Fusion Model Using a Neural Network and MLE for a Single Snapshot DOA Estimation with Imperfection Mitigation

Marcio L. Lima de Oliveira
Computer Architecture for Embedded Systems
University of Twente
Enschede, The Netherlands
m.l.limadeoliveira@utwente.nl

Marco J. G. Bekooij
Department of Embedded Software and Signal Processing
NXP Semiconductors, University of Twente
Eindhoven, The Netherlands
marco.bekooij@nxp.com

***Abstract*—In this paper, we discuss the fusion of deep neural networks with the Maximum Likelihood Estimation (MLE) algorithm for estimating both the Direction Of Arrival (DOA) of MIMO radars and the number of sources for a single-snapshot scenario using a Minimum Redundancy antenna Array (MRA) with imperfections and evaluating it with synthetic and real-world data. The combination of Deep Learning (DL) with the classical MLE seems to be a viable solution for this problem, as it is less computationally intensive than a classical MLE while not losing generalization and being better at estimating the number of sources. In both our experiments, using synthetic and real-world data, our method performs close to MLE and appears to be a deployable solution for real scenarios. Besides reducing the computational load of MLE, the novelty of our model lies in the fact that the deep learning model learns the gain and phase errors as a function of the direction of arrival instead of applying a simple calibration.**

***Index Terms*—Direction of Arrival, Maximum Likelihood Estimation, Deep Learning, Machine Learning, Antenna Imperfection Mitigation**

I. Introduction

The estimation of the Direction Of Arrival (DOA) is a common challenge for various signal processing areas, like acoustics, radars, sonars, wireless communication, and others. For this work, we have considered the problem of calculating multiple DOAs of MIMO radars using a Minimum Redundancy antenna Array (MRA), taking antenna imperfections into account.

Throughout recent years, the employment of Deep Learning (DL) for signal processing has increased exponentially. Fall detection [1], sign-language recognition [2], and sensor fusion [3] are among a few examples of that.

In this paper, we expanded our previous works [4], [5], where we have explored the usage of DL techniques, together with Maximum Likelihood Estimation (MLE), to estimate DOAs using synthetic snapshots and the number of sources. While in our previous works we have only used perfect antennas and synthetic data, in this work, we have introduced imperfections to our models, adapting our previous frameworks accordingly. For evaluating this novel work, we have used both synthetic and real-world data. The investigation of imperfect antennas is important since perfect models are often too simplistic for real applications. The choice of using a non-linear antenna array (MRA) in our previous and current works is to enlarge the resolution without adding extra antennas [6].

In addition, the novelty of our approach lies in the fact that the neural network model learns that the gain and phase errors are, in fact, functions of the direction of arrival. This means that the underline imperfection mitigation applied by our model does not try to calibrate just for the current scenario but considers the angles of the targets, while also estimating the DOAs.

We have utilized the Maximum Likelihood Estimation (MLE) algorithm as a comparison method and within our framework, as proposed in [4]. For this, we utilize a data-driven fusion method to reduce the computational load of MLE.

According to Hacker and Yang [7], for a single-snapshot MRA scenario, MLE is a robust and efficient technique to estimate the DOA, being the most effective they have tested. Subspace-based algorithms tend to have a poor performance when using fewer snapshots and the number of targets we are trying to estimate [8]. In addition, MRA may cause larger side-lobes that can be challenging for such algorithms [6].

The MLE algorithm works by trying all possible angle combinations and choosing the one that would minimize the residual error (difference between estimated and actual snapshots) for the arriving signal, given a previously known antenna model [9]. It is important to note that we have used a traditional implementation of the MLE algorithm for estimating DOAs [10], rather than a different MLE algorithm [11], [12]. The main problem of the traditional implementation of MLE is its computational complexity, which grows exponentially with the increase in the number of targets [13].

Other recent works have also developed various techniques to estimate DOAs using DL [14], [15]. The researchers of [16] have developed a framework called DeepDOA that uses deep learning to obtain a smoother spectrum in order to estimate the DOAs. In the work of [17], they have used a hybrid MB/DD DOA estimation architecture based on the MUSIC algorithm using neural networks. Ma et. al. [18] have used what they called a deep learning DOA classifier to detect up to 11 targets within a given field of view region.

Unlike most of these data-driven DOA estimation techniques, our work demonstrates the capability of our model to estimate DOAs and the number of targets using real-world data, which is a significant improvement as several data-driven DOA estimation techniques lack experimental validation.

II. Signal Model

Consider M antennas not uniformly separated (MRA), with the minimum distance between two antennas being $d = \lambda/2$, in which λ is the signal's wavelength. We can construct an array $D \in \mathbb{R}_+^M$, containing all the distances between the first antenna and the other antennas as $D_m = [0, 1, 4, 6, 13, 14, 17, 19] \times d$, where $m = 1, ..., M$. Moreover, considering that all N sources are at far-field, where $0 \leq N \leq M - 1$. The received snapshot $x(t) \in \mathbb{C}^M$ from a perfect antenna array is given as:

$$x(t) = \mathbf{A}(\theta)s(t) + \eta(t) \; , \tag{1}$$

where $s(t) \in \mathbb{C}^N$ is the signal of the sources, $\eta(t) \in \mathbb{C}^M$ is the added white noise, and $\mathbf{A}(\theta) \in \mathbb{C}^{M \times N}$ is the steering matrix. The steering matrix is a function of the direction of arrival $\theta = \theta_1, ..., \theta_N$, and it can be expressed as:

$$\mathbf{A}(\theta) = \begin{bmatrix} e^{j2\pi\lambda^{-1}D_1 \sin(\theta_1)} & ... & e^{j2\pi\lambda^{-1}D_1 \sin(\theta_N)} \\ ... & & ... \\ e^{j2\pi\lambda^{-1}D_M \sin(\theta_1)} & ... & e^{j2\pi\lambda^{-1}D_M \sin(\theta_N)} \end{bmatrix} . \tag{2}$$

However, eq. 1 assumes that the noise has a Gaussian distribution. In reality, for real antennas, this is rarely the case. Due to imperfections, received snapshots often do not have zero-mean Gaussian error distribution. Some examples of systematic errors are: deviations from the actual position of the sensor element, mutual coupling, and imperfections in the receiving hardware [19]. Calibration techniques are often applied to correct the antenna model [13], [19], [20]. In this work, we have focused on two variables to represent the errors caused by imperfections: gain $\alpha(\theta) \in \mathbb{R}^M$ and phase $\varphi(\theta) \in \mathbb{R}^M$ errors. Considering a diagonal matrix $\mathbf{E}(\alpha(\theta), \varphi(\theta)) \in \mathbb{C}^{M \times M}$ as the error matrix that contains all the gain/phase errors, it is possible to rewrite eq. 1 as:

$$x(t) = \mathbf{E}(\alpha(\theta), \varphi(\theta))\mathbf{A}(\theta)s(t) + \eta(t) \; . \tag{3}$$

The error matrix $\mathbf{E}(\alpha(\theta), \varphi(\theta))$ describes the gain (or amplitude) and phase deviations of the antennas on its main diagonal, while the other elements are equal to zero (because we did not consider the coupling between antennas). It is important to note that both $\alpha(\theta)$ and $\varphi(\theta)$ are angle dependent, as they depend on the array orientation of the beam patterns. In other words, they depend on the DOAs $\theta = \theta_1, ..., \theta_N$ of the targets. Furthermore, the elements of $\mathbf{E}(\alpha(\theta), \varphi(\theta))$ are non-linearly dependent, and the analytical expression for its elements is usually not known. In our experiments, we also did not consider that its elements are somewhat time and temperature dependent. The error matrix can be described as:

$$\mathbf{E}(\alpha(\theta), \varphi(\theta)) = \begin{bmatrix} \alpha_1(\theta)e^{j\varphi_1(\theta)} & ... & 0 \\ ... & & ... \\ 0 & ... & \alpha_M(\theta)e^{j\varphi_M(\theta)} \end{bmatrix} . \tag{4}$$

A. Maximum Likelihood Estimation

The Maximum Likelihood Estimation (MLE) algorithm is a common DOA estimating method. It is a reliable algorithm when just one snapshot is available and when using an MRA configuration [7]. The algorithm works by trial and error as it tries all possible steering matrices containing all possible angle combinations to find the one that produces the least amount of residual error $\varepsilon(\theta)$:

$$\varepsilon(\theta) = x(t) - \mathbf{A}(\theta)\mathbf{A}^+(\theta)x(t) \; , \tag{5}$$

where $\mathbf{A}^+(\theta)$ is the Moore-Penrose pseudo-inverse of the matrix $\mathbf{A}(\theta)$. We can then estimate the DOAs that produce the minimal amount of $\varepsilon(\theta)$:

$$\tilde{\theta} = \underset{\theta}{\operatorname{argmin}} \; ||\varepsilon(\theta)||^2 \; , \tag{6}$$

where $\tilde{\theta} = \tilde{\theta}_1, ..., \tilde{\theta}_N$ are the estimated angles.

The maximum likelihood estimation can also be used for estimating the number of targets in the scene through various iterations. Assuming that the background noise is either known or can be accurately estimated, the MLE algorithm can be used to compare the residual error (calculated using eq. 5) with the noise level. If the difference between the two values exceeds a predetermined threshold, it is probable that there is another target present in the scene, which is contributing to the high residual error estimate. This means that we can start the search for the number of targets by performing a 1D search and assuming that only one target exists. If the difference between the noise floor and the residual noise is higher than a threshold, we can assume that there are 2 targets and perform a 2D search. This search continues until we find the DOAs and the number of targets or until a certain limit is achieved. Since MLE is a computational complexity algorithm, this sort of search can be unrealistic for low-latency scenarios.

B. Antenna Calibration

In order to mitigate the errors of real-world antennas, it is often necessary to apply some techniques. A maximum likelihood setting can be used as a way to find the error matrix $\mathbf{E}(\alpha(\theta), \varphi(\theta))$ [19]. It works as a variation of eq. 6, but with known DOAs and unknown imperfections:

$$\tilde{\alpha}(\theta), \tilde{\varphi}(\theta) = \underset{\alpha(\theta), \varphi(\theta)}{\operatorname{argmin}} \; ||x(t) - \mathbf{P}\,\mathbf{P}^+ x(t)||^2 \; , \tag{7}$$

in which $\tilde{\alpha}(\theta) = \tilde{\alpha}_1(\theta), ..., \tilde{\alpha}_M(\theta)$ are the estimated gain errors, $\tilde{\varphi}(\theta) = \tilde{\varphi}_1(\theta), ..., \tilde{\varphi}_M(\theta)$ are the estimated phase errors, $\mathbf{P} = \mathbf{E}(\alpha(\theta), \varphi(\theta))\mathbf{A}(\theta)$, and the DOAs θ are known.

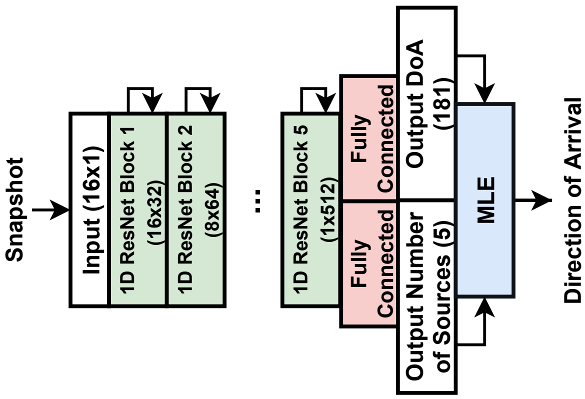


Fig. 1: The architecture of our proposed model is a mix of the architectures of our previous works [4], [5]. Our model takes a snapshot and pre-processes it as input to our neural network. Our DL model has two outputs, one with the most likely DOAs - 181 neurons, a field of view of 180 degrees, a resolution of 1 degree - and another with the most likely number of sources - 5 neurons, 0 to 4 sources. Both outputs are fed to an MLE algorithm to obtain the estimated DOAs.

III. Data and Architecture

In this paper, two kinds of data were used: synthetic and real. First, we wanted to see if a model trained and evaluated with synthetic data, containing known antenna imperfections, would work properly like traditional techniques do. Next, we wanted to know if our model was deployable in the real world with real antennas. To do that, we have created a model that was trained with a large set of synthetic data and fine-tuned it with a small real-world dataset.

The architecture of our model, as seen in Fig. 1, is a mix of the architectures proposed in [4] and [5]. From work [5], we have utilized the concept of using the same neural network body to estimate the DOAs and the number of sources. From work [4], we have used the idea of a framework that fuses a 1D ResNet with MLE to better estimate DOAs with less computational power than required by a traditional MLE.

The ResNet, used in this work, as in our previous works, is a kind of neural network that utilizes skip layers connections to go deeper than regular convolutional neural networks [21].

Our model works as follows. Firstly, the received snapshot is pre-processed in a way that it can be fed to our neural network. Although recent works have shown that complex convolutional neural networks are promising [22], [23], our design relies on real values. The pre-processing step separates and concatenates both real and imaginary parts of the snapshot into a single array that can be processed by the 1D ResNet. In the sequence, the created array passes through the ResNet blocks, where the feature extraction happens. Each block contains batch normalization, ReLU, and three 1D convolutional layers, similar to [4]. At the end of the blocks, the neural network splits into two. Each fully connected segment contains two layers of 512 and 256 neurons, respectively [5]. The first output contains 181 angle bins (180 degrees of field of view and 1 degree of resolution), and gives the neural network's confidence that a target exists at a certain angle. The second output gives the neural network's confidence in the number of sources (varying from 0 to 4 sources). Lastly, by applying a simple threshold, the eight most probable angles are given to a traditional MLE algorithm, together with the most likely number of sources, to finish estimating the DOAs. An example of how the angles are filtered and passed to MLE can be seen in Fig. 2.

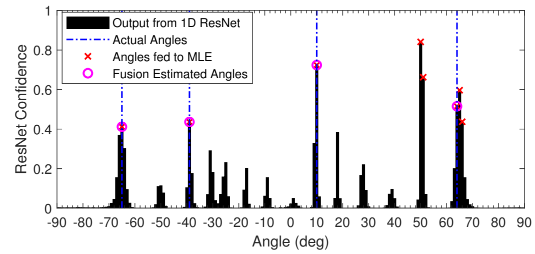


Fig. 2: The output from ResNet, the pre-selected angles, and the output from MLE. Here we can better understand how the fusion works.

According to our previous work [4], the fusion between our deep learning model with MLE seems to improve the accuracy while decreasing the operational complexity of a complete MLE search. We can see in Table IV-B the comparison between the operation complexity of MLE for various numbers of targets and our proposed fusion technique.

IV. Evaluation and Results

We can divide our experiments into two groups: evaluation with synthetic data and real-world data. Both consider imperfect antenna models. Moreover, we have utilized Root Mean Square Error (RMSE) as an evaluation metric as it is commonly used for DOA estimation [14]–[16].

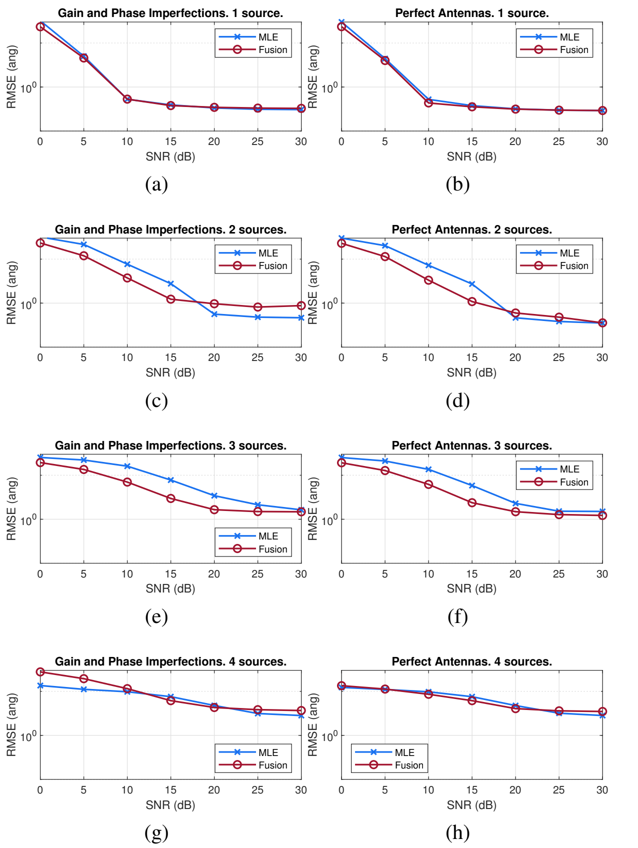
In this work, RMSE is given as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{K \times N} \sum_{k=1}^{K} \sum_{n=1}^{N} (\theta_{n,k} - \tilde{\theta}_{n,k})^2}\,, \tag{8}$$

in which K is the number of snapshots, N is the number of sources, $\tilde{\theta}_{n,k}$ is the estimated angle, and $\theta_{n,k}$ is the actual angle. When the number of sources is unknown, a mismatch between the number of $\tilde{\theta}_{n,k}$ and $\theta_{n,k}$ may happen during

TABLE I: Operational complexity comparison among different techniques.

Model Type	Number of Targets	Operational Complexity
MLE	1	9.2×10^4
MLE	2	8.3×10^6
MLE	3	4.9×10^8
MLE	4	2.2×10^{10}
1D ResNet	4	1.6×10^6
Fusion	4	1.7×10^6



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3: RMSE vs. SNR for MLE and our proposed model (1D ResNet + MLE) from 1 to 4 sources. Antenna imperfections were approximated. For this experiment, the number of sources was previously known. Evaluated with 10,000 synthetically generated snapshots per case. (a, b) 1 source, (c, d) 2 sources, (e, f) 3 sources, and (g, h) 4 sources. For (b, d, f, h) we used perfect antennas as a baseline, while for (a, c, e, g) we used antenna imperfections based on the curves of Fig. 4 .

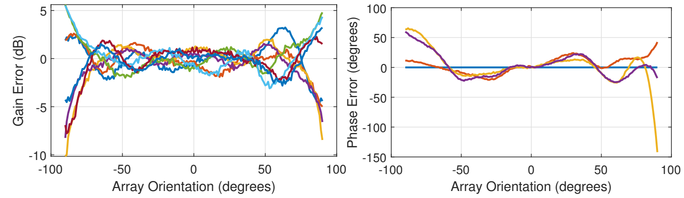
the calculation of RMSE. In this case, we use the smallest difference between the estimated and actual angles.

A. Synthetic Data

As a baseline, we repeated the experiments of our previous work 4, which does not contain any antenna imperfection.

This first experiment was set up to see if our model would perform properly with antenna imperfections, in a scenario using synthetic data with known antenna imperfections and a known number of sources. For this experiment, our deep learning model was trained with 40 million synthetically generated snapshots containing imperfect antennas, as described in eq. 3. Furthermore, we did not consider the estimation of the number of sources.

The curves showing the phase $\varphi(\theta)$ and gain $\alpha(\theta)$ errors are shown in Fig. 4. We have utilized eight gain/phase error curves, one per antenna, considering their array orientations (in degrees) [24]. Furthermore, we have introduced noise in the



(a) (b)

Fig. 4: The deviation errors, when compared to an ideal antenna, that were used for training the DL model with synthetic data. (a) Gain and (b) Phase errors.

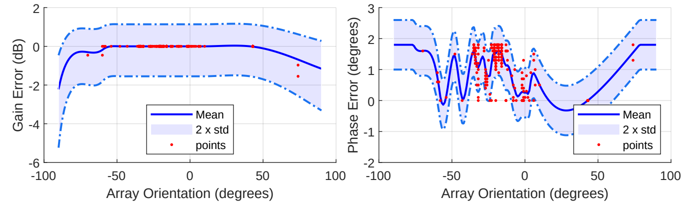
deviation curves used for training as, in most cases, a precise antenna model is often not available. It is important to note that these curves are often calculated considering one known target and, in reality, several targets may exist.

The results seen in Fig. 3 (b, d, f, h) show that our model could handle imperfections when using synthetic data, and it is in line with the results obtained by using MLE, similarly to our results using perfect antennas, as seen in Fig. 3 (a, c, e, g) [4], [5]. In other words, when using synthetic data, the neural network can learn well the gain and phase error curves that are angle dependent. Knowing that our fusion model holds imperfections with generated data, we wanted to find out if it could also handle real data.

B. Real Data

The second experiment was carried out utilizing an FMCW automotive radar with 8 MRA antennas. Two datasets were created with the collected data, one for fine-tuning and another for evaluation. For generating the ground truth, we have utilized the output results from several techniques used by the private sector. This means that our ground truth is not without errors as there are still no perfect techniques to always correctly estimate the angles and the number of targets.

In this experiment, the training was different than the previous one (done purely with synthetic data). First, we had to estimate the gain $\tilde{\alpha}(\theta)$ and phase $\tilde{\varphi}(\theta)$ errors for different array orientations. Since obtaining the gain/phase errors for every single antenna (eight in our case) increases the number of variables to deal with (sixteen variables), we have estimated



(a) (b)

Fig. 5: Estimated deviation errors used during the training of our DL model for real-world data. (a) Gain, and (b) Phase errors.

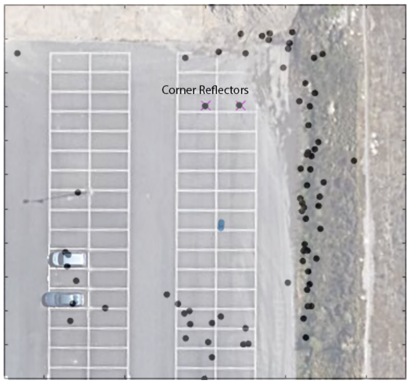


Fig. 6: An overlaying of the DOA and the number of targets estimations with a Google Earth image of the data gathering location using of our model. Unknown number of sources. Unknown DOAs. In this example, a moving bicycle passes through two stationary corner reflectors. The full animation can be found here[1]. For training our model, the estimated curves of Fig. 5 were used.

a single gain and phase error per known angle, following eq. 7.

In Fig. 5, we can see the estimated deviation error curves acquired after fitting the points obtained by solving eq. 7 for different known angles. However, the obtained curves have not been treated as a faithful representation of the radar's imperfections. Instead, for the training of our DL model, we have utilized the obtained curve as the mean of a normal distribution, with the standard deviation fixed at 0.6 dB for the gain and 0.4 degrees for the phase. These values were obtained empirically after several tests.

Moreover, obtaining enough representative labeled radar data to correctly train a DL model would can be quite laborious. For this, we have decided to train our model with synthetic data, utilizing the deviation errors from Fig. 5, and fine-tune our neural network with collected radar data. We have trained our model with 40 million synthetically generated snapshots containing a normal distribution of imperfections that depend on the array orientation. For fine-tuning, we have utilized 2759 real labeled snapshots, 770 augmented snapshots, and 22070 synthetically generated snapshots. In addition, we have used k-fold cross-validation [25] during the fine-tuning process.

The evaluation data was collected in a parking lot close to a hill, as we can see in Fig. 6. In our recording, a bicycle would pass through two corner reflectors positioned at the same distance from the radar but at different angles. In this dataset, we have collected 470 frames, with 29741 snapshots containing one target and 2978 snapshots with two targets.

For this experiment, besides RMSE, we have also used accuracy as a metric. We defined accuracy as the percentage of DOAs that were correctly estimated within the resolution.

In addition to be part of our fusion system, MLE was also used as a comparison method to estimate DOAs and the number of sources.

TABLE II: 1 target. Results using real data.

Type	RMSE	Accuracy within 1°	Accuracy within 5°	Percentage of Sources Estimated Correctly
MLE	14.9°	70.3 %	87.1 %	40.1 %
Our Model	15.2°	70.2 %	86.1 %	76.0 %

TABLE III: 2 targets. Results using real data.

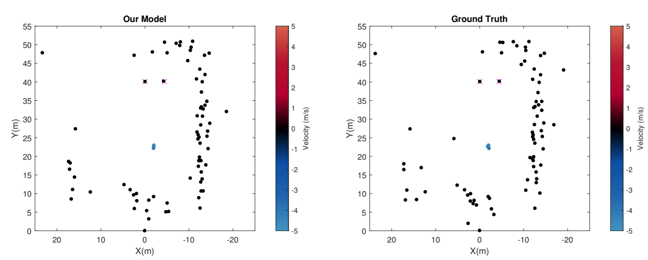
Type	RMSE	Accuracy within 1°	Accuracy within 5°	Percentage of Sources Estimated Correctly
MLE	1.6°	66.5 %	98.7 %	74.1 %
Our Model	2.0°	67.9 %	98.6 %	75.2 %

The results, which can be seen in Tables II and III, show that our technique performs similarly to pure MLE at estimating DOAs, being better at finding the number sources than the classical algorithm, and less computational intense, as seen previously in Table .

It is important to note that the high RMSE and the given percentage when estimating the number of targets are not due to poor performance but due to differences in the estimation of close ground reflections between our data-driven technique and the private company that generated our ground truth. A visual inspection of Fig. 7 or the animation[1] give us some insight that we can correctly estimate the angles for the reflected hill, corner reflectors, cars, and even a light pole. Additionally, when comparing a single frame of our technique with the ground truth, we can see how similar they are.

In other words, if we take the maximum likelihood estimation as the baseline, we can see that our model performs very closely to it, being better at estimating the number of sources correctly, but with reduced operational complexity.

In short, our model can handle both synthetic and real MRA single snapshots while significantly reducing MLE computational burden for estimating DOAs and the number of targets. Unlike many data-driven DOA estimation techniques that lack experimental validation, the fusion model performs closely to the maximum likelihood, and a visual inspection shows that all the important shapes are present in the ground truth and in our output.



(a) (b)

Fig. 7: A comparing example of our model with the used ground-truth. (a) A single frame with its respective angles, distance, and speed using our fusion method. (b) A single frame of our used ground truth.

[1] https://mllimadeoliveira.github.io/Thesis/RFusion.html

V. Conclusion

In this paper, we have proposed a fusion method, using Deep Learning (DL) and Maximum Likelihood Estimation (MLE) for estimating DOAs and the number of sources using a single snapshot, eight Minimum Redundancy Linear antenna Arrays (MRA), and estimating the gain and phase errors.

Besides proving that our model can work with real data instead of only generated synthetic data, our neural network model utilizes an approximated curve of the gain and phase errors within its model as we consider that these errors are angle-depended.

The results of our proposed model look promising as our model seems to perform closely to the classical MLE approach, using much less computational power. Our model also seems to be able to estimate the number of sources slightly better than MLE. With these results, we could conclude that our DL-based DOA estimation approach is an attractive option for practical radar systems.

For future work, we would like to increase our dataset, train with more targets and create our own ground-truth.

References

[1] W. Li, D. Zhang, Y. Li, Z. Wu, J. Chen, D. Zhang, Y. Hu, Q. Sun, and Y. Chen, "Real-Time Fall Detection Using Mmwave Radar," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2022-May, pp. 16–20, 2022.

[2] F. Tavella, A. Galata, and A. Cangelosi, "Phonology Recognition in American Sign Language," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2022-May, pp. 8452–8456, 2022.

[3] K. Aziz, E. De Greef, M. Rykunov, A. Bourdoux, and H. Sahli, "Radar-camera Fusion for Road Target Classification," *IEEE International Radar Conference (RADAR 2020)*, vol. 2020-Septe, 2020.

[4] M. L. Lima de Oliveira and M. J. Bekooij, "Deep-Mle: Fusion Between a Neural Network and Mle for a Single Snapshot Doa Estimation," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2022-May, pp. 3673–3677, 2022.

[5] M. L. Lima de Oliveira and M. J. G. Bekooij, "ResNet Applied for a Single-Snapshot DOA Estimation," *IEEE Radar Conference*, pp. 1–6, 2022.

[6] Z. Szalay and L. Nagy, "ULA and MRA beamforming using conventional and particle filter based algorithms," *International Conference on Applied Electromagnetics and Communications (ICECom)*, pp. 8–11, 2013.

[7] P. Häcker and B. Yang, "Single snapshot DOA estimation," *Advances in Radio Science*, vol. 8, pp. 251–256, 2010.

[8] F. Vincent, O. Besson, and E. Chaumette, "Approximate maximum likelihood estimation of two closely spaced sources," *Signal Processing*, vol. 97, pp. 83–90, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2013.10.017

[9] P. Stoica and K. C. Sharman, "Maximum Likelhood Methods for Direction-of- Arrival Estimation," *IEEE TRANSACTIONS ON ACOUSTICS. SPEECH, AND SIGNAL PROCESSING*, vol. 38, no. 7, pp. 1132–1143, 1990.

[10] H. L. Van Trees, *Optimum Array Processing*, 2002, vol. 4.

[11] B. M. Radich and K. M. Buckley, "Single-snapshot DOA estimation and source number detection," *IEEE Signal Processing Letters*, vol. 4, no. 4, pp. 109–111, 1997.

[12] J. Zeng, Z. He, and B. Liu, "Maximum likelihood DOA estimation using particle swarm optimization algorithm," *CIE International Conference of Radar Proceedings*, no. 5, pp. 0–3, 2007.

[13] H. Krim and M. Viberg, "Two Decades of Array Signal Processing Research," *IEEE Signal Processing Magazine*, pp. 67–94, 1996.

[14] A. M. Elbir, "DeepMUSIC: Multiple Signal Classification via Deep Learning," *IEEE Sensors Letters*, vol. 4, no. 4, 2020.

[15] O. Bialer, N. Garnett, and T. Tirer, "Performance Advantages of Deep Neural Networks for Angle of Arrival Estimation," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2019-May, 2019, pp. 3907–3911.

[16] Z. Yang, P. Chen, R. Geng, and Y. Jia, "DeepDOA: A Novel Deep Learning-Based Method for DOA Superresolution in a Single Snapshot," *2022 IEEE 5th International Conference on Electronics Technology, ICET 2022*, pp. 703–706, 2022.

[17] J. P. Merkofer, G. Revach, N. Shlezinger, and R. J. van Sloun, "Deep Augmented Music Algorithm for Data-Driven Doa Estimation," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2022-May, pp. 3598–3602, 2022.

[18] Y. Ma, Y. Zeng, and S. Sun, "A Deep Learning Based Super Resolution DoA Estimator With Single Snapshot MIMO Radar Data," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 4142–4155, 2022.

[19] B. Chong and C. M. S. See, "Sensor-Array Calibration Using a Maximum-Likelihood Approach," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 6, 1996.

[20] A. Liu and G. Liao, "An eigenvector based method for estimating DOA and sensor gain-phase errors," *Digital Signal Processing: A Review Journal*, vol. 79, no. 12, pp. 116–124, 2018.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[22] A. Fuchs, J. Rock, M. Toth, P. Meissner, and F. Pernkopf, "Complex-valued Convolutional Neural Networks for Enhanced Radar Signal Denoising and Interference Mitigation," in *IEEE Radar Conference*, no. 17774193. IEEE, 2021.

[23] Y. Cao, T. Lv, Z. Lin, P. Huang, and F. Lin, "Complex ResNet Aided DoA Estimation for Near-Field MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 139–11 151, 2020.

[24] G. Sommerkom, D. Hampicke, R. Klukas, A. Richterl, A. Schneider, and R. Thoma, "Reduction of DoA Estimation Errors Caused by Antenna Array Imperfections," *29th European Microwave Conference*, pp. 287–290, 1999.

[25] D. Berrar, "Cross-validation," *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, vol. 1-3, pp. 542–545, 2018.