Road Users Classification with 24 GHz FMCW Radar

Rudi Coppola, Sajid Ahmed, and Mohamed-Slim Alouini

1 Introduction

Our goal is to propose a system for road users classification using an out-of-the-shelf 24 GHz FMCW commercial radar to demonstrate the scalability of our approach. In particular, our research was developed using Infineon's Position2Go sensor [1]. We based our classification model on hand-crafted features independent of the relative speed between the target and the sensor. We aimed at developing a solution for localization and classification which maintains the same performance in a scenario where the sensor is fixed on the ground and in a scenario where the sensor is mounted on a moving platform, such as a car or a quadcopter. The main aspects of our work are:

1. We used a single transmitter and a single receiver.
2. We chose low complexity features based on the micro-Doppler generated by the target's unique kinematics, radial distance, and target reflectivity.
3. We obtained a low prediction time by using only two subsequent frames.

2 Proposed Approach

Our solution is based on two fundamental steps, detection and classification. Remarkably, both steps are independent of each other: this means that the detection of a target can be achieved by different means, for example using multiple receivers, leaving the classification performance unaffected.

2.1 Detection

Under the assumption that the target is moving, we separate the clutter from the target's return using a double stage Moving Target Indication filter [2]; essentially, it consists of a high-pass filter for the slow-time samples. The clutter is static for the sensor while the target is moving: the effect of high-pass filtering results in the suppression of the return of static scatterers. This highlights the moving targets present in the scene. Note that in our experiments, a single moving target was present in the scene. However, our model can classify multiple targets on the scene as long as they are resolved in the range spectrum. When the target is located, the corresponding phase-history is extracted to obtain the Doppler spectrum.

2.2 Feature Extraction

The initial features to be extracted are the target's location and the range spectrum's amplitude. These are immediately available from the detection step and require no additional computation. Next, we consider the Doppler spectrum. The strongest Doppler frequency is associated with the target's ego speed. To capture the unique distribution of the micro-Doppler, disregarding ego speed and the target's reflectivity, in a preliminary step, the Doppler spectrum is shifted to center the ego speed at 0 m/s, and the magnitude is normalized to extend in the interval $[0, 1]$. At this point, we extract details on the micro-Doppler distribution around $0m/s$ using custom weighting functions to highlight different portions of the spectrum. In addition, we divide the magnitude of the spectrum into three intervals. We consider the number of spectrum peaks occurring, simple descriptive statistics as the second and third moment of the distribution of the peaks for each interval, and the variation of these parameters from one frame to the next.

2.3 Classification Model - Gradient Boosted Trees

We used Gradient Boosted Trees (GBTs) to build and train the model. This model is based on Classification and Regression Tree (CART) ensembles. A CART is created by recursive binary splitting of the feature space. In each iteration, the space is divided in two sub-regions according to some threshold that maximizes their purity. The procedure is iterated on each sub-region until no further improvement is achieved or a stop criterion is met [3]. Each branch of the tree terminates with a *leaf*, which contains the label of the class predicted. We used a specific implementation of GBTs provided by [4]: in this new formulation, each new tree added to the ensemble improves the accuracy based on the previous trees' errors. Besides, each tree is trained using a subset of the features available to improve the robustness and the generalization of the model.

3 Experimental Setup

Below we present the experimental setup used to create the video for the Radar Challenge.

3.1 Sensor Configuration

Tables 1, 2 summarize the the firmware configuration, the measurement specifications and the signal processing parameters.

Firmware Configuration		
Parameter	Meaning	Value
f_s	Start Frequency	24.025 GHz
B	Bandwidth	200 MHz
T_c	Up-chirp duration	300 μs
N_s	Samples per chirp	64
T_s	Fast time sampling frequency	4.687 μs
T_{PRI}	Pulse Repetition Interval	500 μs
N_c	Chirps per frame	128
T_{FRI}	Frame Repetition Interval	200 ms

Table 1: Firmware Configuration

Measurement Specifications and Signal Processing Parameters		
Parameter	Meaning	Value
R_{max}	Maximum unambiguous range	48 m
R_{res}	Range resolution	0.75 m
v_{max}	Maximum unambiguous velocity	±6.24 m/s
v_{res}	Velocity resolution	0.01 m/s
N_{pad}^R	Fast-time samples after zero-padding	256
N_{pad}^d	Slow-time samples after zero-padding	512

Table 2: Measurement Specifications

3.2 Data Acquisition

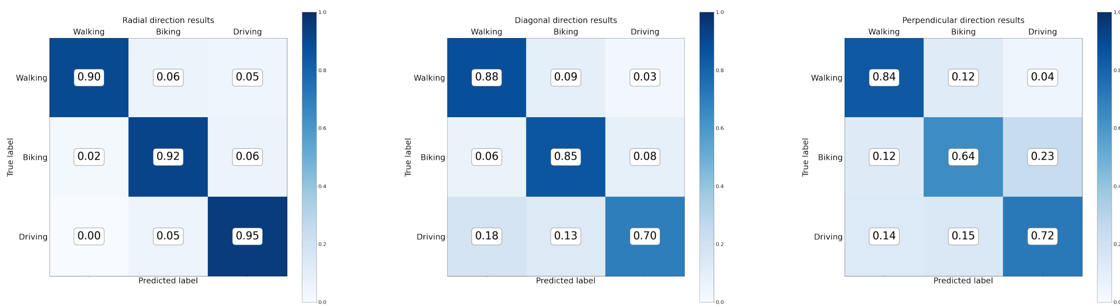
The experiments were conducted in an ample parking lot, as shown in the video, to limit the clutter interference. We started by considering targets moving in the radial direction and then gradually tested our methods on more complicated scenarios such as diagonal and perpendicular motion. A target moving in a perpendicular direction is challenging to classify based on the micro-Doppler, as the radar can measure only movements in the radial direction. Hence in this condition, the distinguishing traits of the kinematics of each target are less prominent. In total, we obtained 5123 frames during which a target was detected, distributed as in Table 3. We limited the detection to a range interval from 5 meters to 25 meters. In the video we start by presenting the results for a model based on radial motion (first column of Table 3) then proceed to show the results for a model based on the entire data set. In both cases, the data is split in 70% for training with a 5-fold stratified cross-validation for hyper-parameter tuning, and 30% for testing. The entire footage shown in the video is part of the test set: this means that the model was never exposed to it during training.

Number of frames			
	Radial	Diagonal	Perpendicular
Pedestrian	669	572	557
Cyclist	396	524	652
Car	310	568	875

Table 3: Frame distribution

3.3 Results

In Figure 1, we summarize only the results for the model trained and tested on the entire data set.



(a) Test on radial data (b) Test on diagonal data (c) Test on perpendicular data

Figure 1: Confusion matrices for the model trained on the complete data set and tested on data on each direction of motion

4 Conclusions

In this work, we presented an approach to classify road users with an FMCW radar, using two frames for each prediction, spaced 200ms apart. In the video we proposed:

1. A model for radial motion, which achieves 98.2%, 97.3% and 98.1% for pedestrians, cyclists, and cars respectively.

2. A model for general motion which achieves, on average, 87.3%, 80.3% and 79% for pedestrians, cyclists and cars respectively.

The device in combination with our model can be directly implemented as a surveillance system or mounted on a vehicle while preserving the same performance levels.

References

[1] Infineon Technologies, "DEMO POSITION2GO." [Online]. Available: https://www.infineon.com/cms/en/product/evaluation-boards/demo-position2go/

[2] M. Richards, *Fundamentals of Radar Signal Processing*, 2012.

[3] T. Hastie and J. H. Friedman, *Springer Series in Statistics The Elements of Statistical Learning*, 2009, vol. 27, no. 2.

[4] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-Augu, 2016, pp. 785–794.