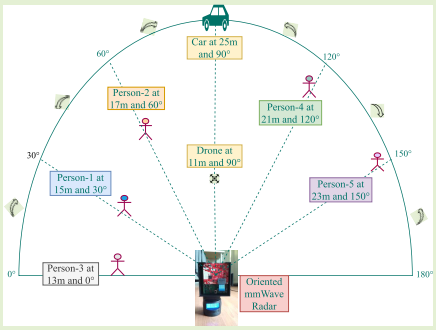
Target Classification by mmWave FMCW Radars Using Machine Learning on Range-Angle Images

Siddharth Gupta, Prabhat Kumar Rai, Abhinav Kumar, *Senior Member, IEEE*, Phaneendra K. Yalavarthy, *Senior Member, IEEE*, and Linga Reddy Cenkeramaddi

***Abstract*—In this paper, we present a novel multiclass-target classification method for mmWave frequency modulated continuous wave (FMCW) radar operating in the frequency range of 77 - 81 GHz, based on custom range-angle heatmaps and machine learning tools. The elevation field of view (FoV) is increased by orienting the Radar antennas in elevation. In this orientation, the radar focuses the beam in elevation to improve the elevation FoV. The azimuth FoV is improved by mechanically rotating the Radar horizontally, which has antenna elements oriented in the elevation direction. The data from the Radar measurements obtained by mechanical rotation of the Radar in Azimuth are used to generate a range-angle heatmap. The measurements are taken in a variety of real-world scenarios with various objects such as humans, a car, and an unmanned aerial vehicle (UAV), also known as a drone. The proposed technique achieves accuracy of 97.6 % and 99.6 % for classifying the UAV and humans, respectively, and accuracy of 98.1 % for classifying the car from the range-angle FoV heatmap. Such a Radar classification technique will be extremely useful for a wide range of applications in cost-effective and dependable autonomous systems, including ground station traffic monitoring and surveillance, as well as control systems for both on-ground and aerial vehicles.**



***Index Terms*—Autonomous systems, azimuth angle, elevation angle, enhanced field of view, FMCW radar, heatmap, machine learning, mmWave Radar, YOLO v3.**

I. Introduction

THE mmWave Radars are frequency modulated continuous wave (FMCW) Radars that are low power, highly compact, and very small in size, with a very high resolution. These mmWave Radars detect the range, velocity, and angle of arrival (AoA) of objects in their direct field of view (FoV). The range and velocity of an object, in particular, can be accurately sensed in millimeters and centimeters per second. These radars operate in the GHz range, with a radio frequency (RF) bandwidth of a couple of GHz available. They can detect distances ranging from a few centimeters to 300 - 400 meters. [1], [2].

FMCW radars have been used for 2-dimensional (2D) synthetic-aperture Radar (SAR) imaging [3] and image reconstruction using 2D FFT or range Doppler plot which can capture the object in a constant FoV due to fixed horizontal and vertical movement. This bounds the user's perspective and capability to capture a wide range of scenarios. In this article, we have used 1D Range FFT plot for classification. With the fixed positioning of radars, getting features like distance [4], [5] is achievable but in a rotating radar capturing 180 degrees FoV, it is difficult to estimate the objects features such as height and width. With our classification and detection model, we are able to recognize and mark the objects in the 180 degrees FoV heatmap as a bounding box. This FoV is adjustable and can be adjusted according to the application requirements.

These mmWave Radars have limited number of transmitting and receiving antennas due to the complexity and the cost. Due to the limit of transceivers antennas, FoV can only be boosted

Manuscript received May 4, 2021; revised June 10, 2021; accepted June 11, 2021. Date of publication June 25, 2021; date of current version September 15, 2021. This work was supported in part by the Indo-Norwegian Collaboration in Autonomous Cyber-Physical Systems (INCAPS) Project of the International Partnerships for Excellent Education, Research and Innovation (INTPART) Program under Grant 287918, in part by the Low-Altitude UAV Communication and Tracking (LUCAT) Project of the Program on ICT and Digital Innovation (IKTPLUSS) Program from the Research Council of Norway under Grant 280835, and in part by the Department of Science and Technology (DST), Government of India under Grant INT/NOR/RCN/ICT/P-01/2018. The associate editor coordinating the review of this article and approving it for publication was Prof. Piotr J. Samczynski. *(Corresponding author: Linga Reddy Cenkeramaddi.)*

Siddharth Gupta and Prabhat Kumar Rai are with the Department of Information and Communication Technology, University of Agder, 4630 Kristiansand, Norway, and also with the Indian Institute of Technology, Hyderabad, Telangana 502285, India (e-mail: ee18mtech01003@iith.ac.in; ee18mtech01005@iith.ac.in).

Abhinav Kumar is with the Indian Institute of Technology, Hyderabad, Telangana 502285, India (e-mail: abhinavkumar@ee.iith.ac.in).

Phaneendra K. Yalavarthy is with the Department of Computational and Data Sciences (CDS), Indian Institute of Science (IISc), Bengaluru 560012, India (e-mail: yalavarthy@iisc.ac.in).

Linga Reddy Cenkeramaddi is with the ACPS Group, Department of Information and Communication Technology, University of Agder, 4630 Kristiansand, Norway (e-mail: linga.cenkeramaddi@uia.no).

Digital Object Identifier 10.1109/JSEN.2021.3092583

in either elevation or azimuth direction. It is mostly designed to increase FoV in azimuth direction as lot of applications depend upon wide FoV in azimuth. mmWave Modules which are used for traffic management systems and installed as ground station requires broad FoV in both azimuth and elevation direction. There is a way to increase the FoV in elevation and azimuth of radar modules by embedding 2-dimensional (2-D) antenna array. But increase in transceiver antennas, increases the complexity, computational latency and cost. A 3-dimensional (3D) view mmWave Radar is implemented in [6] using fan-beam antenna for mobile robotic application. There are limitations in range and velocity measurements in addition to lack of details of angle of arrival. The 3D near field imaging for robotics and security scans has been proposed in [7], they have integrated the LiDAR data, which increases complexity, size, and latency. A search and track mmWave ground station synthetic aperture Radar has also been proposed in [8]. It has several disadvantages of being bulky, complex and lacks the angle estimation of target for rotating Radars. The localization and mapping work is done by rotating FMCW Radar in [9]. The work is applied to sense range and velocity of target. All the above mentioned works have used some kind of mechanical rotation for mmWave Radar, they have not provided the localization of target which is very important if the application of rotating Radar is surveillance. Apart from localization, the reported works do not provide any classification method if different targets are present in the scene. To overcome the above mentioned challenges, we propose a rotating mmWave FMCW Radar which detect the range, angle in FoV (range and angle, combined can be said as localization in 3D map) (Note: Here angle refers to FoV of test area) along with classification of multi-class objects such as humans, car, and an aerial vehicle such as drone, using (YOLO V3) [10], the real time object detection technique based on deep convolutional neural network architecture using skip connections or shortcut layers and routes as shown in Fig. 9. We have used different cases heatmaps for training generated by post processing of FMCW radar raw data.

Object recognition in itself is a complicated task and with decreasing lighting conditions the detection becomes more and more complex. In pitch dark scenarios detecting objects by conventional methods of capturing images using regular cameras will provide very low accuracy [11]. In low light environments the challenging task is to find the recognizable features on which the models can be trained, but due to darkness the difference between the foreground and background is hard to find. Some research works have shown promising results by using night vision camera [12] and Lidar based object detection [13]. The proposed technique in this article is novel and accurate irrespective of the light conditions. The system is robust to work under low lighting or no lighting conditions and still provide higher accuracy.

The remainder of this paper is organized as follows. Section II provides the system description. The details of the experimental setup and formulation of the principle used in order to extract information is described in Section III. Further, in Section III-B, the dataset details, machine learning model and performance evaluation are presented.

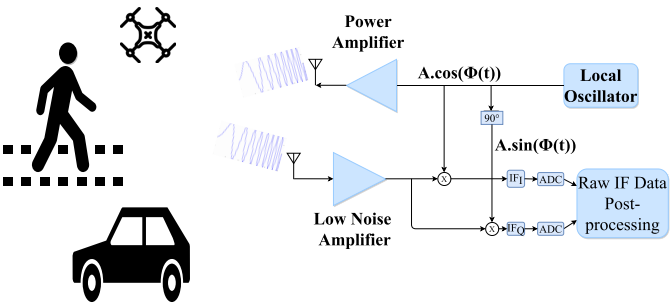


Fig. 1. Radar architecture with complex IF signal.

Finally, concluding remarks and possible future works are discussed in Section VI.

II. System Description

The system consists of Texas Instruments (TI) mmWave Radar, which have 3 transmitters and 4 receivers [14]. It also has a complex baseband mixer, analog to digital converters (ADCs), and an integrated phase locked loop (PLL). The architecture can also be seen in Fig. 1. The received signal is mixed with quadrature phase and in-phase of the transmitted FMCW signal. Hence a complex intermediate frequency (IF) signal is created and then it is passed through analog-to-digital-converters (ADCs) to digitize it. The FMCW Radar is operating in a frequency range of 77-81 GHz and bandwidth of 4 GHz. The other Radar parameters can also be seen in a Table I. We have placed mmWave FMCW Radar as shown in Fig. 2 i.e. we have oriented the Radar by 90°. In this setup, it gives wide FoV in elevation and narrow FoV in azimuth. We have installed this entire setup on a programmable rotor which rotates the setup in 180° for total of 800 frames. 800 frames have been used to collect the raw intermediate frequency (IF) data at every measurement case. The single frame time is chosen of 40 ms. The Rotor has a rotating head which enables it to rotate and spin Radar through 360 degrees. It is battery powered and highly portable. The Rotor has a rotating head which enables it to rotate and spin Radar through 360 degrees. It is battery powered and highly portable. The Rotor's rotation angle, velocity, and direction are all programmable. The rotational velocity is always chosen so that at least one frame per degree is covered. The rotational velocity can be adjusted depending on the number of frames per second in Radar. If the number of frames per second for the Radar is increased, the rotational velocity can also be increased until at least one frame is covered per degree is reached. In mmWave FMCW Radar, the transmitted signal is a signal whose frequency is changing linearly with time. This sweep in frequency is also known as chirp. A set of these chirps forms a Frame. We are using a frame, which consists of 128 chirps. The collected raw data is then post-processed in MATLAB [15]. The measurements are performed by positioning the aerial vehicle/drone, humans and a car at certain distance and angle (FoV) covering several radial ranges from Radar $\approx$ 26 meters for 180° FoV. Each measurement is termed as different case, in which the positions of humans, car and aerial vehicle are puzzled. Such different cases can also be observed in Fig. 3.

TABLE I
RADAR CONFIGURATION PARAMETERS

S.No.	Parameter	Configuration
1	Start Frequency	77 GHz
2	Frequency Slope	29.982 MHz/μs
3	Number of Transmitters	3
4	Number of Receivers	4
5	ADC samples	256
6	Number of frames	800
7	Chirp loops	128
8	Bandwidth	1798.92 MHz
9	Frame periodicity	40 ms
10	Sampling Rate	10 MSPS
11	Drone Size	322×242×84 mm
12	Human Height	172 cm
13	Car Size	4315x1780x1605 mm
14	Measurement Range	upto 26 meters
15	Transmission Power	12 dBm
16	Rx Noise Figure	14 dB (76 to 77 GHz) 15 dB (77 to 81 GHz)

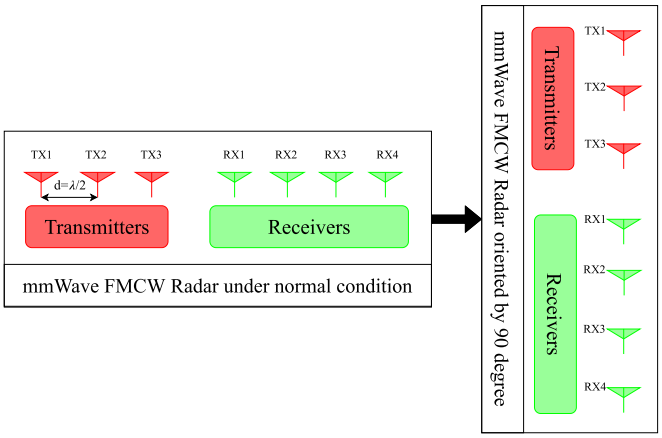


Fig. 2. Radar orientation.

The instantaneous frequency of the transmitted chirp explained above is given by

$$f_{tr}(t) = f_c + St = f_c + \frac{B}{T_{chirp}}t, \quad 0 \le t \le T_{chirp} \tag{1}$$

where, f_c is the start frequency of the chirp, B is the bandwidth, T_{chirp} is the chirp time and S is the chirp slope. The phase of the transmitted chirp is given by

$$\phi_{\text{tr}}(t) = 2\pi f_c t + \pi \frac{B}{T_{\text{chirp}}}t^2, \tag{2}$$

Using (1) and (2), the transmitted signal within a period (T_{chirp}) is given by

$$S_{tr}(t) = A\cos(2\pi f_{tr}(t)t + \phi_{\text{tr}}(t)),$$

where, $f_{tr}(t)$ is frequency of the transmitted chirp and $\phi_{tr}(t)$ is phase of the transmitted chirp [16]. Similarly, the received

Range (m) → Angle (degree) ↓	5	7	9	11	13	15	17	19	21	23	25
0	a-6	h-4, m-2	f-5	c-1	d-1	e-4	i-5, n-2	a-7	g-4	k-2	L-7
30	h-6	b-6, n-1	a-1, L-6	i-6, m-3	c-2	d-2	f-1, k-5	e-5	b-7	g-5	j-7
60	g-1	e-1, k-6	c-6, n-3	a-2, L-3	b-1	i-4	d-3	f-2	j-2	c-7	h-3
90	i-1, n-6	g-2	e-2, m-6	d-6	a-3, k-4	b-2, L-4	c-3	h-1	f-3	j-4	d-7
120	e-6	i-2, L-1	j-1	h-2, n-4	g-7, m-4	a-4	b-3	c-4	d-4	f-4	k-7
150	j-6, k-3	f-6	h-3	g-3	e-3, n-5	f-7	a-5, m-5	b-4, L-5	c-5	d-5	i-7
180	L-2	j-3	g-6	b-5	i-3	m-1	e-7	n-7	k-1	h-7	m-7

Where [a, b, c, d, e, f, g, h, i, j, k, L, m, n] are 14 different cases

-1	Human-1
-4	Human-4
-3	Human-3
-2	Human-2

-5	Human-5
-6	Drone
-7	Car

Fig. 3. Measurement cases.

signal after the reflection from remote target is just the delayed version of the transmitted signal and is given by

$$S_{rx}(t) = S_{tr}(t - \tau),$$

where, $\tau = 2R/c$ is the time delay between the transmitted and the received signal, R is the range of the object from radar, and c is the velocity of light in the vacuum. The reflected signal from the target is passed through a quadrature mixer, in which it is mixed with the in-phase and quadrature phase of the transmitted signal and the complex IF signal is generated also shown in Fig. 1. This IF signal is passed through a low-pass filter and digitized using an ADCs at a sampling frequency of 10 MSPS [16]. The frequency of this IF signal is related to the range of the target in direct-line of sight (D-LOS) that reflects the transmitted chirp is given by (3).

$$f_{IF} = \frac{B.2R}{T_{\text{chirp}}.c}, \quad R = \frac{f_{IF}c}{2S}, \tag{3}$$

where, R, f_{IF}, c, and S are range, intermediate frequency, velocity of light in vacuum and slope of the FMCW chirp, respectively. It is noted that the range measurements are unaffected by the Radar orientation.

III. MEASUREMENTS AND SIGNAL PROCESSING

The measurements are done in a practical outdoor with multi-class-target and environment. Measurement scenes (cases) are labelled with a, b, c, d ... n. Targets in each measurement case are labelled as 1, 2, 3, 4, 5, 6 and 7. The object, a-1 is read as human-1 in measurement case "a" located at 5m range from radar and 0 degrees angle with respect to radar. An example of the measurement setup shown in the Fig. 4 is used for the data collection and dataset creation for further machine learning processing. Measurements are made in an outdoor parking-lot with

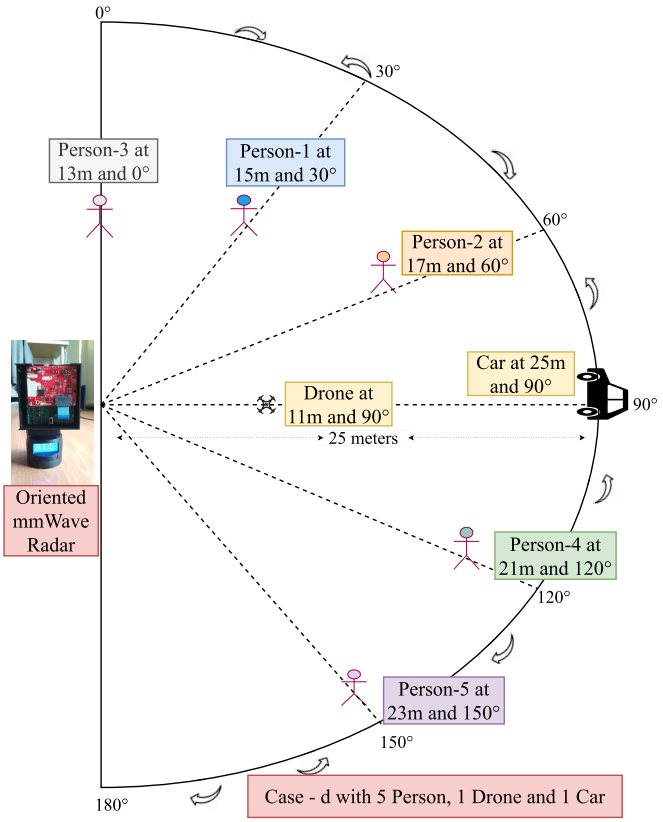


Fig. 4. Test setup for multi-target case - "d".

humans, car and an aerial vehicle at various ranges and angles as shown in Fig. 3. It can be observed from the Fig. 3 that in case-d, person-1 is at 0° and 13 m range, person-2 is at 30° and 15 m range, person-3 is at 60° and 17 m range, person-4 is at 120 degrees and 21 m range and person-5 is at 150° and 23 m range from the Radar. Also aerial vehicle/drone is positioned at 90° and 11 m range, car is positioned at 90° and 25 m range. Arrangement for case-d can also be seen in Fig. 4. The raw IF data is collected from the mmWave Radar during the measurements of all cases. The collected raw data is then post-processed in MATLAB, the file can also be accessed in [17].

A. Range Axis Calculation From ADC Samples

- Firstly, Unwanted zero clutter was removed. Radar clutter is classified into two types: mainlobe clutter and sidelobe clutter [18]. The mainlobe clutter is caused by unwanted ground returns within the radar beamwidth (mainlobe), whereas the sidelobe clutter is caused by unwanted returns from any other direction outside the mainlobe. When the radar is placed at a lower height from the ground, the main lobe / sidelobe intersects the ground. Because the area of ground in the radar beam is often quite large, the ground return can be much larger than the target return. The clutter associated with ground returns close to radar is removed by removing associated components per range bin in range FFT.
- Each chirp contains 256 ADC samples (N) corresponding to frequency bins (fr_{bin}) can be translated into range bins (R_{bins}):

$$fr_{bin} = (0 : N-1)\frac{f_s}{N}, \tag{4}$$

$$R_{bin} = \frac{c fr_{bin}}{2S}, \tag{5}$$

where, $f_s = 10$ MHz, $c = 3 \times 10^8$ m/s, and $S = 29.982$ MHz/μs are the sampling frequency, velocity of light in vacuum, and the slope of chirp, respectively.
- Range is estimated by taking 256-point 1D FFT of the sampled IF signal using (6) and frequency bins are translated to range bins using (R_{bins}) in (5).

$$X_k = \sum_{n=0}^{N-1} Xr_n e^{-j2\pi kn/N}, \tag{6}$$

where, $N = 256$ is number of ADC sample in a chirp, $k = 0 : N-1$ is the element indexing or iterations, Xr_n is the input IF signal of nth index.
- The range bin is estimated for all 128 chirps in a frame then averaged over all of the 128 chirps to get the average range estimation for a frame.
- The range axis is what we get from a frame. This is done for all frames to get the angle axis in range-angle heatmap. The angle axis details are given in the following section.

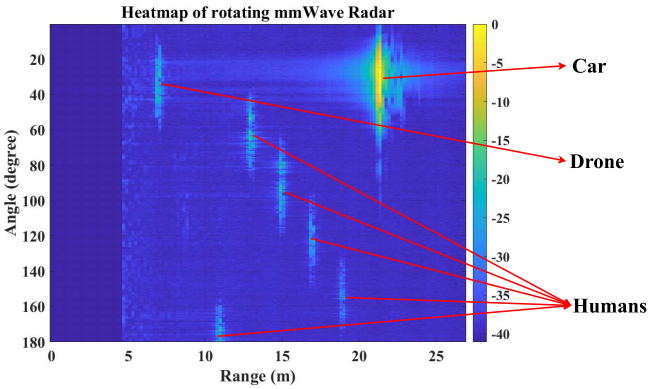


Fig. 5. Angle versus range heatmap for multiclass-target measurement case - "b".

B. Angle Axis Calculation From the Frames

- A rotor is attached to Radar to cover a FoV, θ_{FoV}, in T seconds. In our case, we have FoV of 180°. In this T duration, Radar transmits N_f number of frames. These N_f frames cover the entire FoV, θ_{FoV}, in T seconds. Hence, this entire FoV θ_{FoV} is divided into angle bins, θ_{bin}. The angle bins (θ_{bin}), total FoV (θ_{FoV}), and total number of frames (N_f) are related by

$$\theta_{bin} = \frac{\theta_{FoV}}{N_f}, \tag{7}$$

- For example, in our case we have set θ_{FoV} to 180° and N_f is set to 800, then each frame corresponds to 0.225°, i.e. 4.44 frames per degree. Hence, a heatmap is plotted which shows angle FoV versus range axis in which each bin corresponds to power at that location. Fig. 7 shows

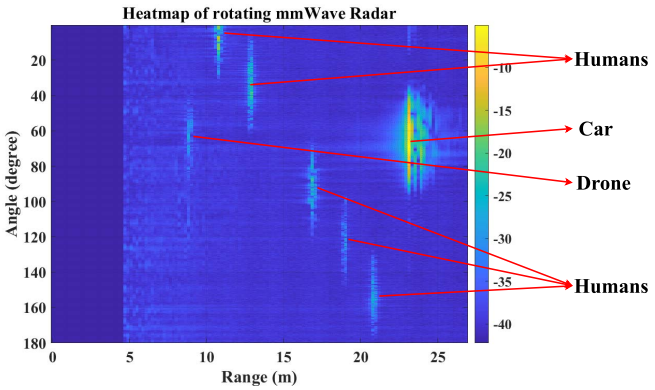


Fig. 6. Angle versus range heatmap for multiclass-target measurement case - "c".

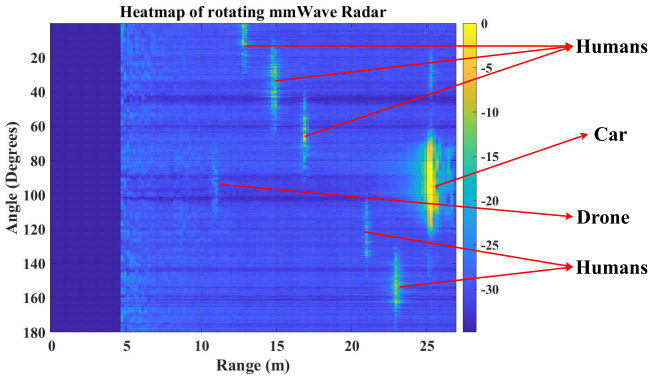


Fig. 7. Angle versus range heatmap for multiclass-target measurement case - "d".

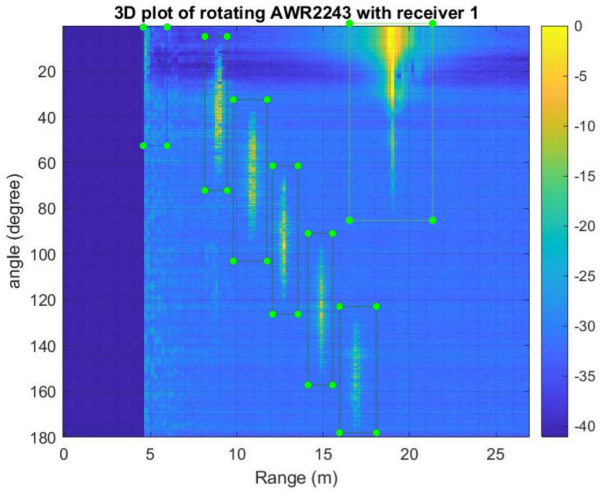


Fig. 8. Marking boxes enclosing objects for creating ground truth.

angle versus range heatmap processed with the proposed algorithm for measurement case-d shown in Fig. 4 for a single receiving antenna.

- we have four receiving antennas and extracted four heatmaps of that particular case.
- Each case's heatmap with four antennas is saved for machine learning vision YOLO model for classification of the objects used in the measurement cases. Here, heatmap is a plot in which the power of data of x-y axes is represented as light colour (high power) to dark colour (low power).

Algorithm 1 Angle vs Range Heatmap Extraction

Require: $range\_angle$, range and angle of Raw data Xr having $num\_frame = 800$, $num\_chirp = 128$ for 4 antennas

Ensure: % % Raw data Xr contains 3-D data, i.e. data corresponding to receiving antennas, having frames $num\_frame = 200$ and each frame has chirps $num\_chirp = 128$ chirps % %

for $n \leftarrow 1$ to num_frame **do**

$x\_ant_1(n) \leftarrow Xr(n,1))$ %% antenna 1

$x\_ant_2(n) \leftarrow Xr(n,2))$ %% antenna 2

$x\_ant_3(n) \leftarrow Xr(n,3))$ %% antenna 3

$x\_ant_4(n) \leftarrow Xr(n,4))$ %% antenna 4

for $m \leftarrow 1$ to num_chirp **do**

Ensure: % % FFT of Raw data x_ant of receiving antenna - 1, 2, 3, 4 of mth chirp, nth frame % %

$Xk_1(n,m) \leftarrow FFT(x\_ant_1(n,m))$

$Xk_2(n,m) \leftarrow FFT(x\_ant_2(n,m))$

$Xk_3(n,m) \leftarrow FFT(x\_ant_3(n,m))$

$Xk_4(n,m) \leftarrow FFT(x\_ant_4(n,m))$

Ensure: % % After zero-clutter removal % %

$X_1(n,m) \leftarrow zero\_clut\_rmv(Xk_1(n,m))$

$X_2(n,m) \leftarrow zero\_clut\_rmv(Xk_2(n,m))$

$X_3(n,m) \leftarrow zero\_clut\_rmv(Xk_3(n,m))$

$X_4(n,m) \leftarrow zero\_clut\_rmv(Xk_4(n,m))$

end for

Ensure: % % Mean range computation for 128 chirps % %

$range_1 \leftarrow mean(X_1(n,[1:num\_chirp]))$

$range_2 \leftarrow mean(X_2(n,[1:num\_chirp]))$

$range_3 \leftarrow mean(X_3(n,[1:num\_chirp]))$

$range_4 \leftarrow mean(X_4(n,[1:num\_chirp]))$

Ensure: % % Angle (FoV) computation from frames %

$angle \leftarrow n * \theta_{FoV}/num\_frame$

end for

Ensure: % % Heatmap for receiver 1, 2, 3 & 4 are stored for a single case as % %

$range\_angle_1 \leftarrow [range_1, angle]$

$range\_angle_2 \leftarrow [range_2, angle]$

$range\_angle_3 \leftarrow [range_3, angle]$

$range\_angle_4 \leftarrow [range_4, angle]$

The range and angle (FoV) is also discussed and thoroughly explained in Algo. 1.

IV. Machine Learning Model and Performance Evaluation

From the above mentioned algorithms and measurements made in multi-target cases, we estimate the angle and range of an object. To classify the object we have used Yolo model on the above collected data with measurements.

A. Dataset

We created a dataset with the following features: Next, we analyse the estimated angle against actual angle. It can be observed from Fig. 11 that estimated angle deviates more as the range of the target increases. This is quite intuitive as the range increases, reflected signal strength falls and that causes

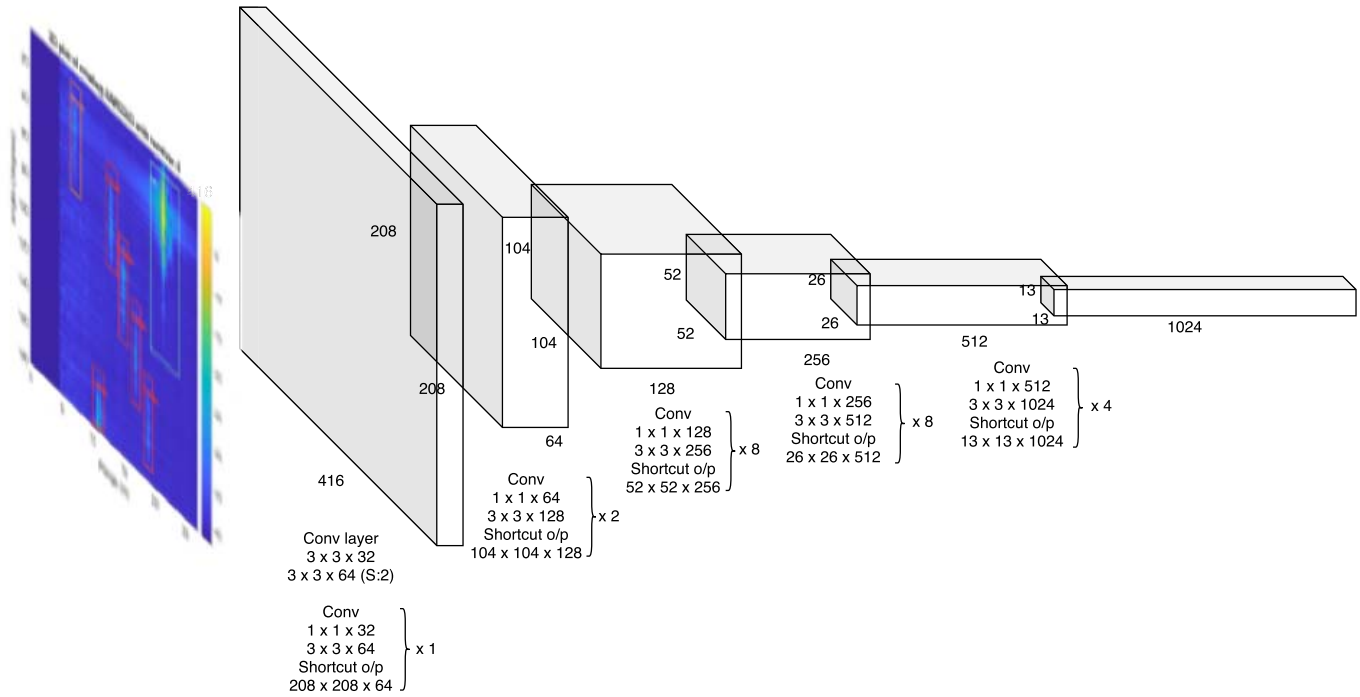


Fig. 9. Darknet 53 architecture used for detection.

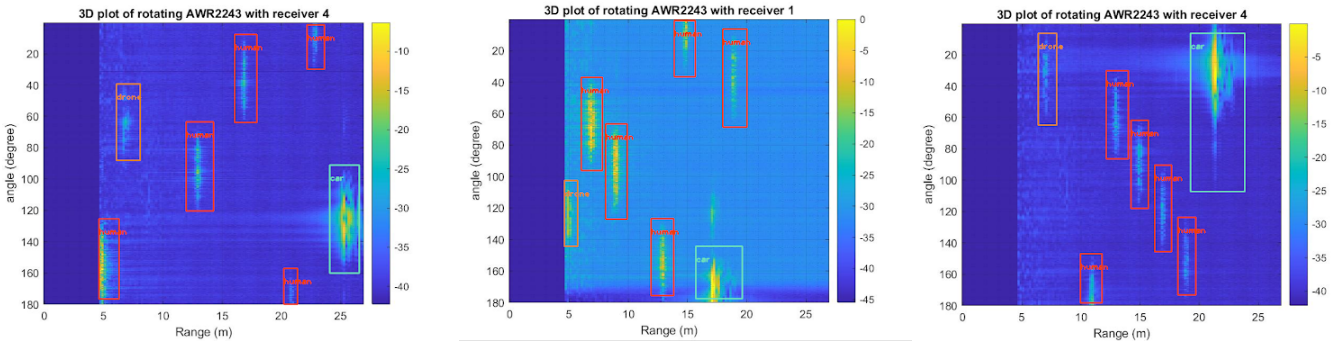


Fig. 10. Detection of car, human and drone in (a) case 1 (b) case 2 (c) case 3.

more spread in angle-range cluster, especially for the targets with smaller cross sections. Angle estimation could be more accurate if targets have larger cross section, for e.g. cars.

B. Preprocessing

For the data to be fed into the Yolo algorithm marking of objects in the image is necessary. These markings will act as a ground truth for our model. The images are marked using LabelImg software. For each object present in the image, a rectangular box is created enclosing it as shown in Fig. 8 and by using these individual boxes' coordinates a corresponding text file is created. Each line in the text file corresponds to an object present in image in 5 variable format representing its class, x coordinate, y coordinate, width and height.

C. Model Parameters

The model is used to classify three classes Humans, Drones and Cars. We have used the batch size of 64 and 16 subdivisions. For three classes, 24 filters in the final layer as per (8).

$$filters = (num\_classes + 5) * 3 \tag{8}$$

We have used learning rate of 0.001 with a 0.9 momentum and 0.0005 decay as mentioned in Table II

D. Model Summary

You Only Look Once (Yolo v3) [10] comes from the family of Convolutional Neural Networks for real time object detection in camera generated images. Yolo is frequently used for vehicle, human, helmet detection and more. These objects have sharp edges which can separate them from the background. In our dataset, the images contain heatmap of objects which has no distinct boundaries, or shape which makes the objects more difficult to recognize. The shape and size can vary as shown in Fig. 10 (a), at 21m the region covered by human is less than half of it covered by human present at 13m or 17m. Similarly for Car in Fig. 10 (a), (b) and (c). In this experiment, we are using the third variation of Yolo expressed as Yolo v3.

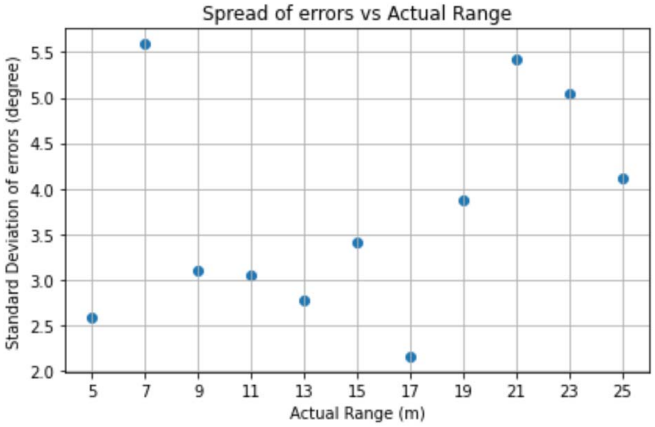


Fig. 11. Spread of errors vs actual range.

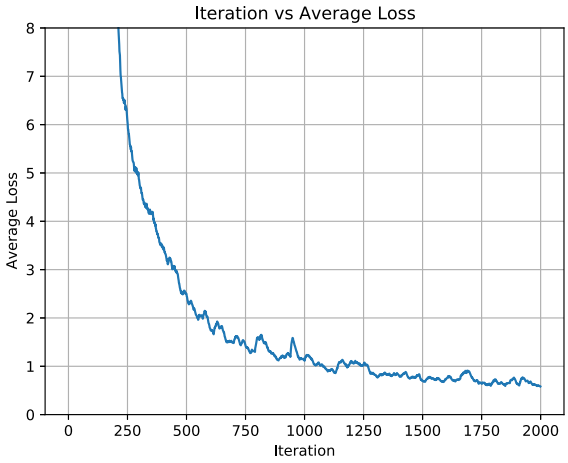


Fig. 12. Iteration vs Average loss curve.

Pre-trained weights are used for transfer learning and faster results.

E. Model Architecture

The underlying architecture of Yolo is Darknet53 [19] as shown in Fig. 9 and is used as feature extractor of our model. Darknet53 has 53 successive 3 × 3 and 1 × 1 convolutional layers. It is a upgraded version of Darknet19 [19] containing residual layers or skip connections like deep Residual networks [20] which used feature maps of previous layer along with and nth backward layer.

$$o_x = \sigma(i_x) + c_x$$
$$o_y = \sigma(i_y) + c_y$$
$$o_w = p_w \exp i_w$$
$$o_h = p_h \exp i_h$$

For each i_x, i_y, i_w and i_h the model predicts out corresponding values, o_x, o_y, o_w, o_h, which represent the bounding box of the object in consideration. Since we are classifying three classes the output vector is larger than these four values. We are using nine anchor boxes and for each anchor box we have 8 output values.

F. Performance Matrics

Precision is a performance metric that computes the ratio of correctly predicted positive observations to all predicted positive observations [21]. This is calculated as:

$$\text{Precision} = (TP)/(TP + FP) \tag{9}$$

The recall performance metric computes the ratio of correctly predicted positive observations to all observations in the actual positive class [21]. This is calculated as:

$$\text{Recall} = (TP)/(TP + FN) \tag{10}$$

The specificity performance metric computes the ratio of true negatives to total negatives [21]. This is calculated as:

$$\text{Specificity} = (TN)/(TN + FP) \tag{11}$$

F1-score is a performance metric that calculates its value using both recall and precision [21]. It is calculated as:

$$\text{F1-score} = 2 * (Recall * Precision)/(Recall + Precision) \tag{12}$$

TABLE II
YOLO V3 CONFIGURATION PARAMETERS

S.No.	Parameter	Configuration
1	Number of Classes	3
2	Batch Size	64
3	Sub division	16
4	Max Batches	6000
5	Filters	24
6	Learning rate	0.001
7	Momentum	0.9
8	Decay	0.0005
9	CUDA version	10.1
10	GPU	Tesla P100-PCIE

Class	Car	Human	Drone	None	Total
Car	380	0	0	7	387
Human	0	924	0	3	927
Drone	0	0	374	9	383

Fig. 13. Confusion matrix.

V. RESULTS

We have run the tests for 10 folds and each fold contains 2000 iterations for weight optimization. In each fold, 20% of cases are removed randomly and the training was done with the remaining 80% cases. We have achieved accuracy ranging from 87.68% for 6th fold upto 97.46% in two cases for 0th and 3rd fold which can be seen in Fig. 14. The average accuracy achieved is 94.53%.

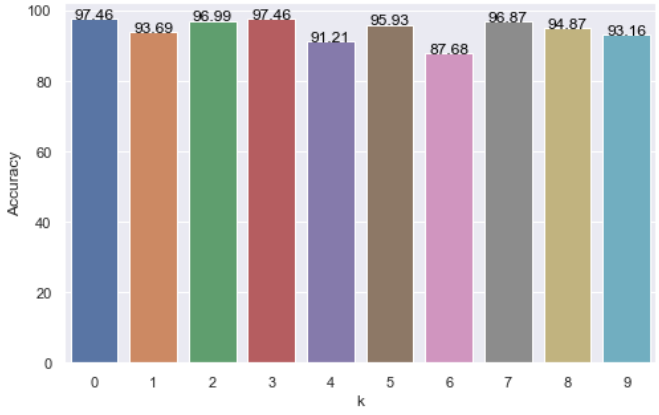


Fig. 14. Accuracy vs k-run.

TABLE III
RECALL, PRECISION AND F1-SCORE EVALUATION METRIC VALUES

	Car	Human	Drone
Precision	0.992	0.997	0.994
Recall	0.989	0.998	0.981
Specificity	0.997	0.997	0.998
F1-Score	0.987	0.997	0.985

We have achieved the best accuracy of 97.46% for the object classification. 1678 objects are rightly classified as their actual class among 1697 samples. As shown in Fig. 10 the bounding boxes are made in different colors signifying different classes car, human and drone. Average loss of convolutional model is dropped significantly and converges really fast in the initial iterations as shown in Fig. 12. The weights used for transfer learning and prediction are of 2000^{th} iteration with 0.794 average loss. As shown in confusion matrix in Fig. 13 we have rightly predicted car 98.1%, human 99.6% and drones 97.6%. All the objects which are not predicted as their actual class are actually missed by the model, hence we counted them as None class. The other performance metrics such as precision, recall, specificity and F1-score are summarized in III. It can be observed that precision, recall and F1-score are high for human compared to car and drone.

When targets get closer, bounding boxes in the range-angle maps overlap. Targets can be resolved so long as these bounding boxes are distinguishable. This eventually increases the processing complexity and degrades the resolution. Fusing the data from additional sensors such as RGB cameras can be used to increase the resolution. As a future work, we plan to fuse the additional sensors with mmWave radars to improve the resolution.

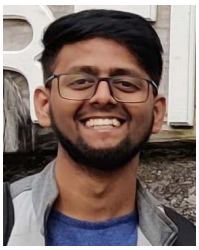
VI. CONCLUSION

The enhancement in both azimuth FoV and elevation FoV of mmWave FMCW Radar has been demonstrated for small scale ground station and traffic management applications. As per state of the art tools and technologies, object detection and recognition in daylight conditions is quite feasible but in a case with low or insufficient illumination, it is still a challenging task. We achieved a high level of predictability about the object's characteristics by combining mmWave radar imaging with advanced computer vision techniques such as Yolo and Faster RCNN. In this article, we achieved the accuracy in the range of 87.68 % - 99.7 % while predicting three different classes in multi-target scenarios using custom range-angle images obtained from mmWave Radar.

REFERENCES

[1] *Advanced Driver Assistance Systems (ADAS)*. Accessed: Sep. 2020. [Online]. Available: https://www.ti.com/applications/automotive/adas/overview.html
[2] L. R. Cenkeramaddi, J. Bhatia, A. Jha, S. K. Vishkarma, and J. Soumya, "A survey on sensors for autonomous systems," in *Proc. 15th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Nov. 2020, pp. 1182–1187.
[3] M. E. Yanik and M. Torlak, "Near-field 2-D SAR imaging by millimeter-wave radar for concealed item detection," in *Proc. IEEE Radio Wireless Symp. (RWS)*, Jan. 2019, pp. 1–4.
[4] S. Ayhan *et al.*, "Millimeter-wave radar distance measurements in micro machining," in *Proc. IEEE Top. Conf. Wireless Sensors Sensor Netw. (WiSNet)*, Jan. 2015, pp. 65–68.
[5] M. Z. Ikram, A. Ahmad, and D. Wang, "High-accuracy distance measurement using millimeter-wave radar," in *Proc. IEEE Radar Conf. (RadarConf)*, Apr. 2018, pp. 1296–1300.
[6] R. Rouveure, P. Faure, and M.-O. Monod, "PELICAN: Panoramic millimeter-wave radar for perception in mobile robotics applications. Part 1: Principles of FMCW radar and of 2D image construction," *Robot. Auton. Syst.*, vol. 81, pp. 1–16, Jul. 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01383402
[7] S. Nowok, S. Kueppers, H. Cetinkaya, M. Schroeder, and R. Herschel, "Millimeter wave radar for high resolution 3D near field imaging for robotics and security scans," in *Proc. 18th Int. Radar Symp. (IRS)*, Jun. 2017, pp. 1–10.
[8] U. Aulenbacher, K. Rech, J. Sedlmeier, H. Pratisto, and P. Wellig, "Millimeter wave radar system on a rotating platform for combined search and track functionality with SAR imaging," *Proc. SPIE*, vol. 9252, pp. 1–11, Oct. 2014, doi: 10.1117/12.2064824.
[9] D. Vivet, P. Checchin, and R. Chapuis, "Localization and mapping using only a rotating FMCW radar sensor," *Sensors*, vol. 13, no. 4, pp. 4527–4552, Apr. 2013, doi: 10.3390/s130404527.
[10] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," Apr. 2018, *arXiv:1804.02767*. [Online]. Available: https://arxiv.org/abs/1804.02767
[11] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," *Comput. Vis. Image Understand.*, vol. 178, pp. 30–42, Jan. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1077314218304296
[12] H. Eum, J. Lee, C. Yoon, and M. Park, "Human action recognition for night vision using temporal templates with infrared thermal camera," in *Proc. 10th Int. Conf. Ubiquitous Robots Ambient Intell. (URAI)*, Oct. 2013, pp. 617–621.
[13] M. Pavelka and V. Jirovský, "LiDAR based object detection near vehicle," in *Proc. Smart City Symp. Prague (SCSP)*, May 2017, pp. 1–6.
[14] *AWR2243 Single-Chip 76- to 81-GHz FMCW Transceiver*. Accessed: Sep. 2020. [Online]. Available: https://www.ti.com/lit/ds/symlink/awr2243.pdf?ts=1600889609815
[15] The Math Works. *MATLAB Version 2019a, Computer Software*. Accessed: Sep. 2020. [Online]. Available: https://www.mathworks.com
[16] *The Fundamentals of Millimeter Wave Sensors*. Accessed: Sep. 2020. [Online]. Available: https://www.ti.com/lit/wp/spyy005/spyy005.pdf?ts=1598299967367
[17] P. K. Rai. *Targets Classification mmWave FMCW Radar*. Accessed: Sep. 2020. [Online]. Available: https://github.com/prabhatrai111/Targets-Classification-mmWave-FMCW-Radar
[18] M. Sanoal and M. Santiago, "Automotive FMCW radar development and verification methods," Ph.D. dissertation, Dept. Comput. Sci. Eng., Chalmers Univ. Technol., Gothenburg, Sweden, 2018.
[19] J. Redmon. (2016). *DarkNet: Open Source Neural Networks in C*. [Online]. Available: http://pjreddie.com/darknet/
[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, pp. 1–12, Dec. 2015. [Online]. Available: http://arxiv.org/abs/1512.03385
[21] A. Dayal, N. Paluru, L. R. Cenkeramaddi, J. Soumya, and P. K. Yalavarthy, "Design and implementation of deep learning based contactless authentication system using hand gestures," *Electronics*, vol. 10, no. 2, p. 182, Jan. 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/2/182

Siddharth Gupta received the B.E. degree in electronics and telecommunication engineering from Jabalpur Engineering College, India, in 2016. He is currently pursuing the M.Tech. degree in electrical engineering from IIT Hyderabad, Hyderabad, India. He is currently working as a Research Intern with the Department of ICT, University of Agder, Kristiansand, Norway. His current research interests are in the different areas of wireless communications, machine learning, smartphone applications, and signal processing.



Prabhat Kumar Rai received the B.Tech. degree in electronics and telecommunication engineering from the Bhilai Institute of Technology, Durg, India, in 2017. He is currently pursuing the M.Tech. degree in electrical engineering from IIT Hyderabad, Hyderabad, India. He is currently working as a Research Intern with the Department of ICT, University of Agder, Kristiansand, Norway. His current research interests are in the different areas of wireless communications, RADAR communication, and signal processing.

Abhinav Kumar (Senior Member, IEEE) received the B.Tech., M.Tech., and Ph.D. degrees in electrical engineering from the Indian Institute of Technology Delhi, in 2009 and 2013, respectively. From September 2013 to November 2013, he was a Research Associate with the Indian Institute of Technology Delhi. From December 2013 to November 2014, he was a Postdoctoral Fellow with the University of Waterloo, Canada. Since November 2014, he has been with the Indian Institute of Technology Hyderabad, India, where he is currently an Associate Professor. His research interest is in the different aspects of wireless communications and networking.

Phaneendra K. Yalavarthy (Senior Member, IEEE) received the M.Sc. degree in engineering from the Indian Institute of Science, Bengaluru, India, and the Ph.D. degree in biomedical computation from Dartmouth College, Hanover, NH, USA, in 2007. He is a Professor with the Department of Computational and Data Sciences, Indian Institute of Science, Bengaluru. His research interests include medical image computing, medical image analysis, and biomedical optics. He is a Senior Member of SPIE and OSA, and serves as an Associate Editor for IEEE TRANSACTIONS ON MEDICAL IMAGING.

Linga Reddy Cenkeramaddi received the master's degree in electrical engineering from the Indian Institute of Technology, New Delhi, India, in 2004, and the Ph.D. degree in electrical engineering from the Norwegian University of Science and Technology, Trondheim, Norway, in 2011. Before joining the Ph.D. Program at NTNU, he worked for Texas Instruments in mixed signal circuit design. After finishing his Ph.D., he worked in radiation imaging for an atmosphere space interaction monitor (ASIM Mission to International Space Station) with the University of Bergen, Norway, from 2010 to 2012. At present, he is the Group Leader of the Autonomous and Cyber-Physical Systems (ACPS) Group and working as an Associate Professor with the University of Agder at Campus Grimstad, Norway. His main scientific interests are in cyber-physical systems, autonomous systems, and wireless embedded systems.