An Image Processing Based Classifier to Support Safe Dropping for Delivery-by-Drone

Assem A. Abdelhak

Technological University Dublin, assem.abdelhak@tudublin.ie

Alan Hicks

Manna Drone Delivery,

Dan Moss

Manna Drone Delivery

See next page for additional authors

Recommended Citation

A. Alsawy, A. Hicks, D. Moss and S. Mckeever, "An Image Processing Based Classifier to Support Safe Dropping for Delivery-by-Drone," 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS), Genova, Italy, 2022, pp. 1-5, doi: 10.1109/IPAS55744.2022.10052868

Funder: the European Union’s Horizon 2020 Research and innovation

Authors
Assem A. Abdelhak, Alan Hicks, Dan Moss, and Susan McKeever

An Image Processing Based Classifier to Support Safe Dropping for Delivery-by-Drone

Assem Alsawy
Technological University Dublin (TU Dublin)
Ahram Canadian University (ACU), Egypt
Dublin, Ireland
assem.abdelhak@tudublin.ie

Alan Hicks
Manna Drone Delivery
Dublin, Ireland
alan@manna.aero

Dan Moss
Manna Drone Delivery
Dublin, Ireland
dan.moss@manna.aero

Susan Mckeever
Technological University Dublin (TU Dublin)
Dublin, Ireland
susan.mckeever@tudublin.ie

***Abstract*—Autonomous delivery-by-drone of packages is an active area of research and commercial development. However, the assessment of safe dropping/ delivery zones has received limited attention. Ensuring that the dropping zone is a safe area for dropping, and continues to stay safe during the dropping process is key to safe delivery. This paper proposes a simple and fast classifier to assess the safety of a designated dropping zone before and during the dropping operation, using a single onboard camera. This classifier is, as far as we can tell, the first to address the problem of safety assessment at the point of delivery-by-drone. Experimental results on recorded drone videos show that the proposed classifier provides both average precision and average recall of 97% in our test scenarios.**

***Index Terms*—Unmanned Aerial Vehicles, UAV, Autonomous drone, Drone delivery, Image processing, Segmentation.**

I. Introduction

Drones are now used in many fields, including security, surveying and mapping, construction, aerial videos recording, energy and utilities, public safety and emergency, infrastructure and transport, mining, and goods delivery [1] [2]. Utilising techniques in computer vision and artificial intelligence enables drones to operate autonomously, akin to intelligent robots, qualified for carrying out duties that normally require human intervention [3].

Commercial drones are now equipped with a number of self-managing autonomous functions, capable of performing tasks such as navigation and landing without human control [4]. However, there are still functions that require human supervision [5]. Autonomous delivery-by-drone refers to the fully automated delivery of packages by drones to planned destinations. It aims to provide a cost effective and environmentally cleaner solution to the last mile problem of delivery to consumer, and is an active area of research and commercial development. The *dropping zone* refers to the ground area where the drone lowers or drops the package in a controlled manner. During the process, the drone flies to the drop zone which is typically pre-defined with the customer, rather than located autonomously. The package is then lowered slowly to the ground, and the connecting rope either pulled up, or dropped with the package. A challenging aspects of automated delivery-by-drone is to ensure that the dropping zone is safe for dropping during this time, such that unforeseen interventions such as vehicles or pets entering at the point of delivery can be detected. A classification process can be used to do the initial inference of the dropping zone as a safe or unsafe area. For the delivery duration, it is then necessary to find out if there is any change in the area of delivery and whether it is necessary to abort the delivery.

This project has received funding from the European Union's Horizon 2020 Research and innovation Programme under the Marie Skłodowska-Curie Co-funding of regional, national and international programmes Grant agreement No: 847402.

A. *Image processing*

Image processing in Computer Vision (CV) is concerned with the analysis of images through the tuning of various features and factors with the image [6]. Image segmentation creates a pixel-by-pixel mask for objects in the image. Binary image segmentation can be considered as the simplest form of image segmentation. In binary segmentation the entire image is split into two categories. The typical use of binary classification is to differentiate foreground objects of an image from the background [7].

Image processing provides simple fast solutions for multiple computer vision problems. Given the limited processing and size of lightweight drone devices, we have looked to image processing for an efficient, flexible, fit-for-purpose solution, rather than a more complex learning based approach.

B. *Overall Aim*

Automated delivery requires a rapid safety assessment of the target dropping zone, and tracking of the safety of the zone until the dropping process completes. However, there is no automated solution for autonomous assessment and overseeing of a safe drone delivery process.

Solving this problem of ensuring safe package dropping/delivery is a worthwhile step towards converting drone delivery into full automation. This work aims to develop a simple and fast classification approach using an image segmentation method integrated with a motion detector to assess the safety of a designated dropping area during the package dropping operation.

II. Related Work

While there is no body of published research investigating the safe dropping for delivery drones to date, there is an overlap between this problem and the problem of safe autonomous drone landing. Therefore, research on safe landing detection are a worthwhile area of related work.

In the autonomous normal drone landing, the drone can recognize a safe landing zone using either visual or non-visual information sources.

A. Visual autonomous landing mechanism

Approaches using visual sources rely on enabling the drone to recognize known markings from images provided by onboard cameras. Landing sites can be identified visually by marking them with certain shapes such as squares, circles, and H-shapes, or by a QR code. In 2017, Y. Huang et al. applied The histogram of oriented gradients (HOG) and linear support vector (SVM) algorithms to discover known landmarks [8]. Also, in 2017, PH Nguyen introduced a marker based-tracking algorithm based on the visible-light camera [9]. In 2020, MFR Lee et al. presented an airborne vision system with a fully convolutional neural network to recognize landing markers [10]. In 2020, C Minhua, et al. proposed an augmented reality tag (ARtag) which be used as the target marker to ensure that the drone can identify the landing area [11]. In 2021, G.Nio et al. proposed an autonomous drone landing scheme to land on a mobile unmanned ground vehicle (UGV) with high accuracy [12].

B. Non-visual autonomous landing mechanism

Using non-visual sources, landing can be on an open area such as a wide open field using GPS information. GPS is commonly used in such cases. Some research approaches supplement GPS information with the use of onboard sensors to more accurately detect a known landing zone. In 2018, Masataka Kan et al. [13] proposed a method to enable the drone to autonomously navigate, hover, and land using GPS information. This research introduced GPS calibration to overcome the deviations found using GPS sensors [13]. In 2019, Patrik et al. [14] used Global Navigation Satellite System (GNSS) in drone navigation and landing to reduce the positional deviation of the landing position between the calculated landing site and the target landing site.

C. The difference between the landing process and the delivery dropping process

Landing research has addressed the issue of navigating to a planned landing area using GPS information, and recognizing the landing area by detecting landing markers or QR codes. These different landing methods are interesting as a mechanism for choosing and finding a predetermined landing point. However, since delivery by drone is adhoc to many destinations, a mechanism that does not rely on pre-marking at ground level is required.

A regular landing, as set out in Figure 1, typically uses GPS information to perform a landing on a landmark. The autonomous drone navigates to the destination point specified by the GPS coordinates, hovers, and then instantly lands or recognizes a landmark to land on.

Fig. 1. Steps of drone normal landing on the landmarks

Figure 2 shows the steps of the dropping operation. In the package dropping scenario, the drone navigates to the dropping point using GPS information and hovers to assess the safety of the dropping area. The dropping algorithm focuses on whether the dropping zone is clear of any obstacle at the beginning of the dropping and during the dropping process.

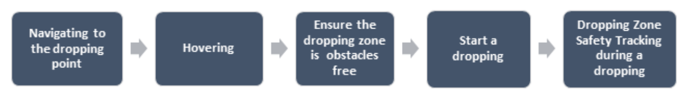

Fig. 2. Steps of drone delivery dropping

III. Image processing based Classifier for Safe Dropping Zone Assessment

In our approach, an onboard camera is mounted on the drone and directed vertically downward. When the drone reaches the destination drop point using the GPS information, the camera starts live streaming the area directly below the drone. To assess the safety of the dropping zone, a three-step assessment is applied to the dropping area image within each frame. Figure 3 demonstrates the sequence of the 3-step assessment: image pre-processing, binary segmentation, and threshold function. A motion detector is also applied to achieve better performance by removing redundant checking where no change has occurred.

Fig. 3. Three-steps assessment

A. Pre-Processing

Image pre-processing operations are used to clean , remove any distortion in the image, and to convert the image to a lower level of abstraction [15]. These operations are applied in the sequence shown in figure 4.

Fig. 4. The sequence of pre-processing operations

The bounded box is defined and cropped in each frame to match the agreed location and size of the target dropping zone. The gray-scale conversion is then used to simplify the computational requirements [16]. Finally, a Gaussian blur transformation is applied to each pixel value to smooth the image and reduce image noise [17].

B. Binary Segmentation

The main goal of the binary segmentation process is to transform the image representation to a two category structure, whereby it is simpler to detect any obstacle or unsafe object from the empty space background. The histogram-based thresholding method is adopted to determine the threshold value β that separates the two categories. Each pixel value in the segmented image is calculated as in equation 1.

$$p_{ij} = \begin{cases} 1 & P_{ij} \geq \beta \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where (P_{ij}) is the value of the pixel in position (i,j)

C. Threshold Function to Determine Empty Space

In this phase, the ratio of the segment classes in the cropped image are calculated. These values specify the ratio of all the objects that may be present in this dropping zone, and thus the ratio of empty space that may be available for dropping. The ratio of empty space (I) can be computed as in equation 2.

$$I = \frac{\sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij}}{N} \tag{2}$$

Where N is a number of pixels in the cropped image, l is a number of rows, k is a number of columns.

Hence, a threshold function is applied to classify the dropping zone in the frame as safe or unsafe, as in equation 3.

$$Output = \begin{cases} \text{Safe} & I \geq \alpha \\ \text{Unsafe} & \text{otherwise} \end{cases} \tag{3}$$

Where α is the minimum acceptable empty space ratio in the dropping region, α can be determine by delivery company according to the package size and the area of dropping zone.

Figure 5 demonstrates the 3-step assessment for a frame in a live stream of a dropping zone: The original frame is shown in (a); (b) shows the image of the selected dropping zone after the first step, pre-processing; (c) represents the segmented image of the dropping zone resulting from the second step; (d) is the step of calculating and using the ratio of empty space within the dropping zone. If this ratio $\geq \alpha$, the zone is considered safe for dropping, otherwise, it is not safe at that moment.

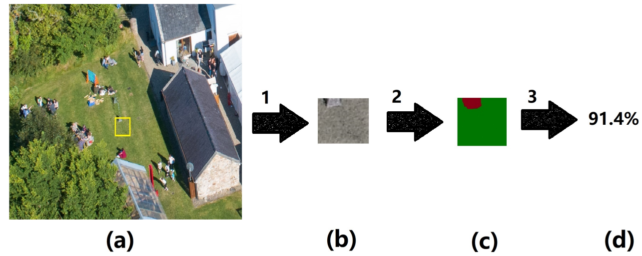

Fig. 5. 3-step assessment of a frame in a live stream Example 1

Figure 6 demonstrates the 3-step assessment for another sample frame, with a higher proportion of non-empty space (and therefore, more likely to exceed the safety threshold for empty space) in a live stream of a dropping zone.

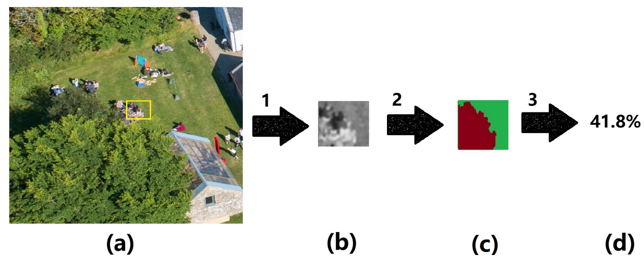

Fig. 6. 3-step assessment of a frame in a live stream Example 2

D. Motion detection

In order to reduce redundant safety assessment, a motion detection function is proposed to check if any changes occurred in the dropping region. The 3- step assessment will only be implemented if a change is detected. The use of motion detection will avoid repeating all previous steps in each frame if no changing has happened.

While the drone is hovering during delivery, motion can be detected by simply detecting changes between every two successive frames. Sum of absolute differences (SAD) is the adopted method, where SAD is a measure of the dissimilarity between two images. It is the absolute numerical value of each pixel P in the location (i, j) in an image is subtracted from the corresponding pixel $P^{'}$ in another image, as in the form $|P_{ij} - P^{'}_{ij}|$.

Equation 4 determines the sum of absolute differences (SAD) to compare two successive frames in the drone camera live stream.

$$SAD = \sum_{i=1}^{k} \sum_{j=1}^{l} |P^{m}_{ij} - P^{m+1}_{ij}| \tag{4}$$

Where m and $m+1$ are consecutive frames numbers, l is the number of rows in each frame, and k is the number of columns in each frame. In the case of identical frames, SAD equals zero which means there is no movement was detected. However, a small value can be tolerated as a difference between frames due to the shake of the mounted camera because of the drone hovering.

Algorithm 1 Safe Drone Delivery Dropping Algorithm

Require: Maximum dropping time in second: T, Prepossessed streaming frames: $f_1, f_2, ...f_T$, The thresholds: α, β, ϵ
Ensure: Dropping area status Safe/ Unsafe

```
t ← 15 * T                                         ▷ One frame is processed every 1/15 second
while t ≠ 0 do
    m ← 15 * T − t + 1
    if m > 1 then
        if Σ_{i=1}^{k} Σ_{j=1}^{l} |P_{ij}^{m} − P_{ij}^{m−1}| < ε then    ▷ Motion detection
            t ← t − 1
            continue                               ▷ Skip to the next frame
        end if
    end if
    for i ← 1 to k do                              ▷ Binary segmentation
        for j ← 1 to l do
            if P_{ij}^{m} > β then
                p_{ij}^{m} ← 1
            else
                p_{ij}^{m} ← 0
            end if
        end for
    end for
    if (Σ_{i=1}^{k} Σ_{j=1}^{l} p_{ij})/N ≥ α then   ▷ Threshold function
        Output ← Safe
    else
        Output ← Unsafe
    end if
    t ← t − 1
end while
```

Algorithm 1 presents a delivery drone safety dropping algorithm based on image processing.

IV. Experimental Tests

To assess the accuracy of the proposed classifier, a test dataset was compiled, with the criteria that it must represent real drone footage, of both safe and unsafe outdoor delivery zones, including transitions between both states. A range of dropping surfaces were required that would reasonably represent dropping spots supplied by a customer for dropping. We used 18 publicly available recorded drone videos. The video have frame rates of 15 frame/second, recorded by drones from different heights (12m, 14m, 16m, 18m). Each video is 30 seconds long, and uses different side lengths of dropping area boxes (2m, 3m, 4m, 5m). Nine recorded videos are for horizontal clear dropping surfaces. The remaining nine are for mottled surfaces such as grass, concrete, and brick surfaces speckled with irregular markings, shades, and colors, Various objects are inserted at random times for 15 separate seconds in each of the 18 videos, to simulate zone changes and unsafe dropping areas.

In each video, the motion detection function is applied to each frame to detect any movement by discovering any change from one frame to the next frame, as per our algorithm. If no motion is detected, it means there is no changes in the safety state. Otherwise, the classifier assesses the safety using the 3-step assessment (pre-processing, binary segmentation method, threshold function). Table I displays the results of applying the proposed classifier to the recorded videos dataset.

The confusion matrix for these experiments in table II demonstrates that:

Precision (Safe) = 100%, Recall (Safe) = 95.19%.
Precision (Unsafe) = 95.41%, Recall (Unsafe) = 100%.

This approach assumes that the user has picked a viable drop spot to begin with, that has been accurately specified - for examples, the spot will not for example be the roof of a house, a sloped surface, a watery surface. Therefore the proposed approach is not concerned with the type of the drop spot surface. The proposed classifier is reliable for assessing the safety of the dropping area, is capable of detecting objects in the area, and can detect any change in the safety state during the whole dropping process.

TABLE II
Confusion matrix

		Predicted Class	
		Safe	Unsafe
Actual Class	Safe	3855	195
	Unsafe	0	4050

TABLE I

ACCURACY OF THE PROPOSED CLASSIFIER APPLIED TO THE RECORDED VIDEOS DATASET, AS NUMBER OF CORRECT FRAMES AGAINST ORIGINAL AND PERCENTAGE ACCURACY

Type	Actual output	no. of frames	no. of correct results	Accuracy
Empty flat horizontal surfaces	Safe	2025	2025	100.00%
Empty mottled surfaces	Safe	2025	1830	90.37%
Flat horizontal surfaces with objects	Unsafe	2025	2025	100.00%
Mottled surfaces with objects	Unsafe	2025	2025	100.00%

V. CONCLUSION

The proposed image processing-based classifier aims to automate the delivery dropping process as a step towards a fully autonomous delivery drone. This classifier is a simple and fast classifier that evaluates the safety of the dropping zone for a delivery drone. Experiments show good results in recognising the safety status of the surface, for a range of horizontal and mottled surfaces, and detects the changes when objects interfere during the dropping operation. The approach currently assumes that a viable dropping zone has been previously designated prior to the drone journey, and that the drone has navigated correctly to this point.

As a future work, it is planned to develop a more sophisticated approach that be able to interact with the environment with a semantic and spatial perception. This will aim to deal with more challenging use cases such as the requirement to autonomously identify a dropping zone, or to relocate to a safe nearby spot if the the primary dropping zone is not safe.

REFERENCES

[1] Mohammed, F., Idries, A., Mohamed, N., Al-Jaroodi, J., and Jawhar, I. (2014, May). UAVs for smart cities: Opportunities and challenges. In 2014 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 267-273). IEEE.

[2] Ayamga, M., Akaba, S., and Nyaaba, A. A. (2021). Multifaceted applicability of drones: A review. Technological Forecasting and Social Change, 167, 120677.

[3] Hussein, M., Nouacer, R., Ouhammou, Y., Villar, E., Corradi, F., Tieri, C., and Castiñeira, R. (2020, August). Key enabling technologies for drones. In 2020 23rd Euromicro Conference on Digital System Design (DSD) (pp. 489-496). IEEE.

[4] Gamulescu, O. M., Musetoiu, O. M., and Leba, M. (2017). THE SELF-PILOTING FUNCTION OF A MULTICOPTER. Annals of'Constantin Brancusi'University of Targu-Jiu. Engineering Series, (4).

[5] Rokhsaritalemi, S., Sadeghi-Niaraki, A., and Choi, S. M. (2018, October). Drone trajectory planning based on geographic information system for 3d urban modeling. In 2018 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 1080-1083). IEEE.

[6] Borstell, H. (2018). A short survey of image processing in logistics. In 11th International Doctoral Student Workshop on Logistics (pp. 43-46). Magdeburg: Universität Magdeburg.

[7] Kumar, A. S. (2021). Image Segmentation and Object Recognition. Journal of Research Proceedings, 1(2), 101-112.

[8] Huang, Y. P., Sithole, L., and Lee, T. T. (2017). Structure from motion technique for scene detection using autonomous drone navigation. IEEE Transactions On Systems, Man, And Cybernetics: Systems, 49(12), 2559-2570.

[9] Nguyen, P. H., Kim, K. W., Lee, Y. W., and Park, K. R. (2017). Remote marker-based tracking for UAV landing using visible-light camera sensor. Sensors, 17(9), 1987.

[10] Lee, M. F. R., Aayush, J., Saurav, K., and Anshuman, D. A. (2020, August). Landing Site Inspection and Autonomous Pose Correction for Unmanned Aerial Vehicles. In 2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS) (pp. 1-6). IEEE.

[11] Minhua, C., and Jiangtao, G. (2021). Robotics Lab 2015 Project: Autonamous Landing Mobile Robotics Lab 2015: Autonamous Landing on a Moving Target. Computer, 1(2022).

[12] Niu, G., Yang, Q., Gao, Y., and Pun, M. O. (2021). Vision-Based Autonomous Landing for Unmanned Aerial and Ground Vehicles Co-operative Systems. IEEE Robotics and Automation Letters, 7(3), 6234-6241.

[13] Kan, M., Okamoto, S., and Lee, J. H. (2018). Development of drone capable of autonomous flight using GPS. In Proceedings of the international multi conference of engineers and computer scientists (Vol. 2).

[14] Patrik, A., Utama, G., Gunawan, A. A. S., Chowanda, A., Suroso, J. S., Shofiyanti, R., and Budiharto, W. (2019). GNSS-based navigation systems of autonomous drone for delivering items. Journal of Big Data, 6(1), 1-14.

[15] Vidal, M., and Amigo, J. M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. Chemometrics and Intelligent Laboratory Systems, 117, 138-148.

[16] Saravanan, C. (2010, March). Color image to grayscale image conversion. In 2010 Second International Conference on Computer Engineering and Applications (Vol. 2, pp. 196-199). IEEE.

[17] Gedraite, E. S., and Hadad, M. (2011, September). Investigation on the effect of a Gaussian Blur in image filtering and segmentation. In Proceedings ELMAR-2011 (pp. 393-396). IEEE.