Urban Bird-Drone Classification with Synthetic Micro-Doppler Spectrograms

White, Daniel; Jahangir, Mohammed; Baker, Chris; Antoniou, Michail

Link to publication on Research at Birmingham portal

Urban Bird-Drone Classification with Synthetic Micro-Doppler Spectrograms

Daniel White, Mohammed Jahangir, Chris J. Baker, *Fellow, IEEE* and Michail Antoniou, *Senior Member, IEEE*

Abstract— In this article, a method for creating highly realistic synthetic drone micro-Doppler spectrograms is presented and its effectiveness of training a bird-drone classifier for real scenario classification is shown via comparisons to a real benchmark. The effect of drone motor speed sampling used when simulating drone micro-Doppler is shown to have a significant impact on the accuracy of synthetic results and variations of this approach are explored. Four synthetic datasets were created differing in motor speed sampling and each were compared in their ability to train a convolutional neural network to classify real data. The highest fidelity synthetic dataset achieved a classification accuracy of 86.6% compared to the real benchmark accuracy of 89.7%. The adverse effect on classifier robustness when reducing the simulation fidelity by altering the motor speed sampling is shown.

Index Terms— UAVs, Avian, Simulation, CNN, Synthetic Training, HERM

I. INTRODUCTION

UN-MANNED aerial vehicles or drones continue to see a proliferation in technological capabilities and consumer interest, but widely available, low-cost drones pose the risk of their illegitimate use and exploitation in sensitive airspaces [1]. As drones are small and can fly at low altitudes, they are difficult to detect and differentiate from birds and current academic and industrial efforts are to improve sensors [1], [2], including radar, for reliable drone classification [3], [4] in difficult operational environments [5] against increasingly complex confuser targets. Birds and drones have similar flight profiles, yet it has been shown tracker feature [6], [7] and spectral feature extraction [8], [9] approaches offer an effective baseline of classifiable features. Using hand-crafted, processed features such as these alone do not provide adequate separability for reliable classification of edge cases, such as for unseen targets or at the radar's operational limit. Spectral feature classification approaches may be over reliant on the mere presence of characteristic micro-Doppler (μDoppler) sidebands drones have been observed to produce [10].

Daniel White, Mohammed Jahangir, Chris Baker and Michail Antoniou are with the School of Electrical, Electronic and Systems Engineering, University of Birmingham, UK. (email: d.white.4@bham.ac.uk; m.jahangir@bham.ac.uk; c.j.baker.1@bham.ac.uk; m.antoniou@bham.ac.uk)

The extraction of data-driven features from a large training set through deep learning is another, highly effective approach for classification of radar signals in many applications, such as in human gait [11], (inverse) synthetic aperture radar target [12], [13], [14], [15] and indeed small airborne target classification, where drone vs confuser target [16], [17], inter-drone model [18], and operational state classification, such as measurements of rotor speeds [19], blade length and blade and motor count [20] or estimations of payload mass [21] have been done. It has been shown that Convolutional Neural Networks (CNNs) can produce very high ceilings of classification accuracy at short range [16], [18], in benign environments [22] and with simulated data [23], [24], [25]. Further, novel deep learning classification schemes have been demonstrated using recurrent models [26], fusion of data inputs [22], [27], and un- [28] and semi-supervised [29] learning approaches.

A lack of quality training data is a key issue preventing validation of deep learning classifiers for deployment in security applications. As a result, exploration of data augmentation methods to bolster the amount of existing data have been done, such as affine transformations of radar data [13] (flipping, rotating), injection of noise into existing data [13], [17] and also the use of transfer learning of existing models trained on optical images [17] or synthetic radar data [11] retrained using a modest amount of real data. Another approach to increase the data available for training has been to include synthetic data created from simulation models [12]. If feasible, this method is advantageous as it allows for a precise control of the output data through exploration of possible target parameters. For drones, this includes various specific parameters, such as the number, size and shape of blades and rotors. More generally, using synthetic data permits great variety of the target setting, such as, setting a desired Signal-to-Noise Ratio (SNR) of returns, specifying target motion and velocity, and the inclusion of any real or synthetic noise background that the simulated target could be injected into.

Synthetic drone radar data can be created using a number of methods [30], [31]. The decomposition of targets into their rotating/moving parts and baseband I/Q simulation of reflected returns has been shown to be a fast and effective simulation approach. This method faithfully reproduces the μDoppler sidebands found in real radar drone data [10], [24], [32] and large-scale creation of such data has allowed investigative studies of deep learning on simulated data for inter-drone classification [19], [20], [24]. Less frequently reported in literature is the effectiveness of synthetic data for the training of classifiers tested on real drone and bird data [33]. Further,

the effect of simulation fidelity on classifier performance is unknown.

This work on radar classification of bird and drone targets assesses an approach of creating simulated drone μDoppler data by testing its efficacy as synthetic training data on a real dataset existing over long ranges and in a difficult, urban operational environment. Through the collection of genuine drone flights and the recording of their trajectory and motor speeds, we produce synthetic data that is equivalent to its real counterpart by using these flight parameters to inform the simulator [34]. The simulation is then subject to different motor speed sampling choices which are shown to affect the spectral output realism significantly. Simulated data is then modified to be alike to their real counterpart by injection into a real noise background, the treatment of sidebands to have a realistic amplitude modulation and the power scaling of body and blade components to match the SNR found in the real data. Four different versions of simulated data are created varying only in motor speed sampling, and alongside a real baseline training set, each were used to train the same CNN classifier and each of these classifiers were tested on the same, real dataset allowing a performance comparison of their utilization as training data. The accuracies, and class recall rates are compared across the synthetic versions elucidating the fidelity of simulation required for reliable and balanced synthetic training. The remainder of this article is organized as follows. Section II shares the drone simulation model that is used as the seed for generating synthetic data and how the output depends on the choice of motor speed sampling of the drone target. Section III explains the experimental setup for obtaining the real data that is used in the investigation of the training via synthetic data and includes details of the radars, environment and bird and drone targets used to create baseline results. Section IV introduces the CNN architecture and training procedure used in the real synthetic training experiments alongside results of the real data classification baseline. Section V will describe a procedure for creating realistic synthetic drone datasets and how the motor speed sampling choice effects the μDoppler fidelity, alongside the results and comparisons of performance using the real and various synthetic datasets tested on real data. Section VI is a discussion of the implications of shared results and, finally, section VII will summarize findings and offer conclusions.

II. MULTI-ROTOR DRONE MICRO-DOPPLER FRAMEWORK

This section describes the basic framework used for simulating drone μDoppler. Characteristic μDoppler sidebands induced by a rotating motor structure are observed in helicopters and drones, and are known as Helicopter Rotor Modulation (HERM) lines [35]. Depending on the radar, the drone target and the range, μDoppler modulations induced by the rotating propeller blades can be detected and measured [36]. The μDoppler pattern presents as numerous sidebands either side of the main body return if the radar integration
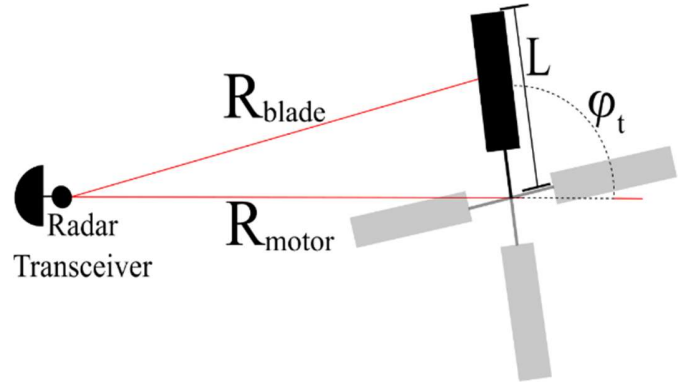

Fig. 1. Schematic of a drone motor and radar transceiver.

period is longer than the rotation period of a drone blade, and as a sinusoidal trace or blade flashes if not [10]. Many have reported on drone μDoppler in regard to its collection and appearance across pulsed [32], [34], [37] and frequency modulated continuous-wave radar [16], [38], [39] in regard to its ability to provide classifiable features [9], [40]. These returns can be mathematically approximated by treating the rotor blades as thin wires rotating about a central motor hub. The model originally created for the modelling of helicopter μDoppler can be extended for drones by considering contributions from more than one motor hub. Each mth motor contains B blades that each contribute to signal modulations. The derivation begins with the definition of geometrical lengths pertaining to the drone-radar interaction, shown in Fig. 1. Considering first a single blade on a single motor, R_{blade} and R_{motor} are the monostatic distances of the transceiver to the drone's motor center and an arbitrary point on its blade. As the radial length of the drone blade, L, is small it is assumed that $L \ll R_{motor}$ allowing for a Taylor series approximation simplifying the analytical form of the equation producing R_{blade}. Then integrating infinitesimal blade contributions from $0 \rightarrow L$, and summing over each blade on each motor yields the following baseband, complex echo signal:

$$X_{motors}(t) = \Sigma_{m=1}^{M} \Sigma_{b=1}^{B} L e^{-j\frac{4\pi}{\lambda}\left(R(t)+\frac{L}{2}\cos\left(\varphi_t^{m,b}\right)\right)} * sinc\left(\frac{L}{2}\frac{4\pi}{\lambda}\cos\left(\varphi_t^{m,b}\right)\right) \quad (1)$$

where λ is the radar wavelength, $R(t)$, is the target range at time, t, and $\varphi_t^{m,b} = \frac{2\pi b}{B} + \Omega_t^m t$ is the phase angle of the bth blade where Ω_t^m is the rotation rate of the mth motor. This formulation ignores any constant amplitude terms. This model assumes no tilt in the rotor plane and full mathematical derivations and expansions can be found in [34], [35], [41], [42].

This work explores how the appearance and training efficacy of the synthetic μDoppler changes using different Ω_t^m choices. The most common choice found in literature is to apply the same estimated rotation speed to each motor [19], [23], [24], [32], but increasingly more convincing results were found using dynamic and different motor speeds across each

motor reverse engineered from a real trajectory [42], and more recently [34] which used real recorded motor speeds from a drone flight. Fig 2. uses equation (1) and three different choices of Ω_t^m to create example μDoppler spectrograms for a pulsed radar system with 7 kHz Pulse Repetition Frequency (PRF) via a Short-Time Fourier Transform (STFT) using 4096 pulses with Blackman-Harris window and 50% overlap is generated, and this processing is used throughout this work when generating the spectrograms. The effective Coherent Processing Interval (CPI) in this case is 0.56 s, which is equivalent to a timestep as observed in the spectrogram. An arbitrary drone radial velocity profile is used and the drone body Doppler returns are not considered. If many full rotations of the blade structure occur withing a CPI, harmonic peaks appear with uniform frequency separation centered around the body velocity with SNR decreasing further from the body return. This is the only configuration of μDoppler that will be concerned in this paper. The parallel spectral lines occur as the motor rotation provides an additional oscillation in the timeseries signal which is brought out in the STFT as multiple harmonic peaks centered around the Doppler frequency of the motor platform (drone body). The number of observed spectral lines depends primarily on the length of the blade and the radiation scattering nature, the target range and the radar's noise profile [36]. From this basic model a whole variety of μDoppler spectra can be generated by varying the model parameters such as M, B and L. In this work, we focus on the effect of rotor speeds, Ω_t^m. Drone motor speeds would be expected to fall between 40 and 250 Hz, depending principally on the propellor length. Fig. 2a uses constant motor speeds across all motors and time, and uniform single sidebands are observed in this instance. Fig. 2b shows the result obtained when each of the four motors are set to a different fixed speed of 85, 90, 95 and 100 Hz. In this case there are now four distinct sidebands within each harmonic grouping that result from the differing speeds of each rotor blade. Fig 2c uses real motor speed values that were extracted from a drone flight [34]. This approach is further developed in Section V where we introduce varying strategies for sampling real rotor speed data to aid the modelling process. The model is refined with reference to real data of drone targets and further steps are included in the model to scale and normalize the data for SNR. Prior to the detailed description of the full simulation model, the radar systems and datasets that form the basis of the experimental validation are described and explained in the following section.

III. RADAR SYSTEM AND TARGET DATASET

A. Radar System

The multi-rotor drone μDoppler defined in Section II is to be used with reference to data recordings from a real radar system. The University of Birmingham (UoB) possesses two pulse-Doppler L-Band staring radars designed for detection, tracking and classification of small
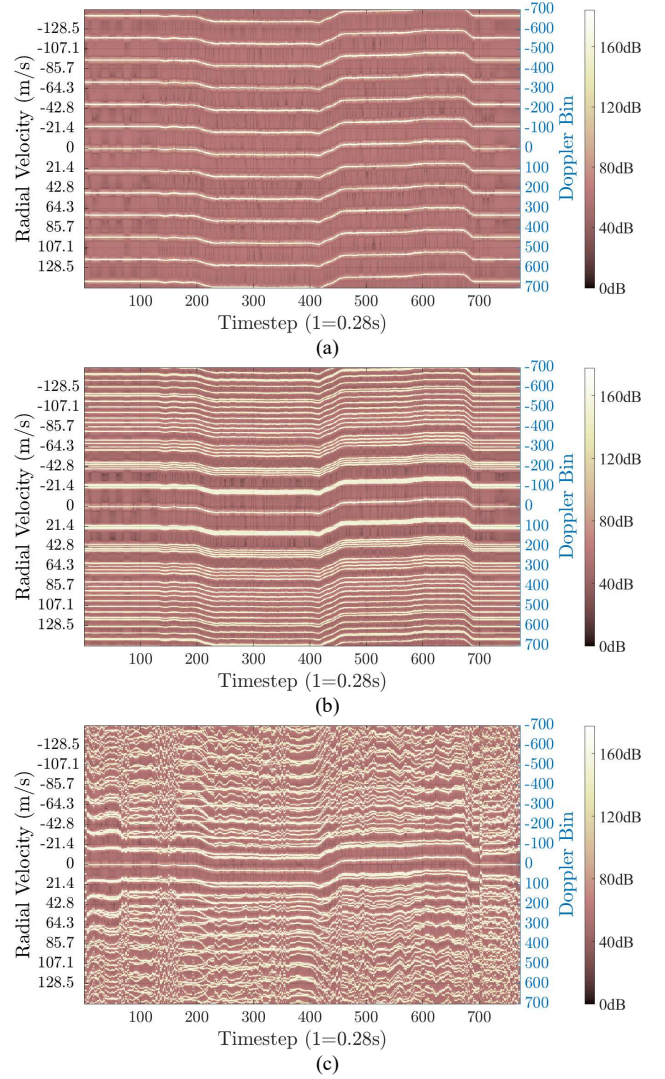
Fig. 2. HERM Line spectrum from (1) simulation model applied to a quadcopter of blade length, L=0.19 m. Motor speeds used are (a) each motor at 100 Hz, (b) four motors at 85, 90, 95 and 100 Hz, (c) using real recorded motor speeds as [34].

targets such as drones and birds [5] with both installed permanently, directed towards the densely urban and suburban Birmingham City center and having a configurable operational range of several kilometers. With these radars a large dataset of bird and drone targets have been collected for data-driven classification research. The staring radar collects range-Doppler data and performs fully digital beamforming over a 90° azimuth and 60° elevation sector altogether producing a 4D datacube (range, Doppler, azimuth and elevation) for each CPI. The radar in operation creates tracks of targets and the tracker output can be accessed for the collection of raw data for opportune targets with no ground truth. Spectrograms can be created for a trajectory within the field of view by collecting and concatenating the timeseries/fast-time samples from the resolution cell determined to contain the target.

TABLE I
OPERATING PARAMETER SUMMARY OF EMPLOYED RADAR

Parameter	Value
Operating Frequency	L-Band
Beamforming Process	Single flood illumination transmitter with coherent beamforming on receive
Bandwidth	~ 2 MHz
Pulsewidth	$\sim 1\mu s$
Transmit Power	2 kW
Receive Channels	4x16
Coverage	30° El, 90° Az
Range Resolution	~ 80 m
PRF	~ 7.3 kHz
CPI Duration	~ 0.5 s
Doppler Resolution	4 Hz
Spectrogram Formation	4096-point STFT with 50% overlap of successive CPIs

TABLE II
DRONE TARGETS USED FOR DATASET GENERATION

Target	Mass (kg)	Body Diagonal (mm)	Blade Radius (mm)	Image
DJI *Inspire* 2	3.4	605	190	
DJI *Matrice* 300	6.3	895	220	
DJI *Mini* 2	0.25	213	60	

Applying a STFT to this yields the frequency domain signal composition which allows for a detailed view of the target profile in radar systems with a high PRF, long CPI and subsequent high Doppler resolution.

Both radars contributed drone and bird data to a collective joint dataset. These commercially designed, prototype radars are the same model but are of different generations and have different gain properties and phase noise profiles. Radar #1 is 176m above mean sea level, has higher power, lower overall phase noise and extracted spectrograms are clean of artifacts and collect strong μDoppler from the large drone targets. Radar#2 is 186m above mean sea level and has lower SNR of target body and μDoppler returns as well as more frequent spectral artifacts and raised noise floors in resolution cells containing strong clutter due to the effects of phase noise [5]. These differences in the relative characteristics of the data from the two individual radar units offers a level of signature variation that provides a useful check of the robustness of the classifier and justifies the use of real data in generating the performance results reported in this paper. These radars can operate as a network [43] but only monostatic data was used in this study. Further details of the radar testbed, urban environment and radar operation can be found in [5], [17] and a parameter summary is displayed in Table I.

B. Experimental Targets

Drone and bird targets were used in this study and the methodology for the collection of their signatures at scale for deep learning purposes are explained. Major differences exist due to the nature of the targets.

1) **Drone Targets**

Table II shows a summary of the three drone targets featured in this work. Two, the DJI Inspire 2 (*Inspire*) and DJI Matrice R300 (*Matrice*), are large drones that produce strong μDoppler from their $\sim20cm$ radius, plastic blades. This measurement is similar to the radar's operational wavelength, so the near resonant scattering of the blades produces μDoppler sidebands that are often large and can come with significant fluctuations. The last drone in the dataset is the popular DJI Mini 2 (*Mini*) that is a tenth of the mass and has rotor blades that are less than a third the length of the other two drones used. This target is less likely to produce strong μDoppler sidebands, and presents a more challenging case when differentiating from bird spectrograms that as reported later are also shown to lack significant μDoppler. Drone data was collected over fourteen measurement campaigns occurring at different locations, at different times of year and in a variety of weather from calm and warm to strong winds of 17 knots and light rain. The flights were loops as large as the area of launch permitted and these were repeated in forward and reverse directions and at different heights of 100m, 80m and 60m if the location had clear line of sight to the radar. The large drones were programmed to follow waypoints and their travelling speeds were varied across different repetitions. The small *Mini* drone did not support automatic flying with waypoints and the pilot manually recreated the planned trajectory. Fig. 3a shows the GPS recording of a drone flight at a location in Birmingham within the radar field of view along with a spectrogram of the data in Fig. 3b from Radar#1 using a monochrome false color scheme for visual interpretability. In this plot the stationary clutter return is at zero Doppler and the strongest non-Zero Doppler return is that from the drone body. Offset from the body Doppler are HERM lines that are clustered in separate sidebands. Within each harmonic grouping of sidebands, up to four distinct lines

(a)

(b)

(c)

TABLE III
COMPOSITION OF TRAIN, TEST AND VALIDATION SETS

Set	Targets	#Flights	#Images	% Radar #1	% Radar #2	Mean SNR± STD (dB)	Mean STD of each flight± STD (dB)
Training	Bird	321	3625	25.8	24.8	29±9	5.5±1.6
	Inspire	46	1444	10.6	9.1	39±12	7.0±2.5
	Matrice	29	1058	7.6	6.8	41±12	7.5±2.2
	Mini	40	1129	5.1	10.4	34±9	5.7±2.0
	3625 Birds	3631 Drones	7256 Total	49.0	51.0		
Testing	Bird	146	1818	25.0	25.7	31±9	5.7±1.7
	Inspire	19	612	7.8	8.8	37±12	7.3±2.2
	Matrice	19	671	10.3	7.9	42±13	8.3±1.5
	Mini	19	535	8.4	6.1	30±11	6.0±1.6
	1818 Birds	1818 Drones	3632 Total	51.5	48.5		
Validation	Bird	24	371	36.7	15.9	32±11	5.6±1.6
	Inspire	4	127	13.8	3.9	46±11	8.0±3.7
	Matrice	4	170	19.0	4.5	50±11	9.3±1.4
	Mini	3	45	2.4	3.9	38±8	6.6±1.5
	371 Birds	342 Drones	713 Total	71.8	28.2		
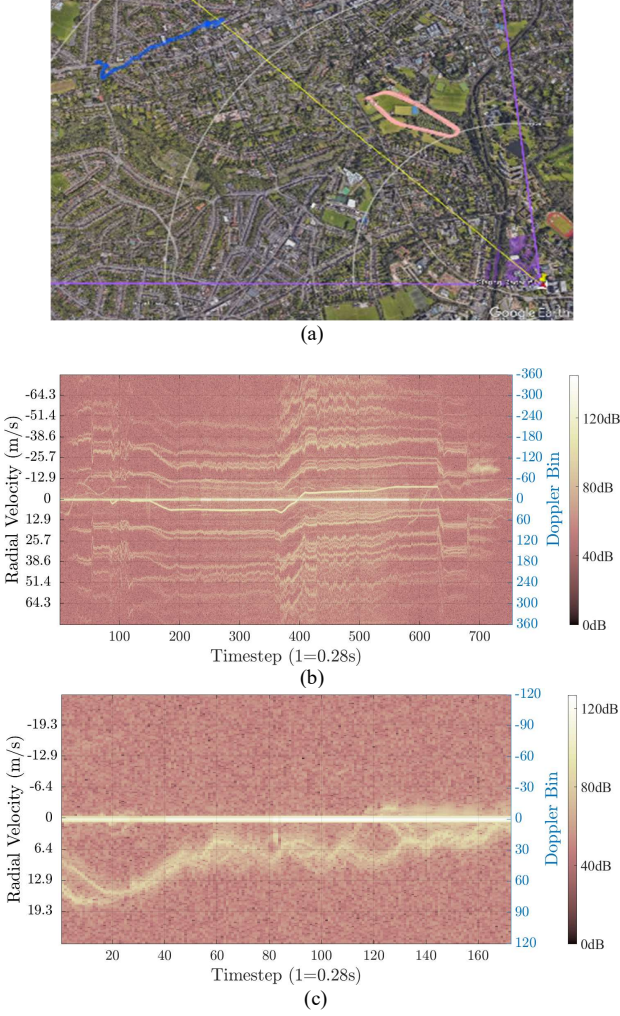
Fig. 3 (a) Example trajectory flown by *Inspire* (red, from GPS) and a bird (blue, from radar tracker) presented in Google Earth (b) resultant drone and (c) bird spectrograms from Radar#1. Radial white lines in (a) mark 1 km range increments.

are observed that are the contributions of each of the four rotors present on the drone. Occasionally all four motor contributions may be individually resolved in our setup. For example, this occurs clearly for the 1st and 2nd sidebands in Fig. 3b at timesteps 400 to 550. The frequency of observing this phenomena for any given drone will depend greatly on the radar's Doppler resolution.

2) **Bird Targets**
Bird data was collected in one of two ways to contribute to the dataset. This included data from controlled trials with GPS-tagged birds [45], as well as opportune birds, with a 3%-97% split. From the radar tracker output, the unclassified positional trajectories were filtered for track duration and target height leaving only suspected bird

targets and excluding ground vehicles. The timeseries data was then extracted for each post-filter track, then the spectrogram was formed, and each was verified by the authors to ensure tracker artefacts did not feature. The variety of bird targets was significant, featuring single birds, pairs, small flocks and occasional large flocks and mixtures of species. Returns from multiple birds in a track feature multiple peaks in the spectrogram which may appear similar to large drones' μDoppler returns. Fig. 3c shows the corresponding spectrogram for the bird target in Fig. 3a. Note the absence of any HERM lines and Doppler sidebands which is very distinct of bird targets. In this example there are at least two dominant body Doppler lines suggesting that the track consists of two birds that are flying in a pair-wise manner.

Fig. 4 shows example spectrograms taken at similar range for all the drone target types used in this study using both radars operating in isolation. Overall, the constructed dataset contains a diverse range of real spectra. The two radars although nominally similar do have some build variations. They have slightly different transmit power and their noise characteristics differ somewhat, therefore providing a greater variability in the dataset compared to that obtained from a single system. Fig. 5 shows example spectrograms of single or multiple birds, for illustration purposes and to demonstrate the variability that can be present in bird spectrograms, and further examples can be found in [44].
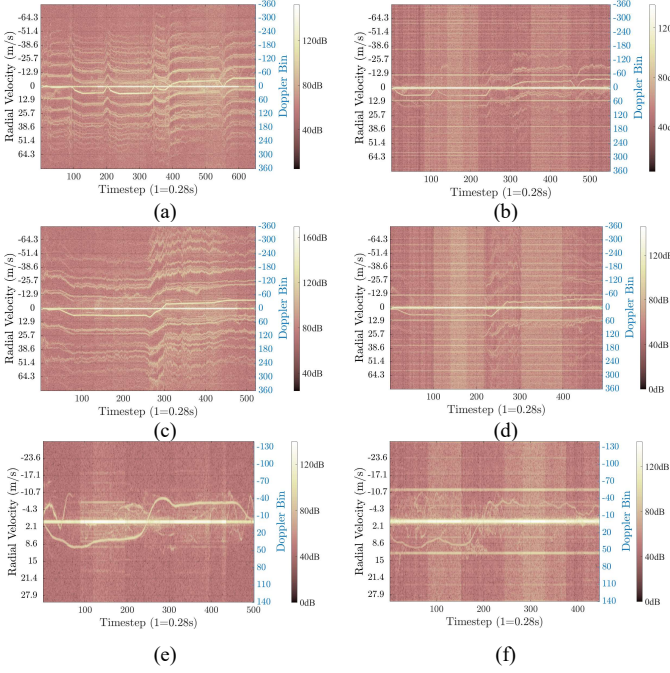
Fig. 4a-f. Example spectrograms of drones used as captured by Radar #1 (left) and Radar#2 (right) for the (a and b) *Matrice*, (c and d) *Inspire* and (e and f) *Mini*.
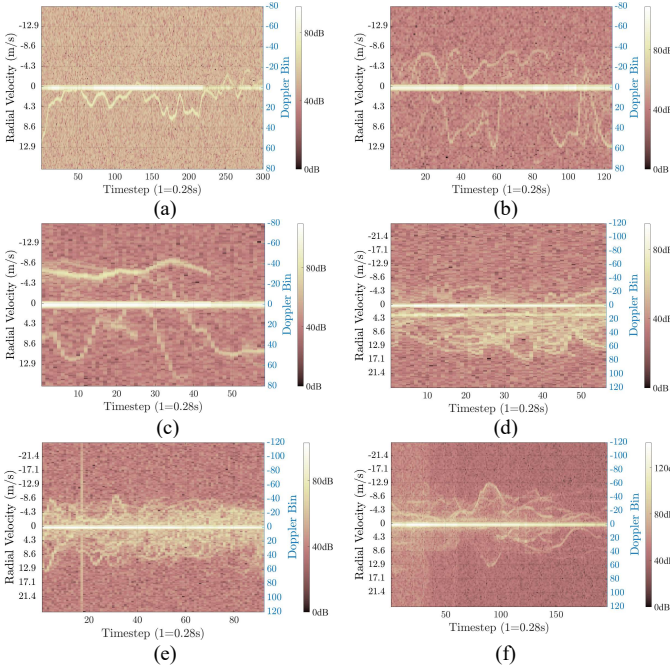
Fig. 5a-d. Example spectrograms of opportune birds from Radar#1. Number of individuals increasing in each flock from: (a) a single bird, (b) a pair, (c) a trio of birds, (d) a small flock, (e) a large, uncountable flock of birds, and (f) many birds entering and leaving tracked resolution cells.

C. The Experimental Dataset

Due to the large power of stationary clutter from the urban test environment, a 15 Hz high-pass filter was used to remove slow-moving clutter for every flight. This was done to maximize the dynamic range of target returns represented in the spectrogram color map. This removed returns with speeds less than 1.65 m/s effecting the central 18 Doppler bins. The training subsets of data were then partitioned such that data from a given flight was not split across train, test or validation subsets. The total number of 20 timestep images per flight was calculated and balanced train, test and validation subsets with a 60:30:10 ratio were formed at random using this information. The duration of each spectrogram as input to the classifier was 5.6 s. Table III shows the quantities of data used for training, validating, and testing the deep learning classifier. The train and test datasets contain nearly equal amounts of bird and drone samples. 31% of the train set and 29% of the test set was *Mini* data, which will not feature any μDoppler. The inclusion of this target in the dataset will prevent the classifier being strongly overfitted to the presence of sidebands indicating drone presence. For the baseline classification results all the data used is real radar data.

IV. BASELINE CLASSIFICATION RESULTS OF REAL DATA

In this experiment, the large CNN model AlexNet [45] was used as it has been previously shown using the same radar system, signal processing and similar targets as used in this report that it performs the best under diminishing SNR compared to other similar, large, ImageNet pretrained models. This reference [17] conducts experiments of training as a function of SNR that will be highly relevant to this study. AlexNet comprises 5 convolutional and 3 fully-connected layers, and was pretrained on a dataset of 1,000 classes and 13 million images. In both these studies, only the fully connected layers post-convolution stage were trained from random initialization. This and other large optical image models are trained on RGB images which requires 3 channels for each pixel of input, thus in this work an RGB false color scheme was used for the spectrogram images. Bayesian optimization [46] was used to find appropriate learning parameters for the training of real data, then these were offset to the nearest sensible round value for the final training. An initial learning rate of $1.5 * 10^{-4}$ with decay of 0.1 every 10 epochs was used, with the stochastic gradient descent with 0.98 momentum and $2 * 10^{-4}$ L2 regularization and a mini-batch size of 16 with samples shuffled every epoch. Training was repeated 9 times with different constant random number seeds. Model training and evaluation was performed in MatLab 2021a.

The mean performance of the trained classifier on the test set from the nine repeats is shown in the confusion matrix in Fig. 6. The real dataset achieved a 89.7 % classification accuracy, that was well balanced in regard to the rate of false positives of the two classes, with probability of false alarms being 8.9% for birds and 11.6% for drones. This demonstrates the CNN's ability to provide a robust feature extraction and classification efficacy on a challenging urban dataset with the inclusion of returns from two

Fig. 6 Confusion matrix showing average performance of real-trained baseline classifier applied to real test set.

non-identical radars. The sub-drone class rate of false alarm for large drones was 94%, and for the smaller *Mini* was 73%. The classifier does possess a bias towards classification of drones with μDoppler than those without which although undesirable is to be expected. These results obtained training the CNN on real data and testing on real data provide the baseline for comparison when the training is repeated after partially replacing the training and validation set with simulated data. Training was stopped after the validation loss exceeded the training loss 10 times, then the models at the completion of each epoch of training were tested with the training set and validation set. The final epoch used had the maximum sum of training and validation scores which ensures that the model is both well trained and minimally overfitted. An identical training procedure is used for the synthetic data investigation presented in the following section.

V. SYNTHETIC DATA GENERATION AND CLASSIFICATION RESULTS

This section details the preparation of a large training set of synthetic drone data for deep learning followed by the synthetic training results using the created datasets. Firstly, the steps to apply different motor speed sampling choices, Ω_t^m, from a real motor speed recording is shown, followed by the steps to convert the basic simulation into realistic synthetic data. This requires setting simulation parameters controlling signal power, such that the entire dataset had plausible SNR that would match the distribution of the real data to a reasonable degree. The following subsections presents a comparison of results when Ω_t^m is varied to produce four different synthetic data versions that are each used to train a CNN with each tested on the same, real dataset.

A. Procedure for Synthetic Data Creation

This subsection lays out a methodology for producing duplicate realistic synthetic spectrograms of large multi-rotor drone μDoppler, extending on work previously proposed in [34]. The construction of the synthetic data can be summarized in the following summation for each timestep, t, in a CPI,
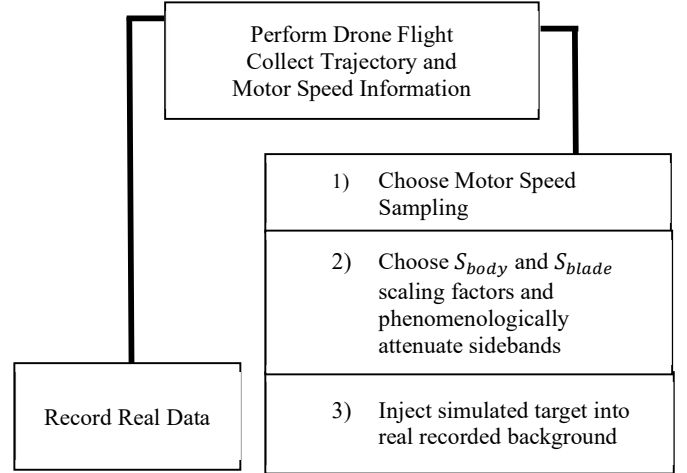
Fig. 7 Procedure for creating a synthetic drone spectrogram from a real flight used in this work.

$$X_{CPI}^{synthetic}[t] = \\ S_{Body}e^{-j\frac{4\pi R[t]}{\lambda}} + \\ S_{Blade} * X_{motors}[t] \\ \qquad * FFT(e^{-\alpha(|v|-v_{body}[t])}) + \\ X_{BG}[t] \qquad (2)$$

where X_{motors} from (1) is summed with an amplitude attenuation envelope using FFT, the fast Fourier transform operation, $X_{BG}[t]$ are the timeseries samples from the recorded background data, S_{Body} and S_{Blade} are scaling factors for the body and blade reflectors respectively, α is a hyperparameter that controls the number of HERM lines present, v is Doppler frequency and v_{body} is the central Doppler frequency of body returns. Fig. 7 shows the order of the steps to creating this data and these steps are elaborated in the following sub-sections.

1) **Motor Speed Sampling**

Equation (1) features Ω_t^m which are the motor speeds of each of the four motors assuming the use of a quadcopter. Fig. 8a displays the real spectrogram image for comparison with following synthetic examples. In flight, the motor speeds are recorded at an approximate constant rate, of 30 Hz for the *Inspire* and 20 Hz for the *Matrice*. The simplest implementation is to resample the recordings to yield a single motor speed per CPI [34] and the resultant motor speed signal is shown in Fig. 8b along with the synthetic spectrogram result in Fig. 8c. We used a piecewise cubic spline interpolation [47] to uniformly under-sample the motor speed signal at each CPI. This approach has limitations in its realism as motor speeds are not coupled across overlapping CPIs, resulting in motor returns jumping discontinuously across Doppler bins in successive CPIs (see timesteps ~100). With comparison to real data (Fig. 8a), the sidebands in the simulation are overly clean and are of uniform spectral width. In effort to improve on this, the motor speeds were sampled to less than the duration of the CPI. At first, the motor speeds were down-sampled (from

(a)

(b)

(c)

(d)

(e)
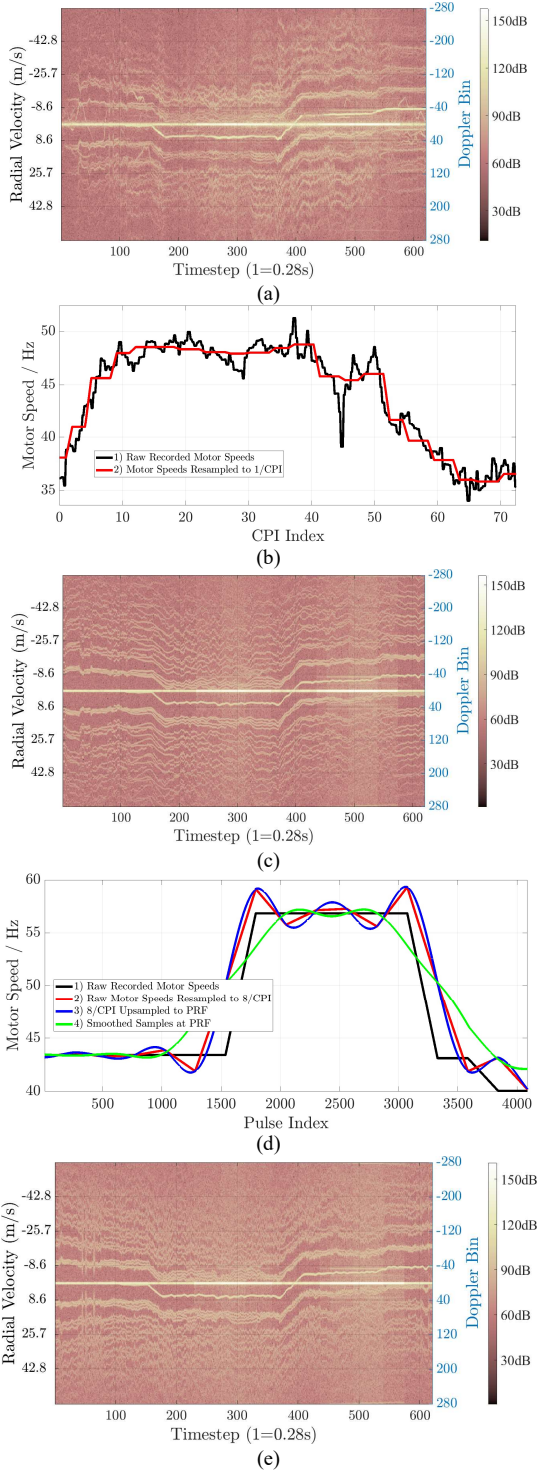
Fig. 8a-e Motor Speed Sampling and Resultant Synthetic Spectrograms. (a) Real *Matrice* Spectrogram, (b) Resampling of the raw motor speed recordings (black) to 1 value per CPI (red) using a cubic interpolating spline, (c) Resultant, Per-CPI Synthetic Spectrogram, (d) Sub-CPI Sampling Procedure. Original samples (black) at 30Hz are decimated to 8/CPI (red) before resampling to PRF (blue) and finally after applying smoothing with a 2500-point rolling average (green). (e) Resultant, Sub-CPI Synthetic Spectrogram.

$30\ Hz$) via decimation to 8/CPI ($\sim14.3\ Hz$), and then up-sampled by an integer factor (of 4096/8 = 128) to PRF ($\sim7.3\ kHz$) via time domain 0-padding to required signal length and application of a low-pass interpolating filter. Appling the present motor speed signal to (1) followed by FFT produced HERM lines that were spread widely over the spectrum with insufficient integration gain. Smoothing these values significantly using a rolling-average with a long window length (between approximately 1,000 to 3,500 compared to the total CPI length of 4096) produced highly convincing sidebands with more realistic bandwidth and better continuity across CPIs. These steps and their effect on the motor speed signal are shown is pictorially in Fig. 8d and the resultant spectrogram is shown in Fig. 8e. These two models will be referred to as the per-CPI and sub-CPI synthetic versions and are both high-fidelity choices for motor speed sampling that will be used in the following synthetic training experiments. Wind will have a significant impact on a drone's motor speeds which will be preserved in the recordings to an extent depending on the drone's sampling rate. Rapid changes in motor speeds over a CPI induced by wind changes or to enable a maneuver, manifested as artifacts in the per-CPI spectrogram as discontinuities in the spectral lines which was improved in the sub-CPI implementation.

2) Artificial Sideband Modulation and Simulated Component Scaling

The model outlined in Section II treats the drone blades as a series of infinite point reflectors. It produces accurate frequency modulation of returns but very poorly captures the amplitude modulation of the sidebands, especially in our setup where the size of the radiation wavelength and drone blade are similar. The thin wire simulation model as used produces sidebands that are too strong and too numerous in comparison with the real data. At short ranges from the radar the two larger drone models exhibit typically 3-5 pairs of sidebands, with a SNR reduction of 5-10 dB compared to the main body and 3-8 dB for each subsequent sideband harmonic, whereas the thin wire simulation shows up to 10 pairs of sidebands with very little fall off in signal power until the final harmonics. To remedy this, the sidebands of the simulation were artificially attenuated to approximately match the number and signal power fall off observed in real data. A back-to-back exponential decay attenuation envelope in frequency centered around the main body return of the form $f(v) = e^{-\alpha(|v|-v_{body})}$ was Fourier transformed to the timeseries domain then summed element wise with the timeseries simulation [34]. The α parameter was used to tune the strength and number of sidebands, reducing their presence as α →0. With this, the factors S_{Body} and S_{Blade} were applied to ensure the body and blade SNR in the simulated data was similar to that of the experimental data.

3) Injection into Real Background

A simple way to add realistic noise background to the simulated spectrogram is to form a spectrogram using real data with the same flight trajectory as the simulated target but from a time instance when there is no target present in the real data.
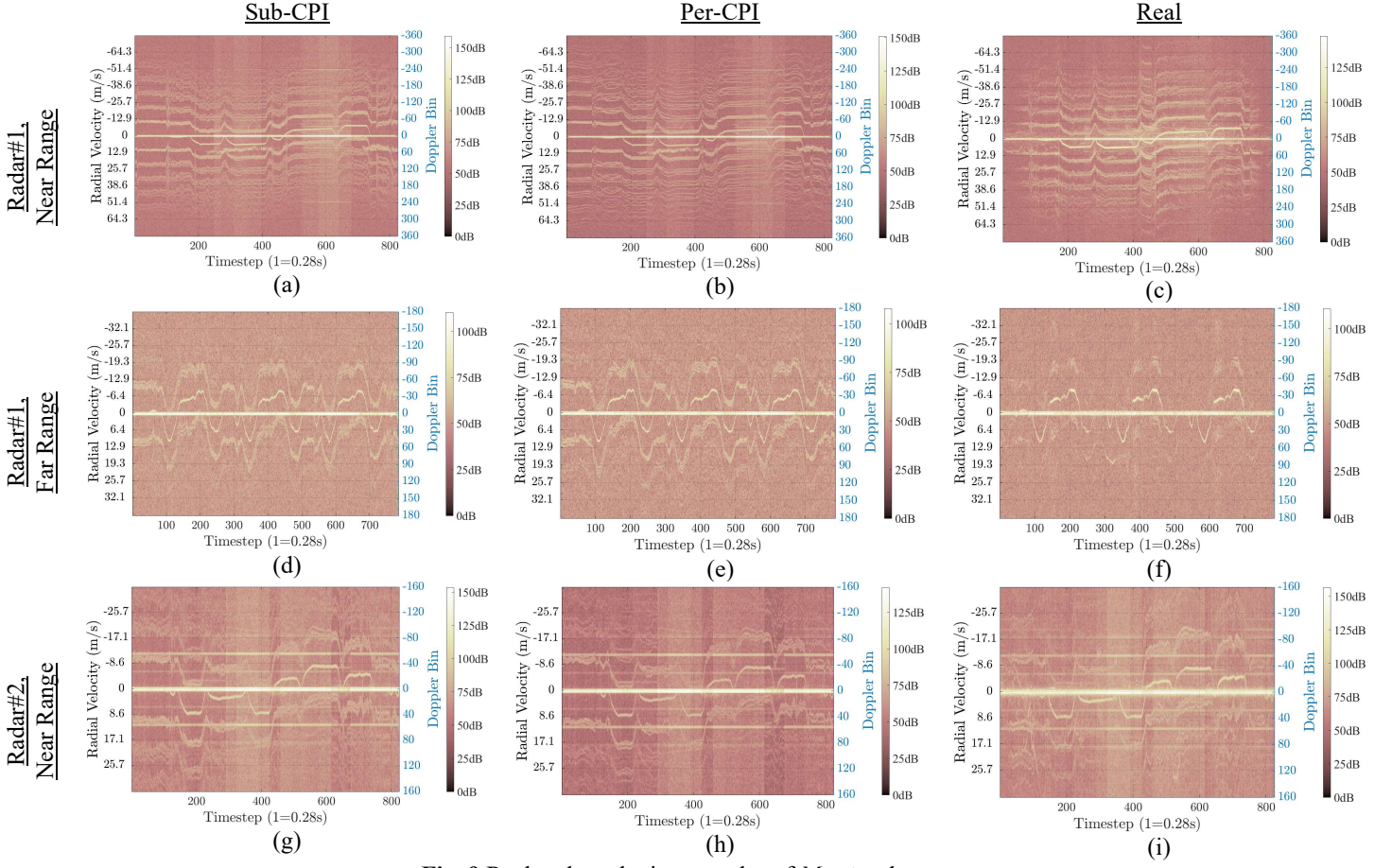
Fig. 9 Real and synthetic examples of *Matrice* drone.
(a-c) Near range at 1.4 km with Radar#1; (a) Sub-CPI (b) Per-CPI (c) Real
(d-f) Far range at 2.7 km with Radar#1; (d) Sub-CPI (e) Per-CPI (f) Real
(g-i) Near Range at 1.4 km with Radar#2; (g) Sub-CPI (h) Per-CPI (i) Real

TABLE IV
SIMULATION PARAMETERS USED, WHERE SHORTHAND $2E^{-3} \equiv 2 \times 10^{-3}$

		Matrice	*Inspire*	*Matrice*	*Inspire*	*Matrice*	*Inspire*
Synthetic Version	Motor Sampling Function, Ω_t^m	S_{body} Radar#1 Min/Max, Radar#2 Min/Max		S_{blade}, Radar#1 Min/Max, Radar#2 Min/Max		α Radar#1 Min/Max, Radar#2 Min/Max	
Sub-CPI	Real Speeds sampled to PRF/8 & Long Window Rolling Average	$E^{-5}/2.5E^{-1}$ $2E^{-2}/.25$	$2E^{-2}/1$ $2E^{-2}/2E^{-1}$	$5/40$ $5/35$	$7/32$, $5/32$	$5.5E^{-3}/1.1E^{-3}$, $7E^{-3}/1.7E^{-2}$	$5.5E^{-3}/1.1E^{-2}$, $5.5E^{-3}/8E^{-2}$
Per-CPI	Real Speeds sampled to CPI	$E^{-5}/1.4E^{-3}$, E^{-10}/E^{-7}	$E^{-13}/5E^{-2}$, $E^{-12}/2E^{-7}$	$5/32$ $7/30$	$7/30$, $10/30$	$5.5E^{-3}/$ $1.1E^{-2}$, $2.7E^{-3}/E^{-2}$	$E^{-2}/6E^{-3}$, $E^{-2}/7E^{-3}$
Single-Motor	As sub-CPI but only one motor	As sub-CPI					
Fixed-Speeds	$\Omega_t^m = \Omega$ 60 *Hz Matrice* / 80 *Hz Inspire*	As per-CPI					

(a)

(b)

(c)

(d)

(e)
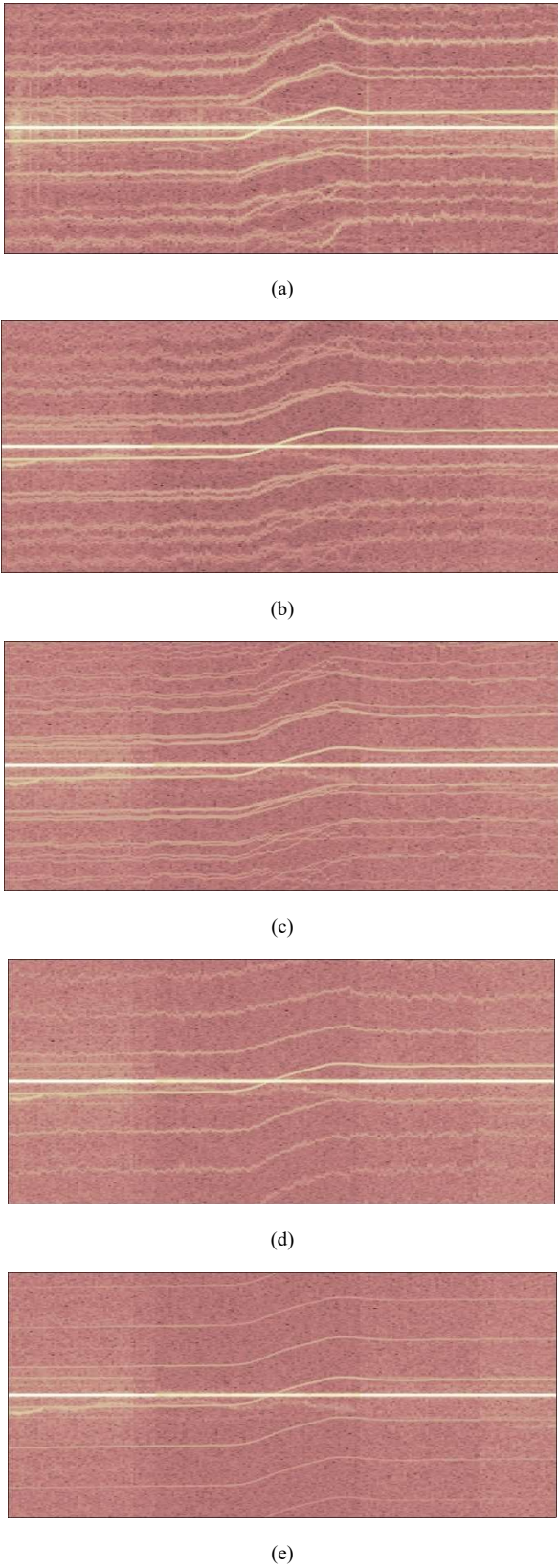
Fig. 10 Fifty-Six seconds of Real and Synthetic models for the *Inspire* at range 2.5 km with Radar#1. The central 480 Doppler bins are shown. (a) Real (b) Sub-CPI (c) Per-CPI (d) Single Motor (e) Fixed Speeds.

(a)   (b)   (c)   (d)   (e)

(f)   (g)   (h)   (i)   (j)
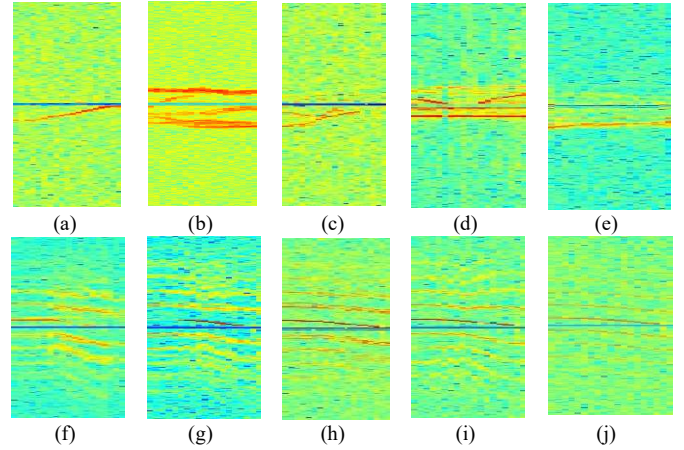
Fig. 11. Example high-SNR images from Radar #1 as input to the CNN Classifier. (a-e) Five Real Bird examples, (f) Real Drone, (g) Sub-CPI, (h) Per-CPI, (i) Single-Motor (j) Fixed-Speeds

Thus, the background for the synthetic target was extracted by time-shifting the real flight to a time when there was no control targets present. Extracting the raw data in the same way produces a spectrogram with the same local (but not time-dependent) clutter properties as the real flight for each timestep of the flight. The simulated target's 1-D timeseries returns are summed elementwise with the background signal - extracted, injecting the simulated target into a real background. This approach was taken to maximize the similarity of the real and synthetic data background properties, but this is limited as the real data came from regular intervals over a two-year period and the background data came from a single period in March 2022. The variation of weather and wind on clutter is present in the real data set but not as much for the synthetic cases. The weather conditions and locations could be altered to yield varied backgrounds for training data but this lies outside the scope of the current study.

B. Experimental Synthetic Datasets

For an accurate assessment of synthetic drone μDoppler for training of CNNs compared to a real baseline, a dataset was constructed so that it was as close to the collected real returns as possible. In the ideal case, the SNR of sidebands would be dictated by the blade's radar cross section multiplied by the number of blades and the coherent integration factor with range scaling to the negative fourth power. However, drone μDoppler as found in our collected data does not always behave as expected. For example, Figures 3b and 9f have a large fluctuation regarding the symmetry and strength of the sidebands along the flight, while by contrast Fig. 9c shows clean and consistent μDoppler for the whole flight. With the goal of creating a dataset that contained plausible synthetic μDoppler within the bounds of the real dataset, the solution implemented in this work was to scale the power of body and the power and count of sidebands to be between the nearest and furthest range real examples. This was done by setting α to match the number of sidebands as found in real data, then modifying S_{Blade} to match the real and synthetic SNR of the
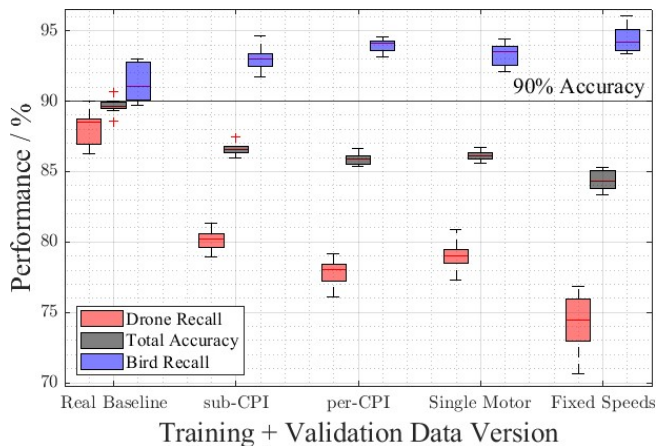
Fig. 12. Total accuracy and bird/drone recall performances on the real test set achieved with the real and four synthetically trained models presented via a box plot. Total accuracy (black), bird recall (blue) and drone recall (red) are represented in the boxes, where the top and bottom- edges of the box are the upper and lower quartiles of values represented. The red line within each box is the median result. Results that fall beyond 1.5 times the interquartile range above or below the nearest quartile are marked as outliers with a red plus, otherwise the connected horizontal bar indicates the most extreme result.

first sideband to within 1-2dB, repeating until both criteria were satisfied. This process was repeated for both the large drone targets, for both the per-CPI and sub-CPI versions, and for both radars. The values yielded from this approach are presented in Table IV. This method has limitations that will prevent perfectly matching SNR across all sidebands for the real and synthetic datasets. Imperfections in the synthetic SNR recreation are mitigated in the following experiment by using the RGB colormap, and from the dividing and shuffling of samples for training. Thus, we have preserved the frequency modulation of the HERM lines to be the largest difference between any of the datasets. The sub-CPI version also had an intra-CPI rolling average length factor to be decided. This was varied between different flights randomly between 1500 and 3000 which were found to produce sidebands matching those of the corresponding real spectrograms. Fig. 9 shows the real and synthetic per-CPI and sub-CPI comparison of the *Matrice* drone, at long and short range for Radar#1, and at short range for Radar#2. At the longest ranges, Fig. 9d-f, the real sidebands fluctuated in their appearance and power. The simulation model presented in this paper does not attempt to model such fluctuations and synthetic parameter values were set so that sidebands were always visible in the synthetic data. As well as the sub-CPI and per-CPI models, two other lower fidelity synthetic datasets were constructed that are simpler representations of the motor speeds. Firstly, the sub-CPI model was simplified to have just one rotor instead of four. This is referred to as the Single-Motor model. The second simplification was to use a single speed for all four rotors. This last model provides context for this simulation approach

as reported in the literature [19], [24], [32], and will be referred to in this paper as the Fixed-Speeds model. The fixed rotational frequency was constant across the flight and sensible values within a range typical of the drone model was randomly chosen for each flight. The five classifiers trained in the following section have their spectrograms shown for the *Inspire* and Radar#1 in Fig. 10a-e, and Table IV lists the parameters and motor sampling used in their creation.

C. Synthetic Training Results and Comparison

The CNN classifier trained on real data and data from the four synthetic models all followed identical training procedures as described in section 2.B. Fig. 11 shows ten exemplar images as were used to train the classifier. The duration of each is 20 timesteps which equates to 5.8 s. The repeated training on the real and different synthetic versions each occurred with the same random seeding as each other and when training on real data. The results on the tests with real data are summarized for all training instances via box plot in Fig. 12. The total accuracy is the ratio of correct to total classifications. The recall is the ratio of true positives to true positives plus false negatives and is a measure of how trustworthy a classification result of a particular target is, i.e., a 90% drone recall means 90% of predicted drones were truthfully drones. A key metric to observe is the difference of the drone and bird recalls which are present in Fig. 12 in red and blue respectively either side of the accuracy. Instantly we observe that for all the datasets, the recalls are significantly unbalanced towards birds. This means that models were biased to classify challenging edge-case samples in the test set as birds to differing degrees as discussed below. Table V shows training result figures of merit for each simulation version and the mean accuracies of the large, with μDoppler, drone targets, and the small, μDoppler absent, *Mini* drone target.

The real baseline observed a mean value of 89.7% for total accuracy, and drone recall (88.0%) was worse than the bird recall (90.5%) by 2.5% which is a reasonably balanced performance, slightly biased towards selecting edge cases as birds. The best performing synthetic model was the sub-CPI version, suffering a total accuracy decrease of 3.1% to 86.6%. The per-CPI and Single-Motor synthetic versions were ~1.5% lower at 85.9% and 86.1% respectively, and the Fixed-Speeds model's mean total accuracy fell a further 1.4% to 84.4%. The median and spread of accuracy of the synthetic models are approximately identical for the multi rotor sub-CPI and per-CPI and single motor sub-CPI model, but the performance of

TABLE V
SUMMARY OF SYNTHETIC TRAINING RESULTS

Training and Validation Data	Mean Accuracy	Mean Drone Recall	Mean Bird Recall	Mean Large Drone Accuracy	Mean Small Drone Accuracy
Real Baseline	**89.7 ± 0.5**	**88.0 ± 1.2**	**90.5 ±1.3**	**94.0 ± 1.1**	**73.3 ±2.2**
Sub-CPI	86.6 ± 0.5	80.1 ± 0.8	93.0 ± 0.9	85.4 ±0.7	67.3 ±1.4
Per-CPI	85.9 ± 0.4	77.8 ± 1.0	93.9 ± 0.5	83.0 ±1.0	65.4 ±2.1
Single-Motor	86.1 ± 0.1	79.0 ± 1.0	93.3 ± 0.8	84.2 ±1.0	66.6 ± 1.8
Fixed-Speeds	84.4 ± 0.7	74.3 ± 2.0	94.5 ± 1.0	78.3±2.6	64.8 ± 2.2

the fixed-speeds fell significantly further demonstrating that motor speed variation is an essential component of training. Observing the drone recall as a function of the μDoppler fidelity, the reliability of drone classifications decreased rapidly to a mean value of 74.5% for the least accurate, fixed speeds model. The drone recalls of the two multi-rotor models incorporating all four motor speeds medians fell from the real baseline of 88.4% to 80.1% (sub-CPI) and 78.8% (per-CPI). The Single-Motor version's drone recall was between these values at 79.0%. Each of these models had mean drone recalls to within 2.1% of each other, with the final, fixed speeds model considerably worse at 3.5% from the next best recorded recall.

The decrease in drone recall against training data fidelity signifies the model is becoming overfitted to the shape of the synthetic sidebands and struggles to recognize real sidebands as effectively. The degradation of drone recall with reduced synthetic data fidelity shows that the more realistic the simulated data, the more accurate the classifier is, crucially producing less drone false positives. Bird recall increased for the two multi-speed multi-rotor models from the mean real baseline of 90.5% to 93.0% for sub-CPI and 93.9% for per-CPI versions. Again, the Single-Motor version sat between these values at 93.3%, with the lowest fidelity Fixed-Speeds version ~1% above the alternatives. This indicates that synthetic training with both motor speed variation and a multi-rotor model minorly reduces the models bias towards birds. As the synthetic fidelity decreases, the bird recall gets higher, thus showing that the model is biased towards predicting difficult targets as birds, which is an operationally ineffective outcome. Table V also displays the accuracies of the small and large drones in our dataset. The *Mini* drone was always misclassified significantly more frequently than the large μDoppler producing targets. We observe that the small drone accuracy decreases across the datasets at a similar rate as the overall drone recall despite itself not changing in the training sets.

VI. Discussion

When creating synthetic drone μDoppler, we have highlighted the importance of the selection of motor speeds. Visually, accessing recordings of real motor speed values and implementing them in simulation makes a great improvement and subsequent comparisons with equivalent real data become highly informative compared to using constant motor speed values. For synthetic training, we have shown that motor variation is imperative, as classifier accuracy and balance were worse when using the Fixed-Speeds model. We presented per-CPI and sub-CPI motor speed sampling as two choices if using real or estimated motor speeds. Sub-CPI sampling improved the appearance of individual sidebands visually to have a non-uniform bandwidth and better continuity in time. It was shown synthetic training with the sub-CPI sampling outperformed the per-CPI very slightly. This was a first step towards even more realistic HERM frequency modulation, but further work is required to understand the phenomenology of sideband power fluctuations and asymmetricities around the body return which was not attempted in this work. Such effects could be simulated with some further signal processing

choices, for example, setting sideband attention to be stochastic and non-symmetrical.

In this study, we strove to keep the SNR of real and synthetic data as similar to each other as possible to understand the utility of directly replacing real data with realistic synthetic data. Having demonstrated the promise of multi-speed multi-rotor synthetic data, future work can explore the benefit of varying the synthetic parameters beyond the extent used here. To robustly classify edge-case and reliably classify unseen drone models, greater variation of the simulation and injected background would train classifiers with a wider scope for interpolating unseen measurements amongst its training samples. Further research demonstrating synthetic training using a wide and varied synthetic data will be required to inform operators of the trade-off between access to large volumes of synthetic data and how varied the signatures must be compared to the use of small, real datasets. Indeed, how performance is affected when using a mix of real and synthetic data is not known but would be expected to offer a significant performance improvement with respect to the limited quantities of real data available as the variety of training data will be larger.

VII. Conclusion

In this work, we presented a method for simulating realistic μDoppler returns for a rotary winged drone and showed the effect of simulation fidelity on classifier performance using a large, challenging dataset with comparison to an equivalent real benchmark of bird and drone targets. Synthetic datasets with increasing degrees of realism recreating large drone flights were produced using the trajectory and motor speed information recorded during the real flight, including setting SNR and involving the effects of urban clutter. Four synthetic versions were used to illustrate synthetic training capabilities and these differed in the motor speed sampling and motor configuration used in their creation. The best result achieved with the synthetic dataset was 86.6% accuracy compared to the real data benchmark that stood at 89.7%. The accuracy and balance between predicted classes of the synthetically trained classifier diminished as the fidelity of the simulation decreased. The largest performance deficit came from the Fixed-Speeds model at 84.4% accuracy. In conclusion, it has been found that multi-rotor modelling and motor speed variation are both required to prevent sub-optimal training producing a classifier with a high rate of false positives.

References

[1] J. Wang, Y. Liu and H. Song, "Counter-unmanned aircraft system(s) (C-UAS): State of the art, challenges, and future trends," in *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 3, pp. 4-29, 1 March 2021, doi: 10.1109/MAES.2020.3015537.

[2] J. A. Besada, I. Campaña, D. Carramiñana, L. Bergesio, and G. de Miguel, "Review and simulation of counter-UAS sensors for unmanned traffic management," *Sensors*, vol. 22, no. 1, p. 189, Dec. 2021, doi: 10.3390/s22010189.

[3] A. Coluccia, G. Parisi, and A. Fascista, "Detection and classification of multirotor drones in radar sensor networks: A review," *Sensors*, vol. 20, no. 15, p. 4172, Jul. 2020, doi: 10.3390/s20154172.

[4] C. Clemente, F. Fioranelli, F. Colone and G. Li, "Radar UAV and bird signature comparisons with micro-Doppler," in *Radar Countermeasures*

for Unmanned Aerial Vehicles, London, SciTech Publishing, 2021, ch. 9, pp. 279–314.

[5] M. Jahangir et al., "Networked staring radar testbed for urban surveillance: status and preliminary results," in *Proc. Int. Conf. Radar Syst.*, Edinburgh, UK, 2022, pp. 471-476, doi: 10.1049/icp.2022.2363.

[6] N. Mohajerin, J. Histon, R. Dizaji and S. L. Waslander, "Feature extraction and radar track classification for detecting UAVs in civillian airspace," in *IEEE Radar Conf.*, Cincinnati, OH, USA, 2014, pp. 674-679, doi: 10.1109/RADAR.2014.6875676.

[7] M. Ritchie, F. Fioranelli, H. Borrion, and H. Griffiths, "Multistatic micro-Doppler radar feature extraction for classification of unloaded/loaded micro-drones," *IET Radar, Sonar Navig.*, vol. 11, no. 1, pp. 116–124, 2017, doi: 10.1049/iet-rsn.2016.0063.

[8] M. Jahangir, B. I. Ahmad and C. J. Baker, "Robust drone classification using two-stage decision trees and results from SESAR SAFIR trials," in *IEEE Int. Radar Conf.*, Washington, DC, USA, 2020, pp. 636-641, doi: 10.1109/RADAR42522.2020.9114870.

[9] L. Pallotta, C. Clemente, A. Raddi and G. Giunta, "A feature-based approach for loaded/unloaded drones classification exploiting micro-Doppler signatures," in *Proc. IEEE Radar Conf.*, Florence, Italy, 2020, pp. 1-6, doi: 10.1109/RadarConf2043947.2020.9266458.

[10] P. Klaer et al., "An investigation of rotary drone HERM line spectrum under manoeuvering conditions," *Sensors*, pp. 11–15, 2020, doi: 10.3390/s20205940.

[11] M. S. Seyfioglu, B. Erol, S. Z. Gürbüz and M. G. Amin, "DNN transfer learning from diversified micro-Doppler for motion classification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2164-2180, Oct. 2019, doi: 10.1109/TAES.2018.2883847.

[12] N. Inkawhich et al., "Bridging a Gap in SAR-ATR: Training on Fully Synthetic and Testing on Measured Data," in *IEEE J. Sel. Topics Appl. Earth Observ. and Remote Sens.*, vol. 14, pp. 2942-2955, 2021, doi: 10.1109/JSTARS.2021.3059991.

[13] J. Ding, B. Chen, H. Liu, and M. Huang, "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 364–368, 2016, doi: 10.1109/LGRS.2015.2513754.

[14] S. Zaied, A. Toumi, and A. Khenchaf, "Target classification using convolutional deep learning and auto-encoder models," in *4th Int. Conf. Adv. Technol. Signal Image Process.*, pp. 1–6, 2018, doi: 10.1109/ATSIP.2018.8364502

[15] B. Xue and N. Tong, "Real-world ISAR object recognition using deep multimodal relation learning," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4256–4267, 2020, doi: 10.1109/TCYB.2019.2933224.

[16] S. Rahman and D. A. Robertson, "Classification of drones and birds using convolutional neural networks applied to radar micro-Doppler spectrogram images," *IET Radar, Sonar Navig.*, vol. 14, no. 5, pp. 653–661, 2020, doi: 10.1049/iet-rsn.2019.0493.

[17] H. Dale, C. Baker, M. Antoniou, M. Jahangir, G. Atkinson, and S. Harman, "SNR-dependent drone classification using convolutional neural networks," IET Radar, Sonar Navig., vol. 16, no. 1, pp. 22–33, 2022, doi: 10.1049/rsn2.12161.

[18] A. Karlsson, M. Jansson and M. Hämäläinen, "Model-aided drone classification using convolutional neural networks," in *Proc. IEEE Radar Conf.*, New York City, NY, USA, 2022, pp. 1-6, doi: 10.1109/RadarConf2248738.2022.9764194.

[19] B. I. Ahmad, J. Grey, M. Newman and S. Harman, "Low-latency convolution neural network for estimating drone physical parameters with radar," in *Int. Conf. on Radar Sys.*, Edinburgh, UK, 2022, pp. 1-6, doi: 10.1049/icp.2022.2282.

[20] N. Regev, I. Yoffe, and D. Wulich, "Classification of single and multi propelled miniature drones using multilayer perceptron artificial neural network," *IET Conf. Publ.*, vol. 2017, no. CP728, pp. 1–5, 2017, doi: 10.1049/cp.2017.0378.

[21] J. S. Patel, C. Al-Ameri, F. Fioranelli, and D. Anderson, "Multi-time frequency analysis and classification of a micro-drone carrying payloads using multistatic radar," *J. Eng.*, vol. 2019, no. 20, pp. 7047–7051, 2019, doi: 10.1049/joe.2019.0551.

[22] B. K. Kim, H. S. Kang, and S. O. Park, "Drone classification using convolutional neural networks with merged doppler images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 38–42, 2017, doi: 10.1109/LGRS.2016.2624820.

[23] B. Choi and D. Oh, "Classification of drone type using deep convolutional neural networks based on micro-Doppler simulation," in *2018 Int. Symp. Antennas Propag.*, Busan, Korea (South), 2018, pp. 1-2.

[24] D. Raval, E. Hunter, S. Hudson, A. Damini, and B. Balaji, "Convolutional neural networks for classification of drones using radars," *Drones*, vol. 5, no. 4, p. 149, Dec. 2021, doi: 10.3390/drones5040149.

[25] S. Yoon *et al.*, "Efficient protocol to use FMCW radar and CNN to distinguish micro-Doppler signatures of multiple drones and birds," in *IEEE Access*, vol. 10, pp. 26033-26044, 2022, doi: 10.1109/ACCESS.2022.3155776.

[26] J. Duan, L. Zhang, Y. Wu, Y. Zhang, Z. Zhao and X. Guo, "Classification of birds and drones by exploiting periodical motions in Doppler spectrum series," *in J. Sys. Eng. and Electron.*, vol. 34, no. 1, pp. 19-27, February 2023, doi: 10.23919/JSEE.2023.000002

[27] Y. Gao, Y. Zhou, Y. Wang and Z. Zhuo, "Narrowband radar automatic target recognition based on a hierarchical fusing network with multidomain features," in *IEEE Geosci. and Remote Sens. Letters*, vol. 18, no. 6, pp. 1039-1043, June 2021, doi: 10.1109/LGRS.2020.2993039.

[28] N. Ashush, S. Greenberg, E. Manor, and Y. Ben-Shimol, "Unsupervised drones swarm characterization using RF signals analysis and machine learning methods," *Sensors*, vol. 23, no. 3, p. 1589, Feb. 2023, doi: 10.3390/s23031589.

[29] M. Bauw, S. Velasco-Forero, J. Angulo, C. Adnet, and O. Airiau, "Near out-of-distribution detection for low-resolution radar micro-Doppler signatures," arXiv:2205.07869 [eess.SP], Jul. 2022.

[30] M. Moore, D. A. Robertson and S. Rahman, "Simulating UAV micro-Doppler using dynamic point clouds," in *Proc. IEEE Radar Conf.*, New York City, NY, USA, 2022, pp. 1-6, doi: 10.1109/RadarConf2248738.2022.9764284.

[31] L. Lehmann and J. Dall, "Simulation-based approach to classification of airborne drones," in *Proc. IEEE Radar Conf.*, Florence, Italy, 2020, pp. 1-6, doi: 10.1109/RadarConf2043947.2020.9266405.

[32] Y. Cai, O. Krasnov and A. Yarovoy, "Simulation of radar micro-Doppler patterns for multi-propeller drones," in *Proc. Int. Radar Conf.(RADAR)*, Toulon, France, 2019, pp. 1-5, doi: 10.1109/RADAR41533.2019.171372.

[33] N. Rojhani, M. Passafiume, M. Sadeghibakhi, G. Collodi and A. Cidronali, "Model-based data augmentation applied to deep learning networks for classification of micro-Doppler signatures using FMCW radar," *IEEE Trans. Microw. Theory and Techn.*, vol. 71, no. 5, pp. 2222-2236, May 2023, doi: 10.1109/TMTT.2023.3231371.

[34] D. White et al., "Multi-rotor drone micro-Doppler simulation incorporating genuine motor speeds and validation with L-band staring radar," in *Proc. IEEE Radar Conf.*, New York City, NY, USA, 2022, pp. 1-6, doi: 10.1109/RadarConf2248738.2022.9764352.

[35] J. Martin and B. Mulgrew, "Analysis of the theoretical radar return signal form aircraft propeller blades," in *Proc. IEEE Int. Conf. Radar*, Arlington, VA, USA, 1990, pp. 569-572, doi: 10.1109/RADAR.1990.201091.

[36] J. Markow and A. Balleri, "Examination of drone micro-Doppler and JEM/HERM signatures," in *Proc. IEEE Radar Conf.*, Florence, Italy, 2020, pp. 1-6, doi: 10.1109/RadarConf2043947.2020.9266342.

[37] P. Beasley et al., "Multistatic radar measurements of UAVs at X-band and L-band," in *Proc. IEEE Radar Conf.*, Florence, Italy, 2020, pp. 1-6, doi: 10.1109/RadarConf2043947.2020.9266444.

[38] S. Zulkifli and A. Balleri, "FMCW radar prototype development for detection and classification of nano-targets," in *Proc. IEEE Int. Radar Conf.*, Washington, DC, USA, 2020, pp. 738-743, doi: 10.1109/RADAR42522.2020.9114824.

[39] V. V. Reddy and S. Peter, "UAV micro-Doppler signature analysis using FMCW radar," in *Proc. IEEE Radar Conf.*, Atlanta, GA, USA, 2021, pp. 1-6, doi: 10.1109/RadarConf2147009.2021.9454978.

[40] R. I. A. Harmanny, J. J. M. de Wit and G. P. Cabic, "Radar micro-Doppler feature extraction using the spectrogram and the cepstrogram," in *Proc. 11th Eur. Radar Conf.*, Rome, Italy, 2014, pp. 165-168, doi: 10.1109/EuRAD.2014.6991233.

[41] C. Bennett, S. Harman and I. Petrunin, "Realistic simulation of drone micro-Doppler signatures," in *Proc. 18th Eur. Radar Conf.*, London, United Kingdom, 2022, pp. 114-117, doi: 10.23919/EuRAD50154.2022.9784488.

[42] J. Gérard, "Drone recognition with deep learning," Ph.D. dissertation, LISN, Univ. Paris-Saclay, 2022.

[43] D. Griffiths, M. Jahangir, J. Kannanthara, C. J. Baker, M. Antoniou and Y. Singh, "Direct signal synchronization for staring passive bistatic radar," in *Proc. Int. Conf. Radar Sys.*, Edinburgh, UK, 2022, pp. 220-225, doi: 10.1049/icp.2022.2319.

[44] G. Atkinson et al., "The role of target signatures in bird-drone classification," in *Proc. IEEE Radar Conf.*, San Antonio, TX, 2023. To be published.

[45] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM 60*, pp. 84–90, doi: 10.1145/3065386.

[46] J. Wu, X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, and S. H. Deng, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *J. Electron. Sci. Technol.*, vol. 17, no. 1, pp. 26–40, 2019, doi: 10.11989/JEST.1674-862X.80904120.

[47] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26 , no. 6, pp. 508–517, 1978, doi: 10.1109/TASSP.1978.1163154.

Daniel White received an MSci Degree in Physics from the University of Birmingham, Birmingham, UK in 2020, and is presently undergoing a Ph.D. degree with the department of Electronic, Electrical and Systems Engineering at UoB as part of the Microwave Integrated Systems Laboratory (MISL) group. His research interests include machine learning, drone technology and radar systems. Mr. White received 3rd place in the student paper competition at their first conference attendance at the IEEE International Radar Conference 2022.

Mohammed Jahangir is the lead senior research fellow at the University of Birmingham Quantum Technology Hub Timing and Navigation programme. He is responsible for the development of a fully coherent quantum enabled networked staring radar. Having graduated with 1st class honors from Imperial College he joined the UK MoD research centre in Worcestershire to embark on a research career on surveillance radars. He obtained his PhD from UCL in 2000 on coherent radar clutter statistics. He has three decades experience of working on many aspects of surveillance systems for both defence and the civil sector. He has widely collaborated with academia and industry in his pursuit to drive innovation in radar techniques. In 2013 he joined Aveillant Limited in Cambridge to be the R&D lead for the staring holographic radar and was recognised as a world leading expert on wind farm mitigation and counter drone surveillance. His current focus is on multi-function high-fidelity quantum-enabled networked radar that harnesses persistent dwell to track difficult targets in complex clutter environment. He has several patents, authored three book chapters and published over 60 journal and conference papers. He has served as TPC and track chairs on several of the international radar conferences and has organised special sessions and delivered tutorial and workshop on drone detection and surveillance. He is the chair of the EMSIG UK Radar Society ASPIRE focus group, serves as a PRI for the IET Chartership registration and is a keen STEM ambassador.

Chris J. Baker (Fellow, IEEE) was the Ohio State Research Scholar in Integrated Sensor Systems with The Ohio State University, Columbus, OH, USA. He held the position of the Chief Technology Officer for Aveillant at Thales company, France. He held the Chair of Intelligent Radar Systems at the Thales-Royal Academy of Engineering, University College London, London, U.K. Until 2011, he was the Dean and the Director of the College of Engineering and Computer Science, The Australian National University (ANU), Canberra ACT, Australia. He holds the Chair of Intelligent Sensor Systems at the University of Birmingham, Birmingham, U.K. He has been actively engaged in radar systems research since 1984 and is the author of over 300 publications. His research interests include coherent radar techniques, radar signal processing, radar signal interpretation, electronically scanned radar systems, radar imaging, and natural and cognitive echo locating systems. Mr. Baker is a fellow of the IET. He was a recipient of the IEE Mountbatten premium (twice) and the IEE Institute premium.

Michail Antoniou (Fellow, IEEE) received the B.Eng. degree (Hons.) in electronic and communications engineering and the Ph.D. degree in radar sensors and systems from the University of Birmingham, Birmingham, U.K., in 2003 and 2007, respectively. From 2006 to 2011, he was a Research Fellow with the Microwave Integrated Systems Laboratory, University of Birmingham. Since 2011, he has been with the Department of Electronic, Electrical and Systems Engineering, University of Birmingham, where he is a Reader in RF Sensing Systems. His research interests include passive radar, synthetic aperture radar (SAR) processing, multistatic and multiple-input and multiple-output (MIMO) radar, and cognitive radar, where he has some 100 publications in peer-reviewed international journals and conferences.