Discrimination of stationary from moving targets with recurrent neural networks in automotive Radar

Conference Paper · April 2018

DOI: 10.1109/ICMIM.2018.8443525

6 authors, including:

Christopher Grimm
HELLA Group

Tai Fei
HELLA Group

Ernst Warsitz
HELLA Group

Reinhold Haeb-Umbach
Universität Paderborn

Discrimination of stationary from moving targets with recurrent neural networks in automotive Radar

Christopher Grimm, Tobias Breddermann, Ridha Farhoud, Tai Fei, Ernst Warsitz
Hella KGaA Hueck & Co, 59555 Lippstadt, Germany
Email: {christopher.grimm, tobias.breddermann, ridha.farhoud, tai.fei, ernst.warsitz}@hella.com

Reinhold Haeb-Umbach
Department of Communication Engineering
University of Paderborn
33098 Paderborn, Germany
Email: haeb@nt.uni-paderborn.de

***Abstract*—In this paper, we present a neural network based classification algorithm for the discrimination of moving from stationary targets in the sight of an automotive radar sensor. Compared to existing algorithms, the proposed algorithm can take into account multiple local radar targets instead of performing classification inference on each target individually resulting in superior discrimination accuracy, especially suitable for non rigid objects, like pedestrians, which in general have a wide velocity spread when multiple targets are detected.**

***Index Terms*—Automotive radar, classification, pedestrian, stationary, Recurrent Neural Network.**

I. Introduction

In the development of Advanced Driver Assistant Systems (ADAS) for partial or full autonomous cars, a full scene understanding is a desired objective and the motivation for multiple publications in the past. On the way to full scene understanding, automotive radar sensors could play an import role, since they are mostly insensitive to weather and lighting conditions, so can operate in almost any relevant environment condition other than e.g. cameras and LIDARs. For ADAS functions the classification of surrounding objects is an essential task, hence different traffic objects require special control functions in critical situations [1]. The ADAS functions therefore need extensive information about the surrounding objects like their size, position, class and trajectory.

Since radar signal processing offers just a limited measurement space with geometric (relative distance and angle of arrival) and kinematic (relative velocity) states, an appropriate model for accurate classification is necessary. One example where kinematic information is being used in radar, is the human gait detection, where slight doppler-shifts („Micro-Doppler“), introduced by the different oscillating body parts of a human when walking through a radar beam, are detetcted, see [2] [3]. Publications known to the authors, dealing with micro-doppler doppler classification approaches in general have full access to the complete doppler spectrum. Since current automotive radar sensors are very restricted in memory and processing time, they require significant information compression, where the full doppler spectrum cannot be stored during the complete signal processing time for exactly this reason. One of the earliest signal processing stages for automotive radar is target detection, where individual point target are detected in the noisy radar reflections. These targets, in general represent a subset of the range-doppler domain and therefore can be used for target classification, once the access to the whole range-doppler spectrum is not longer possible for subsequent processing algorithms. In previous publications [7], [8] we proposed a hypothesis based algorithm for the instantaneous discrimination of individual targets into the classes stationary or moving. Due to the discrimination of targets individual, the algorithms struggles to annotate very slow moving object reflections from pedestrian as moving, which appear occasionally like ground touching legs. For subsequent processing steps, this could reduce system performance significantly. To solve this issue, we present a neural network based method for the discrimination of stationary and moving targets in the vehicle environment, which investigates targets in a cluster and thus is able to understand local target relationship for improved inference. The proposed algorithm can thus be understood as the primary stage of more complex classification approaches. It estimates a confidence of its classifications, which support adjacent post processing steps in a way, that classical statistical data processing can be utilized. In contrast to this publication, there are also other classification algorithms which provide inference by investigating the relative movement of the object over longer observation time via tracking [4], [5] and therefore require longer period for accurate inference.

The paper is structured as follows. In section II the model for stationary targets is derived to motivate the used features for the proposed neural network algorithm. Crucial for neural network parameter adjustment is also the utilized training data, which in this paper is gathered in simulation and described in detail. Also the proposed model architecture is given. In section III the performance of the proposed algorithm is evaluated and compared to our previous algorithm. In section IV the paper is concluded.

II. Stationary Target Discrimination

A. Basic Model

Modern automotive radar sensors are capable of resolving the relative distance ${}^{R}\vec{r}_{P}(t)$, the relative velocity ${}^{R}\dot{\vec{r}}_{P}(t)$ and the Direction-of-Arrival (DoA) ${}^{R}\phi_{S}$ of a targets P in sight. S describes the sensor coordinate system with ${}^{S}x$ being sensor normal and R the coordinate system which is pointing towards the target P.

Here the radial velocity v_R of a target is measured by the radar sensor. The velocity v_{Ego} of the ego-vehicle can be obtained by the vehicles wheel encoders or estimated based on previously detected stationary targets, see [8]. The relative velocity equation is then fully described by

$$v_R = v_{\text{Ego}} \cdot \cos(\mu_{\Phi}), \tag{1}$$

where the right hand term can be interpreted as the expected relative velocity assuming a stationary target has been observed.

However, since these variables are in general infected by noise, the equations becomes an approximation and moving targets can be detected as outliers in the data, see [7], [6]. The variables are modeled as random and targets must disagree to equation 1 significantly to be detected as outliers and thus moving. Due to this significance very slow moving targets are often unintentionally classified as stationary, see [8].

B. Annotated data generation

1) Training data: Since training neural networks demands sufficient amount of annotated data, which is hard and expensive to gather on real world data, we create artificial data for training. Since the pysical relationship for stationary radar targets hold according to eq. 1 it is possible to Monte-Carlo sample stationary targets in the $\cos(\phi)$ vs. v_r plane. For stationary targets $[-180, 180)° = \{\phi \in \mathbb{R} | -180° \leq \phi < 180°\}$ and $[0, 30]\mathrm{m\,s^{-1}} = \{v_{\text{Ego}} \in \mathbb{R} | 0\,\mathrm{m\,s^{-1}} \leq v_{\text{Ego}} \leq 30\,\mathrm{m\,s^{-1}}\}$ are drawn from equal distributions and all variables are then corrupted with gaussian noise, which variance was identified for real world radar sensor in [8]. Moving targets, represented by driving cars, are drawn similar to stationary targets extended by their longitudinal velocity $[4, 30]\mathrm{m\,s^{-1}} = \{v_{\text{cars}} \in \mathbb{R} | 4\,\mathrm{m\,s^{-1}} \leq v_{\text{cars}} \leq 20\,\mathrm{m\,s^{-1}}\}$ and their heading angle $[-180, 180)° = \{\psi \in \mathbb{R} | -180° \leq \psi < 180°\}$.

In real world observations the collected radar targets are in general dominated by stationary targets and just a few moving targets are observed, however, in this work we draw equal number of moving and stationary targets for every frame in order not to shift the attention of the trained network more towards stationary targets. Beside stationary and moving targets, also targets from pedestrians are generated which imitate real world observations from pedestrian targets. Here we first draw the pose of pedestrians center in $x$vs. y-plane, draw the mean walking velocity from the pedestrian as uniformly distributed from $[1, 3]\mathrm{m\,s^{-1}} = \{v_{\text{ped.}} \in \mathbb{R} | 1\,\mathrm{m\,s^{-1}} \leq v_{\text{ped.}} \leq 3\,\mathrm{m\,s^{-1}}\}$, the number of the observed targets for each pedestrian $n_{\text{ped.,targets}} \in \{1, 2, 3, 4\}$ and their placement on the torso or extremity. Then the noise corrupted relative velocity, range and DoA for each pedestrian targets is computed.

2) Real world testing data: For testing the proposed algorithm performance, the same dataset which was used in our previous publication [7], [8] is used. In this dataset, the ego-vehicle drives on a straight trajectory with varying velocity and passes pedestrians walking parallel to the ego-vehicle with also varying velocities. This scenario was chosen, since walking pedestrians have little velocity difference to stationary targets and thus makes them hard to detect as moving targets. This scenario is typical for urban areas, where walkways often border to driveways, so the scenario has practical relevance for ADAS. As a radar sensor, a $77\,\mathrm{GHz}$ automotive development radar was used, which utilizes phase-monopulse as DoA estimation technique and provides range resolution of $0.1875\,\mathrm{m}$ up to $70\,\mathrm{m}$ in distance.

C. System Design

The weak point of the classification algorithm from [7] is, that it classifies each target individually and do not pay attention to specific data constellations of different target object, like pedestrians. We have already investigated in [7], that occasionally misclassification of radar targets associated to pedestrians appear to very slow movement like periodic disappearing velocity from pedestrians legs. In that case, the hypothesis test will favor towards stationary targets. In further processing steps these misclassified targets can lead to some confusion for example in stationary environment mapping, where the ground touching legs, misclassified as stationary, would result in a stride of the moving pedestrian in the map. Therefore it is desired to classify all targets belonging to moving objects as moving, even when they have negligible relative velocity to stationary objects.

In order to achieve radar targets constellation specific target classification, we first extract radar targets, which we believe could be associated to one object and thus form a subset of the object corresponding micro-doppler spectrum and then perform inference for target classification, as shown in fig. 1.

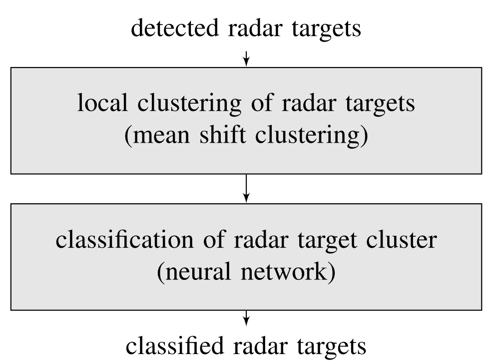

Figure 1: Flow chart of the proposed algorithm

1) Local clustering of radar targets: In the clustering step, we utilize mean shift clustering in the x vs. y domain ($\rightarrow$ top view on ego-car), since objects can easily described here and most likely do not overlap with other targets, whereas in range-doppler domain clustering algorithm need a more elaborated structure due to highly overlapping of object signatures. Mean shift clustering with radial basis function with bandwith parameter set to an average step with of pedestrian of $0.7\,\mathrm{m}$ according to [9], was chosen to collect radar targets likely stemming from pedestrians. One example for the clustering is shown in fig. 2, 2nd row.

2) Classification of radar target cluster: Since the detected clusters in general consist of a non constant number of targets, the classifier must take the radar target sequence length into account. To deal with this specific requirement, recurrent neural networks (RNN) come into spotlight. For exactly that reason, these types of function approximation tools are used for example in speech recognition and text-to-text transcriptions. In our application the radar targets are streamed into the net in arbitrary arrangement and give an class probability posterior estimate after completed input, thus doing multi-to-single prediction. For processing, we allow a maximum sequence length of 10 targets per cluster, exceeding leads to discarding of the following targets. However, on real world data we observed a maximum number of targets per cluster of 5, so it will not conflict. To allow the network self modulating the hidden states and forgetting, we choose long short-term memory (LSTM) in favor of standard RNN's since RNN's tend to unintentionally favor younger input sequence over older ones to do classification. However, we want the network to remember essential features of the cluster, no matter of radar target sequence. The LSTM is designed in such a way, that it receives the input vector X

$$X = [v_r, v_{\text{Ego}}, \cos(\phi), v_{\text{Ego}}\cos(\phi), R, \phi] \tag{2}$$

where v_r is relative velocity, v_{Ego} is ego-velocity, ϕ is the DoA, $v_{\text{Ego}}\cos(\phi)$ is expected relative velocity for a stationary target and R is range. Here we give network multiple transformed instances of ϕ to reduce linearization effort for the net. Based on eq. 1 the network is provided with all relevant features plus also polar-coordinates to learn local connection of the targets. From this 6-by-1 feature vector the network creates a 32-by-1 hidden state vector which is fully connected (weight + bias) to the 2-by-one logit vector and output (class prediction) is given by softmax function over the logits. In training we utilize cross entropy cost function with Adam [10] optimization algorithm in backpropagation through time for parameter tuning.

III. Results

A. Artificial data

While training the artificial dataset is generated on the fly, which means, that the network likely never sees a specific realization multiple times and therefore we do not introduce further regularization techniques. A subjective overview of the prediction performance including visual representation of clustering is given in fig. 2. Here it can be observed, that radar targets from pedestrians have been successfully clustered, illustrated by equally colored rings around the purple colored radar targets in fig. 2, row 2. Mostly the predicted class match to the ground truth, see fig. 2, row 2 vs. fig. 2, row 4.

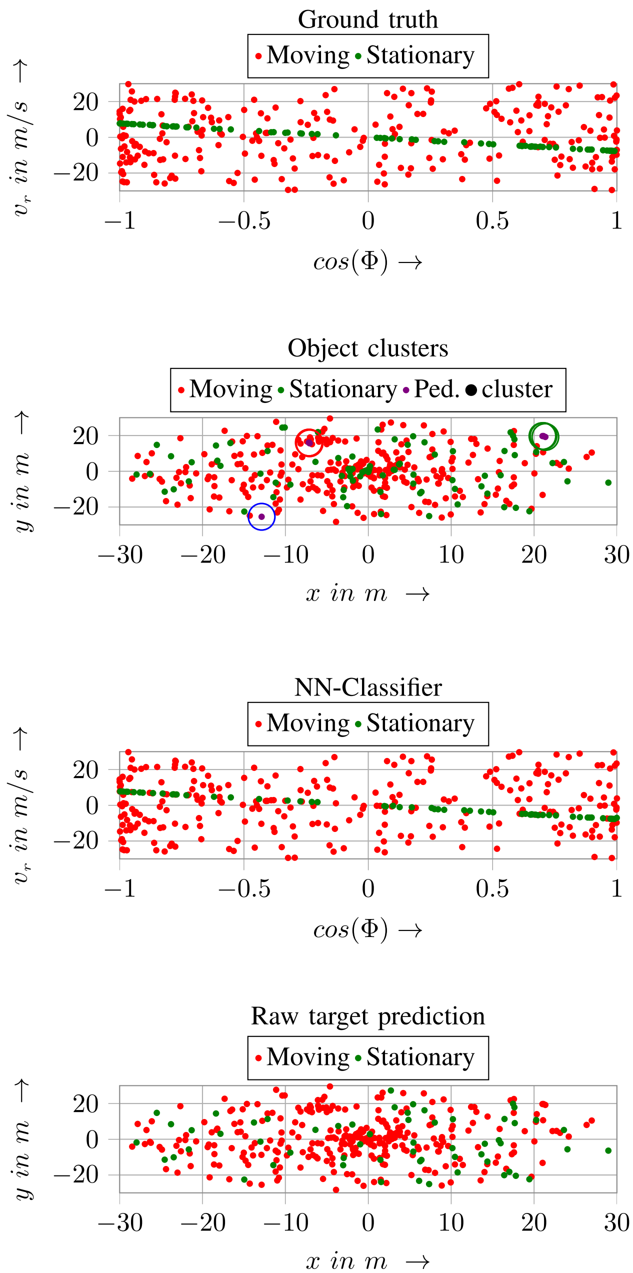

Figure 2: 1st row) ground truth annotation, 2nd row) ground truth annotation & cluster assignment for pedestrians in bird view, 3rd row) predictions, 4th row) predictions in bird view

To give statistical classification performance, the confusion matrix is given in tab. I, where the prediction is stated vs. the actual cluster annotations of moving, pedestrian and stationary.

Table I: Confusion matrix - proposed LSTM based classifier on artificial data

Actual \ Prediction	Moving	Stationary
Cars (moving)	99.7%	0.3%
Pedestrian	73.0%	27.0%
Stationary	0.7%	99.3%

As a reference, the confusion matrix of the hypothesis test from [7] is given in tab II.

Table II: Confusion matrix - classifier from [7] on artificial data

Actual \ Prediction	Moving	Stationary
Cars (moving)	91.0%	9.0%
Pedestrian	55.2%	44.8%
Stationary	0.5%	99.5%

Comparing these confusion matrices, it can be observed, that the proposed algorithm achieved in general superior classification performance for pedestrian compared to the hypothesis test. Also can be observed, that the classification accuracy for pedestrians as moving targets increased to 73% coming from 55.2%. Also the classification accuracy for moving targets increased to 97.1% coming from 91%. The superior classification results reflects as well in the prediction overview in fig. 2. Here the pedestrian radar target clusters have been classified as moving instead of stationary as intended.

B. Real world data

To quantify the real world performance of the proposed classification algorithm, we tested the accuracy on real world data and provide the results as confusion matrix in tab. III and in IV for the reference algorithm from [7].

Table III: Confusion matrix - proposed LSTM based classifier on real world

Actual \ Prediction	Moving	Stationary
Cars (moving)	99.5%	0.5%
Pedestrian	98.6%	1.4%
Stationary	0.4%	99.6%

Table IV: Confusion matrix - classifier from [7] on real world data

Actual \ Prediction	Moving	Stationary
Cars (moving)	99.0%	1.0%
Pedestrian	88.0%	12.0%
Stationary	6.2%	93.8%

The superior performance of the proposed classifier can be observed in a accuracy gain of over $\approx 10\%$ for pedestrians correctly classified as moving and $\approx 5\%$ for stationary targets. In contrast to the real world data, both classifiers achieve worse accuracy on the artificial dataset, which can be justified by more challenging scenarios in artificial dataset, like arbitrary heading movements of the moving/pedestrian targets in the simulation, which was desired for sensitive neural network parameter tuning.

IV. Conclusion

The intention of this paper is to provide a classification framework, which is able to discriminate moving radar targets from stationary ones accommodating the local structure of the radar targets and achieve better discrimination performance especially for pedestrian targets. Therefore a clustering of radar targets was performed to extract some local radar target information and transform it into a proper class prediction. Since the gathered radar targets per cluster are of uncertain number, a recurrent neural network type was utilized to perform classification. The proposed classifier achieved a accuracy gain of $\approx 10\%$ and $\approx 5\%$ for pedestrian targets and stationary targets respectively compared to a hypothesis based discrimination framework from prior publication.

References

[1] The Washington Post, *Google patents a way for self-driving cars to understand a cyclists hand signals*, https://www.washingtonpost.com/news/innovations/wp/2015/05/04/google-patents-a-way-for-self-driving-cars-to-understand-a-\\cyclists-hand-signals/?utm_term=.11181ff9b1bf, "[Visited; 11. April 2017]", 2016.

[2] P. van Dorp and F.C.A. Groen, *Feature-based human motion parameter estimation with radar*, IEEE Transactions on Radar, Sonar & Navigation, 2008.

[3] P. Molchanov, *Radar Target Classification by Micro-Doppler Contributions*, Dissertation Univ. Tampere, 2014.

[4] C. Lundquist and U. Orguner and P. Schoen, *Tracking Stationary Extended Objects for Road Mapping using Radar Measurments*, in Proceedings of the IEEE Intelligent Vehicles Symposium, 2009.

[5] L. Danielsson, *Tracking and radar sensor modelling for automotive safety systems*, University dissertation from Chalmers University of Technology, 2010.

[6] D. Kellner and M. Barjenbruch and K. Dietmayer and J. Klappstein, *Instantanous Lateral Velocity Estimation of a vehicle using doppler radar*, in Proceedings of the 16th International Conference on Information Fusion, 2013.

[7] C. Grimm and Breddermann, T. and R. Farhoud and T. Fei and E. Warsitz and R. Haeb-Umbach, *Hypothesis test for the detection of moving targets in automotive radar*, in Proceedings of the IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems, 2017.

[8] C. Grimm and R. Farhoud and T. Fei and E. Warsitz and R. Haeb-Umbach, *Detection of moving targets in automotive Radar with distorted ego-velocity information*, in Proceedings of the IEEE MICROWAVES, RADAR AND REMOTE SENSING SYMPOSIUM (MRRS), 2017.

[9] S. Heuel, *Fußgängererkennung im Straßenverkehr mit 24GHz Radarsensoren*, Dissertation, TU Hamburg-Harburg, Germany, 2013.

[10] Kingma, D. and Ba, J., *Adam: A Method for Stochastic Optimization*, arXiv:1412.6980 [cs], Jan. 2017.