IET Radar, Sonar & Navigation

Special Issue: Innovative Radar Detection, Tracking and Classification for Small UAVs as an Emerging Class of Targets

IET Journals
The Institution of Engineering and Technology

DopplerNet: a convolutional neural network for recognising targets in real scenarios using a persistent range–Doppler radar

ISSN 1751-8784
Received on 29th June 2019
Revised 8th November 2019
Accepted on 6th February 2020
E-First on 3rd March 2020
doi: 10.1049/iet-rsn.2019.0307
www.ietdl.org

Ignacio Roldan[1] ✉, Carlos R. del-Blanco[2], Álvaro Duque de Quevedo[3], Fernando Ibañez Urzaiz[3], Javier Gismero Menoyo[3], Alberto Asensio López[3], Daniel Berjón[2], Fernando Jaureguizar[2], Narciso García[2]

[1]*Advanced Radar Technologies (ART), Madrid, Spain*
[2]*Grupo de Tratamiento de Imágenes, Information Processing and Telecommunications Centre, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain*
[3]*Grupo de Microondas y Radar, Information Processing and Telecommunications Center, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain*
✉ *E-mail: ignacioroldanmontero@gmail.com*

Abstract: In the past few years, the commercial use of drones has exploded, since they are a safe and cost-effective solution for many kinds of problems. However, this fact also opens the door for malicious use. This work presents a novel system able to detect and recognise drones from other targets, allowing the police and security agencies to deal with this new aerial thread. The proposed system only uses a persistent range–Doppler radar, avoiding the restrictions of the optical sensors, usually required for the recognition part. The processing is based on constant false alarm rate detection stage, followed by a convolutional neural network that performs the recognition. This network takes as input raw range–Doppler radar data and predicts their class (car, person, or drone). For this purpose, an extensive controlled trial test campaign has been performed, resulting in a novel dataset with more than 17,000 samples of drones, cars, and people, acquired in real outdoor scenarios. As far as authors' knowledge, this is the first range–Doppler radar database for the recognition of drones and other targets. The high-accuracy results (99.48%) suggest that this system could be successfully used in security and defence applications to discriminate between drones and other entities.

1 Introduction

In the past few years, the commercial use of unmanned aerial vehicles (UAVs), also known as drones, have exploded. Current studies estimate that the global market will hit $127 billion in 2020 [1]. A significant increase in the number of UAVs will be expected, as well. According to the Federal Aviation Administration of the US, there were registered over 670.000 drones only in the US in 2017, and it is prospected that over seven million drones will be sold in 2020 [2]. The main reason for this success is that in almost every sector, drones can be used as a safe and cost-effective solution for many kinds of problems. Even more, successful use cases have been achieved in agriculture, security, emergency management, and media. However, the fact that they are widely available, easy to use, and cheap, also opens the door for malicious use. There have already been cases of privacy violations, smuggling of illegal substances, collision hazards, and deployment of explosive weapons [3–5]. Therefore, police and security agencies must have the right equipment to deal with this new aerial thread.

Radar systems have been proved to be a good solution for early detection of these kinds of threats [6, 7], being able to detect drones from several kilometres. However, these systems usually report only the position of the target, and objects such as vehicles, people, or birds will be reported in the same way as drones (i.e. they will be indistinguishable among them). This classification is very important in order to take countermeasures against drones, such as turning on jamming systems that cannot radiate continuously for security reasons. Owing to this, radars usually rely on optical sensors to achieve classification using supervised machine-learning techniques, which exploits the rich information contained in visual or infrared imagery. More specifically, the artificial neural networks (NNs) have grown in popularity in the recent years due to the good results achieved in many different tasks, such as autonomous car driving [8], human–machine interface [9], object classification in videos [10] etc. Broadly speaking, NNs are designed for spotting patterns in data. However, optical sensors have serious limitations for distant drone detection and recognition: the operational range is much lower, and the performance is severely degraded in adverse weather conditions.

Taking this into account, this publication presents a novel system that uses only a persistent range–Doppler radar for target detection and recognition, avoiding the restrictions imposed by the optical sensors. The recognition capability using a persistent range–Doppler radar is theoretically feasible because it provides additional information about the radar detections, such as the power in each Doppler frequency or its evolution over time, also known as Doppler signature. However, this information can be difficult to interpret for a human operator. For this reason, a deep-learning-based automatic recognition system has been developed, which will provide high-level interpretations (target type and location) to the operator. More specifically, a convolutional NN (CNN) has been implemented, which takes as input arrays of range–Doppler radar data, and predicts their class (a car, a person, or a drone). It is known that the proper performance of the CNN depends on the availability of high-quality training data, which should be collected and labelled in real and representative scenarios. For this purpose, an extensive controlled trial test campaign has been done, resulting in a dataset with more than 17,000 samples of drones, cars, and people. This database, called Real Doppler RAD-DAR (RDRD) [11], has been uploaded to Kaggle and it is publicly accessible to anyone who wants to test their classification algorithms with real radar data. As far as authors' knowledge, this is the first public range–Doppler radar database for the recognition of unmanned aerial system (UAS) and other targets.

In the next section, a brief summary of the current state of the art about drone classification is presented. In Section 3, an introduction to the radar system used to gather the data is done. Section 4 explains the field test campaign and presents the analysis of the data. Section 5 shows the developed CNN architecture, and

finally, Section 6 presents the experiments and results with performance metrics.

2 State of the art

The current drone boom goes hand in hand with the upward trend in the number of publications about characterisation and classification of UAVs. In 2017 and 2018, more than 40 and 50 publications, respectively, were published about this topic, whereas fewer than ten were published per year in previous years [12]. Although there have been numerous works that examine the radar cross-section (RCS) characteristics for various drones, this section will be focused on publications addressing UAV classification.

Works in the literature addressing UAV classification and recognition based on radar information can be split into two groups: based on handcrafted features (where features are manually selected through oriented signal processing techniques) and based on deep-learning features (the own classification algorithm computes the best possible features for the target task).

2.1 Handcrafted feature techniques

In [13], Ren and Jiang propose a robust classification framework to distinguish between UAV and non-UAV targets. The proposed signal representation, called two-dimensional (2D) complex regularised spectral analysis, is based on the magnitude and phase information of the Fourier transform, producing a complex log spectrum. Also, classical spectrogram, cepstrogram, and cadence velocity diagram (CVD) are added to the feature vector to increase its discriminability. To remove the unreliable feature dimensions, a subspace reliability analysis (SRA) is proposed and compared with standard principal component analysis (PCA). The approach was applied to data provided by a low-power continuous-wave radar from Thales Asia operating at the X-band. The training database was composed of 854 s of recording for UAVs and 204 s for non-UAVs, mostly containing birds. The results show that the SRA algorithm outperforms the previous works for the same dataset, obtaining a low error and a very good false alarm rate. It is important to remark that this work used phase information, usually discarded for not being useful for a human operator in a 2D visual plot.

In [14], Torvik *et al.* aim to show the potential of using polarimetric parameters to distinguish between large birds and UAVs of comparable size. They used a dataset of more than 8000 samples containing four classes: drones with carbon fibre rotor blades, drones with plastic rotor blades, flapping wings birds, and gliding birds. Up to 12 polarimetric features were extracted from every radar signature sample across multiple domains, obtaining accuracy around 99%.

Other works have focused on determining whether a drone is carrying a payload or not. In [15], Ritchie *et al.* state that the increase in the mass causes different dynamics and inertia response on the drone, which carefully treated can lead to an accurate classification. Particularly, two different feature representations were used: based on the centroid analysis of the micro-Doppler spectrum and based on singular value decomposition (SVD). These features are used along with two different classifiers, a Naïve Bayesian classifier and a random forest. The database was acquired by a multi-static radar system called NetRad, developed by the University College of London, which is a netted coherent pulse radar with three nodes that operate at 2.4 GHz. The results show that the Naïve Bayesian classifier has the best performance for drone hovering, with a classification accuracy of 97% for the case of using micro-Doppler centroid features and only 1 s of data for feature extraction. On the other hand, the SVD approach obtained the best accuracy for flying micro-drones, above 95%.

In [16], Zhang *et al.* proposed a method to classify three different types of drones: quadcopters, helicopters, and hexacopters. Two different types of radars were used, one operating in the K-band and another operating in the X-band. First, the measured data is chopped in samples of 0.5 s, and then a short-time Fourier transform (STFT) is computed. PCA is applied to the resulting spectrograms and a support vector machine is used as the classifier. The used dataset contained 720 samples for each radar sensor and for each drone. The results showed that mixing information from both radars could increase the classification performance up to an overall accuracy of 97.7% across all three classes. Zhang *et al.* [17] performed an analysis with the same drones flying together in the same scene, but in different combinations, to see if it was possible to classify the type of drone with its CVD signature. An average accuracy of 94.7% was obtained over seven different scenarios, using a *k*-means clustering as the classification technique.

There has been also some effort in using wavelet transform (WT) as a feature for classification. In contrast with the Fourier transform, the WT preserves the time information. This allows to simultaneously analyse frequency and time variations in a multiresolution space, as opposed to STFT, where the resolution is determined by the depth of the FT. Rahman and Robertson [18] applied WT to the data captured by a phase-coherent W-band radar, obtaining more prominent features in comparison with other spectrogram techniques. Thus, it was concluded that an improved Doppler resolution does indeed provide more useful information, which benefits the potential classification opportunities involving small drones.

2.2 Deep-learning feature techniques

In contrast to handcrafted features approaches (where most of the effort is the selection of adequate features to represent the different classes), there has been some research focused on learning from data the optimal features for the target application. For example, Kim *et al.* [19] used the pre-trained 'GoogleNet' CNN to process the spectrograms and CVDs collected from two flying drones. They carried out two different experiments, one in an anechoic chamber and another in an outside scenario with hovering drones at 50 and 100 m. The resulting classification accuracies were 94.7 and 100%, respectively. The system to capture the data was a frequency-modulated continuous wave (FMCW) radar operating in the Ku-band, and the database was composed of 60,000 images.

In [20], Mendis *et al.* used a deep belief network (DBN) to classify three different types of micro-drones. The data, captured with an S-band CW radar, is pre-processed to obtain the spectral correlation function (SFC), which is the Fourier transform of the correlation function. The resulting SFC is used as the input feature for the DBN. The results are very dependent on signal-to-noise ratio, but it achieves accuracies around 97% for the three classes.

In the area of autonomous driving, deep learning has been used for ground target classification to avoid accidents in autonomous driving systems. For example, in [21], Kim *et al.* propose several convolutional recurrent NNs to classify among pedestrians, cyclists, and vehicles using range–velocity images captured with an FMCW radar working at 77 GHz. The best performing network obtains an average accuracy of 94.7% using 1505 samples.

Finally, Choi and Oh [22] simulate 500 spectrograms that model each part of the drone as a primitive shape, such as cylinders, ellipsoids, and spheres, whose RCSs are analytically well-characterised. Each of the samples are simulated by varying the number of rotors, the rotor speed, and the orientation of the drone. These spectrograms are used as feature vectors to train a CNN, designed to classify several types of drones that differed in their number of rotors. The obtained accuracy was 93%.

2.3 Contributions of the proposed system regarding the state of the art

Research in drone classification has been successful using both classical signal processing and modern techniques, such as deep NN. However, most of the discussed work make some assumptions that may be unrealistic in practical and real scenarios:

- Targets are visible for the radar enough time and they are also close enough to extract segments of micro-Doppler signatures.
- Drones are flying in close range so that the quality of the signal is near perfect.
- The simulated data (ideal scenarios) is similar enough to that generated by the real reflections of the drone.

Table 1 Waveform parameters

radar frequency (f_1)	8.75 GHz (X-band)
bandwidth (Δf)	200 MHz
ramp period (T_m)	350 μs
ramp frequency (f_m)	2.86 kHz

Table 2 Acquisition parameters

number of samples per ramp (N_s)	8192
number of integrated ramps (N_d)	512
number of channels (N_a)	8
number of range bins (N_r)	4096
dwell time (T_d)	0.1792 s
number of synthesised beams	5 (−40°, −20°,0°,20°,40°)
sample rate (f_s)	32 MHz
maximum beating frequency (f_{bmax})	16 MHz
mean-time gap between cubes	0.53 s

Table 3 Radar performance

range resolution (ΔR)	0.878 m
maximum unambiguous range	3.598 km
Doppler resolution	5.58 Hz, 0.34 km/h
Doppler ambiguity	1.4 kHz, 88.1 km/h

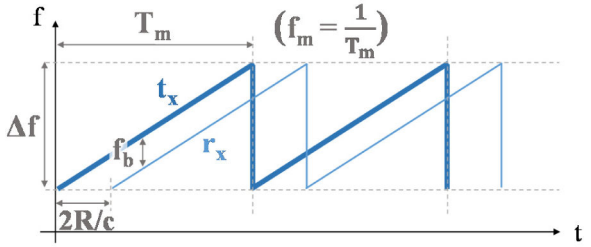

Fig. 1 *FMCW waveform*

The solution proposed in this work does not make such strong and unrealistic assumptions. It is more oriented to real operating conditions and scenarios. For this purpose, real data has been collected, captured by a real radar sensor, and targets distances are much further than those in other works, up to 3 km.

3 Radar system

This section summarises the main aspects of the sensor used to collect the data: a 2D ubiquitous or persistent radar system developed by the microwave and radar group from the Universidad Politecnica de Madrid [23, 24] called RAD–DAR (Radar with Digital Array Receiver). Ubiquitous radars are those that can light up the surveillance scene all the time with a wide beam. They also use multiple simultaneous beams generated on reception by applying digital beamforming. Since there is no scanning, an optimal trade-off between the dwell time and refreshing rate of target data is achieved. This is especially interesting when the targets are low RCS objects with slow dynamics such as drones. Tables 1–3 summarise the main features of the radar.

3.1 Hardware description

The radar uses an FMCW on a frequency band centred at 8.75 GHz with a maximum bandwidth of 500 MHz (see Fig. 1). A programmable signal generator (direct digital synthesiser + X-band phase-locked loop) provides this waveform, clock, and trigger signals. Next, a transmitter amplifies the signal up to 5 W and generates the local oscillator sample for subsequent demodulation.

A block diagram of the system can be seen in Fig. 2. Antenna subsystems have been designed in microstrip technology, ensuring

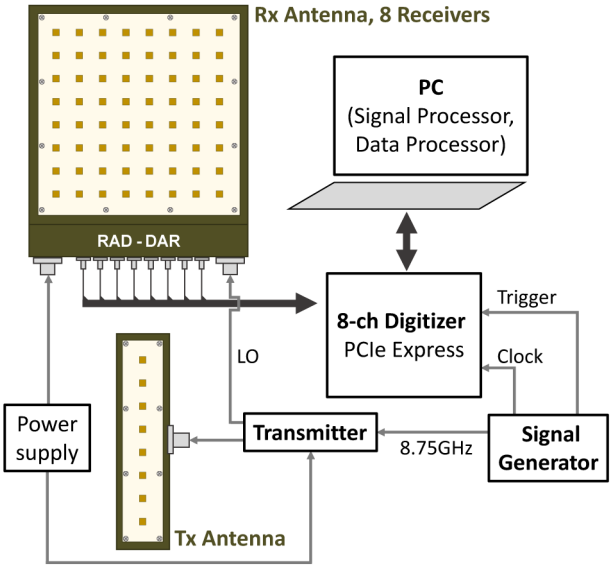

Fig. 2 *Radar block diagram*

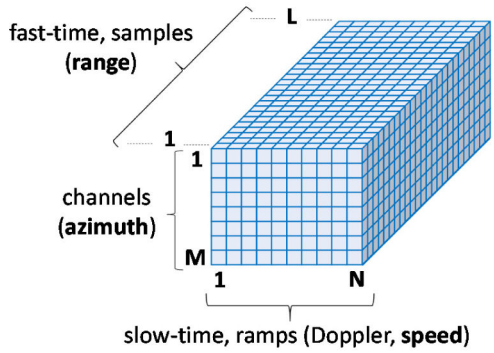

Fig. 3 *Radar data cubes*

a half-power beamwidth of 10° in the elevation plane and 90° in the azimuth plane. The system is composed of a transmitting antenna and eight receiving antennas, each of them composed of an array of eight elements.

3.2 Radar data generation

The signal processing chain uses the raw tensors acquired by the hardware to obtain a matrix for each scene, where the rows are the distance cells and the columns are the Doppler frequencies (see Fig. 3). This is performed by a 3D processing that uses 2D fast Fourier transform (FFT), beamforming, and monopulse techniques.

The target range information is drawn from the beating frequency since the system uses an FMCW waveform. This frequency is the difference between transmitted and reflected frequencies. Mathematically, range information is estimated as in the equation below:

$$R = \frac{f_b \cdot c \cdot T_m}{2\Delta f} \tag{1}$$

where R is the range in metres and f_b is the beating frequency, T_m is the signal period, c is the speed of light, and Δf is the bandwidth of the signal. The beating frequency is, in turn, obtained from applying the FFT to the fast-time axis (see Fig. 3). Next, a second FFT is applied in the slow-time axis to obtain the Doppler information of the target.

Then, the signal processor synthesises five reception beams pointing to $\varnothing_n$ degrees with the information from the eight channels, corresponding to pointing angles between −40° and + 40°. This is carried out by applying a phase increment to each received signal as in the equation below:

$$\varnothing_n = \frac{2\pi}{\lambda} \cdot n \cdot \Delta x \sin\theta \tag{2}$$

where θ is the steering direction, n is the channel, $\Delta \boldsymbol{x}$ is the separation between receiving antennas in the azimuthal plane (in this case 0.0185 m), and λ is the wavelength (0.0343 m).

By adding these eight phase-shifted signals, a range–Doppler matrix is generated for each pointing angle. These matrices, properly trimmed and selected, will be the input for the CNN-based classification algorithm. An example of these matrices can be seen in Fig. 4.

Once the matrices for each pointing angle have been generated, a monopulse technique is applied to measure the target's azimuth and reject detections that do not correspond to a processed beam. This technique consists in generating a difference signal for each pointing angle by subtracting the channels 5–8 from the channels 1–4. The quotient between the magnitude of the difference signal (Δ) and the sum signal (Σ) calculated with the beamforming, provide the monopulse ratio e [25] according to the equation below:

$$e = \frac{|\Delta|}{|\Sigma|} \cdot \cos\delta \tag{3}$$

where δ is the phase angle between Δ and Σ. The azimuth information is then computed by comparing the monopulse functions values between those obtained from the data cubes with those obtained from anechoic-chamber data [25].

As a result, a 4092×512 matrix for each scene is obtained, where the rows are the distance cells, the columns are the Doppler frequencies, and the values are in dBm.

3.3 Radar target detection

A CA-cell-averaging-constant false alarm rate [26] detector is applied to obtain the radar plots. This is carried out by comparing the signal power at a range bin with an adaptive threshold calculated as the average signal power of its adjacent cells. Several guard cells are fixed to avoid errors in the power estimation of adjacent cells caused by target echoes. Each target detection is turned into a plot (a single-point detection) by computing its centre of mass, usually referred to as the centroid. As a result, a list of plots is obtained, where each one has information about range, speed, azimuth, received signal power, and time.

The previous list of plots is used to trim the distance–Doppler matrices (for example, the one shown in Fig. 3) to obtain submatrices that are related to the potential targets, such as the ones shown in Fig. 5. In this last figure, an example of trimmed matrices for each of the considered types of targets (cars, drones, and people) is shown, where the dimensions have been experimentally set to 11×61 cells, after the analysis of the collected radar data. This matrix size amounts to a range of 9.658 m in distance and 20 km/h around the target detection, according to the radar parameters in Table 3.

Note the visually different characteristics of each target. First, drones are low RCS targets, and therefore, the received echoes have low power. Also, they are physically small and compact, which makes that the received power is focused on a few cells. On the other hand, cars are big metal targets with high RCS, which induces more dispersion in distance cells. They are also very compact with no moving parts at different speeds, and hence the energy is compressed in a couple of Doppler bins. Finally, people echoes are medium power and less disperse in distance. However, they are spread among more Doppler bins due to the independent leg and arm motion regarding the body in the walking action. (the periodical movement of the arms and legs is sometimes faster than the body and sometimes slower). The periodical change of the energy distribution among different Doppler bins can be used also to distinguish among classes. Fig. 6 shows three time-consecutive matrices, which are separated 400 ms. It can be seen how drones and cars have a more stable shape than people, who tend to have a periodic shrinking and expanding movement due to the arms and legs motion.

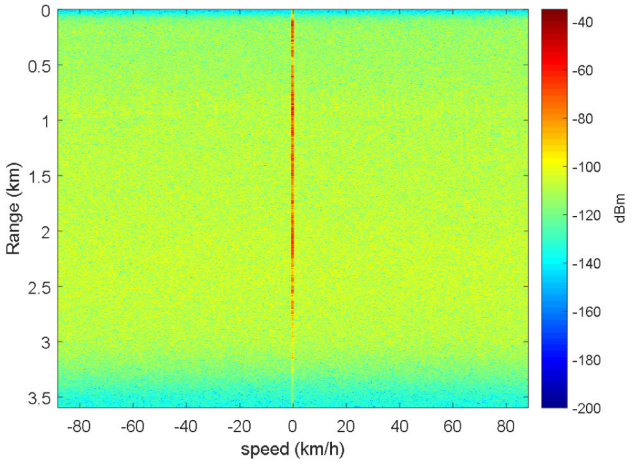

Fig. 4 *Range–Doppler matrix*

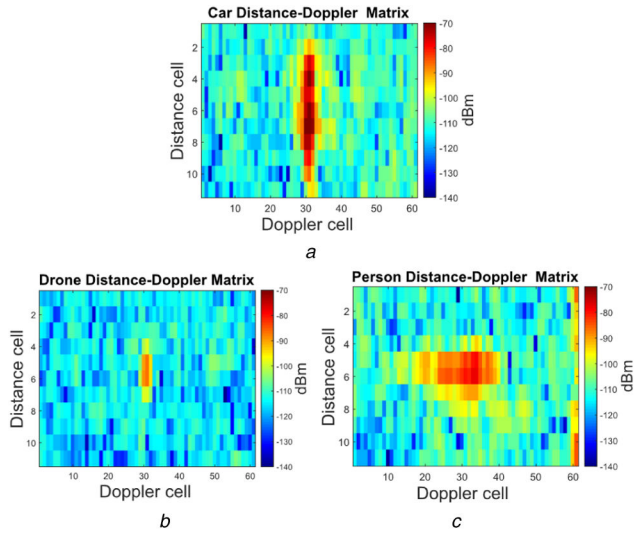

Fig. 5 *Trimmed distance–Doppler matrices*
(a) Car matrix, ***(b)*** Drone matrix, ***(c)*** Person matrix

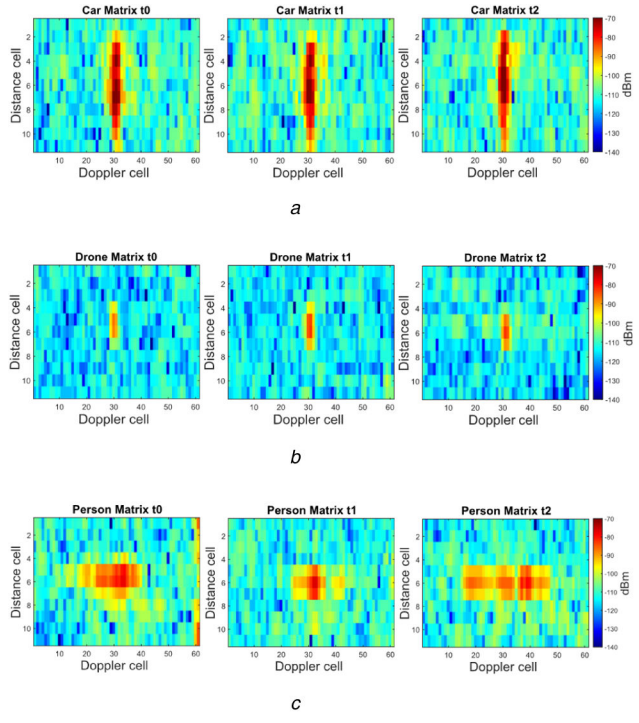

Fig. 6 *Matrices evolution over time*
(a) Car matrices, ***(b)*** Drone matrices, ***(c)*** Person matrices

Fig. 7 *Field test set-up*

Table 4 DJI Phantom four main specifications

weight (propellers and battery included)	1380 g
diagonal length (propellers not included)	350 mm
maximum speed	72 km/h
maximum flying height (above sea level)	6000 m
positioning system	global positioning system/ Glonass
maximum flying time (battery life)	28 min
maximum range (remote control range)	5 km

To take advantage of the previous time dependency, the input data to the designed CNN will be a tensor with multiple stacked trimmed matrices of different time steps. More specifically, the input for the proposed CNN will be a $11 \times 61 \times 3$ (distance–Doppler–time) tensor.

4 RDRD database

The performance of the classification algorithms is directly connected to the quality of the data. Therefore, a controlled trial test campaign has been done in order to create a database, called RDRD [11], by collecting reliable labelled data of drones, people and cars under different situations. This database is publicly available in Kaggle under the licence CC BY 4.0, and therefore, everyone can use it and modify it for testing their algorithms with real radar data. The RDRD database is composed of thousands of CSV files, where each of them contains an 11×61 matrix representing the power in dBm of each of the cells, such as the ones presented in Fig. 5. Moreover, the samples of each target are organised in a directory structure, where the root folder is the name of the class (i.e. drone, car, or person) and the subfolder contains all the data belonging to a time-correlated experiment. As far as authors' knowledge, it is the first publicly available persistent range–Doppler radar database composed by real data recorded in a real scenario, not synthetic.

The selected scenario for the recoding of the database is a farm located in Riocabado, a village in the province of Avila, Spain. This place counts with an entirely unobstructed line of sight extended up to 5 km. Fig. 7 shows the radar system installed for the acquisition.

To guarantee the diversity in the recorded data, several exercises have been performed trying to have measurements of each of the targets in different distances, moving at different speeds and with different orientations. The ground targets are limited to the allowed paths and direct line of sight is needed in order to detect them, therefore, a visibility map calculated using the digital

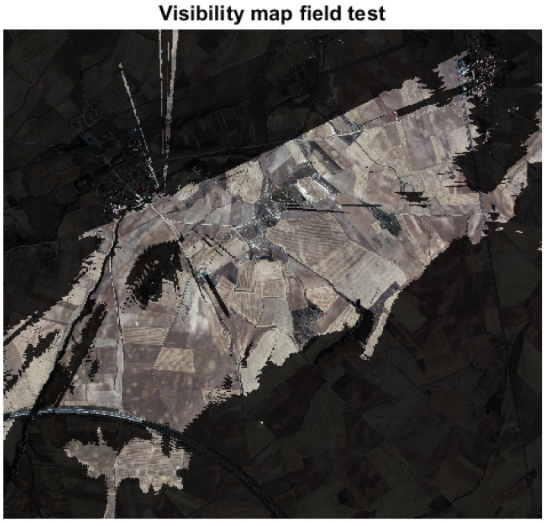

Fig. 8 *Visibility map*

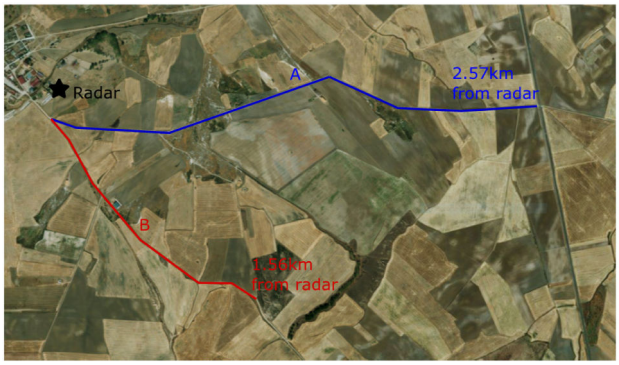

Fig. 9 *Ground routes*

elevation model of the terrain has been used to determine the best routes (see Fig. 8).

Considering the information provided by the visibility map, only two routes can be used for the ground targets. Route A, in blue shown in Fig. 9, has a total length of 2.7 km and a maximum distance from the radar of 2.5 km. In this route, the targets are moving tangentially at the beginning and radially by the end, and therefore, the orientation is going to change over time. On the other hand, route B, in red shown in Fig. 9, has a total length of 1.7 km with a maximum distance of 1.5 km. This route is on a slope, which will help in the variety of orientations captured. Furthermore, each path has been repeated at different speeds, for the purpose of distributing the power in different Doppler bins.

On the other hand, drones are less limited by the environment, but all the flights have been done at the east of the radar to avoid flying above the village. During all the tests performed a commercial micro-drone, DJI Phantom 4 [27] designed for both private and professional operation have been used. Table 4 shows its main features.

Three different exercises have been recorded with the drone. First, different radial trajectories up to 3.1 km from the radar have been performed using the autopilot in the return way. Second, a circular trajectory with a radius of 200 m centred at a range of 800 m, and finally, a free flight changing speed and altitude manually. The result of this campaign is a novel labelled dataset with a total of 17,485 samples belonging to three different classes: 5065 drones, 6700 people and 5720 cars. The data have a distribution of 32.71% of the samples belonging to cars, 28.97% corresponding to drones, and 38.32% to people, which are well-balanced and suitable for training NNs.

5 CNN classifier

The developed CNN is oriented to run in embedded systems to provide real-time target recognition for the radar operators. This

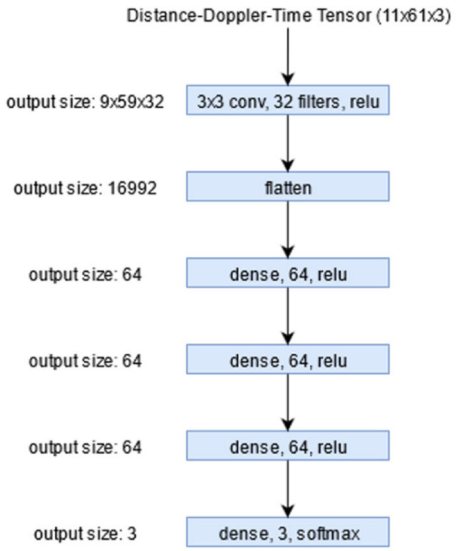

Fig. 10 *CNN structure*

restricts the total number of network parameters (and, therefore, layers) to consume few computational, power, and memory resources.

The architecture of the proposed network (see Fig. 10) is described as follows. The first layer is a convolutional layer with a filter size of 3 × 3 and a total number of 32 filters. The result of this layer is a 3D array of values that represents the processing of the input data. Each 2D output slice of this 3D array is called feature map. Moreover, to simulate non-linear behaviours of the input data, a rectified liner unit (ReLu) [28] function is applied to each of the convolutional layer's outputs. The next layer is a flatten one, which reshapes the previous feature map into a vector, allowing to stack dense layers (also called fully connected layers) in the next steps. More specifically, four of these dense layers are concatenated, the first three ones using a ReLu activation function, and the last one a SoftMax function. The output of SoftMax is composed of three probability values, one per every considered target to be recognised. All these layers gradually transform the input data into richer and more semantic features of radar information. In another to tune the hyperparameters of the network, such as the number of filters or the kernel size, a grid search has been performed. This method evaluates several configurations of the hyperparameters from a given range and obtains the best combination of them. The results presented in the next section are obtained with the best hyperparameter combination.

6 Results

This section starts showing the effect of using multiple consecutive frames in the recognition performance, and finishes comparing the proposed CNN with other state-of-the-art networks on the created radar database.

A test has been carried out using networks inputs with an increasing number of time-consecutive frames (that are concatenated) to measure the effect on the recognition accuracy. Fig. 11 shows the results of this test. Note that increasing the number of frames has two effects on the system. First, the number of parameters of the network will be greater, and therefore, the computational and memory costs of the system will be higher. Second, every extra added frame will delay the classification around 400 ms (this is the gap between two consecutive frames). Considering these two factors and the small increase in performance when adding more than two frames, three consecutive frames have been selected for the system as a reasonable trade-off.

Next, two state-of-the-art embedded networks are compared with the proposed CNN: MobileNetV2 [29] and NasNet [30]. These are specifically designed for running in embedded systems and have demonstrated good performance.

MobileNetV2 is based on depthwise separable convolutions that allow very memory-efficient inference. It has obtained very good

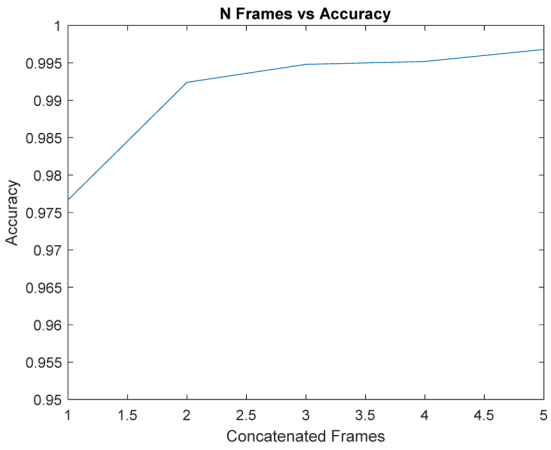

Fig. 11 *Concatenated frames versus accuracy*

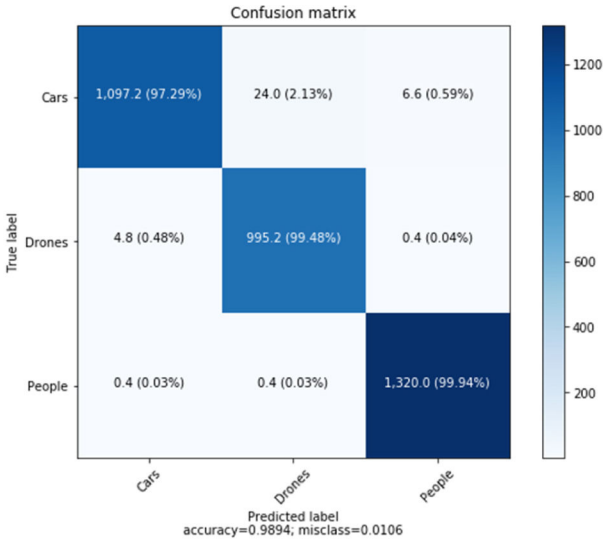

Fig. 12 *MobileNetV2 confusion matrix*

performance on the ImageNet dataset [31]. NasNet has been also used for image classification and proposes a new regularisation technique, called ScheduledDropPath, to outperform the generalisation capability.

All the previous networks have been tested on the RDRD database. A five-fold validation process has been performed, where the dataset has been split randomly into five sub-datasets containing 20% of the data each. The training process has been repeated five times, using each time a different combination of the four sub-datasets for training and only one for testing. Also, a randomly selected 10% of the training samples have been used for validation in training phase. Finally, the data have been normalised to zero mean and a standard deviation of one before the training process. The three tested network architectures have been trained under the same conditions: no pre-trained weights, 25 epochs, and using Adam [32] as the optimisation algorithm.

Figs. 12–14 show the confusion matrices of MobileNetV2, NasNetMobile, and the proposed network model, respectively. The confusion matrices shown are the mean of the five confusion matrices obtained with the five-fold validation process. They get a mean accuracy of 0.9894 for MobileNetV2, 0.9769 for NasNetMobile and 0.9948 for the proposed network, DopplerNet. Although the recognition performance is quite satisfactory in all the cases, the proposed network model outperforms the others. However, accuracy is not the only criterion for judging the network. Since a false positive of a drone could lead to triggered unnecessary (and sometimes dangerous) countermeasures, the precision is an important metric to consider. Moreover, a false negative means an undetected intrusion that can have disastrous consequences, so recall is also very relevant. Finally, another criterion is the hardware complexity required to implement the

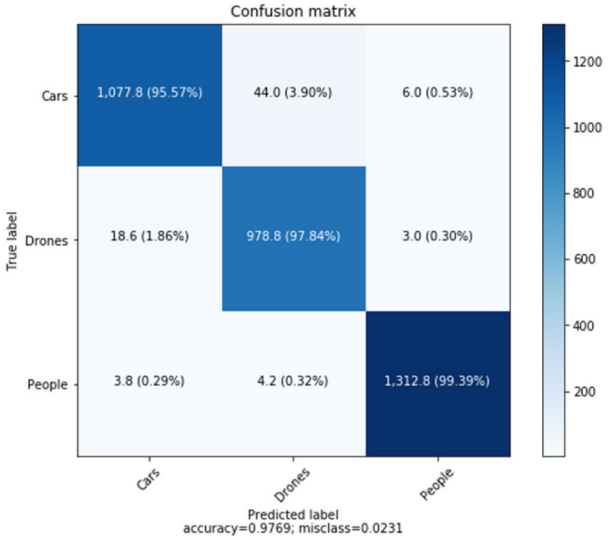

Fig. 13 *NasNetMobile confusion matrix*

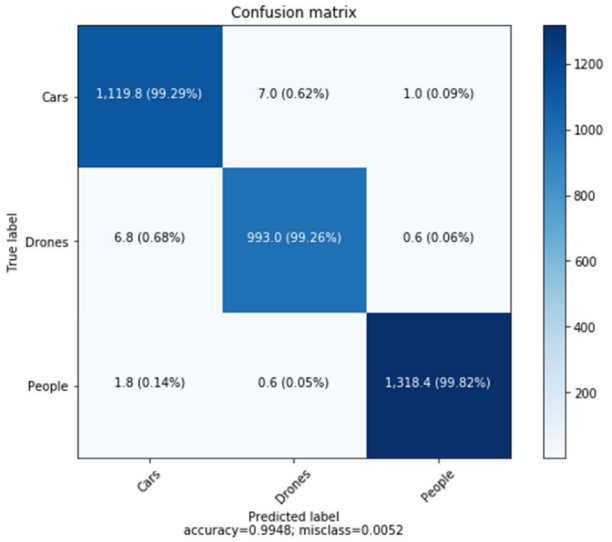

Fig. 14 *DopplerNet confusion matrix*

Table 5 Models metrics

Model	Accuracy	Precision	Recall	Number of parameters
NasNetMobile	0.9769	0.9771	0.9769	4.272.887
MobileNetV2	0.9894	0.9895	0.9894	3.325.043
DopplerNet	0.9948	0.9948	0.9948	3.818.755

network, which is dependent on the number of its parameters. This should be as small as possible. Table 5 shows all these metrics for the three proposed networks. As it can be seen, the presented CNN model outperforms in accuracy, precision, and recall both pre-designed networks. Regarding the number of parameters, the custom CNN has 12.92% more parameters than MobleNetV2, but 11.89% less than NasNetMobile. In any case, all of them have the same magnitude of parameters.

7 Conclusions

One of the main contributions of this paper has been to generate the RDRD database including drones, cars, and people that are publicly available on the Kaggle webpage.

Another contribution is the comparison of three different NNs, two pre-designed to classify images, and another designed for this application. The accuracy results show that is not only possible to have a reliable classification, but also with high accuracy. The proposed CNN has obtained the highest accuracy: 98.99%. These results support that radar systems can make use of these algorithms to provide information about the nature of the target to the operator. This is especially interesting when the targets are UAVs, and countermeasures, such as jamming systems, can be triggered autonomously when the threat is classified as a drone.

8 Acknowledgments

Advanced Radar Technologies have played a key role in the development of this publication providing support during the field tests and allowing the use of their facilities. Moreover, this work has been partially supported by the Ministerio de Economía, Industria y Competitividad (AEI/FEDER) of the Spanish Government under project TEC2016–75981 (IVME) and also by the Ministerio de Ciencia, Innovacion y Universidades within the projects TEC2014-53815-R and TEC2017-87061-C3-1-R (CIENCIA/AEI/FEDER, UE).We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

9 References

[1] 'Clarity from above – PricewaterhouseCoopers'. Available at https://www.pwc.pl/en/publikacje/2016/clarity-from-above.html, accessed May 2019

[2] 'Drone registrations are still soaring – Jonathan Vanian'. Available at http://fortune.com/2017/01/06/drones-registrations-soaring-faa/, accessed May 2019

[3] 'Drones seized over HMP Pentonville carrying drugs and phones – BBC'. Available at https://www.bbc.com/news/uk-england-london-37152665, accessed May 2019

[4] 'Heathrow plane is near miss with drone – BBC'. Available at https://www.bbc.com/news/uk-30369701, accessed May 2019

[5] 'Apparent attack in Venezuela highlights risk of drone strikes – Reuters'. Available at https://www.reuters.com/article/us-venezuela-politics-drones/apparent-attack-in-venezuela-highlights-risk-of-drone-strikes-idUSKBN1KQ0MG, accessed May 2019

[6] 'Drone Sentinel – advanced radar technologies'. Available at http://www.advancedradartechnologies.com/products/art-drone-sentinel/, accessed May 2019

[7] 'Anti-UAV defence system – Blighter'. Available at http://www.blighter.com/products/auds-anti-uav-defence-system.html, accessed May 2019

[8] Loquercio, A., Maqueda, A.I., Del-Blanco, C.R., *et al.*: 'DroNet: learning to fly by driving', *IEEE Robot. Autom. Lett.*, 2018, **3**, (2), pp. 1088–1095

[9] Bao, P.J., Maqueda, A.I., Del-Blanco, C.R., *et al.*: 'Tiny hand gesture recognition without localization via a deep convolutional network', *IEEE Trans. Consum. Electron.*, 2017, **63**, (3), pp. 251–257

[10] Maqueda, A.I., del Blanco, C.R., Jaureguizar, F., *et al.*: 'Structured learning via convolutional neural networks for vehicle detection'. SPIE Commercial + Scientific Sensing and Imaging, Anaheim, CA, USA, April 2017, (SPIE Proceedings Vol. 10223, Real-Time Image and Video Processing 2017), pp. 1–9

[11] 'RDRD database – Kaggle'. Available at https://www.kaggle.com/iroldan/real-doppler-raddar-database, accessed June 2019

[12] Patel, J.S., Fioranelli, F., Anderson, D.: 'Review of radar classification and RCS characterisation techniques for small UAVs or drones', *IET Radar Sonar Navig.*, 2018, **12**, (9), pp. 911–919

[13] Ren, J., Jiang, X.: 'Regularized 2-D complex-log spectral analysis and subspace reliability analysis of micro-Doppler signature for UAV detection', *Pattern Recognit.*, 2017, **69**, pp. 225–237

[14] Torvik, B., Olsen, K.E., Griffiths, H.: 'Classification of birds and UAVs based on radar polarimetry', *IEEE Geosci. Remote Sens. Lett.*, 2016, **13**, (9), pp. 1305–1309

[15] Ritchie, M., Fioranelli, F., Borrion, H., *et al.*: 'Multistatic micro-Doppler radar feature extraction for classification of unloaded/loaded micro-drones', *IET Radar Sonar Navig.*, 2017, **11**, (1), pp. 116–124

[16] Zhang, P., Yang, L., Chen, G., *et al.*: 'Classification of drones based on micro-Doppler signatures with dual-band radar sensors'. Progress in Electromagnetics Research Symp., Singapore, November 2017, pp. 638–643

[17] Zhang, W., Li, G., *et al.*: 'Detection of multiple micro-drones via cadence velocity diagram analysis', *IET Electron. Lett.*, 2018, **54**, (7), pp. 441–443

[18] Rahman, S., Robertson, D.: 'Time-frequency analysis of millimeter-wave radar micro-Doppler data from small UAVs'. Sensor Signal Processing for Defence Conf. SSPD, London, UK, December 2017, pp. 1–5

[19] Kim, B.-K., Kang, H.-S., Park, S.-O.: 'Drone classification using convolutional neural networks with merged Doppler images', *IEEE Geosci. Remote Sens. Lett.*, 2017, **14**, (1), pp. 38–42

[20] Mendis, G.J., Randeny, T., Wei, J., *et al.*: 'Deep-learning-based Doppler radar for micro UAS detection and classification'. Proc. IEEE Military Communications Conf. MILCOM, Baltimore, MD, USA, November 2016, pp. 924–929

[21] Kim, S., Lee, S., Doo, S., *et al.*: 'Moving target classification in automotive radar systems using convolutional recurrent neural networks'. European Signal Processing Conf., Rome, Italy, September 2018, pp. 1482–1486

[22] Choi, B., Oh, D.: 'Classification of drone type using deep convolutional neural networks based on micro-Doppler simulation'. Int. Symp. Antennas and Propagation (ISAP), Busan, South Korea, , October 2018, (1), pp. 1–2

[23] Ibañez Urzaiz, F., Duque de Quevedo, Á., Martín Ayuso, A., *et al.*: 'Design, implementation and first experimental results of an X-band ubiquitous radar system'. IEEE Radar Conf., Oklahoma City, OK, USA, April 2018, pp. 1150–1155
[24] Duque de Quevedo, Á., Ibañez Urzaiz, F., Gismero Menoyo, J., *et al.*: 'Drone detection and radar cross-section measurements by RAD–DAR', *IET Radar Sonar Navig.*, 2019, **13**, pp. 1437–1447
[25] Skolnik, M.: '*Radar handbook*' (McGraw-Hill, New York, NY, USA, 2008)
[26] Richards, M.A., Scheer, J.A., Holm, W.A.: '*Principles of modern radar: basic principles*' (SciTech Publishing, Raleigh, NC, USA, 2010)
[27] 'DJI'. Available at https://www.dji.com/es/phantom-4, accessed May 2019
[28] Nair, V., Hinton, G.E.: 'Rectified linear units improve restricted Boltzmann machines'. Proc. 27th Int. Conf. Machine Learning (ICML-10), Madison, WI, USA, June 2010, pp. 807–814
[29] Sandler, M., Howard, A., Zhu, M., *et al.*: 'Mobilenetv2: inverted residuals and linear bottlenecks'. Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, June 2018, pp. 4510–4520
[30] Zoph, B., Vasudevan, V., Shlens, J., *et al.*: 'Learning transferable architectures for scalable image recognition'. Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, June 2018, pp. 8697–8710
[31] 'ImageNet'. Available at http://www.image-net.org/, accessed May 2019
[32] Kingma, D.P., Ba, J.: 'Adam: a method for stochastic optimization', 2014, pp. 1–15. Retrieved from http://arxiv.org/abs/1412.6980, accessed May 2019