

Received February 11, 2022, accepted March 16, 2022, date of publication March 28, 2022, date of current version April 14, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3162866

DroneSense: The Identification, Segmentation, and Orientation Detection of Drones via Neural Networks

STIRLING SCHOLES¹, ALICE RUGET¹, GERMÁN MORA-MARTÍN², FENG ZHU¹,
ISTVAN GYONGY², AND JONATHAN LEACH¹

¹School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K.

²School of Engineering, The University of Edinburgh, Edinburgh EH9 3FF, U.K.

Corresponding author: Jonathan Leach (j.leach@hw.ac.uk)

This work was supported in part by the Defence Science Technologies Laboratory under Project Dstlx-1000147352 and Project Dstlx-1000147844, and in part by Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/T00097X/1 and Grant EP/S026428/1.

ABSTRACT The growing ubiquity of drones has raised concerns over the ability of traditional air-space monitoring technologies to accurately characterise such vehicles. Here, we present a CNN using a decision tree and ensemble structure to fully characterise drones in flight. Our system determines the drone type, orientation (in terms of pitch, roll, and yaw), and performs segmentation to classify different body parts (engines, body, and camera). We also provide a computer model for the rapid generation of large quantities of accurately labelled photo-realistic training data and demonstrate that this data is of sufficient fidelity to allow the system to accurately characterise real drones in flight. Our network will provide a valuable tool in the image processing chain where it may build upon existing drone detection technologies to provide complete drone characterisation over wide areas.

INDEX TERMS Convolutional neural network, drones, orientation detection, pose, segmentation.

I. INTRODUCTION

The proliferation of semi-autonomous aerial vehicles, i.e. drones, into the consumer and industrial spaces, combined with the growing number of drone related incidents (infractions into commercial airspace, [1], [2] or the use of drones by militant groups, [3], [4]) has raised concerns over the ability of existing aerial detection systems to accurately characterise such vehicles [5]–[7]. Specifically, many existing air-space monitoring technologies are optimized to detect the presence of a vehicle, identify its type, and, track its position over time but they lack the resolution to determine target specific features. This, in conjunction with drones ability to decouple their motion in space from their assigned task e.g. simultaneously translate and rotate to keep a subject in frame whilst filming, means that presence, type and position are often insufficient to accurately identify the intent of a vehicle.

To accurately assess the intent of a drone it is necessary to fully characterize its ‘pose’ i.e., not only identify its type but also segment it into functional components and identify the

orientation of these components in 3D space. Fig. 1 conceptually illustrates this process showing a DJI Mavic 2 drone segmented into colour-coded components and placed within a 3D Gimbal corresponding to its orientation.

To address the problem of drone characterization a wide variety of machine learning assisted drone detection systems have been developed. For example, radio based methods, which eavesdrop on the communications between drones and pilots and apply the statistical analyses of control signals [8]–[11], Convolutional Neural Networks (CNNs) analysing the spectrogram [12]–[15], K-Nearest Neighbours (KNNs) [16] clustering of signals, cyclostationary feature extractors [17], decision trees [18] and random forest techniques [19], bit-analysis [20], and, residual [21], recurrent [22] and hierarchical networks [23]. Additionally, acoustic based methods analysing the noise of a drones motors and propellers have also been developed using Mel Frequency Cepstral Coefficients (MFCC) [24]–[29] or by converting the signal to a spectrogram [27], [30], [31]. Once obtained, the MFCC or spectrogram feature set can be used to train Long-Short Term Memory (LSTM) models [24], or Convolution type models such as CNNs [31]–[36], Recurrent

The associate editor coordinating the review of this manuscript and approving it for publication was Felix Albu¹.

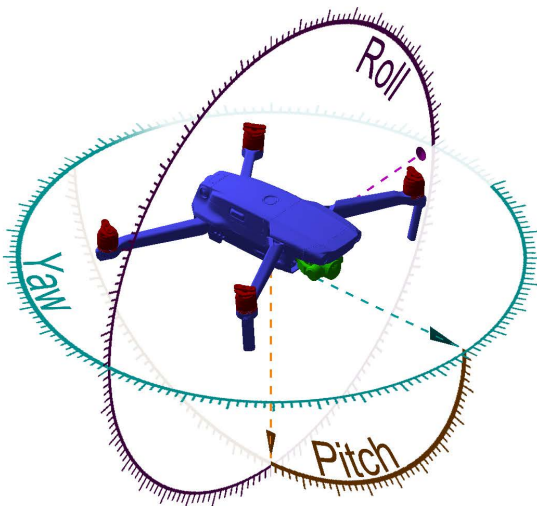


FIGURE 1. A conceptual representation of drone pose. A drone (here represented by a DJI Mavic 2) is identified and divided into its components, for instance, body (blue), engines (red), and camera (green). Further, the orientation of the drone in 3D space as represented by the roll (magenta), pitch (orange) and yaw (cyan) gimbals is identified.

Neural Networks (RNNs) [32]–[34] which incorporate temporal dependence and Convolutional-RNNs (CRNNs) [33], [34]. The feature set can also be used to train vector type models including, Support Vector Machines (SVMs) [27], [28], [30], [35], Gaussian Mixture Models [32] and KNNs [37] or, retrain existing models such as random forests [38], ResNet [25] and LeNet [39].

Despite the relative efficacy of acoustic and radio based systems the introduction of quiet micro-drones and fully autonomous drones (which do not require radio commands) has rendered them progressively less versatile and has necessitated the development of radar and optical based sensor systems. Radar in particular has seen extensive development including pulsed systems [40]–[42], Doppler systems [43]–[47], and Frequency Modulated Continuous Wave (FMCW) systems [48]–[50] all at multiple wavelengths [51]–[60]. The reader is directed to Refs [61]–[64] for a comprehensive review. Whilst radar based systems are able to monitor a large area and are robust to atmospheric conditions, their reliance on micro-Doppler information for drone type identification and poor transverse resolution has prevented their application to problems beyond target detection and tracking. Hence, in parallel to radar systems, machine learning assisted optical drone detection systems have been developed. Such systems have been extensively used to identify the presence of drones in an image and construct bounding boxes at ranges comparable to that of radar systems [65], [66].

The most common approach to optical drone detection is to train existing CNN based networks such as You Only Look Once (YOLO) [67], [68] and ResNet [69], [70] on colour camera images. These networks include coupling to pan-tilt and zoom camera mounts to track moving objects [71], using multi-camera systems to increase the field of view [72], [73],

utilising the high speed nature of YOLO to identify drones at video frame rates [74], comparing the performance of YOLO v2 and YOLO v3 on drones at short range against static backgrounds [75], examining the effect of incorrect images labels on YOLO [76] and, modified YOLO implementations [77].

More complex optical CNN architectures have also been developed where features in the image (such as moving objects) are enhanced before being sent to a second network for identification. These multi-stage networks have proven to generally be more effective at discriminating drones from drone-like objects in images such as birds [78]–[81]. Such networks have been developed using background subtraction with image stabilization [82] and CNNs [83], [84], subtracting successive frames and clustering using an SVM [85], HAAR filters [86] for edge and feature detection, foreground background separation [87], ResNet for feature extraction and SVMs for classification [88], Kalman filters and ResNet [89], Faster-RCNN and ResNet [90], and, using trajectory mapping to suppress erroneous YOLO identifications [91]. Additionally, several other networks have been used for drone identification. These include identifying regions of interest in an image [92] using Histogram of Gradient (HOG) descriptors with thresholding or Fourier descriptors [93], simultaneous image upsampling and downsampling [94], Inception Net v3 [95], generic Fourier descriptors [96], [97], Faster-RCNN [98], and, TIB-Net with CenterNet, lightweight networks optimised for speed of processing [99], [100].

Finally, a number of more niche applications have also been investigated such as, controlling the flight of a drone based on external camera observations [101] and, using cameras mounted on multiple drones to track and even intercept hostile drones [102]–[105]. For a review of the different machine learning implementations listed above the reader is directed to Refs [106]–[108]. Despite the numerous optical systems developed to date characterisation of drones beyond presence, location and type remains rare with demonstrations limited to determining if a drone is carrying a payload [109] or the identification of key points on a single drone at short range [110].

A promising avenue for the more complete characterisation of drones is given by sensor fusion in which multiple sensors are combined. For example, using a large field-of-view low resolution sensor to direct a small field-of-view high resolution sensor with such systems seeing improvements in performance of up to 15% [72], [111], [112]. In the case of optical drone detection systems one such example is the development of depth sensing time-of-flight systems such as LIDARS. LIDARS active illumination allows them to operate when no passive light source is available (such as at night), detect targets which themselves emit no thermal radiation, and, operate to a limited degree through obscurity. Scanning LIDARS have been shown to be effective at drone detection at ranges up to 2 km when coupled with a Variable Radially Bounded Nearest Neighbour (V-RBNN) network to analyse the point cloud [113], [114]. Further, flash LIDAR

systems such as those employing Single Photon Avalanche Detector (SPAD) array cameras allow for the simultaneous capture of a ‘traditional’ high transverse resolution intensity image as well as a lower transverse resolution depth image (where ‘depth’ refers to the distance between the camera and the object for each pixel). Such systems have been shown to be effective at identifying the pose of objects at short range [115], [116] but have yet to be applied to the problem of drone characterization.

Here, we present a CNN which provides the complete characterization of drones. The network takes as input an intensity image as well as depth data and outputs: the identity of the drone i.e., the type of drone in the data; the segmentation of the drone in which each pixel in the intensity image is classified according to the drone component it represents; and, the orientation, the angle of the drone about its three principle axes of rotation, yaw, pitch, and, roll. We examine the performance of the network in multiple scenarios including, different drones, different ranges of motion and different data inputs. We assume that our system is being used in an image processing chain where supplementary systems such as radars would have already distinguished the drone from drone-like objects (e.g. birds) and would be able to direct a small field-of-view camera at the drone. We outline a system for producing large quantities of accurately labelled simulated data on which we train our network. To verify both our network structure and our simulated training data we demonstrate the ability of our network to accurately characterize an image of a real DJI Mavic 2 Zoom drone in flight as captured by a Qintic 4 × 4 SPAD camera [117]. The SPAD camera represents a state-of-the-art sensor fusion system combining a functional transverse resolution of 80 × 240 pixels for intensity and 20 × 60 pixels for depth. Further, each depth pixel outputs a depth histogram with 500 picosecond temporal resolution. Finally, the architecture of the chip has the potential for the alternating acquisition of visible spectrum intensity images and depth histograms at rates in excess of 1000 frames per second [115].

II. NETWORK ARCHITECTURE

We present a network architecture built on a decision tree coupled with an ensemble network. The decision tree identifies the type of drone after which a set of drone-specific pretrained networks are applied in parallel to perform the orientation and segmentation operations. Specifically, the orientation is determined by three identical networks each trained to identify a single axis (roll, pitch or yaw) while the segmentation is performed by an additional U-Net [118] type network. This structure allows multiple drone parameters to be identified simultaneously through network parallelization whilst allowing each network to be optimized on a specific parameter yielding superior overall performance.

The lack of high quality drone image training datasets remains an obstacle for machine learning assisted drone classification. To address this, several publications have examined data augmentation [119] techniques such as,

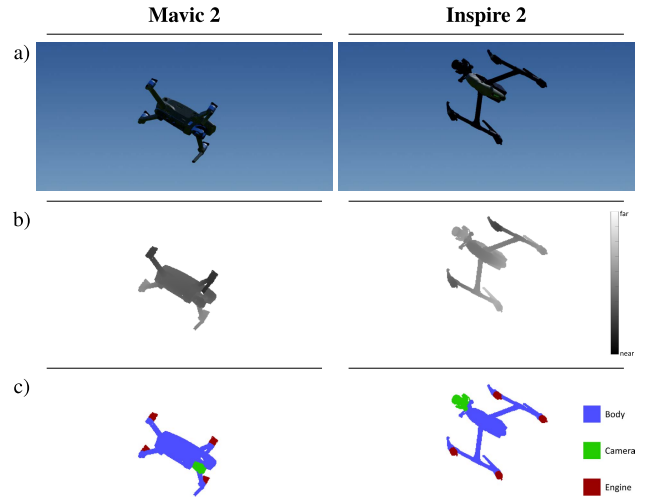


FIGURE 2. Examples of the unreal engines ability to produce realistic, accurately labelled, intensity, depth and segmentation data. a) Intensity images generated by the unreal engine of a DJI Mavic 2 and (an upside down) DJI Inspire 2 drone in flight. b) Unreal engine depth images corresponding to the drones in the top panels. c) Segmentation labels from the Unreal engine for the drones in the top panels showing the body (in blue), the engines (in red), and, the cameras (in green).

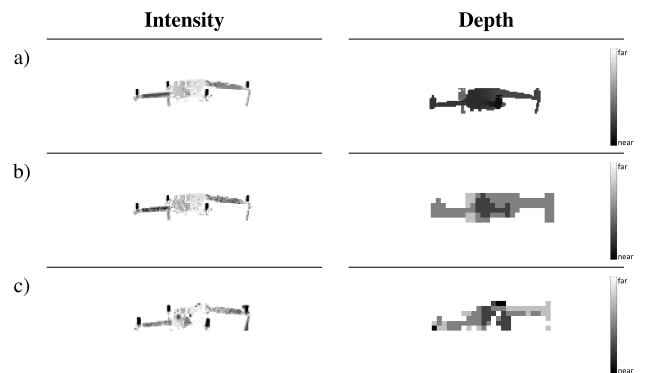


FIGURE 3. Processed images from unreal engine compared to quantic 4 × 4 SPAD camera images. a) The intensity and depth images produced by the unreal environment. b) The data used to train the network. The intensity image is noised with a poisson filter while the depth image is down-sampled and converted to a histogram of depths (visualised here as a depth image). c) Images captured by a quantic 4 × 4 SPAD camera of a real drone in flight. Note that the intensity images have been enhanced in contrast for better visualization.

super-imposing drone images onto unrelated backgrounds [120], super-resolution upscaling [121] and, generating new images from Generational-Adversarial Networks (GANs) [122]. Here, we leverage the capability of the Unreal Engine video game development environment to rapidly produce a large set of photo-realistic, accurately labelled training data as illustrated by Fig. 2. This approach allows us to explore the parameter space of drone types, orientation limits (e.g. the upside down Inspire 2 in Fig. 2), lighting conditions and image qualities to an extent which would be impractical experimentally. Further, our model could be readily extended to include numerous different backgrounds and weather conditions. The Unreal code is publicly available and can be

found at <https://github.com/HWQuantum/DroneSense>. Fig. 3 shows examples of the images processed by the network. The simulated images produced by the Unreal environment (Fig. 3a)) are processed before being passed to the network. The simulated intensity image is noised with a Poisson filter (Fig. 3b)) and resized to 80×240 pixels, while the depth is downsampled and converted to a histogram with a dimensionality of $20 \times 60 \times 15$. Fig. 3c) shows the images produced by a Quantic 4×4 SPAD array camera of a real drone in flight which the simulated data is designed to mimic. We stress that the image sizes used in the simulated data were selected only to match the physical parameters of the Quantic 4×4 SPAD sensor, and the network can be reshaped to any dataset with both intensity and depth information. The images generated by the model could easily be adapted to match those obtained with a different camera. Additionally, the ability for the SPAD to isolate a volume of space using time-of-flight gating means that the background of the images may be neglected.

Fig. 4 shows a summary of the identification, orientation and segmentation networks. At the core of these networks is the Drone Feature Encoder (DFE) which reduces the input data to a latent feature space. The DFE takes as input both a histogram of depths (of size r_h rows, c_h columns, and p_h pages) and an intensity image (of size (r_i, c_i)). The histogram is passed twice through two 3D convolutional layers (each with 32 filters) and axial max-poolings to extract its depth features and reduce it to a dimensionality of $(r_h, c_h, 1)$. The intensity image is passed through two 2D convolution layers (each with 32 filters) and a max-pooling such that it is reduced to a dimensionality of $(r_h, c_h, 1)$. The intensity and depth tensors are then concatenated and passed twice through a set of two 2D convolutions (each with 32 filters) and max-poolings ultimately reducing the network inputs to a latent space of $1 \times 3 \times 32$ filters. The DFE is identical in all the networks with each network distinguished by how it handles the data in this latent space.

In the case of the identification network which defines the decision tree, the latent space is flattened to a 96 element vector and connected to a dense layer with 64 neurons. These neurons are in turn connected to the single output node with a Sigmoid activation. This network is trained using cross-entropy as a loss function such that it outputs an integer corresponding to the type of drone in the image. The orientation networks are identical in structure to the identification network, but the final neuron uses a ReLu activation. ReLu activation allows the neuron to output a continuous value corresponding to the angle in a given axis. Further, the orientation networks are trained using the loss function given in Eqn. 1 which allows them to correctly account for the cyclic nature of angle prediction and handle the discontinuity in prediction between 360° and 0° .

$$\text{Loss} = \min[\text{abs}(l - p), \text{abs}(l - p - 360^\circ)]^2, \quad (1)$$

where l is the label, p is the networks prediction, and, abs is the absolute value function. The segmentation network

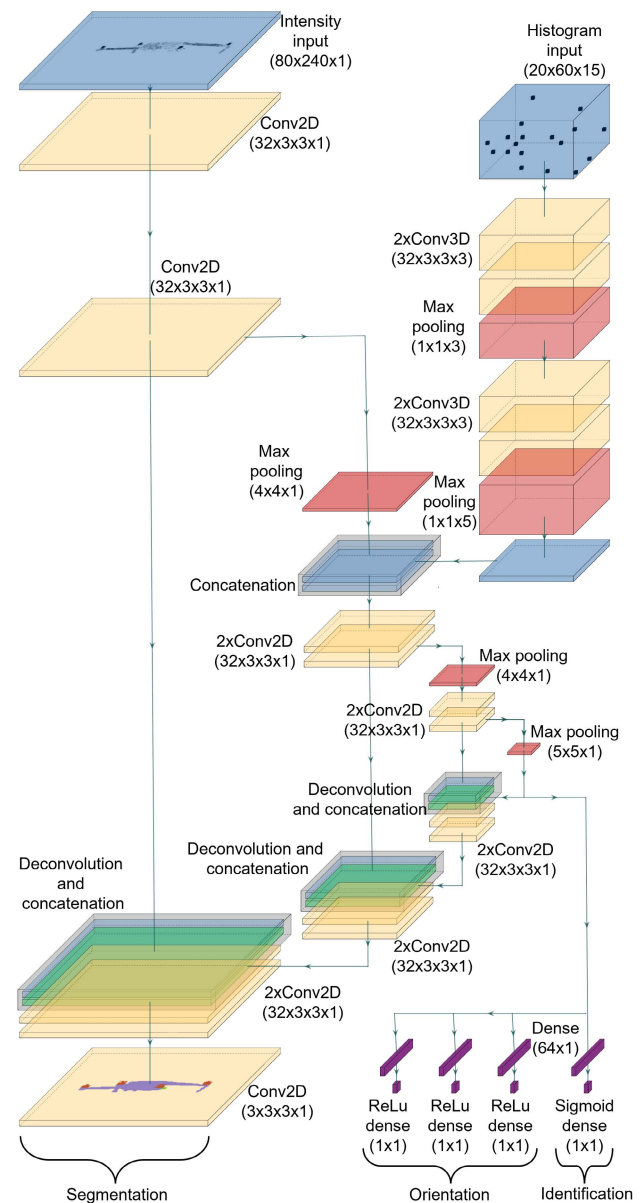


FIGURE 4. A summary of the ensemble network structure, here the common components of the networks have been drawn together while in practice each network in the ensemble is distinct. The networks take in a high transverse resolution intensity image and a low transverse resolution histogram of depth. Using convolution, pooling, and concatenation the inputs are reduced to a dense latent space. The identification network connects this latent space to a dense layer and then to a single Sigmoid activated neuron for drone type classification. By contrast, the three orientation networks use an identical structure but employ ReLu activation in the final neuron to output a continuous value corresponding to the angle in a given axis. Segmentation is performed by up-sampling the latent space to a final convolution with filters corresponding to the components being identified.

attaches a U-Net to the DFE. This U-Net up-samples the latent space to a set of segmentation predictions of size (r_i, c_i, n) where n corresponds to the number of components being identified. Each layer of the U-Net mirrors the DFE, undoing the max-pooling and using skip connections to concatenate the tensors. These concatenated tensors are then

TABLE 1. Summary of network prediction accuracies for both drones in the full angle and reduced angle regimes.

Mavic 2		
Metric	Full angle	Reduced angle
Orientation	(accuracy \pm std)(%)	(accuracy \pm std)(%)
Roll	88.3 \pm 13.0	99.6 \pm 0.4
Pitch	99.2 \pm 0.8	99.7 \pm 0.3
Yaw	92.3 \pm 9.7	96.3 \pm 5.0
Segmentation		
Body	97 \pm 1	96 \pm 1
Engine	86 \pm 9	86 \pm 5
Camera	85 \pm 13	85 \pm 16
Identification	100 \pm 0	100 \pm 0
Inspire 2		
Metric	Full angle	Reduced angle
Orientation	(accuracy \pm std)(%)	(accuracy \pm std)(%)
Roll	95.3 \pm 5.6	98.8 \pm 0.9
Pitch	99.6 \pm 0.4	97.5 \pm 1.4
Yaw	92.4 \pm 8.5	96.0 \pm 4.1
Segmentation		
Body	91 \pm 2	92 \pm 2
Engine	82 \pm 9	84 \pm 5
Camera	81 \pm 10	88 \pm 6
Identification	100 \pm 0	100 \pm 0

passed through two 2D convolutions each with 32 filters. The network was trained using binary cross-entropy with the ADAM optimizer and a learning rate of 0.001 with no drop out. The final output is a single convolutional layer with (in this case) three filters corresponding to the three components being identified; the body of the drone, the engines of the drone, and, the camera on the drone.

III. RESULTS

A. RESULTS ON SIMULATED DATA

Two drones were used for testing, a DJI Mavic 2 Zoom and a DJI Inspire 2. High fidelity models of these drones were placed in the Unreal environment and a total of 72 000 simulated SPAD images generated. The images feature the drones at random positions within the field-of-view, at random orientations, and, at random distances from the SPAD camera, ensuring sufficient variation in the data. From the training images, 10% were reserved for network testing. We do not make use of any image augmentation, although this could be used to increase the total number of training images. The networks were trained until the loss converged and the networks with the best performance on the testing data saved. These models were then validated on a separately generated set of 3600 unseen validation images. This ensured no chance of the network overfitting to the validation data. A summary

of the results for the identification, segmentation and orientation networks is presented in Table 1. The final trained parameters of the model are specific to the images that we use for training, and these images are closely matched to those generated by the Quantic 4 \times 4 sensor. Images collected with a different sensor could be used with this model, however, the optimal performance will always be achieved if the model is retrained with the appropriate images.

To ensure non-negative angular values in all drone orientations a coordinate system was established in which level flight facing away from the camera corresponded to, yaw = 180°, roll = 180° and pitch = 90°. Within this coordinate system two angular regimes were examined, the ‘full angle’ regime and the ‘reduced angle’ regime. In the full angle regime the drone models had the following range of motion: yaw \in [0°, 360°]; roll \in [0°, 360°]; and, pitch \in [0°, 180°] (with pitch limited to [0°, 180°] to negate gimbal-lock). In the reduced angle regime the models were constrained to within the manufacturer’s limits specifically: yaw \in [0°, 360°]; roll \in [140°, 220°]; and, pitch \in [140°, 220°]. By examining these regimes we evaluate the network’s ability to characterise drones flying in both conventional flight modes and more exotic flight modes (such as upside down).

Fig. 5 displays the predictions of the orientation networks for the full angle and reduced angle regimes. The theta coordinate represents the angle and the radial coordinate represents the error with -180° error at the center and $+180^\circ$ error at the circumference. The solid red ring indicates the ground truth. Network under and over predictions fall inside of and outside of the red ring respectively. Predictions made by the network trained on the full range of angles are shown as blue triangles. Predictions made by the network trained on the reduced range of angles (indicated by the shaded region) are shown as green dots.

By examining the radial distribution of the predictions, the accuracy of the networks in each axis and in each regime can be compared and the following observations made. First, the accuracy of the networks is contingent upon the number of images-per-angle the network is given to train on. In the full angle regime where the pitch is restricted to half the range of the roll and yaw the network accuracy improves significantly since for the same number of total training images the number of examples-per-degree is doubled to ~ 400 . This is also why in the reduced angle regime where the roll is restricted its accuracy matches that of the pitch, while the yaw does not, even when the same total number of training images is used. Second, the accuracy of the networks is coupled i.e., for a reduced range of motion in one axis the accuracy of the remaining axes will increase. While this effect is less pronounced than that of examples-per-degree it can be observed in Table. 1 where a 4° increase in yaw accuracy is observed for both drones in the reduced angle regime. This despite the range of motion in that axis remaining constant. The improvement can be attributed to the reduced variance (roll and pitch range) in the images which the yaw network must learn. Third, the accuracy of the networks is somewhat

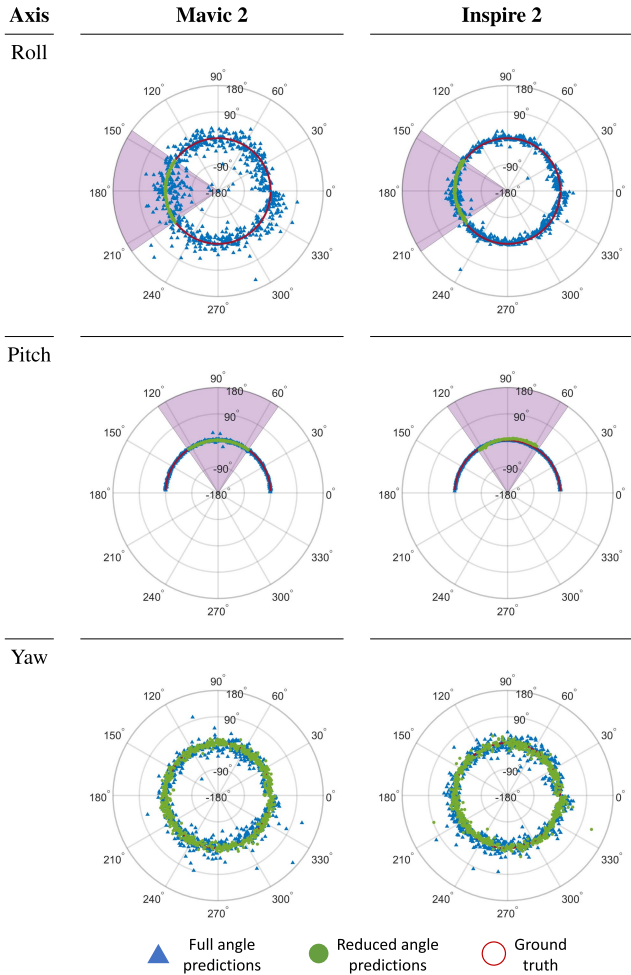


FIGURE 5. The results of the orientation prediction networks for the two drones in the full angle and reduced angle regimes. The theta coordinate represents the angle with the solid red ring indicating the ground truth. The radial coordinate represents the error (up to a maximum of $\pm 180^\circ$) with network under and over predictions falling inside of and outside of the red ring respectively. Predictions made by the network trained on the full range of angles are shown as blue triangles. Predictions made by the network trained on the reduced range of angles (indicated by the shaded region) are shown as green dots.

contingent on the symmetry of the drone. Specifically, the Mavic 2 is nearly perfectly symmetric about its roll axis, consequently the accuracy of the Mavic 2 roll network in the full angle regime is the worst. This is because there are the fewest features to unambiguously identify the roll at angles outside of a 90° to 270° range.

Examining the Intersection over Union (IoU) scores in Fig. 6 it is apparent that the networks can effectively segment both drones into their components regardless of their orientation. The score relating to the ‘body’ label is the highest in all cases indicating that the network is most accurate at predicting this component. This is likely because it is the most prevalent in terms of pixels in the image. Additionally, the fact that the rows and columns of the IoU scores do not sum to 100 indicates a conservative predictor. This means the network leaves some pixels (particularly around the perimeter

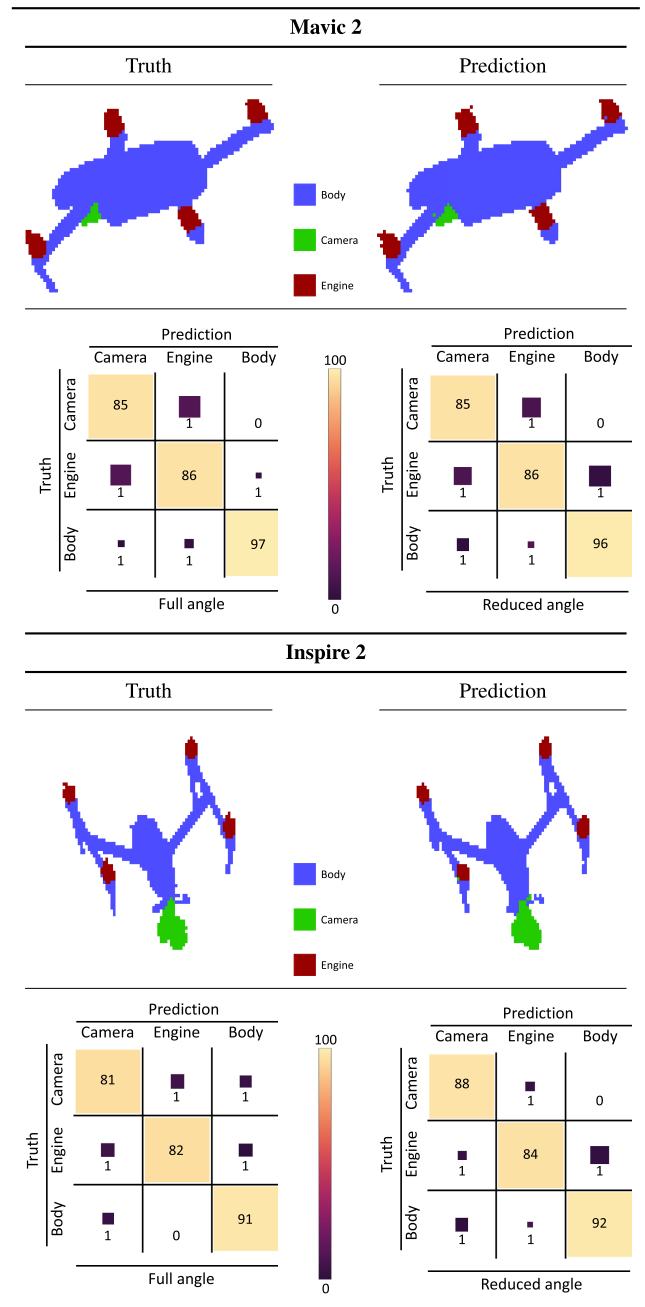


FIGURE 6. Qualitative and quantitative analyses of the segmentation networks. The quantitative analyses uses the IoU percentage for the two drones in the full angle and reduced angle regimes. The number in, as well as the size of each coloured region corresponds to the nearest integer IoU percentage. Note that the sizes of the regions have been scaled logarithmically for clearer representation.

of the drone) unclassified, reducing the total accuracy but also minimising misclassification.

B. REDUCED INPUT RESULTS

To further examine the functioning of the networks an ablation study was conducted. Specifically, the effect of removing one input channel, either the histograms or the intensity was quantified. Given that all networks share the DFE it was

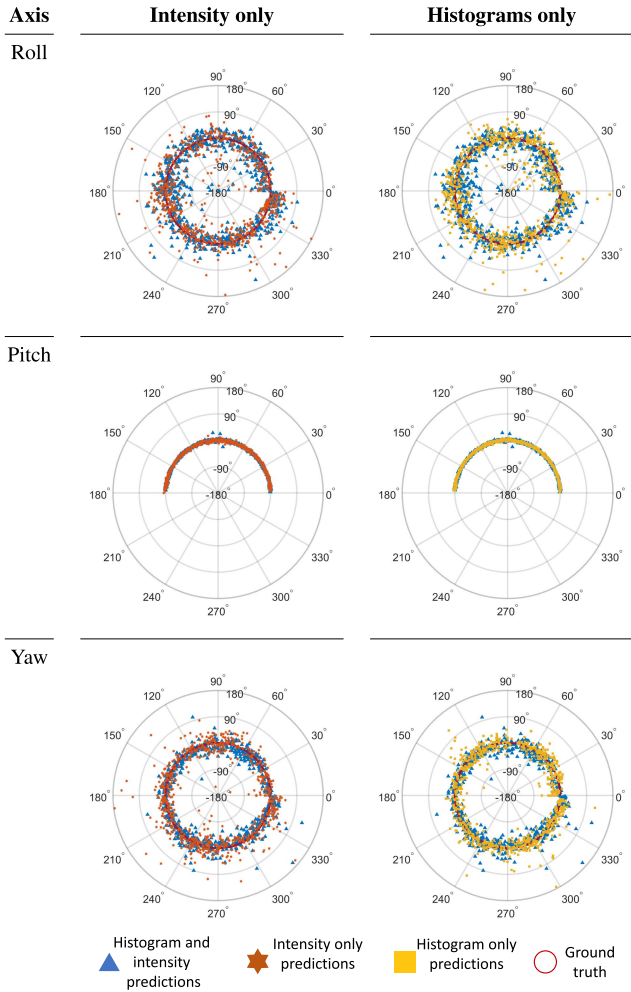


FIGURE 7. The results of the orientation prediction networks for the Mavic 2 drone in the full angle regime when trained using only an intensity or depth input. The theta coordinate represents the angle with the solid red ring indicating the ground truth. The radial coordinate represents the error (up to a maximum of $\pm 180^\circ$). Network under and over predictions fall inside of and outside of the red ring respectively. Predictions made by the networks trained on both inputs are shown as blue triangles. Predictions made by the networks trained on only intensity or depth data are shown as orange stars and yellow squares respectively.

determined to be sufficient to retrain only the orientation network for the Mavic 2 in the full angle regime since changes in performance in this network would be indicative of changes in all networks. Table 2 presents a summary of the findings with the network predictions visualized in Fig. 7. Table 2 and Fig. 7 indicate that the orientation of a drone can be more accurately determined from a depth input than an intensity input although the relative improvement is small. It should be noted however, that the images on which the network was trained do not contain a background. In real world cases where drones could be optically camouflaged the ability for depth sensing devices to isolate volumes of space ahead of background objects using time-of-flight gating may significantly enhance their robustness in orientation detection.

Additionally, given that the segmentation network can only reliably produce images up to the size of its largest input

TABLE 2. Summary of the Mavic 2’s orientation network accuracy when trained using only an intensity input or a depth input in the full angle regime.

Metric	Histogram only	Change
Orientation	(accuracy \pm std)(%)	(accuracy ; std)(%)
Roll	87.4 \pm 13.7	-0.9 ; +0.7
Pitch	99.1 \pm 0.9	-0.1 ; +0.1
Yaw	91.9 \pm 10.1	-0.4 ; +0.4

	Intensity only	Change
Orientation	(accuracy \pm std)(%)	(accuracy ; std)(%)
Roll	86.6 \pm 14.2	-1.7 ; +1.2
Pitch	98.8 \pm 1.2	-0.4 ; +0.4
Yaw	88.6 \pm 13.1	-3.6 ; +3.4

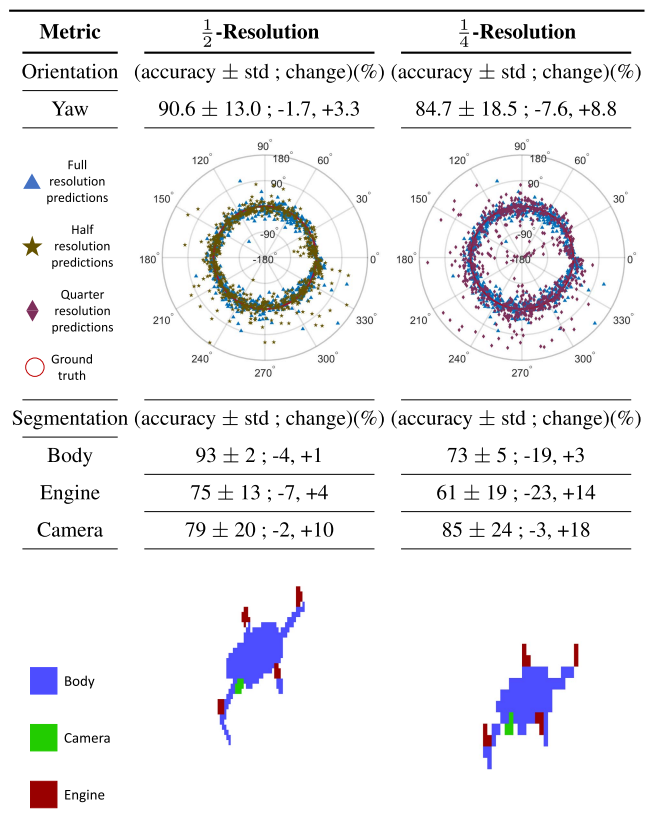


FIGURE 8. Summary of the Mavic 2’s orientation and segmentation network accuracy when trained using inputs at one half and one quarter resolution. The colour panels provide a qualitative illustration of network performance while the numbers report the accuracy and standard deviation as well as the change in those quantities with respect to the full resolution results in Table 1. Generally, as input resolution is reduced network performance worsens particularly in respect to the segmentation of small components on the drone, such as the engines.

(due to its U-Net structure) there is a benefit to providing the network with a high transverse resolution image. This benefit is illustrated in Fig. 8. Fig. 8 shows the degradation in accuracy of the yaw orientation and segmentation networks when trained on inputs at one-half and one-quarter of the





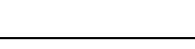
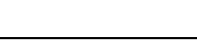
Metric	Ground truth	Prediction; Accuracy (%)
Identification	Mavic 2	Mavic 2; 100
Orientation		
Yaw	31°	19°; 93
Roll	180°	177°; 96
Pitch	90°	92°; 96
Metric	Reference	Prediction
Segmentation		
Body		
Camera		
Engine		

FIGURE 9. The predictions of the network trained on simulated data when applied to Quantic 4 × 4 data of a real drone in flight. The network correctly predicted the drone type and identified the roll, pitch and yaw with an average accuracy of 95%. The segmentation panels show the networks component prediction overlaid onto the intensity image (with the contrast enhanced intensity image shown alongside for reference). The network accurately segmented the drone into the body and engine components whilst not erroneously identifying a camera.

original resolution, as may be the case for a drone which is further away (or smaller) or a lower resolution sensor. As the resolution decreases both the accuracy and the precision (as shown by the increase in standard deviation) of the network decreases. This effect is particularly apparent in the segmentation of small features such as the engines and camera where low resolution images fail to retain the component-specific features on which the network relies for identification. Fig. 8 further illustrates the benefit of sensor fusion approaches which combine depth information with high transverse resolution images.

C. RESULTS ON REAL DATA

To demonstrate the real world applicability of our system, we applied the reduced angle network (trained only on simulated data) to an image of a real DJI Mavic 2 Zoom drone captured in flight using a Quantic 4 × 4 SPAD camera. Fig. 9 summarizes the predictions made by the networks and highlights their ability to fully characterise drones in real world conditions. The network correctly identified the drone type and suffered only a small loss in accuracy when performing the segmentation and orientation operations. This reduction in accuracy can be attributed to the reduction in quality between the simulated data and the input data from the Quantic 4 × 4 (as seen in Fig. 3 c))

IV. CONCLUSION

We present a CNN using a decision tree and ensemble structure to fully characterise i.e., determine the type, orientation and segmentation of drones in flight with accuracies in excess of 90%. We provide a system for the rapid generation of large quantities of accurately labelled photo-realistic training data

and demonstrate that this data is of sufficient fidelity to allow the system to accurately characterise real drones in flight. Our network provides a valuable tool in the image processing chain and can be used in combination with existing drone detection technologies to provide complete drone characterisation over wide areas. Finally, our approach may be readily extended to multiple 3D imaging and sensor fusion systems enabling pose detection for a wide range of vehicles.

REFERENCES

- [1] D. Gettinger and A. H. Michel, "Drone sightings and close encounters: An analysis," Center Study Drone, Bard College, Annandale-on-Hudson, NY, USA, Tech. Rep. 1, 2015.
- [2] A. M. Mehta, L. Tam, D. A. Greer, and K. Letheren, "Before crisis: How near-miss affects organizational trust and industry transference in emerging industries," *Public Relations Rev.*, vol. 46, no. 2, Jun. 2020, Art. no. 101886.
- [3] A. Rossiter, "Drone usage by militant groups: Exploring variation in adoption," *Defense Secur. Anal.*, vol. 34, no. 2, pp. 113–126, Apr. 2018.
- [4] T. Hammes, "The democratization of airpower: The insurgent and the drone," *War Rocks*, vol. 18, pp. 1–2, Oct. 2016.
- [5] D. Schneider, "Regulators seek ways to down rogue drones: Growing antidrone industry offers radar, remote ID, and other tools—[News]," *IEEE Spectr.*, vol. 56, no. 4, pp. 10–11, Apr. 2019.
- [6] P. Nguyen, T. Kim, J. Miao, D. Hesselius, E. Kenneally, D. Massey, E. Frew, R. Han, and T. Vu, "Towards RF-based localization of a drone and its controller," in *Proc. 5th Workshop Micro Aerial Vehicle Netw., Syst., Appl. (DroneNet)*, 2019, pp. 21–26.
- [7] J. O'Malley, "The no drone zone," *Eng. Technol.*, vol. 14, no. 2, pp. 34–38, Mar. 2019.
- [8] I. Bisio, C. Garibotto, F. Lavagetto, A. Sciarrone, and S. Zappatore, "Unauthorized amateur UAV detection based on WiFi statistical fingerprint analysis," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 106–111, Apr. 2018.
- [9] I. Bisio, C. Garibotto, F. Lavagetto, A. Sciarrone, and S. Zappatore, "Improving WiFi statistical fingerprint-based detection techniques against UAV stealth attacks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, New York, NY, USA, Dec. 2018, pp. 1–6.
- [10] H. Zhang, C. Cao, L. Xu, and T. A. Gulliver, "A UAV detection algorithm based on an artificial neural network," *IEEE Access*, vol. 6, pp. 24720–24728, 2018.
- [11] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, "Micro-UAV detection and classification from RF fingerprints using machine learning techniques," in *Proc. IEEE Aerosp. Conf.*, New York, NY, USA, Mar. 2019, pp. 1–13.
- [12] M. S. Allahham, T. Khattab, and A. Mohamed, "Deep learning for RF-based drone detection and identification: A multi-channel 1-D convolutional neural networks approach," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 112–117.
- [13] D. Shorten, A. Williamson, S. Srivastava, and J. C. Murray, "Localisation of drone controllers from RF signals using a deep learning approach," in *Proc. Int. Conf. Pattern Recognit. Artif. Intell. (PRAI)*, New York, NY, USA, 2018, pp. 89–97.
- [14] S. Al-Emadi and F. Al-Senaid, "Drone detection approach based on radio-frequency using convolutional neural network," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 29–34.
- [15] H. Chen, Z. Wang, and L. Zhang, "Collaborative spectrum sensing for illegal drone detection: A deep learning-based image classification perspective," *China Commun.*, vol. 17, no. 2, pp. 81–92, Feb. 2020.
- [16] M. U. Sheikh, F. Ghavimi, K. Ruttik, and R. Jantti, "Drone detection and classification using cellular network: A machine learning approach," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, New York, NY, USA, Sep. 2019, pp. 1–6.
- [17] A. Bello, B. Biswal, S. Shetty, C. Kamhoua, and K. Gold, "Radio frequency classification toolbox for drone detection," *Proc. SPIE*, vol. 11006, May 2019, Art. no. 110061Y.
- [18] H. Ryden, S. B. Redhwan, and X. Lin, "Rogue drone detection: A machine learning approach," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, New York, NY, USA, Apr. 2019, pp. 1–6.

- [19] A. Shoufan, H. M. Al-Angari, M. F. A. Sheikh, and E. Damiani, "Drone pilot identification by classifying radio-control signals," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 10, pp. 2439–2447, Oct. 2018.
- [20] D. Mototolea, R. Youssef, E. Radoi, and I. Nicolaescu, "Non-cooperative low-complexity detection approach for FHSS-GFSK drone control signals," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 401–412, 2020.
- [21] S. Basak, S. Rajendran, S. Pollin, and B. Scheers, "Drone classification from RF fingerprints using deep residual nets," in *Proc. Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2021, pp. 548–555.
- [22] A. Gumaedi, M. Al-Rakhami, M. M. Hassan, P. Pace, G. Alai, K. Lin, and G. Fortino, "Deep learning and blockchain with edge computing for 5G-enabled drone identification and flight mode detection," *IEEE Netw.*, vol. 35, no. 1, pp. 94–100, Jan. 2021.
- [23] I. Nemer, T. Sheltami, I. Ahmad, A. U. Yasar, and M. A. R. Abdeen, "RF-based UAV detection and identification using hierarchical learning approach," *Sensors*, vol. 21, no. 6, p. 1947, 2021.
- [24] H. Kolamunna, T. Dahanayaka, J. Li, S. Seneviratne, K. Thilakarathne, A. Y. Zomaya, and A. Seneviratne, "Droneprint: Acoustic signatures for open-set drone detection and identification with online data," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 1, pp. 1–31, 2021.
- [25] S. Song, Y. Son, and Y. Kim, "Flying drone classification based on visualization of acoustic signals with deep neural networks," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2020, pp. 546–548.
- [26] J. García-Gómez, M. Bautista-Durán, R. Gil-Pita, I. Mohino-Herranz, M. Aguilar-Ortega, and C. Clares-Crespo, "Cost-constrained drone presence detection through smart sound processing," in *Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods*, 2019, pp. 1–7.
- [27] M. Z. Anwar, Z. Kaleem, and A. Jamalipour, "Machine learning inspired sound-based amateur drone detection for public safety applications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2526–2534, Mar. 2019.
- [28] B. Yang, E. T. Matson, A. H. Smith, J. E. Dietz, and J. C. Gallagher, "UAV detection system with multiple acoustic nodes using machine learning models," in *Proc. 3rd IEEE Int. Conf. Robotic Comput. (IRC)*, New York, NY, USA, Feb. 2019, pp. 493–498.
- [29] S. Salman, J. Mir, M. T. Farooq, A. N. Malik, and R. Haleemdeen, "Machine learning inspired efficient audio drone detection using acoustic features," in *Proc. Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, New York, NY, USA, Jan. 2021, pp. 335–339.
- [30] Y. J. He, I. Ahmad, L. Shi, and K. Chang, "SVM-based drone sound recognition using the combination of HLA and WPT techniques in practical noisy environment," *KSH Trans. Internet Inf. Syst.*, vol. 13, no. 10, pp. 5078–5094, 2019.
- [31] G. Ciaburro and G. Iannace, "Improving smart cities safety using sound events detection based on deep neural network algorithms," *Informatics*, vol. 7, no. 3, p. 23, 2020.
- [32] S. Jeon, J.-W. Shin, Y.-J. Lee, W.-H. Kim, Y. Kwon, and H.-Y. Yang, "Empirical study of drone sound detection in real-life environment with deep neural networks," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, New York, NY, USA, Aug. 2017, pp. 1858–1862.
- [33] S. Al-Emadi, A. Al-Ali, and A. Al-Ali, "Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks," *Sensors*, vol. 21, no. 15, p. 4953, 2021.
- [34] S. Al-Emadi, A. Al-Ali, A. Mohammad, and A. Al-Ali, "Audio based drone detection and identification using deep learning," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, New York, NY, USA, Jun. 2019, pp. 459–464.
- [35] M. Ohlenbusch, A. Ahrens, C. Rollwage, and J. Bitzer, "Robust drone detection for acoustic monitoring applications," in *Proc. 28th Eur. Signal Process. Conf. (EUSIPCO)*, New York, NY, USA, Jan. 2021, pp. 6–10.
- [36] Y. Seo, B. Jang, and S. Im, "Drone detection using convolutional neural networks with acoustic STFT features," in *Proc. 15th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, New York, NY, USA, Nov. 2018, pp. 1–6.
- [37] J. Kim, C. Park, J. Ahn, Y. Ko, J. Park, and J. C. Gallagher, "Real-time UAV sound detection and analysis system," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, New York, Mar. 2017, pp. 1–5.
- [38] N. Siriphun, S. Kashihara, D. Fall, and A. Khurat, "Distinguishing drone types based on acoustic wave by IoT device," in *Proc. 22nd Int. Comput. Sci. Eng. Conf. (ICSEC)*, Nov. 2018, pp. 1–4.
- [39] S. Yoo and H. Oh, "Analysis of commercial drone sounds and its identification," in *Proc. Int. Conf. Res. Adapt. Convergent Syst.*, Oct. 2020, pp. 47–52.
- [40] A. E. Ananenkov, D. V. Marin, V. M. Nuzhdin, V. V. Rastorguev, and P. V. Sokolov, "Possibilities to observe small-size UAVs in the prospective airfield radar," in *Proc. 20th Int. Conf. Transparent Opt. Netw. (ICTON)*, New York, NY, USA, Jul. 2018, pp. 1–6.
- [41] T. Al-Nuaim, M. Alam, and A. Aldowesh, "Low-cost implementation of a multiple-input multiple-output radar prototype for drone detection," in *Proc. Int. Symp. ELMAR*, New York, NY, USA, Sep. 2019, pp. 183–186.
- [42] A. Aldowesh, T. BinKhamis, T. Alnuaim, and A. Alzogaiby, "Low power digital array radar for drone detection and micro-Doppler classification," in *Proc. Signal Process. Symp. (SPSymo)*, Sep. 2019, pp. 203–206.
- [43] Y. Zhao and Y. Su, "The extraction of micro-Doppler signal with EMD algorithm for radar-based small UAVs' detection," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 3, pp. 929–940, Mar. 2020.
- [44] Y. Zhao and Y. Su, "Cyclostationary phase analysis on micro-Doppler parameters for radar-based small UAVs detection," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 9, pp. 2048–2057, Sep. 2018.
- [45] H. Sun, B.-S. Oh, X. Guo, and Z. Lin, "Improving the Doppler resolution of ground-based surveillance radar for drone detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 6, pp. 3667–3673, Dec. 2019.
- [46] A. Huang, P. Sevigny, B. Balaji, and S. Rajan, "Radar micro-Doppler-based rotary drone detection using parametric spectral estimation methods," in *Proc. IEEE SENSORS*, New York, NY, USA, Oct. 2020, pp. 1–4.
- [47] J. F. Ren and X. D. Jiang, "A three-step classification framework to handle complex data distribution for radar UAV detection," *Pattern Recognit.*, vol. 111, Mar. 2021, Art. no. 107709.
- [48] S. Zulkifli and A. Balleri, "Design and development of K-band FMCW radar for nano-drone detection," in *Proc. IEEE Radar Conf. (RadarConf20)*, New York, NY, USA, Sep. 2020, pp. 1–5.
- [49] W. Y. Zhang and G. Li, "Detection of multiple micro-drones via cadence velocity diagram analysis," *Electron. Lett.*, vol. 54, no. 7, pp. 441–442, 2018.
- [50] G. Sacco, E. Pittella, S. Pisa, and E. Piuze, "A MISO radar system for drone localization," in *Proc. 5th IEEE Int. Workshop Metrol. Aerosp. (MetroAeroSpace)*, New York, NY, USA, Jun. 2018, pp. 549–553.
- [51] C. Özdemir, "Radar cross section analysis of unmanned aerial vehicles using predicys," *Int. J. Eng. Geosci.*, vol. 5, no. 3, pp. 144–149, Oct. 2020.
- [52] V. Semkin, M. Yin, Y. Hu, M. Mezzavilla, and S. Rangan, "Drone detection and classification based on radar cross section signatures," in *Proc. Int. Symp. Antennas Propag. (ISAP)*, New York, NY, USA, Jan. 2021, pp. 223–224.
- [53] A. D. De Quevedo, F. I. Urzaiz, J. G. Menoyo, and A. A. Lopez, "Drone detection and RCS measurements with ubiquitous radar," in *Proc. Int. Conf. Radar (RADAR)*, New York, NY, USA, Aug. 2018, pp. 1–6.
- [54] S. Rzewuski, K. Kulpa, B. Salski, P. Kopyt, K. Borowiec, M. Malanowski, and P. Samczynski, "Drone RCS estimation using simple experimental measurement in the WiFi bands," in *Proc. 22nd Int. Microw. Radar Conf. (MIKON)*, May 2018, pp. 695–698.
- [55] M. Guo, Y. Lin, Z. Sun, and Y. Fu, "Research on monostatic radar cross section simulation of small unmanned aerial vehicles," in *Proc. Int. Conf. Millim. Wave Technol. (ICMMT)*, New York, NY, USA, May 2018, pp. 1–3.
- [56] R. Nakamura and H. Hadama, "Characteristics of ultra-wideband radar echoes from a drone," *IEICE Commun. Exp.*, vol. 6, no. 9, pp. 530–534, 2017.
- [57] R. Nakamura, H. Hadama, and A. Kajiwara, "Ultra-wideband radar reflectivity of a drone in millimeter wave band," *IEICE Commun. Exp.*, vol. 7, no. 9, pp. 341–346, 2018.
- [58] T. Mizushima, R. Nakamura, and H. Hadama, "Reflection characteristics of ultra-wideband radar echoes from various drones in flight," in *Proc. IEEE Top. Conf. Wireless Sensors Sensor Netw. (WiSNET)*, New York, NY, USA, Jan. 2020, pp. 30–33.
- [59] M. Passafiume, N. Rojhani, G. Collodi, and A. Cidronali, "Modeling small UAV micro-Doppler signature using millimeter-wave FMCW radar," *Electronics*, vol. 10, no. 6, p. 747, 2021.
- [60] J. Ochodnický, Z. Matousek, M. Babjak, and J. Kurty, "Drone detection by Ku-band battlefield radar," in *Proc. Int. Conf. Mil. Technol. (ICMT)*, May 2017, pp. 613–616.
- [61] J. S. Patel, F. Fioranelli, and D. Anderson, "Review of radar classification and RCS characterisation techniques for small UAVs or drones," *IET Radar, Sonar Navigat.*, vol. 12, no. 9, pp. 911–919, 2018.
- [62] S. A. Musa, R. Abdullah, A. Sali, A. Ismail, N. E. A. Rashid, I. P. Ibrahim, and A. A. Salah, "A review of copter drone detection using radar systems," *Def. S&T Tech. Bull.*, vol. 12, no. 1, pp. 16–38, 2019.

- [63] 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, 2020.
- [64] B. Taha and A. Shoufan, "Machine learning-based drone detection and classification: State-of-the-art in research," *IEEE Access*, vol. 7, pp. 138669–138682, 2019.
- [65] A. Carrio, S. Vemprala, A. Ripoll, S. Saripalli, and P. Campoy, "Drone detection using depth maps," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, New York, NY, USA, Oct. 2018, pp. 1032–1037.
- [66] J. Ryu and S. Kim, "Small infrared target detection by data-driven proposal and deep learning-based classification," vol. 10624, May 2018, Art. no. 106241J.
- [67] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.
- [68] D. K. Behera and A. Bazil Raj, "Drone detection and classification using deep learning," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 1012–1016.
- [69] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [70] X. D. Zhang and K. Chandramouli, "Critical infrastructure security against drone attacks using visual analytics," in *Computer Vision Systems (Lecture Notes in Computer Science)*, vol. 11754. Cham, Switzerland: Springer, 2019, pp. 713–722.
- [71] J. Park, D. H. Kim, Y. S. Shin, and S.-H. Lee, "A comparison of convolutional object detectors for real-time drone tracking using a PTZ camera," in *Proc. 17th Int. Conf. Control, Autom. Syst. (ICCAS)*, New York, NY, USA, Oct. 2017, pp. 696–699.
- [72] H. Liu, F. C. Qu, Y. J. Liu, W. Zhao, and Y. T. Chen, "A drone detection with aircraft classification based on a camera array," in *Proc. IOP Conf., Mater. Sci. Eng. Bristol, U.K.*, vol. 322, 2018, Art. no. 052005.
- [73] E. Unlu, E. Zenou, N. Riviere, and P.-E. Dupouy, "Deep learning-based strategies for the detection and tracking of drones using several cameras," *IPSPJ Trans. Comput. Vis. Appl.*, vol. 11, no. 1, pp. 1–13, Dec. 2019.
- [74] M. Wu, W. Xie, X. Shi, P. Shao, and Z. Shi, "Real-time drone detection using deep learning approach," in *Proc. Int. Conf. Mach. Learn. Intell. Commun.* New York, NY, USA: Springer, 2018, pp. 22–32.
- [75] S. A. Hassan, T. Rahim, and S. Y. Shin, "Real-time UAV detection based on deep learning network," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, New York, NY, USA, Oct. 2019, pp. 630–632.
- [76] A. Koksall, K. G. Ince, and A. Aydin Alatan, "Effect of annotation errors on drone detection with YOLOv3," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 1030–1031.
- [77] K. Madasamy, V. Shanmuganathan, V. Kandasamy, M. Y. Lee, and M. Thangadurai, "OSDDY: Embedded system-based object surveillance detection system with small drone using deep Yolo," *EURASIP J. Image Video Process.*, vol. 2021, no. 1, p. 14, Dec. 2021.
- [78] A. Schumann, L. Sommer, J. Klatte, T. Schuchert, and J. Beyerer, "Deep cross-domain flying object classification for robust UAV detection," in *Proc. 14th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, New York, NY, USA, Aug. 2017, pp. 1–6.
- [79] A. Coluccia *et al.*, "Drone-vs-bird detection challenge at IEEE AVSS2019," in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, New York, NY, USA, Sep. 2019, pp. 1–7.
- [80] C. Craye and S. Ardjoune, "Spatio-temporal semantic segmentation for drone detection," in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, New York, NY, USA, Sep. 2019, pp. 1–5.
- [81] A. Coluccia, A. Fascista, A. Schumann, L. Sommer, A. Dimou, D. Zarpalas, M. Mendez, D. de la Iglesia, I. Gonzalez, J. P. Mercier, G. Gagne, A. Mitra, and S. Rajashekar, "Drone vs. bird detection: Deep learning algorithms and results from a grand challenge," *Sensors*, vol. 21, no. 8, p. 27, 2021.
- [82] L. Du, C. Q. Gao, Q. Feng, C. Wang, and J. Liu, "Small UAV detection in videos from a single moving camera," in *Computer Vision (Communications in Computer and Information Science)*, vol. 773. Singapore: Springer-Verlag, 2017, pp. 187–197.
- [83] U. Seidaliev, D. Akhmetov, L. Ilibayeva, and E. T. Matson, "Real-time and accurate drone detection in a video with a static background," *Sensors*, vol. 20, no. 14, p. 3856, 2020.
- [84] A. Sharjeel, S. A. Z. Naqvi, and M. Ahsan, "Real time drone detection by moving camera using COROLA and CNN algorithm," *J. Chin. Inst. Engineers*, vol. 44, no. 2, pp. 128–137, Feb. 2021.
- [85] P. Neduchal, F. Berka, and M. Zelezny, "Stationary device for drone detection in urban areas," in *Interactive Collaborative Robotics (Lecture Notes in Artificial Intelligence)*, vol. 10459. Cham, Switzerland: Springer, 2017, pp. 162–169.
- [86] D. Lee, W. Gyu La, and H. Kim, "Drone detection and identification system using artificial intelligence," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, New York, NY, USA, Oct. 2018, pp. 1131–1133.
- [87] P. A. Prates, R. Mendonca, A. Lourenco, F. Marques, J. P. Matos-Carvalho, and J. Barata, "Vision-based UAV detection and tracking using motion signatures," in *Proc. IEEE Ind. Cyber-Phys. Syst. (ICPS)*, May 2018, pp. 482–487.
- [88] K. Abbasi, A. Batool, M. A. Asghar, A. Saeed, M. J. Khan, and M. ur Rehman, "A vision-based amateur drone detection algorithm for public safety applications," in *Proc. U.K./China Emerg. Technol. (UCET)*, Aug. 2019, pp. 1–5.
- [89] A. Fernandes, M. Baptista, L. Fernandes, and P. Chaves, "Drone, aircraft and bird identification in video images using object tracking and residual neural networks," in *Proc. 11th Int. Conf. Electron., Comput. Artif. Intell. (ECAI)*, New York, NY, USA, Jun. 2019.
- [90] A. J. Garcia, J. Min Lee, and D. S. Kim, "Anti-drone system: A visual-based drone detection using neural networks," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2020, pp. 559–561.
- [91] S. Srigrarom, N. J. L. Sie, H. Cheng, K. H. Chew, M. Lee, and P. Ratsamee, "Multi-camera multi-drone detection, tracking and localization with trajectory-based re-identification," in *Proc. 2nd Int. Symp. Instrum., Control, Artif. Intell., Robot. (ICA-SYMP)*, New York, NY, USA, Jan. 2021, pp. 1–6.
- [92] Q. Dong and Q. Zou, "Visual UAV detection method with online feature classification," in *Proc. IEEE 2nd Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, New York, NY, USA, Dec. 2017, pp. 429–432.
- [93] Z. Wang, L. Qi, Y. Tie, Y. Ding, and Y. Bai, "Drone detection based on FD-HOG descriptor," in *Proc. Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discovery (CyberC)*, New York, NY, USA, Oct. 2018, pp. 433–436.
- [94] H. Sun, W. Geng, J. Q. Shen, N. Z. Liu, D. Liang, and H. Y. Zhou, "Deeper SSD: Simultaneous up-sampling and down-sampling for drone detection," *KSII Trans. Internet Inf. Syst.*, vol. 14, no. 12, pp. 4795–4815, 2020.
- [95] N. Shijith, P. Poornachandran, V. G. Sujadevi, and M. M. Dharmana, "Breach detection and mitigation of UAVs using deep neural network," in *Proc. Recent Develop. Control, Autom. Power Eng. (RDCAPE)*, New York, NY, USA, Oct. 2017, pp. 360–365.
- [96] E. Unlu, E. Zenou, and N. Riviere, "Generic Fourier descriptors for autonomous UAV detection," in *Proc. 7th Int. Conf. Pattern Recognit. Appl. Methods*, Setubal, Portugal, 2018, pp. 550–554.
- [97] E. Unlu, E. Zenou, and N. Riviere, "Using shape descriptors for UAV detection," *Electron. Imag.*, vol. 2018, no. 9, pp. 128-1–128-5, 2018.
- [98] M. Nalamati, A. Kapoor, M. Saqib, N. Sharma, and M. Blumenstein, "Drone detection in long-range surveillance videos," in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, New York, NY, USA, Sep. 2019, pp. 1–6.
- [99] H. Sun, J. Yang, J. Shen, D. Liang, L. Ning-Zhong, and H. Zhou, "TIB-net: Drone detection network with tiny iterative backbone," *IEEE Access*, vol. 8, pp. 130697–130707, 2020.
- [100] L. Tao, T. Hong, Y. Guo, H. Chen, and J. Zhang, "Drone identification based on CenterNet-TensorRT," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, New York, NY, USA, Oct. 2020, pp. 1–5.
- [101] A. Hernandez, C. Copot, R. De Keyser, T. Vlas, and I. Nascu, "Identification and path following control of an AR.Drone quadrotor," in *Proc. 17th Int. Conf. Syst. Theory, Control Comput. (ICSTCC)*, New York, NY, USA, Oct. 2013, pp. 583–588.
- [102] C. Shinde, R. Lima, and K. Das, "Multi-view geometry and deep learning based drone detection and localization," in *Proc. 5th Indian Control Conf. (ICC)*, New York, NY, USA, Jan. 2019, pp. 289–294.
- [103] L. Wang, J. Ai, L. Zhang, and Z. Xing, "Design of airport obstacle-free zone monitoring UAV system based on computer vision," *Sensors*, vol. 20, no. 9, p. 2475, 2020.
- [104] D. T. Wei Xun, Y. L. Lim, and S. Srigrarom, "Drone detection using YOLOv3 with transfer learning on NVIDIA Jetson TX2," in *Proc. 2nd Int. Symp. Instrum., Control, Artif. Intell., Robot. (ICA-SYMP)*, New York, NY, USA, Jan. 2021, pp. 1–6.
- [105] C. Ruiz, X. L. Chen, and P. Zhang, "Poster abstract: Hybrid and adaptive drone identification through motion actuation and vision feature matching," in *Proc. 16th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, New York, NY, USA, Apr. 2017, pp. 327–328.

- [106] D. H. Lee, "CNN-based single object detection and tracking in videos and its application to drone detection," *Multimedia Tools Appl.*, vol. 80, pp. 34237–34248, Oct. 2021.
- [107] M. Saqib, S. Daud Khan, N. Sharma, and M. Blumenstein, "A study on detecting drones using deep convolutional neural networks," in *Proc. 14th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, New York, NY, USA, Aug. 2017, pp. 1–5.
- [108] H. M. Oh, H. Lee, and M. Y. Kim, "Comparing convolutional neural Network(CNN) models for machine learning-based drone and bird classification of anti-drone system," in *Proc. 19th Int. Conf. Control, Autom. Syst. (ICCAS)*, New York, NY, USA, Oct. 2019, pp. 87–90.
- [109] U. Seidaliev, M. Alduraibi, L. Ilipbayeva, and A. Almagambetov, "Detection of loaded and unloaded UAV using deep neural network," in *Proc. 4th IEEE Int. Conf. Robotic Comput. (IRC)*, Nov. 2020, pp. 490–494.
- [110] R. Jin, J. Q. Jiang, Y. H. Qi, D. F. Lin, and T. Song, "Drone detection and pose estimation using relational graph networks," *Sensors*, vol. 19, no. 6, p. 1479, 2019.
- [111] S. Jamil, M. Rahman, A. Ullah, S. Badnava, M. Forsat, and S. S. Mirjavadi, "Malicious UAV detection using integrated audio and visual features for public safety applications," *Sensors*, vol. 20, no. 14, p. 3923, 2020.
- [112] H. Liu, Z. Wei, Y. Chen, J. Pan, L. Lin, and Y. Ren, "Drone detection based on an audio-assisted camera array," in *Proc. IEEE 3rd Int. Conf. Multimedia Big Data (BigMM)*, New York, NY, USA, Apr. 2017, pp. 402–406.
- [113] B. H. Kim, D. Khan, C. Bohak, W. Choi, H. J. Lee, and M. Y. Kim, "V-RBNN based small drone detection in augmented datasets for 3D LADAR system," *Sensors*, vol. 18, no. 11, p. 3825, 2018.
- [114] B. H. Kim, D. Khan, W. Choi, and M. Y. Kim, "Real-time counter-UAV system for long distance small drones using double pan-tilt scan laser radar," *Proc. SPIE*, vol. 11005, May 2019, Art. no. 110050C.
- [115] I. Gyongy, S. W. Hutchings, A. Halimi, M. Tyler, S. Chan, F. Zhu, S. McLaughlin, R. K. Henderson, and J. Leach, "High-speed 3D sensing via hybrid-mode imaging and guided upsampling," *Optica*, vol. 7, no. 10, pp. 1253–1260, 2020.
- [116] G. Mora-Martín, A. Turpin, A. Ruget, A. Halimi, R. Henderson, J. Leach, and I. Gyongy, "High-speed object detection with a single-photon time-of-flight image sensor," *Opt. Exp.*, vol. 29, no. 21, pp. 33184–33196, 2021.
- [117] S. W. Hutchings, N. Johnston, I. Gyongy, T. Al Abbas, N. A. W. Dutton, M. Tyler, S. Chan, J. Leach, and R. K. Henderson, "A reconfigurable 3-D-stacked SPAD imager with in-pixel histogramming for flash LIDAR or high-speed time-of-flight imaging," *IEEE J. Solid-State Circuits*, vol. 54, no. 11, pp. 2947–2956, Nov. 2019.
- [118] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2015, pp. 234–241.
- [119] Y. Wang, Y. Chen, J. Choi, and C.-C.-J. Kuo, "Towards visible and thermal drone monitoring with convolutional neural networks," *APSIPA Trans. Signal Inf. Process.*, vol. 8, no. 1, pp. 1–13, 2019.
- [120] Y. Chen, P. Aggarwal, J. Choi, and C.-C.-J. Kuo, "A deep learning approach to drone monitoring," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, New York, NY, USA, Dec. 2017, pp. 686–691.
- [121] V. Magoulaniotis, D. Ataloglou, A. Dimou, D. Zarpalas, and P. Daras, "Does deep super-resolution enhance UAV detection?" in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Sep. 2019, pp. 1–6.
- [122] S. R. Yellapantula, "Synthesizing realistic data for vision based drone-drone detection," M.S. thesis, Dept. Comput. Eng., Virginia Polytech. Inst. State Univ., Blacksburg, VA, USA, 2019.



ALICE RUGET received the M.Sc. degree in electrical engineering from the CentraleSupélec, Gif-sur-Yvette, France, in 2017, and the M.Sc. degree in biomedical engineering from ETH Zürich, Switzerland, in 2019. She is currently pursuing the Ph.D. degree in computational imaging for ultrafast imaging in three dimensions with the HW Quantum Group, Heriot-Watt University, Edinburgh, U.K. Her research interest includes the development of machine learning algorithms to enhance the image quality for different imaging systems, such as single-photon sensitive detectors.



GERMÁN MORA-MARTÍN received the B.Sc. degree in nanoscience and nanotechnology from the Universitat Autònoma de Barcelona (UAB), Cerdanyola del Vallès, Spain, in 2017, the M.Sc. degree in physique, spécialité optique, photonique, and signal et image from Aix-Marseille Université, Marseille, France, and the M.Sc. degree in optics and photonics from Karlsruher Institut für Technologie (KIT), Karlsruhe, Germany, in 2019, through Europhotonics, an Erasmus Mundus Program. He is currently pursuing the Ph.D. degree with The University of Edinburgh under the supervision of Istvan Gyongy, working on high-speed 3D imaging with SPAD sensors. He develops algorithms for object detection and super-resolution using the data produced by SPAD sensors.



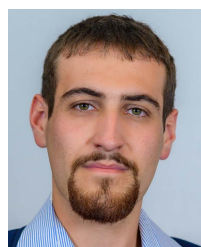
FENG ZHU received the B.Sc. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2011 and 2017, respectively. Since 2017, he has been a Research Associate with the Department of Physics, Heriot-Watt University. His research interests include quantum and classical optics and imaging.



ISTVAN GYONGY received the M.Eng. and Ph.D. degrees from the University of Oxford, U.K. Following a period in industry, where he worked on processors for smartphones and on a cloud-connected activity tracking systems for dairy farms, he joined The University of Edinburgh, U.K. He is currently a Lecturer and is developing single-photon avalanche diode (SPAD) cameras and exploring applications in 3D capture as well as in the life sciences.



JONATHAN LEACH received the M.Sc. degree in physics and the Ph.D. degree from the University of Glasgow, in 2002 and 2004, respectively. He was a Senior Research Associate with the Quantum Photonics Group, University of Ottawa, Ottawa, ON, Canada. He joined Heriot-Watt University, Edinburgh, U.K., in 2012, to establish a research program in experimental quantum optics, where he is currently an Associate Professor. He has written over 100 peer-reviewed articles in scientific journals. His research interests include applying classical and quantum optics techniques to solve problems in information and imaging science.



STIRLING SCHOLES received the B.Sc. degree (Hons.) in physics and the M.Sc. degree (Hons.) in optics and photonics from the University of the Witwatersrand, Johannesburg, South Africa, in 2018 and 2020, respectively. He is currently pursuing the Ph.D. degree in applied imaging systems with the Quantum Optics Group, Heriot-Watt University, with a focus on data fusion approaches for high-speed tracking and identification of objects using 3D time-of-flight technology.