The End-to-End Segmentation on Automotive Radar Imagery

Conference Paper · April 2022

DOI: 10.23919/EuRAD50154.2022.9784524

3 authors, including:

Yang Xiao
University of Birmingham
5 PUBLICATIONS **54** CITATIONS
SEE PROFILE

Marina Gashinova
University of Birmingham
217 PUBLICATIONS **2,345** CITATIONS
SEE PROFILE

The End-to-End Segmentation on Automotive Radar Imagery

Yang Xiao, Liam Daniel, Marina Gashinova

MISL Lab, University of Birmingham, UK

YXX752@student.bham.ac.uk

***Abstract* — Segmentation and classification of surfaces and objects in automotive radar imagery are key techniques to identify the passable region for path planning in autonomous driving. The end-to-end segmentation on automotive radar imagery is proposed in this paper, where the input B-scope automotive radar map is processed to output the segmented radar map with labeled area classes. The algorithm discussed in this paper is the extension of our previous published work [1], where we proposed two-stage segmentation processed including (i) pre-segmentation using watershed transformation (WT), and (ii) the region classification based on the Multivariate Gaussian Distribution (MGD) classifier and the extracted distribution features. In the current paper, we use the B-scope radar map representation to simplify the coordinate transformation procedure as compared to PPI image representation. Secondly to improve classification of low-return regions two-tier segmentation process is introduced, where after the first classification of regions of high return, the more subtle classification is made between classes of low returns. Radar test dataset, collected in outdoor driving scenarios and labeled according to optical ground truth is used for assessing the Jaccard similarity coefficient (JSC) performance of segmentation results, which show higher accuracy of classification than in our previous algorithm [1].**

***Keywords* — Automotive sensor, radar imaging, image segmentation, surface classification.**

I. INTRODUCTION

Path planning and obstacle avoidance are the key tasks for achieving autonomous driving [2]. One of the approaches to identify the passible and impassable region from the sensed imagery is to segment the image and classify the surfaces and objects. Image segmentation technique has already been widely applied in optical imagery [3]. However, until recently, image segmentation has not been considered as a suitable technique in automotive radar as the resolution of conventional automotive radar systems could not provide the image format similar to optical imagery. The utilization of mm-wave and sub-THz radar allows both high resolution and high sensitivity to texture, providing the necessary contrast to distinguish between different regions and will therefore facilitate full deep scene reconstruction in future automotive imaging sensors.

The technique of image segmentation has been developed in remote sensing for the processing of SAR imagery [4]. The general methods of image segmentation on SAR imagery includes: 1) the edge-based segmentation such as fuzzy model [5, 6] and watershed transformation (WT) [7]; 2) the region-based methods such as the region merging, region growing and region splitting [8]; 3) hybrid method which combine edge-based and region-based methods; 4) semantic methods such as support vector machine and Convolution Neural Network [9].

For the automotive imagery, which has significant differences compared to long-range large-scale scene mapping of remote sensing, we utilize hybrid method, which includes the edge-based segmentation using WT, and the region-based classification based on distribution features of radar returns from regions of different class, after calibration to account for changing grazing angle of scene map in automotive scenario. The algorithm discussed in our previous work [1, 10] utilizes PPI representation of input radar map for pre-segmentation. The complicated coordinate transformation procedure between 'raxels' of PPI map and radar cells is involved to compensate the size difference of radar cells over ranges.

In this paper, B-scope radar map representation is used for pre-segmentation to simplify the coordinate transformation and improve the performance of end-to-end segmentation as originally proposed in [1, 10]. The JSC is used as a metric of performance of the approach and comparison of JSC values of new proposed approach and that of published in [1] showed that significant improvement on segmentation of regions labelled according to characteristics of 4 general classes is achieved. This achievement is accomplished through the utilization of B-scope radar map, as well as two-tier segmentation process where better classification is made for high return regions initially.

II. AUTOMOTIVE RADAR IMAGE DATASET

In this section, we describe the automotive radar image dataset utilized in this paper. The data collection is conducted based on the high-resolution 79 GHz FMCW real aperture radar, designed by the University of Birmingham and ELVA-1 [11]. The segmentation algorithm introduced in [1] includes the pre-segmentation based on WT and the region classification based on the distribution feature of radar return power. These two steps utilize two kinds of information, where radar imagery is treated as both RGB color scheme map, to allow application of algorithms used in computer vision sampled by pixels and the radar return power map, which carries physical information on scattering properties of regions positioned in range and azimuth w.r.t radar, where sampling is defined by radar resolution cells. The appropriate mapping between pixels and radar cells needs to be developed to connect these two image representation formats and achieve the end-to-end automatic segmentation.

PPI map and B-scope map are two different coordinate scheme representations of radar map. The segmentation

algorithm in [1] has utilized the PPI radar map for WT pre-segmentation. A PPI radar map is presented in Fig. 1 (a) as an example, which is a polar coordinate map of the area sensed by the radar system. The advantage of using the PPI data format is to provide a better view of the position and shape information of objects. The radar raster image pixels of the PPI map with three color channels were defined as 'raxels' in [1]. After initial pre-segmentation with RoIs defined by raxels, the classification of segmented regions was achieved by extracting distribution features of corresponding intensity values of equivalent regions mapped in terms of radar cells. Since the size of radar cells is range-dependent, the complicated coordinate transformation between raxels of PPI map and radar cells was developed to compensate the size difference.

B-scope map, which will be used in this paper, is a Cartesian representation of mapped scene as shown in Fig. 1(b), where the size of radar cells is uniform across the range. The pixels of B-scope map with three color channels are equivalent of radar cells. The axis information is removed from B-scope map to keep the same size.

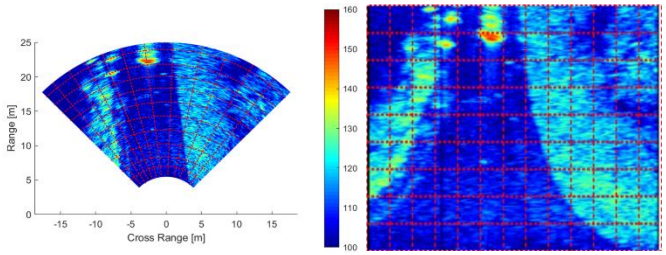

Fig. 1 (a) The PPI radar map; (b) The B-scope radar map. The radar cells are represented using red dashed lines.

The calibrated and uncalibrated radar maps are utilized for feature extraction and region classification as discussed in [1]. In the current paper both the B-scope and PPI radar map are based on the calibrated return power values. The data labelling is required for feature extraction of training dataset [1] and labelled PPI maps in training dataset can be used for both approaches (PPI and B-scope) for the scene under test.

III. The End-to-End Segmentation on B-scope Radar Map

The end-to-end segmentation is the sequence of processing steps, where initially regions of interest (RoI) are discerned by combining pixels into super-pixels, based on the similarity of their characteristics, and finally identified regions are labelled and classified based on comparison with training dataset. The block diagram of the end-to-end segmentation algorithm is presented in Fig. 2. The input is B-scope raw radar map, the output is the segmented B-scope map with the regions labelled according to defined classes. The highlight parts are the main improved parts of segmentation algorithm compared to our previous algorithm in [1].

A. The Pre-segmentation of B-scope Radar Map

Initially the RoIs are discerned in B-scope map by using WT in OpenCV [12]: surface classes – tarmac, grass, which are defined as passable regions, but requires adjustment of car control when traversing from one into other, and obstacles classes: object, shadow. Unclassified regions will be assigned to class "unknown". The result of such pre-segmentation is shown in Fig. 3 (a). The further steps of two-tier of classification will be introduced in next section.

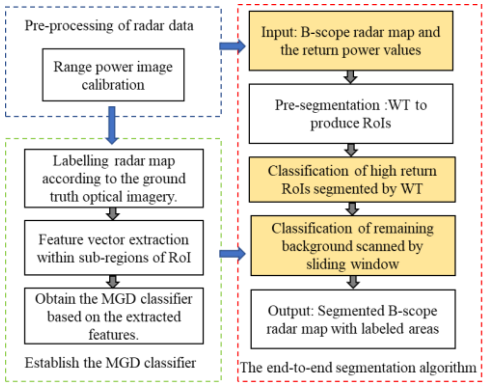

Fig. 2 The block diagram of the end-to-end segmentation on automotive radar imagery.

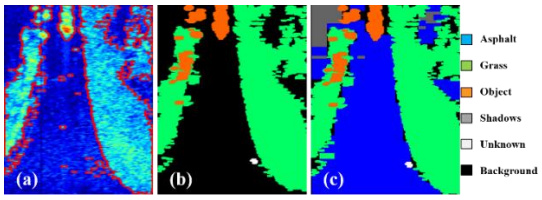

Fig. 3 (a) The result of pre-segmentation; (b) The segmentation mask after region classification of areas with higher contrast (mainly contain grass and objects); (c) The segmentation mask after region classification of the remaining background which are mostly asphalt and shadows.

B. The Region Classification based on MGD Classifier

The classification of pre-segmented color-coded RoIs is based on analysis of statistical parameters of return, in other words-distribution features of such regions, inputted in MGD classifier [1]. Two advantages can be achieved when using B-scope map as the input: simplicity of mapping between two formats of image (RGB) and data (radar map) representation and improved classification due to unambiguous correspondence between radar cells and pixels.

Indeed, when extracting the radar intensity values based on pre-segmented PPI map, the double counting of extracted intensities might happen in the overlapping areas of two regions. This impacts the accuracy of positioning of test feature vector in the space of features defined at the training stage and leads to confusion in area classification. It is worth stressing here, that avoiding of coordinate transformation w.r.t to the method with PPI input helps to reduce the computation complexity of procedure of distribution feature extraction.

Second improvement presented in this paper relates to introduction of two-tier of classification procedure. First layer will focus on limited number of classes with expected highest power return, such as grass and object regions used as an example in this paper. The classification result of this stage is presented in Fig.3 (b). RoIs after initial pre-segmentation based on WT are initially sorted by area size. Three kinds of regions

are defined here, which are small region (SR), with area size $s_a < S_1$, the general size region (GR), with area size $S_2 > s_a > S_1$ and the large size region (LR), with area size $s_a > S_2$. S_1 and S_2 are somewhat arbitrary, trail-based thresholds of region size to avoid over- and under-segmentations associated with deviations of features in the areas of the same class. The SRs are generally related to higher variations in the region of the same class and have size smaller than decorrelation spatial parameter. They are discarded as individual regions and merged with GRs and LRs. The GRs are classified by MGD classifier based on the extracted distribution features of return power values. The LRs may be related to combined regions, for instance large grass area can contain individual objects such as trees, or sign post. Therefore, we apply another layer of WT on the LRs to segment merged regions missed in the initial WT procedure. The classification of all the RoIs considered at this stage is based on the methodology of second-stage of classification in work [1] since the area shape produced by WT is irregular.

WT could not produce the contours of the remaining background due to the lower contrast between asphalt and shadows. Therefore, we aim to classify first and then remove regions of grass and objects in the first tier of classification, and use a sliding window to scan the remaining background region to produce the RoIs. The remaining background region is presented as black in Fig. 3 (b), and will be segmented and accordingly classified in the second tier of classification. The size of sliding window is 30×30 and is determined by the number of return power values required to produce one histogram plot. The classification of the remaining background is based on the results of first stage of classification discussed in our previous work [1]. The areas of shadows and asphalt are properly segmented in this stage, and the final segmentation result of B-scope radar map is shown in Fig. 3 (c). Most of the objects and surfaces are correctly segmented and classified using end-to-end segmentation.

IV. The Performance Estimation of End-to-End Segmentation on Automotive Radar Image

In this performance estimation part, eight frames of segmentation results (Fig. 4, frame 1(c) to 8(c)) are presented for scene analysis. The segmentation results are presented in the format of PPI map, which is transformed from segmented B-scope map, for the sake of visual comprehension of scene segmentation. The corresponded optical image and radar map are presented in Fig. 4, 1(a)-8(a), and 1(b)-8(b), respectively, which showed the annotated ground objects and surfaces.

The objects annotated in the ground truth optical imagery includes: 1) F, H, I and K, which are pedestrians; 2) C, G and J which are the trees; 3) M and P which are the vehicles. Most of the objects showed up in the ground truth imagery are correctly segmented and classified in the segmentation results. The only confused object in the segmentation results is C in Fig. 4, 1(b), which is a tree on the grass region. This can be explained that object C is regarded as the extremely small region and discarded for classification. This confusion of results can be optimized by improving the threshold for determining the small region size.

The surfaces annotated in the ground truth optical images includes: 1) E, which is the passible asphalt region; 2) D and O, which are the grass areas along the road. The surfaces are correctly classified and the segmented edges showed good agreement with the radar map. One of the possible confusions in the surface segmentation is that at longer ranges asphalt may be classified as shadow due to the fact that its power return becomes comparable to the noise floor of radar. It therefore has the same characteristics as the shadow, which is absence of the return, and will assume the statistics of the radar noise floor. For example, region B in Fig. 4, 1(c), which is the asphalt area at the range of 20 m in ground truth imagery, is confused as shadow. The noise floor hitting range are determined by factors such as the normalized radar cross section (NRCS), the area of resolution cell and the free space propagation loss. The method of estimating the noise floor hitting range of different areas will be discussed in our full publication.

The segmentation performance is estimated by calculating the JSC parameter based on the area size of labelled PPI radar map and the segmented PPI map [13]. We conduct the JCS estimation based on the test date set which is composed of 120 frames of radar imagery. Table 1 shows the comparison between the JSC estimation of the end-to-end segmentation discussed in this paper, (J_1), and the JSC estimation presented in our previous work [1], (J_2). This comparison shows that the segmentation performance of grass and objects have been significantly improved by implementing end-to-end segmentation algorithm. The mean JSC of all four areas has been increased 0.11.

Table 1. Comparison between the JSC estimation of the end-to-end segmentation and the JSC presented in previous work [1].

	Asphalt	Grass	Shadows	Objects	Mean JSC
J_1	0.8	0.84	0.81	0.87	0.83
J_2	0.8	0.64	0.79	0.64	0.72

The improvement in processing time after using the B-scope radar map for pre-segmentation has also been estimated. To do this we compare the computation times t_1 and t_2 taken to extract the return power values of one frame of a PPI map and one frame of a B-scope map respectively. The obtained values were $t_1 = 0.64$ s and $t_2 = 0.000156$ s; the processing time has been improved 4000 times by using the B-scope map.

V. Conclusion

In this paper, we have developed the end-to-end segmentation for high resolution automotive radar imagery based on previous algorithm proposed in [1]. To achieve the end-to-end segmentation, first pre-segmentation on B-scope radar map is used and then the classification of segmented RoIs is made based on MGD classifier and the extracted distribution features. It has been shown that the utilization of B-scope radar map will reduce computational complexity and improve accuracy of classification by avoiding overlap in regions due to coordinate transformation between scene maps sampled in pixels (for image processing) and radar cells (for classification, based on physical feature extraction) compared to using PPI radar map as an input.

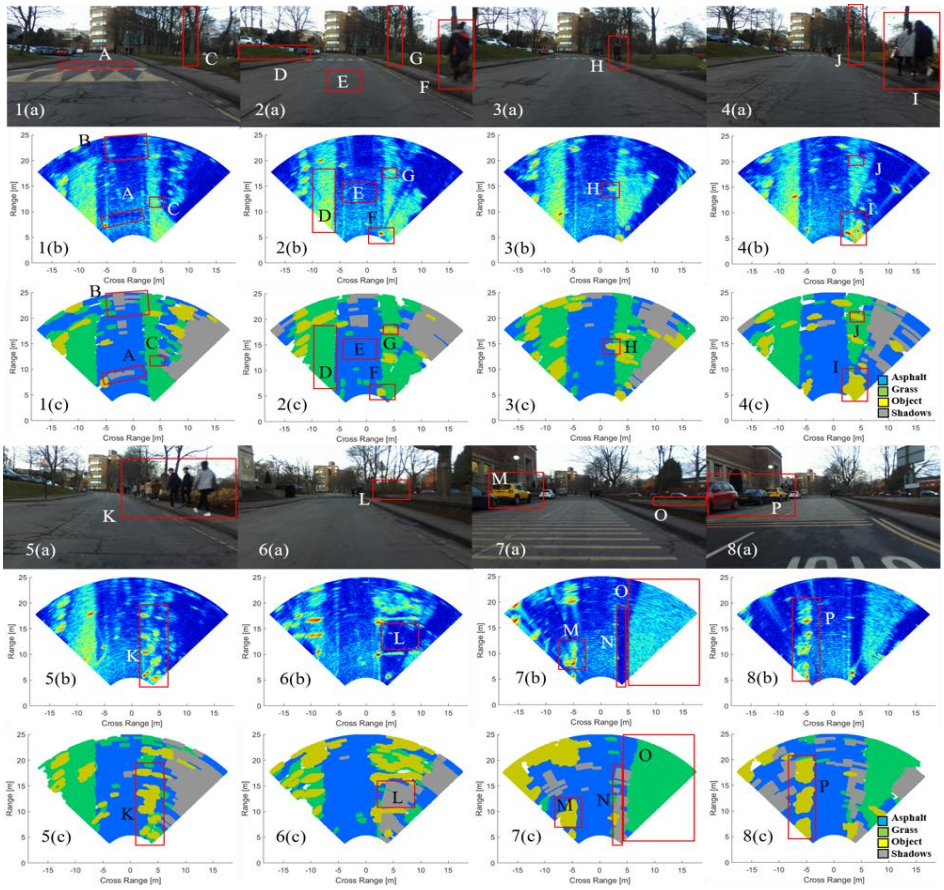

Fig. 4 Snapshots of scenes seen by the moving vehicle (radar is installed at the back). 1(a)-8(a) are ground truth optical images of scenes, 1(b) and 8(b) are radar maps and 1(c) and 8(c) are segmented maps.

To tackle the problems of region merging, over-segmentation, and confusion between similar classes we have introduced tiered (layer by layer) pre-segmentation process where regions of high return and high contrast are processed first to avoid masking the regions with low return and contrast, which will be then segmented and classified at the second tier of processing.

It has been shown that as a result of proposed approach the JSC-based metric of accuracy of classification is higher than that previously reported in [1]. Indeed the average JSC for all four areas has been increased by 0.11.

REFERENCES

[1] Y. Xiao, L. Daniel, and M. Gashinova, "Image Segmentation and Region Classification in Automotive High-Resolution Radar Imagery," *IEEE Sensors Journal*, vol. 21 no. 5, pp. 6698-6711, 2020.

[2] D. Feng *et al.*, "Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1341-1360, 2020.

[3] M. Aladem and S. A. Rawashdeh, "A single-stream segmentation and depth prediction CNN for autonomous driving," *IEEE Intelligent Systems*, pp. 1-1, 2020.

[4] M. D. Hossain and D. Chen, "Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 115-134, 2019.

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

[6] M. M. Trivedi and J. C. Bezdek, "Low-level segmentation of aerial images with fuzzy clustering," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 16, no. 4, pp. 589-598, 1986.

[7] 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 Transactions on Image Processing*, vol. 19, no. 10, pp. 2781-2787, 2010.

[8] F. Lang, J. Yang, D. Li, L. Zhao, and L. Shi, "Polarimetric SAR image segmentation using statistical region merging," *IEEE geoscience and remote sensing letters*, vol. 11, no. 2, pp. 509-513, 2013.

[9] R. Kemker, C. Salvaggio, and C. Kanan, "Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning," *ISPRS journal of photogrammetry and remote sensing*, vol. 145, pp. 60-77, 2018.

[10] Y. Xiao, L. Daniel, and M. Gashinova, "Feature-based Classification for Image Segmentation in Automotive Radar Based on Statistical Distribution Analysis," in *2020 IEEE Radar Conference (RadarConf20)*, 2020: IEEE, pp. 1-6.

[11] "http://elva-1.com/." (Accessed: 28th March 2021).

[12] "https://docs.opencv.org/master/d3/db4/tutorial_py_watershed.html." (Accessed: 28th March 2021)

[13] S. Niwattanakul, J. Singthongchai, E. Naenudorn, and S. Wanapu, "Using of Jaccard coefficient for keywords similarity," in *Proceedings of the international multiconference of engineers and computer scientists*, 2013, vol. 1, no. 6, pp. 380-384.