PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Multiple drone type classification using machine learning techniques based on FMCW radar micro-Doppler data

Joshua Bernard-Cooper, Samiur Rahman, Duncan Robertson

Joshua Bernard-Cooper, Samiur Rahman, Duncan Robertson, "Multiple drone type classification using machine learning techniques based on FMCW radar micro-Doppler data," Proc. SPIE 12108, Radar Sensor Technology XXVI, 121080A (27 May 2022); doi: 10.1117/12.2618026

SPIE.

Event: SPIE Defense + Commercial Sensing, 2022, Orlando, Florida, United States

Multiple Drone Type Classification Using Machine Learning Techniques Based on FMCW Radar Micro-Doppler Data

Joshua Bernard-Cooper[a], Samiur Rahman[a], Duncan A. Robertson[a]
[a]University of St Andrews, SUPA School of Physics & Astronomy, St Andrews, Scotland, UK

ABSTRACT

Systems designed to detect the threat posed by drones should be able to both locate a drone and ideally determine its type in order to better estimate the level of threat. Previously, drone types have been discriminated using millimeter-wave Continuous Wave (CW) radar, which produces high quality micro-Doppler signatures of the drone propeller blades with fully sampled Doppler spectra. However, this method is unable to locate the target as it cannot measure range. By contrast, Frequency Modulated Continuous Wave (FMCW) data typically undersamples the micro-Doppler signatures of the blades but can be used to locate the target. In this paper we investigate FMCW features of four drones and if they can be used to discriminate the models using machine learning techniques, enabling both the location and classification of the drone. Millimeter-wave radar data are used for better Doppler sensitivity and shorter integration time. Experimentally collected data from Ttree quadcopters (DJI Phantom Standard 3, DJI Inspire 1, and Joyance JT5L-404) and a hexacopter (DJI S900) have been. For classification, feature extraction based machine learning was used. Several algorithms were developed for automated extraction of micro-Doppler strength, bulk Doppler to micro-Doppler ratio, and HERM line spacing from spectrograms. These feature values were fed to classifiers for training. The four models were classified with 85.1% accuracy. Higher accuracies greater than 95% were achieved for training using fewer drone models. The results are promising, establishing the potential for using FMCW radar to discriminate drone types.

Keywords: Micro-Doppler, Millimeter wave radar, UAV, drone, machine learning, feature extraction, classification

1. INTRODUCTION

The potential misuse of widely available consumer drones presents a major security issue[1]. Drones can be used to transport contraband, engage in terrorist activity, and invade the privacy of unsuspecting individuals, motivating the development of systems that are able to detect, track, and classify them[2].

Detection system candidates include visual, acoustic, passive Radio Frequency (RF), thermal, and radar[3]. Radar is resistant to all weather conditions, can operate during day and night, and works on drones that do not emit signals unlike passive RF, so provides an advantageous method of drone detection. The incorporation of machine learning into a radar-based detection system allows targets to be identified rapidly, accurately, and without a human operator.

For a detection system to be reliable, it needs to have a high probability of detecting a target and a low probability of giving a false alarm. Birds and drones both fly slowly and at low altitudes and have a low Radar Cross-Section (RCS), meaning birds are likely to produce false alarms for a drone-detecting security system if it is based on RCS for target discrimination[4]. As a result of this, many early studies of machine-learning in radar-based drone detection systems focused on discriminating birds from drones with micro-Doppler signatures. This has now been achieved with great success.

Molchanov et al.'s 2014 study used Principal Component Analysis (PCA), calculating the correlation matrix of spectrograms for eleven classes including planes, birds, quadcopters, stationary rotors, and helicopters[4]. Eigendecomposition was used to find the eigenvectors of this matrix, which were Fourier transformed to obtain feature values that were used in conjunction with Support Vector Machine (SVM) and Naïve Bayes Classifier methods for

Radar Sensor Technology XXVI, edited by Kenneth I. Ranney, Ann M. Raynal, Proc. of SPIE
Vol. 12108, 121080A · © 2022 SPIE · 0277-786X · doi: 10.1117/12.2618026

classification[4]. SVM achieved an accuracy of 95% for classifying birds, and an accuracy of 96% for one drone model. A Naïve Bayes Classifier (NBC) achieved an overall accuracy of 88.42%[4].

In 2014, de Wit et al. utilised Singular Value Decomposition (SVD) to extract features[5]. In 2017, Rahman and Robertson investigated blade flashes and Helicopter Rotation Modulation (HERM) lines, comparing these for CW and FMCW radar[6]. This study demonstrated the advantage of using millimeter-wave radar to obtain robust micro-Doppler signatures, which became significant for later studies discriminating multiple drone types from CW radar data[7].

Fuhrman et al. (2017) utilised SVD to extract micro-Doppler features from the spectrograms of six drone types in laboratory scenarios, and from spectrograms of simulated bird flight[8]. Using SVM as a classifier, 100% accuracy was achieved. The study demonstrated that drones and birds could be discriminated reliably using micro-Doppler signatures but had not done so in a realistic outdoor scenario[8]. In 2019, Rahman and Robertson gathered data from outdoor drone and bird flights[7]. A combination of the authors' own algorithms and SVD were used on a small dataset to extract micro-Doppler spread, strength, and periodicity. SVM achieved 91.7% accuracy, whereas Linear Discriminant Analysis (LDA) achieved 100%[7].

Beyond the basic drone/bird classification problem, the subsequent goal of micro-Doppler-based drone classification is to determine the model of drone detected. This may allow a system to aid in determining a target's threat level. Drone model classification based on propeller micro-Doppler signatures has been achieved with millimeter wave Continuous Wave (CW) radar[9,10], but that method is unable to locate targets as it cannot measure range.

Zhang et al.'s 2017 study used PCA to extract micro-Doppler features of three drone types from Continuous Wave (CW) micro-Doppler data, with an SVM classifier achieving 94.7% accuracy[9]. Rahman and Robertson (2020) also classified multiple drones with micro-Doppler features extracted from CW radar data[10]. The authors compared the classification accuracy of supervised machine learning techniques with that of a Convolutional Neural Network (CNN). GoogLeNet, the CNN, achieved an accuracy of 99.74%, whereas SVM and Linear Discriminant Analysis achieved only 75%[10]. Those studies have shown that experimental radar data can be used to classify individual drone models based on fully sampled micro-Doppler data.

Unlike CW radar, Frequency Modulated Continuous Wave (FMCW) radar can locate targets, but due to hardware Constraints, micro-Doppler signatures of the very fast rotating propeller blades are under typically sampled[11]. We are not aware of previous studies that have achieved discrimination between drone models based on FMCW micro-Doppler signatures and that is the objective of this work.

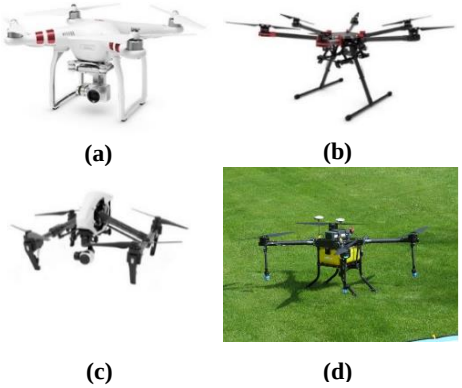
This paper aims to develop feature extraction algorithms for four drone models, the DJI Phantom Standard 3, DJI Inspire 1, DJI S900 Hexacopter, and Joyance JT5L-404. These are a small quadcopter, a medium quadcopter, a large hexacopter, and a large quadcopter respectively. Section 2 presents the feature extraction algorithms, Section 3 describes the classification training results, and conclusions are given in Section 4.

2. PROPOSED FEATURE EXTRACTION ALGORITHMS

2.1 Training Dataset

The four drone models used are shown in Figure 1. Data was collected with the 'T-220' 94 GHz FMCW radar in staring mode[16], with a range to stationary hovering targets between 33 m and 138 m. Specifications of the radar used are shown in Table 1. In total, 14.1 GB of raw time series radar data, corresponding to approximately 210 s of drone flight were analysed. For each of the four drone types, range bins occupied by the drones during episodes of interest were manually selected from each collected time series data run for the labelled dataset creation, totalling 90.87 seconds of drone flight. For the S900, three runs with a total of 13.65 seconds of flight time were divided into 13 episodes. For the Joyance, six runs with a total of 43.02 seconds of flight time were divided into 12 episodes of interest. For the Phantom, three runs with

a total of 5.35 seconds of flight time were divided into 10 episodes of interest. For the Inspire, six runs with a total of 28.85 seconds of flight time were divided into 12 episodes of interest. For each of these episodes, a spectrogram was generated.



(a) **(b)** **(c)** **(d)**

Figure 1 – The four drone models discriminated: (a) DJI Phantom Standard 3[12], (b) DJI S900 Hexacopter[13], (c) DJI Inspire 1[14], (d) Joyance JT5L-404[15]

These spectrograms were analysed to identify candidate features for drone type classification. Algorithms were developed to extract the identified features automatically from each spectrogram generated for subsequent classification. Features were ultimately chosen for classification training based on their separation of drone types in feature space and their impact on classification accuracy. The features chosen were micro-Doppler strength, bulk to micro-Doppler ratio, and HERM line spacing, as discussed in the next section.

These feature extraction algorithms were then tested on a small FMCW micro-Doppler dataset, with the extracted features used in conjunction with supervised machine learning techniques to classify the four drone models. As relatively few data points were used in training, focus is placed on proving the validity of the developed algorithms for extracting features that successfully discriminate drone models.

Table 1 – Parameters of the T-220 FMCW radar[16]

Operating frequency	94 GHz
Operating mode	FMCW, staring
Bandwidth	150 MHz
Antenna beamwidth	0.9° azimuth, 3° elevation
Polarization	Circular, odd-bounce
Chirp time	51.2 μs
Chirp period / chirp repetition frequency	80.489 μs / 12.4 kHz
Sampling rate	10 MHz, 512 fast time samples, 256 range bins
Maximum unambiguous velocity	±9.93 ms^{-1}
Short Time Fourier Transform (STFT) length	512-4096 samples, 41.2-329.6 ms
STFT window	Gaussian, width factor 0.4942-0.4950
STFT overlap	95%

2.2 Micro-Doppler Strength

The first feature extracted was the average micro-Doppler strength, derived from spectrograms with a window length, or coherent processing interval (CPI), of 512 samples (41.2 ms) for each episode of interest. The feature value was created by calculating the mean dBm value of the micro-Doppler sidebands, with any values below the selected noise floor of -80 dBm discarded. As the average Doppler noise floor is approximately -86 dBm, a 5-6 dB margin is used to ensure only Doppler signal is selected and not noise. To ensure that the bulk Doppler was not considered when calculating this value, it was suppressed below the noise floor so that it was discarded in calculation. This process is illustrated by Figure 2 and shown by Algorithm 1.

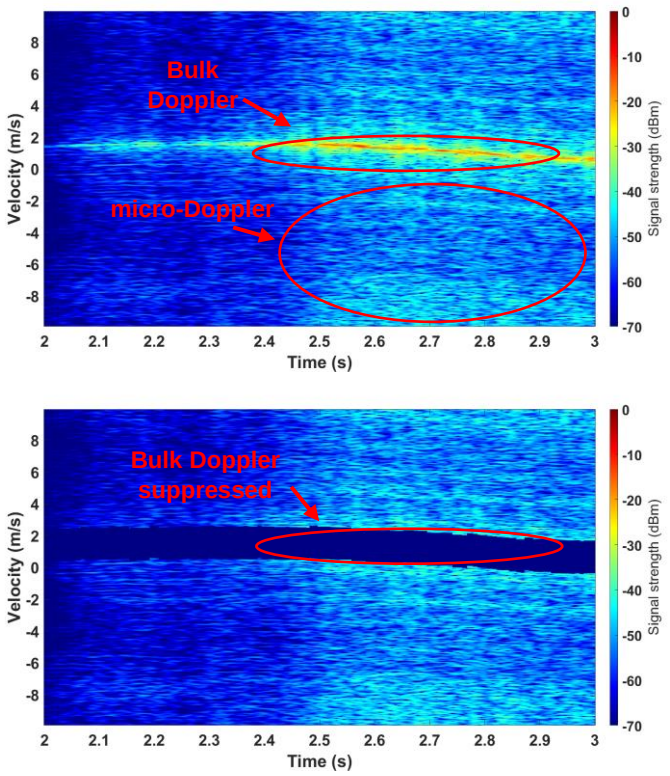


Figure 2 – Two spectrograms generated from the same second of S900 flight, showing before (upper) and after (lower) the bulk Doppler suppression algorithm was applied

Algorithm 1 Suppress bulk Doppler
begin procedure
input: spectrogram
Bulk suppression width = w
Suppression interval width = w*round(window length/velocity range)
Suppression value = s
for i = 1 to Number of spectrogram CPIs
    Select Bulk Doppler index
    Calculate suppression interval start and stop points
    for n = Suppression interval start to stop
        nth data point of suppression interval of Bulk-suppressed spectrogram's ith CPI = s
    end for
end for
output: bulk-suppressed spectrogram
end procedure

The mean was then taken of all values in the spectrogram matrix above -80 dBm (as per Figure 2 (lower)), giving a single average value for the amplitude of the micro-Doppler signature during the selected episode of interest. These micro-Doppler strength values were normalised for the power fall-off with range of the drone using radar calibration data when added to the feature value dataset. An additional feature, the micro-Doppler density, was calculated as the number of elements in the matrix with a value above the noise floor, divided by the total number of elements in the matrix. These processes are shown in Algorithm 2.

Algorithm 2 Calculate micro-Doppler strength and density

```
begin procedure
    input: spectrogram
    for i = 1 to Number of Doppler bins
        for n = 1 to Number of spectrogram CPIs
            if the point in the ith Doppler bin of the nth spectrogram CPI ≥ noise floor
                Total micro-Doppler strength += value of point
                Data points above the noise floor +=1
            end if
            Number of data points +=1
        end for
    end for
    micro-Doppler strength = Total micro-Doppler strength/Number of data points
    micro-Doppler density = Data points above the noise floor/Number of data points
    output: micro-Doppler strength and micro-Doppler density
end procedure
```

Micro-Doppler strength was used as a feature value but micro-Doppler density did not show discrimination between drone types, so was not used as a feature. This is likely because the under-sampling of micro-Doppler in FMCW radar means the Doppler sideband energy is aliased and spread right across the Doppler spectrum, so this measure of Doppler 'occupancy' within the spectrogram does not vary significantly with drone type.

2.3 Bulk to Micro-Doppler Ratio

The second feature calculated for these spectrograms was the bulk to micro-Doppler ratio. Using the same initial spectrograms generated for calculating the micro-Doppler strength (Figure 2 (upper)), the values of the maximum amplitude for each CPI were summed then divided by the number of CPIs, to give a mean value for the bulk Doppler. As dBm values scale logarithmically, a ratio is obtained through subtraction. The previously obtained value of the micro-Doppler strength was subtracted from the calculated value of the spectrogram's bulk Doppler, giving a ratio that was used as a second feature value. Algorithm 3 shows this process.

Algorithm 3 Calculate bulk to micro-Doppler ratio

```
begin procedure
    input: spectrogram
    for i = 1 to Number of spectrogram CPIs
        Bulk Doppler = max(spectrogram's ith CPI)
        Total bulk Doppler strength += Bulk Doppler
    end for
    Bulk Doppler strength = Total bulk Doppler strength/Number of spectrogram CPIs
    Bulk to micro-Doppler ratio = Bulk Doppler strength – micro-Doppler strength
    output: bulk to micro-Doppler ratio
end procedure
```

2.4 HERM Line Spacing

As HERM line spacing depends on the rotation rate and blade length of the observed propellers, which may be characteristic of drone type, this provided another primary candidate for feature value extraction. Spectrograms were generated with a long window length / CPI of 4096 (329.7 ms) so that the HERM lines were easily visible. For each spectrogram CPI, a region with a width of ±7 m/s about the bulk Doppler was extracted using the array of bulk Doppler index values. This was done to ensure that the same HERM lines were tracked through the spectrogram, should the drone move.

The number of peaks in the region with a minimum prominence of 10 dBm were counted so that insignificant peaks in the micro-Doppler signature between HERM lines were not considered. Prominence is measured by first extending a horizontal line to the left and right of the peak until it crosses the signal due to a higher peak or reaches the end of the signal. The minimum of the signal is then measured for the left- and right-hand instances, which is either a valley or a signal endpoint. The higher of these two minima is the reference level from which the height of the peak above this is its prominence. A minimum prominence value can be set in MATLAB's findpeaks function. The peak selection is demonstrated by Figure 3. The mean number of peaks per CPI was calculated by dividing the number of peaks by the number of CPIs in the spectrogram. The width of the region (14 m/s) was divided by the number of peaks to give the mean spacing of the HERM lines. This is shown in Algorithm 4.

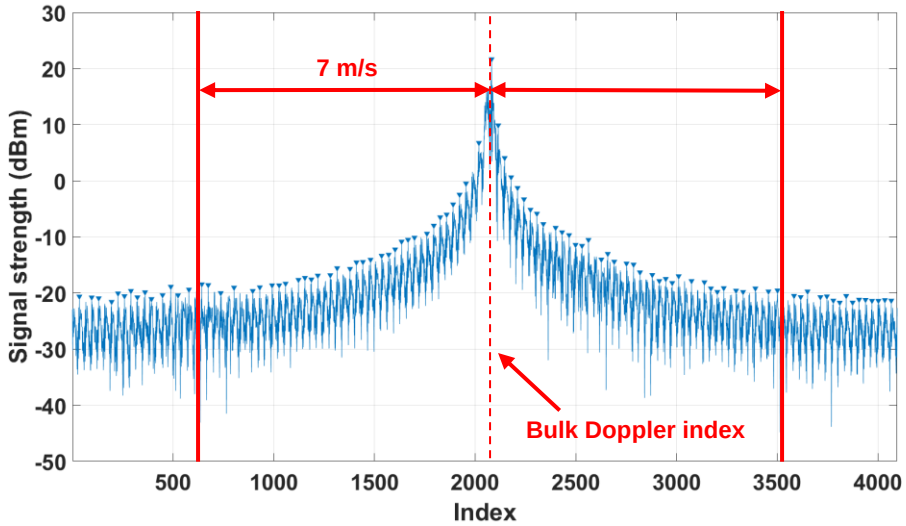


Figure 3 – A CPI showing the HERM line peaks, and the region in which these were selected for calculating HERM line spacing

Algorithm 4 Calculate HERM line spacing
begin procedure
1 **for** i = 1 to Number of spectrogram CPIs
2 Spectrogram CPI = ith column of the spectrogram matrix
3 Select Bulk Doppler index
4 Calculate HERM line interval start and stop points
5 [Peak value array, Peak location array] = findpeaks(HERM line interval)
6 Total number of peaks += length of Peak value array
7 **end for**
8 Average peaks per CPI = Total number of peaks/Number of CPIs
9 HERM line spacing = Width of CPI/Average peaks per CPIs
end procedure

2.5 Other Feature Extraction Algorithms

Two further feature extraction algorithms were developed with the objective of finding more features that might discriminate between drone types. The first measured the symmetry of the HERM lines in each generated spectrogram, inspired by Bennet et al. 2020[17]. The second used SVD, as this had been successful in discriminating between drone and bird micro-Doppler signatures[5]. Neither of these algorithms produced feature values that discriminated between drone models, so these were not used for classification training.

3. CLASSIFICATION TRAINING RESULTS

The three features used for classification training were micro-Doppler strength, HERM line spacing, and the micro-Doppler to bulk ratio. A data set of these feature values derived from the 47 selected spectrograms was created in Microsoft Excel, allowing them to be imported into MATLAB's Classification Learner app. Each set of three feature values was labelled with the associated drone model, providing four target classes for classification (Phantom, Inspire, S900, and Joyance). 47 values for each feature were used, corresponding to 4 drones. The Inspire had 12 values for each feature, the Phantom had 10, the Joyance had 12, and the S900 had 13. In total 141 feature values were used; the Inspire had 36 feature values, the Phantom had 30, the Joyance had 36, and the S900 had 39.

Classification was carried out for 4 classes, as well as results for all possible combinations of 3 and 2 classes created by excluding drone models and their feature values. For each classification type, classification training was carried out with all three features (micro-Doppler strength, bulk to micro-Doppler ratio, and HERM line spacing) and all possible combinations of two features. All 24 classifiers available in MATLAB's Classification Learner app were tested for each of these combinations. Five-fold cross-validation was used to prevent overfitting the generated models to the data.

Classification was also tested with Principal Component Analysis (PCA) enabled. With a small number of features, PCA produced a universal reduction in classification accuracy as the model depended on all features used for sufficient training. In most scenarios, only two features were used, so eliminating one of them affected the training significantly.

The greatest accuracy achieved for each classification method is shown in Table 2, including the features and classifier that produced this result. Given that prediction speed is an important consideration in security applications, this is also included. Classification was performed on an Intel i5-4690K processor running at 3.50 GHz clock speed with 8 GB of RAM. The results obtained for 4-, 3- and 2-class classification are discussed further in the following sections. It was observed that the first two features, micro-Doppler strength and HERM line spacing, are most separated in feature space. This is why in all but two cases these are the features used in classification

As the dataset used was small, with 10-13 feature values per drone type, these results are proof of principle for the developed feature extraction algorithms, validating their potential in future work to create a larger dataset that could be used to locate and discriminate drone models in an FMCW radar-based drone detection system.

Table 2 – The results of classification training with the chosen feature values, micro-Doppler strength (M), HERM line spacing (H), and bulk to micro-Doppler ratio (R). For each classification type, the combination of features and classifier that gave the highest accuracy is show in in the table. The prediction speed of each classifier is also shown, as this is an important consideration.

Classification Type	No. of features	Classifier	Accuracy (%)	Prediction Speed (ms)
4-Class	2 (M, H)	Quadratic SVM	85.1	0.23
3-Class (Inspire, S900, Joyance)	2 (M, H)	Subspace Discriminant	97.3	5.00
3-Class (Inspire, Phantom, Joyance)	2 (M, H)	Kernel Naïve Bayes	97.1	0.59
3-Class (Phantom, S900, Joyance)	2 (M, H)	Quadratic Discriminant	88.6	0.43
3-Class (Inspire, Phantom, S900)	2 (M, H)	Kernel Naïve Bayes	77.1	0.56
2-Class (S900, Joyance)	2 (M, H)	Quadratic Discriminant	100	0.63
2-Class (S900, Inspire)	3 (M, H, R)	Logistic Regression	100	1.00
2-Class (Joyance, Phantom)	2 (M, H)	Quadratic SVM	100	0.59
2-Class (Joyance, Inspire)	2 (M, H)	Kernel Naïve Bayes	95.8	0.71
2-Class (Phantom, Inspire)	2 (M, R)	Quadratic SVM	95.5	0.48
2-Class (S900, Phantom)	2 (M, H)	Kernel Naïve Bayes	73.9	0.71

4.1 4-Class Classification

For the 4-Class classification type, in which a model was trained to discriminate between all four drones, the best results were observed using micro-Doppler strength and HERM line spacing as features. A Quadratic SVM classifier produced the highest accuracy of 85.1%. This classifier also produced the highest accuracy for two 2-Class classification types. Figure 4 shows the predictions made by the model for each point in feature space. An SVM classifier was expected to give a high accuracy, given its success in previous studies discriminating drones from birds[4].

The feature values of the S900 and those of the Phantom occupy a similar region of the plot, making them difficult to discriminate for the model, creating the most inaccuracy. As shown by the confusion matrix in Figure 4, the greatest number of incorrect predictions occurred where a Phantom was classified as an S900. Training 3-Class models, in which one drone model is excluded, illustrates this further.

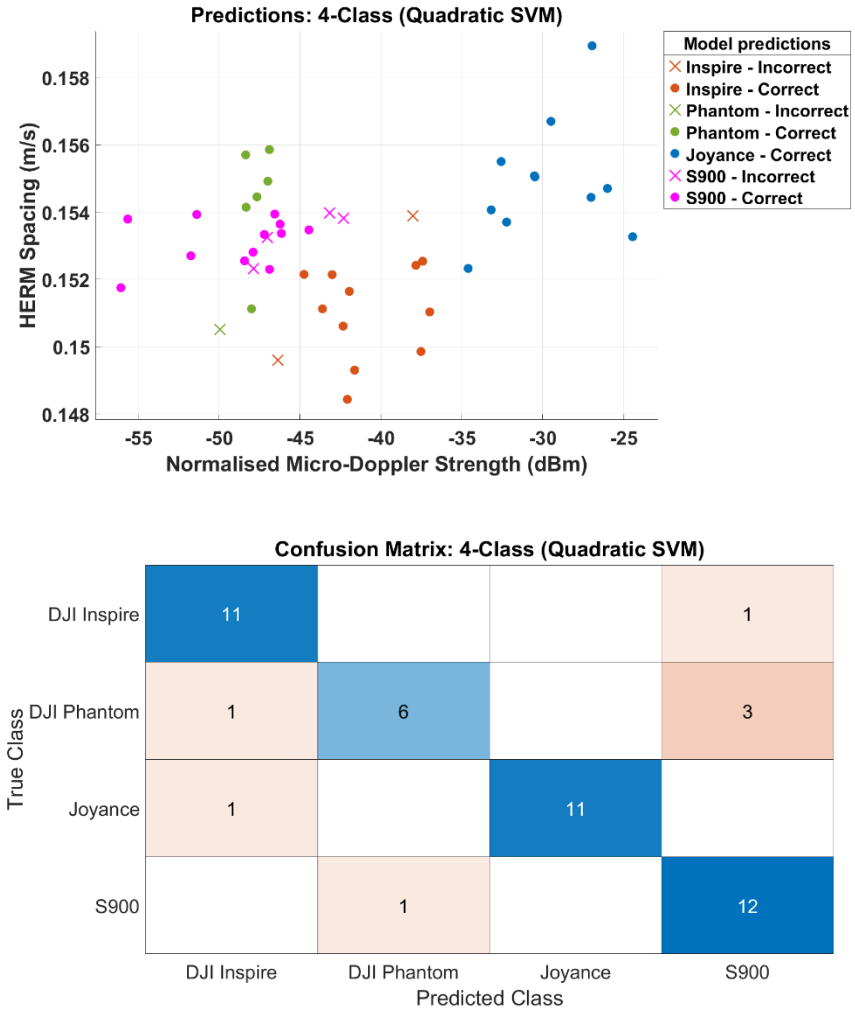


Figure 4 – Scatter plot and confusion matrix for 4-Class classification. The scatter plot shows that the datapoints for the S900 and the Phantom occupy a similar area of the plot, causing the greatest source of confusion for the model.

4.2 3-Class Classification

As with 4-Class classification, 3-Class classification also produced the best results using micro-Doppler strength and HERM line spacing as features. The results of these classifications support the suggestion that the S900 and Phantom were the greatest confusers in 4-Class classification, and therefore the greatest source of inaccuracy. The two 3-Class classification types excluding the S900 and Phantom scored 97.1% and 97.3% accuracy, respectively.

Using the Inspire, Phantom, and Joyance as the three classes, a Kernel Naïve Bayes classifier produced the highest accuracy of 97.1%. Figure 5 shows the predictions made by the model for each point in feature space; there is significantly less overlap in the regions occupied by the classes than in 4-Class classification. This classifier also gave the highest accuracy when classifying the S900, Inspire, and Phantom, as well as for two 2-Class classification types. Like SVM, Naïve Bayes classifiers also showed success in previous studies[4], so were expected to perform well for drone model discrimination.

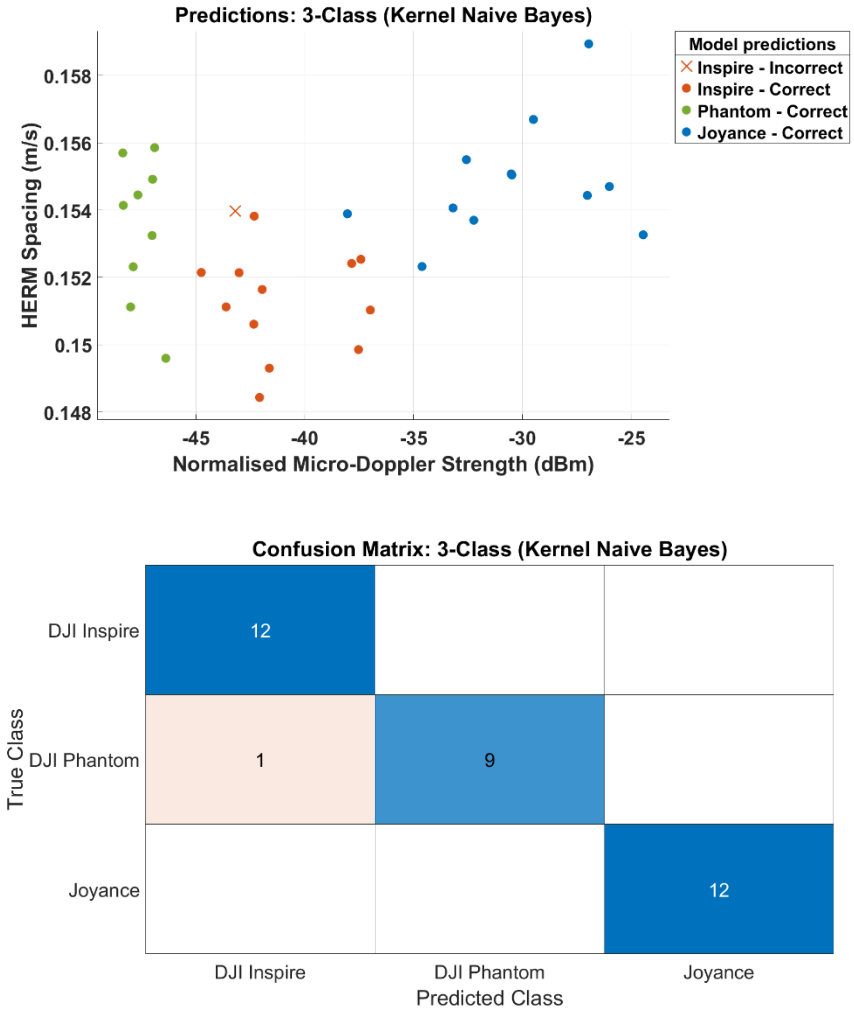


Figure 5 – Scatter plot and confusion matrix for 3-Class classification with the Inspire, Phantom, and Joyance. The scatter plot shows that each class occupies a relatively distinct area in feature space compared to the 4-Class classification in Figure 4.

Using the Inspire, S900, and Joyance as the three classes, a Subspace Discriminant classifier achieved the highest accuracy of 97.3%. Figure 6 again shows significantly less overlap in the regions occupied by the classes than in 4-Class classification. The Subspace Discriminant classifier is a classification ensemble provided by MATLAB's Classification Learner app. As a result, the Subspace Discriminant classifier had the longest prediction speed, at 5 ms. Although this seems fast, allowing for 200 classifications per second, what is considered fast or slow depends on the application requirement. In the case of security, where a drone classification system would be deployed, real-time, low-latency classification is needed to launch a counter-measure. The Subspace Discriminant classifier is five times slower than the next fastest classifier and may therefore delay the deployment of such measures.

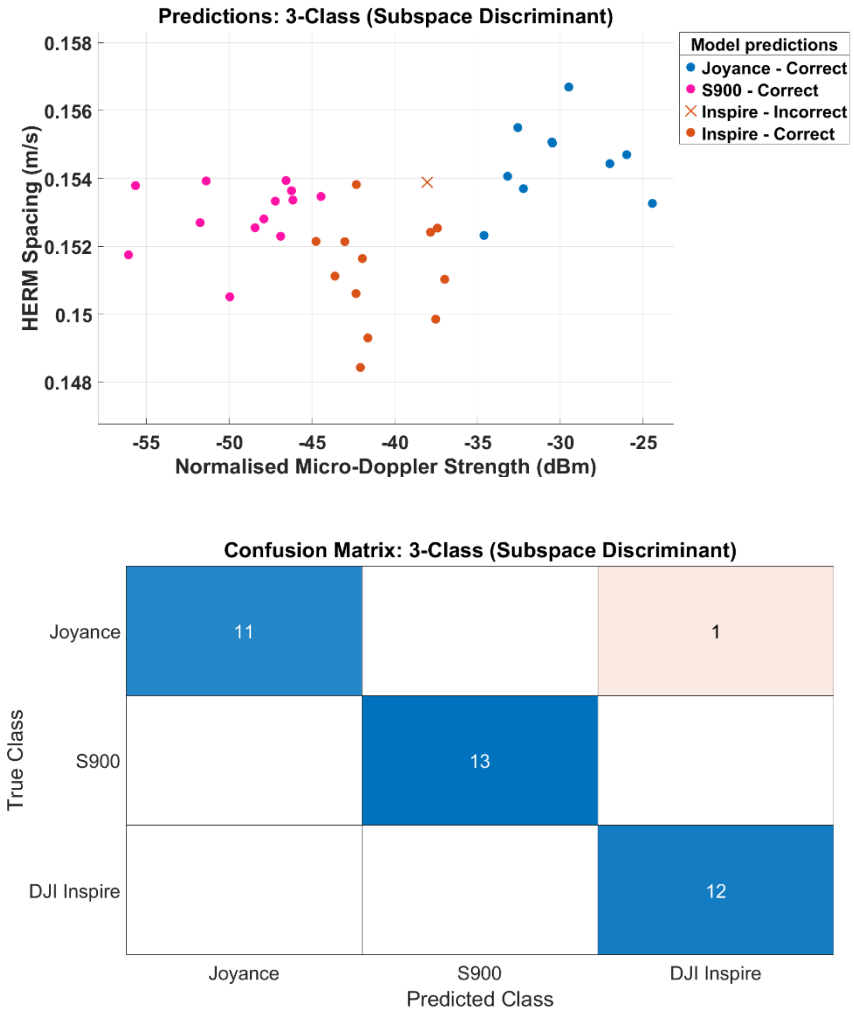


Figure 6 - Scatter plot and confusion matrix for 3-Class classification with the Inspire, S900, and Joyance. The scatter plot shows that each class occupies a relatively distinct area in feature space compared to the 4-Class classification in Figure 4.

Comparatively, the two 3-Class classification types including the Phantom and S900 only scored 77.1% and 88.6% accuracy. These results were obtained with a Kernel Naïve Bayes classifier and a Quadratic Discriminant classifier, respectively. The significant drop in classification accuracy indicates that these drone models are the cause of greatest confusion. The reason for the similarity in feature values between the Phantom and S900, shown by Figure 7, is unclear and requires further investigation in the future. A possible cause is the similar length of these models' propeller blades. A larger dataset in future may reveal more information and provide more clarity as to why a quadcopter and hexacopter produce such similar feature values. The Quadratic Discriminant classifier also produced the highest accuracy with one 2-Class classification type.

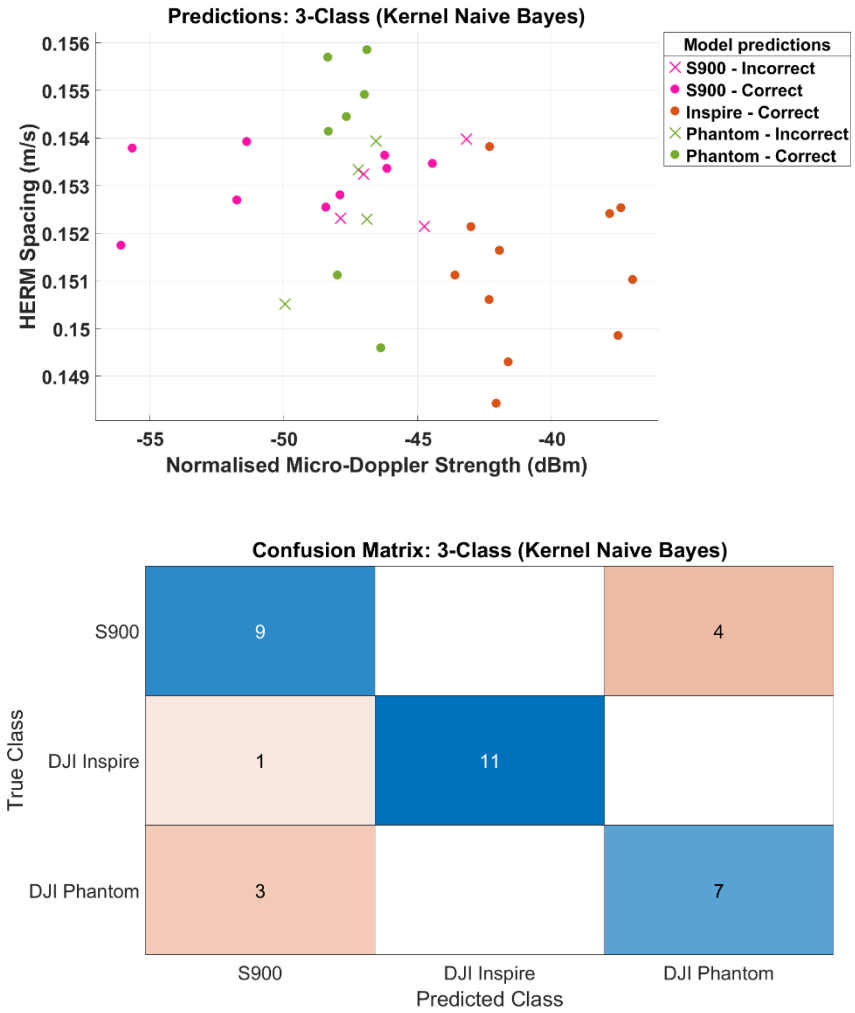


Figure 7 - Scatter plot and confusion matrix for 3-Class classification with the Inspire, S900, and Phantom. The scatter plot shows that the datapoints for the S900 and Phantom occupy a similar region of the feature space, as in the 4-Class classification in Figure 4, leading to the greatest source of confusion.

4.3 2-Class Classification

100% accuracy was achieved for three 2-Class classification types. Among these was the classification of the S900 and Inspire, which was the only classification type to achieve the highest accuracy using all three features. The classifier used was Logistic Regression. This level of accuracy indicates that these drone models are very well separated in feature space. This suggests that it is possible to group drones into separate categories, where each category will contain multiple drones having similar feature space values. Even though 100% accuracy may be due to the small dataset, and may be reduced when more datapoints are used, it definitely shows the potential of the model to discriminate between the three combinations of 2 drones with very high accuracy.

Two of the three remaining 2-Class classification types also achieved a high accuracy, including the classification of the Phantom and Inspire. This was the only classification type to achieve the highest accuracy using only micro-Doppler strength and bulk to micro-Doppler ratio as features.

The final 2-Class classification type, classifying the Inspire and S900 Hexacopter, achieved only 73.9% accuracy. This further supports the observation that these drone models act as the greatest source of confusion in classification types with a greater number of drone models, and are not discriminated well by the defined features.

4. CONCLUSION

In this paper, three feature extraction algorithms have been developed to train machine learning classifiers for multiple drone type classification based on a small set of FMCW radar data in which the micro-Doppler signatures are under-sampled. Approximately 14 GB of raw experimental data of drones in flight, spanning ~210 seconds, from a 94 GHz millimeter-wave radar were analysed in this work. From the raw data, a total of 47 episodes of interest were extracted (10-13 for each drone type), spanning about 90 seconds of flight, from which spectrograms were generated for feature extraction. from each spectrogram, the three feature values were extracted and labelled then used for classification. . The results are very promising, with the accuracy achieved validating the developed algorithms as having the potential for use in FMCW radar systems for locating and discriminating drone types.

4-Class classification achieved an accuracy of greater than 85%. 3-Class classification also achieved high accuracies, with two classification types achieving greater than 97% and one achieving over 88%. The exception to this was the classification type including the S900 and Phantom, which achieved 77.1%. 2-Class classification achieved accuracies over 95% in all cases except discriminating between the S900 and Phantom. From this, it can be concluded that these two drone types are the greatest confusers and have similar feature values. 'This might be due to the similarity of their blade lengths, but further investigation is required (especially with a larger dataset) for a better understanding.

To improve this classification and verify that the algorithms are suitable for use in a real-world system, a larger and more robust dataset that contains more datapoints per drone model is needed but that was outside of the authors' scope and is therefore suggested as a focus for future work. Further work should also address the need to create a more diverse dataset including flying data and data acquired from radars with different parameters, allowing for optimisation of threshold factors. The creation of a large dataset would also provide the opportunity to test CNNs for classification of drone types based on FMCW micro-Doppler spectrogram images, as this may produce higher classification accuracies as seen with CW data in Rahman and Robertson (2020)[10].

REFERENCES

[1] Yaacoub, J.-P., Noura, H., Salman, O. and Chehab, A., "Security analysis of drones systems: Attacks, limitations, and recommendations," Internet of Things 11(2), 1–39 (2020).
[2] Markarian, G. and Staniforth, A., [Countermeasures for Aerial Drones], Artech House, Norwood, 1-42 (2020).
[3] Taha, B. and Shoufan, A., "Machine Learning-Based Drone Detection and Classification: State-of-the-Art in Research," IEEE Access 7(1) (2019).
[4] Molchanov, P., Harmanny, R. I., de Wit, J. J., Egiazarian, K. and Astola, J., "Classification of small UAVs and birds by micro-Doppler signatures," International Journal of Microwave and Wireless Technologies 6(3-4), 435-444 (2014).
[5] de Wit, J. J., Harmanny, A., R. I. and Molchanov, P., "Radar Micro-Doppler Feature Extraction Using the Singular Value Decomposition," 2014 IEEE International Radar Conference, 2 (2014).
[6] Rahman S. and Robertson D.A., "Millimeter-wave micro-Doppler measurements of small UAVs," Proc. SPIE 10188, Radar Sensor Technology XXI, 101880T, 5-10 (2017).
[7] Rahman, S. and Robertson, D., "Millimeter-wave radar micro-Doppler feature extraction of consumer drones and birds for target discrimination," Proc. SPIE 11003, 1-7 (2019).
[8] Fuhrmann, L., Biallawons, O. Klare, J., Panhuber, R., Klenker and Ender, J., "Micro-Doppler analysis and classification of UAVs at Ka band," Proceedings of IEEE Radar Symposium, 2 (2017).
[9] Zhang, P., Yange, L., Chen, G., and Li, G., "Classification of Drones Based on Micro-Doppler Signatures with Dual-band Radar Sensors," Progess in Electromagnetics Research Symposium PIERS, 638-643 (2017).

[10] Rahman, S. and Robertson, D., "Multiple drone classification using millimeter-wave CW radar micro-Doppler data," Proc. SPIE. 11408., 5-8 (2020).
[11] Rahman, S., and Duncan, R., "Radar micro-Doppler signatures of drones and birds at K-band and W-band," Scientific Reports, 17396(8) (2018).
[12] DJI. (2021). Phantom Standard 3. Retrieved from DJI: https://www.dji.com/uk/phantom-3-standard
[13] DJI. (2021). Spreading Wings S900. Retrieved from DJI: https://www.dji.com/uk/spreading-wings-s900
[14] DJI. (2021). Inspire 1. Retrieved from DJI: https://www.dji.com/uk/inspire-1
[15] Joyance. (2022). Retrieved from Joyance: https://joyance.tech/
[16] Robertston, D.A, Brooker, G.M. and Beasley, P.D.L., "Very low-phase noise, coherent 94GHz radar for micro-Doppler and vibrometry studies," Proc. SPIE 9077, Radar Sensor Technology XVIII, 907719, 3-5 (2014);
[17] Bennett, C., Jahangir, M, Fioranelli, F., Ahmad, B.I. and Kernec, J.L., "Use of Symmetrical Peak Extraction in Drone Micro-Doppler Classification for Staring Radar," 2020 IEEE Radar Conference (RadarConf20), 1-6 (2020).