# Vision-based UAV Detection for Air-to-Air Neutralization

Fidel González[1], Rafael Caballero[1], Francisco J. Pérez-Grau[1] and Antidio Viguria[1]

***Abstract*— The widespread availability of Unmanned Aerial Vehicles (UAVs) poses potential threats for people and properties on the ground, and other airspace users. This work introduces the design, development and validation of a UAV neutralization system that is based on another UAV with a capture device. The operation is fully autonomous, and only relies on data captured by two cameras onboard the captor UAV: one for long-range detections up to 40m, and another one for short-range accurate estimations prior to the actual capture. The approach has been extensively validated in field experiments, proving robustness and computational efficiency.**

## I. INTRODUCTION

The use of Unmanned Aerial Vehicles (UAVs) has risen dramatically in recent years, making it a widely accessible technology that has become quite popular among the general public. The development of low-cost, easy-to-fly UAVs has resulted in an increase in the amount of incidents in which the safety and security of people and property on the ground, as well as other airspace users, has been compromised. When used with malicious intentions, these platforms can pose a threat to society, which is why law enforcement agencies are interested in UAV neutralization systems.

The challenge of neutralizing a potentially malicious UAV, from now on the intruder, can generally be divided into two stages: the first one involves detecting, localizing, tracking, and classifying it; the second one involves its neutralization to prevent it from achieving its goal. Our main interest is to minimize the damage caused by the possible impact of the intruder when intercepted. In this way, we propose use another UAV, from now on the captor, with a mechanical system to catch the intruder, as Fig. 1 shows. End-users such as law enforcement agencies are also interested in conducting an investigation to analyze intruder's information (owner, payload...). In addition, this allows avoiding possible side effects if the intruder carries a volatile load.

Our main contribution is the design, development and validation with field experiments of a completely autonomous UAV neutralization system, carried out by another UAV. The main focus of this paper is on the first stage of the system, i.e. the strategy and algorithms for the detection and localization of the intruder, using information provided exclusively by onboard sensors on the captor.

The rest of the paper is structured as follows. Section II describes existing systems for UAV detection and interception. Section III presents a description of the system, providing details of both the hardware and software architectures.

[1]Fidel González, Rafael Caballero, Francisco J. Pérez-Grau and Antidio Viguria are with the Advanced Center for Aerospace Technologies (CATEC), Seville, Spain. fgonzalez@catec.aero

Fig. 1: Picture of the capture moment.

Section IV discusses the approach to detect and approach the intruder. Then, Section V covers the experiments carried out to validate the detection strategy. Finally, the conclusions and future lines of research are summarized.

## II. RELATED WORK

There are already commercial products and systems under development for UAV neutralization purposes, using a variety of technologies [1]. Currently, most of these developments involve jamming techniques for introducing noise, typically to Global Positioning System (GPS) signals [2], in order to hinder or even prevent the flight. The main drawback of this approach is that the neutralized UAV will most likely cause an uncontrolled accident [3]. Other solutions are even more aggressive and involve shooting down the UAV via high-power laser systems [4].

The detection of the intruder can be performed from the ground or from another UAV, using different types of sensors. Radar sensors can be used for detection from the ground [5]. They provide long-range detection and their operation is not dependent on adverse weather conditions. Although these sensors are the most widely used for detecting UAVs from the ground, they present difficulties in detecting small UAVs, which can be mistaken for birds [6], and besides they are usually very expensive. Nevertheless, some authors have recently used low-cost radar to guide a pursuing behavior [7].

Acoustic sensors are another way to detect intruder UAVs. Standard propellers emit a sound that can be characterized and detected with a network of acoustic sensors placed for that purpose, such as [8] [9] [10]. Although this method has advantages such as the detection of UAVs without communications link, it suffers from weather conditions or a short detection range, not to mention its great uncertainty in general, since different UAVs exhibit different sounds. Moreover, the sensitivity to external noise makes it unfeasible for detecting a UAV from another UAV.

LIDARs bring the benefit of providing rich information about the shape and distance to the target [11], but small UAVs may not be detected in long range scenarios due to the required resolution. Some other approaches make use of event cameras [12], which are suitable for ground-based detection systems, but their low resolution prevents them from constituting an effective detection system. Moreover, an onboard detection system based on this technology would provide an unmanageable amount of events, due to the highly dynamic nature of onboard images.

More suitable for onboard detection are color cameras [13], probably the most widespread sensing option for UAVs due to their low size and cost. Some approaches apply classic vision computer techniques such as color and contour analysis [14], which have a limited range due to image resolution and motion blur. Most of the recent research on UAV detection from images is mainly dominated by Deep Learning approaches using trained Convolutional Neural Networks (CNNs) [15] [16], but detections are usually noisy and not consistent in consecutive frames.

## III. SYSTEM DESCRIPTION

Unlike most of the solutions proposed so far, our approach is based on image analysis using two cameras in order to cover a longer detection range. A 5MP Basler camera with a 6mm lens was selected for long-range detections, allowing the algorithm to perform detections from up to 40m between the intruder and the captor. For short-range detections, an Intel RealSense D435i RGB-D camera was chosen because of its good performance outdoors and low baseline, producing reliable and accurate depth images at up to 10m. The onboard computer is a NVIDIA Jetson TX2, which performs all the computations without requiring any ground-based processing system.

To carry out the pursue and interception, a custom aerial platform was designed, as shown in Fig. 2. The platform consist of a cross-structure made of carbon fiber tubes with an additional carbon fiber tube to hold a net of 2x4 meters for the physical capture. The autopilot is a DJI A3 with its GPS antenna and compass.

Regarding the software, two main frameworks have been used: Matlab/Simulink and ROS. The whole software overview is depicted in Fig. 4. Starting with the Matlab/Simulink algorithms, the *Control* block implements a velocity-based pure pursuit control, since the smoothest and more robust output produced by both detectors are the normalized Line Of Sight (LOS) vector indicating the

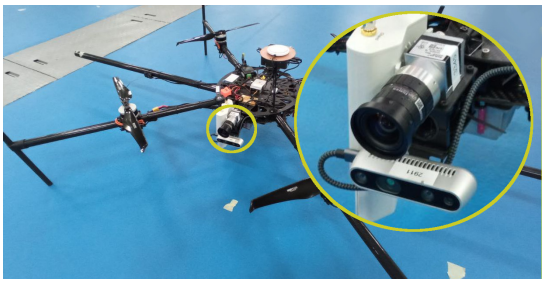

Fig. 2: Platform with camera detail.

direction of the intruder with respect to the captor. This LOS vector goes through a *Target Pose Estimation* algorithm, which thanks to a Kalman Filter, is responsible for feeding a constant rate reference to the control, filtering detection outliers and noise, and also increase the reference's rate from the approximately 10 Hz provided by *Vision Algorithms* to 50 Hz, more appropriate for control algorithms.

In the ROS-based algorithms, it is important to mention the *State machine* (see Fig. 3) which maintains the control logic of the guidance, providing the intelligence behind the fully autonomous capture. There are safety transitions that are not represented for clarity, going to manual and landing states from each of the other states.

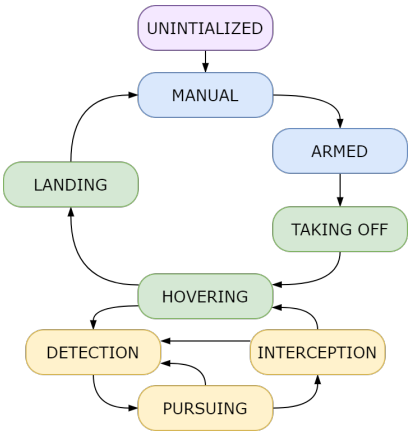


Fig. 3: State machine main states and main transitions.

Moreover, the *Safety manager* algorithm is included in this architecture for checking periodically that every critical subsystem is working properly. For example, it checks the battery state so if its level is lower than a specific value, notifies to the *State machine* which immediately switches to landing state. Also, it ensures that all required data between software modules is available at proper rates.

The next key software piece of this system is the *DJI SDK*, which is in charge of publishing the essential UAV data (position, attitude, battery state...), receiving control commands (velocity commands in our case), and communicating this information to the DJI *Autopilot* via serial port. The *Vision Algorithms* block, highlighted in green, is the main focus of this paper, and will be discussed in detail in Section IV. There is a different algorithm for each camera, but both

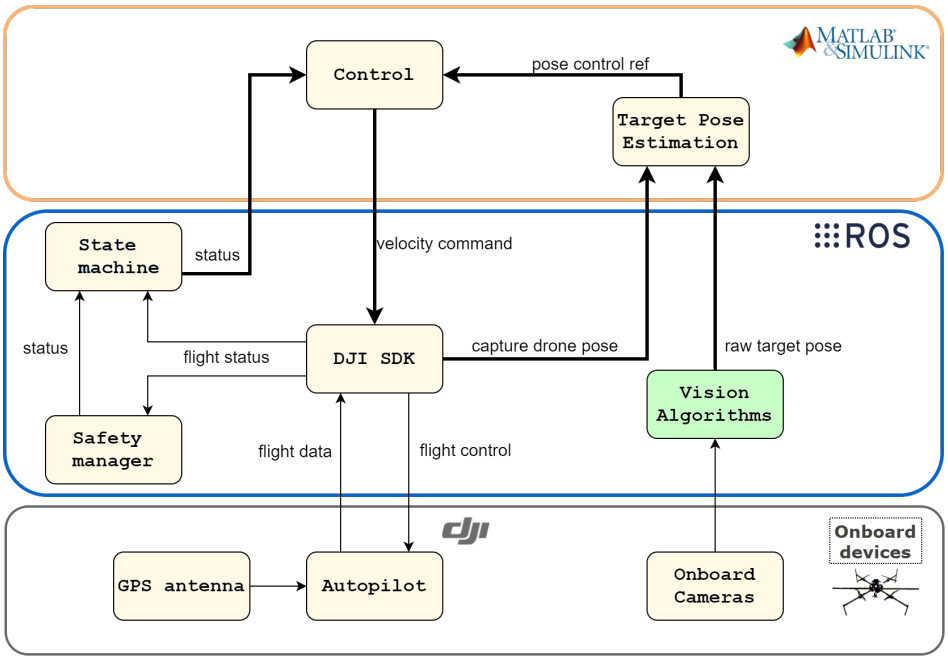


Fig. 4: Software architecture and devices overview.

provide a LOS vector connecting the captor to the estimated position of the intruder, as well as the director vector, which is the result of normalizing the LOS vector. This is the actual information used for the UAV control.

## IV. DETECTION STRATEGY

This section provides a description of the two algorithms used for the detection strategy. As mentioned before, one was designed to provide a long-range detection for target approaching, and the other a short-range detection with higher accuracy for the actual interception.

### A. Long-Range Detection

The core of this algorithm is a CNN, in particular YOLOv3-tiny [17]. The implementation was made using the open-source Darknet framework. The chosen model is not the best available but is good enough to provide continuous UAV detections up to 40m using the onboard TX2 computer at the camera framerate.

Fig. 5: Long-Range Detection: sample full frame with UAV detection detail.

Regarding the dataset used to train the network, it must be taken into account that vision-based air-to-air UAV detection is a very specific application, so there are not many available datasets for that purpose. Therefore, a custom dataset was created during test flights in different scenarios with different aerial platforms and illumination conditions. These recorded images have been manually labelled, and then data augmentation was performed (random flipping, illumination and contrast changes) in order to increase the data variability and the number of images. At the end, a total of 88.193 labelled images were collected. They were split into 90%/10% for training/evaluation.

The models generated gave a mean Average Precision (mAP) of 83%, and were also tested using completely new images from a different scenario and sensor, obtaining surprising results with a true positive rate of 70.15%.

The long-rage detection algorithm steps are basically four, as Fig. 7 shows. The first step is to gather the input data: an RGB image, the camera intrinsic parameters, the UAV position in local Cartesian coordinates (used for debugging) and the transformation between the camera and the UAV center of mass. Then, the inferencing step outputs the coordinates in pixels of the center of the detection, as well as the width and height of the bounding box that encapsulates the detection. Hereafter, the distance to the intruder is estimated using the width and height values obtained and the assumed proportion between the area and the distance by multiplying the area of the detection box by this parameter. This depth is used to obtain a director vector, or *LOS*, in the localization step. Since this vector may be noisy due to the NN inference, the long-range detection is used only to approach to the intruder until the short-range detection can provide robust estimations. The last step is shared by both the long- and short-range detection algorithms, therefore it will be addressed separately in Section IV-B.

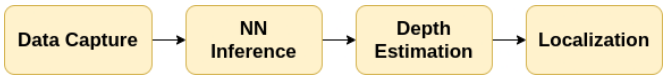


Fig. 7: Long-range detection pipeline.

### B. Short-Range Detection

The short-range detection algorithm makes the assumption that in the flight zone, the intruder UAV is the only object that satisfies the algorithm conditions, since the neutralization procedure will most likely take place in open air.

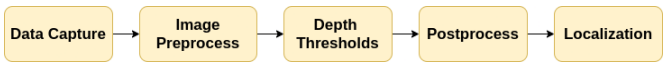


Fig. 8: Short-range detection pipeline.

The algorithm can be divided into five main steps (see Fig. 8). Firstly, the depth and RGB images, camera transformation and UAV position information are gathered, as in the long-range detection. Then, the depth image is pre-processed by setting the unknown pixels to the maximum value (*max depth*), and a morphological opening is performed with a predefined *kernel_size*. Afterwards, the pre-processed depth image is thresholded $N$ (1) times, from 0 to *max_depth* with a specific *depth_step*, in order to obtain multiple depth masks. For each mask, contours are extracted, and their shape is analyzed so only those which are within *max/min_area*, *max/min_circularity* and *max/min_convexity* bounds are considered valid candidates for intruder detections.

$$N = \frac{max\_depth\_value}{depth\_step} \tag{1}$$

Later, the valid candidates are post-processed to check if their depth masks are closer than *max_shift_between_candidates*, and the amount of candidates is higher than *min_candidates_size*. Last, the localization step is described hereafter.

### C. Localization

The localization step takes place once the intruder position on the image and its depth are estimated, in order to obtain the $LOS$ vector used as input for the captor control.

The director vector and the relative position of the target with respect to the camera is calculated using the pinhole camera model (2 and 7) and the plumb bob distortion model (3, 4 and 5), where $(u, v)$ is the center in pixels of the detected contour, $(c_x, c_y)$ is the principal point in the image, $(f_x, f_y)$ is the focal length in both axis and $(k_1, k_2, k_3, p_1, p_2)$ are the distortion parameters.

$$x' = \frac{u - c_x}{f_x}; y' = \frac{v - c_y}{f_y} \tag{2}$$

$$r = \sqrt{x'^2 + y'^2}; f = 1 + k_1 * r^2 + k_2 * r^4 + k_3 * r^6 \tag{3}$$

$$x'' = x' * f + 2 * p_1 * x' * y' + p_2 * (r^2 + 2 * x'^2) \tag{4}$$

$$y'' = y' * f + 2 * p_2 * x' * y' + p_1 * (r^2 + 2 * y'^2) \tag{5}$$

$$Z_{\text{corrected}} = depth * \frac{1}{\sqrt{|1 + x''^2|}} * \frac{1}{\sqrt{|1 + y''^2|}} \tag{6}$$

$${}^{\text{Cstd}}T_{\text{target}} = \begin{bmatrix} x'' * Z \\ y'' * Z \\ Z \end{bmatrix} \tag{7}$$

It is important to note that long- and short-range detectors differ on how $Z$ is calculated. The short-range algorithm considers $Z$ directly as the distance provided by the depth image (depth perpendicular to image plane), while the long-range algorithm estimates the distance from the bounding box area and represents the diagonal depth, instead of the perpendicular depth. Thus, 6 is only applied to the long-range detector to correct the diagonal angles of the object with respect to the center of the image through some trigonometric identities. Once the pose of the object is determined, the estimated LOS vector is normalized.

## V. EXPERIMENTAL RESULTS

Different experiments have been carried out to test the proposed capture system. In these experiments, the intruder UAV was a DJI F450 platform carrying an onboard RaspberryPi 4 for recording ground-truth telemetry provided by its DJI N3 autopilot. Remotely operated flights performing random trajectories were performed with the intruder for testing the autonomous capture system. The estimated depths of both algorithms were analyzed separately, while the estimated *LOS* vectors were compared together.

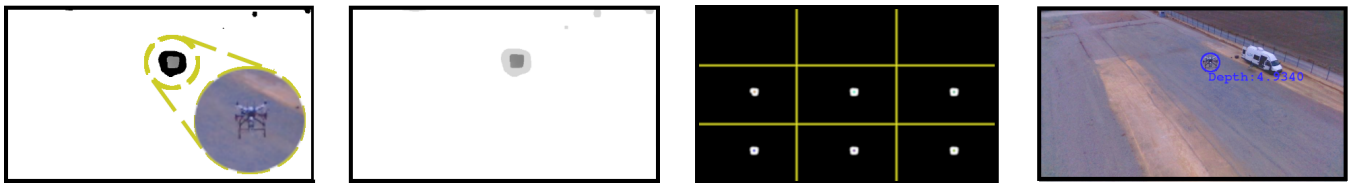

Fig. 6: Short-Range Detection: from left to right: input depth image (with color image detail), pre-processed image, thresholded binary masks, and final detectionresult.

### A. Long-Range Detection

The presented test flight starts with the intruder at around 43m from the captor, and an approach maneuver until the relative distance was 10m. As expected, the long-range detection is inaccurate in terms of relative position calculation, since the estimated depth is not completely reliable. This does not pose major problems to the capture strategy, since the valuable output of the long-range detector is the *LOS* vector, not this depth.

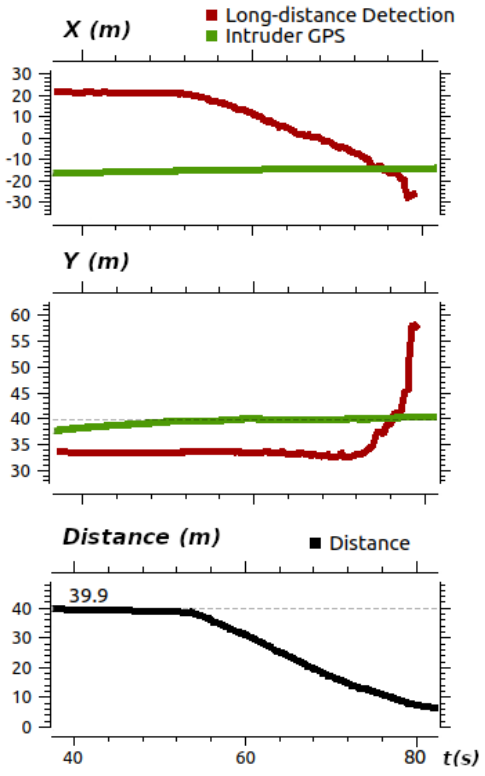


Fig. 9: Long-range detection maximum distance.

It can be observed in Fig. 9 that, even with a small platform as the intruder, once it gets closer than 39.9 m, the detection is constant despite of its position error. This range and continuity could be affected by complex backgrounds or occlusions of the intruder, but proves to be very stable when the background is uniform (e.g. the sky), allowing low altitude flight strategies being optimal to succeed in finding the intruder.

### B. Short-Range Detection

A fragment of a successful capture is shown to demonstrate the maximum range for this algorithm. Once the target is closer than 8m, detections are very stable and accurate.

However, it can be observed in Fig. 10 that the detections start when the distance between UAVs is 11.58 m, providing accurate estimations throughout the full final approach before capture. The slight error in X-axis is due to the fact that there must be a 2m separation between the captor and the intruder in order to effectively acquire the intruder platform with the net that is hanging from the captor.

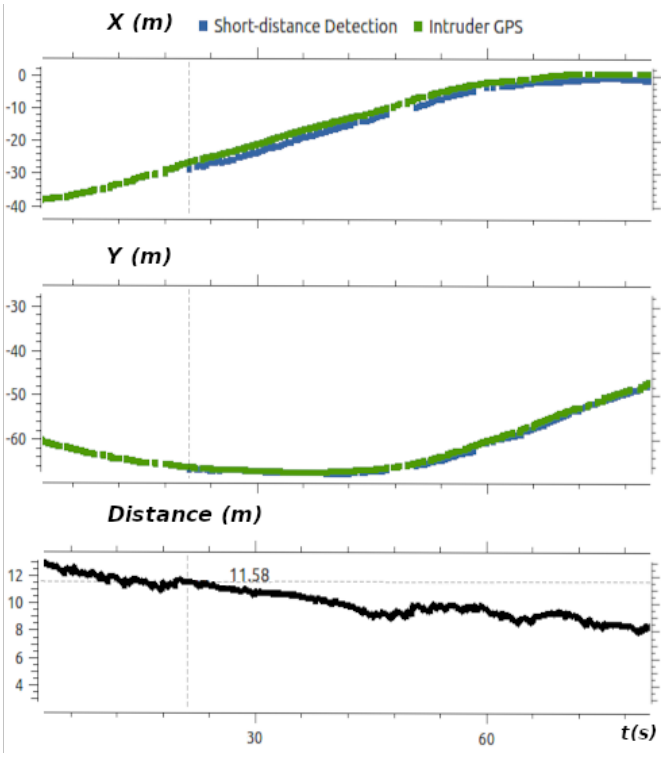


Fig. 10: Short-range detection maximum distance.

### C. LOS Vector Comparison

As the short-range detection proves to have enough accuracy to precisely determine the position of the intruder, its *LOS* vector will be used in order to analyze the long-range precision in terms of relative direction. A fragment of a flight performing the pure pursuit of the intruder UAV ending with its capture is used to compare the outputs, when both detectors are providing estimations.

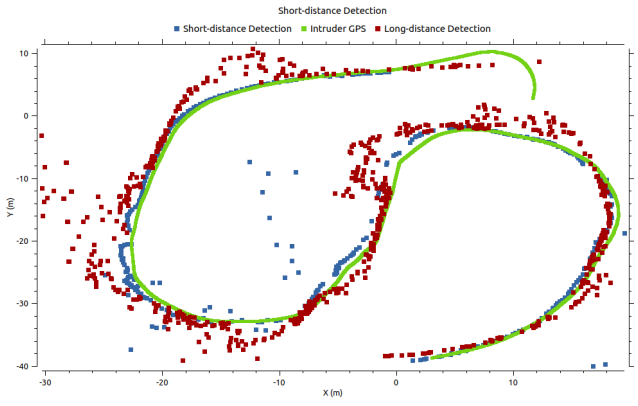


Fig. 11: Estimated intruder positions with long- and short-range distance detectors.

The resulting *LOS* director vector, which is adimensional, shows that the direction of the target estimated by long-range detection is close to the accurate vector produced by short-range detection. Both positions (Fig. 11) and directions (Fig. 12) become really close, validating that the detection strategy

is precise enough to serve as control input for the proposed aerial interception system.

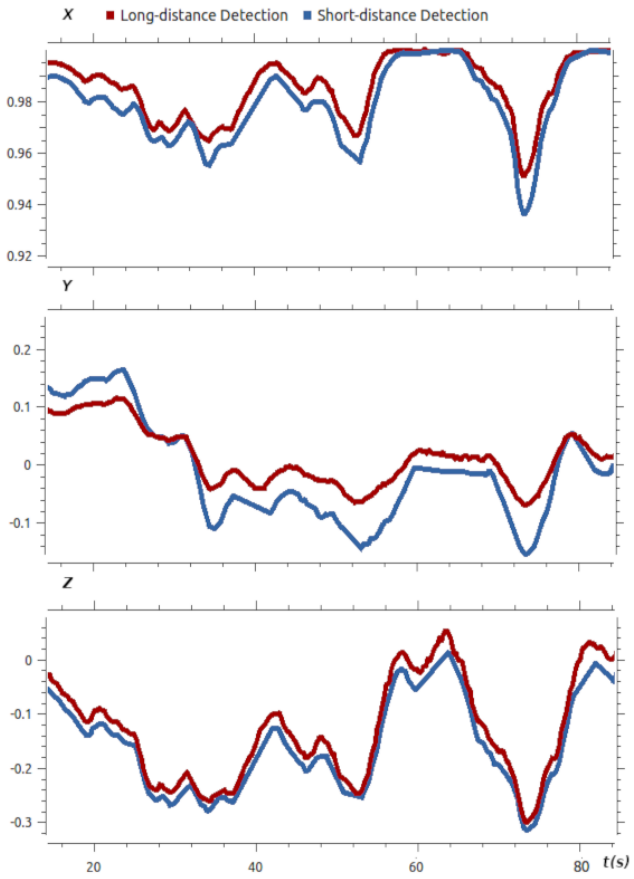


Fig. 12: Estimated director vectors for long- and short-range detectors.

## VI. CONCLUSIONS

This work demonstrates that an autonomous aircraft capable of detecting and capturing potentially malicious UAVs is achievable. The importance of properly choosing the devices (sensor suite, onboard computer...), in addition to the suitable control and detection strategies is crucial to succeed in this kind of complex tasks. Moreover, including Deep Learning algorithms to complex problems keeps demonstrating the power of this emerging technology by outperforming classic approaches. The proper selection of a synergistic detection-control couple is one of the key elements that have led to promising results.

Regarding the neural network, more work can be done in gathering more generalized images for the intruder dataset. In this way, more training strategies could be tested and parameters could be tuned to improve the network. This can be achieved by performing more experiments where to record the data, or using simulation environments. In addition, other interesting research line is to model the detection uncertainty of the neural network in order to include such information as input in the tracking algorithm.

## ACKNOWLEDGMENT

This work has been partially supported by the ASSISTANCE project funded by the EU H2020 programme under grant agreement 832576, and DURABLE project funded by the Interreg Atlantic Area Programme through the European Regional Development Fund (ERDF). Thanks also goes to Manuel Garcia, Ines M. Lara and Manuel Barbero for their support in the experiments.

## REFERENCES

[1] Georgia Lykou, Dimitrios Moustakas, and Dimitris Gritzalis. Defending airports from uas: A survey on cyber-attacks and counter-drone sensing technologies. *Sensors*, 20(12):3537, 2020.

[2] R. Mitch, Ryan C. Dougherty, M. Psiaki, S. Powell, B. O'Hanlon, J. Bhatti, and T. Humphreys. Signal characteristics of civil gps jammers. 2011.

[3] David Hambling. drone crash due to gps interference in u.k. raises safety questions.

[4] Boeing. Boeing's compact laser weapons system tracks and disables uavs.

[5] Michael Jian, Zhenzhong Lu, and Victor C. Chen. Drone detection and tracking based on phase-interferometric doppler radar. In *2018 IEEE Radar Conference (RadarConf18)*, pages 1146–1149, 2018.

[6] Gong Jiangkun, Yan Jun, Li Deren, Kong Deyong, and Hu Huiping. Interference of radar detection of drones by birds. *Progress In Electromagnetics Research*, 81:1–11, 2019.

[7] Sedat Dogru and Lino Marques. Pursuing drones with drones using millimeter wave radar. *IEEE Robotics and Automation Letters*, 5(3):4156–4163, 2020.

[8] Andrea Bernardini, Federica Mangiatordi, Emiliano Pallotti, and Licia Capodiferro. Drone detection by acoustic signature identification. *Electronic Imaging*, 2017(10):60–64, 2017.

[9] Joël Busset, Florian Perrodin, Peter Wellig, Beat Ott, Kurt Heutschi, Torben Rühl, and Thomas Nussbaumer. Detection and tracking of drones using advanced acoustic cameras. In *Unmanned/Unattended Sensors and Sensor Networks XI; and Advanced Free-Space Optical Communication Techniques and Applications*, volume 9647, page 96470F. International Society for Optics and Photonics, 2015.

[10] Jorge Mariscal-Harana, Víctor Alarcón, Fidel González, Juan José Calvente, Francisco Javier Pérez-Grau, Antidio Viguria, and Aníbal Ollero. Audio-based aircraft detection system for safe rpas bvlos operations. *Electronics*, 9(12):2076, 2020.

[11] Marcus Hammer, Marcus Hebel, Martin Laurenzis, and Michael Arens. Lidar-based detection and tracking of small uavs. In *Emerging Imaging and Sensing Technologies for Security and Defence III; and Unmanned Sensors, Systems, and Countermeasures*, volume 10799, page 107990S. International Society for Optics and Photonics, 2018.

[12] Anton Mitrokhin, Cornelia Fermüller, Chethan Parameshwara, and Yiannis Aloimonos. Event-based moving object detection and tracking. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–9, 2018.

[13] Amy R Wagoner, Daniel K Schrader, and Eric T Matson. Towards a vision-based targeting system for counter unmanned aerial systems (cuas). In *2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pages 237–242. IEEE, 2017.

[14] Manuel García, Rafael Caballero, Fidel González, Antidio Viguria, and Aníbal Ollero. Autonomous drone with ability to track and capture an aerial target. In *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 32–40. IEEE, 2020.

[15] Dongkyu Lee, Woong Gyu La, and Hwangnam Kim. Drone detection and identification system using artificial intelligence. In *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1131–1133, 2018.

[16] Widodo Budiharto, Alexander A S Gunawan, Jarot S. Suroso, Andry Chowanda, Aurello Patrik, and Gaudi Utama. Fast object detection for quadcopter drone using deep learning. In *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*, pages 192–195, 2018.

[17] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018.