http://eprints.gla.ac.uk/184987/

Deposited on 7 May 2019

Effective Ground-Truthing of Supervised Machine Learning for Drone Classification

Jacob Sim
University of Glasgow, UK
2123659s@student.gla.ac.uk

Mohammad Jahangir
Aveillant Limited, UK & University of Birmingham
mohammad.jahangir@aveillant.com &
m.jahangir@bham.ac.uk

Francesco Fioranelli
University of Glasgow, UK
francesco.fioranelli@glasgow.ac.uk

Chris J Baker
University of Birmingham, UK
c.j.baker.1@bham.ac.uk

Holly Dale
University of Birmingham, UK
HAD449@student.bham.ac.uk

Abstract — It has already been shown that multibeam staring radar is able to detect and track low observable targets such as drones due to its high sensitivity [1]. Due to this level of sensitivity, targets that have a similar RCS to drones are also detected and tracked. These are predominantly birds. Birds and drones are similar in several ways such as flight altitude, velocity and manoeuvrability [2] such that discrimination between them is challenging. Hence, there is a need to look for high performing methods of classification, for example, machine learning. Supervised training of machine learning classifiers requires accurately labelled training data. For control targets, such as drones, truth data from the on-board GPS logging can be used for data labelling. However, opportune bird targets require a separate data collection method that enables association with the radar output for a classifier to be effectively trained. This paper shows a method of collecting and displaying ground-truth for small targets onto GoogleEarth so that the radar data can be appropriately used to create accurate training data for a machine learning, drone and bird classifier. Results of classification performance are presented showing high performance that is aided by the availability of more effective truth data.

Keywords — staring radar, machine learning, training data ground-truth, drones, birds, classification, opportune target

I. INTRODUCTION

Radars that can detect and track small targets such as drones are steadily becoming more common with several companies, including Aveillant, producing radar systems specifically designed for counter-UAS applications [3]. The need for these systems has seen a rapid increase due to recent events including airport and military facility trespassing and attempted assassinations using weaponized drones. The number of drone incidents involving near-misses with aircraft has reached an all-time high with around 120 reports in 2018 alone in the UK [4]. Radar still seems to be the most tried and tested method of detecting, tracking and classifying this relatively new class of targets, but other methods exist that can aid the radar in detecting and classification. To detect these small targets, a radar system requires a greater sensitivity than conventional radars as drones typically have an RCS that can be much less than 1m^2. By increasing the sensitivity, there is a greater number of "confuser targets" detected compared with traditional ATC radar systems. These targets mainly come in the form of birds that have a similar RCS to drones, and hence there is a need to discriminate between these two target classes, otherwise the radar output would become completely congested with airborne targets. Aveillant's Gamekeeper 16U (Figure 1) is a radar system designed for UAS applications and is the system used to collect and test the methods presented in this paper. This radar uses a single L-band transmitter and an array of static planar receivers to provide a 90-degree azimuth field of view and a 30-degree elevation view providing a full 3-D position and track of a target onto a display in real-time.

Fig. 1. Aveillant L-Band Gamekeeper 16U counter-UAS multi-beam staring radar.

Raw radar data is processed to detect and track the target followed by the classifier that uses a multitude of features for target recognition. Some target classes are easier to discriminate, for example, aircraft and drones. The differences between these two are seen in almost every feature evaluated as aircraft generally fly much faster, higher and have a considerably larger RCS. Propeller powered aircraft also produce a greater range of micro-Doppler components as they typically travel at a greater radial velocity compared to the velocity of a drone's propellers. The real classification challenge comes in the form of discriminating between birds and drones as there is little separability between individual feature values. Micro-Doppler components have shown to aid classification of small targets [5, 6]. However, classification performance is very variable as propeller detection range is dependent on echo strength, which might be a further 10 to 20 DBs less than the body echo strength. Due to these similarities between drone and bird classes with no clear singular feature to tell them apart, there is a need search for combinations of features that together create a high performing and robust classifier. Further, there could be more than one set of features used to classify each of these targets at any one time. This could be addressed using a decision tree classifier [5]. Due to the complexity of this form of classifier it is also beneficial to use a machine learning algorithm to build the decision tree. This uses a set of accurate training data collected from real radar data containing the correct class labels of each of the

targets to be considered. Accurate class labelling of this training file also requires accurate truth data to be collected.

The truth data can come in many forms with one of the most accurate and convenient being GPS. For all target classes, GPS can be used to provide near-exact positional data. Specific trials can be conducted to collect data from aircraft and drones easily just by attaching almost any GPS device to them [7]. For the greatest accuracy, a barometrically corrected GPS device is used. Using specialist GPS equipment, trained birds can also be used to collect data but there would be a limit to the breed of birds and the behaviour exhibited due to their captivity. Hence, there is a requirement for a GPS free system that can collect data that can be compared with the radar echo history to provide class labels of the radar tracks. This system would not necessarily require a definitive 3-D position but could be an observer on the ground providing their position and orientation to a target along with the target class label which would be enough to relate it to a recorded track. This system could hence be used for any opportune target that presented itself during a trial such as birds, aircraft and ground targets. This system would also ideally be easy to use and have little opportunity for human error which could corrupt the training data and lead to inaccuracies in the classifier.

The methods developed for collecting ground-truth data from opportune targets, and for creating a training file before training and testing a decision tree classifier are all detailed in this paper.

II. METHODOLOGY FOR GROUND-TRUTHS

The first stage of creating a machine learned classifier is to collect training data with accurate labelling. For control targets, such as drones and aircraft, labelling was achieved using GPS tracks association. This used an algorithm to compare the GPS truth data with any radar track that appeared within a set volume of space around the truth track and had a similar velocity and heading.

For opportune targets, a whole new method was created due to the deficiencies of existing methods. The method used records an observer position and orientation towards the specific target by utilizing an "iPad" tablet and the "*Theodolite*" application. This combination allows an observer to use the built-in camera to capture an image of the opportune target and input, via a custom note, the targets class label e.g. bird or, ideally, type of bird. The image was then saved to the tablet by the application along with meta-data stored within the image. This meta-data contained all the positional and orientation information of the tablet at the time the image was taken, using the on-board tablet sensors. This information includes the time, GPS position, altitude, bearing angle, elevation angle as well as the custom note input by the observer containing the class label. Once proficient in the use of the tablet, an initial measurement of a target could be recorded in under 5 seconds with consecutive measurements taken every second due to the application saving the last entered custom note class label. After the first measurement, a target and related ground-truth information could be recorded at a typical rate of 1Hz.

After all the ground-truth data had been recorded as a compilation of images, an example of which is shown in Figure 2, they are easily transferred to a computer for comparison with the radar tracks. A simple script then extracts the meta-data from each image and creates a KMZ file for viewing on GoogleEarth. Each measurement taken is displayed at the exact time of capture by first plotting the position of the tablet, using the GPS position and altitude, then projecting a line in the direction of the target by using the bearing and elevation angles. Each line is then associated with the extracted image data and given the name of the custom note input by the observer. This provides a visual check that the note input by the observer corresponds to the captured image.

Fig. 2. Opportune bird target "bird67" captured using tablet and application overlayed with tablet sensor data.

The custom note also allowed for a navigable folder structure as this string of characters is also analysed in the same script. For instance, if the observer had input a custom note of "bird123" where the numbers "123" corresponded to the one hundredth and twenty third opportune bird target observed, then each image with the custom note "bird123" could be put into a folder which corresponds to all the measurements taken of the same target. This folder is then also put into a folder along with any others containing the same string of "bird". All this information is contained within a single KMZ file and is subsequently easy to associate with the radar tracks once converted into GoogleEarth format.

An example of an opportune bird target with ground-truth compared with the radar track is shown in Figure 3. This method of providing ground-truth data of an opportune target was an improvement on traditional methods that are fraught with the possibility of human-error. The new method has allowed for a greater number of radar tracks to be given a class label, significantly improving classifier training. As well as its main function, the tablet also provides a range of applications such as a view of the live display of the radar output for monitoring targets in the field. This acts as an early warning system for the observer to prepare for approaching targets before they are directly visible. In addition, it provides a useful tool during trials as it monitors a large percentage of other opportune targets.

To indicate the degree of improvement made to the ground-truthing, compared to predecessor methods, an opportune target measurement took over a minute to record (due to the number of separate measurement devices). This has now been hugely improved to a rate of 1Hz, greatly improving utility. The previous method utilized handwritten logs which were transferred manually onto a computer to generate a KML file. Hence, the reliability of target measurement was lower due to possible human error in

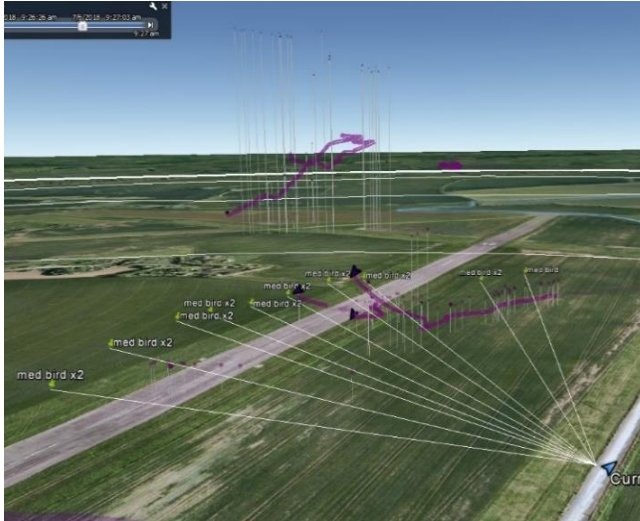making inputs and the longer recording time resulting in fewer target track associations and their labelling.

Fig. 3. Opportune bird target "med bird x2" images displayed on GoogleEarth showing orientation lines correlating with a radar track.

III. GROUND-TRUTH RESULTS

The ground-truthing method described in the previous section was used to provide labelling of real radar tracks obtained from measurements carried out at the trials site based in Deenethorpe, Northamptonshire, UK. Gamekeeper radar data was collected for several control targets flights using a DJI Inspire drone. The radar data was processed, and the tracks exported to GoogleEarth which were then cross referenced against the ground-truth information and track IDs were registered corresponding to drone and bird targets. Table I and Table II lists a selection of tracks that were labelled in this manner.

TABLE I.   BIRD RADAR TRACK IDS LABELLED USING GROUND-TRUTH

Target Label	Track ID	Truth ID	Start Time	End Time
Bird	ID97202/98392	Bird 1	08:58	09:01
Bird	ID102853/101826/101515	Bird 2	09:01	09:03
Bird	ID106333	Bird 3	09:03	09:04
Bird	ID114773/116339/116974	Bird 4	09:04	09:07
Bird	ID116447	Bird 5	09:07	09:08
Bird	ID121327	Bird 6	09:08	09:09
Bird	ID16024	Bird 7	09:24	09:24
Bird	ID16448	Bird 8	09:24	09:26
Bird	ID26618/27020	Bird 9	09:27	09:30
Bird	ID38183/50574	Bird 10	09:52	09:56
Bird	ID42661	Bird 11	09:54	09:54
Bird	ID44782	Bird 12	09:54	09:56
Bird	ID46543	Bird 13	09:54	09:55
Bird	ID53496	Bird 15	09:57	09:58
Bird	ID68221/70167	Bird 20	10:04	10:07
Bird	ID85558	Bird 21	10:09	10:11
Bird	ID110186/112294	Bird 22	10:21	10:26
Bird	ID127769	Bird 27	10:29	10:31
Bird	ID130205	Bird 28	10:30	10:32
Bird	ID133061	Bird 29	10:32	10:34
Bird	ID132035	Bird 31	10:33	10:34

TABLE II.   DRONE RADAR TRACK IDS LABELLED USING GROUND-TRUTH

Target Label	Track ID	Truth ID	Start Time	End Time
Drone	ID35390	Drone 1	09:51	09:53
Drone	ID52286	Drone 2	09:57	10:00
Drone	ID73354	Drone 3	10:06	10:07
Drone	ID78387	Drone 3	10:08	10:09
Drone	ID81606	Drone 3	10:09	10:10
Drone	ID86308	Drone 3	10:10	10:11
Drone	ID107693	Drone 4	10:20	10:23
Drone	ID126795	Drone 5	10:29	10:33

The track labels are then used to extract trajectory-based features from the radar track data. The trajectory information is also used to generate spectrograms of tracks from the raw radar data. Figure 4 and Figure 5 show the labelled trajectories and spectrogram for two targets, labelled using the procedure adopted here. Figure 4 shows a drone with its distinct micro-Doppler signatures clearly visible. In Figure 5 the target is a bird and whilst there is evidence of some very faint Doppler side bands, they exhibit a lack of periodicity and are quite different in character from that of a drone. The labelled trajectory data is used in the machine learning classifier. Previously work used a far less reliable process of labelling of unknown tracks of opportune targets which would then limit the ability of the machine learning classifier to utilised labelled data from opportune targets [8].
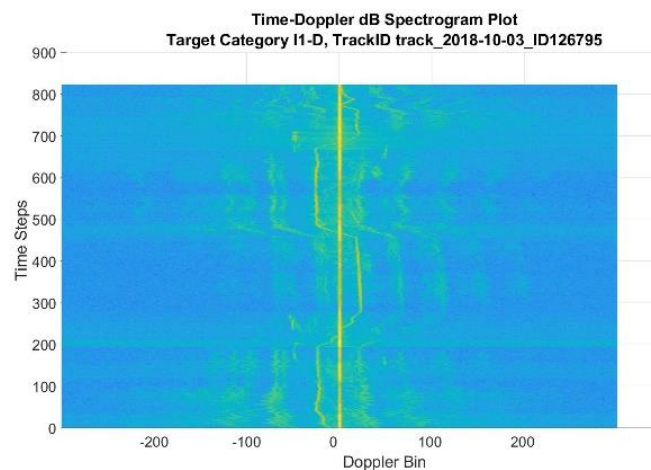

Fig. 4. (top) Radar tracks with a selected tracked (ID 126795) labelled as drone (bottom) Spectrogram for track ID 126795.

IV. MACHINE LEARNING CLASSIFICATION

Once radar, truth, and ground-truth data had all been recorded, the next stage was to create an accurate classifier for the radar system. The classification algorithm is a decision tree based on supervised machine learning. To use machine learning, a training file was created using truth data collected from drone consisting of samples of tracked targets collected from the radar and containing the parameters recorded of each feature. The classifier simply has 2 classes; (i) drone and (ii) non-drone. Each sample requires its class label to be recorded. This allows the machine learning algorithm to find any differences between the collective class features and hence enables discrimination. This is done either using a feature on its own (unlikely for drones and birds) or by combines features together to create a case-by-case classifier.

Features that the classifier selects need to avoid over-tuning to the training data (overfitting). For instance, if trials were conducted in a particular region and the x-y position was used, then the classifier could determine that only a target of a particular class would appear in that region. Consequently, here, the classifier is only be allowed to use the positional feature of height. Derivatives of position are used such as velocity, acceleration and jerk as well as the number of micro-Doppler components associated to a target. As well as the instantaneous features, statistical analysis over a rolling window is used to analyse the feature history of a target. The mean, minimum and maximum of features are used. Other features are also exploited such as the age of the track and the target RCS.
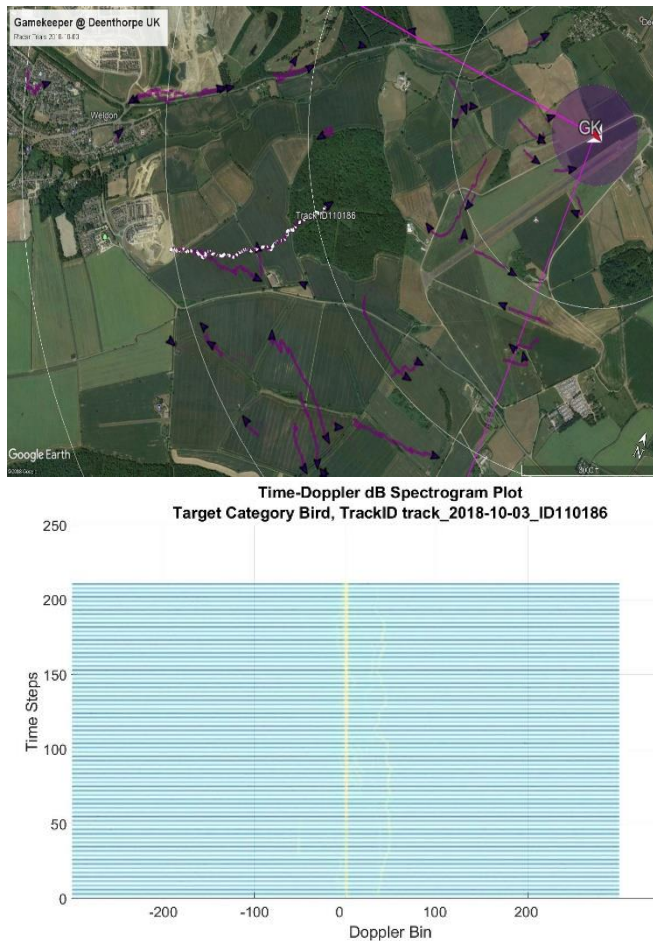

Fig. 5. (top) Radar tracks with a selected tracked (ID 110186) labelled as bird (bottom) Spectrogram for track ID 110186.

The decision tree classifier was created in MATLAB using the "*fitctree*" function from the "Machine learning and statistics" toolbox. This function called for a 1-D array of class labels relating to each sample and a 2-D array containing the feature parameters values for each sample with each column relating to a different feature. These arrays were both extracted directly from the training file and the function would then use these arrays to produce a decision tree classifier. It also allowed for the option of k-fold optimisation and pruning to prevent over-fitting to the provided training data. The output consisted of a 2-D array of the decision tree parameters with each row relating to a node and columns for node number, positive child node number, negative child node number, feature used, split value, data size and majority class.

A typical result from a machine learned 2-class drone decision tree is shown in Figure 6 along with the associated confusion matrix. Class 1 (blue) is the non-drone class and class 2 (orange) is the drone class. This shows a correct classification of 94.3% and a false classification of 0.5%. This result shows what is potentially possible, even with a limited set of training data but is still flexible enough to provide similar results for multiple drone models and separate radar sites.
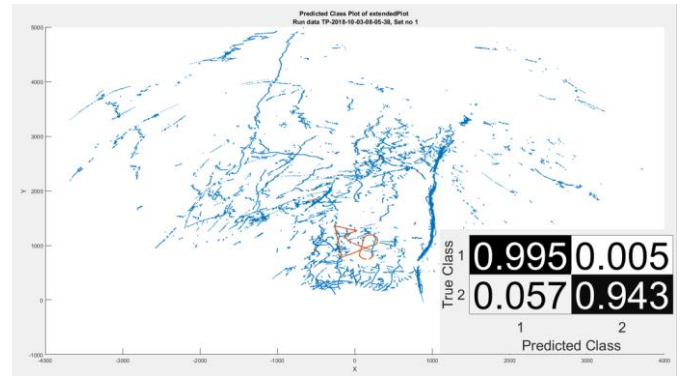

Fig. 6. Typical results for a machine learned decision tree classifier showing the non-drone (blue) and drone (orange) class with confusion matrix.

A decision tree can also be produced to include more than two classes and is only dependant on the training data labelling supplied. If a decision tree only discriminated between two classes such as drones and non-drones as in this case, then the radar classification algorithm could also use another two-class decision tree classifier of bird and non-bird in parallel to add in this class to enable it to be a multi-class classifier. The process of creating a classifier using this method can be easily automated so a new classifier can be produced whenever needed. This could be when new specific class data needs to be added due to it falling outside previous class feature parameters or if changes were made to either previous sections of the radar system's algorithm chain or to the hardware. This could result in even a small parameter change to propagate through to the classification algorithm and result in classification error.

V. CONCLUSIONS

This paper presents an approach to quickly and accurately collect ground-truth experimental data of opportune targets, mostly birds, recorded at radar drone trials. This method enables the recording of accurate information on these targets of opportunity and compare them with the radar output to provide a greater number of class labels to be associated with the radar tracks. The combination of the tablet and application

allows for measurements to be recorded quickly and accurately with a low opportunity of human error input. This then allows for better and more complex classifiers to be created using machine learning by providing more accurate and significant training data. In the final paper, further results will be presented showing the effects of limited truth data versus the more complete set that can now be collected with the method detailed in this paper.

VI. ACKNOWLEDGEMENT

REFERENCES

[1] Jahangir M. & Baker C. J., L-band staring radar performance against micro-drones, *International Radar Symposium IRS 2018*, Bonn, Germany, June 2018.

[2] Ritchie, M., et al., Monostatic and bistatic radar measurements of birds and micro-drone, *IEEE Radar 2016*, Philadelphia, US, May 2016.

[3] Jahangir, M., Baker, C., Robust Detection of Micro-UAS Drones with L-Band 3-D Holographic Radar, SSPD 2016, Edinburgh, UK, Sep. 2016.

[4] "Airprox involving UAS Drones". In: (2018). URL: https://www.Airproxboard.org.uk/Reports-and-analysis/Statistics/Airprox-involving-UAS-Drones/

[5] Jahangir, M., Baker, C., Persistence Surveillance of Difficult to Detect micro-drones with L-band Holographic Radar, CIE 2016 Int. Radar Conf., Guangzhou, China, Oct. 2016.

[6] Molchanov, P., et al., Classification of small UAVs and birds by micro-Doppler signatures, EuRad 2013, Germany, Oct. 2013.

[7] M. Jahangir and C. Baker, Characterisation of low observable targets with a multi-beam staring radar, *IEEE Radar 2017*, Belfast, Oct. 2017.

[8] Jahangir, M., Baker, C. J. & Oswald G. K. A., Doppler characteristics of micro-drones with L-band multibeam staring radar, *IEEE RadarCon 2017*, Seattle, US, May 2017.