drones

MDPI

Letter

5G-Enabled Security Scenarios for Unmanned Aircraft: Experimentation in Urban Environment

Erina Ferro [1], Claudio Gennaro [1,*], Alessandro Nordio [2], Fabio Paonessa [2], Claudio Vairo [1], Giuseppe Virone [2], Arturo Argentieri [3], Andrea Berton [4] and Andrea Bragagnini [5]

1 Institute of Information Science and Technologies of CNR (CNR-ISTI), 56124 Pisa, Italy; erina.ferro@isti.cnr.it (E.F.); claudio.vairo@isti.cnr.it (C.V.)
2 Institute of Electronics, Computer and Telecommunication Engineering of CNR (CNR-IEIIT), 10129 Torino, Italy; alessandro.nordio@ieiit.cnr.it (A.N.); fabio.paonessa@ieiit.cnr.it (F.P.); giuseppe.virone@ieiit.cnr.it (G.V.)
3 Institute of Applied Sciences and Intelligent Systems of CNR (CNR-ISASI), 73100 Lecce, Italy; arturo.argentieri@cnr.it
4 Institute of Clinical Physiology of CNR (CNR- IFC), 56124 Pisa, Italy; bertonandrea@ifc.cnr.it
5 TIM Services Innovation, 10148 Torino, Italy; andrea.bragagnini@telecomitalia.it
* Correspondence: claudio.gennaro@isti.cnr.it; Tel.: +39-050-3153077

Received: 1 May 2020; Accepted: 9 June 2020; Published: 12 June 2020

Abstract: The telecommunication industry has seen rapid growth in the last few decades. This trend has been fostered by the diffusion of wireless communication technologies. In the city of Matera, Italy (European capital of culture 2019), two applications of 5G for public security have been tested by using an aerial drone: the recognition of objects and people in a crowded city and the detection of radio-frequency jammers. This article describes the experiments and the results obtained.

Keywords: drone; security; jammer

1. Introduction

Starting with 1G, born in the '80s, until the current 5G, each mobile network generation was characterized by peculiar wireless technologies, data rates, modulation techniques, capacities, and features. The performance goals envisioned for the upcoming 5G include high data rate (more than 100 Mb/s in uplink and 1 Gb/s in down-link), low latency (<1 ms), energy savings, cost reduction, increased system capacity and massive device connectivity, thus enabling the development of the Internet of Things (IoT) concept. Different from 4G, where carrier frequencies range from 800 MHz to 2.6 GHz, 5G frequencies are divided into two groups, i.e., below 6 GHz and in the millimeter-wave spectrum. Several other types of wireless technologies share the radio range below 6 GHz, potentially causing interference; to partially attenuate this problem, 5G uses the OFDM (orthogonal frequency-division multiplexing) modulation [1,2].

All these new features of 5G allow for the development of innovative digital services, such as virtual and augmented reality, streaming and multiplayer working on smart cellular, faster real-time communications between wireless devices, and real-time analysis of a huge amount of data. At present, Italy is gradually introducing 5G coverage and a significant market share is expected by 2020.

Public security has become a relevant aspect, especially in large urban areas. The possibility to act promptly in case of dangers or alarms can be of fundamental importance in determining the favorable outcome of the interventions. For example, identifying and tracking an individual or a suspicious vehicle that moves in an urban context may require a significant deployment of forces, with costs that sometimes can make the interventions ineffective.

A threat to public security also comes from jammers [3], electronic devices capable of disturbing and inhibiting the operation of common communication technologies, including mobile phones, GPS, and radios. Because of their potential harmful effects for criminal purposes and the widespread use of potentially targeted wireless devices, jammers are illegal.

To cope with these problems, we tested applications of 5G for public security in the city of Matera (Italy), European cultural capital in 2019. The experiments involved aerial drones and exploited the 5G mobile network provided by the Bari-Matera 5G project consortium led by TIM, Fastweb (two Italian mobile telecommunication operators), and Huawei. Through these experiments, we were able to show that the 5G wireless communication technology can be successfully integrated with the drone technology, in order to create a synergy for the development of innovative and low-cost public security applications. In particular, we considered two scenarios related to security in densely populated areas: (i) the recognition of objects and persons from video recording, and (ii) the detection of radio-frequency jammers.

For both scenarios, we used a DJI Matrice 600 drone owing to its high lifting capabilities (see Figure 1). The 5G connectivity was obtained through a TIM 5G Gateway, permanently mounted on the drone. The real challenge of the experiment was to test the two scenarios in the worst flying conditions of the drone (strong wind, high altitude of the drone, distance from the target) and to verify the accuracy of the algorithms used and the measurements done.

Figure 1. 5G aerial drone public security application in Matera, Italy.

This article does not discuss issues about the privacy of individuals related to facial detection and the European laws of the General Data Protection Regulation (GDPR). For more details about this topic, the reader is referred to [4].

The paper is organized as follows. Section 2 contains references to works related to the use of drones for both the security scenarios we used. In Section 3 the security scenario using a 2D camera mounted on the drone is described, while the jammer detection scenario is depicted in Section 4. Sections 5 and 6 present the results of the experimentation in both the scenarios, respectively. Section 7 presents our ideas for future work, while conclusions are reported in Section 8.

2. Related Work

The problem of detecting, recognizing, and tracking people from a drone has become of paramount importance in many video surveillance applications.

One of the first works that investigated the issues and the challenges of performing face recognition with drone-captured images was [5]. The work in [6] introduced the case of UAV-based crowd surveillance applying facial recognition tools. To perform face recognition, the simple Local Binary Pattern Histogram method from the Open Source Computer Vision (OpenCV) was used. In [7], Hsu et al. investigate how altitudes, distances, and angles of depression influence the performance of face detection and recognition by drones. In the same article, the authors present a dataset for the evaluation of facial recognition algorithms from drone images, called DroneFace.

In [8], authors present a system able to track a specific person in the case of multi-person interference. Such a system automatically takes a photo of the target's frontal face, without the use of facial recognition. After boxing out the target on the image received by the ground station, the drone (a Parrot Bebop2) flies around the target person, mounting onboard an RGB camera, which automatically keeps shooting in the horizontal direction with a resolution of 640 $\times$ 480 pixels. In their experiment, the authors of [8] set the flying height at about 2 m, while the distance between the target person and the drone ranged from 4 to 5 m. Such conditions are very different from ours, as will be explained in the following.

In [9], the authors used a camera mounted on board of a drone while the algorithms for face recognition running on a computer on the ground. After scanning the faces and comparing them to the provided database, the output obtained has the individual's face highlighted in a square with a color-coding sequence (white, green, or red). However, no information about the height of the drone is given, neither about the weather conditions. Moreover, no detail is given on the algorithm used for facial recognition.

Another face dataset collected by drone is presented in [10]. The dataset includes 58 different identities. For each identity there are four high-resolution images in four different poses, and a lot of low-resolution face images (in the following referred to as probe images) cropped from the videos acquired by the drone.

The works [11,12] both investigate the challenges and perform experiments for executing the person re-identification from aerial images.

In [13], the authors provide a dataset for human action detection captured from UAVs flying at different altitudes and at different angles. It consists of 43 min-long fully-annotated sequences with 12 action classes such as walking or sitting. Another dataset for human action recognition is proposed in [14]. The authors provide a dataset recorded in an outdoor setting by a free-flying drone. The dataset consists of 240 high-definition video clips for a total of 66,919 frames and it captures 13 dynamic human actions. The videos contained in the dataset were recorded from low-altitude and at low speed and the corresponding frames are at high resolution.

The authors in [15,16] provide a drone captured benchmark for vehicles, composed of 10 h of raw videos from which about 80,000 representative frames are selected. Each frame is fully annotated with bounding boxes as well as up to 14 kinds of attributes, like weather conditions, flying altitude, camera view, vehicle category, and occlusion. The benchmark can be used for three computer vision tasks: object detection, single-object tracking, and multiple object tracking.

Very recently, in [17] the authors present a review of the challenges of applying the vision algorithm to drone-based images and they survey the currently available drone captured datasets.

They also provide a drone captured dataset themselves, VisDrone. The dataset consists of 263 video clips with 179,264 frames and 10,209 static images. It includes image object detection, video object detection, single-object tracking, and multi-object tracking.

In the case of disaster events, search operations for survivors can easily be carried out by autonomous drones capable of searching disaster areas that are difficult to reach by land-based vetches. This topic was studied by Al-Naji et al. [18], who developed techniques for detecting the presence of life signs from humans lying in many poses on the ground by using standard cameras and consumer drones.

Conceptually, the simplest type of jammer consists of a device that transmits strong radio signals within the frequency band of a particular service, to decrease the Signal-to-Noise Ratio (SNR) of the target device, thus impeding its operation. More advanced jammers can implement complex modulation algorithms in order to enhance the interfering effect with the target service [19,20].

The problem of jammer detection and localization is of great interest not only for public security but also for the radio authorities and network operators [21]. At present, jammers are detected by either permanent stations or mobile units equipped with RF instrumentation [22]. In both cases, proper radio-frequency equipment consists of different instruments. A spectrum analyzer is required in order to monitor the radio frequency spectrum and detect the presence of evident interfering signals. However, scalar spectrum analyzers hardly provide information on the jammer typology, for which a proper receiver is instead required, i.e., a device capable of demodulating the received signal. Thanks to the advances in digital electronics and signal processing, Software Defined Radios (SDRs) have nowadays become powerful and miniaturized. These instruments represent versatile general-purpose hardware through which the operator can virtually implement any demodulation protocol [23].

Different jammer detection algorithms have been surveyed in [24] using an SDR platform connected to a PC. In [25], Abdessamad et al. implemented a direction-finding algorithm conceived for spectrum monitoring against interference in a mobile network. The proposed hardware used a set of four ground-based SDR platforms. In [26], Jagannath et al. developed instead a portable jammer detection and localization device. However, the localization capability was achieved by a central ground-based processing unit that gathered the data from several devices exploiting the multi-node information. The possibility to carry out jammer detection and localization with the aid of aerial drones would lead to significant improvements in terms of mobility, system versatility, and portability. Koohifar et al. in [27] described a cooperative algorithm for UAV swarms to localize RF transmitters by using a multitude of omnidirectional sensors.

In this work, we integrated on a UAV the RF instrumentation required to perform jammer detection with localization. The realized proof-of-concept exploits the 5G high-speed and low-latency data connection to allow a remote operator to both examine and elaborate the data and to interact in real-time with the system.

3. The Video Security Scenario

Intelligent cameras have recently seen a large diffusion and represent a low-cost solution for improving public security in many scenarios. Moreover, they are light enough to be lifted by a drone. The use of drones equipped with intelligent cameras for face/object recognition has already been reported in the literature. However, this is the first time that such a solution is tested in a 5G environment. The experiments have shown the validity of the approach despite the very harsh conditions in which our drone operated, due to the strong winds blowing at a height of 40 m from the ground. We flew at 40 m because with the camera used in our setup we saw that with a height of more than 50 m the bounding boxes of the detected faces are too small for having a good face recognition.

In our applications, the visual information recorded by the camera (photo and video) is first transmitted to the ground by using the 5G wireless network and then is processed by using Artificial Intelligence technologies based on deep learning. We point out that recognizing in real-time a face

of a person or a specific object in motion, is now within the reach of a home computer, thanks to the processing power of graphics cards equipped with the most recent Graphics Processing Units.

The process of detecting a person's identity through his/her face is often referred to as facial recognition. Typically, facial recognition in images is performed through two distinct phases. In the first phase, usually referred to as face detection, the system detects a face appearing within an image. Specifically, when an original image is processed, a set of new images is produced as a result. Each of these contains the face of a person appearing in the original image. The second phase is devoted to the recognition of the identities of the people to whom the identified faced belong. This is usually the most complex procedure and requires both a database, with which the faces detected have to be compared, and an algorithm for the face comparison. In the literature, this process is often referred to as face verification and has been the subject of great developments in recent years thanks to deep learning technologies. Although facial recognition has been intensively studied for many years, only recently it was applied to images recorded by drones. Indeed, the interest in computer vision applications applied to images taken by drones is constantly growing, due to their low-cost and the enormous technological progress they have had in recent years.

The equipment we used in Matera for the 2D Camera Recognition includes:

- a 2D camera mounted on the drone; it acquires a video stream and sends it in real-time to a local server;
- a server able to analyze, through deep learning techniques, the frames of the video and to perform face detection and identity recognition by comparing the faces detected with a database of "well-known" people. In case of a match, the server sends a report to the law enforcement agency.

In the server, the face detection algorithm is executed to localize and crop the faces from the captured image. Then, each detected face is compared to the faces recorded in the database in order to identify possible matches. These tasks are carried out by two distinct deep neural networks: the Single Shot Detector (SSD [28]) ResNet-10 model, provided by OpenCV that performs face detection, and the VGG2 provided by the VGG group (http://www.robots.ox.ac.uk/~vgg).

Since for each face detected in a frame we have to proceed to the recognition phase, the total processing time increases linearly with the number of faces detected. In order to satisfy the real-time constraints, we must limit the number of faces to be recognized, with some criteria. Since the execution time for each facial recognition (using the latest generation GPUs) is about 8ms, and imposing a processing constraint of five frames per second, we obtained a maximum of five faces to be recognized for each frame.

To perform facial recognition and to determine whether a person is among those sought, we find the closest face in the database of wanted persons, on the base of similarities with features extracted by using a deep neural network. To improve the effectiveness of facial recognition, several facial images (taken in different circumstances or poses) are associated with each identity to be recognized. To this end, the feature vectors extracted from each image are aggregated by computing the mean feature vector so as to provide a unique template for the identity of the person. This template is used during the facial recognition phase and the template closest to the query feature reveals the recognized identity, using cosine similarity (i.e., the normalized dot product) as a measure of similarity between templates and the query feature. We use the similarity value of the person who has the model closest to the query (i.e., the one who is thought to be the recognized person) as a measure of confidence for the recognition itself. Confidence ranges from 0 to 1 because similarity ranges from 0 to 1 (0 minimum similarity, i.e., minimum confidence, being 1 the maximum similarity, i.e., maximum confidence).

During the recognition, we set a confidence threshold to decide the tolerance level on the recognition error. When this threshold is higher than a preset value, recognition is considered reliable and the identity can be revealed to the operator, otherwise, the face is considered unknown. The threshold can be chosen depending on the scenario: by using smaller thresholds, we accept more false positives (i.e., an unknown person is recognized as one of those recorded in the database) while

by using higher thresholds we accept more false negatives (i.e., a person belonging to the database has escaped control).

4. The Jammer Detection Scenario

The presence of a malicious device in Matera was simulated by using a legal radio device working in the 5.8 GHz ISM band (5.725–5.875 GHz). The chosen device was a CE-certified ImmersionRC 25 mW A/V transmitter, configured to transmit a continuous-wave signal at 5.74 GHz and 14 dBm of output power. A true jammer (absolutely illegal) would produce a higher power level and/or a larger band occupation, thus being more easily detectable.

The following instruments were mounted onboard the drone (see Figure 2):

- A log-periodic antenna (Aaronia HyperLog 7060 X), which provides both directionality and wide operating bandwidth (from 700 MHz to 6 GHz).
- Filters to attenuate both the 5G signal (3.7–3.8 GHz) and the drone telemetry (2.4 GHz).
- A vector spectrum analyzer (Triarchy VSA6G2A), which works as a tunable receiver with 6 GHz maximum frequency and 1.35 MHz passband in real-time mode (FFT-mode). This device also performs a set of different analog and digital demodulations (e.g., FM, PSK, QPSK).
- A single-board computer (SBC) (LattePanda), running the application software to interface the analyzer to the 5G gateway.
- The 5G gateway, to interface the system to the 5G network (permanently mounted in all the experiments).

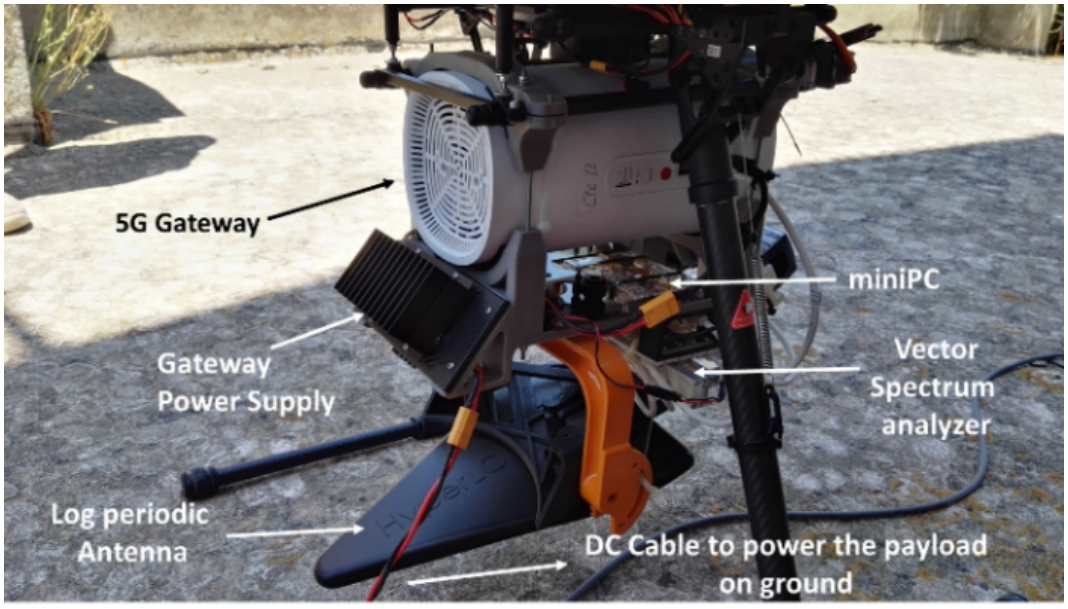

Figure 2. Drone payload with 5G Gateway and jammer detection hardware.

Thanks to the 5G wireless connection to the ground station, the operator interacted in real-time with the instruments onboard the drone. The high data transfer of 5G was used to send both spectrum and I/Q baseband data to ground for further elaboration.

5. Video Security Experimentation Results

In this scenario, the challenge was to recognize people and objects by maintaining the drone in the worst flight conditions, at an altitude of 40 m over the roof of a church (see Figures 3 and 4). The drone was required to never fly directly over the people while facing strong wind blowing at 35–38 km/h

and using the zoom in order to capture better images of the subjects. In these conditions, the drone was swinging considerably.

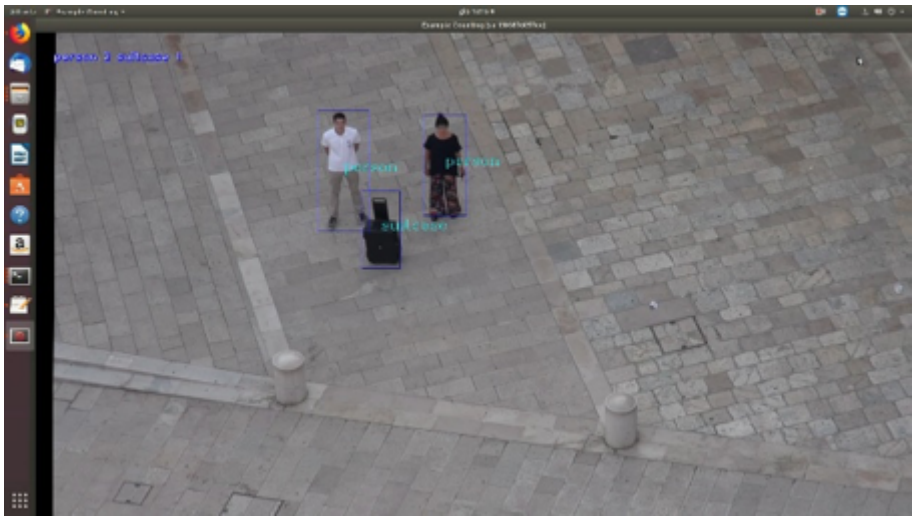

Figure 3. Persons and objects detection.

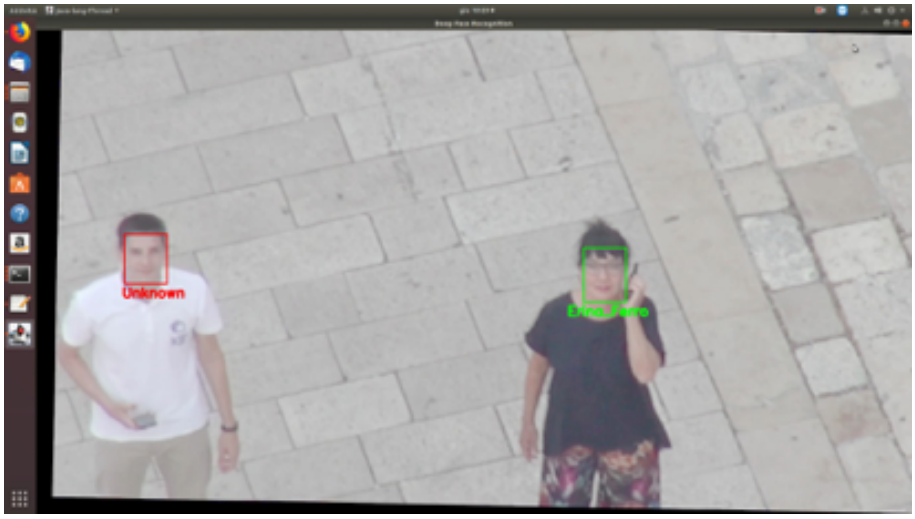

Figure 4. Facial recognition from zoomed in shots.

The experiments were conducted by using a database of known people composed of 440 images of faces belonging to 44 different people working in our research institute. The queries were created by using the faces detected in a video (see Figure 4) recorded by our drone in Matera. Figure 5 shows the result of this experimentation. In particular, we show two metrics:

- TP (True Positive): the fraction of known people, correctly recognized by the system among the whole set of queries.
- FP (False Positive): the fraction of unknown people incorrectly identified as belonging to the dataset.

For the particular experiment in Figure 4, we used a video of length 180 s, in which two subjects appear, one belonging to the dataset (who, therefore, should be recognized by the system) and one extraneous to the dataset (i.e., an unknown). The system detected 135 faces of the known person and 135 faces of the unknown person. As previously mentioned, by varying the value of a specific threshold, it is possible to tune the degree of selectivity of the system. However, there is a trade-off: if we want to reduce the number of false positives, we must accept a lower number of true positives.

For example, for an acceptable result in our scenario we set the threshold to 0.37, thus obtaining a number of true positives of about 10% against a number of false positives of 1%. The reason for this result is that the images were very blurry due to the use of the zoom of the camera and to the strong oscillations of the drone induced by the wind.

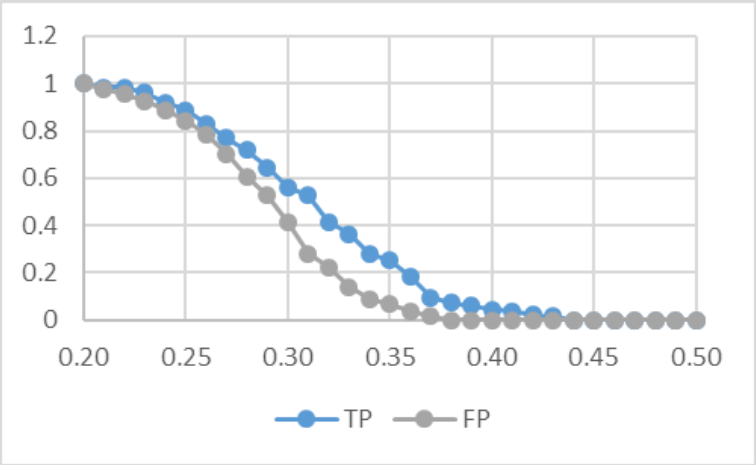

Figure 5. Results of the face recognition experiment.

Experiments on DroneSURF Dataset

We also performed an experiment of face recognition on the DroneSURF dataset [10]. The dataset is composed of 200 videos including 58 different subjects, from which over 786K face annotations have been extracted. It is very challenging for face recognition due to the effects of motion, variations in pose, illumination, background, altitude, and different resolutions of the detected faces.

The dataset is composed of two parts: the high-resolution images (12Mpx) used as ground truth, and the low-resolution images used for test (in the following we refer to these images as *probe*). The high-resolution images are four for each identity in different poses. The probe images are the crops of the detected face of the videos acquired by the drones. The probe images are split in two subsets corresponding to two different acquisition scenarios: *Active* and *Passive* video surveillance, of respectively 333,047 and 379,841 images. The difference between the two scenarios is that in the Active one the drone is actively following the subject, while in the Passive one the drone is monitoring an area without focusing in particular on a specific subject.

In the experiment, we used each probe image as a query towards the set of high-resolution images in order to find the most similar face and, therefore, the corresponding person. These experiments can be considered as a case where the weather conditions are optimal, i.e., in the absence of wind and drone oscillations. To do so, we extracted a visual feature from every image in the dataset (both high and low-resolution images) by using the SeNET architecture [29] and, in particular, the model pre-trained on faces provided in [30]. Then, we computed the Euclidean distance between the feature of the query probe image and all the features of the high-resolution images.

Since in the paper [10] it was not specified how the face detection on the high-resolution images has been executed, we used three different face detectors to detect and crop the bounding boxes of the detected faces: the dlib implementation of the classic Histogram of Oriented Gradients (HOG) face detector (called dlib in the experiment) [31], the MTCNN face detector (called MTCNN in the experiment) [32] and the OpenCV implementation of the SSD framework (Single Shot MultiBox Detector) with a reduced ResNet-10 model (called OpenCV-DNN in the experiment) [33]. Table 1 reports the results of this experiment. As expected, the dataset is very challenging for face recognition, and with all the three face detectors the results are quite similar. In the Active scenario, the accuracy ranges from 22.88% with dlib detector to 24.25% with MTCNN detector. In the Passive scenario, which is even more challenging, the accuracy ranges from 1.7% of the dlib detector to 2.61% of the

OpenCV-DNN detector. We can interpret the accuracy value in the best case (MTCNN) as the true positives for this scenario, i.e., 24.25%, and if we also evaluate the false positives we get a value of 1.3%, which represents a clear improvement compared to the case of Matera with more than double the number of true positives (corresponding to a similar value of false positives).

Table 1. Face recognition accuracy (in %) between probe images and high-resolution images of DroneSURF.

Scenario	SENet		
	dlib	**MTCNN**	**OpenCV-DNN**
Active	22.88	24.25	24.07
Passive	1.7	2.43	2.61

6. Jammer Detection

The spectrogram of a short acquisition (120 s) is shown in Figure 6. It represents the acquired spectra as a function of time. Stronger signals appear in yellow/green. The intermittent jammer contribution is clearly visible at 5.74 GHz. The narrow dots in the spectrogram are, instead, produced by common wireless systems operating in the 5.8 GHz ISM band, e.g., Wi-Fi devices, and hardly exceed the level of −50 dBm.

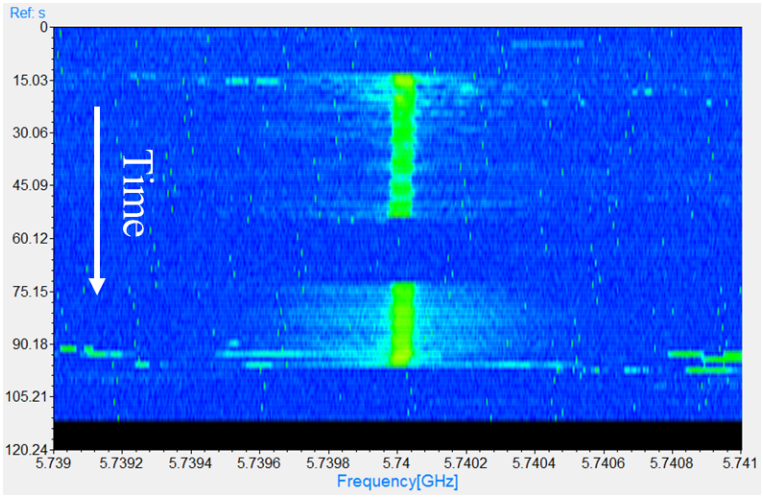

Figure 6. Spectrogram of a 120-seconds acquisition with the jammer detection hardware.

Besides the spectral analysis, the system was used to estimate the direction of arrival of the jammer signal. Thanks to the direction-dependent response of the onboard antenna, a rotation of the drone around its vertical axis produces a significant variation on the detected signal. Figure 7 shows two spectra acquired in the presence of the jammer with the drone facing the malicious device (yellow curve) and the opposite direction (gray curve). The jamming signal at 5.74 GHz is visible in both cases, with a level of about −42 dBm and −67 dBm, respectively. An increase of about 25 dB in the received signal has been observed when the drone/antenna tip was oriented toward the direction of the jammer.

Figure 8 shows the measured signal level of 16 spectra acquired during an azimuthal rotation of the UAV from −50 degs to 250 degs. Even if localized oscillations produced by occasional wind gusts are present, the maximum and minimum signal levels can be observed at about 0 degs and 180 degs, respectively.

As far as the system sensitivity is concerned, the signal level is about 14 dB above the noise floor when the UAV faces the opposite direction with respect to the jammer (see the gray curve of Figure 7). Such conditions represent the worst-case for the direction of arrival estimation. Considering that the

used jammer had an output power of 14 dBm, the same measurement setup would be able to localize a device with an output power as low as 0 dBm in the same test-conditions, i.e., about 50 m of distance between the UAV and the jammer. Equivalently, the same device could be detected at about 5 times the distance, i.e., about 250 m. It should also be pointed out that a true jammer device would easily have a greater output power than the legal A/V transmitter used. For example, a 30-dBm transmitter would be localizable up to a distance of 1500 m.

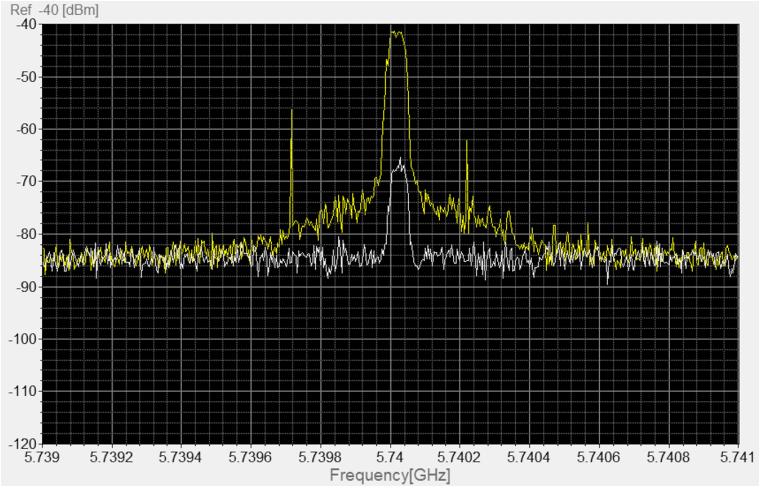

Figure 7. Acquired spectra in presence of the jammer, with the drone facing the jammer (yellow curve) and the opposite direction (gray curve).

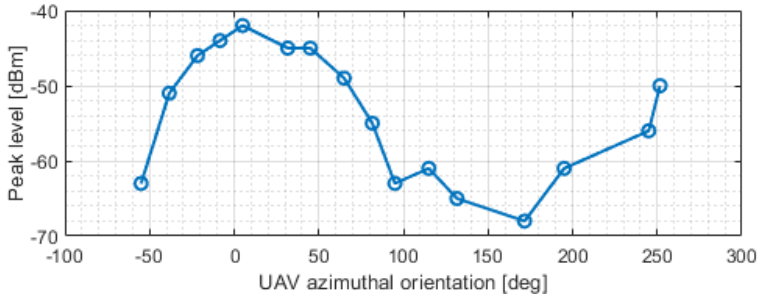

Figure 8. Measured jammer level as a function of the UAV azimuthal orientation with respect to the jammer direction.

7. Future Work

In Matera, in addition to the tests here described for jammer detection and face/object recognition, we also equipped the drone with an omni-directional camera (3D-camera) capable of sending in real-time a 360° video to the ground station through the 5G connection. The 360° video was viewed either through an ordinary monitor, by using commands to rotate the view, or through a modern VR headset, such as a Google Cardboard or Samsung Gear VR. The advantage of such an equipment lies in its ability to self-stabilize thanks to the presence of gyro sensors in the omnidirectional camera, thus eliminating the problems of swaying and vibration of the drone due, for example, to the wind. For these reasons, we are working to utilize the stream of the 3D camera to recognize in real-time faces and objects when the drone is very far from the subject. Furthermore, it is also under study the tracking of the identified object/person in such a way that the artificial intelligence algorithms directly drive the drone in its movements. Another experiment in progress in Matera is the use of the drone for detecting obstacles in an area where an autonomous machine must operate. The drone, still equipped with the 3D-camera and a Raspberry for the elaboration of the images, is used to fly over a vast agricultural field, covered by 5G technology, in order to detect any obstacles to be communicated (with the relevant geo-localization coordinates) to the autonomous agricultural machine in order to

update its prescription map in such a way to avoid the obstacles. In the sector of smart agriculture, another application of the drone we are investigating is the study of the different water requests in large cultivated fields according to the geo-localization of the various sectors of the area, their sun exposure, the weather conditions, the type of cultivation, the presence or not of trees, and so on.

As far as the jammer detection is concerned, future developments will take advantage of the basic demodulation capabilities of the spectrum analyzer and real-time availability of raw I/Q data to the ground station. In this way, a ground computer can perform more complex and demanding demodulation to detect the jammer, for example exploiting the potentialities of neural networks [28,34].

8. Conclusions

In this paper, we have shown the results of our experiments for public security that we tested in the city of Matera, Italy, which was the European capital of culture 2019. In particular, we used a drone to experiment with two security scenarios using the TIM's 5G connection: the recognition of objects and people in a crowded city and the detection of radio-frequency jammers. By equipping the drone with the appropriate instruments, we identified examples of weak radio-frequency interfering signals; we detected objects and recognized people, by means of Deep Learning techniques, from an altitude of 40 m over the ground, flying in adverse weather conditions. Both experiments were successful, despite the strong wind at 35–38 km/h and the drone oscillations. In order to further validate our approach, we also tested out solution for face recognition on DroneSURF, a publicly available face dataset with images captured by a flying drone. During the demo in Matera, at the presence of people from the Ministry of the Economic Development, about 10 persons required to test the reliability of the recognition algorithm with the use of the drone. They were recognized as "persons" but, correctly, their faces resulted "unknown", since they were not in any algorithm training database. For obvious privacy reasons, no photos of their faces were taken. However, the experiment was a complete satisfaction for the ministry staff.

Author Contributions: Conceptualization, E.F., C.G., and G.V.; methodology, E.F., C.G., F.P., G.V., and C.V.; software, A.A., A.N, A.B. (Andrea Berton), F.P., and C.V.; validation, A.N., E.F., C.G., G.V, and C.V.; resources, A.A., and A.B. (Andrea Berton); writing—original draft preparation, A.N., E.F., C.G., F.P., G.V., and C.V.; writing—review and editing, E.F., C.G., F.P., G.V., and C.V.; visualization, A.A., A.B. (Andrea Berton), E.F., C.G., F.P., G.V., and C.V; supervision, E.F.; project administration, E.F.; funding acquisition, A.B. (Andrea Bragagnini). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by TIM Italy S.p.A. in the framework of the contract protocol number 0085423/2018 del 13/12/2018.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Osseiran, A.; Boccardi, F.; Braun, V.; Kusume, K.; Marsch, P.; Maternia, M.; Queseth, O.; Schellmann, M.; Schotten, H.; Taoka, H.; et al. Scenarios for 5G mobile and wireless communications: The vision of the METIS project. *IEEE Commun. Mag.* **2014**, *52*, 26–35. [CrossRef]
2. Andrews, J.G.; Buzzi, S.; Choi, W.; Hanly, S.V.; Lozano, A.; Soong, A.C.; Zhang, J.C. What will 5G be? *IEEE J. Sel. Areas Commun.* **2014**, *32*, 1065–1082. [CrossRef]
3. GPS, Wi-Fi, and Cell Phone Jammers Frequently Asked Questions (FAQs). Available online: https://transition.fcc.gov/eb/jammerenforcement/jamfaq.pdf (accessed on 16 October 2019).
4. Barnoviciu, E.; Ghenescu, V.; Carata, S.V.; Ghenescu, M.; Mihaescu, R.; Chindea, M. GDPR Compliance in Video Surveillance and Video Processing Application. In Proceedings of the 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), Timisoara, Romania, 10–12 October 2019; pp. 1–6.
5. Hsu, H.J.; Chen, K.T. Face recognition on drones: Issues and limitations. In Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use, Florence, Italy, 18-22 May 2015; pp. 39–44.

6. Motlagh, N.H.; Bagaa, M.; Taleb, T. UAV-based IoT platform: A crowd surveillance use case. *IEEE Commun. Mag.* **2017**, *55*, 128–134. [CrossRef]
7. Hsu, H.J.; Chen, K.T. DroneFace: An open dataset for drone research. In Proceedings of the 8th ACM on Multimedia Systems Conference, Taipei, Taiwan, 20–23 June 2017; pp. 187–192.
8. Shen, Q.; Jiang, L.; Xiong, H. Person Tracking and Frontal Face Capture with UAV. In Proceedings of the 2018 IEEE 18th International Conference on Communication Technology (ICCT), Chongqing, China, 8–11 October 2018; pp. 1412–1416.
9. NS, S.R.; Varghese, J.T.; Pandya, F. Unmanned Aerial Vehicle for Human Tracking Using Face Recognition System. In Proceedings of the 2019 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, UAE, 26 March–10 April 2019; pp. 1–5.
10. Kalra, I.; Singh, M.; Nagpal, S.; Singh, R.; Vatsa, M.; Sujit, P. Dronesurf: Benchmark dataset for drone-based face recognition. In Proceedings of the 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), Lille, France, 14–18 May 2019; pp. 1–7.
11. Layne, R.; Hospedales, T.M.; Gong, S. Investigating open-world person re-identification using a drone. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 225–240.
12. Bindemann, M.; Fysh, M.C.; Sage, S.S.; Douglas, K.; Tummon, H.M. Person identification from aerial footage by a remote-controlled drone. *Sci. Rep.* **2017**, *7*, 1–10. [CrossRef] [PubMed]
13. Barekatain, M.; Martí, M.; Shih, H.F.; Murray, S.; Nakayama, K.; Matsuo, Y.; Prendinger, H. Okutama-action: An aerial view video dataset for concurrent human action detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 28–35.
14. Perera, A.G.; Law, Y.W.; Chahl, J. Drone-Action: An Outdoor Recorded Drone Video Dataset for Action Recognition. *Drones* **2019**, *3*, 82. [CrossRef]
15. Du, D.; Qi, Y.; Yu, H.; Yang, Y.; Duan, K.; Li, G.; Zhang, W.; Huang, Q.; Tian, Q. The unmanned aerial vehicle benchmark: Object detection and tracking. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 370–386.
16. Yu, H.; Li, G.; Zhang, W.; Huang, Q.; Du, D.; Tian, Q.; Sebe, N. The Unmanned Aerial Vehicle Benchmark: Object Detection, Tracking and Baseline. *Int. J. Comput. Vis.* **2019**, 1–19. [CrossRef]
17. Zhu, P.; Wen, L.; Du, D.; Bian, X.; Hu, Q.; Ling, H. Vision Meets Drones: Past, Present and Future. *arXiv* **2020**, arXiv:2001.06303
18. Al-Naji, A.; Perera, A.G.; Mohammed, S.L.; Chahl, J. Life signs detector using a drone in disaster zones. *Remote Sens.* **2019**, *11*, 2441. [CrossRef]
19. Mpitziopoulos, A.; Gavalas, D.; Konstantopoulos, C.; Pantziou, G. A survey on jamming attacks and countermeasures in WSNs. *IEEE Commun. Surv. Tutorials* **2009**, *11*, 42–56. [CrossRef]
20. Grover, K.; Lim, A.; Yang, Q. Jamming and anti-jamming techniques in wireless networks: A survey. *Int. J. Hoc Ubiquitous Comput.* **2014**, *17*, 197–215. [CrossRef]
21. RF Interference Hunting—Why Remote Spectrum Monitoring is Becoming a Must. Available online: https://anritsu.typepad.com/interferencehunting/2016/10/rf-interference-hunting-why-remote-spectrum-monitoring-is-becoming-a-must.html (accessed on 18 December 2019).
22. Tools for Network Operators to Protect Spectrum Investment from Costly Interference. Available online: https://anritsu.typepad.com/interferencehunting/2016/12/tools-for-network-operators-to-protect-spectrum-investment-from-costly-interference.html (accessed on 18 December 2019).
23. Buracchini, E. The software radio concept. *IEEE Commun. Mag.* **2000**, *38*, 138–143. [CrossRef]
24. Bhojani, R.; Joshi, D.R. An Integrated Approach for Jammer Detection using Software Defined Radio. *Procedia Comput. Sci.* **2016**, *79*, 809–816. [CrossRef]
25. Abdessamad, W.; Nasser, Y.; Artail, H.; Chazbek, S.; Fakher, G.; Bazzi, O. An SDR platform using direction finding and statistical analysis for the detection of interferers. In Proceedings of the 2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Lisbon, Portugal, 18–20 October 2016; pp. 43–48.
26. Jagannat, A.; Jagannath, J.; Sheaffer, B.; Drozd, A. Developing a Low Cost, Portable Jammer Detection and Localization Device for First Responders. In Proceedings of the 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 11–14 January 2019; pp. 1–4.
27. Koohifar, F.; Guvenc, I.; Sichitiu, M.L. Autonomous Tracking of Intermittent RF Source Using a UAV Swarm. *IEEE Access* **2018**, *6*, 15884–15897. [CrossRef]

28. Pietrow, D.; Matuszewski, J. Objects detection and recognition system using artificial neural networks and drones. In Proceedings of the 2017 Signal Processing Symposium (SPSympo), Jachranka, Poland, 12–14 September 2017; pp. 1–5.
29. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. *arXiv* **2017**, arXiv:1709.01507.
30. Cao, Q.; Shen, L.; Xie, W.; Parkhi, O.M.; Zisserman, A. VGGFace2: A dataset for recognising faces across pose and age. In Proceedings of the International Conference on Automatic Face and Gesture Recognition, Xi'an, China, 15–19 May 2018.
31. dlib Histogram of Oriented Gradients (HOG) Face Detector. Available online: http://dlib.net/face_detector.py.html (accessed on 12 March 2020).
32. Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett.* **2016**, *23*, 1499–1503. [CrossRef]
33. OpenCV-DNN Face Detector—SSD Framework with ResNet-10 Model. 2017. Available online: https://github.com/opencv/opencv/blob/3.4.0/samples/dnn/resnet_ssd_face_python.py (accessed on 12 March 2020).
34. Nawaz, T.; Campo, D.; Mughal, O.; Marcenaro, L.; Regazzoni, C. Jammer detection algorithm for wide-band radios using spectral correlation and neural networks. In Proceedings of the 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 26–30 June 2017; pp. 246–251.