



UAV Landing Platform Recognition Using Cognitive Computation Combining Geometric Analysis and Computer Vision Techniques

J. A. García-Pulido¹ · G. Pajares² · S. Dormido¹

Received: 5 July 2019 / Accepted: 30 October 2021
© The Author(s) 2022

Abstract

Unmanned aerial vehicles (UAVs) are excellent tools with extensive demand. During the last phase of landing, they require additional support to that of GPS. This can be achieved through the UAV's perception system based on its on-board camera and intelligence, and with which decisions can be made as to how to land on a platform (target). A cognitive computation approach is proposed to recognize this target that has been specifically designed to translate human reasoning into computational procedures by computing two probabilities of detection which are combined considering the fuzzy set theory for proper decision-making. The platform design is based on: (1) spectral information in the visible range which are uncommon colors in the UAV's operating environments (indoors and outdoors) and (2) specific figures in the foreground, which allow partial perception of each figure. We exploit color image properties from specific-colored figures embedded on the platform and which are identified by applying image processing and pattern recognition techniques, including Euclidean Distance Smart Geometric Analysis, to identify the platform in a very efficient and reliable manner. The test strategy uses 800 images captured with a smartphone onboard a quad-rotor UAV. The results verify the proposed method outperforms existing strategies, especially those that do not use color information. Platform recognition is also possible even with only a partial view of the target, due to image capture under adverse conditions. This demonstrates the effectiveness and robustness of the proposed cognitive computing-based perception system.

Keywords Image color segmentation · Landing platform · UAV · Pattern recognition · Decision-making · Recognition probability · Perception system · Artificial intelligence · Cognitive computation

Introduction

Problem Definition

Autonomous landing of UAVs is a complex task that requires particular attention during the final phase of such where the GPS is ineffective. A novel cognitive computation approach based on perception and decision-making is proposed to recognize a colored man-made platform (target) on which the quadrotor, Fig. 1, must land safely.

The strategy combines computer vision and pattern recognition techniques along with a geometric analysis to make any final decisions. The cognitive approach outperforms several existing methods, including the one described in [1], which uses a monochrome landing platform [2].

Other groups have designed markers and platforms with various algorithms to address this problem. Shape recognition algorithms [3], based on UAV vision (perception) systems [4, 5] provide a clue as to the approach to the recognition task.

The images of the platform must frequently be acquired under extreme conditions, causing blind areas due to sunshine/reflections, or perspective distortion, because of camera pose.

This requires the application of specific intelligent techniques to transfer human reasoning (perception, recognition, and decision-making) to the computational cognitive system which, to the best of our knowledge, has not to date been addressed.

✉ J. A. García-Pulido
jgarcia5695@alumno.uned.es

¹ Department of Computer Science and Automatic Control, Higher Technical School of Computer Science Engineering ETSII (Escuela Técnica Superior Ingeniería Informática), Universidad Nacional de Educación a Distancia, 28040 Madrid, Spain

² Instituto del Conocimiento (Knowledge Institute), Universidad Complutense, 28040 Madrid, Spain



Fig. 1 The UAV flying during the testing process

With such purpose, we address the same problem as in [1], i.e., the identification of the landing platform with the perception (vision) system on board the UAV but while overcoming the problems mentioned above, while an effective and robust cognitive system is achieved. The same quadcopter md-400 UAV (Fig. 1), along with a Samsung Galaxy S6 mobile device, has been used.

Related Works

Different active and passive methods, as based on computer vision strategies, have been proposed as approaches to precise autonomous UAV landing [6, 7]. For guidance during landing, active methods require external signals such as lamps [8] or infrared (IR) emitters [9, 10]. IR camera-based systems installed on the ground detect a near-IR laser lamp fixed to the nose of the UAV [11]; stereo IR vision systems have also been proposed [12–14]. Alternatively, passive systems only require the use of perception-based techniques on board. Some approaches attempt to recognize the UAV in flight, which is guided and landed from a control ground station [15, 16]. Other methods determine a secure landing area by analyzing the terrain [17, 18], as based on machine learning [19], by mounting either optical flow sensors [20, 21], or IR light-emitting diode (LED) lamps [22], by 3D terrain reconstruction [23–26], or based on 3D positioning with two cameras on the ground for fixed-wing UAVs [6]. Another common practice is the use of markers on the ground, i.e., landing platforms with specific designs, where tailored image processing techniques can recognize these markers even the UAV is in motion [27–32], including on board ships [33, 34]. We have designed a landing platform that uses colored markers, i.e., that includes spectral information. These markers are recognized by the cognitive

system using a monocular RGB-based camera built into the mobile device.

Methods using mixed black/white markers or with unique color on platforms are classified as follows:

1. *Based on specific geometric figures* including circles, ellipses, polygons (squares, pentagons): In our previous work [1], we designed an expert system which provides both the angle of orientation and the degree of confidence in the recognition of a black and white platform [2]. Lange et al. [35, 36] proposed a figure similar to that proposed in [37] based on several concentric white rings on a black background. Nguyen et al. [7] used three concentric circles defining eight black/white areas. Polvara et al. [38] applied convolutional neural networks (CNNs) to detect a black circle enclosing a black cross on a white background.

Cocchioni et al. [39] designed a new marker based on a large circle and a second, much smaller circle, together with two small equilateral triangles of different sizes. Li et al. [40] also created a marker (pattern) that used two concentric circles with a geometric figure inside. Chen et al. [41] proposed a landing platform based on two concentric circles, black inside and light gray outside, with an irregular pentagon inside the black circle. The recognition algorithm used the well-known faster regions with convolution (Faster R-CNN) technique to recognize the figure based on its feature map. Sharp et al. [42] also used black-white squares with the aim of applying corner detection as a descriptor, encountering identical difficulties and limitations. The use of complex figures, based on curvatures, causes distortions because of image perspective projection. Moreover, the absence of relevant spectral information (color) produces saturated areas with high reflectivity due to intense illumination sources (e.g., the sun). Missed areas ensue where descriptors of full figures, corners, or feature maps fail.

2. *Based on letter-shapes*: Wang et al. [43] described a monochromic H-shaped figure [44] enclosed in a circle. The image was binarized and segmented to detect the corner (interest) points of the figure and reveal the geometric correlation between them. Zhao and Pei [45] described a single green H-shaped figure by applying the speeded-up robust features (SURF) descriptor [46, 47] with matching, which is invariant to small rotations and scale changes. Saripalli et al. [37, 48] used a white H-shaped figure and applied the following computer vision techniques: filtering, thresholding, segmentation, and labeling of related components. Guili et al. [49] designed a T-shaped gray figure painted with high-emissivity black powder for use with infrared systems, which allows the determination of the relative orientation angle between the platform and the UAV. These letter-shaped markers are based on a large figure with a unique color (green or gray), but do not prevent saturation, and consequently blind regions, because of reflections or distortions.

3. *Based on fiducial tags*: ‘AprilTags’ [50, 51] were used by several authors [30, 52–54], which were based on edge detectors and designed with low image resolution and intended to deal with occlusions, rotations, and lighting variation. Subsequent improvements have been described [54]. ArUco markers, similar to AprilTags, were used in OpenCV [55]. Chaves et al. [56] used this type of marker with a Kalman filter to guide the drone during the final phase of the landing sequence to land UAVs on ships. However, it is well known that edges are noise-sensitive features where incorrect detection can result from broken edges. Araar et al. [57] used the AprilTags system to generate markers that are printed over the surface of the landing platform. Since these markers are based on black and white, the missing parts problem stubbornly persists.

In summary, several approaches [37, 41, 43, 44, 49] used a single figure producing missing regions in the image due to saturation from direct illumination that causes reflections. Other methods [1, 30, 35, 40, 52–55, 58] used (intertwined or interspersed) figures or complex figures based on circles [59] and/or ellipses to address the missing parts problem, some of which [1] allow a certain plane distortion, but others, like [35], assume the UAV is parallel to the ground during the landing phase, ignoring the angle of inclination. Edges and interest points are also considered with high noise sensitivity [42, 43, 45, 49].

Contributions

Table 1 classifies and summarizes the previously described methods. Active or passive approaches are distinguished according to category and sub-category (I, II, and III). A description with analysis and comments describing drawbacks are shown in Table 1. The proposed new approach to autonomous landing addresses these shortcomings by making the following contributions:

1. The use of spectral information based on different colors, which reduces the likelihood of intensity saturation due to direct/indirect sunlight and compensates for high variability in outdoor environments.
2. Regions that are more robust than edges, from the point of view of the missing parts problem: these are designed to deal with blind spots. Indeed, the recognition of the platform is still possible when only half of the markers are identified (3), even if they are only partially perceived.
3. Combination of spectral and geometric information under a decision-making algorithm that uses the combination of a number of figures detected according to their color and relative positions. According to the number of figures detected for each region, a probability is computed based on the geometric relationships.

4. Recognition can be carried out at different distances and inclination angles between the vision system and the platform. This is not affected by distortions in the figures because of the perspective projection.

To develop our method, using color-based techniques, the following image segmentation methods have been studied:

1. Energy optimization under different approaches: (A) minimizing a higher-order color model built from a normalized color foreground and background histograms [60]; (B) minimizing the conductance of a graph consisting of nodes (pixels) and edge weights representing image intensity changes [61]. In the latter case, the user selects seeds belonging to foreground and background. Our platform is built with different foreground colors and is placed in complex outdoor environments with multiple colors around the target. Thus, two color-based histograms become ineffective, requiring a multi-classification approach [61] without user interactivity; (C) high order minimization is also applied for stereo image segmentation [62], where the energy function utilizes the corresponding relationships of the pixels between the stereo image pairs. The proposed cognitive approach is monocular, and no stereo-pairs can be obtained while flying the drone during the landing phase; hence, the recognition of the platform must guide this operation and the drone stabilization is not guaranteed; (D) in [63], seeds are initially interactively selected. By then applying the lazy random walk algorithm, superpixels and their boundaries are defined and compacted using energy minimization, and these boundaries are adjusted to existing objects in the scene.
2. Motion in videos, assuming that any movement of the platform appears in the image sequence during landing. In this regard, some approaches [64] apply sub-modular optimization by selecting a small number of initial trajectories for each entity. Color or texture information is extracted for each moving entity to build the energy and moving clusters. The apparent movement of objects in the sequence during landing is generated by the drone, but not for moving entities within the scene. The static platform, once it has been identified and located in the image must guide the drone. This is the opposite problem to that described in [64].

With the proposed cognitive approach, candidate regions, that is, those potentially belonging to the platform, are identified with their corresponding bounding boxes, which enclose the background (platform) and the figures inside. For each region, a probability is computed. In this way, the method can be used in the future as a proposal region approach in the context of convolutional neural networks. This is the

Table 1 Summary of comparisons of previous and proposed designs

Type	Category	Sub-category I	Sub-category II	Sub-category III	Description/analysis	Comments/drawbacks
Active (on-board devices)	With landing platform	Single color-based/black and white based on visible camera	Multiples markers	Mandatory markers [1, 7]	<ul style="list-style-type: none"> Specific figures for recognition Geometric figure relations Single figure split into sub-figures 	<ul style="list-style-type: none"> Complex figures (circles/ellipses) very sensitive to image perspective distortion When principal figures are missing, no starting point can be detected for identification
			Main figure/single marker [27, 29, 34, 37, 38, 41, 43–45, 48, 49, 59, 66]	Marker size difference [35, 36, 40, 42, 65]	<ul style="list-style-type: none"> Squares or circles with different sizes or radii Main figure built with sub-figures (squares/circles) Recognition based on the different sizes for each sub-figure 	<ul style="list-style-type: none"> Simple squares may be confused by natural items in outdoor environments (straw cubes cut by a mower) Eccentricity depends on the perspective. Circles are affected by distortions, being perceived as ellipses
Passive (ground-fixed devices)	Fixed-wing/Quadrotor UAVs	Multiple colors-based on visible light camera	Multiples figures	No figures restrictions (proposed cognitive method)	<ul style="list-style-type: none"> Synthetic square markers (AprilTags, ArUco) Computer vision-based techniques to identify black/white parts inside a wide black border and an inner binary matrix 	<ul style="list-style-type: none"> Sensitivity of squares to distortions Although some methods replicate the codes to different sizes, distance remains a problem
			Thermal camera	Main figure/single marker	Letter-shaped marker [9, 10]	<ul style="list-style-type: none"> Recognition based on specific characteristic of regions in the marker (T-shaped, H-shaped, X-shaped, or a single geometric figure like a circle, square, etc.) Binarization, thresholding, segmentation, SURF, etc. are used for feature extraction with descriptors
Passive (ground-fixed devices)	Fixed-wing/Quadrotor UAVs	Visible light camera	Runway	Camera calibration is not required [8]	<ul style="list-style-type: none"> Recognition with only half of the figures Relationships between sub-figures allowing missing parts Robust method to work under adverse conditions 	<ul style="list-style-type: none"> Sub-figures with light colors (Fig. 2) could disappear because of reflections Training required to adjust colors (Table 2)
			Without landing platform (do not use a specific platform)	Anywhere/not specific target [18]	Database with landing areas [19, 67]	<ul style="list-style-type: none"> Landing areas are defined as smooth textures Texture descriptors allow the identification (by learning) of textures areas for similarity with reference patterns
Passive (ground-fixed devices)	Fixed-wing/Quadrotor UAVs	Based on pan-tilt unit (PTU)	Stereo-vision	Infrared cameras [6, 12, 13]	<ul style="list-style-type: none"> 3D terrain reconstruction is applied based on a gray-scale matrix vision camera (mvBlueFOX) The relative position between the target landing and the UAV is obtained based on the information provided by a PTU vision-system at different positions 	<ul style="list-style-type: none"> Appropriate textured areas are required with sufficient sizes for landing A textured database is required Flat terrain is required High computational cost Assumable 3D reconstruction errors are required
			Visible light cameras	Trinocular system	FireWire connection [15]	<ul style="list-style-type: none"> Color (red, yellow, green, and blue) landmarks are used for orientation and position determination CamShif algorithm is applied for detection and tracking the landmarks from the three images
Passive (ground-fixed devices)	Fixed-wing UAV	NIR cameras	Based on array system	Laser lamp [11]	<ul style="list-style-type: none"> A fixed-wind UAV (high flight speed) is equipped with a near infrared laser lamp Camera array for detecting the near infrared laser 	<ul style="list-style-type: none"> Two NIR cameras (with a complex and precise setup process for accuracy) arranged on the ground are required The cameras are placed on each side of the runway requiring a wide baseline
			Based on pan-tilt unit (PTU)	Stereo-vision	Visible light camera [14, 16]	<ul style="list-style-type: none"> Two separate PTUs with two cameras (to avoid limitations on the cameras baseline) are placed on both sides of the runway AdaBoost is used in [16] for tracking and 3D determination by triangulation Three-dimensional localization is applied in [14] for target detection

case in the quadruplet model, inspired by Siamese networks [68], where four network branches are defined, consisting of the exemplar, instances, positive, and negative branches according to their inputs. Comparing this approach with our method, our proposed regions should be considered as exemplars whose probabilities will determine positive or negative branches, along with random instances. Additionally, such bounding boxes can also be considered as crop candidates in the attention box prediction network and supplied to the aesthetics assessment network; both were proposed in [69]. In the aesthetics network, probabilities can be used to assign binary labels considering probabilities (ranging between 0 and 1) less or greater than 0.5, respectively.

Proposed Cognitive Method

General

The premises described above have been addressed by mapping the human reasoning scheme to generate the computational cognitive process, based on perception, recognition and final decision-making:

1. Landing platform containing markers in the form of specific geometric and colored shapes (Fig. 2). Humans have high color perceptual ability.
2. The combination of color identification approaches and shape descriptors based on Hu moments [70], which are invariant to translations, rotations, and scale changes. Humans identify figures under different appearances.

3. A recognition process that uses a combination of color, providing isolated figures, and shapes with geometrical relations based on a *Decision-Making* (final phase of human reasoning) approach together with the Euclidean Distance Smart Geometric Analysis (EDSGA).

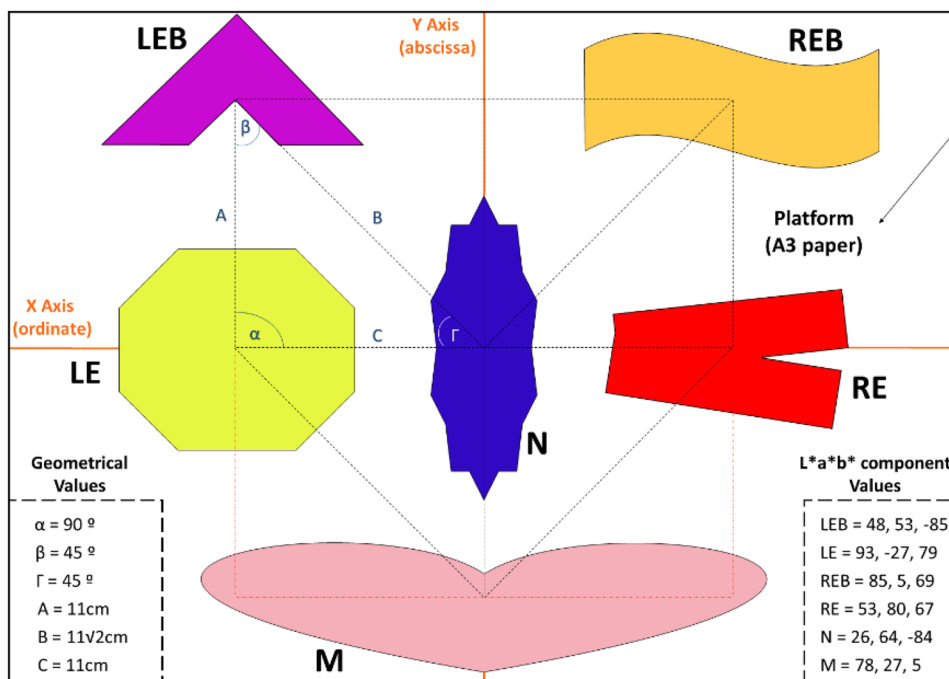
Landing Platform Design and Characteristics

The new landing platform shown in Fig. 2 contains six uniquely shaped and colored figures labeled LEB, LE, REB, RE, N and M. These markers are printed on white A3 paper. Each marker is filled with a different and unique color to facilitate the region detection procedure, and to address the problems that occur when regions are missing.

The center of mass (centroid) of the blue region (N) is placed in the middle of the platform. This is the point where the eight identical right-angled triangles (marked with black dotted lines) converge, arising as a result of dividing the squares into equal parts. This geometric pattern was designed to help determine the orientation of the image and to simplify the geometric relationships required to recognize the landing platform based on a computed probability.

When using RGB-based devices, color is an excellent perception attribute in computer vision [71]. The use of colors on the landing platform is essential because it provides two important advantages. Firstly, using different colors increases the probability of image processing techniques extracting several, or all, regions (more regions can be seen); if some regions (represented by colors) are not clearly detected, then others should be properly seen and recognized by the system, even if the whole platform is only partially

Fig. 2 Landing platform with labeling and geometrical markers



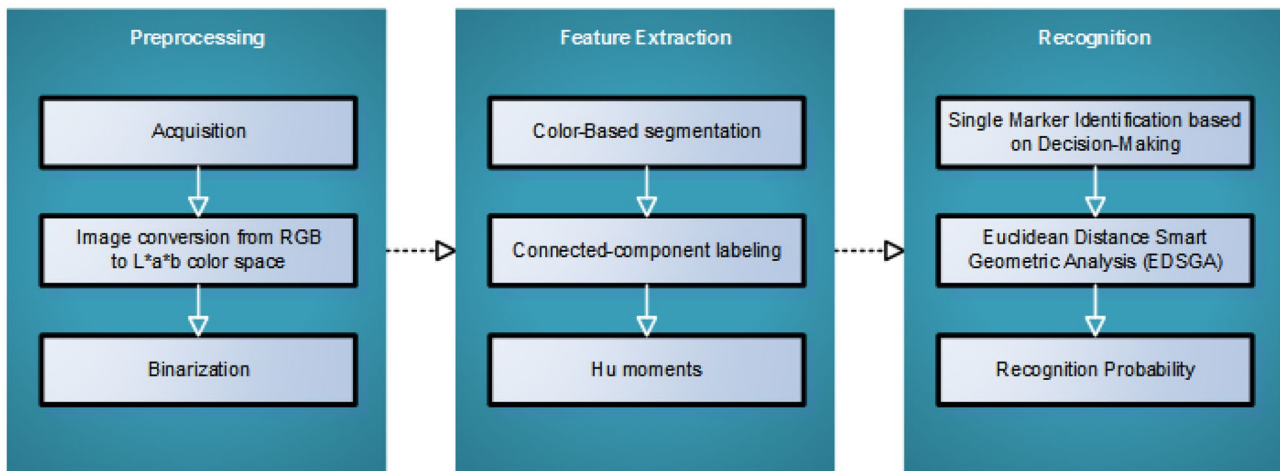


Fig. 3 Design and architecture of the proposed image recognition algorithm

visible in the image. Secondly, the use of a different color for each marker allows the system to perform an additional classification (described below) which is invariant to plane distortion due to orientation, rotation, or scaling, in addition to being invariant due to image perspective projection.

Proposed Solution

The proposed algorithm follows the architecture shown in Fig. 3 and is based on human reasoning, as expressed above, according to the standard imaging segmentation procedures of preprocessing, feature extraction, and recognition. Finally, the recognition probability is calculated by analyzing the geometric relations between the set of markers which have been identified. The processes shown in Fig. 3 that represent novel implementations are:

1. Computation of Hu moments. These are invariant to rotation, translation, and scaling. They are robust

descriptors to deal with marker distortions that may result from image perspective projection and/or if the UAV deviates from the vertical inclination, as well as deviations from the zenithal position on the platform, all of which are typical during landing.

2. A new recognition process which mixes both the $L^*a^*b^*$ color discrimination with the Hu moments. This is based on a *Decision-Making* technique [72] that uses distance measurements together with the EDSGA algorithm. This avoids complex difficulties like those shown in Fig. 4.

Preprocessing

During the landing phase, the platform is imaged with the digital camera of the mobile phone (Samsung Galaxy S6) located onboard the UAV. The standard, high-resolution image is scaled down to 640×480 , 1024×768 and 1707×1280 pixels in each of the three RGB spectral channels. The lowest

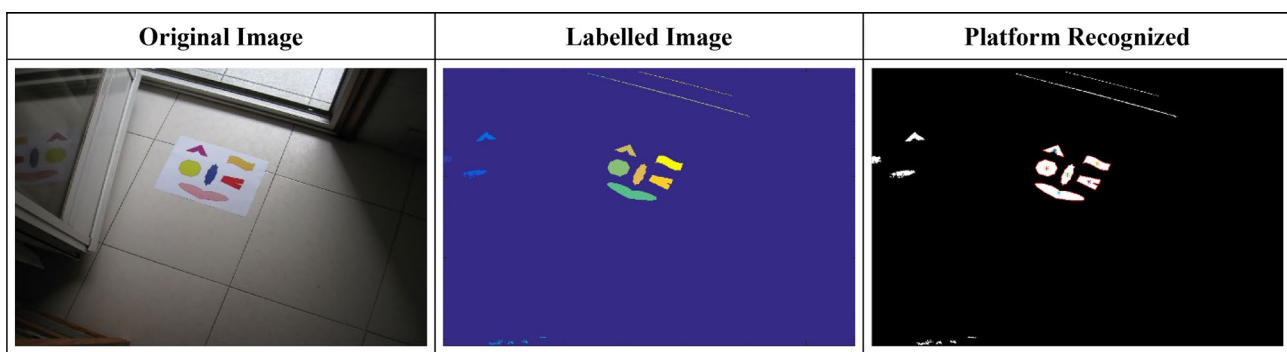


Fig. 4 Illustrative example to solve adverse problems, even with specular reflections. Note that the markers' colors depicted in the labelled image are for display purposes only

Table 2 Numerical values for each marker displayed in Fig. 2 and used during the color image segmentation process in order to obtain a binary image from the CIELAB color image

Marker (r)	a^* (Tl_a)	b^* (Tl_b)	Threshold (Tl)
LEB	45.0	-35.0	23.0
LE	-16.0	33.5	23.6
REB	-3.0	30.68	23.6
RE	48.0	6.0	28.0
N	15.0	-42.0	26.0
M	22.6	-6	19.8

resolution is used first, and if identification fails — as evidenced by a recognition score $< 50\%$ — the entire process is repeated using the next-highest level of resolution, and so on. After image reduction, the RGB image is transformed into the $L^*a^*b^*$ color space. This color space contains luminosity as a separate component, which effectively minimizes the impact of reflection difficulties caused by intense light sources such as the sun, which is the most problematic.

The a^* and b^* components are sufficiently separated in their respective color spaces to avoid overlaps during the segmentation process. Each marker is defined with a unique color to compensate for any parts of the landing platform image that might be missing. CIELAB is categorized as a uniform color space where changes in the color coordinates correspond to identical or similar recognizable changes in visible color tones and color saturation. This was designed to facilitate color measurement in accordance with the Munsell color order system [71]. Moreover, CIELAB is device-independent and has been sufficiently tested to deal with similar spectral component values during the segmentation processes. This occurs in adverse lighting conditions, such as those found in outdoor environments, where high levels of illumination lead to the saturation that causes low contrast in image colors. Low contrast also appears at low lighting levels. According to these considerations, several color spaces were analyzed during experimentation. The experiments showed that the best results were for the CIE 1976 $L^*a^*b^*$, which was, therefore, chosen as the best color space for the segmentation.

Table 3 Averaged values for the seven Hu moments for each marker

Marker (r)	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
LEB	-1.2837	-3.7150	4.3979	-6.8821	-12.8082	-8.9511	-13.4887
LE	-1.8081	-7.1000	-14.7351	-18.2484	-35.1690	-22.4804	-35.5544
REB	-1.2838	-3.0417	-12.3622	-14.6235	-28.6950	-16.8482	-28.3050
RE	-1.5137	-4.0205	-7.0933	-10.0107	-18.5817	-12.0440	-20.6703
N	-1.4001	-3.3519	-15.0324	-16.3327	-32.0414	-18.1191	-33.5053
M	-0.8704	-2.1921	-5.3414	-7.5967	-14.8185	-8.7103	-15.2624

Feature Extraction

The aim of this stage is to segment the image, identifying a group of regions and their measurements. These regions will be used in the recognition stage to match a single region with each marker, where the group of regions will be used as *candidate markers*. Feature extraction can be broken down into several steps:

1. *Color-based segmentation* (described in Algorithm 1) based on the selected CIELAB color space [73] with foreground segmentation [74]. The underlying idea in this algorithm is the consideration of a measure of similarity in the color space at the pixel level. This is also a common concept in color-based image segmentation, as expressed in [75], where color distances are used to measure similarities between the pixel, which is being labeled, adjacent pixels, and seeds that guide the segmentation process.

Algorithm 1 returns a binary image that is used in the connected-component labeling process. Threshold (Tl) and estimated values (Tl_a , Tl_b) used for each figure appear in Table 2. These averaged values are obtained by applying the supervised naïve Bayes training approach with 500 images obtained under different attitudes (distances and inclination angles) of the UAV with respect to the platform, and under different and adverse lighting conditions (sunny, cloudy and alternating clouds with sun). Several rotations and scaling with respect to the UAV were also considered.

For this estimation process, each color marker was manually identified by human inspection of random samples for each figure, computing the averaged values and covariance matrices. The threshold, Tl , was manually obtained by carrying out a heuristic and supervised method through a fine-tuning process during the training stage. To support extreme changes in color shades, the best limit according to Algorithm 1 was found. Correct recognition can now be expected even under extreme and disparate conditions.

1. *Image labeling*: this is based on the 8-connected-components approach [76, 77]. For each binary region, we compute four properties — area, centroid, orientation,

and bounding box — where some segmented regions are potential markers.

2. *Hu moments*: for each binary region, the seven Hu-invariant moments are obtained [78]. As before, the naïve Bayes estimation approach is used with the same 500 images to compute the averaged values and covariance matrices for each moment and candidate region. These are displayed in Table 3.

Recognition

Based on the properties of the binary regions, the goal now is to identify each unique marker ($r = \text{LEB, LE, REB, RE, N, and M}$). This is part of the discrimination procedure derived from human reasoning, and which belongs to the cognitive process.

Single Marker Identification

Using a combination of CIELAB (color) and Hu (shape) moment properties, the marker identification is performed by comparing the segmented candidate markers (binary regions) with the values provided in Tables 2 and 3. In this way, each candidate marker is identified and assigned to the label with the highest similarity, as based on a minimum distance criterion (Euclidean) [79].

$$\begin{aligned}
 Dc_a(r) &= \left| \bar{c}_a - TI_a(r) \right| \\
 Dc_b(r) &= \left| \bar{c}_b - TI_b(r) \right| \\
 Dc(r) &= \sqrt{(Dc_a(r))^2 + (Dc_b(r))^2}
 \end{aligned} \tag{1}$$

Candidate markers must meet the following constraints:

$$\begin{aligned}
 Dc_a(r) &< Tl(r) \\
 Dc_b(r) &< Tl(r) \\
 Dc(r) &< Tl(r)
 \end{aligned} \tag{2}$$

Compute the shape distance D_{Hu} with the reference Hu moments for each region, (r) (see Table 3), as follows:

$$DHu(r) = \sqrt{\sum_{i=1}^7 (Hu - \phi_i(r))^2} \tag{3}$$

Compute the total distance of the candidate binary regions with respect to each marker r :

$$D(r) = \left(\frac{\left(\frac{Dc_a(r)}{\max(Dc_a(r))} \right) + \left(\frac{Dc_b(r)}{\max(Dc_b(r))} \right)}{2} \right) + \left(\frac{DHu(r)}{\max(DHu(r))} \right) \tag{4}$$

Algorithm 1 Color image segmentation. Foreground extraction.

```

Input: IRGB image.
Output: Ibinary image (1: foreground, 0: background).
Convert RGB image into L*a*b* color space to obtain L* (IL), a* (Ia) and b* (Ib) channels (IL is not used)
while select TIa, TIb and TI for each region (LEB, LE, REB, RE, N, M) according to values in Table 2, and for each pixel located at (x,y) do:
    da ← |Ia(x,y) - TIa|; db ← |Ib(x,y) - TIb|; d ← √(da2 + db2)
    if (d < TI) and (da < TI) and (db < TI) or (Ibinary(x, y) = 1) then
        Ibinary(x, y) ← 1; else Ibinary(x, y) ← 0
    end if
end while
return Ibinary
A morphological opening [80] operation is performed in Ibinary to remove spurious regions and isolated pixels
    
```

Color and shape distortions often appear due to adverse outdoor environmental conditions, leading to failures during the marker identification process. To minimize these frequent misclassifications, color and shape properties are combined considering target colors (Table 2) and shapes (Table 3). We compute similarity/dissimilarity values using the following *Decision-Making* process (important part of the human reasoning):

For each binary region with an area ranging from 50 to 80,000 pixels, compute average spectral values \bar{c}_a and \bar{c}_b for channels **a*** and **b***, respectively, and the seven Hu moments (Hu).

Compute the spectral distance D_C , with $TI_a(r)$ and $TI_b(r)$ reference values for each marker provided in Table 2, as follows:

The minimum $D(r)$ value with respect to r allows the classification of the unknown candidate binary region as one of the possible labeled markers, i.e., $r = \{\text{LEB, LE, REB, RE, N, M}\}$. From now on, the markers have been labeled.

It is anticipated that each marker will have none, one, or several candidate regions associated with it. None indicates that a given candidate has not been detected. When several regions exist, the region with the lowest $D(r)$ value will finally be assigned to the marker r because a smaller distance means greater similarity. For a better understanding of the above, consider the following pedagogical example: LEB: {**1**, 2, 3}; LE: {**10**, 11}; REB: {**22**, 23, 24, 25}; RE: {**37**, 38, 39}; N: {}; M: {**45**}.

The number in **bold** indicates the candidate region with a minimum $D(r)$ value with respect to r for each marker.

So, LEB has candidate regions {1, 2, 3}, LE has candidate regions {10, 11}, N could not be detected, and so on. Hence, for this example, the base group (G_B) of regions obtained is: $G_B = \{\text{LEB: 1, LE: 10, REB: 22, RE: 37, M: 45}\}$. The N marker does not appear because it is not detected in this example and the total number of elements for the grouping (tM) is 5.

Euclidean Distance Smart Geometric Analysis (EDSGA)

Once the markers have been identified, groups of candidate markers are built to identify group coherences, which are compatible with the full set of figures drawn on the platform. This is carried out by computing compatibilities between the centroids of the grouped markers in the image and comparing to the grouping of markers on the platform. Geometric distances between centroids are obtained for comparison. This approach assumes that the platform is made up of a group of markers.

This process of determining group associations was designed as follows:

Groups of markers: build groups using the candidate regions for each, applying the following rules:

The length of each group will always be the same and must coincide with the total markers, tM . There is no minimum limit for the tM value as far as groupings are concerned, but the maximum value is 6.

At least $(tM/2) + 1$ members belonging to each group must be equal to G_B . Hence, each group must contain at least $(tM/2) + 1$ regions with the minimum $D(r)$ value regarding its respective marker. Hence, we build as many groups as permutations that can be obtained with the regions labeled as specific markers, i.e., $(tM/2) - 1$ that do not match the elements in G_B . Considering the pedagogical example above, each group must contain at least three members belonging to G_B and permutations for the remainder. Some valid groups could be:

$$G_1 = \{\text{LEB: 1, LE: 10, REB: 22, RE: 38, M: 45}\}$$

$$G_2 = \{\text{LEB: 1, LE: 10, REB: 22, RE: 39, M: 45}\}$$

$$G_3 = \{\text{LEB: 2, LE: 11, REB: 22, RE: 37, M: 45}\}$$

$$G_4 = \{\text{LEB: 2, LE: 10, REB: 23, RE: 37, M: 45}\}$$

Normal font character indicate candidate regions matching G_B , while **bold** characters refer to the candidate regions that are permuted as not belonging to G_B but are labeled as specific markers.

A candidate region must belong to only one marker. The group must not contain repeated regions.

We compute and sum the Euclidean distances (total distance) between all pairs of centroids of all regions belonging to each group.

The group with the minimum total distance is finally chosen as the most likely candidate to be the platform.

Once the most likely group (G_C) is chosen, a reassignment is still feasible up to a maximum of $(tM/2) - 1$ regions for this group, i.e., all regions with discrepancies from G_B with respect to G_C are replaced by those of the latter.

For example, if G_C is $G_2 = \{\text{LEB: 1, LE: 10, REB: 22, RE: 39, M: 45}\}$, considering that $G_B = \{\text{LEB: 1, LE: 10, REB: 22, RE: 37, M: 45}\}$ the region of marker RE is reassigned, changing from 37 to 39.

This EDSGA approach greatly improves the accuracy of the recognition process and, therefore, the probability of recognition. It works well even in adverse situations. Indeed, EDSGA can correctly recognize groups of markers even if there are regions with similar colors and shapes.

This is demonstrated in Fig. 4 where both the standard and reflected platform images are close to each other because of their proximity to a glass door. The algorithm identifies the original image and rejects the reflected one. In this example, the LEB and M regions appear twice due to the reflection in the glass. Without EDSGA, the recognition would have detected 5 figures correctly, but would have failed in the LER marker, having detected the reflection to the left. However, after applying this smart analysis, all elements of the group are compared with the real figures. Hence, the region related to the LEB marker has been reassigned to the region number set in the EDSGA group. Without this step, the recognition would have been performed with a score of 80% instead of 100%.

This illustrative example demonstrates the utility of EDSGA. However, the scope of the application is most evident when hundreds of candidate regions are generated (Fig. 6A); some regions find matches with markers but are partially blind or significantly distorted. In this case, the EDSGA approach ensures that those regions with high similarity scores that can no longer be discerned are correctly assigned.

Recognition Score Computation

The objective of this stage is to provide a score in the range [0,1] that describes the recognition probability of the landing platform as a group. Therefore, we must determine whether the total number of markers, and their geometric characteristics/relationships are similar to the expected values, considering G_C after the reassignment.

Metric Descriptors

In the recognition stage, a series of analyses based on metric descriptors are performed to study the geometric relationships between the detected objects (selected markers) shown in Fig. 2. The following descriptors are used: area, distances between centroids, and the angles between the straight lines connecting the centroids with respect to the base of the image, i.e., the bottom horizontal line of the image. The

areas and distances are strongly affected by the image resolution, caused by the varying distance between the camera and the platform.

Therefore, instead of performing the analysis using only absolute values, relative values are established between the figures using a technique analogous to the comparative partial analysis (CPA) introduced in [1]. However, in this instance, the probability function using these similar measurements must also consider two independent events based on the independence of the probability theory.

The method considers all possible combinations between regions to ensure an adequate number of similar measurements. In this way, sufficient measurements will be obtained even if there are undetected markers, as was the case with marker N in the example above. Therefore, the proposed cognitive method is more robust and accurate. Each combination generates a measure of similarity which has two inequalities. All inequalities are shown in Tables 4, 5, and 6 in the “Appendix”.

In these inequalities, all numeric values were calculated using the zenithal position of the camera relative to the platform. Additional flexibility is required to withstand distortions in the markers due to image perspective projection. This is achieved by considering three thresholds, ArT , DT , AgT , as estimated by hundreds of tests and a trial-and-error procedure under different attitudes and distances between the vision system and the platform. All the values for these thresholds are reported in Tables 4, 5, and 6, respectively.

Table 4 shows a comparison of the area between regions, Table 5 compares the distances between centroids of the regions, and Table 6 compares the difference in the angles generated by the intersection of the straight lines connecting the centers of the regions with respect to the horizontal x -axis of the image. Tables 4, 5, and 6 contain all possible relationships involving areas, distances, and angles for the full set of markers with the flexibility expressed above, based on the three thresholds. Thus, the maximum number of possibilities (relations) when all markers are detected is 15 for the area (Ar_p , $i = 1, \dots, 15$), 105 for the distances between the centroids (Dst_j , $j = 1, \dots, 105$), and 105 for the differences between the angles (Ang_k , $k = 1, \dots, 105$).

Recognition Probability

The final recognition probability function returns a recognition score in the range $[0,1]$ by combining the probabilities of two events (A and B) defined below. In this regard, the terms *recognition score* and *recognition probability* refer to the same concept (herein used interchangeably), which can also be expressed in terms of a percentage. Firstly, the ratio between the total number of markers that could be identified

and those actually identified is considered. This event is represented by A and its probability is computed as:

$$P(A) = \frac{tMR}{tM} \quad (5)$$

where tMR is the total number of markers finally recognized and tM is the total number of possible markers that exist on the landing platform, i.e., six in the proposed design. On the other hand, it also takes into account the probability based on the geometric relationships between the regions belonging to the candidate group, G_C . This probability is computed by applying the geometric relations defined in Tables 4, 5, and 6 (Appendix) and requires the following considerations:

(a) The number of possible relations for each topic (area, distance, angles) is defined by the number of markers detected. Following the example above with respect to the group G_B where the marker N is missing, it can be seen how this marker is involved in five relations in Table 4, so the number of possible relationships in this case is 10. This is applicable to the remaining relationships, i.e., the maximum number of possible relationships is 45, and 45 for the example in Tables 4 and 5, respectively. On this basis, the number of possible relationships for each group is identified as aT_p (area), dT_p (distance), and agT_p (angles).

(b) The number of relationships that are met for each topic among all the possibilities is defined by aT_s (area), dT_s (distance), and agT_s (angles).

The probability of event B is therefore defined as follows:

$$\begin{aligned} P(B_a) &= \frac{aT_s}{aT_p}; P(B_d) = \frac{dT_s}{dT_p}; P(B_{ag}) \\ &= \frac{agT_s}{agT_p} P(B) = \frac{P(B_a) + P(B_d) + P(B_{ag})}{3} \end{aligned} \quad (6)$$

The overall recognition probability from events A and B is modeled as the intersection under the assumption that these are two independent events, i.e., on the basis of probability theory [81], the final probability of detection of the landing platform is computed as follows,

$$P_d = P(A \cap B) = P(A) \cdot P(B) \quad (7)$$

The assumption of independence is sufficient for this approach, although the events could be considered partially dependent; if a marker is missing, this affects event A and also event B since a different number of relations (equations) will be used. Inspired by fuzzy set theory [82] we have modeled $P(A)$ and $P(B)$ as membership degrees to combine them as t -norms and t -conorms for different combinations, including drastic, Einstein, or Hamacher product and sum without apparent improvement with respect to the results provided in (7).

Results

The test strategy is based on 800 images and follows the same test strategy designed in [1]. Thus, 80 of the 800 images were captured using the same settings for angles, distances, and lighting. The other 720 tests were similar. Both series of tests were carried out in the same environment and under the same considerations.

Because failures and limitations were previously observed with 80 images in [1], we started by reproducing similarly adverse test conditions. Further 720 new images, tested under very different conditions, were also used for the tests. In both cases, we set out to analyze the performance in terms of distances, inclination angles, and lighting conditions.

We also tested the following platforms and methods: circles/ellipses [7]; square markers [53, 54]; T/H/X-shaped [43, 48] under identical considerations to those in [1], i.e., without the richer color information, verifying the extent to which this new approach outperforms previous works [1]. A direct comparison is not possible because no additional information is provided about the three main parameters used to evaluate this approach: distance, inclination angle, and lighting conditions.

These useful parameters verify the novelty of our proposed cognitive method, demonstrating effective and robust recognition of images obtained under extreme conditions. However, the superior performance of the proposed cognitive approach in relation to [1] also reveals better performance in relation to the abovementioned methods.

For each image, the same visual supervision process is carried out as defined previously [1]. An outcome is considered successful when the user, by observation, determines that there are at least three red regions identified in the original image and that they match the correct markers on the platform. In Figs. 4, 6, 8, and 10 (see “Platform Recognized” columns), the boundary of each detected region is colored in red for debugging.

Image Acquisition Environment

All the images were acquired in a real test environment using a mobile phone (Samsung Galaxy S6) on board a quadrotor with four 1000 rpm/V brushless engines each driving a 10-inch propeller and powered by its own battery. In addition to acquiring the images, the telephone acts as a control unit that manages sensors, engines, and other quadrotor components. The experiments carried out to test the proposed cognitive approach were performed using the following:

1. Landing platform. The markers shown in Fig. 2 are printed on a white background on A3 paper.
2. All images were acquired using the integrated mobile phone camera (16MP, F/1.9) in automatic mode without use of zoom or flash.
3. 800 pictures were acquired in different environments and conditions: distances, inclination, and lighting (outdoor: sunny, cloudy, sun and shadow; indoor: artificial light).
4. Each image was captured using the following settings: 3 bytes (24 bits) for color depth; a resolution of 4 MP with dimensions of 2560×1944 pixels; 1 byte (8 bits) per RGB channel; and JPEG using standard compression.
5. The test UAV flights were conducted in the courtyard of a building within a residential area.
6. The implementation of this test stage was carried out using the image processing toolbox provided with MATLAB Drive [83], which is used by the MATLAB Mobile App, i.e., with the code running \in the cloud (Drive). The Galaxy mobile platform was connected to the cloud via WiFi or 3G wireless networks and was running under Android 5.0.2, Octa-core (4×1.5 GHz Cortex-A53) CPU, and Mali-T760MP8 GPU, with internal memories of 128 GB and 3 GB RAM.

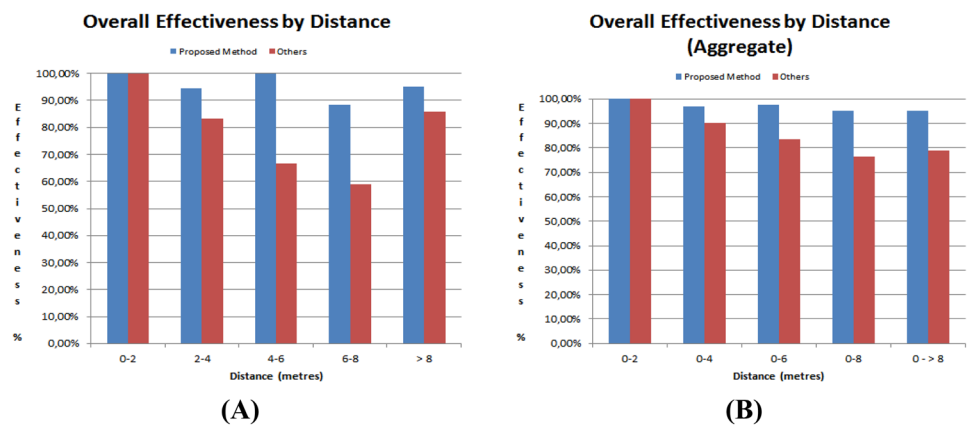
The aim of the tests is to determine the robustness and accuracy of the proposed cognitive method under different conditions in terms of the distance between the quadrotor and the platform, the angles of inclination under different perspectives, and the illumination.

From the point of view of computational cost, three internal resolutions were considered for each image, which varied between 640×480 , 1024×768 and 1707×1280 pixels. In this regard, if any lower resolution is successful at recognition, the others are not tested. If all downscaled images fail, a final attempt using the native image resolution is performed.

Distance Test

Eight hundred images were used when distances from the UAV to the platform were in the range of 0.6 to 12 m. The results obtained are summarized in Fig. 5, where the percentage of effectiveness (i.e., probabilities of detection) against distance in meters are graphically displayed and compared between the proposed cognitive method and the method described in [1]. In Fig. 5A, the distance intervals are unique (i.e., 0–2 m, 2–4 m, 4–6 m, 6–8 m and > 8 m). In Fig. 5B the distance intervals are aggregated (0 to 2 m, 0 to 4 m, 0 to 6 m, 0 to 8 m and > 8 m). The proposed cognitive approach achieves, on average, an acceptable and improved performance (95%) compared to the previous strategy [1], labeled as “Others” in the graphs. Generally, we observe that better performance is achieved at shorter distances, as expected.

Fig. 5 Effectiveness against distance (m) for the proposed cognitive method and the method in [1]: **(A)** exclusive intervals, **(B)** aggregate (cumulative) intervals



Although distance has a significant impact on effectiveness, the proposed cognitive method successfully recognizes an acceptable number of images up to 12 m, the maximum distance we considered. The proposed cognitive approach is more effective than the one described in [1]. Additionally, it can be concluded that only the lower image resolution is needed at shorter distances, as expected.

Illustrative Example Related to Distance

In Fig. 6, we compare two images that demonstrate the versatility of the proposed cognitive method in terms of distance. Both images were obtained under similar lighting conditions, but with distances of 8.94 m and 0.67 m in A and B, respectively. The platform was successfully

recognized with scores of 91.11% and 100%, respectively. In contrast, the method in [1] fails at distances > 8 m.

Inclination Angle Test

Eight images were captured and evaluated with a combination of inclination angles of 1° – 68° , and distances, as described previously. This angle is defined, in degrees, between the vertical axis in the image and the imaginary straight line connecting the center of the camera lens to the centroid of the N marker. As before, the results obtained are displayed graphically in Fig. 7 with the percentage of effectiveness against angle of inclination for the proposed cognitive method, and that described in [1]. In Fig. 7A, intervals of angles are mutually exclusive (i.e., 0 – 10° , 10 – 20° ,

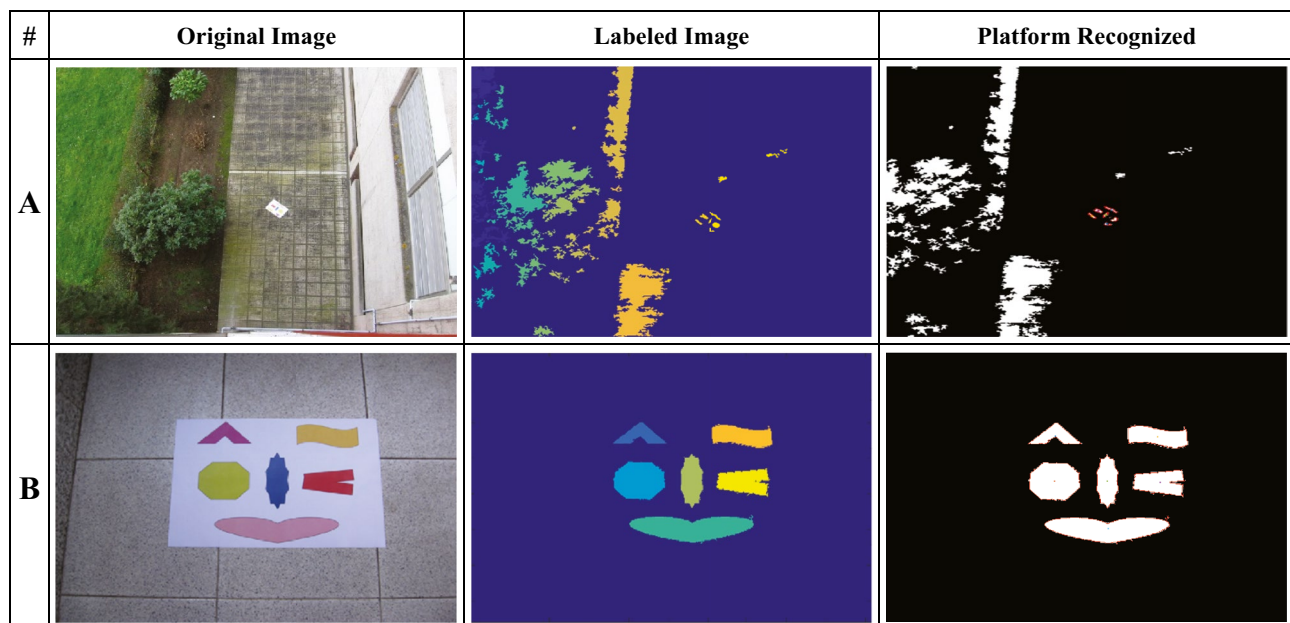
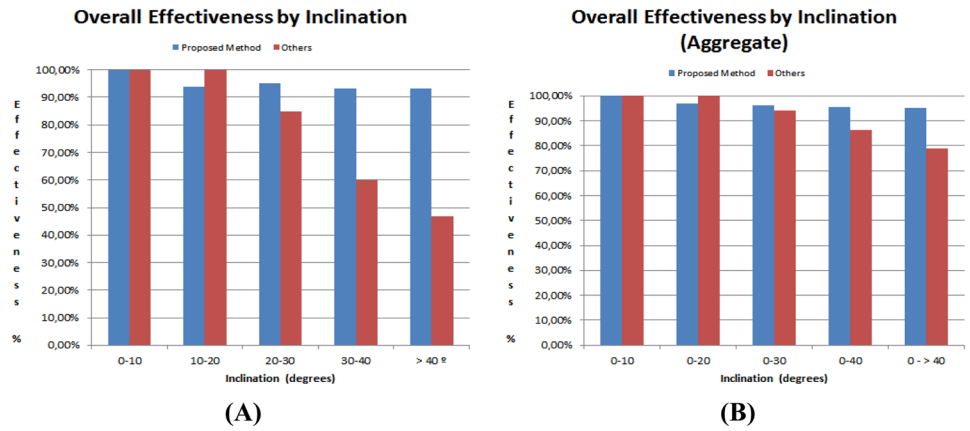


Fig. 6 Illustrative example at two very different distances under identical lighting conditions: **(A)** at 8.9 m, **(B)** at 0.67 m

Fig. 7 Effectiveness against inclination angles (degrees) for the proposed cognitive method and the method in [1]: (A) exclusive intervals, (B) aggregate (cumulative) intervals



20–30°, 30–40°, and > 40°). However, in Fig. 7B the intervals of angles are aggregated (0° to 10°, 0° to 20°, 0° to 30°, 0° to 40°, and > 40°). Again, acceptable performance is achieved with the proposed cognitive approach (95%), which outperforms the strategy in [1]. In the main, better performance is achieved at lower inclination angles.

Illustrative Examples for the Inclination Angle

Figure 8 contains two different illustrative images to show the versatility of the proposed cognitive method in terms of high inclination angle (A: 59.78° and B: 61.17°), where the effect of the image perspective projection is acutely observed. In both cases, the proposed cognitive approach

successfully recognizes the platform. This outperforms the approach in [1], where the experiments failed at these angles. The results of this example are generally representative of what was commonly observed.

Lighting Condition Test

Eight hundred images were acquired on different days and under different daylight conditions in outdoor environments (sunny, cloudy, and shade), and indoors with artificial lighting. These images were acquired under the criteria of the quadrotor’s operator during different flight experiments. Figure 9 displays the averaged values over the total number of images used for this test graphically. It is clear that the

#	Original Image	Labeled Image*	Platform Recognized*
A			
B			

Fig. 8 Meaningful sample of high inclination angles due to the distortion in the plane caused by the perspective. In both cases, we achieved a successful recognition with a score of (A) 96.19% and

(B) 84.44%, respectively. *Both labeled and platform recognized images were zoomed for better illustration

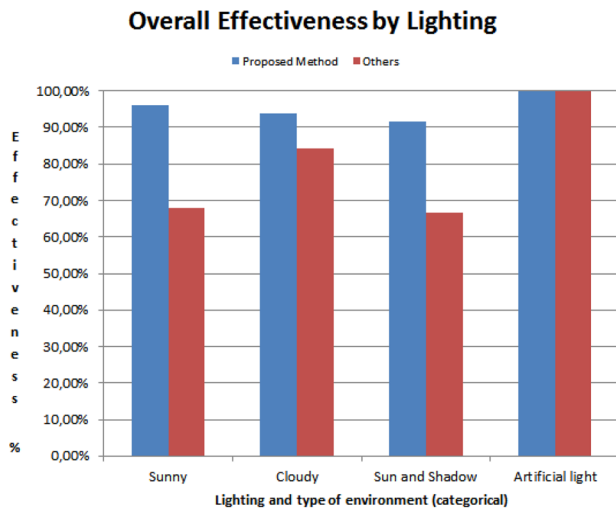


Fig. 9 Effectiveness against different lighting conditions

proposed cognitive approach outperforms the method in [1] in outdoor environments, with similar results to those in [1] for indoor environments. This is because the brightness indoors is constant and allows for sufficient contrast between the black figure on the white background, thus compensating for the additional contribution of color markers in outdoor environments.

Illustrative Examples Under Different Lighting Conditions

Two illustrative images are shown in Fig. 10 to demonstrate the proposed cognitive method's versatility in terms of lighting conditions. Figure 10A is acquired in an artificially lit

indoor environment, where the reflections caused by this type of light and the high inclination angle produce an excess of luminosity that hinders the recognition of the landing platform. Figure 10B, acquired in an outdoor environment, is also affected by the reflection caused by the intensity of the sunlight and the high inclination angle.

In Fig. 10A, all the markers of the landing platform are correctly identified, and the platform was detected with a probability of 83.49%. Best performance was not obtained because marker M is partially blind. A lower recognition score in Fig. 10B results because markers RE and M are totally blind and could not be detected. For this case, the maximum recognition probability value was 4/6 ($\approx 66\%$), i.e., despite these missing markers, the score obtained for recognition is the maximum possible, and the landing was successfully achieved. This result, which was generally observed in different experiments, demonstrates the robustness of the proposed cognitive method, even when markers are missing.

Overall Assessment

It is important to check the robustness and reliability of the recognition feature of the system. The score returned by the probability function should be consistent with the input image. This is achieved by checking that the region identified and associated with each marker correctly matches the expected marker.

To carry out this process, and, therefore, to decide the likelihood of success, we have performed a manual, visual "supervised" human inspection. This supervised process requires a visual inspection of each result to determine

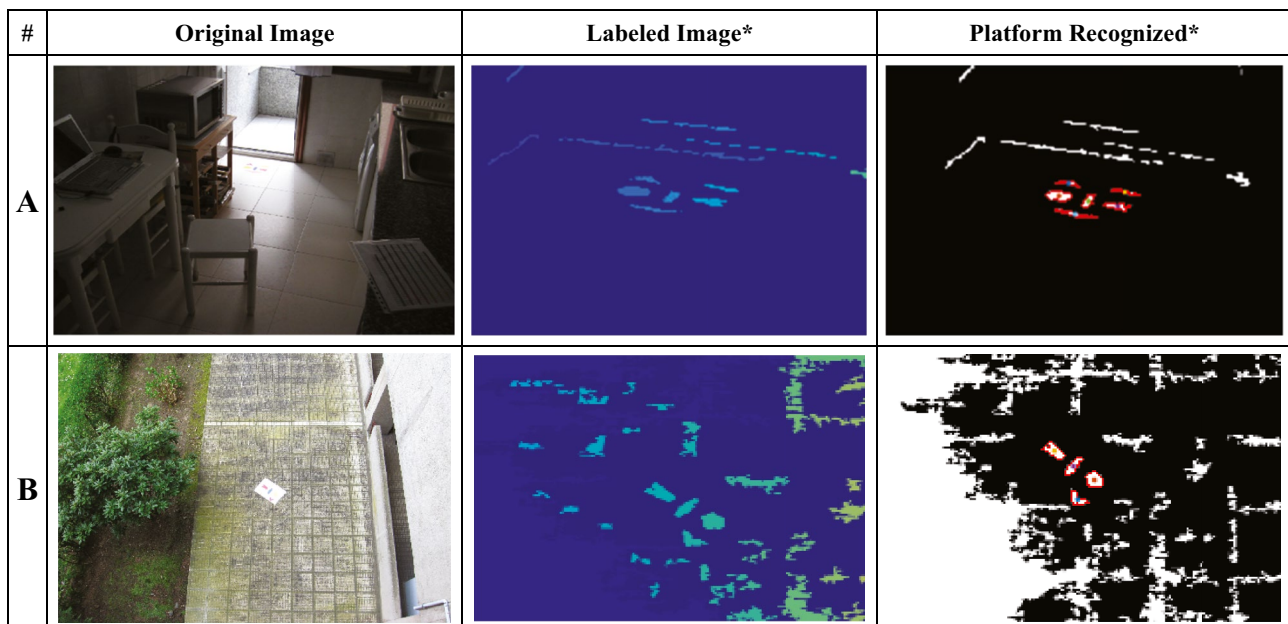


Fig. 10 Illustrative examples under different lighting conditions in indoor (A) and outdoor (B) environments. *Both labeled and platform recognized images were zoomed for better illustration

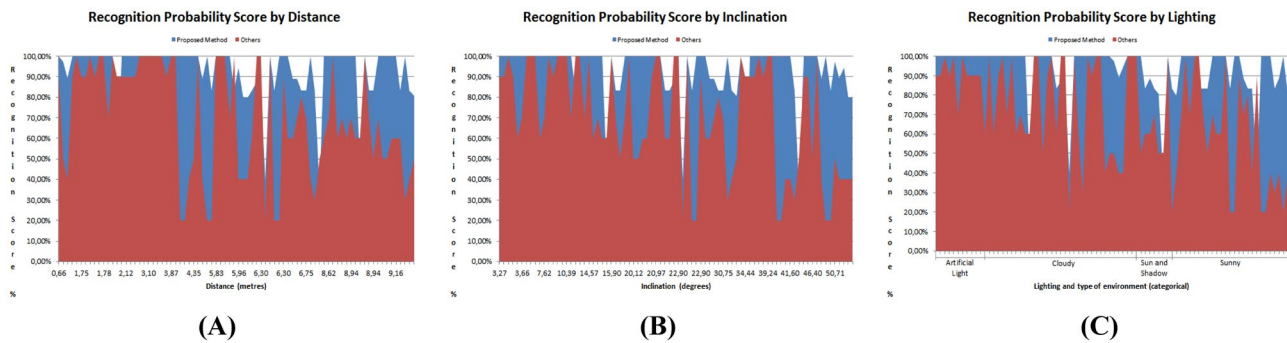


Fig. 11 Score returned by the recognition probability function. (A) Score by distance, (B) Score by inclination angle, (C) Score by lighting conditions

whether the recognition system has detected at least three regions (edges marked in red), and whether these regions were matched with the correct markers.

The conventional state-of-the-art for trackers [84, 85] is that a minimum of two frames indicates a successful outcome. It compares the positive example that is manually provided in the first frame with the following frames. However, our system works in a more efficient fashion since it only needs a single frame for assessment. Several criteria can be used to make a decision about what is really a successful outcome; these criteria have already been applied in [1] and allow an objective evaluation of system behavior.

In addition to the performance reported above, we also provide details about the recognition scores obtained according to Eq. (7) for different distances, angles and lighting, as before.

As displayed in Fig. 11, the minimum score obtained is 35%, while the average was 92.94%. A visual inspection of the result was carried out to determine whether the recognition system detected at least three markers. As a result of this process, it was observed that images with a score lower than 50% resulted from the detection of two markers or less, while those with higher values always correctly detected at least three markers.

From the above, it can be inferred that the result returned by the probability function is coherent with the visual inspection and, therefore, the probability function is reliable because it is consistent with observations made during the flight tests. In addition, images with more extreme lighting issues, blind areas, and large deformations due to perspective have lower scores, which is again consistent with the images with inferior results in [1].

Conclusions and Future Trends

The proposed cognitive approach can recognize a landing platform in a robust manner in conditions that are representative of real-world scenarios. The average recognition

time per image is 0.5 s, 1 s faster than previous work [1]. The cognitive method operates across a broad variety of conditions: different distances, inclination angles, any lighting conditions (sunny, cloudy, sun, and shadow, etc.), complex environments (indoor, outdoor), and with blind regions that result from intense light sources and reflections. Robustness and accuracy are the main characteristics that define the cognitive method.

Cognitive computation, which is the applied paradigm involving visual perception and decision-making, consistently outperforms the previous approach [1], and by extension those that were previously evaluated [1]. With the new design, 760 of 800 new images were successfully recognized (95%), while before [1], 63 images were correctly identified from a total of 80 (78.75%). Although the sensitivity to the inclination angle due to perspective may still be a limiting factor for the practical exploitation of this method, the improvements are still so great that we strongly believe that the proposed algorithm, along with this new platform design, will be highly effective for both rotary- and fixed-wing UAVs under somewhat restricted circumstances of approach angle.

As in [1], the closer the drone is to the landing platform the lower the image resolution that is needed. For short distances (0–4 m), the system works better when using low resolution (1024×768). As the distance increases, the best results are obtained using high resolution (2560×1944). If high resolution is used at short distances, the number of regions obtained through the segmentation process increases and reduces the likelihood of success. Using low resolutions for long distances may cause the platform to be segmented into a single region, which prevents recognition. In addition, the more regions there are, the more important EDSGA is.

Thresholds are considered optimal because the training set used to obtain the Tl values was carried out from many different perspectives, and various weather conditions and environments. In addition, the images used in the results section to evaluate the performance of the method were obtained in environments that were totally different from

those of the training set. From these premises and the support provided by EDSGA for decision-making, along with the good results obtained, we can infer that values shown in Tables 2 and 3 are optimal and effective for complex environments and diverse weather conditions. Hence, no additional calibration is needed (unsupervised method).

It has also been proven that recognition can be carried out using a single frame with difficult images that present several blind areas. In the most extreme cases, where half of the markers (3) were totally blind, the system was still able to carry out the recognition with a score of 50%, given that the remaining markers could be correctly identified. The system also properly recognized images with angles of inclination of up to 68° with respect to the horizontal (axis OX). It was also evidenced that this new approach overcomes some challenging aspects of an angled approach, such as the presence of shadows or severe occlusions in the scene, as well as the overexposure of the images, and distortions.

In the future, one of the main objectives is to use additional recognition approaches, perhaps similar to those used for facial recognition [86] but as applied to the landing platform. In this regard, the face detection feature of the mobile device attached to the UAV frames the landing platform in a single rectangle in the same way as it does for human faces. In this case, more sophisticated training processes will be needed.

As mentioned in the “Contributions”, neural networks can be used for video and image processing [68, 69] and applied to perform autonomous landing. This would extend this method to identify a family of targets, thinking about the distribution of goods in an aerial way. However, convolutional neural networks, like any neural network model, are computationally expensive. In addition, they also have many hyperparameters that need to be adjusted to train them well. Hence, the number of the test images needed by the neural network will be huge compared to the current approach. The presented method complies with the objective of identifying a unique landing pattern in a faster, simpler, and more efficient way than a deep learning-based approach.

In later stages of development, an upgrade of the mobile device is planned. The aim is to obtain better performance with a lower power usage and to upgrade to 4G networks.

Appendix

Tables 4, 5, and 6 show the equations for area, distance, and angles used in the recognition probability function. To be considered a hit, the equation must be met. All numeric values for each similarity measurement are calculated by using zenithal perspective shown in Fig. 2. Additional flexibility is provided by an available threshold which has been estimated using hundreds of tests and a trial and error method.

Table 4 Fifteen area measurements used for recognition probability function. These similarity measurements compare the area between regions shown in Fig. 2. A threshold of 60% is used to perform recognition in any kind of perspective

Area-based relations with ArT = 0.6	
Ar01	$ LEB.Area / LE.Area - 0.4078 \leq 0.4078 * ArT;$
Ar02	$ LEB.Area / REB.Area - 0.5093 \leq 0.5093 * ArT;$
Ar03	$ LEB.Area / RE.Area - 0.5990 \leq 0.5990 * ArT;$
Ar04	$ LEB.Area / N.Area - 0.6516 \leq 0.6516 * ArT;$
Ar05	$ LEB.Area / M.Area - 0.3002 \leq 0.3002 * ArT;$
Ar06	$ LE.Area / REB.Area - 1.4001 \leq 1.4001 * ArT;$
Ar07	$ LE.Area / RE.Area - 1.4946 \leq 1.4946 * ArT;$
Ar08	$ LE.Area / N.Area - 1.7910 \leq 1.7910 * ArT;$
Ar09	$ LE.Area / M.Area - 0.7359 \leq 0.7359 * ArT;$
Ar10	$ REB.Area / RE.Area - 1.1760 \leq 1.1760 * ArT;$
Ar11	$ REB.Area / N.Area - 1.2792 \leq 1.2792 * ArT;$
Ar12	$ REB.Area / M.Area - 0.5612 \leq 0.5612 * ArT;$
Ar13	$ RE.Area / N.Area - 1.0878 \leq 1.0878 * ArT;$
Ar14	$ RE.Area / M.Area - 0.4924 \leq 0.4924 * ArT;$
Ar15	$ N.Area / M.Area - 0.4387 \leq 0.4387 * ArT;$

Table 5 One hundred and five distance measurements used for recognition probability function. These similarity measurements compare the distance between the centroid of the regions shown in Fig. 2. A threshold of 55% is used to perform recognition in any kind of perspective

Relations with the distance between centroids with DT = 0.55	
Dst01	$ d(LEB_LE) / d(LEB_RE) - 0.5 \leq 0.5 * DT$
Dst02	$ d(LEB_LE) / d(LEB_RE) - 0.4444 \leq 0.4444 * DT$
Dst03	$ d(LEB_LE) / d(LEB_N) - 0.7071 \leq 0.7071 * DT$
Dst04	$ d(LEB_LE) / d(LEB_M) - 0.4444 \leq 0.4444 * DT$
Dst05	$ d(LEB_LE) / d(LE_REB) - 0.4444 \leq 0.4444 * DT$
Dst06	$ d(LEB_LE) / d(LE_RE) - 0.5 \leq 0.5 * DT$
Dst07	$ d(LEB_LE) / d(LE_N) - 1 \leq 1 * DT$
Dst08	$ d(LEB_LE) / d(LE_M) - 0.7071 \leq 0.7071 * DT$
Dst09	$ d(LEB_LE) / d(REB_RE) - 1 \leq 1 * DT$
Dst10	$ d(LEB_LE) / d(REB_N) - 0.7071 \leq 0.7071 * DT$
Dst11	$ d(LEB_LE) / d(REB_M) - 0.4444 \leq 0.4444 * DT$
Dst12	$ d(LEB_LE) / d(RE_N) - 1 \leq 1 * DT$
Dst13	$ d(LEB_LE) / d(RE_M) - 0.7071 \leq 0.7071 * DT$
Dst14	$ d(LEB_LE) / d(N_M) - 1 \leq 1 * DT$
Dst15	$ d(LEB_REB) / d(LEB_RE) - 0.8888 \leq 0.8888 * DT$
Dst16	$ d(LEB_REB) / d(LEB_N) - 1.4142 \leq 1.4142 * DT$
Dst17	$ d(LEB_REB) / d(LEB_M) - 0.8888 \leq 0.8888 * DT$
Dst18	$ d(LEB_REB) / d(LE_REB) - 0.8888 \leq 0.8888 * DT$
Dst19	$ d(LEB_REB) / d(LE_RE) - 1 \leq 1 * DT$
Dst20	$ d(LEB_REB) / d(LE_N) - 2 \leq 2 * DT$
Dst21	$ d(LEB_REB) / d(LE_M) - 1.4142 \leq 1.4142 * DT$
Dst22	$ d(LEB_REB) / d(REB_RE) - 2 \leq 2 * DT$
Dst23	$ d(LEB_REB) / d(REB_N) - 1.4142 \leq 1.4142 * DT$
Dst24	$ d(LEB_REB) / d(REB_M) - 0.8888 \leq 0.8888 * DT$
Dst25	$ d(LEB_REB) / d(RE_N) - 2 \leq 2 * DT$

Table 5 (continued)

Relations with the distance between centroids with DT=0.55	
Dst26	$ld(LEB_REB) / d(RE_M) - 1.4142 \leq 1.4142 * DT$
Dst27	$ld(LEB_REB) / d(N_M) - 2 \leq 2 * DT$
Dst28	$ld(LEB_RE) / d(LEB_N) - 1.5910 \leq 1.5910 * DT$
Dst29	$ld(LEB_RE) / d(LEB_M) - 1 \leq 1 * DT$
Dst30	$ld(LEB_RE) / d(LE_REB) - 1 \leq 1 * DT$
Dst31	$ld(LEB_RE) / d(LE_RE) - 1.125 \leq 1.125 * DT$
Dst32	$ld(LEB_RE) / d(LE_N) - 2.25 \leq 2.25 * DT$
Dst33	$ld(LEB_RE) / d(LE_M) - 1.5910 \leq 1.5910 * DT$
Dst34	$ld(LEB_RE) / d(REB_RE) - 2.25 \leq 2.25 * DT$
Dst35	$ld(LEB_RE) / d(REB_N) - 1.5910 \leq 1.5910 * DT$
Dst36	$ld(LEB_RE) / d(REB_M) - 1 \leq 1 * DT$
Dst37	$ld(LEB_RE) / d(RE_N) - 2.25 \leq 2.25 * DT$
Dst38	$ld(LEB_RE) / d(RE_M) - 1.5910 \leq 1.5910 * DT$
Dst39	$ld(LEB_RE) / d(N_M) - 2.25 \leq 2.25 * DT$
Dst40	$ld(LEB_N) / d(LEB_M) - 0.6285 \leq 0.6285 * DT$
Dst41	$ld(LEB_N) / d(LE_REB) - 0.6285 \leq 0.6285 * DT$
Dst42	$ld(LEB_N) / d(LE_RE) - 0.7071 \leq 0.7071 * DT$
Dst43	$ld(LEB_N) / d(LE_N) - 1.4142 \leq 1.4142 * DT$
Dst44	$ld(LEB_N) / d(LE_M) - 1 \leq 1 * DT$
Dst45	$ld(LEB_N) / d(REB_RE) - 1.4142 \leq 1.4142 * DT$
Dst46	$ld(LEB_N) / d(REB_N) - 1 \leq 1 * DT$
Dst47	$ld(LEB_N) / d(REB_M) - 0.6285 \leq 0.6285 * DT$
Dst48	$ld(LEB_N) / d(RE_N) - 1.4142 \leq 1.4142 * DT$
Dst49	$ld(LEB_N) / d(RE_M) - 1 \leq 1 * DT$
Dst50	$ld(LEB_N) / d(N_M) - 1.4142 \leq 1.4142 * DT$
Dst51	$ld(LEB_M) / d(LE_REB) - 1 \leq 1 * DT$
Dst52	$ld(LEB_M) / d(LE_RE) - 1.125 \leq 1.125 * DT$
Dst53	$ld(LEB_M) / d(LE_N) - 2.25 \leq 2.25 * DT$
Dst54	$ld(LEB_M) / d(LE_M) - 1.5910 \leq 1.5910 * DT$
Dst55	$ld(LEB_M) / d(REB_RE) - 2.25 \leq 2.25 * DT$
Dst56	$ld(LEB_M) / d(REB_N) - 1.5910 \leq 1.5910 * DT$
Dst57	$ld(LEB_M) / d(REB_M) - 1 \leq 1 * DT$
Dst58	$ld(LEB_M) / d(RE_N) - 2.25 \leq 2.25 * DT$
Dst59	$ld(LEB_M) / d(RE_M) - 1.5910 \leq 1.5910 * DT$
Dst60	$ld(LEB_M) / d(N_M) - 2.25 \leq 2.25 * DT$
Dst61	$ld(LE_REB) / d(LE_RE) - 1.125 \leq 1.125 * DT$
Dst62	$ld(LE_REB) / d(LE_N) - 2.25 \leq 2.25 * DT$
Dst63	$ld(LE_REB) / d(LE_M) - 1.5910 \leq 1.5910 * DT$
Dst64	$ld(LE_REB) / d(REB_RE) - 2.25 \leq 2.25 * DT$
Dst65	$ld(LE_REB) / d(REB_N) - 1.5910 \leq 1.5910 * DT$
Dst66	$ld(LE_REB) / d(REB_M) - 1 \leq 1 * DT$
Dst67	$ld(LE_REB) / d(RE_N) - 2.25 \leq 2.25 * DT$
Dst68	$ld(LE_REB) / d(RE_M) - 1.5910 \leq 1.5910 * DT$
Dst69	$ld(LE_REB) / d(N_M) - 2.25 \leq 2.25 * DT$
Dst70	$ld(LE_RE) / d(LE_N) - 2 \leq 2 * DT$
Dst71	$ld(LE_RE) / d(LE_M) - 1.4142 \leq 1.4142 * DT$
Dst72	$ld(LE_RE) / d(REB_RE) - 2 \leq 2 * DT$
Dst73	$ld(LE_RE) / d(REB_N) - 1.4142 \leq 1.4142 * DT$
Dst74	$ld(LE_RE) / d(REB_M) - 0.8888 \leq 0.8888 * DT$
Dst75	$ld(LE_RE) / d(RE_N) - 2 \leq 2 * DT$

Table 5 (continued)

Relations with the distance between centroids with DT=0.55	
Dst76	$ld(LE_RE) / d(RE_M) - 1.4142 \leq 1.4142 * DT$
Dst77	$ld(LE_RE) / d(N_M) - 2 \leq 2 * DT$
Dst78	$ld(LE_N) / d(LE_M) - 0.7071 \leq 0.7071 * DT$
Dst79	$ld(LE_N) / d(REB_RE) - 1 \leq 1 * DT$
Dst80	$ld(LE_N) / d(REB_N) - 0.7071 \leq 0.7071 * DT$
Dst81	$ld(LE_N) / d(REB_M) - 0.4444 \leq 0.4444 * DT$
Dst82	$ld(LE_N) / d(RE_N) - 1 \leq 1 * DT$
Dst83	$ld(LE_N) / d(RE_M) - 0.7071 \leq 0.7071 * DT$
Dst84	$ld(LE_N) / d(N_M) - 1 \leq 1 * DT$
Dst85	$ld(LE_M) / d(REB_RE) - 1.4142 \leq 1.4142 * DT$
Dst86	$ld(LE_M) / d(REB_N) - 1 \leq 1 * DT$
Dst87	$ld(LE_M) / d(REB_M) - 0.6285 \leq 0.6285 * DT$
Dst88	$ld(LE_M) / d(RE_N) - 1.4142 \leq 1.4142 * DT$
Dst89	$ld(LE_M) / d(RE_M) - 1 \leq 1 * DT$
Dst90	$ld(LE_M) / d(N_M) - 1.4142 \leq 1.4142 * DT$
Dst91	$ld(REB_RE) / d(REB_N) - 0.7071 \leq 0.7071 * DT$
Dst92	$ld(REB_RE) / d(REB_M) - 0.4444 \leq 0.4444 * DT$
Dst93	$ld(REB_RE) / d(RE_N) - 1 \leq 1 * DT$
Dst94	$ld(REB_RE) / d(RE_M) - 0.7071 \leq 0.7071 * DT$
Dst95	$ld(REB_RE) / d(N_M) - 1 \leq 1 * DT$
Dst96	$ld(REB_N) / d(REB_M) - 0.6285 \leq 0.6285 * DT$
Dst97	$ld(REB_N) / d(RE_N) - 1.4142 \leq 1.4142 * DT$
Dst98	$ld(REB_N) / d(RE_M) - 1 \leq 1 * DT$
Dst99	$ld(REB_N) / d(N_M) - 1.4142 \leq 1.4142 * DT$
Dst100	$ld(REB_M) / d(RE_N) - 2.25 \leq 2.25 * DT$
Dst101	$ld(REB_M) / d(RE_M) - 1.5910 \leq 1.5910 * DT$
Dst102	$ld(REB_M) / d(N_M) - 2.25 \leq 2.25 * DT$
Dst103	$ld(RE_N) / d(RE_M) - 0.7071 \leq 0.7071 * DT$
Dst104	$ld(RE_N) / d(N_M) - 1 \leq 1 * DT$
Dst105	$ld(RE_M) / d(N_M) - 1.4142 \leq 1.4142 * DT$

Table 6 One hundred and five angle measurements used for recognition probability function. These similarity measurements compare the difference of the angles generated by the intersection of the straight line that joins the centroids of the regions regarding the base of the image, i.e., the bottom horizontal line of the image. They range in [0, 360). A threshold of 40° is used to cope with distortion of perspective projection

Relation distances between angle with AgT=40°	
Ang01	$lang(LEB_LE, OX) - ang(LEB_REB, OX) - 90^\circ \leq AgT$
Ang02	$lang(LEB_LE, OX) - ang(LEB_RE, OX) - 63.43^\circ \leq AgT$
Ang03	$lang(LEB_LE, OX) - ang(LEB_N, OX) - 45^\circ \leq AgT$
Ang04	$lang(LEB_LE, OX) - ang(LEB_M, OX) - 26.18^\circ \leq AgT$
Ang05	$lang(LEB_LE, OX) - ang(LE_REB, OX) - 116.64^\circ \leq AgT$
Ang06	$lang(LEB_LE, OX) - ang(LE_RE, OX) - 90^\circ \leq AgT$
Ang07	$lang(LEB_LE, OX) - ang(LE_N, OX) - 90^\circ \leq AgT$
Ang08	$lang(LEB_LE, OX) - ang(LE_M, OX) - 45^\circ \leq AgT$
Ang09	$lang(LEB_LE, OX) - ang(REB_RE, OX) - 0^\circ \leq AgT$

Acknowledgements The authors would like to thank to the physicist Felipe Seoane for reviewing the document and supporting us in our transcription of the research work to this document in the most appropriate and straightforward way possible. The authors would also like to thank Rebeca Cuiñas for her assistance during the design of the landing platform. Finally, we would also give special thanks to Dr. Nicola Tonge and to Dr Mike Bernstein, both NMR (Nuclear Magnetic Resonance) scientists, for their linguistic and stylistic assistance. Special thanks also to Dr Mark Watkins for his proofreading and proof-editing services. The authors thank the anonymous referees for their very valuable comments and suggestions.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- García-Pulido JA, Pajares G, Dormido S, de la Cruz JM. Recognition of a landing platform for unmanned aerial vehicles by using computer vision-based techniques. *Expert Syst Appl*. 2017;76:152–65.
- Cruz JM, Sánchez B, Pajares G. System for guiding an unmanned vehicle towards a platform using visual analysis. Patent no. 2012;201001592:2013.
- Chen S, Zhao X, Sun Z, Xiang F, Zhixin S. Shape recognition with recurrent neural network. *International Conference on Artificial Intelligence and Security*. ICAIS 2019;341–350.
- Tusor B, Takáč O, Molnar A, Gubo S, Varkonyi-Koczy A. Shape recognition in drone images using simplified fuzzy indexing tables. *IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMII)*, 2020;(129–134).
- Unlu E, Zenou, E, Riviere N. Using shape descriptors for UAV detection. *Electronic Imaging*, 2018;1–5.
- Kong W, Hu T, Zhang D, Shen L, Zhang J. Localization framework for real-time UAV autonomous landing: An On-Ground Deployed Visual Approach. *Sensors*. 2017;2017(17):1437.
- Nguyen PH, Kim KW, Lee YW, Park KR. Remote marker-based tracking for UAV landing using visible-light camera sensor. *Sensors*. 2017;2017(17):1987.
- Anitha G, Kumar RNG. Vision based autonomous landing of an unmanned aerial vehicle. *Procedia Eng*. 2012;2012(38):2250–6.
- Xu G, Zhang Y, Ji S, Cheng Y, Tian Y. Research on computer vision-based for UAV autonomous landing on a ship. *Pattern Recognit Lett*. 2009;2009(30):600–5.
- Xu G, Qi X, Zeng Q, Tian Y, Guo R, Wang B. Use of land's cooperative object to estimate UAV's pose for autonomous landing. *Chin J Aeronaut*. 2013;2013(26):1498–505.
- Yang T, Li G, Li J, Zhang Y, Zhang X, Zhang Z, Li Z. A ground-based near infrared camera array system for UAV auto-landing in GPS-denied environment. *Sensors*. 2016;2016(16):1–20.
- Kong W, Zhang D, Wang X, Xian Z, Zhang J. Autonomous landing of an UAV with a ground-based actuated infrared stereo vision system. In *Proceedings of the IEEE/RSJ Int Conf Intel Robot Sys*, Tokyo, Japan, 3–7 November 2013;2963–2970.
- Zhou D, Zhong Z, Zhang D, Shen L, Yan C. Autonomous landing of a helicopter UAV with a ground-based multisensory fusion system. *Seventh International Conference on Machine Vision (ICMV 2014)*. *Int Soc Optics and Photon*. 2015;94451R.
- Tang D, Hu T, Shen L, Zhang D, Kong W, Low KH. Ground stereo vision-based navigation for autonomous take-off and landing of UAVs: a Chan-Vese model approach. *Int J Adv Robot Sys*. 2016;13(2) 67:2016.
- Martínez C, Campoy P, Mondragón I, Olivares-Méndez M.A. Trinocular ground system to control UAVs. In *Proceedings of the IEEE/RSJ Int Conf Intel Robot Sys St. Louis, MO, USA, 10–15 October 2009*;3361–3367.
- Kong W, Zhou D, Zhang Y, Zhang D, Wang X, Zhao B, Yan C, Shen L, Zhang J. A ground-based optical system for autonomous landing of a fixed wing UAV. In *Proc. IEEE/RSJ Int Conf Intel Robot Sys (IROS 2014)* 2014;14–18, Chicago, IL, USA, 1–8.
- Asadzadeh M, Palaiahnakote S, Idris MYI, Anisi MH, Lu T, Blumenstein M, Noor NM. An automatic zone detection system for safe landing of UAVs. *Expert Syst Appl*. 2019;122:319–33.
- Patterson T, McClean S, Morrow P, Parr G, Luo C. Timely autonomous identification of UAV safe landing zones. *Image Vis Comput*. 2014;32(9):568–78.
- Li X. A software scheme for UAV's safe landing area discovery. *AASRI Procedia*. 2013;(4):230–5.
- Al-Sharman MK, Emran BJM, Jaradat A, Najjaran H, Al-Husari R, Zweiri Y. Precision landing using an adaptive fuzzy multi-sensor data fusion architecture. *Appl Soft Comput*. 2018;69:149–64.
- Al-Sharman M, Al-Jarrah M.A, Abdel-Hafez M. Auto takeoff and precision terminal-phase landing using an experimental optical flow model for GPS/INS enhancement. *ASCE-ASME J Risk and Uncertainty in Engineering Systems, Part B: Mech Eng*. 2018.
- Gui Y, Guo P, Zhang H, Lei Z, Zhou X, Du J, Yu Q. Airborne vision-based navigation method for UAV accuracy landing using infrared lamps. *J Int Robot Sys*. 2013;72,(2)197.
- Forster C, Faessler M, Fontana F, Werlberger M, Scaramuzza D. Continuous on-board monocular-vision-based elevation mapping applied to autonomous landing of micro aerial vehicles. *Proc. in Robotics and Automation (ICRA)*, 2015 IEEE Int Conf on. IEEE, 2015;111–118.
- Johnson A, Montgomery J, Matthies L. Vision guided landing of an autonomous helicopter in hazardous terrain. *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2005.
- Bosch S, Lacroix S, Caballero F. Autonomous detection of safe landing areas for an UAV from monocular images. *IEEE/RSJ Intl Conf Intel Robot Sys (IROS)*. 2006.
- Desaraju V, Michael N, Humenberger M, Brockers R, Weiss S, Matthies L. Vision-based landing site evaluation and trajectory generation toward rooftop landing. *Autonomous Robots* 2015;39,3,445–463.
- Davide F, Alessio Z, Alessandro S, Jeffrey D, Scaramuzza D. Vision-based autonomous quadrotor landing on a moving platform. *J Int Robot Sys* February 2017;85,2(369–384).

28. Lee D, Ryan T, Kim H.J. Autonomous landing of a VTOL UAV on a moving platform using image-based visual servoing. *Robotics and Automation (ICRA)*, 2012 IEEE International Conference on. IEEE, 2012;971–976.
29. Lee H, Jung S, Shim D.H. Vision-based UAV landing on the moving vehicle. In *Proceedings of the International Conference on Unmanned Aircraft System*, Arlington, MA, USA, 2016;1–7.
30. Feng Y, Zhang C, Baek S, Rawashdeh S, Mohammadi A. Autonomous landing of a UAV on a moving platform using model predictive control. *Drones* 2018;2,34.
31. Line V. Autonomous landing of a multirotor UAV on a platform in motion. Master Thesis, Norwegian University of Science and Technology, Norway. Available online: <https://brage.bibsys.no/xmloi/handle/11250/2558185>. (accessed on May 2021). 2018.
32. Rodríguez-Ramos A, Sampedro C, Bavle H, Milosevic Z, García-Vaquero A, Campoy P. Towards fully autonomous landing on moving platforms for rotary unmanned aerial vehicle. *Unmanned Aircraft Systems (ICUAS) 2017 Int Conf on*. 170–178. 2017.
33. Polvara R, Sharma S, Wan J, Manning A, Sutton R. Towards autonomous landing on a moving vessel through fiducial markers. *IEEE European Conf Mobile Robot (ECMR)*. IEEE. 2017.
34. Lin S, Garratt MA, Lambert AJ. Monocular vision-based real-time target recognition and tracking for autonomously landing an uav in a cluttered shipboard environment. *Autonom Robot*, 2017;41,(4)881–901.
35. Lange S, Sünderhauf N, Protzel P. Autonomous landing for a multirotor uav using vision. *Workshop Proc. of SIMPAR 2008 Intl. Conf. on Simulation, Modelling and Programming for Autonomous Robots*, 2008;482–491.
36. Lange S, Sunderhauf N, Protzel P. A vision based onboard approach for landing and position control of an autonomous multirotor uav in gps-denied environments. *Advanced Robotics*, 2009. *ICAR 2009. International Conference on IEEE*. 2009;2009:1–6.
37. Saripalli S, Montgomery JF, Sukhatme GS. Vision-based autonomous landing of an unmanned aerial vehicle. *IEEE International Conference on Robotics and Automation*. 2002;11–15:2799–804.
38. Polvara R, Patacchiola M, Sharma S, Wan J, Manning A, Sutton R, Cangelosi A. Autonomous quadrotor landing using deep reinforcement learning. *arXiv preprint*. 2017.
39. Cocchioni F, Mancini A, Longhi S. Autonomous navigation, landing and recharge of a quadrotor using artificial vision. *Int Conf Unman Aircraft Sys (ICUAS)*. 2014;418–429.
40. Li Y, Wang Y, Luo H, Chen Y, Jiang Y. Landmark recognition for UAV autonomous landing based on vision. *Application Research of Computers*. 2012;07:2780–3.
41. Chen J, Miao X, Jiang H, Chen J, Liu X. Identification of autonomous landing sign for unmanned aerial vehicle based on faster regions with convolutional neural network. *IEEE Int Conf Chinese Autom Congress (CAC)*, 2017;2019–2114.
42. Sharp CS, Shakernia O, Sastry SS. A vision system for landing an unmanned aerial vehicle. *IEEE Int Conf on Robot and Automation*, 2001;1720–1727.
43. Wang L, Yang Q, Guo X. Recognition algorithm of the apron for unmanned aerial vehicle based on image corner points. *Laser Journal*. 2016;08:71–4.
44. Nyein EE, Tun HM, Naing ZM, Moe WK. Implementation of vision-based landing target detection for VTOL UAV using raspberry Pi. *Int J Sci Technol Res*. 2015;4(8):184–8.
45. Zhao YJ, Pei HL. An improved vision-based algorithm for unmanned aerial vehicles autonomous landing. *Appl Mechan Mater*. 2013;273(560–565).
46. Bay H, Ess A, Tuytelaars T, Van Gool L. SURF: speeded up robust features. *Comput Vis Image Underst*. 2008;110(3):346–59.
47. Saavedra-Ruiz M, Pinto-Vargas AM, Romero-Cano V. Detection and tracking of a landing platform for aerial robotics applications. *IEEE 2nd Colombian Conf Robot Automation (CCRA)*, Barranquilla. 2018;1–6.
48. Saripalli S, Sukhatme G. Landing on a moving target using an autonomous helicopter. *Field and service robotics*. Springer Tracts in Advanced Robotics. 2006;2006:277–86.
49. Guili X, Yong Z, Shengyu J, Yuehua CH, Yupeng T. Research on computer vision-based for UAV autonomous landing on a ship. *Pattern Recogn Lett*. 2009;30(6):600–5.
50. Olson E. AprilTag: a robust and flexible visual fiducial system. *IEEE Int Conf Robotics and Automation (ICRA)*. 2011.
51. AprilTag. Available online: <https://april.eecs.umich.edu/software/apriltag.html>. (accessed on May 2021).
52. Ling K. Precision landing of a quadrotor uav on a moving target using low-cost sensors. Master's Thesis, University of Waterloo, Canada. UWSpace. <http://hdl.handle.net/10012/8803>. Available on-line: <https://uwspace.uwaterloo.ca/handle/10012/8803>. (accessed on May 2021). 2014.
53. Kyristis S, Antonopoulos A, Chanielakis T, Stefanakis E, Linardos C, Tripolitsiotis A, Partsinivelos P. Towards autonomous modular UAV missions: the detection, geolocation and landing paradigm. *Sensors* 2016;16(11)1844.
54. Garrido-Jurado S, Muñoz-Salinas R, Madrid-Cuevas FJ, Marín-Jiménez MJ. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Reco* 2014;47:6(2280–2292).
55. Detection of ArUco markers. Available online: http://docs.opencv.org/trunk/d5/dae/tutorial_aruco_detection.html. (accessed on May 2021).
56. Chaves SM, Wolcott RW, Eustice RM. NEEC research: toward GPS-denied landing of unmanned aerial vehicles on ships at sea. *Nav Eng J* 2015;127(23–35).
57. Araar O, Aouf N, Vitanov I. Vision based autonomous landing of multirotor uav on moving platform. *J Int Robot Syst* 2017;85(369–384).
58. Sani F, Karimian G. Automatic navigation and landing of an indoor AR. drone quadrotor using ArUco marker and inertial sensors. *Inte Conf Comp Drone Appl (IConDA)*, 2017;102–107.
59. Rabah M, Rohan A, Talha M, Nam K, Kim S. Autonomous vision-based target detection and safe landing for UAV. *Int J Control Autom Syst*. 2018;16:3013–25.
60. Shen J, Peng J, Dong X, Shao L, Porikli F. Higher order energies for image segmentation. *IEEE Trans Image Processing*. 2017;26(10):4911–22.
61. Shen J, Du Y, Li X. Interactive segmentation using constrained Laplacian optimization. *IEEE Trans. Circuits and Syst for Video Techn*. 2014;24(7),1088–1100.
62. Peng J, Shen J, Li X. High-order energies for stereo segmentation. *IEEE Trans Cybernetics*. 2016;46(7):1616–27.
63. Shen J, Du Y, Wang W, Li X. Lazy random walks for superpixel segmentation. *IEEE Trans Image Processing*. 2014;23(4):1451–62.
64. Shen J, Peng J, Shao L. Submodular trajectories for better motion segmentation in videos. *IEEE Trans Image Processing*. 2018;27(6):2688–700.
65. Lôbo-Medeiros FL, Faria-Gomes VC, Campos de Aquino MR, Geraldo D, Lopes-Honorato ME, Moreira-Dias LH. A computer vision system for guidance of VTOL UAVs autonomous landing. *Br Conf Intel Syst (BRACIS)*, 2015;333–338.
66. Nguyen PH, Kim KW, Lee YW, Park KR. Remote marker-based tracking for uav landing using visible-light camera sensor. *Sensors* 2017;17(9)1987.
67. Cesetti A, Frontoni E, Mancini A, Zingaretti P, Longhi S. A vision-based guidance system for uav navigation. *Journal of Intelligent and Robotic Systems*, 2010;57(1–4),233–257.
68. Dong X, Shen J, Wu D, Guo K, Jin X, Porikli F. Quadruplet Network with One-Shot Learning for Fast Visual Object Tracking. *IEEE Trans Image Process*. 2019;28(7):3516–27.

69. Wang W, Shen J, Ling H. A Deep Network Solution for Attention and Aesthetics Aware Photo Cropping. *IEEE Trans Pattern Anal Machine Intell.* 2019;41(7):1531–44.
70. Hu MK. Visual Problem recognition by Moment Invariant. *IRE Trans. Inform. Theory, IT-8,* 1962;179–187.
71. Koschan A, Abido M. *Digital Color Image Processing.* John Wiley & Sons; 2008.
72. Bohanec M. Decision making: A computer-science and information-technology viewpoint. *Interdiscip Descr Complex Syst.* 2009;7(2):22–37.
73. The MathWorks. Color-Based Segmentation Using the L*a*b* Color Space. <https://es.mathworks.com/help/images/examples/color-based-segmentation-using-the-l-a-b-color-space.html>. (accessed on May 2021). 2019.
74. Campos Y, Sossa H, Pajares G. Spatio-temporal analysis for obstacle detection in agricultural videos. *Appl Soft Comput.* 2016;45:86–97.
75. Shen J, Hao X, Liang Z, Liu Y, Wang W, Shao L. Real-Time Superpixel Segmentation by DBSCAN Clustering Algorithm. *IEEE Trans Image Processing.* 2016;25(12):5933–42.
76. Dillencourt MB, Samet H, Tamminen M. A general approach to connected-component labeling for arbitrary image representations. *Journal of the Association for Computing Machinery.* 1992;39(2):252–80.
77. Sedgewick R. *Algorithms in C, 3rd Ed.,* Addison-Wesley, 11–20. 1998.
78. Mercimek M, Gulez K, Mumcu TV. Real object recognition using moment invariants. *IEEE Transactions on Pattern Analysis.* 2005;30:765–75.
79. Theodoridis S, Koutroumbas K. *Pattern Recognition.* Elsevier Academic Press 4th edition. 2009.
80. Dougherty ER, Lotufo RA. *Hands-on Morphological Image Processing.* SPIE Tutorial Texts in Optical Engineering Vol. TT5. SPIE Publications. 2003.
81. Duda RO, Hart PE, Stork DG. *Pattern classification (2nd ed.).* New York; New Delhi: Wiley. 2006.
82. Zimmermann HJ. *Fuzzy set Theory and its applications.* Norwell, USA: Kluwer Academic Publishers; 1991.
83. The MathWorks. The Matlab. <https://es.mathworks.com/products/matlab>. (accessed on May 2021). 2019.
84. Kalal Z, Mikolajczyk K, Matas J. Tracking–learning–detection. *IEEE Trans Pattern Anal Mach Intell.* 2012;2012(34):1409–22.
85. Henriques JF, Caseiro R, Martins P, Batista J. High–speed tracking with kernelized correlation filters. *IEEE Trans Pattern Anal Mach Intell.* 2015;2015(37):583–96.
86. Viola P, Jones V. Rapid Object Detection using a Boosted Cascade of Simple Features. *Proceedings of the 2001 IEEE Comp Soc Conf on Comp Vision and Pattern Recognition.* 2001;1(511–518).

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