ALKU Journal of Science 2024, Sayı 6(2): 165-173
e-ISSN: 2667-7814

Original Article

Adapting Object Detection Models for Multi-Target Detection Utilizing Radars

İbrahim Rıza Hallaç[1*], Deniz Akbaba[1], Gökhan Gökçe[1], S. Gokhun Tanyer[2,3], Peter F. Driessen[3]

[1]*Department of Computer Engineering, Alanya Alaaddin Keykubat University, Antalya, Türkiye.*
[2]*MulticoreWare Inc. San Jose, CA, USA, INNOVTECH, Victoria, Canada*
[3]*Department of Electrical and Computer Engineering, University of Victoria, Victoria BC, Canada.*
**ibrahim.hallac@alanya.edu.tr*

Abstract

This paper investigates the use of Deep Learning (DL) in multiple input multiple output (MIMO) radar target detection, focusing on azimuth and elevation estimation. Traditional methods face challenges like interference and reflections, especially in multi-target scenarios. Feature extraction conventionally relies on range correlation, Doppler filtering, and angle beamforming, followed by detection after constant false alarm rate (CFAR) processing. However, early data sparsification by bin selection often leads to information loss, particularly with large data cubes required for practical implementation. DL techniques offer an alternative, specifically in azimuth and elevation detection at earlier stages of radar data processing. We developed a convolutional neural network (CNN) model that achieved Mean Square Errors (MSE) of 0.149 for azimuth and 0.168 for elevation on single-target data from 5,000 samples. The model's performance in dual-target scenarios showed MSEs ranging from 0.838 to 1.845, tested on 8,000 samples from a dataset of 72,000. This paper details the model development process, its impact on radar target detection, and potential future research directions involving the substitution of multi-bin DL blocks with traditional methods.

Keywords: Radar processing pipeline, MIMO radar, Multi-target detection, Machine learning, Convolutional neural network.

Radarla Çoklu Hedef Tespiti için Nesne Tanıma Modellerinin Uyarlanması

Özet

Bu makale, çoklu giriş çoklu çıkış (MIMO) radar hedef tespitinde için Derin Öğrenme tekniğinin uygulamasını, özellikle azimut ve yükseklik tahminine odaklanarak ele almaktadır. Geleneksel yöntemler, özellikle çoklu hedef senaryolarında parazit ve yansıma gibi zorluklarla karşı karşıya kalmaktadır. Özellik çıkarımı, genellikle menzil korelasyonu, Doppler filtreleme, açı demetleme ve sabit yanlış alarm oranı (CFAR) işleminden sonra tespit adımlarını içeren klasik radar sinyal işleme hattına dayanmaktadır. Ancak, erken aşamada veri seyreltilmesi, pratik uygulamalar için gereken büyük veri küplerinde bilgi kaybına yol açabilmektedir. Derin Öğrenme teknikleri, azimut ve yükseklik tespiti için alternatif bir yaklaşım sunmaktadır. Geliştirdiğimiz evrişimsel sinir ağı (CNN) modeli, 5000 örnekten oluşan tek hedefli veri üzerinde azimut için 0.149 ve yükseklik için 0.168 Ortalama Kare Hata (MSE) değerleri ile yüksek performans göstermiştir. İki hedefli senaryolarda ise model, 72.000 örneklik veri setinden 8000 test örneği üzerinde 0.838 ile 1.845 arasında MSE değerleri elde etmiştir. Bu makale, model geliştirme sürecini, radar hedef tespitindeki etkisini ve Derin Öğrenme ile geleneksel yöntemlerin entegrasyonuna yönelik potansiyel gelecek araştırma yönlerini detaylı bir şekilde ele almaktadır.

Anahtar Kelimeler: Radar işleme hattı, MIMO radar, Çoklu hedef tespiti, Makine öğrenimi, Evrişimsel sinir ağı.

Geliş/Received: 23 Mayıs/May 2024
Kabul Ediliş/Accepted: 5 Haziran/June 2024

1. INTRODUCTION

Radar stands for Radio Detection and Ranging. The history of radar technology is lengthy and fascinating. The concept of using radio waves to detect objects was first proposed during the late 1880s by German physicist Heinrich Hertz, but it wasn't until the 1930s that serious developmental work on radar began. Radar technology was essential to military operations during World War II because it allowed for the identification and tracking of adversarial ships and aircraft. The advancement of radar technology after the war led to its widespread usage in civil applications including weather forecasting and air traffic control. The radar data was processed using traditional signal processing techniques to extract environmental data. However, the limits of traditional signal processing techniques utilized in radar systems have come to limit more recently. These constraints make it difficult to detect things in congested situations and make it difficult to discern between weak and strong signals. Additionally, traditional methods require significant expertise and manual tuning of parameters, which can be time-consuming and error-prone.

In object detection, deep learning algorithms are typically divided into two main categories: two-stage models that provide higher accuracy, and one-stage models that offer faster processing [1]. However, the application of these advanced algorithms, such as Mask R-CNN [2] and YOLO [3], to radar target detection is challenging due to the fundamental differences in the type of data collected. Radar sensors gather complex numbers that lack spatial locality, which is different from the pixel-based information these algorithms are built to process. In this study, we focus on using CNNs and machine learning to identify radar targets, specifically aiming to predict their azimuth and elevation. Our objective is to develop a target detection system using CNNs that overcomes the limitations of traditional signal processing techniques and improves radar performance.

Radar transmits radio waves and detects their reflection on objects to determine their properties. MIMO (Multiple Input Multiple Output) radar uses multiple transmit and receive antennas to improve radar performance. Normally, complex formulas are used to extract features from radar data which is called traditional signal processing. The captured radar data is often called the radar-data-cube since it is formed in a 3D matrix of (time-signal samples, pulses, and virtual receivers). Here we are interested in the third step which is finding a target's azimuth, elevation, magnitude, and phase with Deep Learning Techniques as a "proof of concept".

Traditional signal processing methods are outdated and have problems that formulas cannot overcome. These are interference which is detecting a weak target next to a strong target and reflection which is when signals scatter and create blind spots in detection. While reviewing the literature we noticed that this is typically done in 3 blocks using pre-compiled models to detect targets from the radar data. However, they are just detecting the targets and classifying them. We have elected to focus on predicting the azimuth and elevation of a target to integrate with traditional methods.

We started to build a model for one-target samples. After deciding on the structure of the model, which is a convolutional neural network (CNN), and fine-tuning hyperparameters, we achieved Mean Square Errors (MSE) of 0.149 for azimuth and 0.168 for elevation on 5000 test samples. We then built and trained a CNN using a dataset of 72,000 samples, each with two targets, achieving MSEs ranging from 0.838 to 1.845 on 8000 test samples. Our study explains why target detection cannot be performed using only the last phase of MIMO radar data and presents models that can be used instead, along with their hyperparameters. This will also shed light on future studies. Finally, cross-checking the calculations made using the data obtained here with traditional methods will open the way for studies that will obtain more accurate results.

Artificial intelligence and machine learning have been effective techniques in recent years for enhancing radar performance. In particular, Convolutional Neural Networks (CNNs) have demonstrated promise in the detection and classification of objects in radar data.

The process of radar target detection can be broken down into three fundamental stages: detection, classification, and tracking. This method utilizes a radar data cube dataset to effectively train a machine learning (ML) model. Each radar data cube contains a 3D matrix of (time-signal samples, pulses, and virtual receivers) for extracting the target features such as size, velocity, range, elevation, and azimuth. Target detection entails finding prospective targets in the radar data, target classification entails categorizing each target (e.g., an airplane, a bird, etc.), and target tracking entails monitoring each target's movement over time. The common approach when using AI make use of all this information and forgo traditional signal processing methods to learn entirely from scratch. Instead, we have opted to combine AI with traditional approaches with the context-free part of the whole data cube, even though there has been a lot of study on utilizing AI for target detection and classification.

In our method, we use the part of the radar data cube that consists of virtual receivers to train a CNN model with deep learning. The reason for this is that conventional methods are more efficient at other steps. Azimuth and elevation of the target are output by the model. To achieve high accuracy in predicting the azimuth and elevation of the target, we fine-tuned the hyperparameters in our model. We tested the performance of our CNN model on data sets of different sizes and complexity.

The incorporation of machine learning and artificial intelligence is the latest chapter in the long history of continuous innovation in radar technology. Our goal is to improve the accuracy and robustness of radar systems by focusing on the prediction of azimuth and elevation of the target, paving the way for further research in this field. We believe that our approach could be useful for various applications such as autonomous vehicles, air traffic management, and military surveillance.

To summarize, this work presents a novel method for radar target detection using CNNs, especially for determining the azimuth and elevation of a target as an example. Our work demonstrates the potential of CNN models to improve radar performance and helps in the development of more reliable and accurate radar systems, especially for cases where target power is distributed over multiple bins.

2. LITERATURE REVIEW

Radar target detection is an important task in a variety of applications, including FMCW radar systems for motor vehicles [4], maritime traffic surveillance, and air traffic control. It is a system that uses radio waves to detect objects, their distance, direction, and speed. In recent years, machine learning techniques such as artificial neural networks and deep learning have been applied to improve the accuracy and efficiency of radar target detection.

A comprehensive review of machine learning techniques used in radar signal processing was provided by Lang et al. (2020), covering a wide range of topics such as target detection, classification, and tracking [5]. Similarly, Jiang et al. (2022) provided a review of the various artificial neural networks and deep learning techniques used in radar target detection [6]. These studies highlight the potential of machine learning techniques to improve the accuracy and efficiency of radar target detection.

Several studies have proposed adapting the YOLO (You Only Look Once) model for radar target detection. Kim et al. (2020) proposed a YOLO-based approach for simultaneous target detection and classification in FMCW radar systems for vehicles [4] and achieved high accuracy and real-time performance. Zhou et al (2019) proposed YOLO-RD, a lightweight object detection network for range-Doppler radar imaging [7]. Lira-YOLO, a lightweight model for ship detection in radar images [8], was proposed by Long et al. (2020), which achieves high accuracy while maintaining real-time performance.

Pica et al. (2021) proposed a new SAR target detection approach based on YOLO and very deep multicanonical correlation analysis [9], which achieves high accuracy and robustness to variations in target orientation and aspect angle.

In addition, Baird et al. (2020) proposed a CNN-LSTM network for improving target detection in real maritime long-range surveillance radar data [10], which achieves high accuracy and robustness to variations in target orientation and aspect angle. Kumar and Kumar (2021) proposed a deep convolutional neural network-driven neuro-fuzzy system for moving target detection from radar signals [11], which achieves high accuracy and robustness to variations in target motion.

Zhang (2022) provided a survey of the various convolutional neural network-based approaches that have been applied to target detection in SAR images [12]. The study covers a range of topics, including target detection, classification, and tracking, and provides an overview of the state of the art in this field. These studies demonstrate the potential of machine learning techniques, especially YOLO-based approaches, to improve the accuracy and efficiency of radar target detection.

In summary, machine learning techniques, especially those based on YOLO, utilize the entirety of the data to generate localized outcomes, showing considerable promise in enhancing the accuracy and efficiency of radar target detection across various applications. However, it is important to note that these methods are not directly comparable to our approach, which integrates machine learning CNN models with traditional techniques to accurately pinpoint the target's position. Additionally, the specific data used in our project does not support target localization in the manner typical of such images.

This study includes many carefully planned and conducted empirical experiments. Rather than detailing the findings of each experiment, summaries of the results are provided in their respective sections. This allows readers to easily follow and understand the specific details of each experiment. In the conclusion, the most important findings are discussed in depth, according to the main goals of the study. This approach clearly presents the overall framework and key findings of the study, while also giving readers access to detailed analyses of the experiments.

3. DATA SYNTHESIS

Training and test data are synthesized using a monostatic MIMO radar with 16 receivers and 16 transmitter elements without loss of generality. The received signal pattern for the antenna array at the direction (u, v) can be calculated by

$$g(u, v) = \boldsymbol{br} \tag{1}$$

where the received signal at the virtual receivers (VRXs) form the received signal (column) vector $\boldsymbol{r}$ of length P, and $\boldsymbol{b}$ is the steering (column) vector evaluated at(u, v). One can oversample (u, v) uniformly for the received signal pattern

$$u_m = \frac{2m'}{Mq_\phi} - 1, \qquad \text{for} \quad 0 \le m' < Mq_\phi, \tag{2}$$

$$v_n = \frac{2n'}{Nq_\theta} - 1, \qquad \text{for} \quad 0 \le n' < Nq_\theta. \tag{3}$$

and where the received signal is the superposition due to each target

$$r_{rx,s}(p) = \sum_{t=1}^{T} \sigma_{c,t}\, e^{j\pi(mu_t + nv_t)} \tag{4}$$

where T is the total number of targets in the far-field, and $p = 1, 2, \ldots P_s < MN$ are the VRXs created by available M TX and N RXs located on the grid points (m, n). Equation (4) provides the received signal at the VRX elements and (1) provides the synthesized received pattern to measure the radar performance metrics, namely, the peak-to-side-lobe ratio (PSLR) and the beam width (BW) calculated in the usable field of view (uFOV) of the antenna. The SNR of signals is calculated using normally distributed noise at the antenna elements.

4. DATASET

We have utilized two distinct types of datasets, each comprising sensor data, target azimuth, elevation, absolute amplitude, and SNR. Sensor data consists of 192 complex numbers representing raw radar data values. Azimuth and elevation denote the polar coordinates of the target. Absolute amplitude measures the target's reflectivity, and SNR indicates the noise level in the data.

One approach to handling radar signal data in deep learning involves using neural networks specifically designed to process complex-valued inputs, weights, and activations. This method does not require additional preprocessing and is straightforward but remains underexplored. In our exploration of using deep learning techniques for traditional radar signal processing in object detection, our primary goal is to leverage state-of-the-art object detection models, such as YOLO, which are typically applied to 2D or 3D image inputs. Consequently, we have chosen to propose baseline CNN models for this purpose. Accordingly, we have implemented basic preprocessing of the complex numbers. This preprocessing involves treating the real and imaginary parts of each number as separate features and exploring additional properties that can be derived from complex numbers, such as their angles.

The general structure of the datasets remains consistent, differing primarily in the number of targets and noise levels. Specifically, the single-target dataset contains 180,000 samples with mixed noise values, while the two-target dataset includes 90,000 samples, each with a 20dB noise level. We have partitioned each dataset into training, validation, and test sets, with ratios of 64%, 16%, and 20%, respectively.

5. EXPERIMENTS AND RESULTS

This section describes the experiments performed with two different data sets and the results obtained. First, we focused on training CNN models and optimizing hyperparameters to improve their performance. Building on the findings from these initial experiments, we further developed our CNN models to improve their effectiveness in dual-target scenarios. The following subsections describe the methodology applied, the experimental setups, and the results of these efforts, highlighting how the iterative refinements have incrementally improved the accuracy and effectiveness of our target detection models.

5.1 Experiments on Single-Target

We began our experiments using a simple CNN architecture to determine if deep learning could effectively learn azimuth and elevation values, compared to traditional signal processing methods. Our initial model, shown in Figure 1, used one-dimensional convolutional layers and dense layers. We experimented with the model by adding more convolutional and dense layers and placing max pooling layers between them for enhanced performance. We also included dropout layers to help prevent overfitting.

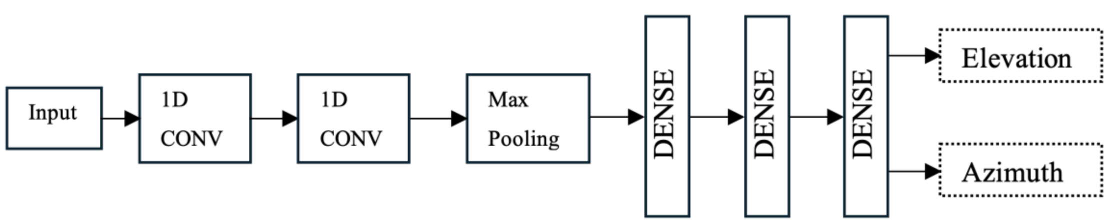

Figure 1. Initial CNN model with 1D convolutions

As we refined our model, several critical questions emerged regarding the optimal setup for our specific problem. Key considerations included:

1. Choosing additional input features beyond the real and imaginary parts, with options ranging from 192x2 to 192xN.
2. Deciding on the input shape, with possibilities such as 192x3, 3x192, or 571x1 to manage three features from 192 sensors.

3. Selecting between one-dimensional and two-dimensional convolution types.
4. Investigating the impact of various layer types such as pooling and dropout layers on model performance.

Instead of conducting extensive experiments to finalize these parameters, we used an intuitive approach to determine the most effective settings for our dataset of single-target samples. In our investigations into the optimal configuration for our CNN model tailored for a single-target dataset, we arrived at several key conclusions from the considerations outlined above. The decisions that yielded the best results are summarized as follows:

- **Input Shape:** The input shape of 192 x 3 (real part, imaginary part, and angle) proved to be the most effective.
- **Convolution Layers:** Using two convolutional layers provided the best performance. Two-dimensional (2D) convolution was found to be slightly more effective than one-dimensional (1D) convolution.
- **Pooling Layers:** Contrary to expectations, max-pooling layers negatively impacted our model's performance. While these layers generally reduce overfitting and computational load by compressing feature maps, they also lead to the loss of critical information, which is especially detrimental in precise tasks such as single-target detection.
- **Dropout Layers:** The inclusion of dropout layers was beneficial. These layers helped prevent overfitting by randomly dropping units during the training process.

According to these settings, the results of the best-performing model with 2D convolution layers are shown in Table 1. In addition to Mean Square Error (MSE), we used an additional metric called 'Angle Error' for our evaluations. This metric is calculated by taking the square root of the sum of the squares of the azimuth and elevation errors for each prediction.

Table 1. Results for single-target dataset

	Azimuth	Elevation	Angle
MSE	0.149	0.168	0.253
Std	0.132	0.159	0.171
50%	0.112	0.128	0.219
Max Error	1.082	2.457	2.458

5.2 Experiments on Two Targets with Equal Rcss

Moving to a two-target dataset, we adapted our best model from the single-target experiments to handle multiple targets without changing the hyperparameters. This required reworking the output layer to perform regression on multiple azimuth and elevation values simultaneously. To better understand the input structure, Figure 2 provides an overview of the sensor and target data and illustrates a single example from the dataset.

Table 2. Sample values from the two-target dataset

		Sensor Data	Num. of Targets	Target-1 Azimuth	Target-1 Elevation	Target-2 Azimuth	Target-2 Elevation
Data-1	1	-0.393112, 0.254718	2	-10.424340	-4.574590	79.002089	-8.630493
	2	0.656807, 0.418274	…	…	…	…	…
	…	…	…	…	…	…	…
	191	-1.289922, -0.102665	…	…	…	…	…
	192	-0.504360, 0.784353	…	…	…	…	…

In our investigation of the two-target dataset, we evaluated various modifications to our CNN model. In this section, we summarize the results, focusing on which modifications improved accuracy and which did not. We explain the successful techniques that improved the performance of our model, as well as the modifications that fell short of expectations. These findings are important for refining our approach to CNN architectures for complex radar signal processing tasks with multiple targets. In the following, we divide the results into two categories to provide a clear analysis of our experimental results.

Improvements in Accuracy:

- **Changing Input Shape and Parameters**: Adding the magnitude of real and imaginary values and Cartesian x and y values to the input parameters, along with changing the filter size and adding an extra dense layer, improved the model's accuracy.
- **Increased Output in Dense Layer**: Modifying the output layer to include Angle1 and Angle2 improved accuracy, though results were still not optimal.
- **Optimization Adjustments**: Switching to the relu activation method, adam optimizer, and mae for the loss function were determined as the best parameters.
- **Training Duration**: Extending training to 200 and 300 epochs showed slight improvements.

That Did Not Help:

- **Max Pooling Layer**: Adding max pooling layers negatively affected the model.
- **Additional Convolutional and Dense Layers:** Initially adding more layers increased the error rate without reaching desired levels.
- **Change in Activation Methods:** Switching to elu and selu activation methods, and later mixed methods (tanh and sigmoid), did not yield positive effects.
- **L1 and L2 Kernel Regularizers:** Both L1 and L2 regularizers did not improve the model significantly.
- **Extended Training Beyond 300 Epochs:** Further extending training led to overfitting, without significantly lowering the maximum error rate.
- **Swapping Target Data in Problematic Subsets:** Attempting to address issues in a subset of the data by swapping azimuth and elevation of targets did not resolve the high error rates.

Table 3. Results for two-target dataset

	Azimuth1	Elevation1	Angle1	Azimuth2	Elevation2	Angle2
MSE	0.838	1.845	2.211	0.934	0.968	1.875
Std	1.324	1.619	1.895	1.229	1.463	1.637
50%	0.504	1.412	1.682	0.611	0.560	1.447
Max Error	14.730	14.372	14.965	14.954	14.573	13.482

Transitioning to a two-target dataset, we adapted our best model from the single-target experiments to handle multiple targets.

6. CONCLUSION AND FUTURE WORK

In this study, we proposed and evaluated CNN models for azimuth and elevation estimation in radar target detection, focusing on both single and dual-target scenarios. Our experiments demonstrated that using two-dimensional convolution layers, a specific input shape (including the angle, magnitude of real and imaginary values, and Cartesian x and y values), and dropout layers significantly improved model performance. While we aimed to apply deep learning object detection architectures like YOLO for multiple target detection, the nature of our data—consisting only of complex signal numbers—prevented us from converting it into spatial information suitable for region proposal and object localization. Our approach shows promise compared to traditional radar processing methods and offers a potential direction for future research in radar signal processing.

In future work, we plan to extend our research to the following areas:

- **In-Cabin Detection:** Identifying the presence of children or pets inadvertently left inside a vehicle.
- **Fall Detection:** Monitoring for fall incidents to enhance the safety of the elderly.
- **Posture Detection:** Analyzing posture for health and safety applications, and others.

Additionally, we will explore multi-target scenarios, integration with traditional methods, and the potential of incorporating an autoencoder model into our input layer to better represent radar signals.

7. AUTHOR CONTRIBUTIONS

Ibrahim Riza Hallac: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing.

Deniz Akbaba and Gökhan Gökce: Equally contributed to the development of the methodology, conducting experiments, and reporting the results.

S. Gokhun Tanyer: Proposed the idea, administered the project and supervision, created the datasets, and contributed to review & editing.

Peter F. Driessen: Project administration, funding acquisition, and contributed to review & editing.

REFERENCES

[1] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE Trans. neural networks Learn. Syst., vol. 30, no. 11, pp. 3212–3232, 2019.

[2] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961–2969.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.

[4] W. Kim, H. Cho, J. Kim, B. Kim, and S. Lee, "YOLO-based simultaneous target detection and classification in automotive FMCW radar systems," Sensors, vol. 20, no. 10, p. 2897, 2020.

[5] P. Lang et al., "A comprehensive survey of machine learning applied to radar signal processing," arXiv Prepr. arXiv2009.13702, 2020.

[6] W. Jiang, Y. Ren, Y. Liu, and J. Leng, "Artificial neural networks and deep learning techniques applied to radar target detection: A review," Electronics, vol. 11, no. 1, p. 156, 2022.

[7] L. Zhou, S. Wei, Z. Cui, and W. Ding, "YOLO-RD: A lightweight object detection network for range doppler radar images," in IOP Conference Series: Materials Science and Engineering, 2019, vol. 563, no. 4, p. 42027.

[8] Z. Long, W. E. I. Suyuan, C. U. I. Zhongma, F. Jiaqi, Y. Xiaoting, and D. Wei, "Lira-YOLO: A lightweight model for ship detection in radar images," J. Syst. Eng. Electron., vol. 31, no. 5, pp. 950–956, 2020.

[9] M. Amrani, A. Bey, and A. Amamra, "New SAR target recognition based on YOLO and very deep multi-canonical correlation analysis," Int. J. Remote Sens., vol. 43, no. 15–16, pp. 5800–5819, 2022.

[10] Z. Baird, M. K. Mcdonald, S. Rajan, and S. J. Lee, "A cnn-lstm network for augmenting target detection in real maritime wide area surveillance radar data," IEEE Access, vol. 8, pp. 179281–179294, 2020.

[11] M. Bharat Kumar and P. Rajesh Kumar, “Deep Convolutional Neural Network driven Neuro-Fuzzy System for Moving Target Detection Using the Radar Signals,” J. Inf. \& Knowl. Manag., vol. 21, no. 01, p. 2250010, 2022.

[12] Y. Zhang and Y. Hao, “A survey of SAR image target detection based on convolutional neural networks,” Remote Sens., vol. 14, no. 24, p. 6240, 2022.