

Range-Doppler Detection in Automotive Radar with Deep Learning

Weichong Ng
*School of Electrical and Electronic
Engineering*
Nanyang Technological University
Singapore
e170123@e.ntu.edu.sg

Guohua Wang
Algorithm Team
Hertzwel
Singapore
guohua.wang@hertzwel.com

Siddhartha
Algorithm Team
Hertzwel
Singapore
siddhartha@hertzwel.com

Zhiping Lin
*School of Electrical and Electronic
Engineering*
Nanyang Technological University
Singapore
ezplin@ntu.edu.sg

Bhaskar Jyoti Dutta
Hertzwel
Singapore
bhaskar.dutta@hertzwel.com

***Abstract*—This paper presents a comprehensive study on radar detection using deep learning with application to automotive vehicles. Automotive radars face complex target scenarios consisting of both small point targets and large extended targets. However, the current works on automotive radar detection mainly focus on point target detection. Moreover, those works use the complex range-Doppler data for detection. In this paper, a deep learning-based method for extended target detection is presented that takes advantage of augmented data for neural network training and prediction. Extensive simulations have been conducted to evaluate the proposed detection method and the results show performance improvement over a recent related method.**

I. INTRODUCTION

Automotive radar can detect the positions and speeds of objects in the vicinity of the cars. It is playing a more and more important role in ensuring the safety of automotive driving and has been regarded as one of the enabling sensor technologies to realize the fully autonomous navigation of cars [1-2]. One of the fundamental problems in automotive radar is radar target detection [3]. In automotive radar applications, signal processing flowchart starts from range and Doppler processing by Fast Fourier Transform (FFT), which produces a Range-Doppler (RD) map [3]. After that, target detection is performed on the RD map to identify those Range-Doppler bins that are most likely due to potential targets. One popular detection methods is the constant false alarm rate (CFAR) detection. In general, the CFAR detection in RD domain takes a probabilistic approach to set a threshold and compares it with the range-Doppler bin data in order to determine the presence of a target. Meanwhile, the threshold is set in such a way that the False Alarm Rate (FAR) is constant, and the probability of the detection is proportional to the signal-to-noise ratio (SNR). The probabilistic approach of CFAR usually relies on the assumptions on the distribution of

noise and target, which may be easily violated in practical scenarios. Whenever this happens, the detection may deviate from the expected performance. Among the CFAR class of detection approaches, there are many different algorithms for setting the threshold to deal with issues such as target masking, clutter variation, etc. Two most commonly used CFAR detection methods are the Cell-Averaging CFAR (CA-CFAR) and Ordered-Statistics CFAR (OS-CFAR) [4,14], which can be applied to RD detection in the two-dimensional domain. CA-CFAR has lower computational complexity but inferior performance as compared to OS-CFAR. The main problem in automotive radar target detection lies in that there are multiple targets of different physical sizes and Radar Cross Section (RCS) coexisting in the same RD map. For example, big buses and trucks may occupy tens of resolution bins in range and Doppler while small objects like bicycle and human may only occupy a few bins. Meanwhile, in order to achieve good performance, the parameters of CA-CFAR and OS-CFAR such as the window size, guard cell size need to be tuned to suit certain scenarios. This complex situation and nature of conventional probabilistic approaches make it challenging to design a common detector for all targets in a given RD map.

Deep learning (DL) is a powerful approach to a wide range of problems and has been applied in many different areas, such as target classification, tracking, and detection in the point cloud domain in automotive radar signal and data processing [5-7]. In [8], target detection using a fully-connected feed-forward network with the help of minimax normalization, can achieve a comparable or a slightly lower detection performance as compared to the traditional CA-CFAR. However, by using this neural network, it could obtain a much lower FAR. Although improving the detection rate is an important factor, lowering the FAR can be beneficial to many practical radar setups also. In [9], U-Net was used as the neural network for target detection. It was trained using data consisting of complex raw data of RD map. In [9], a better

detection performance using deep learning was obtained as compared to CA-CFAR. However, there are some limitations in [9]. First, only single target was considered in [9]. Second, only point target was investigated and the detection was claimed when the predicted target was within a 3-by-3 grid centered at the true position, which is not realistic when practical applications are concerned. When it comes to the point target detection, we consider multiple targets and more importantly, extended targets. Multiple targets and extended targets are more challenging due to the mutual interference effect among targets. In order to improve the detection performance, this paper proposes a new augmented data method for training the deep neural network. In particular, we train the U-net using augmented data set that consists of both the complex valued data directly from the range-Doppler processing, as well as the magnitude value of the RD map. The rationale behind this data augmentation is that the CNN has excellent performance in image processing. By augmenting the complex data with the magnitude value, we exploit the augmented data structure to make better use of CNN for improved detection performance. Simulation studies based on realistic scenarios are carried out to illustrate the performance of the proposed deep learning detection approach.

The novelty of the paper mainly lies in the following:

1. Data augmentation is exploited in the range-Doppler detection using deep learning. The existing work on the RD detection is using only complex data. Instead, our data augmentation method makes use of both the complex data and the absolute value of the RD map to take advantage of the CNN in image processing.
2. Detection of extended target is studied with extensive simulations.
3. Detection of multiple targets is studied.

The rest of this paper is organized as follows. In section II, we introduce the background of automotive radar and the detection problem. In section III, DL based detection approach is described, followed by extensive simulation studies presented in section IV. Finally, section V concludes this paper.

II. BACKGROUNDS ON RD DETECTION IN AUTOMOTIVE RADAR

A. Automotive Radar

Frequency Modulated Continuous Wave (FMCW) radar is a type of radar sensor which radiates continuous transmission power. FMCW radar relies on the de-chirp processing on FMCW in the mixing stage that leads to significantly reduced receiver bandwidth and processing complexity. This advantage as compared with other radar waveform types makes FMCW radar a popular choice for automotive vehicles such as UAV's [10-11]. In a typical FMCW radar system, there is a waveform synthesizer to generate a frequency-modulated continuous wave signal that is transmitted by the transmit antenna. The receiver antenna receives the reflected signal which is then mixed with the Synthesizer signal to produce an Intermediate-Frequency (IF) output. IF signal is

detection of extended targets, the performance of the existing method may drop significantly, though this was not investigated in [9].

In this paper a comprehensive study on the deep learning based RD detection is presented. Instead of focusing on single

then passed to a low pass filter, Analog-to-Digital Converter (ADC) and Digital Signal Processor (DSP) for processing. Fig. 1 illustrate a simple FMCW radar setup.

The IF signal has a beat frequency that is proportional to the distance of the object from the radar. By applying a Fast Fourier Transform (FFT) onto the IF signal, the distance of the target can be obtained. If there is a moving target, the IF signal also has a Doppler component that depends on the relative velocity between the radar and the target. The Doppler component can be obtained by taking a second FFT over all the pulses received.

Multiple Input Multiple Output (MIMO) FMCW radar refers to a FMCW radar with multiple transmit and multiple receive antennas. MIMO FMCW radar as compared with conventional FMCW radar has improved angular resolution due to the extended virtual aperture size enabled by waveform diversity from the transmit side [12]. In a typical processing scheme for MIMO FMCW radar, the 2D FFT is performed for each transmit-receive pair to generate a RD map. Each RD map corresponds to one virtual antenna.

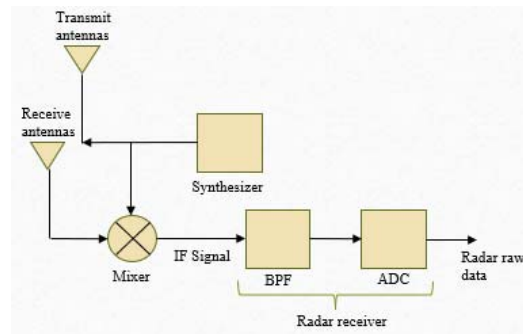


Fig. 1. A simple FMCW radar

There are some parameters that are required to specify in MIMO FMCW radar systems, such as the number of antennas, range resolution, maximum range and maximum velocity. The definitions of those parameters are given in below.

A pulse of FMCW is also called a chirp which is characterized by a start frequency (f_c), Bandwidth (B) and duration (T_c). Range resolution d_{res} , refers to the ability to resolve two closely spaced targets given by [13]:

$$d_{res} = \frac{c}{2B}$$

where c is speed of light and B is the bandwidth of the chirp. Maximum range d_{max} , is the maximum radar detection range given by [13]:

$$d_{max} = \frac{F_s c}{2S}$$

where F_s is the ADC sampling rate and S is the slope of the chirp. Maximum velocity V_{max} , is the maximum radar detection velocity given by [13]:

$$V_{max} = \frac{c}{4f_c T_c}$$

where f_c is the chirp start frequency, and T_c is the chirp duration. RCS σ , is the measure of a target's ability to reflect radar signals in the direction of the radar receiver given by [14]:

$$\sigma = \frac{4\pi r^2 S_r}{S_t}$$

where r is the range of the target, S_r is the scattered power density in the range r and S_t is the power density that is intercepted by the target.

In the simulations to be presented in this paper, we are using 2 transmit antennas and 4 receive antennas that form an array of 8 virtual antennas. Range resolution is set at 1m, maximum range is 100m and maximum speed is ± 73 m/s.

B. Range Doppler Detection in Automotive Radars

The purpose of RD detection is to determine which RD cell in the RD map is due to target. Traditional radar detection CFAR has a detection strategy of maintaining a fixed FAR and maximising the probability of detection [15]. However, there are a handful of parameters that the user must predefine such as the threshold, the number of guard cells and the number of training cells, as shown in Fig. 2. The performance of the traditional radar detectors can be affected if those parameters are not carefully tuned. In the real-world applications, the radar may be required to operate in different scenario which makes it difficult to tune the parameters adaptively. Meanwhile, as demonstrated in [8-9], those traditional radar detection methods have inferior performance as compared to existing deep learning methods. Therefore, this paper is working on enhanced deep learning technique to achieve adaptive radar RD detection with improved performance.

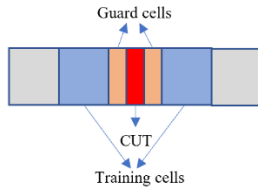


Fig. 2. 1-D CA-CFAR detector

III. DEEP LEARNING RD DETECTION

A. Range Doppler Data

In this work, radar range resolution d_{res} , radar maximum range d_{max} and radar maximum velocity V_{max} are set at 1m, 100m and ± 73 m/s, respectively. Simulation data is generated with the help of MATLAB Phased Array System Toolbox [16]. It simulates a 77 GHz millimeter wave radar with a 2-element transmit array and a 4-element receive array to form 8 virtual arrays. Both 2-element transmit array and 4-

element array are uniform linear array. Target distance, speed and RCS σ , are set at random between, 1 to 100m, -73 to 73m/s and 1 to 100 m², respectively. For multi-target detection, the number of targets is chosen randomly between 1 to 6. Complex baseband noise is added to the input signal by changing the noise figure randomly between 0 to 40dB to each data frame. Noise figure describes the performance of a receiver, and it is the dB equivalent of the noise factor. The noise factor is defined as the ratio of the input signal-to-noise ratio to the output signal-to-noise ratio. The dimension of the raw data frame is $N_s \times N_{Ant} \times N_C$, where N_s is the number of samples, N_{Ant} is the number of virtual arrays and N_C is the number of chirps. Windowing is used on the raw data frame to reduce the side-lobes in both range and Doppler domain. In particular, Taylor windows are used.

2D-FFT is applied to the raw data frame and the output has a dimension of $N_R \times N_D \times N_{Ch}$, where N_R is the number of range FFT length, N_D is the Doppler FFT length and N_{Ch} is the number of channels. It is used as an input to RD detection network. Ground Truth (GT) corresponding to actual target position and velocity is generated in the dimension of $Row \times Column$, where Row represents the range and Column represents the Doppler. The target in the GT is represented by a single dot, having a pixel value of 255. While the non-target has a pixel value of 0. RD detection is formulated as segmentation, and the target in the radar data frame is represented by a single white dot on the GT.

Radar data frames are complex data. Hence, data of one virtual antenna can be split into 2 channels directly, i.e., the real component and the imaginary component. This is the data used in [9]. In this paper, we augment this complex data with the magnitude value. Meanwhile, the complex data component is normalised by dividing with the largest magnitude value. To improve the detection, Logarithmic functions is applied to the magnitude component.

To compare the effect of adding the magnitude value, 2 types of data are generated, one is using only the complex data and N_{Ch} is $2 \times 8 = 16$, which was used by the existing method in [9]. The proposed augmented data consists of both complex data and magnitude value. Hence, N_{Ch} is $3 \times 8 = 24$.

B. U-Net

U-Net is a convolutional network architecture that is mostly used for image segmentation. A U-Net has ‘‘U’’ shape and it contains two paths. The first path is the contraction path, which consists of stacks of convolution followed by activation function and max pooling layers for down sampling. Max pooling operation reduces the dimension and helps to understand the class information [17]. The second path is the expanding path, consisting of an up-sampling of the feature map followed by convolution together with activation function, and a concatenation with a corresponding cropped feature map from the contracting path. The up-sampling process provides precise localization [18].

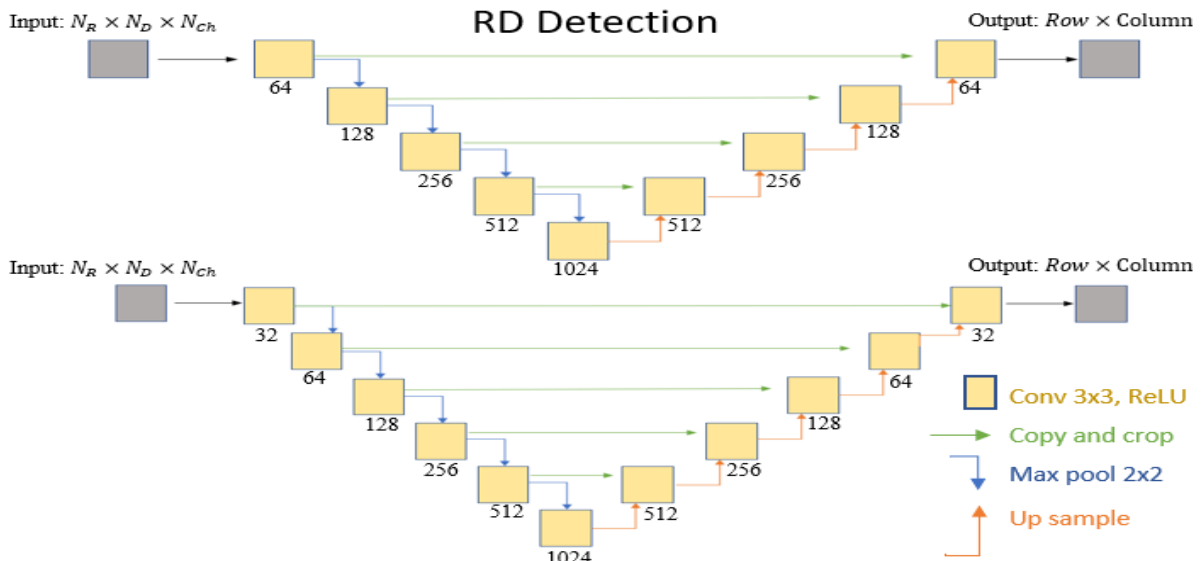


Fig. 3. RD Detection Network architecture use by case studies 1 and 2 (top), RD Detection Network architecture use by case studies 3 and 4 (bottom)

C. U-Net Training

In this paper, RD Detection is using 2D U-Net as the network architecture. The network is trained from scratch, it did not load any pre-trained model. U-Net is using 3×3 convolution filters with Rectified Linear Unit (ReLU) as the activation function, 2×2 max pooling filters with stride 2 and 2×2 up-sampling filters. In total, the network for case studies 1 and 2 has 23 convolution layers and the network for case studies 3 and 4 has 28 convolution layers. The network architecture can be found in Fig. 3. Since the RD Detection is a binary problem, binary cross entropy dice loss is used.

The RD Detection operates with a class-imbalance data, the number of non-targets is far greater than the number of the targets. F1 Score is used as the metrics for better evaluation of the model training process. F1 Score is taking a balance between Precision and Recall. Precision is how accurate the model is out of those predicted positive, how many of them are actual positive. Recall calculates how many of the actual positive the model capture through labelling it as positive. Precision, Recall and F1 Score are generated from the confusion matrix [19] which is shown in table 1. Class 1 is referring to the target and Class 2 is referring to the non-target.

TABLE 1. CONFUSION MATRIX

	Predicted Class 1	Predicted Class 2
Actual class 1	True Positive (TP)	False Positive (FP)
Actual class 2	False Negative (FN)	True Negative (TN)

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1\ Score = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

Overfitting is a modelling error that occurs when the model is trained to closely fit with the training data set. The model does not generalize well when it is tested against the unseen data. The following steps are taken to reduce overfitting [20, 21]:

- Dropout is set at 0.2.
- Early stopping will monitor the validation F1 Score. If the validation F1 Score does not increase after 8 epochs, then the training will be stopped. The number of epochs is set at 400. With the enabling of Early stopping, the model does not have to training the full 400 epochs.

All of our network models are trained with the Adam optimizer, with initial learning rate 0.001, betas 0.9, 0.999 [9]. If the validation F1 Score does not increase after 4 epochs, the learning rate will be decreased at a factor of 0.1. Batch size of 32 was used. It was implemented using Keras [22].

$$New\ Learning\ Rate = learning\ rate \times factor$$

D. Dataset for Deep Learning

In order to train each model, 26300 radar data with multiple target have been generated. These 26300 data have been split into [14728, 3682, 7890] which is for training, validating and testing. In this paper, there are 4 case studies. In each of the case studies, comparison is done between the existing method in [9] and the proposed method. Since we are using 8 virtual antennas, the existing method [9] will have $N_{Ch} = 16$ and the proposed method will have $N_{Ch} = 24$.

- Models are trained and tested on single target data.
- 10000 Single radar target data are generated for noise figure at 0dB, 10dB, 20dB, 30dB, 40dB each. These data are used to evaluate the models that are trained in the above case study.

- Models are trained and tested on extended target data. It is also evaluated when the noise figure is at 0dB, 10dB, 20dB, 30dB and 40dB.
- Different types of extended target and multi-target are added together to form a new dataset. Models are trained and tested based on this dataset.

IV. SIMULATION RESULTS

In this section, the proposed deep learning detection method is compared with an existing and related method [9] by considering various scenarios as shown in the following different case studies.

A. Case-1: Single Point Target and Mixed SNR

In this case, we consider a single point target with a range of SNRs. The data is generated according to the description in Section III.

Two models are trained using the existing method [9] and the proposed method. Both training graph can be found in Fig. 4 and Fig. 5. Table 3 shows the models training and testing F1 Score.

A False alarm is an erroneous RD Detection caused by noise or other interfering signals. FAR is calculated for both models, and it can be found in table 2 [19].

$$FAR = \frac{FP}{FP + TN}$$

TABLE 2. FAR OF BOTH METHOD IN [9] AND PROPOSED METHOD

Model	FAR
Method in [9]	4.8591E-06
Proposed method	3.89192E-06

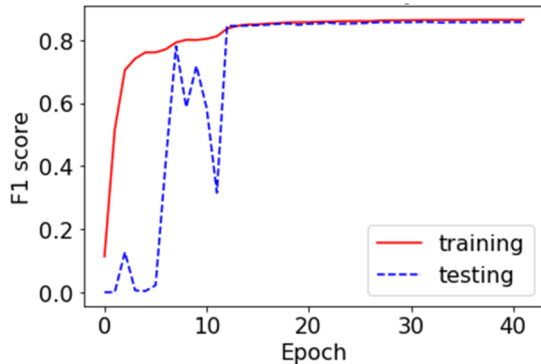


Fig. 4. Method in [9] model training graph

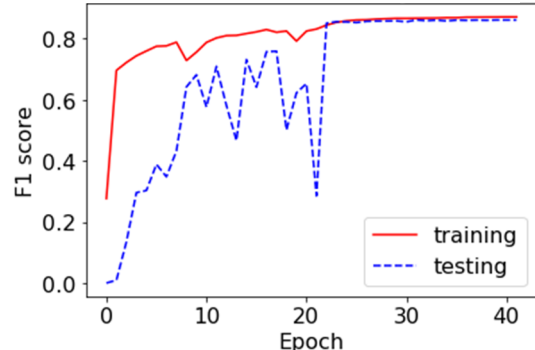


Fig. 5. Proposed method model training graph.

TABLE 3. PERFORMANCE OF MODEL TRAINED BY METHOD IN [9] VS PROPOSED METHOD MODEL

Model	Training F1 Score	Testing F1 Score
Method in [9]	83.220	82.389
Proposed method	88.322	87.611

Since the models training and testing data consist of radar data frames whose noise figure is set at random between 0-40dB, the testing F1 Score for both models are not very high. It can be seen from case study 2 that, both models can perform much better when the noise figure is set between 0-20dB. By comparing the testing F1 Score and the FAR, it can be observed that the model trained by the proposed method has shown that it has a better F1 Score and lower FAR.

B. Case-2: Single Point Target and Multiple SNR

To evaluate the impact of noise on the detection performance, the detection rate is investigated for each SNR. Range, speed and RCS of the target are set at random. Each evaluation is using 10000 data frames with its noise figure set at 0dB, 10dB, 20dB, 30dB and 40dB, respectively. For a single target data, when the noise figure is at 0dB, range is set at 1m, speed is set at 1m/s and RCS is set at $10m^2$, the Signal to Noise Ratio (SNR) is 22.11 dB. Table 4 gives the models evaluation result and Fig. 6. is the corresponding graph. It can be clearly observed that the proposed method trained model has higher F1 Score than the existing method [9] trained model. Therefore, the model trained by the proposed method is more robust to noise.

TABLE 4. MODELS EVALUATION USING DATA WITH DIFFERENT NOISE LEVEL

Noise Figure (dB)	Method in [9] model F1 Score	Proposed method model F1 Score
0	94.5236	96.280
10	91.3493	93.070
20	85.5261	87.3331
30	78.8353	81.2006
40	67.6988	70.960

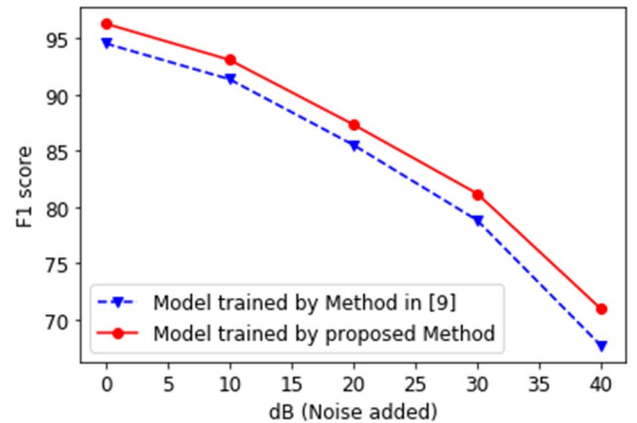


Fig. 6. Model evaluation graph for model trained by Method in [9] vs model trained by proposed method.

C. Case-3: Single Extended Target

In this case study, the performance of the proposed detection method will be compared with the existing method in [9] for the case of a single extended target. Training data are generated using a single $3 \times 3m$ extended target. Its GT is $3 \times 3 \text{ pixel}$ at the corresponding location. Due to hardware limitation, for case 3 and case 4, 18000 data has been used and it was split into [10400, 2600, 5000] which was for training, validating and testing. 5000 data frames were used to evaluate the model at each noise level. The batch size had reduced from 32 to 26. The extended single target range, speed, RCS and the amount of noise added are randomly generated. Table 5 is the models evaluation result and Fig. 7 is the corresponding graph. It is evident that at different noise levels, the F1 Score of the model trained by the proposed method is always higher than the model trained by the existing method [9]. In Table 6, FAR was calculated for both models and it shown that the proposed method trained model has a lower FAR. Fig. 8 shows an example of the models input.

TABLE 5. EXTENDED TARGET MODELS EVALUATION USING DATA WITH DIFFERENT NOISE LEVEL

Noise Figure (dB)	Method in [9] model F1 Score	Proposed method model F1 Score
0	97.7861	98.9039
10	95.8096	97.3486
20	90.6200	92.2321
30	83.7677	85.0992
40	74.8768	76.0139

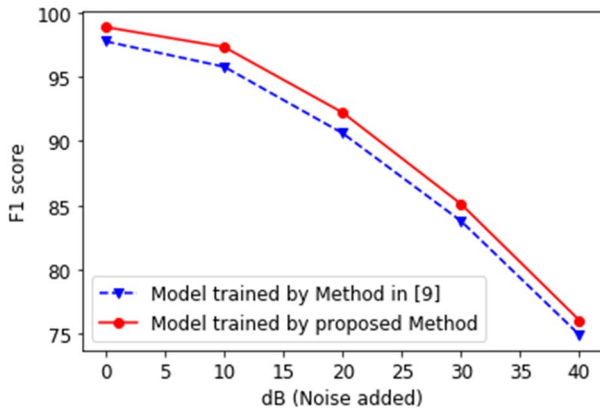


Fig. 7. Extended target models evaluation graph for model trained by method in [9] vs model trained by proposed

TABLE 6. FAR FOR BOTH METHOD IN [9] AND PROPOSED METHOD

Model	FAR
Method in [9]	5.38404E-06
Proposed method	5.10019E-06

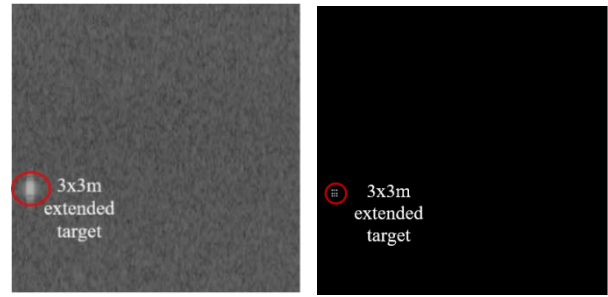


Fig. 8. 256 x 256 pixel resolution RD map of a single $3 \times 3m$ extended target (left). Its GT (right) represented by $3 \times 3 \text{ pixel}$ at the corresponding location with pixel value of 255. We are using only the magnitude value from 1 virtual antenna for displaying purpose.

D. Case-4: Multiple Target

In this case study, multiple target consisting of both point target and extended target are considered. In total there are 3 type of extended target. They are, $3 \times 3m$, $3 \times 9m$ and $9 \times 3m$. Every dataset has an $3 \times 3m$ extended target. The number of point targets are selected between 1 to 6 and they are added randomly with a 50% chance. $3 \times 9m$ and $9 \times 3m$ extended target are also added randomly with a 50% chance. This form a new dataset consisting of 18000 data and it was split into [10400, 2600, 5000] which was for training, validating and testing. Two new models are trained using these new datasets. Table 7 is the models evaluation result and Fig. 9 is the corresponding graph. FAR was calculated in Table 8. It is noticeable that with an additional magnitude component, the performance of the model proposed in this paper is better than using the method in [9]. Fig. 10 shows an example of the models input.

TABLE 7. MODELS EVALUATION USING DATA WITH DIFFERENT NOISE LEVEL

Noise Figure (dB)	Method in [9] model F1 Score	Proposed method model F1 Score
0	93.463	97.2479
10	92.0665	95.677
20	88.34038	91.5572
30	82.79761	85.6726
40	74.1198	77.2874

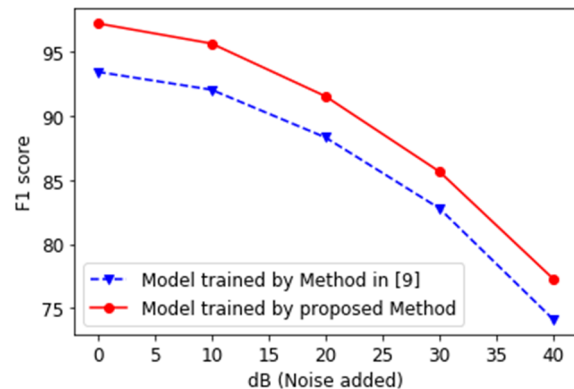


Fig. 9. Models evaluation graph for model trained by Method in [9] vs model trained by proposed method.

TABLE 8. FAR FOR BOTH METHOD IN [9] AND PROPOSED METHOD

Model	FAR
Method in [9]	2.78626E-05
Proposed method	2.70199E-05

For both models, individual extended target and multi-target are extracted from the prediction results and their F1 Score are calculated in Table 9. The proposed method remains to have a higher F1 score as compared to the method in [9].

TABLE 9. INDIVIDUAL F1 SCORE

Type	Method in [9] model F1 Score	Proposed method model F1 Score
3x3 extended target	89.7316	92.7997
3x9 extended target	90.9297	93.7245
9x3 extended target	89.1999	92.858
Multi-targets	71.4473	87.287

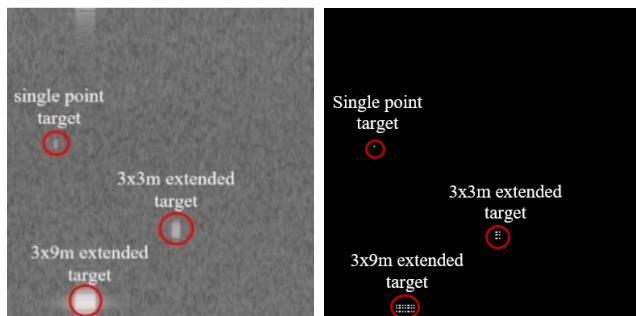


Fig. 10. 256 x 256 pixel resolution RD map of multiple extended target with single point target (left). Its GT (right) represented by corresponding location with pixel value of 255. We are using only the magnitude value from 1 virtual antenna for displaying purpose.

V. CONCLUSION

This paper proposed a RD detection method using deep learning. Unlike the existing work of deep learning-based detection which studied only point target detection [9], this paper proposed a new method for the detection of extended target using augmented data consisting of both complex data and its magnitude values. Extensive simulation study has been performed on realistic scenarios including both extended targets and point targets. The simulation results demonstrated that the new method proposed in this paper can achieve better detection performance with lower FAR and higher detection rate. Future work will focus on the real data evaluation of the method proposed.

ACKNOWLEDGMENT

The first author would like to thank technical staffs from Hertzwell Pte Ltd for providing valuable discussions on this work.

REFERENCES

- [1] Bilik, I., Bialer, O., Villeval, S., Sharifi, H., Kona, K., Pan, M., ... & Geary, K. (2016, May). Automotive MIMO radar for urban environments. In *2016 IEEE Radar Conference (RadarConf)* (pp. 1-6). IEEE.
- [2] Murad, M., Nickolaou, J., Raz, G., Colburn, J. S., & Geary, K. (2012, May). Next generation short range radar (SRR) for automotive applications. In *2012 IEEE Radar Conference* (pp. 0214-0219). IEEE.
- [3] Winkler, V. (2007, October). Range Doppler detection for automotive FMCW radars. In *2007 European Radar Conference* (pp. 166-169). IEEE.
- [4] Blake, S. (1988). OS-CFAR theory for multiple targets and nonuniform clutter. *IEEE transactions on aerospace and electronic systems*, 24(6), 785-790.
- [5] Chen, S., Wang, H., Xu, F., & Jin, Y. Q. (2016). Target classification using the deep convolutional networks for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), 4806-4817.
- [6] Held, D., Thrun, S., & Savarese, S. (2016, October). Learning to track at 100 fps with deep regression networks. In *European Conference on Computer Vision* (pp. 749-765). Springer, Cham.
- [7] Yang, S., Luo, P., Loy, C. C., & Tang, X. (2015). From facial parts responses to face detection: A deep learning approach. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3676-3684).
- [8] Akhtar, J., & Olsen, K. E. (2018, August). A Neural Network Target Detector with Partial CA-CFAR Supervised Training. In *2018 International Conference on Radar (RADAR)* (pp. 1-6).
- [9] Brodeski, D., Bilik, I., & Giryas, R. (2019, April). Deep Radar Detector. In *2019 IEEE Radar Conference (RadarConf)* (pp. 1-6). IEEE.
- [10] Oh, B.-S., Guo, X., & Lin, Z. (2019). A UAV Classification System based on FMCW Radar Micro-Doppler Signature Analysis. *Expert Syst. Appl.*, vol. 132, pp. 239-255.
- [11] Sun, H., Oh, B.-S., Guo, X., & Lin, Z. (2019). Improving the Doppler Resolution of Ground-Based Surveillance Radar for Drone Detection. *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, pp. 3667 - 3673.
- [12] Fishler, E., Haimovich, A., Blum, R., Chizhik, D., Cimini, L., & Valenzuela, R. (2004, April). MIMO radar: An idea whose time has come. In *Proceedings of the IEEE radar conference* (Vol. 2004, pp. 71-78).
- [13] Rao, S. (2017). Introduction to mmWave sensing: FMCW radars. *Texas Instruments (TI) mmWave Training Series*.
- [14] Richards, M. A. (2005). *Fundamentals of radar signal processing*. Tata McGraw-Hill Education.
- [15] Ward, K. D., Watts, S., & Tough, R. J. (2006). *Sea clutter: scattering, the K distribution and radar performance* (Vol. 20). IET.

- [16] MATLAB and Phased Array System Toolbox Release 2019b, The MathWorks, Inc., Natick, Massachusetts, United States.
- [17] Graham, B. (2014). Fractional max-pooling. *arXiv:1412.6071*.
- [18] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [19] Davis, J., & Goadrich, M. (2006, June). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning* (pp. 233-240). ACM.
- [20] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- [21] Caruana, R., Lawrence, S., & Giles, C. L. (2001). Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In *Advances in neural information processing systems* (pp. 402-408).
- [22] Chollet, F. (2015). keras. <https://keras.io/>