Multi-target Angle of Arrival Estimation Using Rotating mmWave FMCW Radar and Yolov3

A. N. Wilson, A. Kumar, A. Jha, L. R. Cenkeramaddi

Abstract: **It is still challenging to accurately localize unmanned aerial vehicles (UAVs) from a ground control station (GCS) using various sensors. The mmWave frequency modulated continuous wave (FMCW) radars offer excellent performance for target detection and localization in harsh environments and low lighting conditions. However, the estimated angle of arrival (AoA) of targets in the captured scene is quite poor. This article focuses on improving AoA estimation by combining cutting-edge machine learning algorithms with a mechanical radar rotor setup. A mmWave FMCW radar system is mounted on a programmable rotor to capture range-angle maps of targets at various locations. The range-angle images are then labeled and trained further with the Yolov3 algorithm. Subsequent testing reveals that for detected target objects, the centroid of the bounding boxes from the detected objects provides accurate AoA estimation with very low root mean square error (RMSE). The results show that the proposed approach outperforms traditional methods in terms of performance and estimation accuracy.**

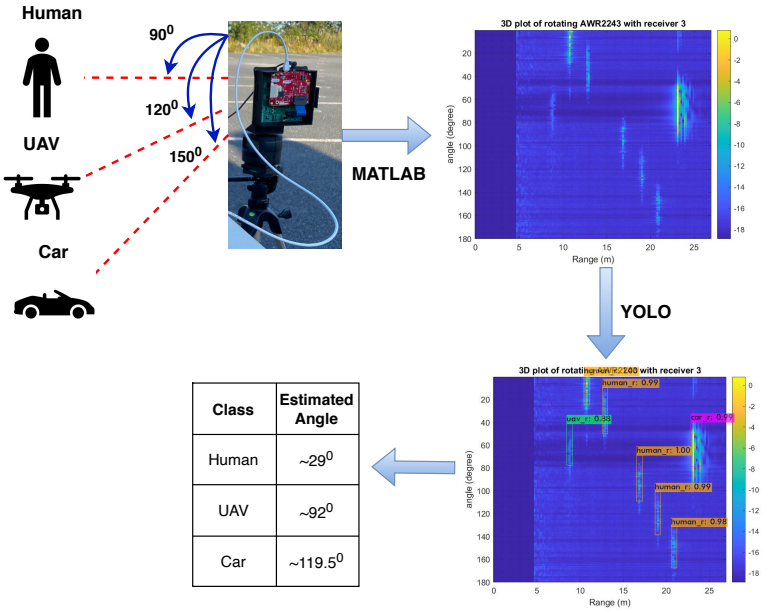


Figure E.1: Block diagram of the proposed multi-target AoA estimation approach.

E.1 Introduction

As the popularity of unmanned aerial vehicles (UAV) increases, there is a growing demand for better detection, localization and tracking techniques. Ground control stations (GCS) are equipped with sophisticated sensor technologies for improved UAV localization. Some of these sensors are RGB cameras, ultrasonic sensors, LiDARs, etc. However most of these sensors fail in adverse weather and lighting conditions. In this regard, the mmWave frequency modulated continuous wave (FMCW) radars have shown superior performance compared to other sensors. The mmWave FMCW radars are small, lightweight, and compact radars that offer excellent performance for target detection. The radars provide a radial distance measurement range from 0.2 meters to 300 meters in addition to excellent velocity estimation performance. Additionally, they offer high range and velocity resolutions. Further, their performance in adverse weather and environmental conditions makes them a suitable choice for UAV and GCS applications.

Despite the above-mentioned advantages, the mmWave FMCW radars suffer from poor angle of arrival (AoA) estimation and AoA resolutions [1]. Accurate AoA estimation from a single target requires atleast one transmitter and two receivers for the radar. Increasing the number of transmitter and receiver pairs can effectively improve the AoA estimation, however this results in a tradeoff with increased size and hardware complexity. Furthermore, estimation of AoA from multiple objects for better target localization is even more challenging and hence it is still an ongoing research topic.

There has been several attempts in the literature to improve the AoA estimation. The authors in [2] propose an adaptive radar signal processor for detection of multiple UAVs in the range-Doppler domain. The radar signal processor operates by initially performing pulse compression (PC), FFT and beamforming on the received signal samples. The processed signal is then evaluated for a possible detection based on constant false alarm rate (CFAR) logic for each range-Doppler angle bin. Both single and multi-target scenarios are considered and the angular parameters of the targets are retrieved. Simulated and real-world experiment results indicate that the proposed method is able to detect and resolve two UAVs placed $\approx 8^\circ$ apart with satisfactory performance. In [3], Aubry *et al.* attempt to solve the problem of adaptive radar detection in the presence of limited training data. The proposed method consists of three generalised likelihood ratio test (GLRT) based detection schemes that is based on the assumption of linear combination for inverse covariance matrices. Reported results show that the proposed GLRT-1 and GLRT-3 exhibit significant performance gain over conventional adaptive detectors in presence of limited training samples. In [4], a novel DoA estimation algorithm based on 2-D spectrum sensing is developed. The proposed approach exploits the intrinsic block-sparsity of the 2-D space-frequency profile to obtain 2-D occupancy awareness. By including a term to account for the block-sparsity, a non-convex regularized maximum likelihood (RML) estimation problem is solved using the block sparse learning via iterative minimization (BSLIM) algorithm to obtain the space-

frequency profile. The BSLIM algorithm exhibits superior performance in terms of detection rate and false alarms over traditional approaches. In [5], the authors devise a novel set of covariance matrix estimators called median matrices that are independent of the probability distribution of the samples. The matrices are constructed by exploiting its positive definite attributes. Based on the output from the new estimators, a generalised inner product (GIP) selection criteria is utilized to discard the secondary outliers in the training data. Reported results indicate that the log-Euclidean median-based estimator outperforms other estimators in terms of selection probability of secondary outliers in the training data. Other covariance estimators that can also be utilized to infer radar disturbances and aid in accurate UAV detection and localization include the fast maximum likelihood (FML) covariance estimator [6] and the multi-class inverse Wishart mixture (MC-IWM) filter [7].

The work by [8], uses a fast iterative adaptive algorithm employing a time-shared method to accurately estimate the AoA. The first phase of the technique involves a coarse azimuth angle estimation that is performed using just a single mmWave FMCW radar. The obtained intermediate results are then fed into a more refined iterative algorithm to accurately determine the direction of arrival (DOA) estimations within a small region. However, this method relies on multiple mmWave FMCW radars during the second phase making it computationally intensive. Another approach by [9] uses linear algebraic techniques to jointly calibrate the antennas and estimate the AoA. The iterative optimization technique provides excellent performance when there is an effective coupling between the antennae elements. Additionally, the algorithm is shown to work well in presence of several sensors and transmitter-receiver pairs. Daegun *et al.* in [10] proposed a joint angle and delay estimation algorithm which utilizes the dual-shift-invariant structure of the received signal to extend the one-dimensional pseudospectrum searching. Monte Carlo simulations were used to gauge the performance of the proposed method. Reported results show that the proposed technique exhibited superior performance with respect to state-of-the-art algorithms. However, the technique is computationally intensive and requires further experimentation with complex real-world scenarios. In [11], the authors combine the advantages of the 2D-FFT and MUSIC algorithms to develop a low complexity estimation method to approximate the various FMCW radar parameters such as AoA. However, this method fails to provide any performance insight when used to detect multiple targets. In [12] and [13], the authors use the notion of rotating the radar about its axis to scan an area. The signals obtained in each frame of the rotating radar are used to construct the range profile for the respective field of view. All the range profiles are then stitched and combined together to get the range-angle maps that offer a 180° view of the scene. However, the work by [12] focuses on target detection and classification and doesn't provide enough information for AoA estimation. Linga *et al.* in their work [13], has provided a novel AoA estimation technique with relatively good performance, however this can be further improved using machine learning techniques. Our work will be an extension to this work, where we will use state-of-the-art machine learning techniques

to estimate the AoA from the range-angle maps. We will be using convolutional neural networks (CNN) based algorithms which will serve to be useful for embedded and internet-of-things (IoT) related applications.

Hence, in this work, we propose a multi-target AoA estimation method that utilizes rotating mmWave FMCW radars. The proposed method employs a mechanical rotor setup for rotating the mmWave FMCW radar along the azimuthal axis followed by Yolov3 for post-processing the range-angle heatmaps. The AoA estimate is calculated from the centroid of the detected bounding box.

The remaining sections of this paper is structured as follows. Section II gives an overview of mmWave FMCW radars and signal processing aspects. Section III gives a brief overview of the machine algorithm that is used on the collected dataset. Section IV defines the various metrics used to measure the performance of the proposed method. Section V provides a high level system description along with the measurement and dataset details. Section VI summarizes the obtained results using the proposed method. And finally, Section VII concludes the paper by providing a brief summary and potential future research directions.

E.2 mmWave FMCW Radars and Signal Processing

The mmWave FMCW radars operate by transmitting high frequency chirp signals on to the scene. The transmitted chirp signal is reflected back upon encountering obstacles in its path. The reflected chirp is captured by the receiving antennae for further processing. The transmitted and received chirp signals are then fed into a mixer to obtain the intermediate frequency (IF) signal. The IF signal is further fed into an ADC which samples the analog signal and provides digital values. The raw IF samples then undergo additional processing to obtain the radial range, radial velocity and AoA estimation of the target [14].

Range estimation involves performing an FFT transform over the captured IF signal samples [15]. The range R is calculated as,

$$R = \frac{cf_{IF}}{2S}, \tag{E.1}$$

where f_{IF} is the frequency of the IF signal, c is the velocity of light in vacuum (3x10^8 m/s), and S is the slope of the radar.

Velocity estimation exploits the phase of the IF signal. Change in distance within a short time may not be accurately captured by the range estimation (due to limitations in range resolution) procedure. However, these changes can be captured by utilizing the phase difference between the received chirps. To obtain the velocity estimation, we initially compute the range profile. Next, we perform a second FFT across the received chirp signals to capture the small phase changes that can provide the velocity estimation values.

AoA estimation involves utilizing the number of receiver antennas on the mmWave FMCW radar. A differential distance exists from an object to each of the receiving antennas. This differential distance corresponds to a phase change that

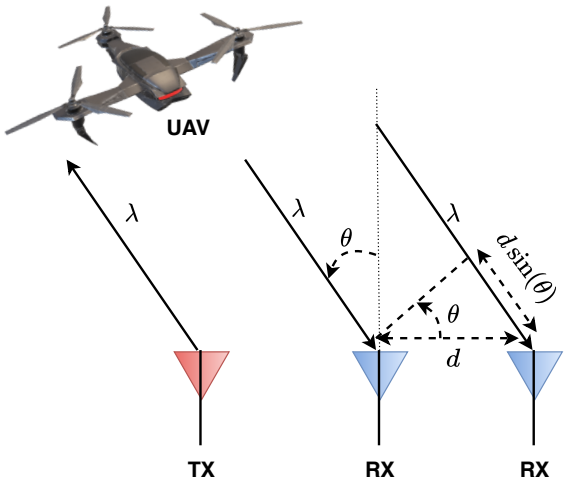


Figure E.2: Angle of arrival estimation.

can be obtained by performing a second FFT on the output of the range profile. The second FFT is applied across the different receiver antennas so as to obtain the angle-FFT. The angle-FFT can then be used to obtain AoA estimation of target objects in the scene. Note that in angle-FFT, the 2D-FFT is performed over the different receiver antennas separated in space whereas for velocity estimation the 2D-FFT utilizes the phase difference between the different chirps that are separated in time [16]. If ϕ is the phase difference between the received chirp signals and d the distance between adjacent receiving antennas, then the following holds true,

$$\phi = \frac{2\pi d \sin\theta}{\lambda}, \tag{E.2}$$

where λ is the wavelength of the chirp signal. Thus the AoA estimation θ is given as,

$$\theta = \sin^{-1}(\frac{\lambda\phi}{2\pi d}) \tag{E.3}$$

Fig E.2. gives a pictorial representation of the estimation of AoA. As seen from (E.2) and (E.3), when $\theta = 0°$, ϕ is more sensitive to changes in θ. As θ increases, the sensitivity of ϕ to θ decreases. This is due to the fact that ϕ and θ share a non-linear relationship as $\phi \propto \sin\theta$ and sensitivity of $\sin\theta$ decreases as θ increases. Thus AoA estimation is more accurate when the target is placed perpendicular to the radar [16].

Similar to AoA estimation, calculation of AoA resolution is also important to correctly distinguish between targets placed close to each other with a small AoA difference. The AoA resolution is defined as the least angle required between two targets so that they can be easily distinguished as separate objects in the angle-FFT [17]. The AoA resolution for mmWave FMCW radars is given as,

$$\theta_{res} = \frac{\lambda}{Nd\cos\theta}, \tag{E.4}$$

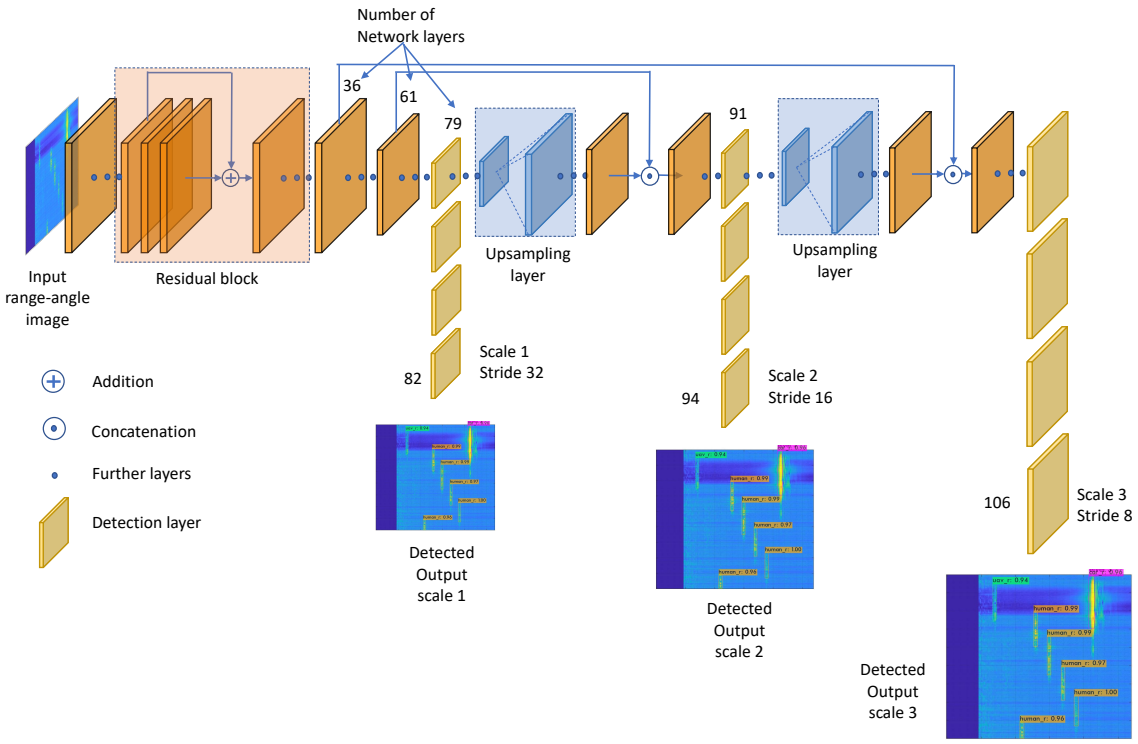


Figure E.3: Yolov3 Darknet Architecture

where N is the number of receiver antennas on the radar. If it is assumed that $d = \lambda/2$ and $\theta = 0$, then the expression for θ_{res} reduces to,

$$\theta_{res} = \frac{2}{N}, \tag{E.5}$$

implying that the AoA resolution improves with increased number of receiver antennas. For example, with 4 receiver antennas, the AoA resolution is 0.5 radians which is equivalent to 28°, a relatively high value. Since increasing the number of antennas can also increase space and hardware complexity, there is active research to improve the AoA resolution with minimum constraints.

E.3 You Look Only Once (Yolo) Architecture and Working

The Yolov3 is a state-of-the-art multi-scale object detection algorithm that can detect and classify objects from images and video with very fast inference time [18]. Just as the name suggests, the algorithm performs only a single-stage forward pass over the whole image to determine the class probabilities and predictions. In this work, we are using the Yolov3 version for our AoA estimation from range-angle maps.

The Yolov3 algorithm is based on the Darknet-53 architecture [19] that uses a combination of convolution layers and skip connections as shown in Fig. E.3. Inspired by ResNet and other architectures [20], Darknet-53 primarily consists of

feature extraction and feature detection stages. The feature extraction stage comprises of 53 convolution layers that are arranged as 3x3 and 1x1 consecutive layers followed by a skip connection. On the other hand, the feature detection stage is made up of 53 layers that helps to enhance the accuracy of the predictions as compared to previous versions. Thus the algorithm employs a sum total of 106 convolutional layers, both for feature extraction and feature detection.

Table E.1: Yolov3 configuration parameters

Parameter	**Value**
Batch Size	64
Number of classes	3
Max Batches	6000
Sub division	16
Learning rate	0.001
Momentum	0.9
Decay	0.0005
Filters	24
GPU version	Tesla V100-SXM3
CUDA version	11.4

The Yolov3 algorithm operates by initially passing the input image through the feature extractor to obtain multi-scale feature embeddings. Yolov3 supports three different feature scales namely, 13x13, 26x26, and 52x52. The obtained feature maps are then fed into the feature detection stage to predict bounding boxes on the detected object. To predict bounding boxes, the each feature map is divided into a collection of grid cells. Each grid cell is capable of predicting three bounding boxes. Each bounding box 'B' consists of '$5 + C$' attributes and class probabilities. Here, 'C' represents the number of classes in the dataset. In this work, $C = 3$, as we are using the model to detect three classes: UAV, car, and humans. The class probabilities provide the probability of existence of a particular class in the respective grid cell. The '5' represents bounding box attributes (t_x, t_y, t_h, t_w) and the objectness scores. Bounding box dimensions in Yolov3 are computed relative to the anchor box dimensions. Anchor boxes are prior boxes that have predefined aspect ratios. The predefined aspect ratios are determined by running a k-means algorithm on the entire dataset prior to training. Lastly, the objectness score denotes the presence of an object in the corresponding grid cell. Feature detection is performed by convolving the downsampled feature maps with a 1x1 detection kernel whose shape is given by 1x1x$(B * (5 + C))$. In Yolov3, feature detection occurs at the 82^{th}, 94^{th}, and 106^{th} layers.

Once the object is detected, a maximum of three bounding boxes per grid can be drawn based on the relative position of the object within the grid cell. To avoid multiple bounding boxes for the same object, a non-maximum suppression is performed. In non-maximum suppression, the bounding box that has the most

overlapping area with the detected object is retained while the other bounding boxes are discarded. A high level Yolov3 architecture diagram is given in Fig E.3. Table E.1 further lists the various parameters used for the Yolov3 algorithm to detect the various classes from the range-angle maps.

E.4 Performance Metrics

In this section, the various metrics used to quantify the performance of the machine learning model are discussed [21]. A machine learning model can predict the class of an object as either true (positive) or false (negative). When the model predicts the class as true and the actual class is positive, then the observation is defined as a true positive (TP). Similarly, when the model predicts the class correctly as negative, then the observation is a true negative (TN). However, when the predicted class is true and the actual class is negative, then the observation is a false positive (FP). Similarly, the observation is defined as a false negative (FN) when the predicted class is false and the actual class is positive.

E.4.1 Precision

Precision represents the fraction of total number of positively classified classes to the total number of positively predicted classes. Precision is given as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{E.6}$$

where TP and FP represents the true positives and false positives from the predicted result.

E.4.2 Recall

Recall is calculated as the ratio of true positive to the total number of positive classes. Recall is computed as:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{E.7}$$

where FN is the false negatives in the predicted results.

E.4.3 F1-score

F1-score provides a means to measure the performance of a machine model by utilizing both precision and recall metrics. F1-score is calculated as follows:

$$\text{F1-score} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{E.8}$$

E.4.4 Accuracy

The accuracy metric provides an indication of the performance of the model across all the classes. It is defined as follows:

$$\text{Accuracy} = \frac{TN + TP}{TN + FP + TP + FN} \tag{E.9}$$

E.4.5 Root mean square error (RMSE)

The RMSE value provides a measure of how much the estimated value has deviated from the actual true value. It is given as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{P}(x_{estimate} - x_{actual})^2}{P}} \tag{E.10}$$

where P is the sample size, $x_{estimate}$ is the estimated value, and x_{actual} is the actual or true value in the experiment. Lower RMSE value implies that the estimated measurement is closer to the true value and hence better model prediction.

E.5 System Overview and Dataset Details

E.5.1 System Description

The AoA estimation of multiple targets using mmWave FMCW radar is challenging. As described in Section E.2, accurate AoA estimation requires an increased number of transmitter and receiver antennas. This can lead to increased hardware complexity. Hence in our setup, the system at the GCS is made up of mmWave FMCW radar that utilizes only one transmitter-receiver antenna to estimate the AoA.

Our setup consists of a mmWave FMCW radar that is mounted firmly on a rotor. The rotor in itself is mounted on a static tripod. The rotor is battery-powered, programmable, and highly portable. The mmWave radar is mounted on the head of the rotor. The head of the rotor can rotate the radar through the entire 360° in the azimuth direction based on requirements.

The parameters of the rotor that can be programmed include the rotational direction, the rotational velocity, and the rotational angle. The rotational velocity of the rotor can be set according to the number of chirp frames transmitted by the radar per second. Here the chirp frame or simply frame is defined as a set of K equally spaced chirp waveforms that are transmitted by the radar [14]. The performance of the AoA estimation is directly influenced by the rotational velocity of the rotor. If the rotational velocity of the rotor is high, then the number of frames per second that are captured will be less. Similarly, if the rotational velocity is low, then the captured number of frames per second from the radar is high. Improved AoA estimation performance requires higher number of frames per second, as higher number of frames can capture more information from multiple targets present in the scene. Hence in our setup, we have programmed the rotational velocity of the rotor to be low so as to capture at least one frame per degree of rotation of the radar.

Table E.2: mmWave FMCW radar AWR1843 and AWR2243 Parameters

Radar Parameter	**Value**
No. of frames (AWR2243)	200
No. of frames (AWR1843)	800
Frame periodicity	40 ms
No. of Tx antennas	3
No. of Rx antennas	4
No. of ADC samples	256
Sampling Rate	10 MSPS
Frequency Range (RF)	$77 - 81$ GHz
Bandwidth (RF)	1798.92 MHz
No. of Chirps	128
Chirp Slope	29.982 MHz/μs
Rx Noise Figure	14 dB (76 to 77 GHz) 15 dB (77 to 81 GHz)
Transmission Power	12 dBm

Table E.3: Experiment Parameters

Experiment Parameter	**Value**
UAV Size	$32.24 \times 24.2 \times 8.4$ cm^3
Human Height	172 cm
Car Size	$431.5 \times 178.0 \times 160.5$ cm^3
Measurement Range	upto 26 meters

Based on the adjusted rotational velocity, let the desired FoV to be covered by the rotor in t seconds be θ_{FoV}. In a t second duration, the radar transmits n_f frames. By considering that the entire θ_{FoV} is divided into smaller angle bins (θ_b), we obtain the following relation,

$$\theta_b = \frac{\theta_{FoV}}{n_f} \tag{E.11}$$

In our experiment scenario, the desired FoV is 180°. The n_f is 800 and 200 frames for the radars AWR2243 [22] and AWR1843 [23] respectively. Each frame is made up of 128 chirp waveforms transmitted by the radar. This corresponds to 0.225° per frame for AWR2243 and 0.9° per frame for AWR1843. To increase the FoV in the elevation, the radar is placed vertical to the ground plane in our setup.

The two mmWave FMCW radars, AWR2243 [22] and AWR1843 [23] that we use in our experiment are manufactured from Texas Instruments. Both radars have identical frequency range spanning between $77 - 81$ GHz. Additionally, they have identical number of transmitter and receiver pairs, RF bandwidth, chirp slope, sampling rate, and ADC samples. However, the number of frames employed is 800 for the AWR1843 and 200 for the AWR2243. A detailed list of the key parameters of the radars can be found in Table. E.2.

Table E.4: Measurement cases for Set1_UavCarHumans (Range in meters and angle in degrees)

Range (m) → / Cases ↓	5	7	9	11	13	15	17	19	21	23	25
AA	U(0°)		H1(30°)	H2(60°)	H3(90°)	H4(120°)	H5(150°)	C(0°)			
BB		U(30°)		H5(180°)	H1(60°)	H2(90°)	H3(120°)	H4(150°)	C(30°)		
CC			U(60°)	H1(0°)	H2(30°)		H3(90°)	H4(120°)	H5(150°)	C(60°)	
DD				U(90°)	H1(0°)	H2(30°)	H3(60°)		H4(120°)	H5(150°)	C(90°)
EE	U(120°)	H1(60°)	H2(90°)		H3(150°)	H4(0°)	C(180°)	H5(30°)			
FF		U(150°)	H5(0°)			C(150°)	H1(30°)	H2(60°)	H3(90°)	H4(120°)	
GG	H1(60°)	H2(90°)	U(180°)	H3(150°)	C(120°)				H4(0°)	H5(30°)	
HH	U(30°)	H4(0°)	H3(150°)	H2(120°)				H1(90°)		C(180°)	H5(60°)
II	H1(90°)	H2(120°)		U(30°)	H3(180°)	H4(60°)	H5(0°)				C(150°)
JJ	U(150°)	H3(180°)	H1(120°)						H2(60°)	H4(90°)	C(30°)
KK	H3(150°)	U(60°)			H4(90°)		H5(30°)		H1(180°)	H2(0°)	C(120°)
LL	H2(180°)	H1(120°)	U(30°)	H3(60°)		H4(90°)		H5(150°)			C(0°)
MM		H2(0°)	U(90°)	H3(30°)	H4(120°)	H1(180°)	H5(150°)				C(180°)
NN	U(90°)	H1(30°)	H3(60°)	H4(120°)	H5(150°)		H2(0°)	C(180°)			

U - UAV, C - Car, H1 - Human 1, H2 - Human 2, H3 - Human 3, H4 - Human 4, H5 - Human 5.

Table E.5: Measurement cases for Set2_UavCar (Range in meters and angle in degrees)

Range (m) → / Cases ↓	5	7	9	11	13	15	17	19	21	23	25
AA		U(0°)									C(60°)
BB			U(0°)					C(60°)			
CC				U(0°)			C(90°)				
DD	U(60°)					C(120°)					
EE			U(60°)				C(150°)				
FF				U(60°)			C(180°)				
GG		U(90°)							C(180°)		
HH		U(120°)								C(150°)	
II			U(120°)			C(0°)					
JJ				U(120°)					C(0°)		
KK			U(150°)				C(0°)				
LL				U(150°)						C(90°)	
MM	U(180°)					C(60°)					
NN		U(180°)							C(120°)		
OO				U(180°)							C(30°)

U - UAV, C - Car.

The measurement scenario is an outdoor parking setting with UAV, car, and humans placed at varying locations with respect to the radar. Each measurement is taken by positioning UAV, car, and humans at different locations within a radial distance of $\approx$ 26 meters and 180° FoV in front of the radar. We limit the radial distance to $\approx$ 26 meters to remain within the permissible measurement area. However, the mmWave FMCW radars can measure targets with range upto 300 meters [22],[23]. Depending upon the radar version used (AWR1843/AWR2243), a total of 800/200 frames containing the raw IF signal is captured for each measurement. Each frame of the radar has a duration of about 40 ms and consists of 128 chirps. Further, each frame of the radar provides raw IF signal data corresponding to a

Table E.6: Measurement cases for Set3_OnlyHumans (Range in meters and angle in degrees)

Range (m) → / Cases ↓	5	7	9	11	13	15	17	19	21	23	25
AA	H1(0°)	H2(30°)	H3(60°)	H4(90°)	H5(120°)						
BB		H3(0°)	H4(30°)	H5(60°)		H1(150°)	H2(180°)				
CC			H5(0°)		H1(90°)	H2(120°)	H3(150°)	H4(180°)			
DD				H1(30°)	H2(60°)	H3(90°)	H4(120°)	H5(150°)			
EE				H2(0°)	H3(30°)	H4(60°)	H5(90°)		H1(180°)		
FF					H4(0°)	H5(30°)		H1(120°)	H2(150°)	H3(180°)	
GG							H1(60°)	H2(90°)	H3(120°)	H4(150°)	H5(180°)
HH						H1(0°)	H2(30°)	H3(60°)	H4(90°)	H5(120°)	
II							H2(0°)	H3(30°)	H4(60°)	H5(90°)	H1(150°)
JJ								H2(0°)	H3(30°)	H4(60°)	H5(90°), H1(120°)
KK	H4(150°)	H5(180°)							H1(0°)	H2(30°)	H3(60°)
LL	H3(120°)	H4(150°)	H5(180°)							H1(0°)	H2(30°)
MM	H2(90°)	H3(120°)	H4(150°)	H5(180°)							H1(0°)
NN	H1(60°)	H2(90°)	H3(120°)	H4(150°)	H5(180°)						
OO	H1(30°)	H2(60°)	H3(90°)	H4(120°)	H5(150°)						

H1 - Human 1, H2 - Human 2, H3 - Human 3, H4 - Human 4, H5 - Human 5.

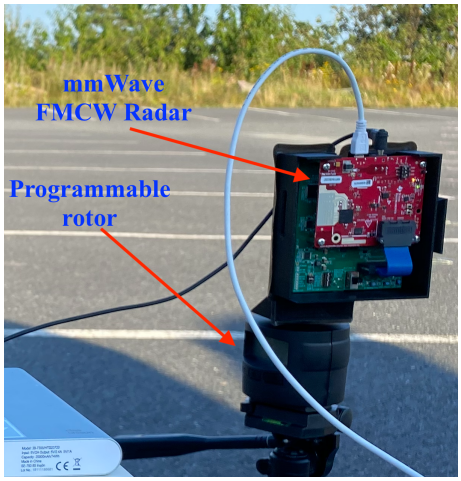


Figure E.4: The rotating mmWave FMCW radar measurement setup.

0.225°/0.9° FoV depending upon the radar model. The collected raw IF signal data is post-processed in MATLAB [24] to obtain the range-profile for each frame. Finally, the range profiles for all the frames are stitched together so as to obtain a 180° FoV range-angle map of the whole measurement scenario that comprises of multiple targets [12]. Fig. E.4 shows the experimental setup of the rotating mmWave FMCW radar that is used to capture the raw IF signals. Other relevant parameters related to the experiment can be found in Table. E.3.

The range-angle images obtained from MATLAB is fed into the Yolov3 algorithm for training. We implement a 10-fold scheme for validation. In this scheme, the training is performed for 10 folds of the experiment. In each fold, only 80% of the dataset images are used for training and the remaining are assigned as test images. We utilize the pretrained weights available from the Darknet-53 repository to initiate the training. Upon training, the optimum weights are obtained indepen-

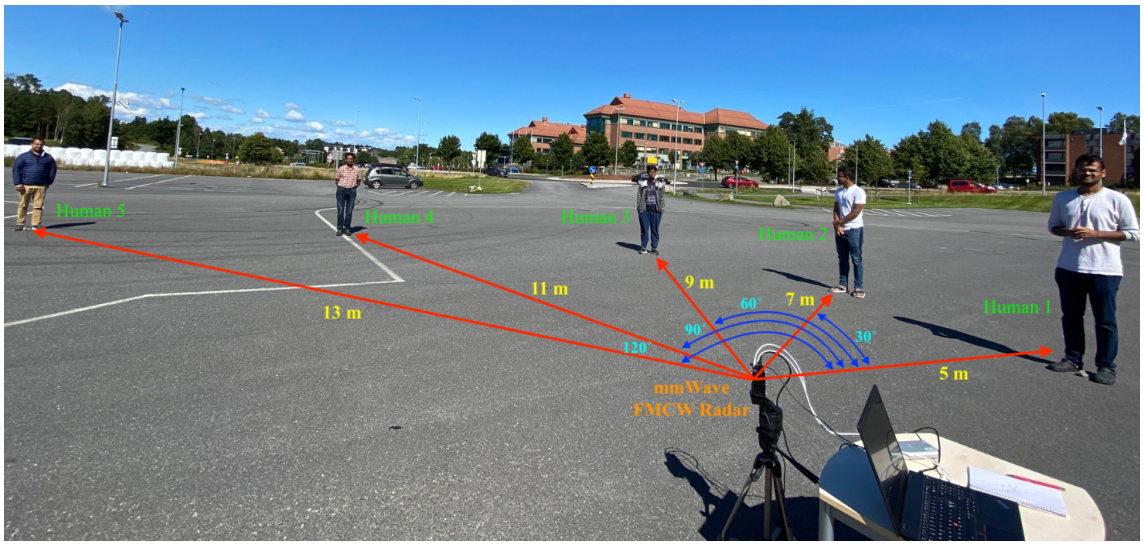


Figure E.5: Measurement setup for case AA from Set3_OnlyHumans.

dently for each fold. Thus the weights of the 1^{st} fold are independent of the 2^{nd} fold, and so on. By using this approach, we are implicitly removing any bias that is associated with the training. The weights obtained after training are used on the test images to obtain the predictions. The predictions from the Yolov3 algorithm are range-angle images that contain bounding boxes on the detected object. It is to be noted that each pixel dimension in the range-angle image corresponds to a specific range and angle in the measurement setup. Hence, the centroid coordinates of the detected bounding box correspond to the target object's estimated range and AoA in the measurement setup. Once the AoA estimation is obtained, the algorithm performance can be calculated by computing the root mean square error between the AoA estimation and the ground truth angle of the target.

E.5.2 Dataset Details

Based on the above measurement setup, we have collected the raw IF signal data from two mmWave FMCW radars, the AWR2243 and the AWR1843 for different target objects. The target objects include combinations of UAV, car, and humans positioned at various locations in front of the radar. We used the AWR2243 mmWave FMCW radar to capture the raw IF signals of UAV, car, and humans. Similarly, we used the AWR1843 mmWave FMCW radar to collect the raw IF signals of humans. The collected raw IF signals are processed in MATLAB to obtain the range profile for each frame. The range profiles for each frame are stitched together to obtain the radar range-angle images. We have created three datasets based on these range-angle images. The first dataset contains UAV, car, and humans while the second dataset comprises of only UAV and car combination. The third dataset contains range-angle images of only human targets. The first, second, and third datasets are named as Set1_UavCarHumans, Set2_UavCar, and Set3_OnlyHumans respectively. The range-angle images obtained after processing from MATLAB have a resolution of 875×656 pixels for Set1_UavCarHumans and Set2_UavCar. Similarly, the range-

Table E.7: RMSE and accuracy values for 10 folds

Sl No.	UAV (RMSE)	Car (RMSE)	Human (RMSE)	Accuracy (%)
1	1.4746°	1.3234°	0.9644°	99.3902
2	0.7869°	1.4576°	0.8635°	98.4939
3	1.3661°	1.0667°	1.0702°	97.9166
4	0.8700°	1.3807°	1.2312°	98.6111
5	0.9305°	1.2518°	1.1380°	98.7012
6	0.7274°	1.2540°	0.8898°	98.4423
7	0.8963°	1.6049°	1.1868°	96.8838
8	1.4128°	1.2237°	1.1005°	98.6413
9	1.3441°	1.2035°	0.9511°	95.5882
10	0.9922°	1.1559°	0.8297°	99.3569

angle images in Set3_OnlyHumans have a resolution of 1167×875 pixels.

The dataset measurement cases are detailed in Table. E.4, E.5, and E.6 respectively. The different cases are labelled as AA, BB, CC ....OO depending upon the dataset. The entries inside the table are represented as an object that is positioned at a specified distance and angle with respect to the radar. The abbreviation for the labels used in the table entries are given towards the bottom of the respective tables. Hence, the first entry in case AA of Table. E.6 depicts a person at 5 m and 0° with respect to the radar. The rest of the entries can be interpreted as follows: a second person at 7 m and 30°, the third person at 9 m and 60°, the fourth person at 11 m and 90°, and the fifth person positioned at 13 m and 120° with respect to the radar. Fig. E.5 shows one particular entry of the measurement setup for Set3_OnlyHumans.

The range-angle images depict the signatures of the different target objects, UAV, car, and humans. To train using the Yolov3 algorithm, we draw bounding boxes on the range-angle images using the LabelImg software [25]. The bounding boxes are drawn such that the centroid of the boxes represents the AoA of the target object.

E.6 Results

The labeled range-angle images are used to train the Yolov3 algorithm for optimum weight parameters. The training is performed for 10 folds of the experiment. In each fold of the experiment, the training is executed for 6000 iterations so that the average loss is minimized. The average loss vs iterations for the 7^{th} fold of the experiment can be observed in Fig E.6. Further, the accuracy for each fold of the experiment is listed in Table E.7. As observed and calculated from Table E.7, we obtain a relatively high average classification accuracy of 98.20%, demonstrating the reliability of our approach. The predicted classification can be observed on the range-angle image as

Table E.8: Comparing the advantages of this work with other techniques

Method	**Number of antennas**	**Targets with same range or angle**	**Target classification**
2D-ESPRIT [26]	1 Tx, 2 Rx	✗	✗
DFT-ESPRIT [27]	1 Tx, 2 Rx	✗	✗
Dual-Smoothing [28]	1 Tx, 2 Rx	✗	✗
Clustered ESPRIT [29]	1 Tx, 2 Rx or more. Rx antennas could be less than number of targets	✓	✗
Rotating Radar [13]	1 Tx, 1 Rx	✓	✗
Ours	**1 Tx, 1 Rx**	✓	✓

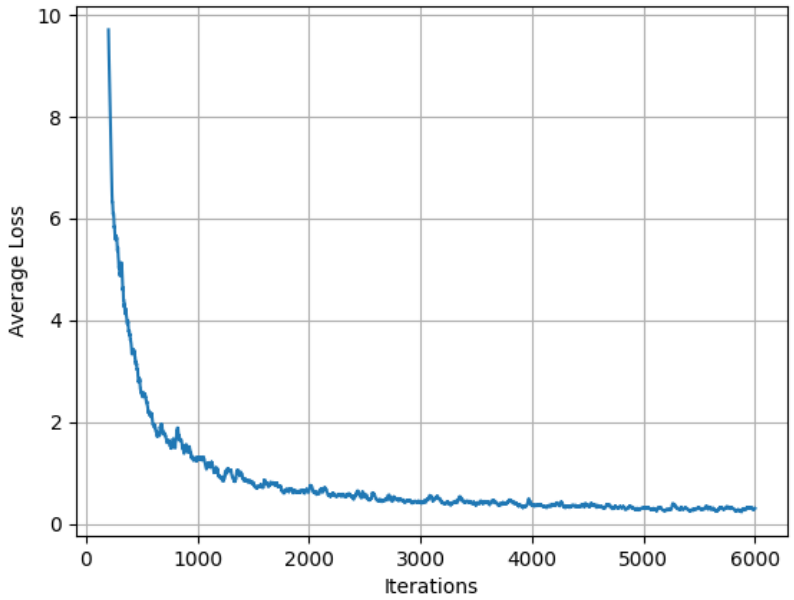
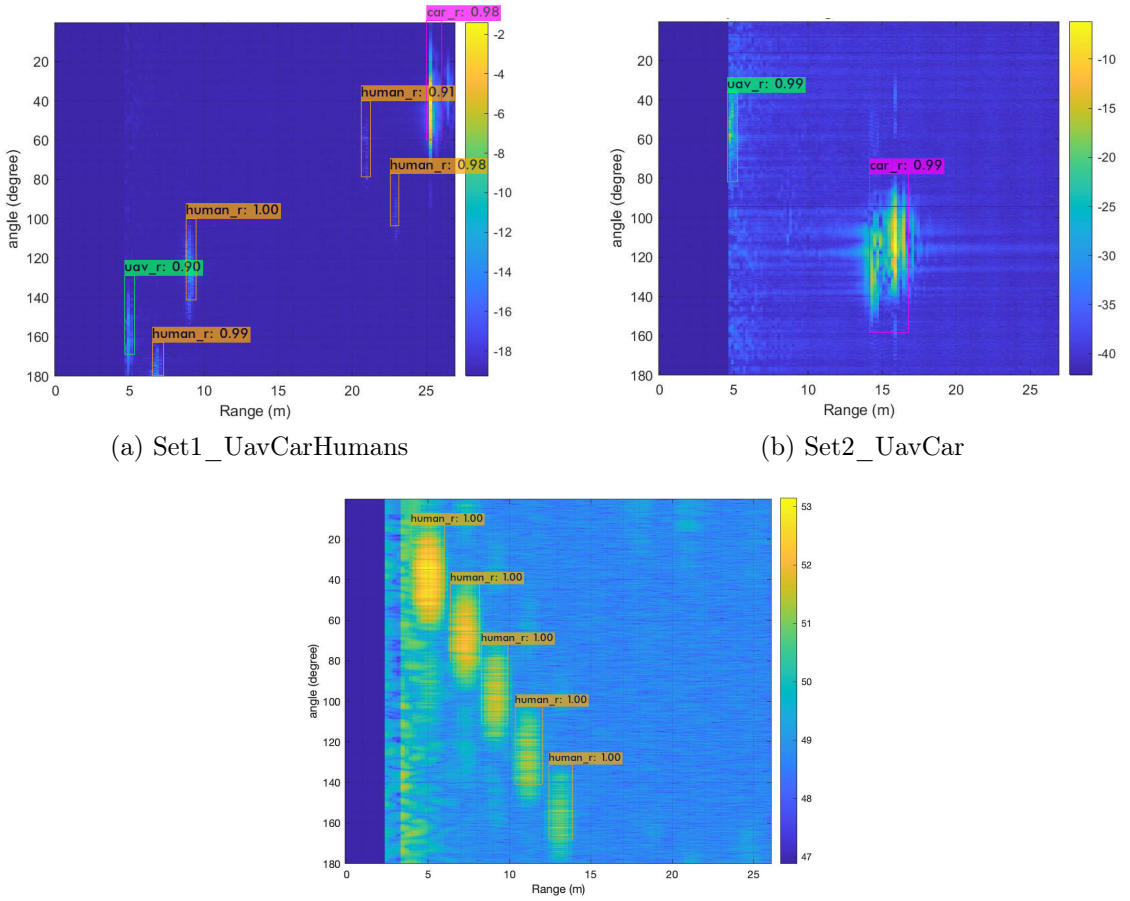


Figure E.6: Average loss vs iteration curve for the 7^{th} fold

shown in Fig. E.7. The prediction statistics can also be visualized with the help of a confusion matrix as shown in Fig E.8. The confusion matrix shows that UAV, car, and humans are classified 97.89%, 99.85%, and 99.28% respectively. The average precision, recall, and F1-score for 10 folds of the experiment are 0.991, 0.992, and 0.991 respectively.

The performance of the algorithm to estimate the AoA is calculated using the RMSE value. Table E.7 lists the different RMSE values obtained for each fold of the algorithm for different test scenarios. It can be seen that the average AoA RMSE value for UAV, car, and humans are 1.0800°, 1.2922°, and 1.0225° respectively.



(a) Set1_UavCarHumans

(b) Set2_UavCar

(c) Set3_OnlyHumans

Figure E.7: Images depicting Yolov3 class prediction for range-angle images from all three datasets.

Class	UAV	Car	Human	None	Total
UAV	650	0	0	14	664
Car	0	669	0	1	670
Human	0	0	1955	14	1969

Figure E.8: Confusion matrix

Interestingly, the average RMSE value of cars is greater than that of UAV and humans. This is expected as it aligns with the fact that the car has more AoA spread due to its large size. The proposed method implies that it is difficult to accommodate the car to a single AoA due to its large size as compared to UAV and human.

Table E.8 shows the advantages of the proposed method over other similar techniques such as 2D-ESPRIT [26], DFT-ESPRIT [27], dual-smoothing [28], clustered ESPRIT [29], and rotating radar [13] in terms of number of antennas, and target classification. It is observed that the proposed method is similar to the rotating radar technique as both utilize just a single transmitter-receiver antenna for detecting mul-

tiple non-interacting targets. However, our method provides an added advantage of target classification along with AoA estimation. Additionally, the proposed method also has a lower RMSE error as compared to the rotating radar technique.

E.7 Conclusion

The AoA estimation using mmWave FMCW radars is not accurate due to limited number of antennas. As a result, target localization is imperfect in critical applications involving UAVs and GCS. This work demonstrated that the FoV and AoA estimation of mmWave FMCW radars can be improved by combining a mechanical rotor setup with cutting-edge computer vision techniques such as Yolov3. The radar setup's range-angle images were fed into the Yolov3 algorithm, which detected, classified, and localized multiple targets in the scene. The proposed method achieved very high accuracy for target classification with low AoA estimation error outperforming traditional techniques. Evaluating the proposed approach's real-time performance on UAVs and GCSs to reduce latency in target detection and localization could be a potential future research direction.

References

[1] A. N. Wilson, Abhinav Kumar, Ajit Jha, and Linga Reddy Cenkeramaddi. Embedded sensors, communication technologies, computing platforms and machine learning for UAVs: A review. *IEEE Sensors Journal*, 22(3):1807–1826, Feb. 2022.

[2] Tianyuan Yang, Antonio De Maio, Jibin Zheng, Tao Su, Vincenzo Carotenuto, and Augusto Aubry. An adaptive radar signal processor for uavs detection with super-resolution capabilities. *IEEE Sensors Journal*, 21(18):20778–20787, 2021.

[3] A. Aubry, Vincenzo Carotenuto, Antonio Maio, and G. Foglia. Exploiting multiple a priori spectral models for adaptive radar detection. *Radar, Sonar & Navigation, IET*, 8:695–707, 08 2014.

[4] Augusto Aubry, Vincenzo Carotenuto, Antonio De Maio, and Mark A. Govoni. Multi-snapshot spectrum sensing for cognitive radar via block-sparsity exploitation. *IEEE Transactions on Signal Processing*, 67(6):1396–1406, 2019.

[5] A. Aubry, Antonio Maio, Luca Pallotta, and Alfonso Farina. Median matrices and their application to radar training data selection. *Radar, Sonar & Navigation, IET*, 8:265–274, 04 2014.

[6] A. Aubry, A. De Maio, and V. Carotenuto. Optimality claims for the fml covariance estimator with respect to two matrix norms. *IEEE Transactions on Aerospace and Electronic Systems*, 49(3):2055–2057, 2013.

[7] Paolo Braca, Augusto Aubry, Leonardo Maria Millefiori, Antonio De Maio, and Stefano Marano. Multi-class random matrix filtering for adaptive learning. *IEEE Transactions on Signal Processing*, 68:359–373, 2020.

[8] Wenqiang Wei et al. DOA estimation of distributed mmWave radar system via fast iterative adaptive approach. In *Proc. International Conference on Control, Automation and Information Sciences (ICCAIS)*, pages 414–418, Xi'an, China, 2021.

[9] Muhammad Z. Ikram, Murtaza Ali, and Dan Wang. Joint antenna-array calibration and direction of arrival estimation for automotive radars. In *Proc. IEEE Radar Conference (RadarConf)*, pages 1–5, Philadelphia, PA, USA, 2016.

[10] Daegun Oh and Jong-Hun Lee. Low-complexity range-azimuth FMCW radar sensor using joint angle and delay estimation without SVD and EVD. *IEEE Sensors Journal*, 15(9):4799–4811, Sept. 2015.

[11] Sangdong Kim and Kyun-Kyung Lee. Low-complexity joint extrapolation-MUSIC-based 2D parameter estimator for vital FMCW radar. *IEEE Sensors Journal*, 19(6):2205–2216, March 2019.

[12] Siddharth Gupta et al. Target classification by mmWave FMCW radars using machine learning on range-angle images. *IEEE Sensors Journal*, 21(18):19993–20001, Sept. 2021.

[13] Linga Reddy Cenkeramaddi et al. A novel angle estimation for mmWave FMCW radars using machine learning. *IEEE Sensors Journal*, 21(8):9833–9843, Apr. 2021.

[14] The fundamentals of millimeter wave radar sensors. [Online]. Available: https://www.ti.com/lit/wp/spyy005a/spyy005a.pdf.

[15] FMCW radars - Range estimation. [Online]. Available: https://training.ti.com/intro-mmwave-sensing-fmcw-radars-module-1-range-estimation.

[16] FMCW radars - Angle estimation. [Online]. Available: https://training.ti.com/intro-mmwave-sensing-fmcw-radars-module-5-angle-estimation.

[17] Introduction to mmwave sensing - FMCW Radars. [Online]. Available: https://training.ti.com/sites/default/files/docs/mmwaveSensing-FMCW-offlineviewing_0.pdf.

[18] Joseph Redmon and Ali Farhadi. YOLOv3: An incremental improvement. *arXiv*, 2018.

[19] Joseph Redmon. Darknet: Open source neural networks in C. [Online]. Available: http://pjreddie.com/darknet/.

[20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *arXiv*, 2015.

[21] Khawaja Moyeezullah Ghori et al. Performance analysis of different types of machine learning classifiers for non-technical loss detection. *IEEE Access*, 8:16033–16048, Dec. 2019.

[22] Texas Instruments AWR2243. [Online]. Available: https://www.ti.com/document-viewer/AWR2243/datasheet.

[23] Texas Instruments AWR1843. [Online]. Available: https://www.ti.com/document-viewer/AWR1843/datasheet/device-overview-x3342#x3342.

[24] MATLAB. [Online]. Available: https://www.mathworks.com/products/matlab.html.

[25] LabelImg software. [Online]. Available: https://github.com/tzutalin/labelImg.

[26] DG Oh, YH Ju, and JH Lee. Subspace-based auto-paired range and DOA estimation of dual-channel FMCW radar without joint diagonalisation. *Electronics letters*, 50(18):1320–1322, Aug. 2014.

[27] Sangdong Kim, Daegun Oh, and Jonghun Lee. Joint DFT-ESPRIT estimation for TOA and DOA in vehicle FMCW radars. *IEEE Antennas and Wireless Propagation Letters*, 14:1710–1713, Apr. 2015.

[28] Daegun Oh, Yeonghwan Ju, Hyunsoo Nam, and Jong-Hun Lee. Dual smoothing DOA estimation of two-channel FMCW radar. *IEEE Transactions on Aerospace and Electronic Systems*, 52(2):904–917, Apr. 2016.

[29] Wen-Hsien Fang and Li-Der Fang. Joint angle and range estimation with signal clustering in FMCW radar. *IEEE Sensors Journal*, 20(4):1882–1892, Oct. 2019.