

## Article

# DoA Estimation for FMCW Radar by 3D-CNN

Tzu-Hsien Sang <sup>1,\*</sup>, Feng-Tsun Chien <sup>1</sup>, Chia-Chih Chang <sup>1</sup>, and Kuan-Yu Tseng <sup>1</sup>, Bo-Sheng Wang <sup>2</sup>, Jiun-In Guo <sup>1</sup><sup>1</sup> Institute of Electronics, National Yang Ming Chiao Tung University<sup>2</sup> Institute of Electronics, National Chiao Tung University

\* Correspondence: tzuhsien54120@faculty.nctu.edu.tw

**Abstract:** A method of direction-of-arrival (DoA) estimation for FMCW radars is presented. In addition to MUSIC and ESPRIT, which are well-known high-resolution DoA estimation algorithms, deep learning has recently emerged as a very promising alternative. It is proposed in this paper to use a 3D convolutional neural network (CNN) for DoA estimation. The 3D-CNN extracts from the radar data cube spectrum features of the region of interest (RoI) centered on the potential positions of the targets, thereby capturing the spectrum phase shift information, which corresponds to DoA, along the antenna axis. Finally, the results of simulations and field experiments are provided to demonstrate the superior performance, as well as the limitations, of the proposed 3D-CNN.

**Keywords:** FMCW Radar; Deep Learning; Three-Dimension Convolution Network; Direction-of-Arrival Estimation

## 1. Introduction

Radar is considered to be one of the key sensor technologies for enabling autonomous driving [1]; in particular mmW radar becomes very popular because it has advantages, especially compared to LiDAR (Light Detection And Ranging), such as small size, low transmit power, excellent range resolution, low clutter sensitivity, as well as it is relatively inexpensive while maintaining certain performance under almost all weather conditions [2,3]. The most commonly used waveform in low-cost mmW radar today is Frequency Modulated Continuous Wave (FMCW). By utilizing the large bandwidth available and the high carrier frequencies in automotive radar bands, good resolution can be achieved in range and speed. However, the angle resolution is limited by the number of antennas of the phased-array. Since the automotive radar sensor is a mass-produced product, cost is an important consideration, so the number of receiving antennas is usually low in the tens [4, 5].

High-resolution DoA estimation has recently become very relevant in the current field of autonomous driving and advanced driver assistant system (ADAS). Conventional high-resolution DoA estimation algorithms, such as MUSIC and ESPRIT [6-9], are based on sub-space methods. In the application of automotive radars, usually only a few numbers of measurement snapshots can be used. However, the sub-space methods can only achieve the best performance when enough snapshots are available. And because they are based on presumed signal models, the performance of these algorithms degrade fast when the signal-to-noise ratio (SNR) becomes low. In the scenarios typical in vehicular transportation, accurate DoA estimation becomes challenging [10, 11]. In addition, when the number of detected targets is unknown, and multiple targets have different radar cross-section (RCS), their performance may also be affected. The correct detection of the number of targets in the receiver signal is also an important issue in array signal processing [12].

The aim of this paper is to present a 3D-CNN for high-resolution DoA estimation for FMCW radar. Up to two targets will be distinguished and their DoAs be estimated. Region of interest (RoI) is used to manage cases with more targets. Two formats of data are used in the process: the raw radar data cube and the statistical average in the form of covariance matrices. Simulation and experiment results are provided to demonstrate the superior

performance of the proposed method over conventional methods such MUSIC as well as emerging deep learning methods [13-20]. Neural networks (NNs) have been shown more effective than MUSIC for estimating DoAs, for instance, when SNRs are relative low. But either no convincing experiment results have been shown to support the claims [14-20], or the angular differences between targets are not small enough to show the supremacy of deep-learning-based methods [20]. In this paper, experiment results in addition to simulation results are presented to demonstrate the superior performance of the proposed 3D-CNN. To overcome the issue of lacking the diversity of range in the training data set, a data augmentation technique is also developed. For scenarios of two targets with small angular differences, our 3D-CNN shows very promising results of resolving the angles for these critical cases. In addition, it is relatively easy to extend the proposed approach for processing the data cube to obtain the target's other parameters in the future.

## 2. Materials and Methods

The task of estimating DoAs can be divided mainly into two stages, the first being signal preprocessing and the second being DoA estimation by CNN. The two stages will be described in the following two subsections. Furthermore, details of data generation and data augmentation in simulation/experiment are also provided.

### 2.1. Signal Model and Signal Preprocessing

In this paper the conventional model of plane waves impinging a Uniform Linear Array (ULA) is adopted [21]. Suppose that plane waves reflected from  $N$  targets with complex amplitudes  $S_n(t)$  and wavelength  $\lambda$  come from directions  $\theta_n$ ,  $n = 1, \dots, N$ . The ULA contains  $M$  antennas spaced by  $d$ . After demodulation, the received signal of the ULA at snapshot time  $t$  can be formulated as

$$\mathbf{x}(t) = A\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

where  $\mathbf{x}(t)$ ,  $\mathbf{s}(t)$ , and  $\mathbf{n}(t)$  are

$$\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T, \mathbf{s}(t) = [S_1(t), \dots, S_N(t)]^T, \mathbf{n}(t) = [n_1(t), \dots, n_M(t)]^T. \quad (2)$$

In the above equations,  $\mathbf{s}(t)$  (an  $N \times 1$  vector) denotes the reflected signals coming from the far-field DoAs  $\theta_1, \dots, \theta_N$ , and  $\mathbf{n}(t)$  is an  $M \times 1$  vector of independent additive white Gaussian noise (AWGN) with zero mean and variance of  $\sigma^2$ .

Define an  $M \times N$  matrix  $A$  whose columns are steering vectors toward  $N$  DoAs:

$$A = [a(\theta_1), \dots, a(\theta_N)], \quad (3)$$

with

$$a(\theta_n) = [1, e^{-j\eta_n}, e^{-j2\eta_n}, \dots, e^{-j(M-1)\eta_n}]^T, \quad (4)$$

where  $a(\theta_n)$  is the steering vector and  $\eta_n = 2\pi d \sin(\theta_n)/\lambda$  is the phase difference between adjacent elements. The demodulated signals contain enough information to estimate the DoAs.

Many DoA estimation methods use the  $M \times M$  covariance matrix of  $\mathbf{x}(t)$ :

$$R_x = E[\mathbf{x}(t)\mathbf{x}^H(t)] = AR_sA^H + R_N, \quad (5)$$

where  $[\cdot]^H$  and  $E[\cdot]$  denotes the conjugate transpose and expectation operator.  $R_s$  and  $R_N$  are the  $K \times K$  signal and noise covariance matrices, respectively. It is assumed that noise components are AWGN and the covariance matrix is

$$R_N = \sigma^2 I, \quad (6)$$

where  $I$  is the identity matrix and  $\sigma^2$  is the noise power. In practice,  $R_x$  is unknown, and the number of sampled signal vectors is limited. The maximum likelihood estimation of the spatial covariance matrix can be obtained from  $L$  snapshots by :

$$\hat{R}_x = \frac{1}{L} \sum_{l=1}^L \mathbf{x}(l) \mathbf{x}(l)^H. \quad (7)$$

The demodulated signal vectors contain sufficient information for DoA estimation. However, the size of data vectors can be huge that the training of NNs is greatly complicated. In radar applications, the signals reflected from multiple targets arrive at the receiver at different times. It is desirable to use this fact to first do a coarse separation of target signals such that the training can be simplified. Therefore, RoI will be first established to limit the coverage of the NN, and target signals with very close ranges and DoAs in the range-azimuth spectrogram will be grouped into the same RoI for further processing.

FMCW radar is chosen to further build the signal model due to its popularity in vehicular applications [1,2], especially the signal framework of TI mmW radar AWR1843BOOST+DCA100EVM is adopted [22]. The process of generating data cube [23] and selecting RoI in the proposed scheme is summarized in Fig. 1 and 2. First, multiple snapshots are collected and each is transformed by DFT to generate range information. Afterwards all snapshots are stacked along the antenna and chirp axis to form a data cube for further processing. The selection of RoI is done by first detecting possible targets on the range axis by using the cell-averaging constant false alarm rate (CA-CFAR) principle [24-26]. Then the  $J$  samples centered at the target will be claimed as RoI. If needed, a coarse DoA/speed interval can also be obtained by conventional DoA/speed estimation methods such as angle/speed DFT along the antenna/chirp axis [23]. The three-dimensional RoI of the data cube can be then cropped and further manipulated as the input of the deep learning framework.

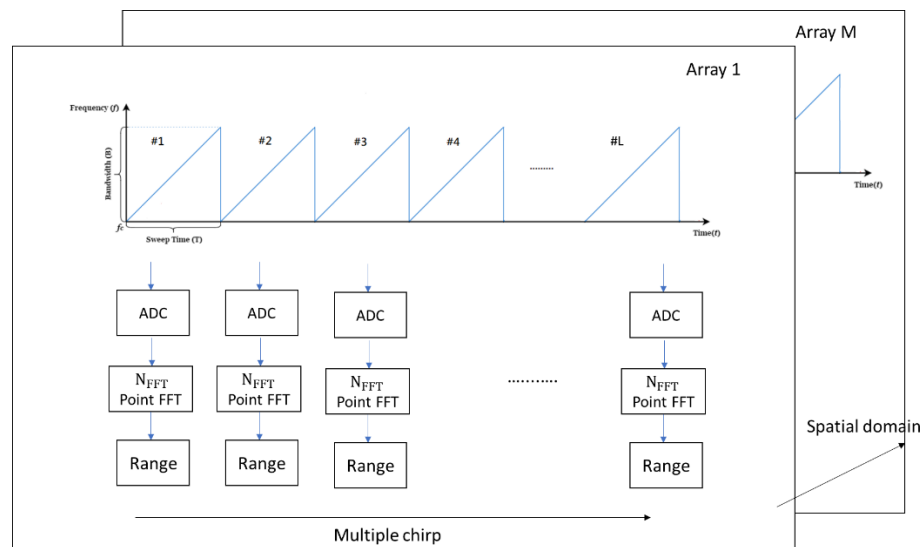
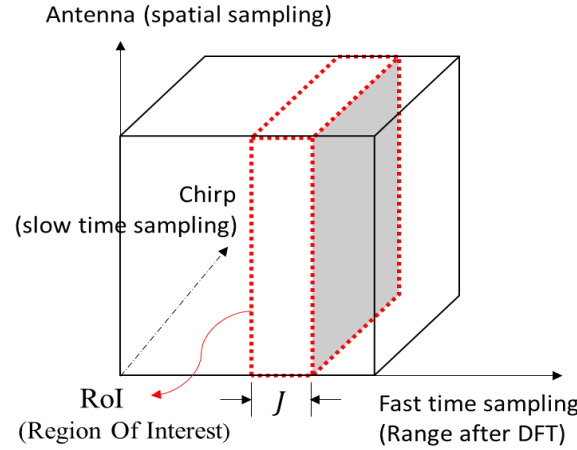


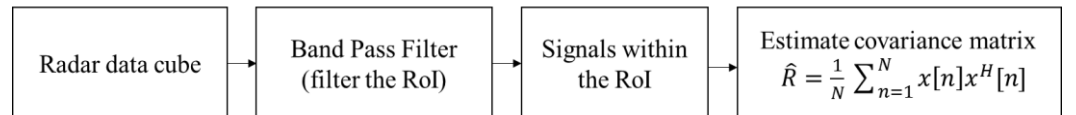
Figure 1. Generation of a data cube.



**Figure 2.** RoI selection in a data cube.

The data cube is complex-valued. The NN will take the absolute value, the real and the imaginary part of the data as the input for DoA estimation. A similar data format has been found to be effective for training neural networks [27]. In addition to the RoI of the data cube, the covariance matrix is also used as the input for determining the number of targets. The details will be given in the next subsection. In order to reduce the impact from noise outside the RoI, the covariance matrix is estimated with only the signal within the RoI. As shown in Fig. 3, the data is first filtered by a band-pass filter fitted to the RoI, then the maximum likelihood estimation  $\hat{R}_x$  in (7) is obtained. Since  $\hat{R}_x$  is Hermitian, the diagonal terms and upper triangular part of it contains all the information. Let  $\hat{R}_x$  be first normalized so that its real and imaginary parts are within  $[-1,1]$  and then be morphed into a vector  $\mathbf{R} \in \mathbb{R}^{M(M-1) \times 1}$  for the input to a NN:

$$\mathbf{R} = [Re(\hat{R}_{12}), Im(\hat{R}_{12}), \dots, Im(\hat{R}_{1M}), \dots, Re(\hat{R}_{23}), \dots, Im(\hat{R}_{(M-1)M})] \quad (8)$$



**Figure 3.** The covariance matrix of RoI.

## 2.2. DoA Estimation by CNN

Deep learning has breached into many application areas where conventional signal processing techniques are facing difficulties; among numerous neural network models, CNNs are known to be very capable of extracting features from matrix- or tensor-formed data [28-31]. Its architecture typically consists of one or more convolution layers followed by fully-connected networks leading to the output. In the proposed scheme, because a 3-D data cube is processed to make high-precision DoA estimation, a 3D-CNN [30] is deployed. But before the 3D-CNN, a simple fully-connected NN is first used for determining the number of targets. The overall signal flow together with the signal preprocessing is summarized in Fig. 4, and more details are itemized in the following.

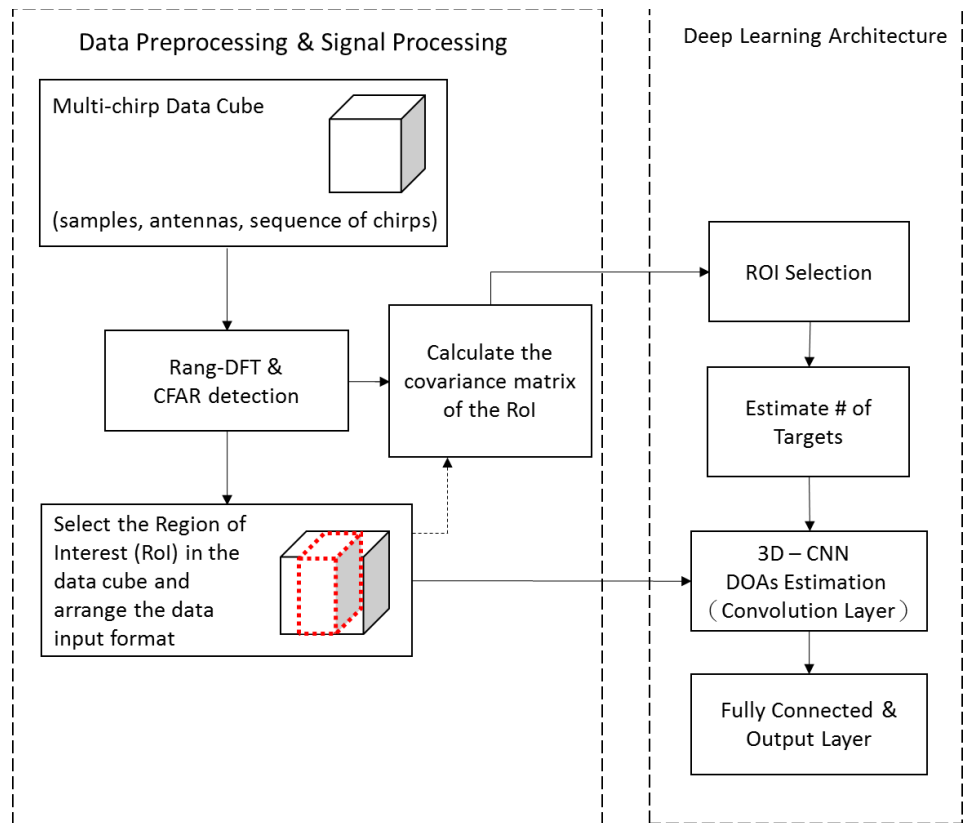


Figure 4. Signal processing and deep learning architecture.

1. Signal preprocessing and the selection of RoI have been described in the previous subsection.
2. An independent small-size fully -connected NN is used to estimate the number of targets. The signal's statistical average is found to be more robust to provide coarse information, such as the number of targets; while the data cube contains subtle information without being averaged out so that it enables high-resolution DoA estimation. The input of the simple fully-connected NN to estimate the number of targets is an  $M \times M$  covariance matrix (in our case,  $M = 8$ ), and the covariance matrix is obtained by (8). The estimated number (0, 1, or 2) will then be used to choose the one-target or two-target 3D-CNN for estimating the DoAs.
3. The 3D-CNN does DoA estimation by treating it as a classification task [32]. There are twenty one classes (corresponding to DoAs  $-10^\circ, -9^\circ, \dots, 10^\circ$ ) to be classified. In Fig. 4, the highlighted RoI is the input to the 3D-CNN and it is a  $J \times M \times C \times 3$  tensor (in our case,  $J = 65, M = 8, C = 4$ , which stand for number of fast-time samples, of antennas, and of chirps respectively). The 3 is the number of data formats that includes the absolute value, the real part, and imaginary part; using the arrangement is shown to yield better performance than only using the real and imaginary part [27]. The first convolutional layer in the 3-D CNN (Conv1) uses  $30 \times 5 \times 3$  kernels with the stride step of  $1 \times 1 \times 1$ , and the number of channels is 256. The same number of channels are passed to the second convolutional layer (Conv2), which uses  $2 \times 2 \times 2$  kernels, and a max pooling (MP) layer follows. The output of feature extraction is then flattened and connected into the form of a matrix, and a four-layer fully connected layer is used for its classification. The loss function used for the classification task is binary cross entropy in Pytorch environment [33].
4. The parameters for the fully-connected NN and 3D-CNN are listed in Table 1 and 2. The term FC stands for Fully Connected and BN for Batch Normalization [34]. In the

training of this fully-connect NN and the 3D-CNN, the popular optimization tool ADAM and the drop-out technique are used [35, 36].

**Table 1.** Fully-connected NN parameters.

Layer	# of Filter	Activation	Layer	# of Filter	Activation
Input	56		FC3+BN	1024	ReLu
FC1+BN	1024	ReLu	FC4+BN	1024	ReLu
FC2+BN	1024	ReLu	FC5+BN	3	Sigmoid

**Table 2.** 3D-CNN parameters.

Layer	Filter	# of filter	Activation	Output
RoI spectrum				$65 \times 8 \times 4 \times 3$
Conv1	(30,5,3)	256	ReLu	
Conv2	(2,2,2)	256	ReLu	
MaxPool 1	(2,2,2)			
FC			ReLu	500
FC			ReLu	200
FC			ReLu	100
FC			Sigmoid	21

2.3. Data Generation and Data Augmentation

Both simulation data and experiment data are acquired for training the NNs and evaluating the proposed DoA estimation scheme. The simulation data is primarily for evaluating and comparing the performance of DoA estimation with the Cramer-Rao Bound (CRB) [37] and existing algorithms for a wide range of SNRs. The experiment data, on the other hand, is primarily for demonstrating the applicability of the proposed DoA estimation scheme.

Simulation data is generated by the MATLAB Phased Array tool box [38] with the parameters delineated below. To simplify the performance analysis, the number of targets is set to be 0, 1, or 2. The radar is an FMCW one with 8 receiving antennas. The resonance frequency is 77 GHz and one snapshot received at each antenna has  $L = 256$  sample points. The DoA of each incident plane wave ranges from  $-10^\circ$  to  $10^\circ$ , and signals under various SNRs ranging from -10 to 15 dB are generated. Detailed simulation parameters are summarized in Table 3. In total, 63k snapshots are generated for training and 2.1k for each testing in different SNR. Note that all the pairs of DoA ( $[-10^\circ, -9^\circ]$ ,  $[-10^\circ, -8^\circ]$ , ...,  $[9^\circ, 10^\circ]$ ) are considered in both training and testing.

In simulation, the model is first trained by the first training data set with SNRs in the range from 15 to 0 dB; then the model is further strengthened by the second training data set with SNRs from 15 to -5 dB. Number of the both training data is 63k. The reason for doing the two-stage training is to make sure the model first learns the correct signal features with relatively high-SNR data then make the model more robust by training it with more noisy data.

**Table 3.** Simulation parameters.



Parameters	Value	Parameters	Value
Transmit signal	FMCW	Target Number	0, 1, or 2
Carrier Frequency $f_c$	77 GHz	Target Range	70 m to 90 m
Frequency Slope $B$	3.476 THz/s	Target Angle	-10° to 10°
Bandwidth	0.15 GHz	Samples per chirp	256
Tx/Rx Antenna	1/8	SNR	-10 dB to 15 dB

Experiments with metal corner reflectors mounted on tripods on a baseball field (please see Fig. 5) are conducted to collect data. The TI mmW radar AWR1843BOOST+DCA100EVM is used [39]; it has 2 Tx and 4 Rx antennas to constitute a virtual 8-antenna receiving array. Detailed experiment parameters are summarized in Table 4. Note that some radar parameters used in simulation and experiments are not exactly the same; it is due to the limitations of both the MATLAB Phased Array tool box and the TI radar platform. The training data set for experiment includes both 21k raw snapshots (directly read from the TI radar) and 28k snapshots with data augmentation, and about 7k raw snapshots are reserved for testing. Details and methods of data augmentation are provided below. The reason of using augmented data is to increase the robustness of our trained model.



**Figure 5.** Experiments with two metal corner reflectors.

**Table 4.** Experiment parameters.

Parameters	Value	Parameters	Value
Transmit signal	FMCW	Target Number	0, 1, or 2
Carrier Frequency $f_c$	77 GHz	Target Range	10 m
Frequency Slope $B$	4.2 THz/s	Range Augmentation	9.7 m to 10.3 m
Bandwidth	3.5 GHz	Target Angle	-10° to 10°
Tx/Rx Antenna	2/4	Samples per chirp	512

In addition to the data collected in the field experiment, data augmentation is deployed to increase the quantity and enhance the range diversity of the training data set. We suspect that the lack of reporting high-resolution results with using NNs on experiment data in literature might be just due to the robustness issue caused by insufficient range diversity. The experiments are conducted with the targets placed only at certain ranges. Nonetheless, real-world targets come with random ranges. Therefore, it is essential to expose the NNs to targets with a variety of ranges. One way to do it without trying

to place targets at a multitude of ranges is implementing data augmentation. Note that the demodulated FMCW signal along the fast-time axis (in a data cube shown in Figure 2) exhibits a beat frequency corresponding to the target's range:

$$p(t) = e^{j2\pi[(\frac{2BR}{cT})t + \frac{2f_c R}{c}]} \quad (9)$$

where  $p(t)$  is the demodulated FMCW chirp signal,  $R$  is the target's range,  $f_c$  is the carrier frequency,  $B$  is the bandwidth,  $T$  is the chirp duration, and  $c$  is the speed of light. Therefore, it only takes a proper frequency term to change the target's range. As a result, we only need to multiply a phase term to signals along the fast-time axis to create virtual targets at the range  $R + \Delta R$ , where  $\Delta R$  is randomly chosen from -0.3 to 0.3 m:

$$p_{\text{aug}}(t) = e^{j2\pi[(\frac{2B\Delta R}{cT})t + \frac{2f_c \Delta R}{c}]} \times p(t). \quad (10)$$

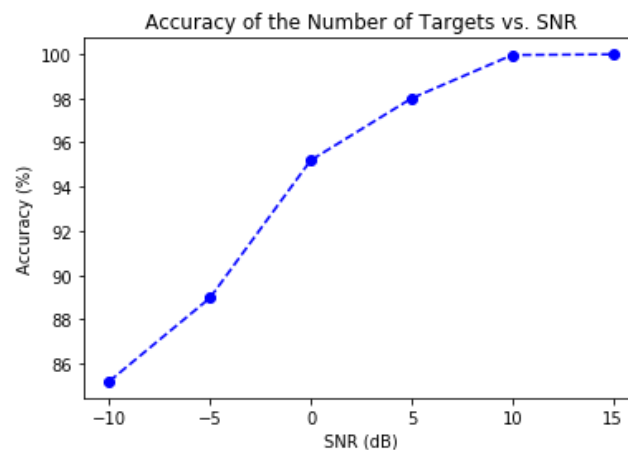
From each angle set, there are several snapshots chosen to be augmented, and 220 different  $\Delta R$ 's are generated to augment the size of data set to about 28k.

### 3. Results

In this section, the simulation results of evaluating and comparing the performance of the proposed DoA estimation method with the Cramer-Rao Bound (CRB) [37] and the DNN (Deep NN) in [19], which has the state-of-the-art performance among deep-learning-based method methods, is first presented. Then the experiment results with up to two corner reflectors are presented to demonstrate the potential of the proposed method in applying to real-world situations. Note that the proposed DoA estimation method consists of mainly two steps: the first is to determine with a fully-connected NN the number of targets in the RoI, while the second is to estimate the targets' DoAs with a 3D-CNN designed either for one target or for two targets. Therefore, for both simulation and experiment, the accuracy of estimating the number of targets is first characterized, followed by the results on DoA estimation.

#### 3.1. Simulation Results

Fig. 6 shows the accuracy of determining the number of targets. As can be seen, the accuracy reaches 100% once the SNR hits 5 dB, and it almost maintains 90% even when the SNR is as low as -10 dB. The highly effective fully-connected NN provides a very ensuring condition for the 3D-CNN to work on DoA estimation.

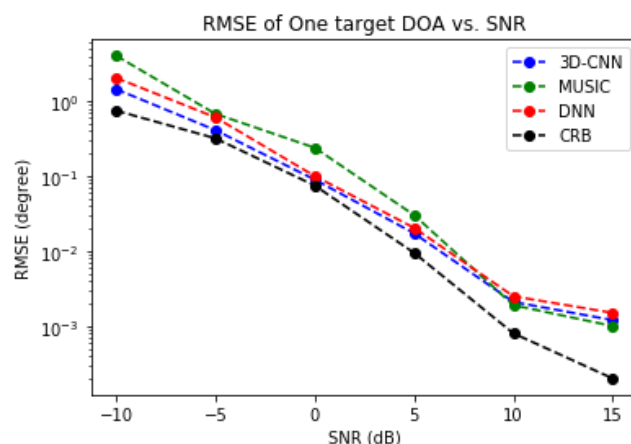


**Figure 6.** The accuracy of determining the number of targets vs. SNR.

Fig. 7 presents the performance of DoA estimation for the case of one target. The root-mean-square error (RMSE) of the proposed method, MUSIC, the DNN in [19], and the CRB are compared. It is clear that both NN methods beat MUSIC and stay close to the

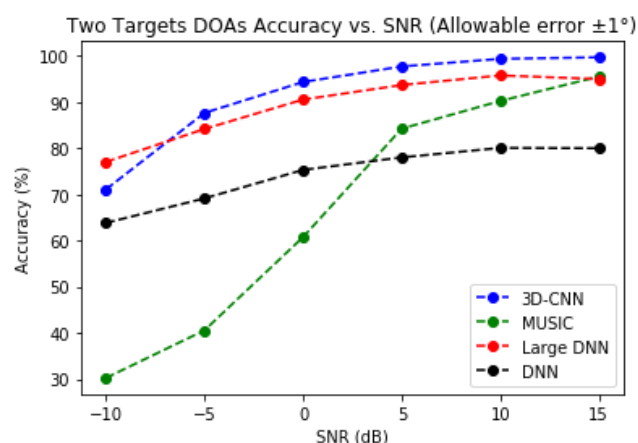


CRB across a wide range of SNRs. The proposed 3D-CNN and the DNN in [19] have similar performance and both perform better than MUSIC in most SNRs. The advantage of the 3D-CNN becomes evident in the case of two targets, as will be shown next.



**Figure 7.** The RMSE of DoA estimation vs. SNR.

Fig. 8 compares the success rate of obtaining correct DoAs for two-target cases. The success is defined with a  $\pm 1^\circ$  estimation error. Four methods are compared. The original DNN proposed in [19] has about 0.135 million parameters, e.g. it is a very small DNN, especially compared to our 3D-CNN. It is no surprise that 3D-CNN has a clear advantage over the whole range of SNRs. However, the performance of the DNN improves when its size is increased. Therefore, we increase its size until the performance tops (after that large DNNs perform poorer due to difficulties in training), and the result is a large DNN with around 11 million parameters. As can be seen from Fig. 8, the 3D-CNN still has the best overall performance, even though the large DNN has diminished the performance gap to 3 dB. It is also worth noting that at SNR = -10 dB, the 3D-CNN performs worse because the training data set does not contain data from this SNR. It shows that 3D-CNN is somewhat less robust than the large DNN.

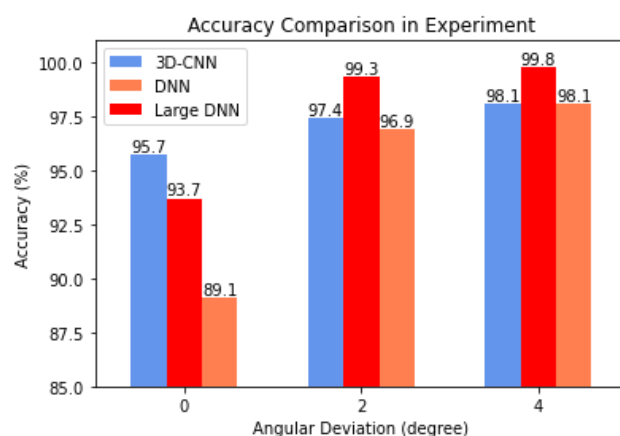


**Figure 8.** Accuracy of two-target DoA estimation, with an allowable error of  $\pm 1^\circ$ .

### 3.2. Experiment Results

One-target and two-target experiments are conducted with the experiment set-up presented in Sec. 2.3. In one-target scenarios, all the three methods (3D-CNN, DNN, and MUSIC) estimate DoA correctly with  $1^\circ$  precision. Fig. 9 shows the comparison of three NNs in terms of the accuracy of DoA estimation. The figure gives out three sets of accuracy, i.e. with angular deviation of  $0^\circ$ ,  $2^\circ$ , and  $4^\circ$ . Notice that MUSIC has been excluded

from the comparison; it is because that MUSIC often fails to identify two DoAs and is not qualified to compete with these methods. In two-target scenarios, with the tolerable angular deviation of  $0^\circ$ , the accuracy of 3D-CNN reaches 95.7%; while for the large DNN and DNN, the accuracy results are 93.7% and 89.1% respectively. When larger angular deviation is allowed, however, the accuracy of large DNN again goes to the top. This again shows that DNN has more robust performance.



**Figure 9.** Accuracy of two-target DoA estimation vs. angular deviation.

The 3D-CNN not only has better accuracy when zero angular deviation is allowed, it also excels in the critical cases where the angular difference between two targets is small. When the angular difference is less than  $4^\circ$ , the 3D-CNN has an accuracy of 97.8 %, while the DNN and the large DNN has 88.6% and 95.6% respectively.

In summary, the 3D-CNN has the best performance, especially when the situation is critical. On the other hand, the slight decline in its performance when facing low-SNR data that is not experienced in the training phase shows that its performance, like many data-driven methods do, highly depend on the quality and quantity of the training data set available.

#### 4. Discussion

DoA estimation is a classical problem for radar data processing. High-resolution methods such as MUSIC have been available for decades. Novel methods based on different types of NNs are widely regarded as the more promising approach for the old problem. In this paper we have demonstrated, as illustrated by Fig. 6-9, that two methods (3D-CNN and DNN) can consistently beat MUSIC with a large margin under both simulated and experiment scenarios. The 3D-CNN also has a better overall performance than the DNN, and it shows the potential of working with the RoI approach to generate high-resolution radar point clouds.

It may be argued that the better performance of 3D-CNN is due to the fact that its input data-format (3-D tensor) retains more information from the raw radar data than the DNN's input (covariance matrix) does. On the other hand, the stability of data benefitting from the averaging operation in estimating the covariance matrix gives a little advantage to the DNN, as demonstrated by the slightly more robust performance of DNN under some conditions.

Another advantage of 3D-CNN is its better potential in future development. For instance, its input data contains the information along the slow time; therefore it is possible to train the same network for estimating target velocity, i.e. using the same 3D-CNN to output target's range, velocity and DoA simultaneously is possible.

Finally, not only the input data format plays an essential role in model learning, data augmentation is essential too. In acquiring the training data set, it is very difficult to conduct experiments covering comprehensively the range of targets. The neural networks,

whether it is 3D-CNN or DNN, would perform poor if the range information in the training set is not diverse enough. Eq. (10) offers a way to cover range diversity in doing data augmentation for preparing a comprehensive training data set.

**Author Contributions:** Conceptualization, T.H.S. and F.T.C.; methodology, T.H.S. and F.T.C.; software, C.C.C.; investigation, all authors; data curation, B.S.W., C.C.C., and K.Y.T.; writing—original draft preparation, T.H.S., F.T.C., and B.S.W.; writing—review and editing, T.H.S. and F.T.C.; visualization, B.S.W., C.C.C., and K.Y.T.; supervision, T.H.S., F.T.C., and J.I.G.; project administration, T.H.S., F.T.C., and J.I.G.; funding acquisition, T.H.S., F.T.C., and J.I.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is partially supported by the “Center for mmWave Smart Radar Systems and Technologies” under the “Featured Areas Research Center Program” within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE), Taiwan, R.O.C., and partially supported under MOST projects with grants MOST 108-3017-F-009-001 and MOST 110-2634-F-009-028.

**Data Availability Statement:** Training and testing data sets for the neural network models are available at: [https://drive.google.com/drive/u/1/folders/1JGfP9RrxgR1IozXMxpdOmnaoUaq\\_vdC](https://drive.google.com/drive/u/1/folders/1JGfP9RrxgR1IozXMxpdOmnaoUaq_vdC)

**Acknowledgments:** The authors would like to thank Jia-Jheng Lin and Jian-Xian Lu of Prof. Jiun-In Guo’s research group at Institute of Electronics, National Yang Ming Chiao Tung University (NCTU) for their generous support in so many aspects.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Wenger, J. Automotive radar - status and perspectives. *IEEE Compound Semiconductor Integrated Circuit Symposium CSIC '05*, **2015**, pp.21 pp.-24.
- Hasch, J.; Topak, E.; Schnabel, R.; Zwick, T.; Weigel, R.; Waldschmidt, C. Millimeter-Wave Technology for Automotive Radar Sensors in the 77 GHz Frequency Band. *IEEE Transactions on Microwave Theory and Techniques*, **2012**, vol. 60, no. 3, pp. 845-860.
- Sang, T.-H.; Tsai, S.; Yu, T., Mitigating Effects of Uniform Fog on SPAD Lidars, *IEEE Sensors Letters*, **2020**, vol. 4, no. 9, pp. 1-4.
- Kukkala, V.K.; Tunnell, J.; Pasricha, S.; Bradley, T. Advanced Driver Assistance Systems: A Path Toward Autonomous Vehicles. *IEEE Consumer Electronics Magazine*, **2015**, vol. 7, no. 5, pp. 18-25.
- Rohling, H. Milestones in radar and the success story of automotive radar systems. *11-th INTERNATIONAL RADAR SYMPOSIUM*, **2010**, pp. 1-6.
- Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, **1969**, vol. 57, no. 8, pp. 1408-1418.
- Barabell, A. Improving the resolution performance of eigenstructure-based direction-finding algorithms. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, **1983**, vol. 8, pp. 336-339.
- Schmidt, R. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, **1986**, vol. 34, no. 3, pp. 276-280.
- Roy, R.; Kailath, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **1989**, vol. 37, no. 7, pp. 984-995.
- Weiss, A. J.; Friedlander, Benjamin. Direction finding in the presence of mutual coupling. *IEEE Transactions on Antennas and Propagation* **1991**, vol. 39, no. 3, pp. 273-284.
- Krim, H.; Viberg, M. Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.*, **1996** vol. 13, no. 4, pp. 6794.
- Chen, P.; Wu, T.-J.; Yang, J. comparative study of model selection criteria for the number of signals. *IET Radar, Sonar & Navigation*, **2008**, vol. 2, no. 3, pp. 180-188.
- Elbir, A. M. DeepMUSIC: Multiple Signal Classification via Deep Learning. *IEEE Sensors Letters* **2020**, vol. 4, no. 4, pp. 1-4.
- Kase, Y.; Nishimura, T.; Ohgane, T.; Ogawa, Y.; Kitayama, D.; Kishiyama, Y. Doa estimation of two targets with deep learning. *IEEE 15th Workshop on Positioning, Navigation and Communications (WPNC)*, **2018**, pp. 1-5.
- Tang, G.; Gao, X.; Chen, Z.; Zhang, Y.; Zhong, H.; Li, M. Deep Neural Network Based Multiple Targets DOA Estimation for Millimeter-Wave Radar. *IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation* **2019**, pp. 433-438.
- Chakrabarty, S.; Habets, E. A. P. Multi-Speaker DOA Estimation Using Deep Convolutional Networks Trained with Noise Signals. *IEEE Journal of Selected Topics in Signal Processing* **2019**, vol. 13, no. 1, pp. 8-21.
- Ling, Y.; Gao, H.; Zhou, S.; Yang, L.; Ren, F. Robust Sparse Bayesian Learning-Based Off-Grid DOA Estimation Method for Vehicle Localization. *Sensors* **2020**, 20, 302.
- Gall M.; Gardill M.; Fuchs J.; Horn T. Learning Representations for Neural Networks Applied to Spectrum-Based Direction-of-Arrival Estimation for Automotive Radar, *2020 IEEE/MTT-S International Microwave Symposium (IMS)*, **2020**, pp. 1031-1034.

19. Chen, M.; Gong, Y.; Mao, X. Deep Neural Network for Estimation of Direction of Arrival with Antenna Array, *IEEE Access* **2020**, vol. 8, pp. 140688-140698.
20. Cong J.; Wang X.; Huang M.; Wan L. Robust DOA Estimation Method for MIMO Radar via Deep Neural Networks, *IEEE Sensors Journal*, vol. 21, no. 6, **2021**, pp. 7498-7507.
21. Stoica, P.; Moses, R. L., *Introduction to Spectral Analysis*, 1st ed.; Prentice-Hall: New Jersey, USA, 1997; pp. 221-247.
22. Rao, S.; Ahmad, A.; Roh, J. C.; Bharadwaj, S. 77GHz single chip radar sensor enables automotive body and chassis applications. *Texas Instruments, Dallas, TX, TI Application Note n. SPRY315*, **2017**.
23. Patole, S. M.; Torlak, M.; Wang, D.; Ali, M. Automotive radars: A review of signal processing techniques. *IEEE Signal Processing Magazine* **2017**, vol. 34, no. 2, pp. 22-35.
24. Rohling, H. Radar CFAR Thresholding in Clutter and Multiple Target Situations. *IEEE Transactions on Aerospace and Electronic Systems* **1983**, vol. AES-19, no. 4, pp. 608-621.
25. Rohling, H.; Mende, R. OS CFAR performance in a 77 GHz radar sensor for car application. *Proceedings of International Radar Conference* **1996**, pp. 109-114.
26. Rohling, H. Ordered statistic CFAR technique – an overview. *12th International Radar Symposium (IRS)*, **2011**, pp. 631-638.
27. Elbir, A.M.; Mishra, K.V. Joint Antenna Selection and Hybrid Beamformer Design Using Unquantized and Quantized Deep Learning Networks. *IEEE Transactions on Wireless Communications* **2020**, 19, pp. 1677-1688.
28. Haykin, S. *Neural Networks and Learning Machines*. Third Edition, Prentice hall New York, **2009**.
29. Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagating errors. *Nature* **1986**, vol. 323, pp.533-536.
30. Ji, S.; Xu, W.; Yang, M.; Yu, K., 3D Convolutional Neural Networks for Human Action Recognition, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2013**, vol. 35, no. 1, pp. 221-231.
31. Sang, T.-H.; Xu, Y.-C., Clipping noise compensation with neural networks in OFDM systems, in *Signals*, **2020**, vol. 1, no. 1, pp. 100-109.
32. El Zooghby, A. H.; Christodoulou, C. G.; Georgiopoulos, M. A neural network-based smart antenna for multiple source tracking. *IEEE Transactions on Antennas and Propagation*, **2000**, vol. 48, no. 5, pp. 768-776.
33. Binary Cross Entropy Documentation. Available online: [https://pytorch.org/docs/master/generated/torch.nn.functional.binary\\_cross\\_entropy.html#torch-nn-functional-binary-cross-entropy](https://pytorch.org/docs/master/generated/torch.nn.functional.binary_cross_entropy.html#torch-nn-functional-binary-cross-entropy) (accessed on 23 May 2021).
34. Ioffe, S.; Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning* **2015**, 37, pp.448-456.
35. Kingma, D. P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, **2014**.
36. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, **2014**, 15, pp. 1929-1958.
37. Stoica, P.; Nehorai, A. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **1989**, vol. 37, no. 5, pp. 720-741.
38. MATLAB Phased Array System Toolbox Documentation. Available online: <https://www.mathworks.com/help/phased> (accessed on 7 April 2021).
39. AWR1843 Single-Chip 77- to 79-GHz FMCW Radar Sensor. Available online: [https://www.ti.com/lit/ds/sym-link/awr1843.pdf?ts=1594818635366&ref\\_url=https%253A%252F%252Fwww.google.com%252F](https://www.ti.com/lit/ds/sym-link/awr1843.pdf?ts=1594818635366&ref_url=https%253A%252F%252Fwww.google.com%252F) (accessed on 7 April 2021).