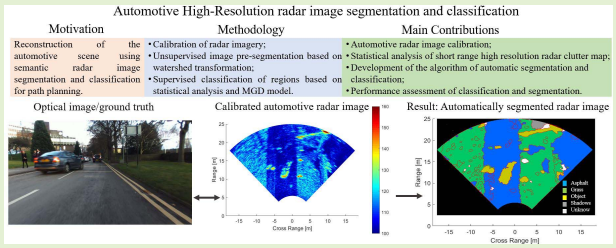
Image Segmentation and Region Classification in Automotive High-Resolution Radar Imagery

Yang Xiao, Liam Daniel, *Member, IEEE*, and Marina Gashinova

***Abstract*—Image segmentation and classification of surfaces and obstacles in automotive radar imagery are the key technologies to provide valuable information for path planning in autonomous driving. As opposed to traditional radar processing, where clutter is considered as an unwanted return and should be effectively removed, autonomous driving requires full scene characterization. Hence, clutter carries necessary information for situational awareness of the autonomous platform and needs to be fully assessed to find the passable areas. In this paper, we proposed a method of automatic segmentation of automotive radar images based on two main steps: unsupervised image pre-segmentation using marker-based watershed transformation, followed by the supervised segmentation and classification of regions containing objects and surfaces based on the use of statistical distribution parameters. Several distributions were considered to characterize returns from specific region types of interest within the scene (denoted as classes) in calibrated radar imagery—the extracted distribution parameters were assessed for their ability to distinguish each class. These parameters were then used as features in a multivariate Gaussian distribution model classifier. Both the performances of the proposed supervised classification algorithm and the automatically segmented results were investigated using F1-score and Jaccard similarity coefficients, respectively.**



***Index Terms*—Automotive radar imagery, image calibration, image segmentation, distribution feature extraction, Weibull distribution, multivariate Gaussian distribution, watershed transformation.**

I. Introduction

AUTOMOTIVE sensors are the backbone of advanced driver assistant systems (ADASs) and sensing systems for self-driving cars since they can provide robust assessment of the proximate environment (up to several hundred meters) for path planning and decision making [1], [2]. The sensors generally involved includes LIDAR, camera, ultrasound, global positioning system and automotive radar, the latter is being reliable all-weather technology.

The development of radar was for decades focused on reliable detection and tracking algorithms for typical road actors such as cars and pedestrians [3]. Some elements of classification were implemented based on the characteristics of radar returns, though it could potentially provide wider scene information e.g. class of surface, traditionally referred to as clutter, which was however subject to suppression. Additionally, due to the low resolution, traditional automotive radar did not possess the capabilities to effectively define, assess and classify clutter areas.

As opposed to traditional ADAS systems, for autonomous driving the identification of passable areas is the key task to achieve automated path planning and obstacle avoidance [3]. One of the approaches to identify the passable region from the sensed data is to segment and classify the surfaces and objects viewed by various sensors within a scene. Therefore, radar should be able to function as "clutter mapping" sensor. There is a substantial research in the area of remote sensing on image segmentation, feature extraction and classification. They are critical steps for object-based image analysis (OBIA) of SAR imagery in a wide range of applications of remote sensing, such as sea ice monitoring [4], land cover classification [5], agricultural crop identification [6] etc.

For automotive sensing to this day, however, the radar sensors have never exceeded the resolution required for traditional ADAS, which is far from that delivered by SAR imagers, so that image segmentation and surface classification were hardly considered in mainstream automotive radar sensing

Manuscript received October 14, 2020; revised November 22, 2020; accepted November 23, 2020. Date of publication December 9, 2020; date of current version February 5, 2021. This work was supported by Jaguar Land Rover and the UK-Engineering and Physical Sciences Research Council (EPSRC) Grant EP/N012372/1 as part of the jointly funded Towards Autonomy: Smart and Connected Control (TASCC) Programme. The associate editor coordinating the review of this article and approving it for publication was Prof. Yudong Zhang. *(Corresponding author: Yang Xiao.)*

The authors are with the Department of Electronic, Electrical, and Systems Engineering, University of Birmingham, Birmingham B15 2TT, U.K. (e-mail: yxx752@student.bham.ac.uk; l.y.daniel@bham.ac.uk; m.s.gashinova@bham.ac.uk).

Digital Object Identifier 10.1109/JSEN.2020.3043586

research. Therefore, the key requirements for provision of autonomous navigation can be summarized as follows: 'radar clutter' must be resolved (imaged), characterized (positions, extent and texture estimated to generate the 'depth' of scene), segmented and labelled, enabling deep scene reconstruction similar to SAR, where four main approaches are used:

1) The edge-based segmentation algorithms which are relatively simple and computationally effective; example algorithms are the fuzzy model and watershed transformation (WT). The fuzzy logic based techniques in remote sensing image segmentation include fuzzy c-means clustering (FCM) [7], [8] and fuzzy stochastic estimation maximization (FSEM) [9]. The main challenge of FCM is the sensitivity to weight adjustments of adopted data attributes and the initialization process [10], and the FSEM method was limited to two classes segmentation. WT, which segments objects using closed boundaries has become the most popular edge-based segmentation methods in the remote sensing community. The WT method based on the typical gradient operator can however easily produce over-segmented areas due to image noise [11]. Therefore, the generation of gradient image is the key to improve the performance of WT which could constrain the initial edges. Multiple techniques to produce the gradient image have been explored, such as the Canny edge-detector [12], marker based edge embedded WT [13] and hierarchical method [14]. Algorithms of edge-based segmentation do not function well on images with smooth transitions and low contract, and they are generally sensitive to noise [15], however could be very effective for initial pre-segmentation of the scene, as will be shown in this paper.

2) The region-based methods. These are based on two basic operations: region merging and splitting. Region merging (or growing) starts from an initial region (or pixel), then merge the regions (or pixels) according to a specific homogeneity criterion which determines whether the area belongs to the growing region or not [16]. In contrast, region-splitting starts from the entire image, and then split the image into segments based on the criterion for inhomogeneity of regions [17]. The criteria used in region-based methods include the properties of spectral and spatial characteristics, shape, texture, size, prior knowledge and context of scenarios. The mathematical description of these properties is based on statistical distribution features [18], variance, Moran's I [19], and F measure etc. Typically region-based segmentation is based on either merging or splitting or combination of both techniques [15], [19]. However, finding the appropriate parameters of the criterion for judging splitting or merging of regions is a significant challenge.

3) Hybrid method (HM) could overcome the limitation of both edge-based and region-based segmentation methods. Most of the studies of HMs start from the edge-based method to create an over segmented image, and then merge the similar segments based on either homogeneity or heterogeneity of regions. The advantages of HMs are (i) the freedom to select the methods for initial segmentation and the following region merging; (ii)less sensitivity to image noise texture.

4) Semantic methods are generally supervised approaches which allot the label to each pixel. Machine learning-based semantic methods include Neural network [20], Support vector machine [21] and Convolution Neural Network [22] etc. Semantic methods showed encouraging results on SAR image segmentation, although require a vast amount of training data and a significant number of parameters for validation [15].

The approach proposed in this paper aims at image segmentation and supervised classification of automotive radar images with multiple classes to be identified, so that each pixel within the image is ultimately labelled according to a determined class. It can be defined as a HM consisting of initial edge-based pre-segmentation using the WT method and a subsequent classification and region merging process on the pre-segmented regions of interest (RoI's) based on the statistical distribution parameters extracted from radar data. These will be used as features in a multi-variate Gaussian distribution (MGD) based supervised classifier.

For effective path planning, we aim to distinguish between tarmac, other traversable surfaces and regions consisting objects. Radar scattering from specific objects and clutters will be investigated as a function of grazing angle and resolution. Several types of distribution functions which are physically meaningful for the kind of clutter in the automotive environment will be investigated, and fitting errors evaluated and compared between region classes. In the results of full segmentation, each pixel is labelled according to its class, grouped into clusters (segments) and color coded for reconstructed scene representation.

The paper is organized as follows: Section II introduces the automotive radar data set used in this part of research and outlines the methodology of image calibration. Section III states the procedure of distribution parameter extraction and discusses the fit errors. Section IV proposes the supervised classification algorithm based on the MGD model using identified features. The effective and redundant factors for the classification improvement are determined from F1-score results of classification. Section V introduces the algorithm of automatic segmentation and discusses the results. Finally, conclusions are formulated, and the further steps are outlined.

II. Radar Image Dataset

Here we describe the radar system used for data collection, the formation and calibration of radar imagery and the dataset used for the analysis in this paper.

A. Radar System, Image Formation and Dataset

The radar system is an experimental high-resolution 79 GHz FMCW real aperture radar, designed by the University of Birmingham and ELVA-1 [31] with parameters presented in Table I.

The radar is quasi-monostatic, has a ranging frequency of 33.6 kHz/m and is mechanically scanned in azimuth. Range profiles are generated at equally spaced positions within the scan to form a radar image over a specified field-of-view (FoV) with an azimuth resolution dictated by the antenna beam width.

The fine azimuth resolution achieved through using narrow beam imaging antennas combined with wide bandwidths/fine

TABLE I
79 GHz FMCW IMAGING RADAR PARAMETERS

Parameter	Value	Units
Centre Frequency	78.5	GHz
Bandwidth	5	GHz
Transmit Power	13	dBm
PRF/PRI	232/4.3	Hz/ms
Chirp Duration	1	ms
Az. Beam Width (2-way)	1.7	°
El. Beam Width (2-way)	7.2	°
Antenna Gain	30	dBi
Polarization	VV	

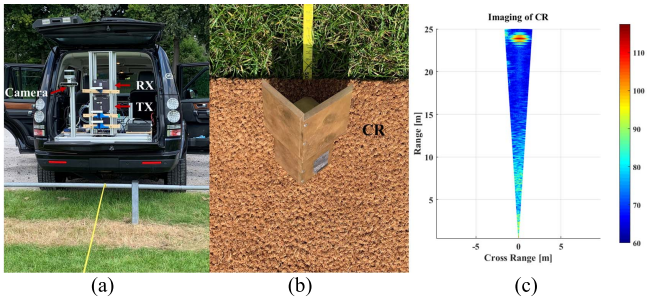


Fig. 2. (a) the experimental setup for measuring radar system response and propagation loss using a CR; (b) CR on ground, supported by a mat to prevent blockage by grass; (c) a radar image of the CR.

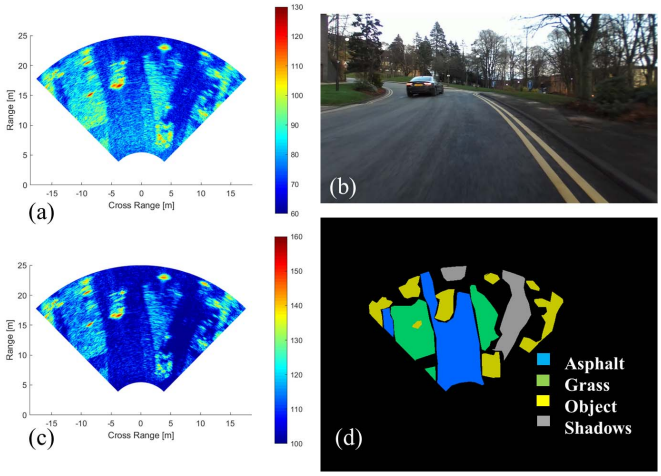


Fig. 1. Automotive radar image set. (a) The uncalibrated radar image in dB power values, (b) corresponding optical image, (c) calibrated radar image and (d) the color coded labeled radar image.

range resolutions provides uniquely high-resolution imagery. Lets' stress here that there is no loss of generality to application of the algorithms applied hereinafter for imagery obtained using synthetic beamforming techniques e.g. MIMO or SAR.

An example radar image can be found in Fig.1 (a). Fig.1 (b) is a corresponding optical image of the scene gathered using a Stereolabs ZED stereo video camera [23] used for ground truth. The radar image here is expressed in uncalibrated dB power units and will be termed an image frame from hereon in. Radar videos are formed using a sequence of image frames from consecutive radar scans.

A dataset containing 330 frames of radar data is used, which was acquired with a scan rate in the order of 1 Hz with 199 azimuth range profiles within a FoV of 90°. The dataset was recorded from a moving vehicle platform within the urban environment of the campus of the University of Birmingham, UK. It contains multiple region types, including asphalt roads, grass areas, and several roadside and urban objects such as buildings, pedestrians, signposts and vehicles [24].

B. Radar Image Composition and Pre-Processing

The mono-static radar image is composed of the backscatter returns from the scene objects, such as cars, tarmac road and grass areas shown in Fig 1 (a), (b). In classical radar applications, these are considered as clutter regions, obscuring detection of target objects (moving cars and pedestrians) and are subject to suppression. For path planning of autonomous platforms however, they are equally as important as targets, and segmentation and classification of all regions is fundamental. The surfaces, road infrastructure and indeed the target objects can be distinguished visually by their intensity contrast. The statistical analysis of these areas forms the basis of our classification procedure and will be discussed in Section III.

For the sake of image segmentation we will treat all regions as clutter, the power return, P_r^s, within the resolution cell of a particular clutter area is governed by the following factors expressed in the form of a radar equation,

$$P_r^s \propto S(R) \cdot L_{\text{atm}}(R) \cdot L_{\text{prop}}(R) \cdot \sigma_0\left(\theta_{\text{g}}, P, h_{\text{rms}}, L_{\text{c}}, \varepsilon_{\text{r}}\right) \cdot A_{\text{cell}}(R, \theta_{\text{az}}, \Delta R) \quad (1)$$

where range dependent terms are: S is the radar response characteristic, L_{atm} is an atmospheric loss, which usually does not exceed 2-3 dB for automotive ranges, L_{prop} is the propagation loss and A_{cell} is the area of the resolution cell, governed by the antenna azimuth beamwidth θ_{az}, and the radar range resolution ΔR. The normalized RCS of the clutter area σ_0 is a function of many factors including grazing angle, θ_g, polarization, P, surface roughness parameters: rms surface height, h_{rms} and correlation length, L_c, and material dielectric permittivity, ε_r [25].

In order to provide a more general segmentation algorithm transferrable between radar systems, all radar response characteristics should be removed from the imagery. Additionally, propagation effects should be compensated leaving solely the effects of the surface scattering. The following sub-section describes the calibration procedure.

1) Image Calibration: A trihedral corner reflector (CR) with an RCS of 17.9 dBsm was placed at ground level at several ranges from the radar and imaged within a narrow FoV of ±5°. The measurement setup is shown in Fig. 2 (a) and (b), an example of the imaged CR is shown in Fig. 2 (c).

The average Tx/Rx height is 1.2 m. At each range within 5-60 m the CR was positioned to ensure maximum return power, and was measured several times. To determine if there was a potential effect from surface type, this was carried out on both grass and asphalt, being two surfaces of significantly different roughness. Using the radar equation and known RCS

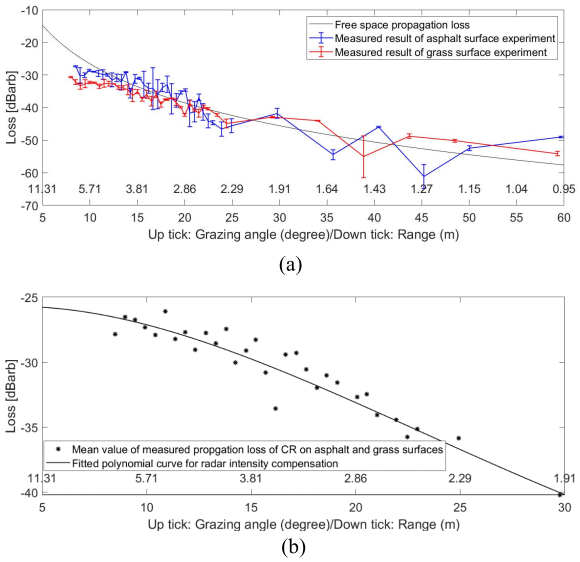


Fig. 3. (a) Measured loss compared with the free space propagation model. (b) Average loss calculated from the grass and asphalt measurements and polynomial loss fit. The label of x-axis shows both range information and the corresponding grazing angles for comparison.

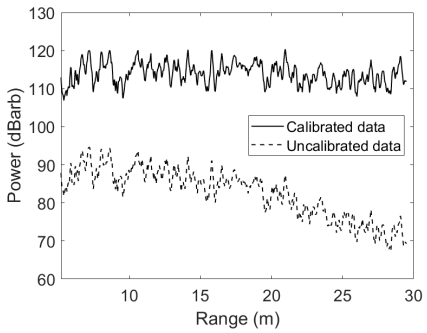


Fig. 4. Comparison of uncalibrated and calibrated radar range profiles.

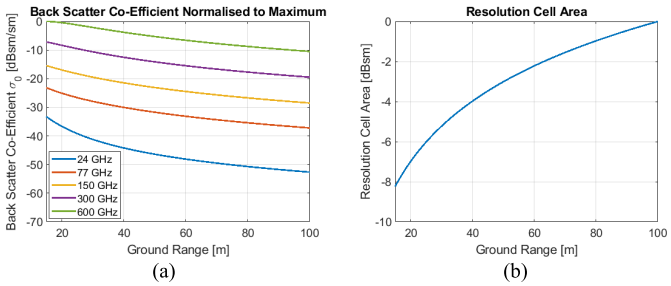


Fig. 5. (a) Back scatter coefficient as a function of frequency and range for a rough asphalt surface (b) resolution cell area as a function of range for experimental radar setup.

of the CR we can evaluate propagation loss as function of range, which relates also to grazing angle. The result is shown in Fig. 3(a), alongside modelled free space propagation loss [36].

It can be seen in Fig. 3(a) that at ranges > 15 m, the loss follows the free space model trend. This is due to the radar intermediate frequency (IF) response being flat in this region and multipath/ground reflection playing a reduced role when considering a ground-based target. At ranges < 15 m, we observe a deviation from the free space model which may be attributed to a combination of drop off in the radar IF response and antenna elevation pattern effects. The important conclusion is that the loss is not significantly different between grass and asphalt measurements and so a common loss curve/calibration can be used for all imagery, not specific to the surface type we are trying to determine. To this end, a polynomial fit must be used to apply the range response corrections as shown in Fig. 3(b), where a 4th order polynomial was used as a close fit to our measurements:

$$L_{prop}^{fit} = -5.7 \times 10^{-6} R^4 + 0.001 R^3 - 0.05 R^2 + 0.36 R - 26.4 \tag{2}$$

Compensation of the radar return w.r.t. range was applied to each range profile within the radar image frame. Fig. 4 shows an example of a range profile before and after the compensation and infers that such compensation results in a "whitening" of the range profile. This effect is shown across all classes of surface in the compensated radar map of Fig. 1(a) which is shown in Fig. 1(c).

2) Backscatter Coefficient: The sheer variation of grazing angle shown on the x-axes upper scale in Fig. 3 is specific to the short-range topology, distinguishing it from other radar imaging systems such as airborne and spaceborne synthetic aperture radar. This variation in grazing angle implies a variation in the rough surface backscatter coefficient σ_0 which in general can be split into 3 regions [26]: linear increase at small grazing angles, plateau at intermediate grazing angles and exponential increase at high grazing angles. The true position and prominence of these transitions are radar and surface specific. Thus, after range loss compensation (Fig. 4) we should either remove influence of the grazing angle on the normalized radar cross section of the clutter, or pre-segment image into regions in range where we expect negligible variation of σ_0. The former requires a priori knowledge of the terrain, while for the latter we need to know the behavior of σ_0 for the chosen classes to define range regions of similar returns. The whitening effect of the range loss compensation indicates, however, that the last two terms of (1.1), σ_0 and A_{res}, must have compensated each other. To explain this effect we have first modelled backscatter coefficient as function of grazing angle using Integral Equation Method [27], which in [28] has shown good agreement between measured and modelled results for asphalt at 94 GHz. Fig. 5(a) shows that simulated backscatter coefficient for multiple frequencies drops as the grazing angle reduces. In Fig. 5(b) a plot of the resolution cell area A_{res} (in dBsm) calculated using radar parameters (Table I) and elevation above the ground. Fig. 5 therefore highlight the fact that for the real-aperture imaging radar an increase in resolution cell size with increase in the range counteracts the falloff in the normalised radar cross section. It should be noted that this discussion of σ_0 relates to our average received power level and the ability to look further ahead of the vehicle, though at the expense of cross-range resolution. It does not account for statistical variations in the

return power between individual resolution cells, which may vary with grazing angle as the scattering mechanism changes. This potential statistical variation with range will be discussed in Section III.

C. Image Format and Data Labelling

Now the data has been described, we discuss its labelling for use in examining potential features for supervised learning. For path planning, we need to distinguish between tarmac, other kinds of terrain, roadside objects and road actors (targets). In terms of image segmentation, this represents the classification into four broad classes such as tarmac (e.g. asphalt), non-tarmac surface e.g. grass (requiring a vehicle response to transit from tarmac), shadow and most general – target object regions, which incorporate all impassable obstacles such as buildings, vehicles, streetlights, traffic signs and trees, which generally have a higher RCS in the imagery. Shadow regions behind objects are of particular interest as their identification and subsequent estimation of shadow depth can provide indirect information about an object's height. Shadow like regions will also be formed from regions of very low backscatter, such as surface water. It should be noted that the methodology proposed in this paper is not limited to these chosen classes. All radar images are labelled using the labelling tool of Liblable [29], [30]. The tool operates on a rectangular grid and so a raster image of the radar scan must be used, not the raw range-azimuth data; corresponding video imagery guides the labelling. Thus mapping between two data representations is required to perform the labelling and corresponding radar data extraction. The representations are: the radar range-azimuth map pixels ('resolution cells') converted and plotted in Cartesian coordinates in a plan position indicator (PPI) representation and the radar raster image pixels with three color channels (defined as "raxels"). The relationship between resolution cells and raxels are shown graphically in Fig. 6. A full description of the transformation may be found in [31], it is very specific to the real aperture radar coordinate map and its representation and is summarised in Section III, A. The labelled image of the radar frame in Fig. 1 (a) is shown in Fig. 1 (d), where the pixels of the radar raster image belonging to a specific class are identified by the same color. The pixels of areas which could not be associated with one of four classes in the optical image due to restricted visibility or uncertainty are shown as black background in the labelled image. These will be treated as belonging to an "unknown" class, which can then be refined in advanced context-based and association approaches.

III. Region Statistical Properties and Feature Extraction

The distribution parameters of the labelled radar clutter are extracted and analyzed to determine potential use as feature vectors in the automatic segmentation and classification procedure. In order to determine which provides the parameters/features with the greatest contrast between region types, distribution fits are made to the probability density histograms of both the radar absolute intensity values, $V_{\text{intensity}}$ and the dB-power values, P_{dB}. Weibull and

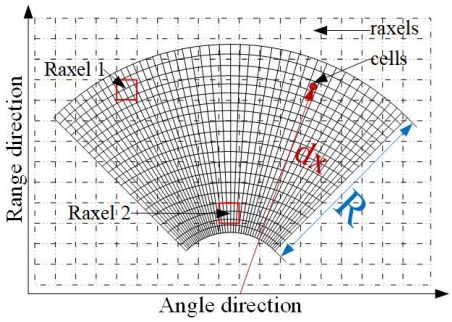


Fig. 6. Relationship between raxels of rasterised radar images and radar map (resolution) cells.

log-Weibull distributions have been widely used in modelling of ground clutter [32], [33]. Both distributions are considered to model radar intensity values, and logarithmic power values. Rayleigh and Rice distributions are considered as reference fits for Gaussian-like clutter and returns consisting of strong components indicating presence of a target. This relationship between Rayleigh and Rician distributions could be mathematically represented as: if the variable $V = \sqrt{X^2 + Y^2}$ where $X \sim N(\upsilon \cos\theta, \sigma^2)$ and $Y \sim N(\upsilon \sin\theta, \sigma^2)$ are independent normal random variables, then $V \sim Rice(|\upsilon|, \sigma)$, and Rayleigh is the special case of Rician distribution that $V \sim Rayleigh(\sigma)$ when $\upsilon = 0$.

The definitions of the PDFs of these four distributions are given as:

$$f_{weibull}(x;\lambda,k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}, \tag{3}$$

$$f_{\text{log-wei}}(x;\lambda,k) = \frac{k}{\lambda}\left(\frac{\ln x}{\lambda}\right)^{k-1} e^{-\left(\frac{\ln x}{\lambda}\right)^k}, \tag{4}$$

$$f_{\text{rayleigh}}(x;c,\sigma) = \frac{(x-c)}{\sigma^2} e^{-(x-c)^2/(2\sigma^2)}, \tag{5}$$

$$f_{\text{rice}}(x;c,\sigma,\nu) = \frac{(x-c)}{\sigma^2} e^{-\frac{(x-c)^2+\nu^2}{2\sigma^2}} I_0\left(\frac{(x-c)\nu}{\sigma^2}\right), \tag{6}$$

where $x \geq 0$, f_{weibull}, $f_{\text{log-wei}}$, f_{rayleigh} and f_{rice} are functions of PDFs of Weibull, log-Weibull, Rayleigh and Rice distributions, respectively; k and λ are the shape and scale parameters of Weibull and log-Weibull distributions; c and σ are the location and scale parameters of Rayleigh and Rice distributions; υ is the shape parameter of Rice distribution which gives the distance between the origin point and the center of the bivariate distribution. $I_0(z)$ represents the Bessel function of the first kind with order zero. Rayleigh distribution with $\lambda_{Rayleigh} = \lambda_{Weibull}/\sqrt{2}$ is the special case of Weibull distribution whose shape parameter $k = 2$.

A. Method to Extract Distribution Features

Initially, we need to extract data from regions where a statistically meaningful assessment can be made.

As the image labelling is performed in raxel space on the raster image, (Fig. 6), the initial task is to map these labelled class regions to their appropriate resolution cell values in the corresponding radar map. The azimuth dimension of each cell depends on its range. Therefore, raxels in the labeled image

at different ranges correspond to a different number of cells in azimuth, e.g. raxel 1 and raxel 2 shown in Fig. 6. A raxel consists of c_{nr}/r_{nr} cells in range direction, where c_{nr} is a number of cells and r_{nr} is the number of raxels within a region of R extent in range. In the case of physical aperture beamforming the relationship between the sizes of raxels and cells in the azimuth direction is a function of distance d_x. In the azimuth direction, the number of cells corresponding to one raxel is, $round\left[\frac{c_{na}R}{2Ac_{nr}d_x\sin(0.5^o)}\right]$, where c_{na} is number of azimuth positions in the FoV of A^o. The analysed radar frames are truncated between ranges of 5-25 m, thus $R = 20$m. The FoV $A = 90°$, $c_{nr} = 668$ and $c_{na} = 199$ ($\approx 0.5°$ azimuth step between range profiles). The rasterised image resolution gives, $r_{nr} = 308$, so one raxel consists of 3×2 cells at the range of 5 m and 3×1 cells at the range of 25 m [31]. In order to simplify the process of cell extraction, we disregard this difference and take the larger 3×2 cells for each raxel whatever the range to the raxel. This guarantees all cell intensities of the RoI are extracted, though some will be extracted multiple times. To remove multiple counts, the co-ordinates of the extracted cells are compared, and repeats are discarded.

It should be noted that the transformation between raxels and cells may be simplified (potentially 1-to-1) if the data was labelled on a rasterized rectangular range-azimuth radar plot (B-scope). The choice to label on the PPI was made due to the direct visual equivalence between this and the ground truth camera imagery, which aids the labeling process.

As discussed in section II-B.2 it is important to investigate variation of statistical characteristics for each class as a function of grazing angle. The grazing angle across the truncated image area varies from 11° to 2.3° and at ranges >25 m the grazing angle variation is insignificant. Due to the significant changes in the grazing angle, range segmentation into regions of 5 m extent will be performed.

Class datasets extracted from all available radar frames are further divided into arrays of size 1000 to produce statistically meaningful histogram plots using a consistent number of intensity values for investigating each class.

B. Distribution Fitting to Region Intensity Statistics

Fig. 7 presents exemplar histograms of I and Q components, and the radar intensity values for the four region classes. Fig. 7(a), (b) and (c), (d) show that the components of the returns from asphalt and shadows area are distributed normally with zero mean, which justify use of Rayleigh and Rice distributions for their $V_{\text{intensity}}$ (Fig. 7 (i), (j)). However, the I and Q components of grass and objects areas shown in Fig. 7 (e), (f) and (g), (h), deviate from normal distribution and the histograms of intensity values in Fig. 7 (k) and (l) are better fit by a Weibull distribution.

The scale ranges of $V_{\text{intensity}}$ (horizontal axes) increase successively from shadows to asphalt to grass, with highest values corresponding to "object" class. The probability density scales (vertical axes) reflect the increasing spread of intensities for these classes with higher median values.

To estimate the intensity histogram fit errors, the normalized root-mean-square-deviation (NRMSD) will be used, which

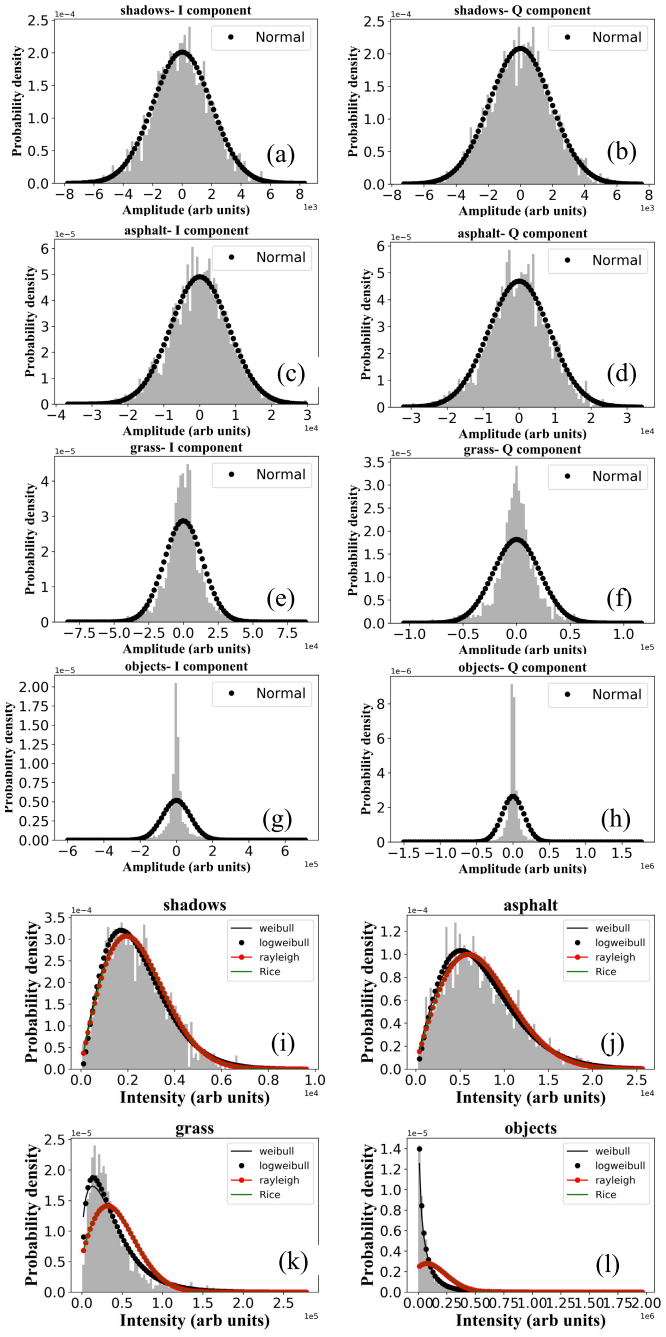


Fig. 7. The probability density histogram plots and distribution fits of I and Q components (a)-(h) and $V_{\text{intensity}}$ (i)-(l), for the four classes.

averages the individual RMSDs over all histogram bins:

$$NRMSD = \frac{1}{N}\sum_{n=1}^{N}\frac{\sqrt{(\upsilon_{PDF}^{(n)} - \upsilon_{hist}^{(n)})^2}}{\left|\upsilon_{hist}^{(n)}\right|} \tag{7}$$

where N is the number of bins, υ_{PDF} and υ_{hist} are the PDF fit value (center of the histogram bin) and the histogram value of each bin respectively.

The NRMSD results are shown in Fig. 8, in which the error bars represent the standard deviation of the fit errors, estimated from fits to all instances of the class regions in the labelled

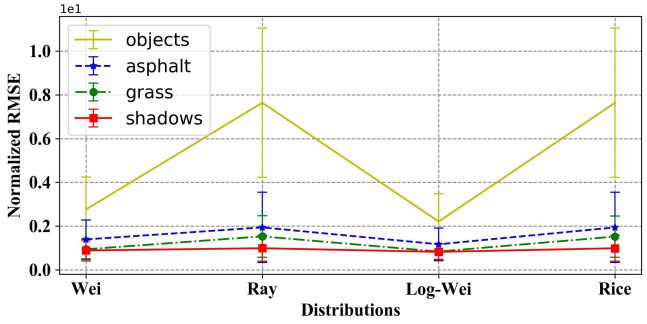


Fig. 8. Results of NRMSDs between PDF fits and the density histograms of intensity values extracted from the four class region types.

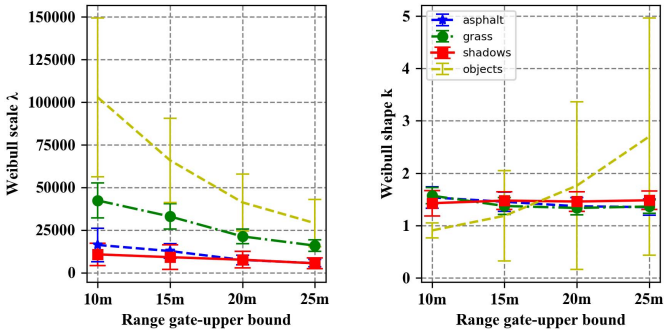


Fig. 9. The distribution parameters obtained from Weibull distribution fit of $V_{\text{intensity}}$ as a function of range gate. Range values indicate the upper value of a range gate.

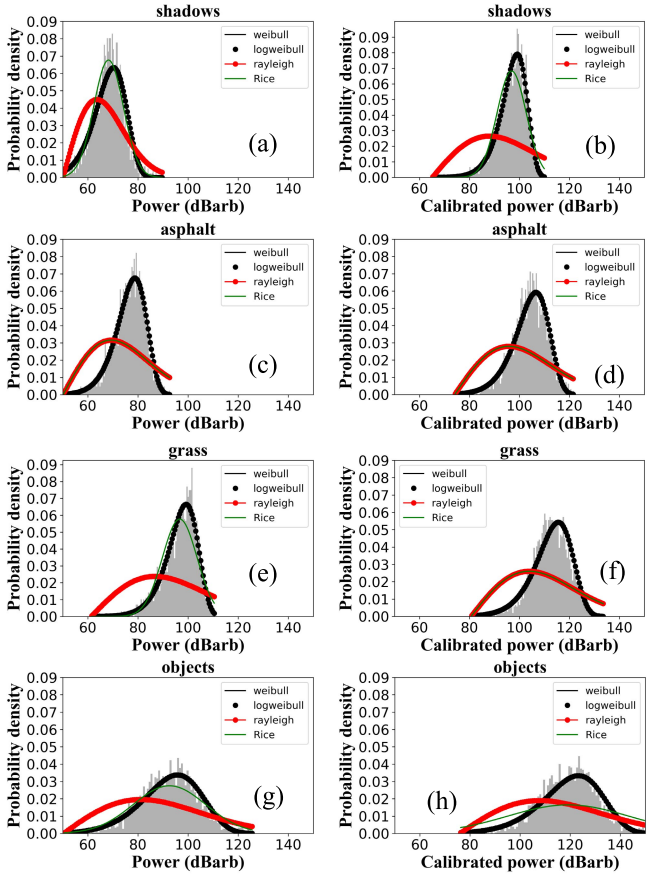


Fig. 10. Density histograms and PDF fits for uncalibrated (left column) and calibrated (right column) radar map dB-power values.

dataset. The result confirms observations in Fig. 7 that Weibull and log-Weibull have smaller fitting error and deviation than Rayleigh and Rician distributions for all four classes. Also the smallest deviation is shown for the shadows and largest for the object class which are intuitively obvious results as the latter is very different from the traditional distributed clutter, while the former should be close to the noise floor of the receiver due to fundamental absence of reflections.

We should stress here, that while finding the 'correct' distribution is important task for characterization of radar returns and essential for understanding of the underlying physical processes, it is not our goal—we want to compare parameters of the same distribution when applied to different classes to determine if there is sufficient parameter contrast to discriminate and classify an RoI within the image. The Weibull fit distribution parameters extracted from $V_{\text{intensity}}$ for the four region types are shown in Fig. 9. Log-Weibull is omitted due to its similar fitting performance, Rayleigh and Rician are omitted due to their poorer fitting performance and thus greater variations in distribution parameters for a given class. Each point represents the mean value of the specific parameter evaluated for all areas of the same class in our dataset within a prescribed range group (discussed in Sections II.2.2 and III.1), the error bar represents the standard deviation. We draw the following conclusions 1) object areas show significant difference when compared to the other classes; 2) most parameters show a monotonic trend of change with increase of range; 3) for all clutter classes, except object, Weibull shape parameter k is close to 2 which shows that conventional clutter returns are close to Rayleigh distribution; 4) There are large overlaps between feature parameters extracted from areas of asphalt, grass and shadows and significant variation in parameter values for the object class. This will result in classification confusion when using distribution parameters derived from the $V_{\text{intensity}}$ data representation.

Next, we will estimate the distribution of class region statistics from both uncalibrated and calibrated dB-power radar maps to determine if this can improve parameter contrast.

C. Distribution Fitting to Uncalibrated and Calibrated Region Power Statistics

Corresponding density histograms and PDF fits for exemplar region dB-power statistics, P_{dB}, are shown in Fig. 10. Shadow has the smallest variation of radar power compared to other classes, as it is defined by the electronic noise of radar transceiver rather than any reflected signal. Comparison of the histograms in Fig. 10 (c)(d) and (e)(f) show that grass area returns follow a slightly wider bell shape distribution than that of asphalt, which can be explained by wider deviations in the height profile of grass within regions. The areas containing target return demonstrate a wide intensity spread and evidence of bimodality. However again let's stress that for the purposes of showing that classes have distinctive differences, we may fit unimodal distributions for all classes. Comparing the uncalibrated and calibrated P_{dB} histograms shows that calibration

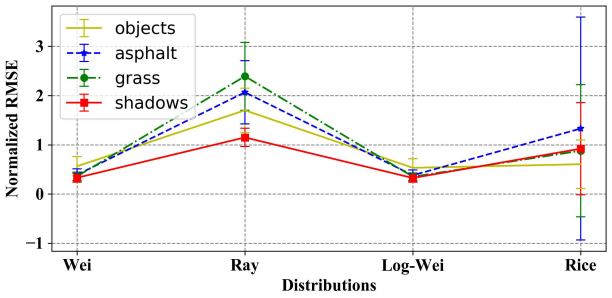


Fig. 11. NRMSEs of PDF fits to the density histograms of uncalibrated radar power map regions for the four classes.

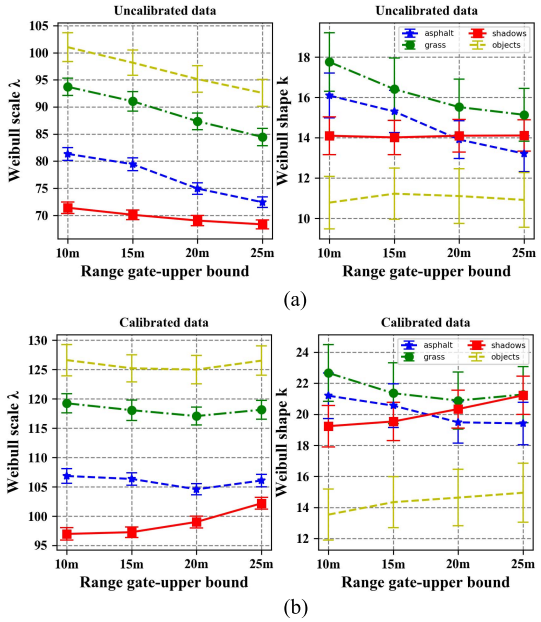


Fig. 12. The parameters obtained from Weibull distribution fit to uncalibrated P_{dB} (a), and calibrated P_{dB}, (b). Range values indicate the upper value of a range gate.

will shift the median and slightly change the distribution scale parameter.

Fig. 11 presents the results of fitting errors estimated based on the P_{dB} of radar clutters. Weibull and log-Weibull distributions show comparable fit errors to the four classes and have smaller fitting errors than Rayleigh and Rician. Thus, Fig. 12 shows the distribution parameters of uncalibrated (a) and calibrated (b) dB-power region statistics resulting from the Weibull fit.

The following conclusions can be drawn:

1) For results of both uncalibrated and range calibrated data, the scale parameter, λ, showed significant contrast between different classes

2) The λ parameter obtained from the uncalibrated P_{dB} decreases with range for all four classes, although for shadows it has distinctively less change than for other classes. Importantly, for classes other than shadow, this decreasing trend is reduced after the "whitening" calibration and therefore we may assume that the same class will have similar statistical parameters across the full range of the calibrated image. By contrast, the calibration leads to an increase in shadow. parameter values which again indicates its essentially different physical nature of shadow—calibration results in an invalid increase of receiver noise floor values, rather than actual calibration of returned powers.

3) The results for Weibull shape parameter, k, mainly stresses the contrast between objects and other classes, with calibration improving the object parameter separation, but reducing that of shadow.

We may conclude that the parameters obtained from the P_{dB} data showed higher contrast than $V_{\text{intensity}}$ and calibrated and uncalibrated data representations both have their own advantages for parameter contrast. Thus, both representations will be tested in discrimination of different classes.

IV. Classification Based on Statistical Distribution Features

Here we propose the supervised classification algorithm based on the MGD model in which the extracted parameters of the Weibull distributions for each class are used as feature parameters. This model has previously been used for unsupervised anomaly detection [34]. At this stage the automotive radar image dataset is separated into training and test data to verify the model performance.

A. Classification Algorithm Based on MGD Model

The general definition of the MGD PDF is denoted as:

$$p(x_t) = \frac{1}{(2n)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t-M)^T \Sigma^{-1}(x_t-M)} \tag{8}$$

where $x_t = [f_1, f_2, \ldots, f_n]$ is a feature vector consisting of n different random variables, M is the vector of their mean values $M = [\mu_1, \mu_2, \ldots, \mu_n]$, Σ is their covariance matrix and $|\Sigma|$ is its determinant. In our case the random variables are the previously identified statistical distribution parameters and will be termed hereon in as feature parameters. To use the MGD formulation for classification we must include a training phase, this is a significantly different methodology than when used solely for anomaly detection. The training stage involves the calculation of one M, and one Σ per class c for the C classes under consideration i.e. M_c and $\sum_c$. This is accomplished using a class training feature parameter set X^c_{tr}, extracted from the training data.

$$X^c_{\text{tr}} = \begin{pmatrix} F^c_{1,1} & \cdots & F^c_{1,n} \\ \vdots & \ddots & \vdots \\ F^c_{m,1} & \cdots & F^c_{m,n} \end{pmatrix} \tag{9}$$

where $F^c_{i,j}$ is an instance of a training feature and $i = \{1 \ldots m\}$ labels the number of samples (observations) of the $j = \{1 \ldots n\}$ feature parameters, i.e. the number of regions of a particular class from which the feature parameters have been extracted. The values μ^c_j in M_c and elements of Σ_c are calculated as follows:

$$\mu^c_j = \frac{1}{m}\sum_{i=1}^{m} F^{\text{c}}_{i,j}$$
$$\Sigma^c_{j,v} = \frac{1}{m}\left(F^{\text{c}}_{i,j} - \mu^c_j\right)\left(F^{\text{c}}_{i,v} - \mu^c_v\right)^{\text{T}} \tag{10}$$

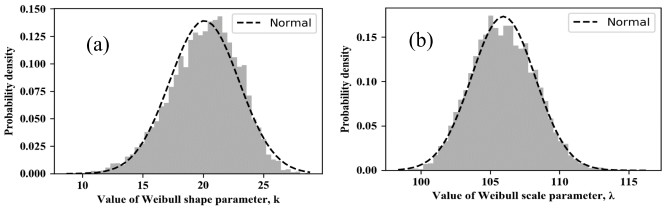


Fig. 13. Probability density histograms of extracted statistical parameters and normal PDF fits for "asphalt" class. (a) Weibull shape parameter λ; (b) Weibull scale parameter *k*.

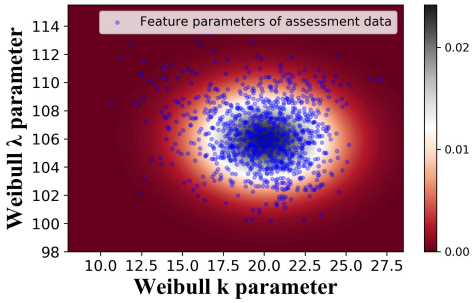


Fig. 14. Comparison between the generated bi-variate (λ and *k* -variates) Gaussian PDF for asphalt (colourmap indicates probability density) and the training feature parameter values (blue markers).

Once MGD's for all classes have been trained, a test vector of feature parameters x_t, from an unknown class region can be evaluated using (8) to obtain a p-value, $p_c\,(x_t)$ from each class MGD. The p-values give the probability that the test set belongs to a particular class, further evaluation of the obtained set of p-values is used to assign a single class type to a region. This involves 2 stages of decision making, which will be described later in this sub-section and expanded upon in Section V.2.

In order for a multi-variate statistic (n -variate in this case) to be Gaussian, each of its n variables should be Gaussian, thus for the MGD to model our feature parameters exactly, the distributions of the parameters should ideally be Gaussian. Examples of the density histograms for Weibull λ and k feature parameters for instances of the asphalt class, are shown in Fig. 13. with corresponding Gaussian fits. Although there are some small observable discrepancies between the data and normal fit, similar to discussion in Section III.2 regarding distribution fits, the MGD framework may still be used and approximates our parameters as Gaussian variables.

An example of the covariance matrix calculated for training feature parameters, λ and k of asphalt areas is shown in (11):

$$\Sigma_{k\lambda} = \begin{bmatrix} 8.22 & -0.21 \\ -0.21 & 5.3 \end{bmatrix} \tag{11}$$

This indicates λ and k have low correlation since the absolute values of non-diagonal elements are close to 0.

Fig. 14 shows the bi-variate Gaussian distribution generated using (8), the covariance matrix in (11) and the corresponding mean values μ_λ and μ_k. The circular shape again highlights the low correlation between features. The training feature parameter values are plotted with blue markers and show good agreement with the generated distribution—this emphasises that the MGD framework approximates our parameter distributions well.

To estimate performance of the MGD classifier, Weibull shape, λ, and scale, k, factors from uncalibrated ("unc") and calibrated ("cal") dB-power data are evaluated as feature parameters in the following combinations: i) $x_t^{(1)} = [w_\lambda^{unc}, w_k^{unc}]$; ii) $x_t^{(2)} = [w_\lambda^{cal}, w_k^{cal}]$; iii) $x_t^{(3)} = [w_\lambda^{unc}, w_k^{unc}, w_\lambda^{cal}, w_k^{cal}]$; the effect of training sample size on performance is also examined.

The full MGD model, combining the individual MGD's for each class, can be represented as:

$$p_{MGD}(x_t) = [p_1(x_t, \Sigma_1, M_1), \ldots, p_c(x_t, \Sigma_c, M_c)] \tag{12}$$

Test RoI's are selected from the labeled data. Each RoI is further divided into equal sample sub-regions. Feature parameters are extracted from each sub-RoI and form the input to the trained MGD model. The class with the highest p-value, $p_c(x_t, \Sigma_c, M_c)$, identifies a sub-RoI as belonging to that class—this is the first stage of classification. Thus, within each initial RoI we may have many sub-RoI's labelled differently according to their individual result of classification. In the second stage of classification, the whole RoI is labelled as the class represented by the largest number of sub-RoI's. In general, more complex metrics can be used for the voting of the class at the second stage of classification.

B. Estimation of Classification Performance

We use the F1-score [35] as a general assessment tool for estimating the performance of classification on the test dataset. The definition of the F1-score calculation is based on the precision and recall estimation of the classification results:

$$F1 = \frac{2 * precision * recall}{precision + recall}, \tag{13}$$

in which, recall and precision are:

$$precision = \frac{T_p}{T_p + F_p}; recall = \frac{T_p}{T_p + F_n}, \tag{14}$$

where T_p, F_p and F_n are the number of true positives, false positives and false negatives obtained from the confusion matrix results. We will evaluate the impacts of (i) the calibration of radar data and (ii) the size of the training dataset on performance of classification, the F1-score results are shown in Fig. 15 and 16. In Fig. 15, the F1-score at each classification stage are plotted for each class presented on x-axis. The results for the three types of feature parameter compositions, $x_t^{(1)}$, $x_t^{(2)}$ and $x_t^{(3)}$ described in Section IV.A are represented using different line types, and results of first and second stage of classification are shown by different markers. The results of first stage of classification showed that F1-scores of asphalt and grass areas are effectively improved by use of features from calibrated data, either solely or in combination with uncalibrated. By contrast, the F1-score of shadows showed the opposite effect where use of $x_t^{(1)}$ results in better performance than that of $x_t^{(2)}$, however the combination of both, $x_t^{(3)}$, showed the best F1-score. These conclusions are consistent with the results and discussion in section III.3, where we stressed the

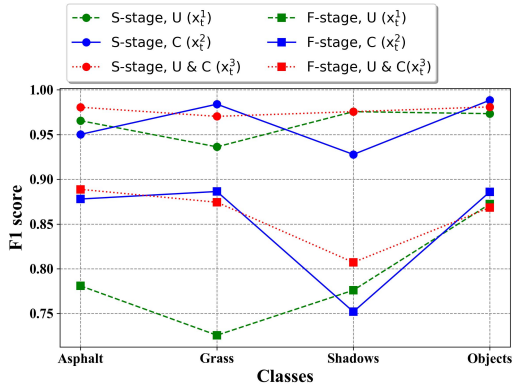


Fig. 15. F1-score results obtained from the classification algorithm. Legend labels F and S indicate results of first and second stages of classification and U and C correspond to the use of uncalibrated and calibrated power data respectively.

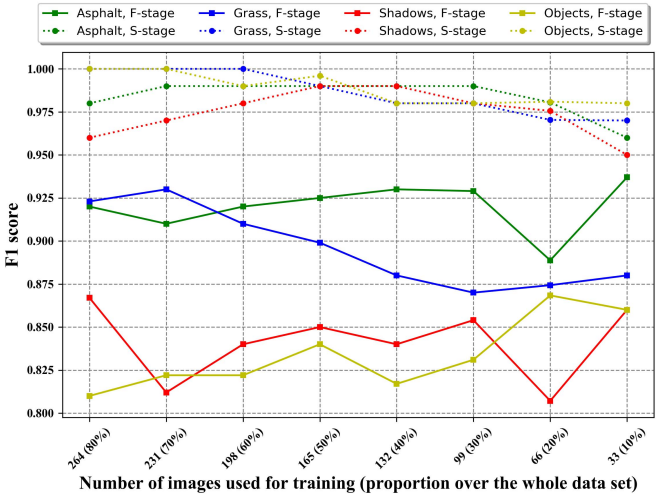


Fig. 16. F1-score results of classification algorithm for varied amounts of training data, using combined features case $x_t^{(3)}$.

physically different mechanisms behind the clutter returns and shadows and the effect of calibration the distribution/ feature parameters. The F1 score results of targets showed similar performance for all $x_t^{(1)}$, $x_t^{(2)}$ and $x_t^{(3)}$ since feature parameters obtained from uncalibrated and calibrated data all demonstrate high contrast with respect to other classes. For the results of second stage of classification, the utilization of features from both uncalibrated and calibrated dB-power data improved the F1-score performance, especially for the asphalt, grass and shadow classes.

In Fig. 16, F1-scores for the combined feature case $x_t^{(3)}$ are plotted as function of the size of the training dataset, which was varied from 80 % of the available dataset (330 frames) to 10 %. Results for different areas are presented using different color. The results showed no obvious monotonic trend for different amounts of training data, however the F1 score of the second stage of classification did drop when using only a small fraction for training. Therefore, this classification algorithm does not significantly depend on the amount of training data, which infers a high consistency of statistical parameters over all instances of class regions in the dataset. The comparison between results of first and second-stage of

TABLE II
CONFUSION MATRIX AFTER FIRST STAGE OF CLASSIFICATION – COMBINED CALIBRATED/UNCALIBRATED FEATURES

		Predicted			
		Asphalt	Grass	Shadows	Objects
Actual	Asphalt	7020	121	1389	37
	Grass	97	3510	0	472
	Shadows	112	0	3149	4
	Objects	1	319	0	2748
Recall		0.82	0.86	0.96	0.90
Precision		0.97	0.89	0.69	0.84
F1 score		0.89	0.87	0.81	0.87

TABLE III
CONFUSION MATRIX AFTER SECOND STAGE OF CLASSIFICATION- COMBINED CALIBRATED/UNCALIBRATED FEATURES

		Predicted			
		Asphalt	Grass	Shadows	Objects
Actual	Asphalt	252	0	7	0
	Grass	1	180	0	5
	Shadows	2	0	180	0
	Objects	0	5	0	257
Recall		0.97	0.97	0.99	0.98
Precision		0.99	0.97	0.96	0.98
F1 score		0.98	0.97	0.98	0.98

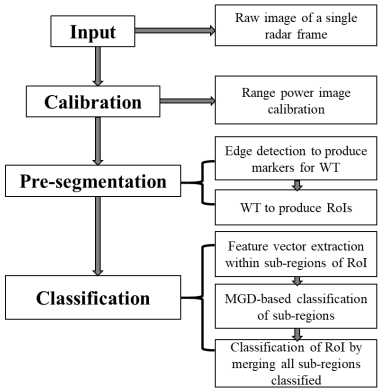


Fig. 17. Block diagram of the automatic segmentation on automotive radar image based on distribution feature extraction and MGD classification model.

classification showed that classification performance is significantly improved in the second stage. This is also highlighted in the corresponding confusion matrices in Tables II and III. The F1-scores of the first stage of classification are not higher than 0.825, and that of the second stage are above 0.95.

V. AUTOMATIC SEGMENTATION OF RADAR IMAGERY

Till this point, we have shown the development and performance of the segmentation algorithm on fully labelled data. Now we move to the discussion of automatic segmentation and classification of a single test frame. The processing flowchart is shown in Fig. 17 and processing steps will be explained in the following sub-sections.

A. Image Pre-Segmentation Using the Watershed Transform

An image pre-segmentation stage is used to define the RoI's within which the classification and region merging stages are applied.

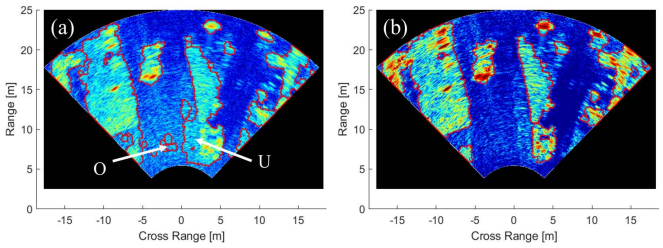


Fig. 18. The pre-segmentation results obtained using (a) uncalibrated and (b) calibrated radar images after single application of WT. 'O' indicates an over-segmented region, 'U' an under-segmented one.

The most widely used methods of image segmentation based on RGB information are thresholding [36], edge detection [37], region growing [38] and WT method [39]. In this paper, the marker-based WT method is used to generate the closed contours of the regions in the radar imagery, which will be subject to classification. The principle of the WT in OpenCV is given in [49] and full description of the method can be found in computer vision literature, e.g. [40]–[42]. Here we will only briefly describe main steps undertaken.

First, a greyscale image generated from the radar image is binarised using a threshold on raxel values. Markers for the WT are produced by subtracting the erosion binary image from the dilated binary image [49], producing a set of broken edges. These are then connected in the WT process to form the contours of sub-regions within the image, generating what we term a pre-segmented image. Results of this process using the uncalibrated and calibrated radar imagery are presented in Fig. 18—the extracted region contours are superimposed over intensity maps. It clearly demonstrates more robust performance of the method on the calibrated image (b), with the uncalibrated image (a) displaying both under-segmented regions containing mixed classes and over-segmented areas.

A single WT on a calibrated image is enough to produce region pre-segmentation, uncalibrated imagery would require multiple applications of the WT—this increases the algorithm complexity and computation time.

B. Region Merging Using MGD-Based Classification Method

After pre-segmentation, which may result in potentially over-segmented RoIs, the identification and merging of similar regions is performed within the classification procedure described in Section IV.A. The input is the pre-segmented test automotive radar image which may include RoIs which do not necessarily correspond to any of the defined classes. Therefore, to stress that for some regions the confidence in finding the correct class is low, another class "unknown" is introduced. This is a class containing the regions which show no obvious bias to any of the other classes in one of two distinct ways, described below.

Firstly, the output p-values of the MGD model evaluated for specific RoI sub-regions are normalized using the softmax function defined as:

$$p_{soft,c} = \frac{e^{p_c}}{\sum_{c=1}^{C} e^{p_c}} \tag{15}$$

TABLE IV
EXAMPLE OF ORIGINAL p-VALUES AND SOFTMAX NORMALIZED VALUES AND THEIR CLASS DETERMINATION

Case	Original class p -values p_c				Softmax p -values $p_{soft,c}$				Class
	p_a	p_g	p_s	p_o	$p_{soft,a}$	$p_{soft,g}$	$p_{soft,s}$	$p_{soft,o}$	
1	2e4	5e-3	2e-33	1e-58	1	0	0	0	Asph
2	0.4	5e-9	2e-3	5e-6	0.33	0.22	0.22	0.22	Asph
3	2e-4	1e-19	4e-40	1e-6	0.25	0.249	0.249	0.249	Unkno wn

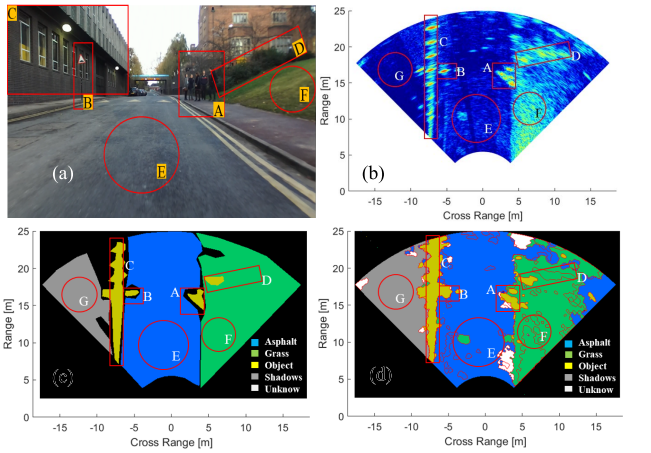


Fig. 19. (a) ground truth optical image; (b) calibrated automotive radar image; the manually labeled image (c) and automatically segmented radar image (d).

in which p_c represents the original p-values obtained from the MGD model and to reiterate, C is the number of classes (4 in this case). The softmax function allows elimination of large differences between the original p values. The "unknown" label will be assigned in the first stage of classification to the sub-ROI's where all $p_{soft,c}$ values are smaller than a threshold t_u. The sub-RoI's for which at least one $p_{soft,c} > t_u$ will be classified according to the highest $p_{soft,c}$ value. In this work we used a threshold $t_u = (1/C) + \varepsilon$ where ε is a small value, in our case 0.01, though this is a subject for further optimization. Table IV gives examples of areas which are classified as either "known" classes (case 1 and case 2) or "unknown" (case 3).

Secondly, the ROI will be classified as an "unknown" area if the highest score in the second stage of classification is not unique.

It should be stressed here that the pre-segmented image may consist of over-segmented areas which could be too small to extract a statistically meaningful number of power values, P_{dB}, upon which to perform a distribution fit and extract parameters. These small areas are extended to a square region of raxels in the size of 30×30 around the center point. Then the feature extraction will be performed within the extended area.

The raxels of the classified RoI will be labeled and colour-coded according to the assigned class, this operation will then be repeated for the next RoI.

C. Results of Automatic Segmentation of Automotive Radar Images

An example of an automatic segmentation result is illustrated in Fig. 19, which shows the annotated ground truth

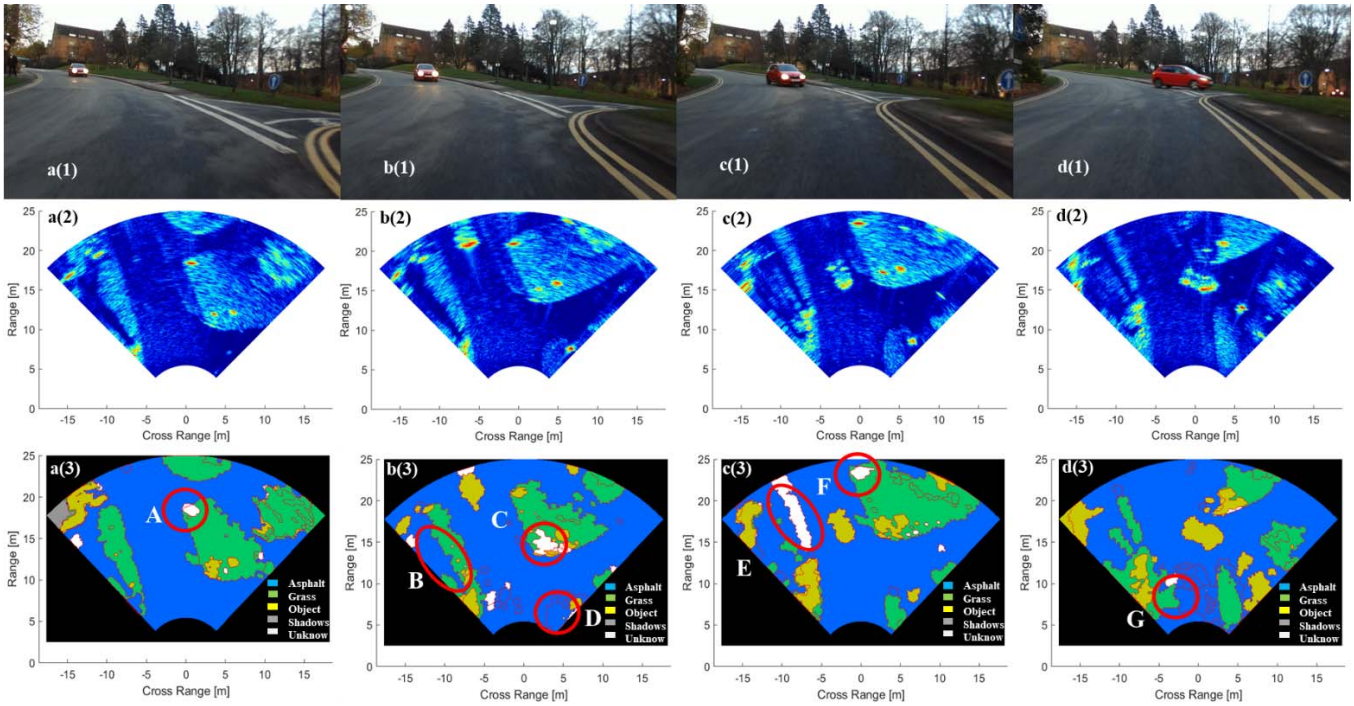


Fig. 20. The automatically segmented results of four consecutive frames of automotive radar imagery. a-d(1) are the optical images; a-d(2) are the calibrated automotive radar images and a-d(3) are the results of automatic segmentation. Region colour coding equivalent to Fig. 19.

image of the scene (a), annotated calibrated radar frame (b), manually labelled image (c) and the automatically segmented and labelled image in (d).

Annotations in the Figure are: A, B, C, D are pedestrians, traffic sign, wall, and bush, respectively and all are of the class 'object'. E is the road surface (class asphalt), F is grassed area (class "grass") and G is a region of shadow of the wall (class shadow), white represents unclassified regions. Square bounding boxes are used for annotation of objects and circles identify surfaces.

Comparing Fig. 19 (c) and (d) one can see that most of the objects and surfaces in the labeled image (c) are classified correctly in the full segmented image (d). Unclassified regions are present e.g. a region of pavement near the annotation E. Pavement regions are currently included in the asphalt class, in (b) however there is noticeable difference in the regions surface returns in contrast to its asphalt surroundings. This may just be an anomalously rough patch of surface, inclusion of a 'pavement' class may remedy the confusion in the classification. Another important feature is the region classified as a patch of grass (within circle E) within the asphalt road surface, this is a manhole cover. Importantly, it has been classified as a surface type feature rather than object and so could still be traversed, but the contrast can warn a vehicle to prepare for a transition. It may also be seen as an anomalous region within the large region of asphalt road and contextually this may aid identification—again additional classes may also be added.

The automatic segmentation results of 4 successive frames of a radar "movie" (radar snapshots) are shown in the bottom row of Fig. 20. Corresponding optical ground truth snapshots and calibrated automotive radar images are on the top and middle row respectively. The imaged scenes are mainly composed of areas of road tarmac, grass, traffic signs and trees on the grass area. The segmented images show that the majority of regions such as the areas of asphalt, grass and objects are properly segmented using the WT method and correctly classified based on the MGD model. The moving vehicle in different frames shown passing through the scene is also identified correctly with well-defined shape.

There are some incorrect classifications of areas. This may be caused by the over-merging of regions in the pre-segmentation. For example, area B in Fig. 20 b (3) shows that the shadow area has merged with the larger asphalt area and is incorrectly classified as asphalt. Similarly area G in Fig. 20 d(3) shows that the asphalt or kerbside merged with the lawn area and is incorrectly classified as grass area. Some regions have been classed as "unknown" areas, such as A in a(3), C in b(3) and F in c(3), which are primarily object regions containing lampposts and signage, and E, an area of grass in c(3). However, each of these unclassified regions are in fact correctly classified in at least one of the 4 segmented images. In the next stages of research, we plan to improve confidence of true positive classification by using frame-to-frame region tracking and associations within a succession of frames.

The numerical estimation of the accuracy of the automatic segmentation and classification is conducted based on a test data set consisting of 120 continuous image frames and utilises the Jaccard similarity coefficient (JSC) [43]. In the performance analysis we exclude segments which, though classified as one of the known classes in automatic segmentation, cannot be confirmed by ground truth and were marked as black areas

TABLE V
THE JSCS OF THE AUTOMATICALLY SEGMENTED RADAR IMAGES AND THE LABELED IMAGES

Areas	Asphalt	Grass	Shadows	Objects
Average JSC	0.81	0.64	0.79	0.64

in the manually labelled images. The JSC for each class is defined as:

$$J_{class} = \frac{A_{fs} \cap A_{label}}{A_{label}}, \tag{16}$$

where A_{fs} is the number of correctly classified raxels overlapping with the labelled data regions of that class and A_{label} is the total number of raxels of corresponding class in the labeled data.

The averages of the JSCs of the four area classes are presented in Table V. Areas of asphalt have the highest JSC of 0.81 and the areas of grass and object show the lowest value of 0.64.

The JSC's for the single frame image segmentation of a relatively complex scene are encouraging. As mentioned previously, improvement will be made by considering results of consecutive frames of the radar movies.

VI. CONCLUSION

An algorithm for automatic segmentation and classification of automotive high-resolution radar images is presented in this paper. The two main steps in this algorithm are: 1) initial image pre-segmentation using WT method; 2) supervised region classification into chosen classes, utilizing statistical parameters of radar image regions as feature parameters in a proposed MGD classifier. The features were based on Weibull distribution parameters extracted from image data represented in dB-power units, this showed better contrast between features/classes than when considering radar intensity values.

As performance metrics, the F1 score and JSC have been used to assess the results of classification/segmentation. They showed good performance of the proposed algorithm to reconstruct the content of single (standalone) radar image frames.

It was shown that the proposed two stage classification process can significantly improve the accuracy of the automatic classification.

A detailed range-power calibration process has been described and shown to vitally enhance the image pre-segmentation process when compared to using uncalibrated data. Segmentation/classification depends highly on the result of pre-segmentation by WT and this should be subject of further studies to determine the feedback mechanism to compensate the occasional inaccuracies of initial segmentation.

It was demonstrated that the calibration tends to reduce the variation of the feature parameters over the extent of the radar imagery. In general the Weibull statistical (feature) parameters for dB-power radar imagery show good separation between classes, the exception to this is the shadow class which has overlap with other classes in the scale factor k parameter, this motivates the inclusion of uncalibrated data features in order to improve classification of all classes.

Further work will also include context-based analysis to enhance classification, for example: anomalous regions within large extents of a particular class.

Finally, we will investigate the full use of multiple consecutive frames of radar movie to improve confidence of classification.

REFERENCES

[1] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 22–35, Mar. 2017.

[2] W. J. Fleming, "New automotive sensors—A review," *IEEE Sensors J.*, vol. 8, no. 11, pp. 1900–1921, Nov. 2008.

[3] L. Daniel, D. Phippen, E. Hoare, M. Cherniakov, and M. Gashinova, "Image segmentation in real aperture low-THz radar images," in *Proc. 20th Int. Radar Symp. (IRS)*, Jun. 2019, pp. 1–8.

[4] B. Scheuchl, D. Flett, R. Caves, and I. Cumming, "Potential of RADARSAT-2 data for operational sea ice monitoring," *Can. J. Remote Sens.*, vol. 30, no. 3, pp. 448–461, Jan. 2004.

[5] N. W. Park and K. H. Chi, "Integration of multitemporal/polarization C-band SAR data sets for land-cover classification," *Int. J. Remote Sens.*, vol. 29, no. 16, pp. 4667–4688, Aug. 2008.

[6] S. C. Steele-Dunne, H. McNairn, A. Monsivais-Huertero, J. Judge, P.-W. Liu, and K. Papathanassiou, "Radar remote sensing of agricultural canopies: A review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 2249–2273, May 2017.

[7] M. M. Trivedi and J. C. Bezdek, "Low-level segmentation of aerial images with fuzzy clustering," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, no. 4, pp. 589–598, Jul. 1986.

[8] R. Krishnapuram and J. M. Keller, "The possibilistic c-means algorithm: Insights and recommendations," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 3, pp. 385–393, Aug. 1996.

[9] H. Caillol, A. Hillion, and W. Pieczynski, "Fuzzy random fields and unsupervised image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 801–810, Jul. 1993.

[10] J. Fan, M. Han, and J. Wang, "Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation," *Pattern Recognit.*, vol. 42, no. 11, pp. 2527–2540, Nov. 2009.

[11] T. Zuva and O. O. Olugbara, "Image segmentation, available techniques, developments and open issues," *Can. J. Image Process. Comput. Vis.*, vol. 2, no. 3, pp. 20–29, 2011.

[12] M. Wang and J. Wang, "A region-line primitive association framework for object-based remote sensing image analysis," *Photogramm. Eng. Remote Sens.*, vol. 82, no. 2, pp. 149–159, Feb. 2016.

[13] D. Li, G. Zhang, Z. Wu, and L. Yi, "An edge embedded marker-based watershed algorithm for high spatial resolution remote sensing image segmentation," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2781–2787, Oct. 2010.

[14] P. Li, J. Guo, B. Song, and X. Xiao, "A multilevel hierarchical image segmentation method for urban impervious surface mapping using very high resolution imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 1, pp. 103–116, Mar. 2011.

[15] M. D. Hossain and D. Chen, "Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," *ISPRS J. Photogramm. Remote Sens.*, vol. 150, pp. 115–134, Apr. 2019.

[16] P. Yu, A. K. Qin, and D. A. Clausi, "Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1302–1317, Apr. 2012.

[17] F. Cheevasuvit, H. Maitre, and D. Vidal-Madjar, "A robust method for picture segmentation based on a split-and-merge procedure," *Comput. Vis., Graph., Image Process.*, vol. 34, no. 3, pp. 268–281, Jun. 1986.

[18] F. Galland, J.-M. Nicolas, H. Sportouche, M. Roche, F. Tupin, and P. Refregier, "Unsupervised synthetic aperture radar image segmentation using Fisher distributions," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2966–2972, Aug. 2009.

[19] J. Chen, M. Deng, X. Mei, T. Chen, Q. Shao, and L. Hong, "Optimal segmentation of a high-resolution remote-sensing image guided by area and boundary," *Int. J. Remote Sens.*, vol. 35, no. 19, pp. 6914–6939, Oct. 2014.

[20] M. N. Kurnaz, Z. Dokur, and T. Ölmez, "Segmentation of remote-sensing images by incremental neural network," *Pattern Recognit. Lett.*, vol. 26, no. 8, pp. 1096–1104, Jun. 2005.

[21] P. Mitra, B. U. Shankar, and S. K. Pal, "Segmentation of multispectral remote sensing images using active support vector machines," *Pattern Recognit. Lett.*, vol. 25, no. 9, pp. 1067–1074, Jul. 2004.
[22] M. Siam, S. Elkerdawy, M. Jagersand, and S. Yogamani, "Deep semantic segmentation for automated driving: Taxonomy, roadmap and challenges," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–8.
[23] *StereoLabs*. Accessed: Dec. 12, 2020. [Online]. Available: https://www.stereolabs.com/zed/
[24] *Automotive Radar Image Dataset*. Accessed: Dec. 12, 2020. [Online]. Available: https://edata.bham.ac.uk/329/
[25] F. T. Ulaby, *Microwave Radar and Radiometric Remote Sensing*. Ann Arbor, MI, USA: Univ. Michigan Press, 2014.
[26] F. E. Nathanson, J. P. Reilly, and M. N. Cohen, "Radar design principles-signal processing and the environment," in *Proc. STIA*, vol. 91, 1991, p. 46747.
[27] A. K. Fung, *Microwave Scattering and Emission Models and Their Applications*. London, U.K.: Artech House, 1994.
[28] E. S. Li and K. Sarabandi, "Low grazing incidence millimeter-wave scattering models and measurements for various road surfaces," *IEEE Trans. Antennas Propag.*, vol. 47, no. 5, pp. 851–861, May 1999.
[29] *Liblabel*. Accessed: Dec. 12, 2020. [Online]. Available: http://www.cvlibs.net/software/liblabel/
[30] A. Geiger, M. Lauer, C. Wojek, C. Stiller, and R. Urtasun, "3D traffic scene understanding from movable platforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 1012–1025, May 2014.
[31] Y. Xiao, L. Daniel, and M. Gashinova, "Feature-based classification for image segmentation in automotive radar based on statistical distribution analysis," presented at the IEEE Radar Conf. (RadarConf), Florence, Italy, Sep. 2020.
[32] M. Sekine *et al.*, "Weibull-distributed ground clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 4, pp. 596–598, Jul. 1981.
[33] R. Vicen-Bueno, M. Rosa-Zurera, M. Jarabo-Amores, and R. Gil-Pita, "Automatic target detection in simulated ground clutter (Weibull distributed) by multilayer perceptrons in a low-resolution coherent radar," *IET Radar, Sonar Navigat.*, vol. 4, no. 2, pp. 315–328, Apr. 2010.
[34] S. Matteoli, M. Diani, and J. Theiler, "An overview of background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2317–2336, Jun. 2014.
[35] M. Kampffmeyer, A.-B. Salberg, and R. Jenssen, "Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2016, pp. 1–9.
[36] T. Kurita, N. Otsu, and N. Abdelmalek, "Maximum likelihood thresholding based on population mixture models," *Pattern Recognit.*, vol. 25, no. 10, pp. 1231–1240, Oct. 1992.
[37] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
[38] A. Tremeau and N. Borel, "A region growing and merging algorithm to color segmentation," *Pattern Recognit.*, vol. 30, no. 7, pp. 1191–1203, Jul. 1997.
[39] G. Hamarneh and X. Li, "Watershed segmentation using prior shape and appearance knowledge," *Image Vis. Comput.*, vol. 27, nos. 1–2, pp. 59–68, Jan. 2009.
[40] Y. Dong, B. Forster, and A. Milne, "Comparison of radar image segmentation by Gaussian- and Gamma-Markov random field models," *Int. J. Remote Sens.*, vol. 24, no. 4, pp. 711–722, 2003.
[41] G. Sheng, W. Yang, X. Deng, C. He, Y. Cao, and H. Sun, "Coastline detection in synthetic aperture radar (SAR) images by integrating watershed transformation and controllable gradient vector flow (GVF) snake model," *IEEE J. Ocean. Eng.*, vol. 37, no. 3, pp. 375–383, Jul. 2012.
[42] T. B. Ijitona, J. Ren, and P. B. Hwang, "SAR sea ice image segmentation using watershed with intensity-based region merging," in *Proc. IEEE Int. Conf. Comput. Inf. Technol.*, Sep. 2014, pp. 168–172.
[43] S. Niwattanakul, J. Singthongchai, E. Naenudorn, and S. Wanapu, "Using of Jaccard coefficient for keywords similarity," in *Proc. Int. Multiconf. Eng. Comput. Sci.*, 2013, vol. 1, no. 6, pp. 380–384.

Yang Xiao received the M.Eng. degree in electronic and communication engineering from Tianjin University, China, in 2016. She is currently pursuing the Ph.D. degree in low-THz automotive radar with the Microwave Integrated Systems Laboratory, University of Birmingham. Her research interests include the automatic segmentation of radar images using statistical feature extraction and deep learning in automotive radar applications, theoretical modeling of propagation properties of low THz signal, and practical experimental work on low-THz radar systems.

Liam Daniel (Member, IEEE) received the M.Sci. degree in theoretical physics from the University of Birmingham, U.K., in 2005, and the Ph.D. degree from the Microwave Integrated Systems Laboratory, School of Engineering, Birmingham. He is a Research Fellow of the University of Birmingham. His research interests include forward scatter radar, passive radar and automotive sensing covering experimental and system design aspects, simulation, and signal processing. His current research interest includes the utilization of low THz imaging radars for autonomous/semi-autonomous platform applications.

Marina Gashinova received the M.Sc. degree in math from St-Petersburg State University in 1991 and the Ph.D. degree in physics and math from St-Petersburg Electrotechnical University, Russia, in 2003. In 2006, she joined the Microwave Integrated System Laboratory (MISL), University of Birmingham (UoB). She is currently the Head of the Pervasive Sensing Group, MISL, leading a number of research projects on automotive sensing and THz sensing. She is also the Chair of Pervasive Sensing with the University of Birmingham. Her area of interests include active and passive radar, imaging THz sensing for ground and spaceborne scene assessments, and cognitive radar, as well as deep neural networks classification.