DOA Estimation for 6G Communication Systems

Haya Al Kassir
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, halkassi@auth.gr

Ioannis T. Rekanos
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, rekanos@auth.gr

Pavlos I. Lazaridis
School of Computing and Engineering
University of Huddersfield
Huddersfield HD1 3DH, U.K.,
p.lazaridis@hud.ac.uk

Traianos V. Yioultsis
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, traianos@auth.gr

Nikolaos V. Kantartzis
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, kant@auth.gr

Christos S. Antonopoulos
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, chanto@auth.gr

George K. Karagiannidis
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, geokarag@auth.gr

Zaharias D. Zaharis
School of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki
54124, Greece, zaharis@auth.gr

2023 12th International Conference on Modern Circuits and Systems Technologies (MOCAST) | 979-8-3503-2107-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/MOCAST57943.2023.10176899

***Abstract*—The objective of this study is to analyze and compare different neural network (NN) architectures as multi-class classifiers to estimate the direction of arrival (DOA) using a uniform linear array (ULA). The study specifically investigates the prediction skills of three NNs: feed forward NN (FFNN), convolutional NN (CNN), and residual NN (ResNet) when estimating incoming signal DOAs in a realistic ULA of (M = 16) elements under noisy conditions. The NNs are trained on a correlation matrix generated by a ULA to estimate the DOAs. The results of the simulations indicate that ResNet performs better than FFNN and CNN in accurately estimating incoming signals.**

Keywords*— neural network (NN), direction of arrival (DOA), uniform linear array (ULA), feed forward NN (FFNN), convolutional NN (CNN), and residual NN (ResNet).

I. Introduction

The direction of arrival (DOA) or angle of arrival (AOA) estimation is essential in many fields, such as wireless communication and radar systems, to accurately identify the origin of a signal and allow source localization and beamforming. As a result, accurate DOA estimation is critical for these systems to operate properly, improve communication system coverage, and security. Several DOA estimation techniques have been developed, including Capon and Bartlett methods [1], as well as subspace techniques such as MUSIC and ESPRIT [2], [3]. These techniques, however, may not be appropriate for certain applications, such as those with high levels of interference and noise, resulting in long response times, significant computing complexity when receiving multiple signals, and mutual coupling between the elements.

Deep learning (DL) has acquired significant importance in DOA estimation in recent years due to the need for precise results in a wide range of applications, where DL algorithms have shown promise in addressing complicated issues in DOA estimation by leveraging the capacity of DL networks to understand complex correlations between incoming signals and DOAs. On the other hand, their ability to be trained end-to-end and handle non-linear data resulted in significant improvements in effectiveness and precision [4]. DL-based DOA has received considerable attention in recent years and has emerged as a popular research area. Numerous studies have concentrated on investigating various NN models to improve DOA applications such as [5], [6], and [7], where convolutional NN (CNN) provided high accuracy in DOA estimation in noisy environments. Furthermore, [8] researched a CNN-based DOA classification technique that reduced complexity while considering mutual coupling, while [9], CNN-based sparse array models achieved precise multi-source DOA modeling. CNN has been also utilized in [10] to estimate DOA based on data that has previously been processed using the MUSIC method. Combining feedforward NN (FFNN) with an autoencoder in [11], resulted in a solution for DOA classification to enhance generalization while considering array constraints. residual NN (ResNet) on the other hand, outperformed conventional approaches in [12] in predicting DOA from diverse directions, while in [13], it addressed mutual coupling by integrating with a generative adversarial network (GAN).

In this paper, we address the topic of estimating DOAs using three DL models, respectively FFNN, CNN, and ResNet, all of which were developed with PyTorch to estimate the DOA of a received signal as a multi-class classification task, with the number of classes determined based on the selected resolution, which will be explained later in this paper. The model is trained using a dataset that contains preprocessed AOA values and their corresponding received signals defined by a correlation matrix.

The content of this paper is organized as follows: Section II introduces the problem formulation, while Section III describes the NNs structures. Section IV contains the simulation results, and finally, the conclusion is presented in section V.

II. Problem Formulation

A Uniform Linear Array (ULA) comprises M elements uniformly spaced apart by a distance of (d = λ/2), where λ represents the wavelength of the signals. In Fig 1, N monochromatic signals ($s_n(k), n = 1, \dots, N$) are transmitted by sources located at angles ($\theta_n(k), n = 1, \dots, N$). The M x M correlation matrix $\mathbf{R}_{xx}$, which represents the input of the NN, can be formulated using the following equation:

979-8-3503-2107-4/23/$31.00 ©2023 IEEE

$$\mathbf{R}_{xx} = \mathbf{A}\,\mathbf{R}_{ss}\,\mathbf{A}^{H} + \mathbf{R}_{nn} \quad (1)$$

where

$$\mathbf{A} = [\boldsymbol{a}(\theta_1)\ \boldsymbol{a}(\theta_2) \dots \boldsymbol{a}(\theta_N)] \quad (2)$$

and

$$\boldsymbol{a}(\theta) = \left[exp\big(j\beta z_1 Cos\theta\big) \dots exp\big(j\beta z_M Cos\theta\big)\right]^T \quad (3)$$

are respectively, the total steering matrix M * N generated by the steering vectors relating to the AoAs of all the incoming signals, and the array steering vector corresponding to the angle θ. $\beta$is the wavenumber in free space ($\beta = 2\pi/\lambda$), superscript T denotes the transpose operation, and H indicates the Hermitian transpose operation.

Assume there is no correlation between any of the incoming signals and any of the noise signals. The noise and signal correlation matrices $\boldsymbol{R}_{nn}$ and $\boldsymbol{R}_{ss}$ respectively can be defined as follows:

$$\boldsymbol{R}_{nn} = P_n\, I_{(M*M)} \quad (4)$$

$$\boldsymbol{R}_{ss} = P_s\, I_{(N*N)} \quad (5)$$

where P_n and P_s are the noise and signal mean power respectively, while $I_{(M*M)}$ and $I_{(N*N)}$ are the M * M and N * N identity matrices respectively.

As a result, the M * M correlation matrix $\mathbf{R}_{xx}$, which represents the input of the NN, can be formulated using the following equation:

$$\mathbf{R}_{xx} = \mathbf{A}\, P_s\, I_{(N*N)}\, \mathbf{A}^{H} + P_n\, I_{(M*M)} \quad (6)$$

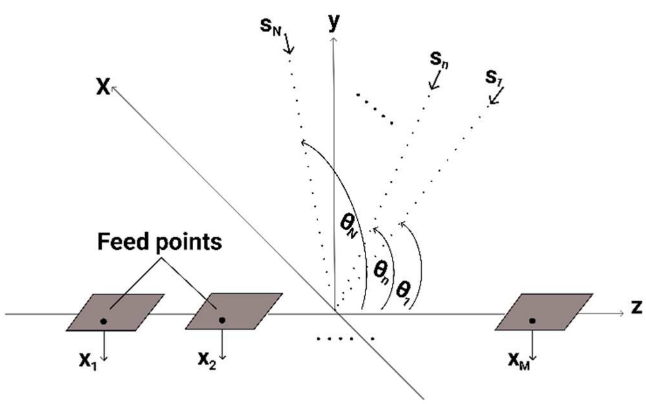

Fig. 1. The structure of the uniform linear array.

To model the correlation matrix, we used data from N incoming signals, each with a distinct AOA and a minimum difference of ($\Delta\ \theta > 6$). We separated the correlation matrix into real and imaginary parts before feeding it into the NN given that it's a complex combination, and NNs are supposed to operate with real-valued inputs. Therefore, the resulting input size is 2*M*M.

III. NEURAL NETWORKS OVERVIEW

The effectiveness of three NN architectures of four layers used as multi-class classifiers for predicting the AOAs of incoming signals is investigated in this study.

a) Feed-forward NN (FFNN)

A traditional NN, also known as a fully connected NN in which information flows only in one direction, from input to output. Although FFNNs are widely used in many applications, they are potentially vulnerable to overfitting and poor performance. Each layer in our FFNN is followed by a dropout layer with a rate of 0.5 and rectified linear unit activation function (ReLU). Given the dimension of the input data is 2*M*M, where M = 16, the input layer utilizes 512 neurons. The sizes of the hidden layers are 1024 and 2048, respectively. Finally, the number of neurons in the output layer corresponds to the number of classes. We set the batch size and learning rate to 100 and 0.0001, respectively, to improve performance and handle the small number of parameters.

b) Convolutional l NN (CNN)

A type of NN that utilizes convolutional layers to classify complex patterns in images, making it particularly effective for classification tasks and object detection. Our CNN architecture consisted of three 2D convolutional layers and one fully connected layer at the output. The ReLU activation function was used in the hidden layers, while max pooling layers were used to reduce the spatial size of the feature maps. The output layer used a SoftMax activation function to classify the input images. The input layer of the CNN comprises 512 neurons, while the hidden layers contain 256 and 128 neurons, respectively. The number of neurons in the output layer is equal to the number of classes. During the training of the CNN, we used a batch size of 1024 and a learning rate of 0.001.

c) Residual NN (ResNet)

One of the most efficient NN in this field since it learns the mapping through residual connections rather than learning it directly from input to output. Furthermore, by utilizing the skip connection, it can avoid overfitting. Our ResNet is comprised of an input convolutional layer, batch normalization, ReLU activation, max pooling, two residual layers, and a fully connected layer that generates the final output. We adopt a similar architecture to the CNN, where we start with 512 neurons at the input, then 256 and 128 neurons utilized by the hidden layers, and finally the number of neurons in the output layer is equal to the number of classes. During ResNet training, we followed the same batch size and learning rate of CNN.

IV. SIMULATION RESULTS

Our research concerns the use of a ULA with M = 16 elements to receive incoming signals from a number of sources (N = 3) in a noisy environment with SNR = 0 dB, implying that the power of the noise is equal to the power of the signal. The AOAs of the signals are recorded with a minimum divergence of ($\Delta\ \theta > 6°$). We use MATLAB to generate 1.1 million records, which we then use to generate a correlation matrix, which represents the input of the NN according to equation (6). 80% of the records are used to train the NN, with the remaining 20%

used for validation. We evaluate the performance of NNs in this work at a resolution of 0.25, which means that the data being evaluated is partitioned into 0.25-degree, with a total of 481 classes defined within a range of arrival angles of [30°-150°]. Therefore, the predicted angles are in the form of [30°, 30.25°, 30.50°,...., 149.75°, 150°]. The simulation results are obtained using two performance metrics: the mean absolute error (MAE) metric, which calculates the absolute differences between predicted and actual values using the following formula:

$$\text{MAE} = \frac{1}{n}\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m}\left|\hat{y}_{ij} - y_{ij}\right| \tag{7}$$

where n is the number of training and validation samples, and m is the number of signals. The predicted angles are represented by $\hat{y}_{ij}$, while the actual angles are represented by y_{ij}.

On the other hand, the second metric is the F1 score, which is a common measure to assess classification model performance. It is defined as the harmonic mean of precision and recall, with values ranging from 0 to 1 and the highest score indicating the best performance. In multi-class classification, the F1 score is calculated as the weighted average of the F1 scores for each class.

As mentioned earlier, our study aims to compare the accuracy of three classifiers, FFNN, CNN, and ResNet by testing their performance at a resolution of 0.25. To evaluate the performance, we conduct training and validation for each of the three NNs classifiers, and the results are based on the two mentioned metrics after 100 epochs. The outcomes of our investigation are presented in Table I, along with Figs 2, 3, and 4.

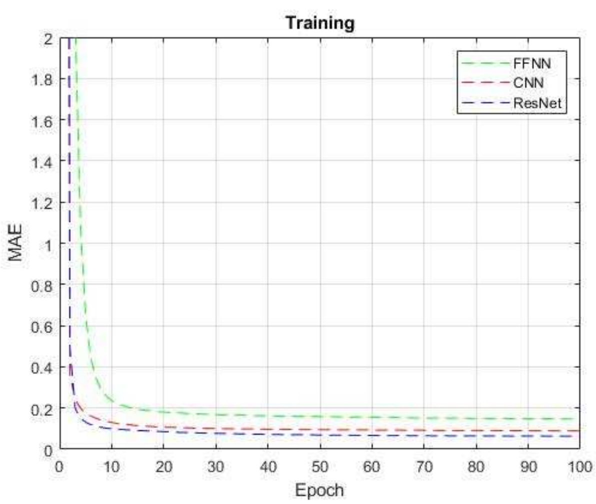

Fig. 2. Training MAE comparison of the three investigated architectures.

Figs 2, and 3 provide a visual representation of how the classifiers performed concerning the MAE metric during the training and validation. ResNet outperformed the other two classifiers based on these representations and what is shown in Table I. The MAE for ResNet was the lowest with 0.063 during training and 0.069 during validation. The CNN classifier performed second best with a training MAE of 0.089 and a validation MAE of 0.094. FFNN exhibited the highest MAE values, with 0.14 during training and 0.15 during validation. The MAE results indicate that ResNet and CNN were better able to minimize the difference between the predicted and actual values, indicating higher accuracy, while FFNN performed the worst. These findings suggest that ResNet and CNN may be more suitable for applications where high accuracy is critical, while FFNN may be used in applications where lower accuracy is acceptable.

The F1 score results in Fig 4, and Table I, also support the finding that ResNet outperformed the other two classifiers in terms of accuracy. The F1 score for ResNet was the highest with 0.989, followed by CNN with 0.985, and FFNN with 0.976. The higher F1 score for ResNet and CNN indicates that they achieved better accuracy in terms of precision and recall, which is desirable in most applications.

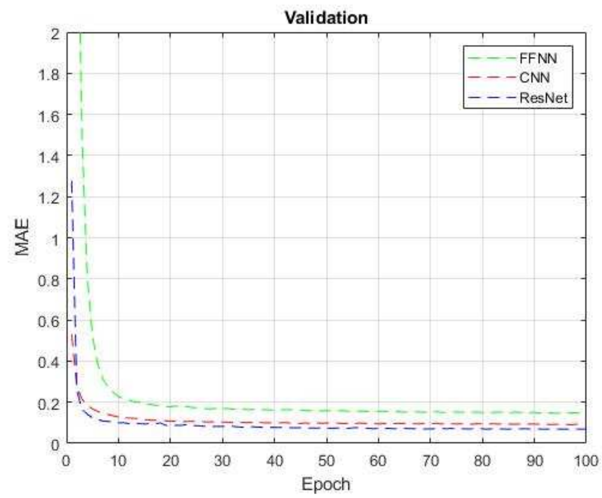

Fig. 3. Validation MAE comparison of the three investigated architectures.

TABLE I.
Performance comparison of the three NN architectures after 100 epochs with 0.25 resolution.

NN	No of neurons per layer	MAE Training/validation	F1
FFNN	512/21024/2048/481	0.14/0.15	0.976
CNN	512/256/128/481	0.089/0.094	0.985
ResNet	512/256/128/481	0.063/0.069	0.989

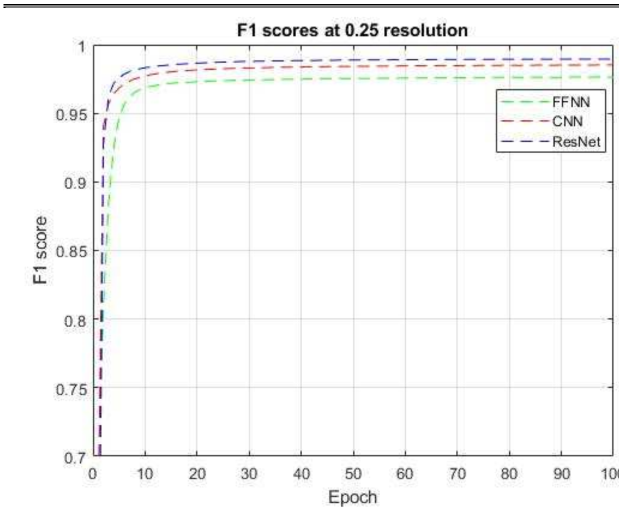

Fig. 4 .F1 score of three investigated architectures at 0.25 resolution.

V. Conclusion

In terms of accuracy, this study compared the performance of various NN architectures for estimating the DOA of incoming signals in a realistic ULA in noisy environment. The investigation compared the performance of FFNN, CNN, and ResNet, as multi-class classifiers based on their prediction skills. The study found that ResNet outperformed the other two classifiers in terms of predictive performance, as demonstrated by lower MAE values and a higher F1 score. Therefore, when taking into account prediction accuracy, ResNet emerges as the most efficient multi-class classifier for our specific task.

Acknowledgment

This research was supported by the European Union, through the Horizon 2020 Marie Skłodowska-Curie Innovative Training Networks Programme “Mobility and Training for beyond 5G Ecosystems (MOTOR5G)” under grant agreement no. 861219.

References

[1] A. Vesa and A. Iozsa, "Direction - of - Arrival estimation for uniform sensor arrays," *2010 9th International Symposium on Electronics and Telecommunications*, 2010, pp. 249-252.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, March 1986.

[3] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984-995, July 1989.

[4] Z. Wang, C. Tang, X. Sima and L. Zhang, "Research on Application of Deep Learning Algorithm in Image Classification," *2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC),* 2021, pp. 1122-1125.

[5] R. Akter, V. -S. Doan, T. Huynh-The and D. -S. Kim, "RFDOA-Net: An Efficient ConvNet for RF-Based DOA Estimation in UAV Surveillance Systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 12209-12214, Nov. 2021.

[6] S. Chakrabarty and E. A. P. Habets, "Broadband doa estimation using convolutional neural networks trained with noise signals," *2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2017, pp. 136-140.

[7] G. K. Papageorgiou, M. Sellathurai and Y. C. Eldar, "Deep Networks for Direction-of-Arrival Estimation in Low SNR," *IEEE Transactions on Signal Processing*, vol. 69, pp. 3714-3729, 2021.

[8] O. J. Famoriji, O. Y. Ogundepo and X. Qi, "An Intelligent Deep Learning-Based Direction-of-Arrival Estimation Scheme Using Spherical Antenna Array With Unknown Mutual Coupling," *IEEE Access,* vol. 8, pp. 179259-179271, 2020.

[9] S. WANDALE and K. ICHIGE, "On the DOA Estimation Performance of Optimum Arrays Based on Deep Learning," *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM*), 2020, pp. 1-5.

[10] C. Liu, W. Feng, H. Li and H. Zhu, "Single Snapshot DOA Estimation Based on Spatial Smoothing MUSIC and CNN," *2021 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Xi'an, China, 2021, pp. 1-5.

[11] Z. -M. Liu, C. Zhang and P. S. Yu, "Direction-of-Arrival Estimation Based on Deep Neural Networks With Robustness to Array Imperfections," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7315-7327, Dec. 2018.

[12] P. Li and Y. Tian, "DOA Estimation of Underwater Acoustic Signals Based on Deep Learning," *2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT),* Shanghai, China, 2021, pp. 221-225.

[13] W. Fang et al., "A Deep Learning Based Mutual Coupling Correction and DOA Estimation Algorithm," *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2021, pp. 1-5