, and camera posture error.

Since direct orientation using GNSS/INS determines the external orientation elements in the images, each error factor is independent of the other factors. Therefore, positioning accuracy is improved by eliminating the error at each stage. The procedure for error identification and elimination is depicted in Fig. 8.

The position error vector V_o of the measured object can be decomposed into the x and y directions of the 2D terrestrial coordinate system. As shown in Eq. (3), V_o can be expressed as a vector V_i owing to the internal expression element of the camera, a vector V_u owing to the position error of the UAV, an error vector V_r owing to the camera posture, and a negligible error vector V_{etc}. In this case, c_x and c_y are the positions of the main point in the image; f_x and f_y are the

Fig. 7. (Color online) Object location measured by VRS-RTK GNSS.

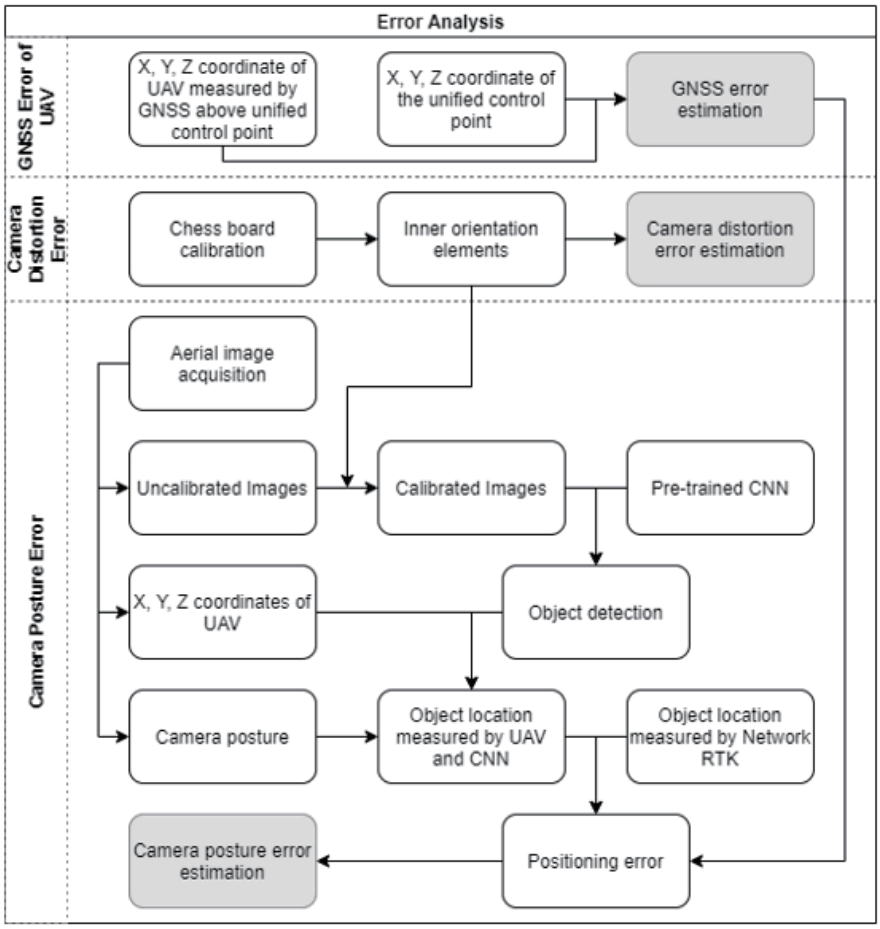


Fig. 8. Process of error analysis for each factor.

focal lengths; k_x and k_y are the radial distortion coefficients; and p_x and p_y are the tangential distortion coefficients in the x and y directions, respectively.

$$V_{o(x_o,y_o,z_o)} = V_{i(c_x,c_y,f_x,f_y,k_1,k_2,p_1,p_2)} + V_{u(x_u,y_u,z_u)} + V_{r(\omega_c,\phi_c,\kappa_c)} + V_{etc} \tag{3}$$

3.3.2 GNSS error of UAV

The GNSS error of the UAV was measured by mounting and flying the UAV at the unified control point (UCP), which was the base point. The position error was estimated by comparing the notified coordinates of the UCP and the coordinates obtained through the GNSS mounted on the UAV (Fig. 9). The UCPs are used as reference points, such as the satellite, level, and gravity reference points, to measure the geographic longitude and latitude, Cartesian coordinates, Earth-centered Cartesian coordinates, height, and gravity. They were installed and managed by NGII of the Republic of Korea, and their planar positioning error was approximately 30 mm. The attributes of the UCPs used in this study are listed in Table 2.

Fig. 9. (Color online) UAV positioning at the UCP.

Table 2
Attributes of UCP used in study.

Item	Value
Name	U-Chuncheon-07
Last updated data	4th February 2020
Latitude	37° 54’ 57.31632”
Longitude	127° 44’ 10.50723”

3.3.3 Camera distortion error

Since only one UAV and camera was used in the study, the error due to the distortion was equal in all experiments. To identify factors such as camera lens distortion and centerline mismatch, the internal orientation elements (c_x, c_y, f_x, f_y, k_1, k_2, p_1, p_2) were calculated using the chessboard method of Zhang.[35] Figure 10 shows some of the images used for the calibration, and Table 3 lists the specifications of the chessboards.

3.3.4 Camera posture error

The error due to the camera posture was estimated after removing the position and camera distortion errors. The camera posture error was analyzed using the data acquired from a testbed.

3.4 Testbed and scenario

3.4.1 Testbed

The testbed for the estimation of camera posture error, which is one of the goals of this study, must satisfy the following conditions:

- It must be outdoors where aerial photography is possible.
- It must be where GNSS measurement is possible, i.e., not in a shaded area.
- A parking space must be divided into sections such that the position of an object can be measured using surveying equipment.

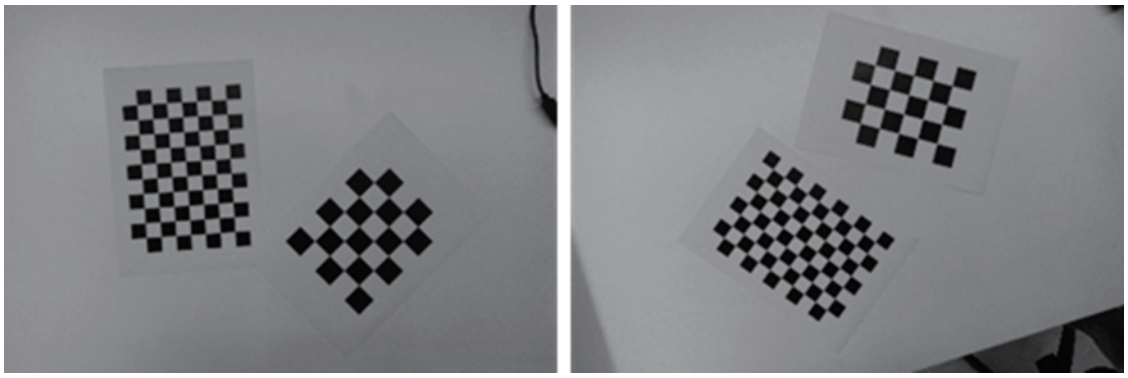


Fig. 10. (Color online) Two chessboard images for camera calibration.

Table 3
Chessboards used for camera calibration.

Item	Chessboard 1	Chessboard 2
Rows	11	6
Columns	8	5
Cell size	21 mm × 21 mm	30 mm × 30 mm
Number of images	36	

(a) (b)

Fig. 11. (Color online) Testbed in Kangwon National University: (a) map and (b) aerial image [Kakao Map,[36] 2022].

- At least 20 of the same vehicles must be present in all the sequential images.

The parking lot in front of the 60th Anniversary Memorial Hall of Kangwon National University (1 Gangwondaehak-gil, Chuncheon-si, Gangwon-do) was set up as a testbed (Fig. 11). Images were acquired on November 17, 2021 from 15:00 to 16:00. The temperature at that time was 12.5 °C, the wind direction was southwest, the wind speed was less than 2.0 m/s, and the pressure at sea level was approximately 1022.0 hPa.

3.4.2 Scenario

The control variables identified in this study were the altitude, yaw, and pitch of the camera. The same area was acquired by the UAV from different locations. Figure 12 illustrates the locations where the images were captured. Table 4 lists the location of each point, the height

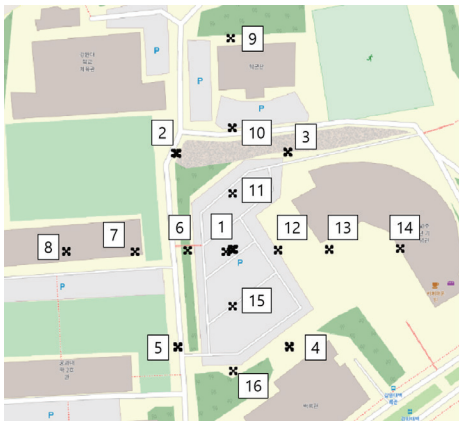


Fig. 12. (Color online) Places where the images were taken.

Table 4
Places where images were taken by scene number and location.

Scene number	UAV location number	Relative altitude (m)	Camera pitch φ_c (°)	Camera yaw* κ_u (°)
1	1	50	−90	0
2	1	100	−90	0
3	1	150	−90	0
4	2	150	−90	0
5	2	100	−90	0
6	2	50	−90	0
7	3	50	−90	0
8	3	100	−90	0
9	3	150	−90	0
10	4	150	−90	0
11	4	100	−90	0
12	4	50	−90	0
13	5	50	−90	0
14	5	100	−90	0
15	5	150	−90	0
16	6	100	−75	90
17	7	100	−60	90
18	8	100	−45	90
19	9	100	−45	180
20	10	100	−60	180
21	11	100	−75	180
22	12	100	−75	−90
23	13	100	−60	−90
24	14	100	−45	−90
25	15	100	−75	0
26	16	100	−60	0

from the ground, and the pitch and yaw of the camera. Images were acquired by programming the waypoint mission in JavaScript object notation (JSON) format in the DJI software development kit. Ten images were supposed to have been taken for each scene; however, the actual number of images taken was between five and ten.

4. Results

4.1 GNSS error of UAV

Figure 13 illustrates the positioning error of the GNSS mounted on the UAV. Since the GNSS receiver mounted on the UAV is a one-frequency receiver, it cannot solve the integer ambiguity problem. Thus, a fixed solution cannot be obtained. As a result, the position error was initially several meters; however, as time passed, the error decreased and stabilized. The localization error was observed to be a decreasing exponential function. Table 5 lists each parameter of the fitted exponential function.

The analysis of the log of the captured images revealed that the interval between images was approximately 4 s, and image stabilization was observed after 35 frames since the UAV moved

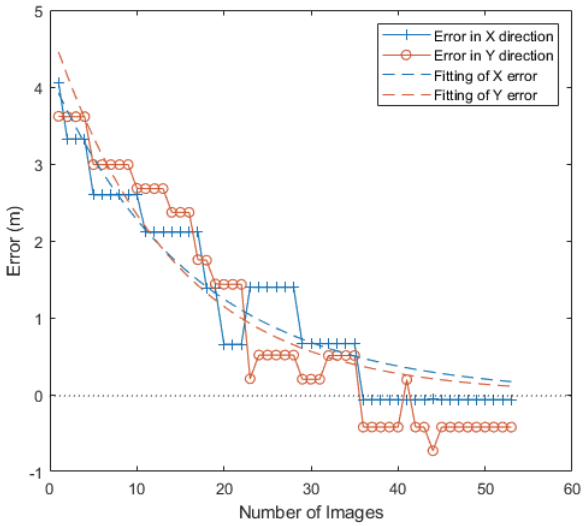


Fig. 13. (Color online) Positioning error of UAV by image frame.

Table 5
Coefficients of fitting curves in *X*- and *Y*-directions.

Direction	Coefficient*	Value	95% confident interval	
			Minimum	Minimum
X-direction	*a*	4.158	3.778	4.539
	b	−0.06028	−0.06803	−0.05252
Y-direction	*a*	4.779	4.175	5.382
	b	−0.07123	−0.08364	−0.05881

*$f(x) = a \cdot e^{b \cdot x}$

to another point. Therefore, by setting a GNSS stabilization time of approximately 2 min after the UAV moved, a position can be obtained within a 1.5 m error, as specified in the DJI document.

4.2 Camera distortion error

The inner orientation elements of the camera are presented in Table 6. Since the difference between the position of the principal point and the center of the image is just 0.5 pixels in both the *X*- and *Y*-directions, the image distortion error near the center is negligible. The ratio of the focal length in the *X*-direction to that in the *Y*-direction is approximately 0.9952. Thus, the scale distortion in each direction is insignificant. According to the radial and tangential distortion coefficients, the distortion is within approximately 100 pixels, even far from the center of the image.

The inner orientation elements of the camera are independent of the UAV position and camera posture. Therefore, the positioning error of the object in the image coordinate system is the same, regardless of the posture. However, when the ground coordinate system was used as a reference, the position error of the object was affected by the camera posture.

4.3 Camera posture error

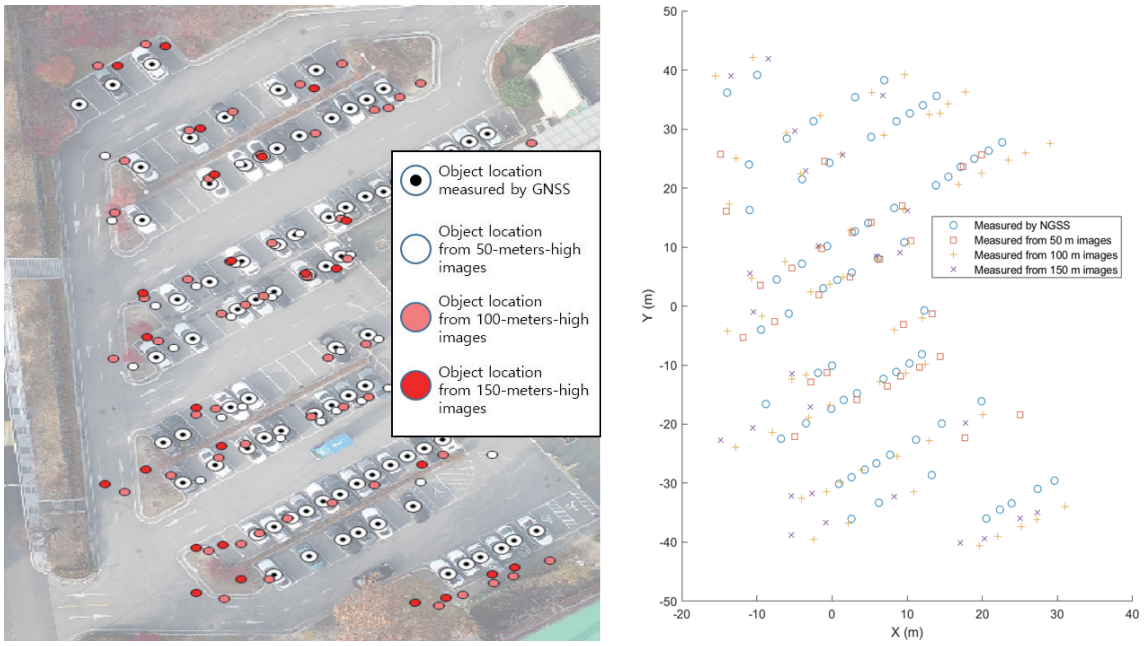
Other error factors were eliminated to estimate the error factors caused by the camera posture. The GNSS position error of the UAV was minimized by allowing sufficient time for the GNSS to stabilize. The camera distortion was calibrated using inner orientation elements.

4.3.1 Variation of posture error with relative altitude

As shown in Fig. 14, the position of the object was calculated from the images. The pitch (φ_c) of the camera was fixed at −90°, the yaw (κ_c) was fixed at 0, and only the relative altitude was changed. The images obtained from Scenarios 1–3 in Table 4 were used. The positions of the objects measured by GNSS and those calculated from the images were analyzed on the basis of the relative altitude.

Table 6
Camera and UAV specifications.[31]

Item	Value
f_x	1026.06271
f_y	1031.52276
c_x	2735.50000
c_y	1823.50000
k_x	−0.00531
k_y	0.00093
p_x	0.00203
p_y	−0.00541

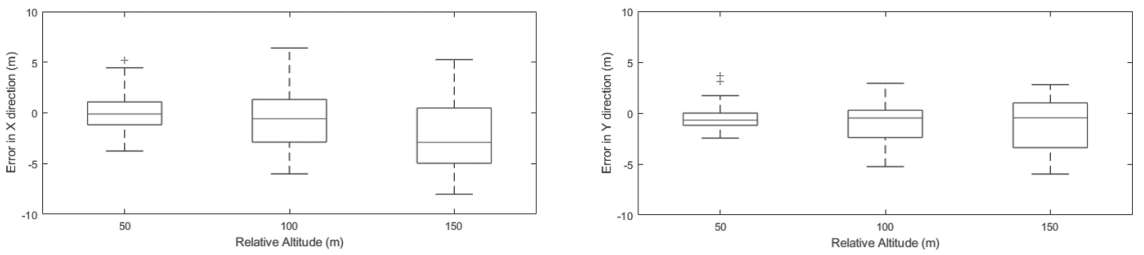


(a) (b)

Fig. 14. (Color online) Positions of objects by relative altitude (a) in reference image and (b) in relative coordinate frame.

Table 7
Errors by relative altitude in both directions.

Relative altitude (m)	Scenario number	Number of samples	Direction	Error	
				Mean (m)	Standard deviation (m)
50	1	174	*X*	0.157	2.148
			Y	−0.429	1.389
100	2	264	*X*	−0.598	2.940
			Y	−0.961	1.874
150	3	115	*X*	−2.479	3.558
			Y	−1.126	2.550



(a) (b)

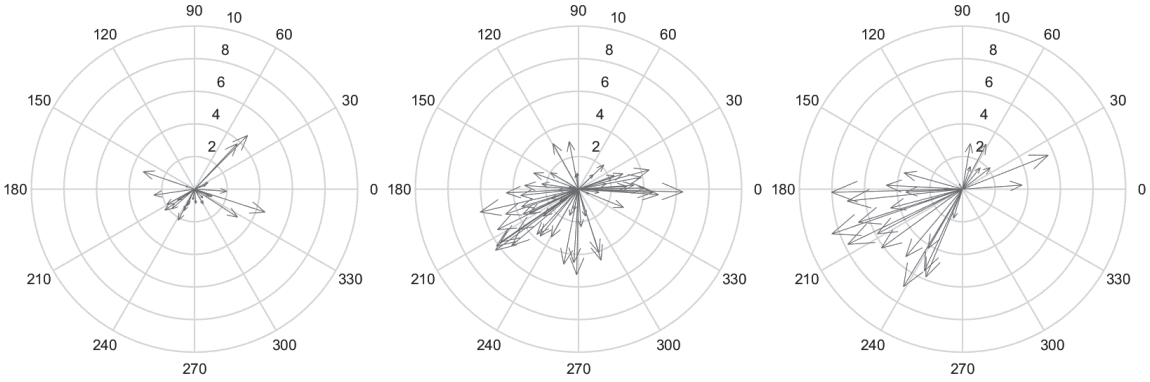
Fig. 15. Box plots of errors in (a) *X*- and (b) *Y*-directions by relative altitude.

Both the size of the errors and the deviation tended to increase with the relative altitude in both axis directions. However, the *x*-axis error was larger than the *y*-axis error (Table 7, Fig. 15).

The axial pixel size of the camera and the focal length in each direction were corrected through calibration. The pixel ratio in each axial direction was 0.9952, which is almost identical to the camera posture error. The direction of the error is related to the yaw (κ_c) of the camera, and the error nonuniformity in each direction is due to the pitch (φ_c) of the camera.

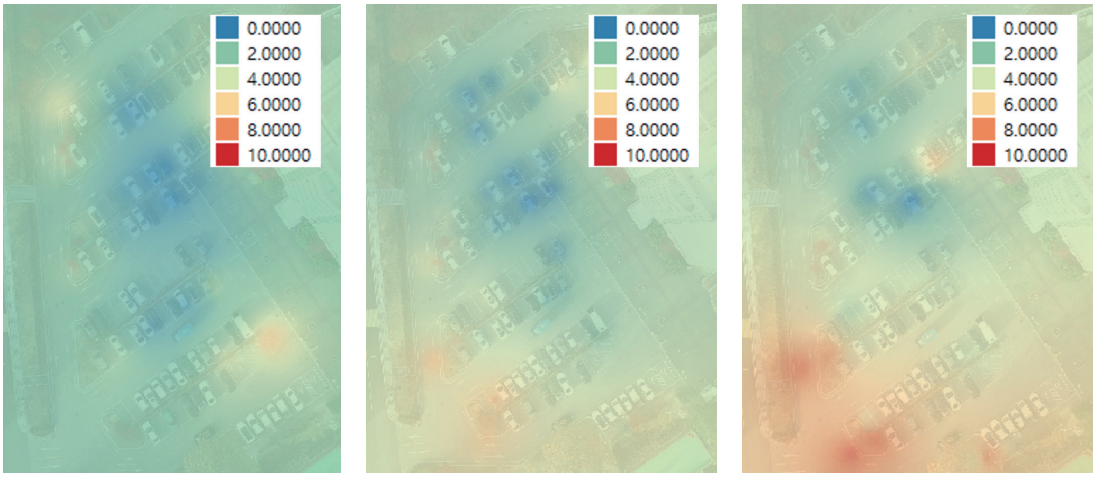
Figure 16 shows the direction and magnitude of the errors. At a relative altitude of 50 m, the direction of the error appears relatively uniform; however, as the altitude increases, the direction of the error tends to be biased toward a specific direction: 270° counterclockwise from the east and southwest directions. It has been observed that the camera posture can be one of the causes of the error through the bias in a specific direction. The results [at 150 m altitude, pitch (φ_c) of the camera of −90°, and yaw (κ_c) of 0 (north)] show an error of approximately 10 m in each direction.

Figure 17 illustrates the estimation error according to the imaging area obtained by interpolation performed by the inverse distance weight (IDW) method using the distance error



(a) (b) (c)

Fig. 16. Direction and size of errors at relative altitudes of (a) 50, (b) 100, and (c) 150 m. The circular unit is the number of degrees counterclockwise from east and the radial unit is meter.



(a) (b) (c)

Fig. 17. (Color online) Error maps estimated by object position using IDW at relative altitudes of (a) 50, (b) 100, and (c) 150 m. The error unit is meter.

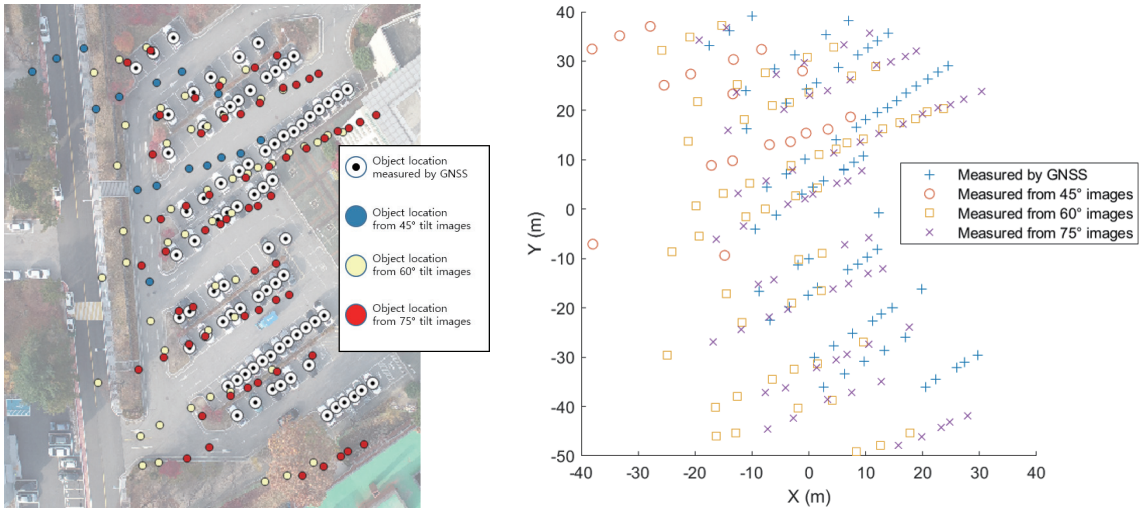
for each position of the object. At all three altitudes, the error is small at the center of the area but large outside the area. When the altitude increases, the external error increases, and in this case, the error is approximately 8 m at 150 m above.

Although the distortion of the camera was corrected, the trend of the distance error in the outer region of the images must be investigated.

4.3.2 Variation of posture error with camera pitch

The error caused by the camera pitch (φ_c) was investigated by setting the relative altitude at 100 m and the yaw (κ_c) at 180°. The locations of the objects were estimated using the image in Fig. 18. This corresponds to Scenarios 19–21 in Table 4. Positions 9–11 (Fig. 12) were the image locations.

Table 8 and Fig. 19 show the error distributions. The *X*-direction error is correlated with the camera pitch, unlike the *Y*-direction error. At the same altitude (100 m), the error in the

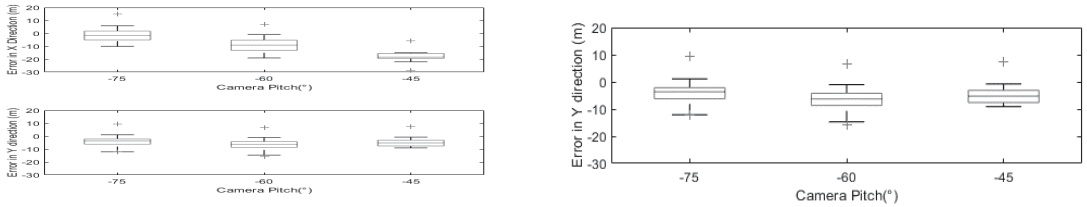


(a) (b)

Fig. 18. (Color online) Positions of objects by camera pitch (a) in reference image and (b) in relative coordinate frame.

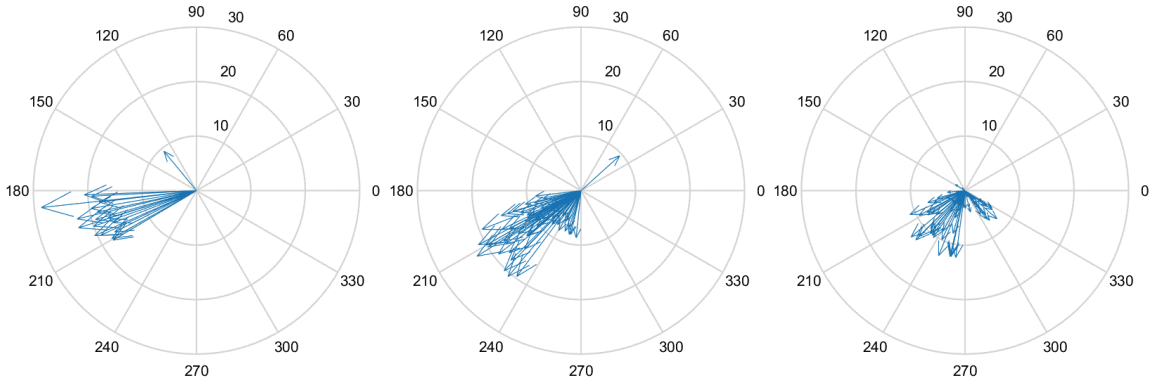
Table 8
Errors by camera pitch in both directions.

Camera pitch φ_c (°)	Scenario number	Number of samples	Direction	Error	
				Mean (m)	Standard deviation (m)
−45	19	63	*X*	−18.056	4.395
			Y	−4.647	3.927
−60	20	292	*X*	−9.004	5.317
			Y	−6.653	3.931
−75	21	355	*X*	1.493	4.581
			Y	−4.268	3.702



(a) (b)

Fig. 19. Box plots of errors in (a) *X*- and (b) *Y*-directions by camera pitch.



(a) (b) (c)

Fig. 20. (Color online) Direction and size of errors for camera pitches of (a) −45, (b) −60, and (c) −75°. The circular unit is the number of degrees counterclockwise from east and the radial unit is meter.

X-direction is approximately 60 cm when the pitch is −90°, indicating a significant effect of the pitch on the error. This trend is illustrated in Fig. 19.

Figure 20 clearly shows the trend in this error. As the pitch of the camera approached 0, the size of the error increased, and the direction of the error was from southeast to east. The closer the camera pitch was to 0, the larger the error was in the yaw direction. This result shows that the pitch of the camera should be kept close to −90° and the yaw should be corrected for accurate object position measurement.

Figure 21 illustrates the variance of the errors according to the object position by the camera pitch. Fig. 21 shows a similar pattern to Fig. 17. The same error modality implies that the same error components occur with the same tendency.

4.3.3 Posture error varying camera yaw

To investigate the error factors caused by camera yaw (κ_c), the relative altitude was fixed at 100 m and the camera pitch (φ_c) was fixed at -75°. Since the same area had to be acquired, the locations from where the images were captured were changed. The locations of the objects were estimated using the images taken from Scenarios 16, 21, 22, and 25 in Table 9, which correspond to locations 6, 11, 12, and 15 in Fig. 12. Figure 22 illustrates the positions of the objects measured by the UAV and CNN using the camera yaw.

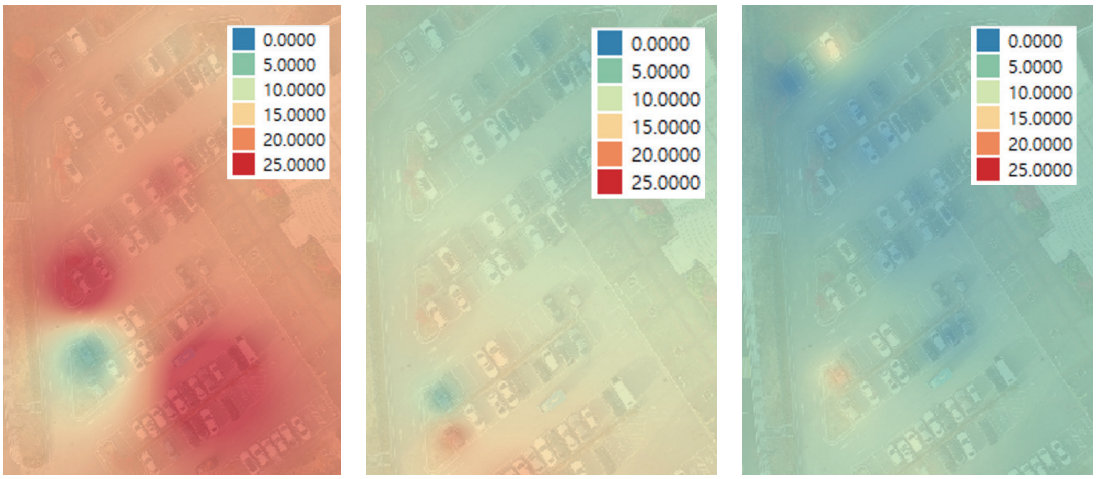


Fig. 21. (Color online) Error maps estimated by object position using IDW at camera pitches of (a) −45, (b) −60, and (c) −75°. The error unit is meter.

Table 9
Errors by camera yaw in both directions.

Camera yaw κ_c (°)	Scenario number	Number of samples	Direction	Error	
				Mean (m)	Standard deviation (m)
−180	6	122	*X*	1.586	4.773
			Y	−3.453	7.380
−90	11	355	*X*	−1.552	4.581
			Y	−4.262	3.702
0	12	139	*X*	−1.604	2.685
			Y	0.620	4.702
90	15	262	*X*	−0.019	3.399
			Y	5.716	3.094

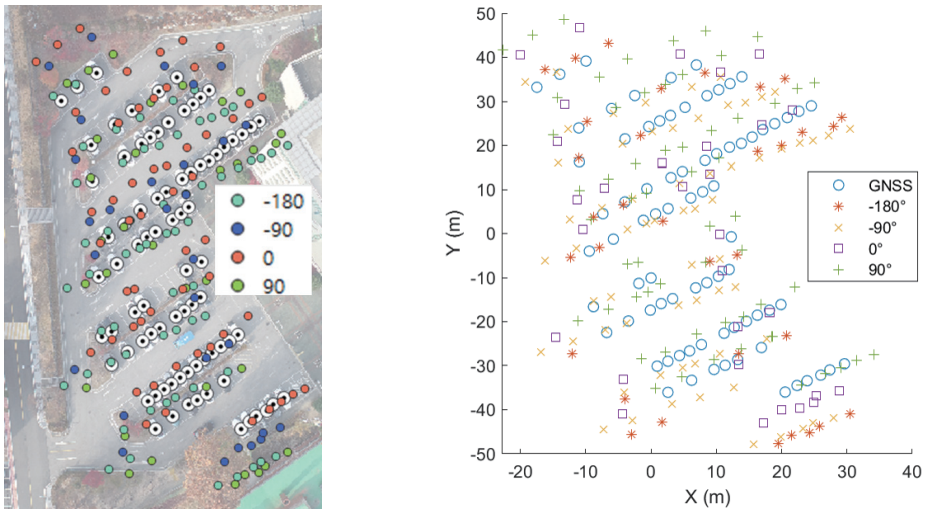
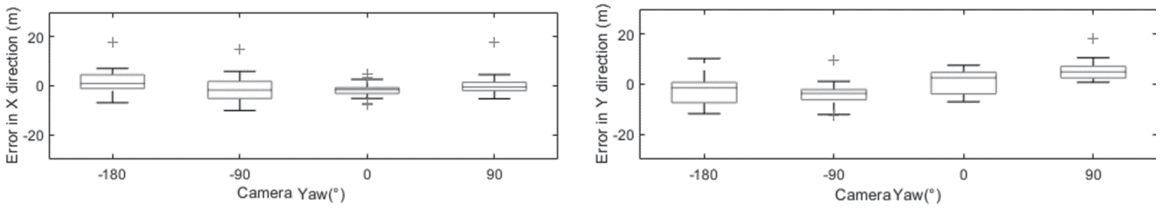


Fig. 22. (Color online) Positions of objects by camera yaw (a) in reference image and (b) in relative coordinate frame.

Table 9 lists the number of samples in each scenario and the mean and standard deviation of the errors by direction. Figure 23 shows a box plot of the errors by camera yaw for each scenario. There is no significant correlation between the camera yaw and the error in Table 9 and Fig. 23.

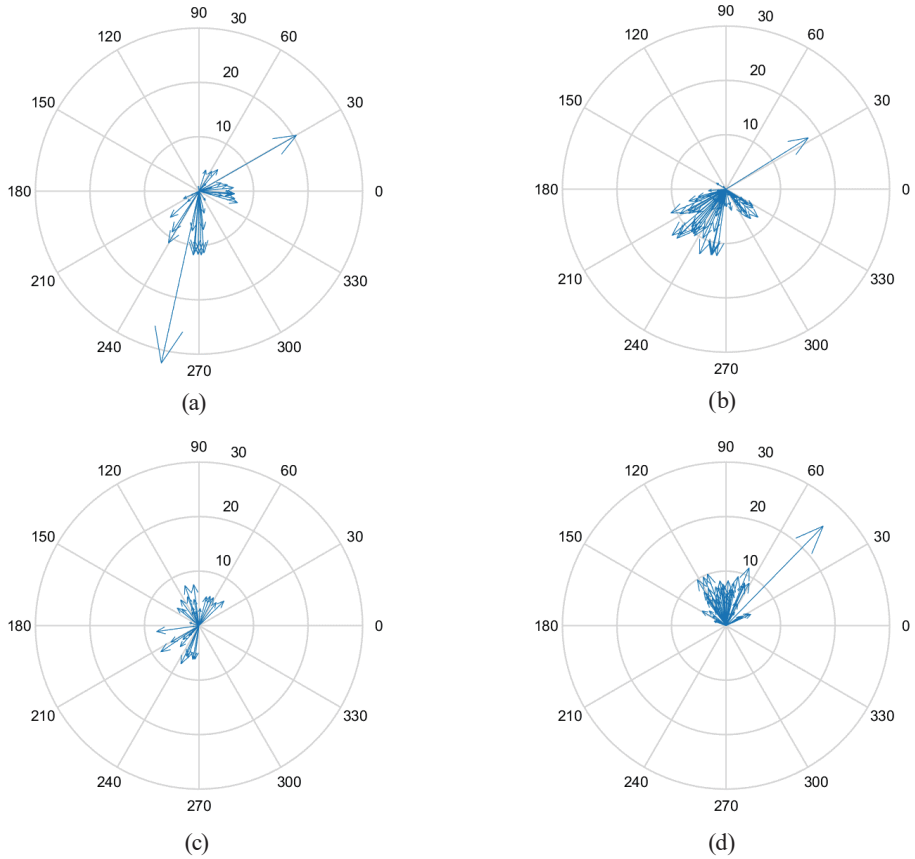
Figure 24 shows the direction and length of the error according to the yaw and verifies the correlation between the error and the camera yaw. The magnitude of the error was approximately 10 m, regardless of the direction of yaw. However, the error was biased in one or two directions. This indicates that the error direction is closely related to the camera yaw.

At a yaw of −180°, errors were primarily in the south direction, although some errors occurred in the east direction. In contrast, at a yaw of −90°, errors were mostly in the southeast direction and partly in the southwest direction. At a yaw of 0°, errors primarily occurred in the



(a) (b)

Fig. 23. Box plots of errors in (a) *X*- and (b) *Y*-directions by camera yaw.



(a) (b)

(c) (d)

Fig. 24. (Color online) Direction and size of errors for camera yaws of (a) −180, (b) −90, (c) 0, and (d) 90°. The circular unit is the number of degrees counterclockwise from east and the radial unit is meter.

northeast and southeast directions with some errors occurring in the northwest direction. In contrast, most errors occurred in the north direction at a yaw of 90°. Therefore, the yaw of the camera and the error vector are correlated.

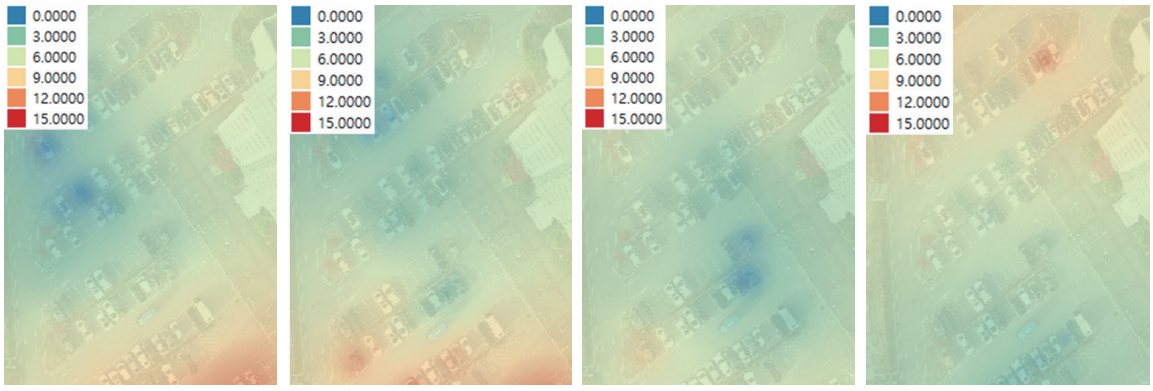
Figure 25 shows that the change in the direction of the error vector affects the distribution of the error magnitude. At yaws of −180 and −90°, the error at the upper left corner of the image is small, and the error at the lower right corner tends to be large. In contrast, for a yaw of 90°, the error is small in the lower part of the image but large in the upper part of the image. Therefore, for the positioning of an object using a UAV and a CNN, an error of 10 m or more may occur depending on the shooting position, yaw, and pitch.

4.4 Correlation between camera distortion error and posture error

The contributions of the position error of the UAV, the distortion error of the camera, and the posture error of the camera to the overall positioning error have a complex relationship. The position error of the UAV is independent of the other errors. However, although the distortion and posture errors of the camera are independent of each other, they affect each other when calculating the position. Therefore, it is necessary to understand not only the correlation between the camera distortion error and the posture error but also the proportion of the camera distortion error relative to the actual error.

The proportion of the camera distortion error relative to the total error can be determined by comparing the position error vectors of the objects measured from one calibrated image and one uncalibrated image. Since the error vector due to camera distortion is the difference between the calibrated and uncalibrated vectors, it was determined that the larger the difference between the two vectors, the larger the camera distortion error.

Therefore, the similarity between vectors and the effect of camera distortion errors are determined through the correlation coefficients, such as the length similarity of the *x*-axis, the length similarity of the *y*-axis, length similarity, and direction similarity. The results of the similarity analysis using 3562 vectors from Scenarios 1 to 26 are shown in Figs. 26 and 27 and Table 10.



(a) (b) (c) (d)

Fig. 25. (Color online) Error maps estimated by object position using IDW for camera yaws of (a) −180, (b) −90, (c) 0, and (d) 90°. The error unit is meter.

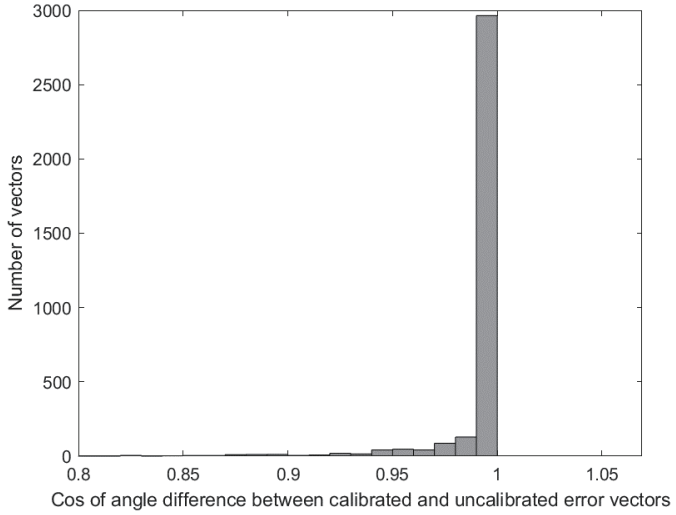


Fig. 26. Histogram of cosine of angle difference between calibrated and uncalibrated error vectors.

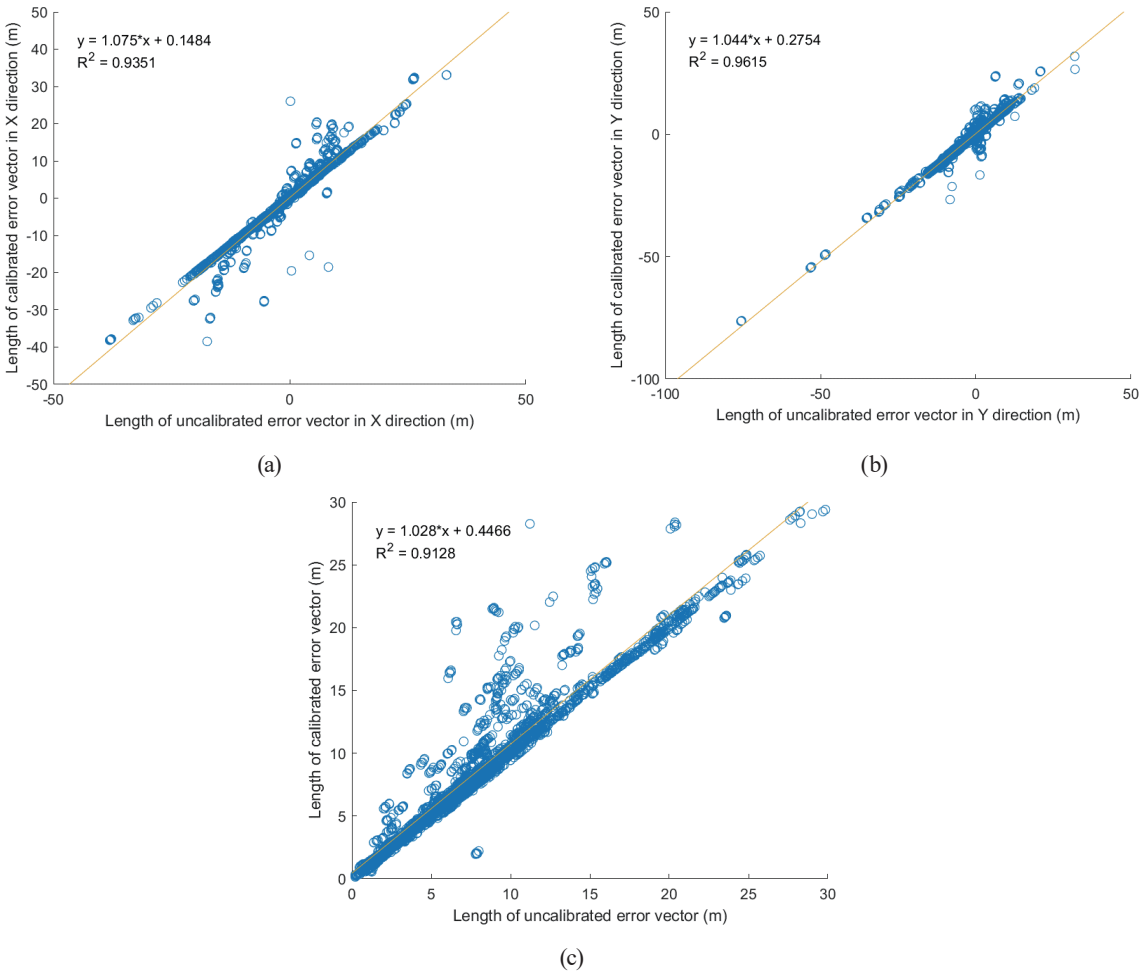


Fig. 27. (Color online) Correlation between calibrated and uncalibrated error vectors according to (a) length of *X*-direction, (b) length of *Y*-direction, and (c) length.

Table 10
Correlation between calibrated and uncalibrated error vectors.

Item	Value
*R^2 of length in *X*-direction	0.9351
R^2 of length in *Y*-direction	0.9615
R^2 of length	0.9128
Covariance	0.9736

*R: correlation coefficient

According to Figs. 26 and 27 and Table 10, there was no significant difference between the error vectors with and without calibration. Therefore, the distortion does not significantly affect the position error of the object, and the proportion of the distortion error relative to the total error is negligible.

5. Conclusions

This study focused on locating objects using UAVs and CNNs. The accuracy and error factors were analyzed, and the following conclusions were drawn:

First, the error factors originating from sensors such as MEMS, GNSS, and CMOS that can affect the direct orientation method using UAVs and CNNs were identified. The location accuracy was also calculated. Second, a method was proposed and verified for quantitatively analyzing the location error of a UAV and minimizing the error due to the UAV and its sensors. Third, error factors due to camera distortion were excluded and errors due to the camera posture were analyzed. The correlation between each element of the camera posture and the error was analyzed and identified. Finally, the effect of camera lens distortion on the overall position error was analyzed, and it was found that the position error due to the actual camera lens distortion was insignificant.

This study has several limitations. First, although the objects were located using a CNN, the effect of the error due to the CNN itself was not analyzed. Second, a detailed analysis of each element constituting the camera posture was not performed. Considering the objectives and scope of this study, the above limitations should be dealt with in a subsequent study.

Despite these limitations, we have presented a verification methodology that can qualitatively analyze errors occurring in the positioning of an object using UAVs and CNNs and proposed a method of improving the positioning accuracy. The results of this study can be utilized in various fields involving the utilization of UAVs and CNNs, as well as for the development and convergence of the two technologies.

Acknowledgments

This paper was supported by “The Ministry of the Interior and Safety” R&D program (20017423).

References

1 P. Zhu, L. Wen, X. Bian, H. Ling, and Q. Hu: Comput. Vision Pattern Recognit. **23** (2018) arXiv:1804.07437. https://arxiv.org/abs/1804.07437
2 P. Mittal, R. Singh, and A. Sharma: Image Vision Comput. **104** (2020) 104046. https://doi.org/10.1016/j.imavis.2020.104046
3 H. Yu, G. Li, W. Zhang, Q. Huang, D. Du, Q. Tian, and N. Sebe: Int. J. Comput. Vision **128** (2020) 1141. https://doi.org/10.1007/s11263-019-01266-1
4 A. A. Gebrehiwot: Flood Extent Mapping in 3D using Deep Learning from High-Resolution Remote Sensing Data: Ph.D. Dissertation, Applied Science and Technology Dept., North Carolina Agricultural and Technical State Univ., Greensboro, North Carolina (2021).
5 R. Zhang, C. Wang, X. Hu, Y. Liu, S. Chen, and B. Su: Int. J. Precis. Agic. Aviat. **3** (2020) 23. https://doi.org/10.33440/j.ijpaa.20200301.63
6 X. Deng, Z. Tong, Y. Lan, and Z. Huang: AgriEngineering **2** (2020) 294. https://doi.org/10.3390/agriengineering2020019
7 L. Wittstruck, T. Jarmer, D. Trautz, and B. Waske: IEEE Geosci. Remote Sens. Lett. **19** (2022) 2503405. https://doi.org/10.1109/LGRS.2022.3141497
8 K. Yu, Z. Hao, C. J. Post, E. A. Mikhailova, L. Lin, G. Zhao, S. Tian, and J. Liu: Remote Sens. **14** (2022) 295. https://doi.org.10.3390/rs14020295
9 L. Du, Y. Sun, S. Chen, J. Feng, Y. Zhao, Z. Yan, X. Zhang, and Y. Bian: Agriculture **12** (2022) 248. http://doi.org/10.3390/agriculture12020248
10 B. Kellenberger, D. Marcos, and D. Tuia: Remote Sens. Environ. **216** (2018) 139. https://doi.org/10.1016/s.rse.2018.06.028
11 M. Lee and S. Yeom: Int. J. Fuzzy Logic and Intell. Syst. **18** (2018) 182. https://doi.org/10.5391/IJFIS.2018.18.3.182
12 D. Biswas, H. Su, C. Wang, and A. Stevanovic: Int. J. of Geo-Inf. ISPRS **8** (2019) 259. https://doi.org/10.3390/igji8060259.
13 L. Liu: J. Am. Soc. Civ. Eng. **147** (2021) 04021017. https://doi.org/10.1061/(ASCE)SU.1943-5428.0000370
14 E. I. Vlahogianni, J. D. Ser, K. Kepaptsoglou, and I. Lana: J. Big Data Anal. Transp. **3** (2021) 1. https://doi.org/10.1007/s42421-021-00038-z
15 H. Guan, X. Lei, Y. Yu, H. Zhao, D. Peng, J. M. Junior, and J. Li: Int. J. Appl. Earth Obs. Geoinf. **107** (2022) 102677. https://doi.org/10.1016/j.jag.2022.102677
16 D. Kang and Y. Cha: Comput.-Aided Civ. Infrastruct. Eng. **33** (2018) 885, https://doi.org/10.1111/mice.12375
17 J. Ding, J. Zhang, Z. Zhan, X. Tang, and X. Wang: Remote Sens. **14** (2022) 663. https://doi.org/10.3390/rs14030663
18 S. Varela, T. L. Pderson, and A. D. B. Leakey: Remote Sens. **74** (2022) 733. https://doi.org/10.3390/rs14030733
19 H. S. Munawar, F. Ullah, D. Shahzad, A. Heravi, S. Qayyum, and J. Akram: Buildings **12** (2022) 156. https://doi.org/10.3390/buildings12020156
20 H. S. Munawar, F. Ullah, A. Heravi, M. J. Thaheem, and A. Maqsoom: Drones **6** (2022) 5. https://doi.org/10.3390/drones6010005
21 M. Altaweel, A. Khelifi, Z. Li, A. Squitieri, T. Basmaji, and M. Ghazal: Remote Sens. **14** (2022) 553. https://doi.org/10.3390/rs14030553
22 H. Saha, S. Basu, S. Auddy, R. Dey, A. Nandy, D. Pal, N. Roy, S. Jasu, A. Saha, S. P. Chattopadhyay, and T. Maity: 2018 IEEE 8th Annual Computing and Communication Workshop and Conf., Las Vagas, NV, USA (IEEE, 2018) 654. https://doi.org/10.1109/CCWC.2018.8301782
23 R. Geraldes, A. Goncalves, T. Lai, M. Villerabel, W. Deng, A. Salta, K. Nakayama, Y. Matsuo, and H. Prendinger: IEEE Access **7** (2019) 122583. https://doi.org/10.1109/ACCESS.2019.2938249
24 W. Zhi, S. Zhang-song, and G. Wen-bin: ACM Int. Conf. Proceeding Series, Xiamen, China (ACM, 2020) 95-99. https://doi.org/10.1145/3390557.3394306
25 X. Liu and Z. Zhang: Wireless Commun. Mobile Comput. (2021) 5565589. https://doi.org/10.1155/2021/5565589
26 R. D. Haamied, B. Q. Ai-Abudi, and R. H. Hassan: Iraqi J. Sci. **62** (2021) 5008. https://doi.org/10.24996/ijs.2021.52.12.37
27 A. Nassar, K. Amer, R. ElHakim, and M. ElHelw: Proc.2018 IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW, 2018) 1594-159410. https://doi.org/10.1109/CVPRW.2018.00201
28 A. Toma, N. Cecchinato, C. Drioli, G. L. Foresti, and G. Ferrin: Towards Drone Recognition and Localization from Flying UAVs through Processing of Multi-Channel Acoustic and Radio Frequency Signals: a Deep Learning Approach, https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-190/MP-IST-190-34.pdf (accessed June 2022).

29 Y. Zhang, K. Gao, and J. Zhu: Proc. 2nd Int. Conf. Mechanical, Electronic, Control and Automation Engineering, Advances in Engineering Research, Qingdao, China (MECAE, 2018) 775-779. https://doi.org/10.2991/mecae-18.2018.136
30 S. Bhoite, N. Ravi, K. Giri, and K. Gupta: Proc. 2021 Asian Conf. Innovation in Technology (ASIANCON), Pune, India (ASIANCON, 2021) 1-6. https://doi.org/10.1109/ASIANCON51346.2021.9544830.
31 Phantom 4 Pro - Product Information – DJI, Available from: https://www.dji.com/phantom-4-pro/info#specs (accessed June 2022)
32 A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam: Comput. Vision Pattern Recognit. (2017). https://doi.org/10.48550/arXiv.1704.04861
33 M. Tan and Q. V. Le: Proc. Int. Conf. Machine Learning, CA, USA (ICML, 2019). https://doi.org/10.48550/arXiv.1905.11946
34 Working Group "Assessment of Quality in Statistics": Item 4.2C: Methodological Documents Glosary: DOC. Eurostat/A4/Quality/03/Glossary, Eurostat, Room Ampere, Bech building, Luxembourg, 2–3 Oct. 2003, https://unstats.un.org/unsd/dnss/docs-nqaf/Eurostat-GLOSSARY_1.pdf (accessed June 2022).
35 Z. Zhang: IEEE Trans. Pattern Anal. Mach. Intell. **22** (2000) 1330. https://doi.org/10.1109/34.888718
36 Kakao Map: https://map.kakao.com (accessed June 2022).