Communication

End-to-End Moving Target Indication for Airborne Radar Using Deep Learning

Yao Gu [1], Jianxin Wu [1,*], Yuyuan Fang [1], Lei Zhang [1] and Qiang Zhang [2]

[1] School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China
[2] Beijing Institute of Tracking Telecommunications Technology, Beijing 100094, China
* Correspondence: wujx65@mail.sysu.edu.cn; Tel.: +86-139-9182-0379

Abstract: Moving target indication (MTI) based on space–time adaptive processing (STAP) has been widely used in airborne radar due to its ability for clutter suppression performance. However, the existing MTI methods suffer from the problems of insufficient training samples and low detection probability in a non-homogeneous clutter environment. To address these issues, this paper proposes a novel deep learning framework to improve target indication capability. First, combined with the problems of target indication caused by the non-homogeneous clutter, the clutter-plus-target training dataset was modeled by simulation, where various non-ideal factors, such as aircraft crabbing, array errors and internal clutter motion (ICM), were considered. The dataset considers various realistic situations, making the proposed method more robust. Then, a five-layer two-dimensional convolutional neural network (D^2CNN) was designed and applied to learn the clutter and target characteristics distribution. The proposed D^2CNN can predict the target with a high resolution to implement an end-to-end moving target indication (ETE-MTI) with a higher detection accuracy. In this D^2CNN, the input was obtained by the clutter-plus-target angle-Doppler spectrum with a low-resolution estimated only by a few samples. The label was given by the target angle-Doppler spectrum with a high-resolution obtained by the target's exact angle and Doppler. Thirdly, the proposed method used a few samples to improve the target indication and detection probability, which solved the problem of insufficient samples in the non-homogeneous clutter environments. To elaborate, the proposed method directly implements ETE-MTI without the support of the conventional STAP algorithm to suppress the clutter. The results verify the validity and the robustness of the proposed ETE-MTI with a few samples in the non-homogeneous and low signal-to-clutter ratio (SCR) environments.

Keywords: moving target indication (MTI); space–time adaptive processing (STAP); radar signal processing

1. Introduction

Radar indication technology is necessary for detecting ground/sea and low-altitude moving targets due to its all-day and all-weather capability. Since ground-based radars are susceptible to occlusion effects and low-altitude blind spots, airborne radar has significant advantages for detecting ground/sea and low-altitude moving targets. Moving target indication (MTI) is one of the most critical tasks in airborne radar. MTI is the presence or absence of a moving target with a certain relative velocity in an interesting scenario, also referred to as the cell under test (CUT). However, it is difficult to detect the target due to the severe ground and sea clutter when the airborne radar is working downward-looking. Moreover, one-dimensional filtering techniques based on the conventional moving target indication and moving target detection (MTD) often suffer from ineffective clutter suppression, especially in the non-homogeneous environments. Therefore, an efficient method for clutter suppression and target indication is needed for target detection.

To suppress the clutter and detect the moving target effectively, space–time adaptive processing (STAP) is proposed. Space–time adaptive processing (STAP) utilizes two-dimensional joint adaptive filtering in the spatial and temporal domains to achieve effective

clutter suppression. Currently, STAP technology has been widely used in airborne radar systems [1–3]. In general, the optimal filters for MTI and STAP require a known clutter and noise covariance matrix (CNCM) of the CUT. Since the clutter covariance matrix (CCM) of the CUT in the optimal filter is unknown, Reed et al. [4] proposed an adaptive STAP filter using the sample covariance matrix (SCM) instead of the real CCM, which is called sample matrix inversion (SMI). To obtain an excellent adaptive clutter suppression performance, the training samples and the CUT need to have the same clutter statistical characteristics and to satisfy the independent and identically distributed (IID) condition. However, due to the non-homogeneous environments, the SMI faces two main challenges in practice. First, the samples of the range cells near the CUT may not satisfy the IID condition, resulting in a large performance loss of the simple training sample selection methods. Second, the number of samples with the IID condition among all available training samples is limited and less than two times the system's degrees of freedom. These problems lead to the degradation of adaptive clutter suppression performance, which in turn causes the loss of target detection performance. Therefore, it is of great theoretical significance and practical application to study the adaptive clutter suppression and target indication techniques in non-homogeneous environments.

To solve the problem of insufficient training samples, researchers propose that the problem's impact can be mitigated or overcome with techniques such as training samples selection and single-sample processing. Such methods are collectively referred to as the non-homogeneous STAP, including the classic methods such as Doppler compensation [5] (DC), angle-Doppler compensation [6] (ADC), and adaptive ADC [7] (A2DC). Although these algorithms can improve clutter suppression performance in non-homogeneous environments, there are also some shortcomings. For example, DC, ADC and A2DC all use a single point as a reference for the mainlobe center compensation, which cannot simultaneously compensate the clutter spectrum in all directions. Therefore, the drawbacks degrade the algorithms' MTI performance. At present, the advanced STAP techniques based on knowledge-aid [8–11] (KA) and sparse recovery [12–16] (SR) are also applied in airborne radar MTI, which can reduce the negative effects caused by the clutter non-homogeneous to a certain extent. Moreover, the KA-STAP techniques aim to improve the performance of the conventional STAP algorithms through prior knowledge of various forms and properties. However, the exact form of prior knowledge is difficult to obtain, resulting in a poor real-time performance. Though SR-STAP can effectively reduce the demand for the IID training samples, it is accompanied by a large amount of computation and grid mismatch. Therefore, the existing STAP techniques in practice have limited ability to suppress the clutter due to the insufficient training samples, thus reducing the detection performance. As a result, the emphasis of STAP-MTI is mainly on breaking the limited IID samples in the CCM estimation.

For image processing in the MTI, different approaches have been investigated [17–19]. For signal data processing, STAP adaptively filters the space–time observation (STO) echo data, while the subsequent constant-false-alarm-rate (CFAR) can be considered a two-class classifier in STAP-MTI. The two classes represent the target-present case or the target-absent case. In addition to STAP-MTI, researchers have recently proposed other alternative MTI methods. The MTI method based on the pattern recognition first transforms the traditional filtering problem into the pattern classification. Khatib et al. [20] proposed a STAP method based on least squares for moving target indication (LI-MTI). The method avoids CCM estimation and constructs a classifier identifier to process the radar space–time echo data. To reduce the moving target energy required by LI-MTI, Khatib et al. [21] constructed a polynomial classifier for target indication (POLY-MTI). However, due to the limited fitting and poor feature extraction ability, the above methods need to be further improved in terms of the non-homogeneous clutter environments and low signal-to-clutter ratio (SCR).

In recent years, deep learning technologies, represented by the convolutional neural network (CNN), have developed rapidly, and have gained extensive attention and great success in the field of computer vision [22–24]. Deep learning automatically learns to

extract the hierarchical and expressive features directly from the STO data. It provides new ideas for problems such as radar image processing [25–28] and radar signal processing [29–31]. Recently, deep learning techniques have been applied to clutter suppression in airborne radar. CNN-STAP [32] utilizes the low-resolution clutter angle-Doppler spectrum to reconstruct the high-resolution clutter angle-Doppler spectrum and then calculates the CCM to derive the STAP weight vector. However, this method is aimed at clutter suppression by CNN. In the field of airborne radar, CNN-MTI [31] uses AlexNet to construct a classifier to achieve effective target indication. However, the CNN-MTI method suffers from a large number of network parameters and a low detection accuracy.

Despite its widespread applications and great advantages, deep learning has rarely been applied to angle-Doppler domain estimation tasks in the field of MTI. We propose an end-to-end moving target indication method based on the D^2CNN to improve the target detection capability with a few training samples. First, the established training dataset considered various realistic situations in the non-homogeneous clutter environments, such as aircraft crabbing, array errors, and internal clutter motion (ICM). Then, a D^2CNN with five layers was built to train and fit the network parameters. Finally, the high-resolution target spectrum after training was used to obtain the velocity and space information. To the best of our knowledge, this paper is the first work to apply deep learning techniques to angle-Doppler spectrum estimation for target indication in non-homogeneous clutter environments.

The main contributions of this paper are as follows:

(1) The proposed method can obtain higher detection accuracy using a few samples, which solves the problem of insufficient samples in non-homogeneous clutter environments. The simulation demonstrates that the proposed ETE-MTI has a much lower computational load and a higher detection accuracy in non-homogeneous and low-SCR environments than the existing CNN-MTI [31] method;

(2) The five-layer D^2CNN was constructed with the requirement of the high resolution, which achieved end-to-end target indication to improve the detection accuracy. The D^2CNN's input was built by the clutter-plus-target angle-Doppler spectrum with a low-resolution estimated by a few samples. The label was constructed by the target angle-Doppler spectrum with a high-resolution obtained by the exact angle and Doppler. Once trained, the D^2CNN can be used to predict the target properly with a high resolution using a few samples in near real-time. We also took into account the spatial–temporal sparsity of the clutter and target, which helps network design and training.

The rest of the paper is organized as follows. In Section 2, the space–time signal model is introduced. In Section 3, the deep learning framework and the principle of the proposed ETE-MTI method are proposed. In Section 4, the simulation results and discussion are provided to demonstrate the proposed method's computational efficiency and target detection performance. The conclusions are presented in Section 5.

Notation: Boldface lowercase letters denote vectors and boldface uppercase letters denote matrices. The transposition and conjugate transposition operations are denoted by superscripts T and H, respectively. The symbols $\otimes, \odot$ and * represent the Kronecker product, Hadamard product and convolution, respectively. $E[\cdot]$ is the notation of the expectation operation. $\|.\|_F$ denotes the Frobenius norm.

2. Signal Model

Assume that the antenna array of the airborne phased array pulse radar system with a uniform linear array (ULA) consisting of N elements is moving with constant velocity v at altitude H. The distance between the two adjacent array elements is equal to the half wavelength. Figure 1 shows the model between the ULA and the ground geometry. The pulse repetition frequency is f_r, and M pulses are transmitted at a constant pulse repetition frequency (PRF) during each coherent processing interval (CPI). Set $O - XYZ$ as the carrier coordinate system, where ULA is placed parallel to the Y-axis, and the angle between v

and the Y-axis is θ_{crab}. P is a clutter patch of a certain range cell on the ground plane. The angle of the clutter patch relative to the antenna array is ϕ, and the azimuth and elevation angles relative to the antenna axis are θ and φ, respectively.

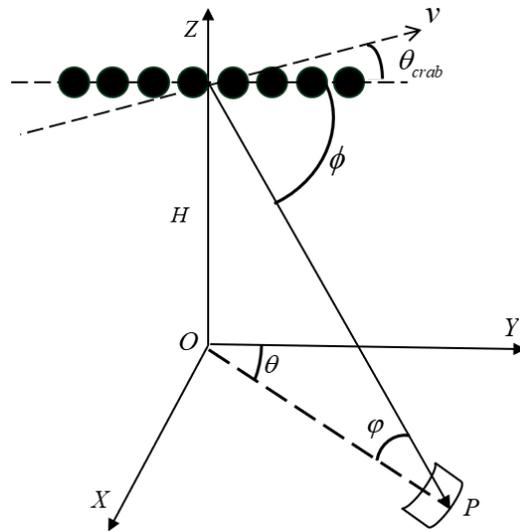**Figure 1.** The geometry of uniform linear array airborne radar.

The space–time snapshot vector **x** can be expressed as:

$$\mathbf{x} = \mathbf{x_c} + \mathbf{x_t} + \mathbf{n}, \tag{1}$$

where $\mathbf{x_t}$ is the target space–time snapshot vector, $\mathbf{x_c}$ is the clutter space–time snapshot vector, and **n** is the complex Gaussian white noise vector.

In the ULA radar system, the target velocity relative to the airborne radar platform is v_t, then the spatial steering vector $v_{s,t}(f_{s,t})$ and the temporal steering vector $v_{d,t}(f_{d,t})$ can be written as:

$$\mathbf{v}_{s,t}(f_{s,t}) = [1, \exp(j2\pi f_{s,t}), \cdots, \exp(j2\pi(N-1)f_{s,t})]^T \tag{2}$$

$$\mathbf{v}_{d,t}(f_{d,t}) = [1, \exp(j2\pi f_{d,t}), \cdots, \exp(j2\pi(M-1)f_{d,t})]^T, \tag{3}$$

where $f_{s,t}(\theta, \varphi) = d\cos(\theta)\cos(\varphi)/\lambda$ and $f_{d,t}(\theta, \varphi) = 2v_t/(\lambda f_r)$ are the normalized spatial frequency (NSF) and normalized Doppler frequency (NDF) of the target, respectively.

The space–time snapshot vector of a single-point target $\mathbf{x}_t$ can be expressed as the multiplication of the complex amplitude σ_t and the corresponding space–time steering vector $\mathbf{v}_t(f_{s,t}, f_{d,t})$ of the target:

$$\mathbf{x}_t = \sigma_t \mathbf{v}_t(f_{s,t}, f_{d,t}), \tag{4}$$

where

$$\mathbf{v}_t(f_{s,t}, f_{d,t}) = \mathbf{v}_{s,t}(f_{s,t}) \otimes \mathbf{v}_{d,t}(f_{(d,t)}). \tag{5}$$

For the clutter scattering point P of a certain range gate, its spatial steering vector $\mathbf{v}_{s,c}(f_{s,c})$ and temporal steering vector $\mathbf{v}_{d,c}(f_{d,c})$ can be described for:

$$\mathbf{v}_{s,c}(f_{s,c}) = [1, \exp(j2\pi f_{s,c}), \cdots, \exp(j2\pi(N-1)f_{s,c})]^T \tag{6}$$

$$\mathbf{v}_{d,c}(f_{d,c}) = [1, \exp(j2\pi f_{d,c}), \cdots, \exp(j2\pi(M-1)f_{d,c})]^T \tag{7}$$

where $f_{s,c}(\theta, \varphi) = d\cos(\theta)\cos(\varphi)/\lambda$ and $f_{d,c}(\theta, \varphi) = 2v\cos(\theta + \theta_{crab})\cos(\varphi)/(\lambda f_r)$ are the clutter patch's NSF and NDF.

Considering the non-ideal factors in non-heterogeneous clutter environments with array errors and internal clutter motion (ICM), the clutter space–time snapshot vectors of all range cells are the accumulation of the echo signal of each clutter block at different

ambiguous ranges. Assuming that each clutter scattering point is statistically independent, the clutter space–time snapshot is defined as:

$$
\begin{aligned}
x_{\mathrm{c}} = \sum_{p=1}^{N_a} \sum_{q=1}^{N_c} a(\theta_q, \varphi_p) \left[\kappa \odot \mathbf{v}_{d,c}(\theta_q, \varphi_p) \right] \\
\otimes \left[\varepsilon_{s,c}(\theta_q, \varphi_p) \odot \mathbf{v}_{s,c}(\theta_q, \varphi_p) \right],
\end{aligned}
\tag{8}
$$

where N_a, N_c, $a(\theta_q, \varphi_p)$ denote the number of ambiguous range rings, the number of spurious scattering points on a single range ring, and the complex scattering amplitude of the qth spurious scattering point on the pth ambiguous range ring, respectively; $\kappa = [\kappa_1, \kappa_2, \ldots, \kappa_M]^T$ represents the real temporal weight vector brought by ICM; $\varepsilon_{s,c}(\theta, \varphi) = [\varepsilon_1(\theta, \varphi), \varepsilon_2(\theta, \varphi), \cdots, \varepsilon_N(\theta, \varphi)]^T$ represents the real spatial weight vector caused by the array errors. $\varepsilon_i(\theta, \varphi)$ obeys the complex Gaussian distribution with mean-zero and variance σ_e^2.

Since each clutter block is statistically independent and $a(\theta, \varphi)$ is a Gaussian random variable with mean-zero and variance $\sigma_c^2(\theta, \varphi)$, the corresponding CCM of this clutter data is defined as:

$$
\begin{aligned}
\boldsymbol{R}_c &= E\left\{ x_c x_c^H \right\} \\
&= \sum_{n=1}^{N_a} \sum_{a=1}^{N_c} \sigma_c^2(\theta_q, \varphi_p) \left[\boldsymbol{T}_d \otimes \boldsymbol{T}_s \right] \odot \left[\mathbf{v}(\theta_q, \varphi_p) \mathbf{v}^H(\theta_q, \varphi_p) \right],
\end{aligned}
\tag{9}
$$

where $\mathbf{v}(\theta, \varphi) = \mathbf{v}_{d,c}(\theta, \varphi) \otimes \mathbf{v}_{s,c}(\theta, \varphi)$ denotes the clutter space–time steering vector; the time autocorrelation matrix is $\boldsymbol{T}_d = \mathrm{Toeplitz}(r_d(0), r_d(1), \cdots, r_d(M-1))$ due to the ICM, where $r_d(m) \triangleq E\{\kappa_{i+m} \kappa_i^*\} = \exp\left\{ -\frac{8\pi^2 \sigma_v^2 m^2}{\lambda^2 f_r^2} \right\}$ $i = 0, 1, \cdots, M-1$; σ_v^2 represents the variance of the spreading of the clutter spectrum caused by the wind speed and λ denotes the wavelength; $\boldsymbol{T}_s = E\{\varepsilon_{s,c}(\theta, \varphi) \varepsilon_{s,c}(\theta, \varphi)^H\}$ denotes the spatial autocorrelation matrix caused by the array errors.

In general, the CCM is unknown, so it is usually obtained by maximum likelihood estimation (MLE) using the adjacent datasets of the CUT as training samples. Hence, the corresponding covariance matrix can be represented by:

$$
\widehat{\mathbf{R}}_{\mathrm{MLE}} = \frac{1}{L} \sum_{l=1}^{L} \mathbf{x}_l \mathbf{x}_l^H,
\tag{10}
$$

where L is the number of training samples. $\mathbf{x}_l$ represents the STO data of the lth training sample.

According to the RMB rule, the number of training samples must be at least twice the number of the system degrees of freedom to keep the loss of SNR within 3 dB. After obtaining the CCM, the space–time adaptive optimal weight vector can be obtained:

$$
\mathbf{w} = \mu \hat{\boldsymbol{R}}_{\mathrm{MLE}}^{-1} \mathbf{v}_t(f_{s,t}, f_{d,t}),
\tag{11}
$$

where $\mu = 1/(\mathbf{v}_t^H \mathbf{R}^{-1} \mathbf{v}_t)$ is the normalization constant. It can be seen that if the estimated CCM is inaccurate, the calculated space–time adaptive filter weight vector and the theoretical STAP optimal filter weight vector have a large gap in the clutter suppression performance, which will affect the performance of subsequent target detection.

Due to the severe clutter, noise and jamming, the moving target is always buried in the interference. The goal of MTI is to detect the moving target's Doppler frequency and spatial frequency from the STO. In this paper, we make use of the D^{2}CNN to learn the distribution characteristics of the clutter and the target. The D^{2}CNN extracts information about the target directly from the clutter-plus-target spectrum. Hence, the proposed method avoids reconstructing the clutter spectrum to achieve the end-to-end target indication for airborne radar.

3. Proposed Method

3.1. Whole Framework of Proposed Method

In essence, ETE-MTI can be viewed as a classification problem, where the pairing of NDF and candidate NSF of the moving target are considered as one class. Furthermore, the clutter and target are separable in the space–time domain. As a result, through the mapping characteristics of deep learning, the target and clutter are distinguished in the space–time domain. Therefore, the clutter is actually filtered out and the target can be better indicated to improve the detection.

The whole framework is shown in Figure 2. In the framework of the proposed method, there are two main steps to obtain the high-resolution target angle-Doppler spectrum.
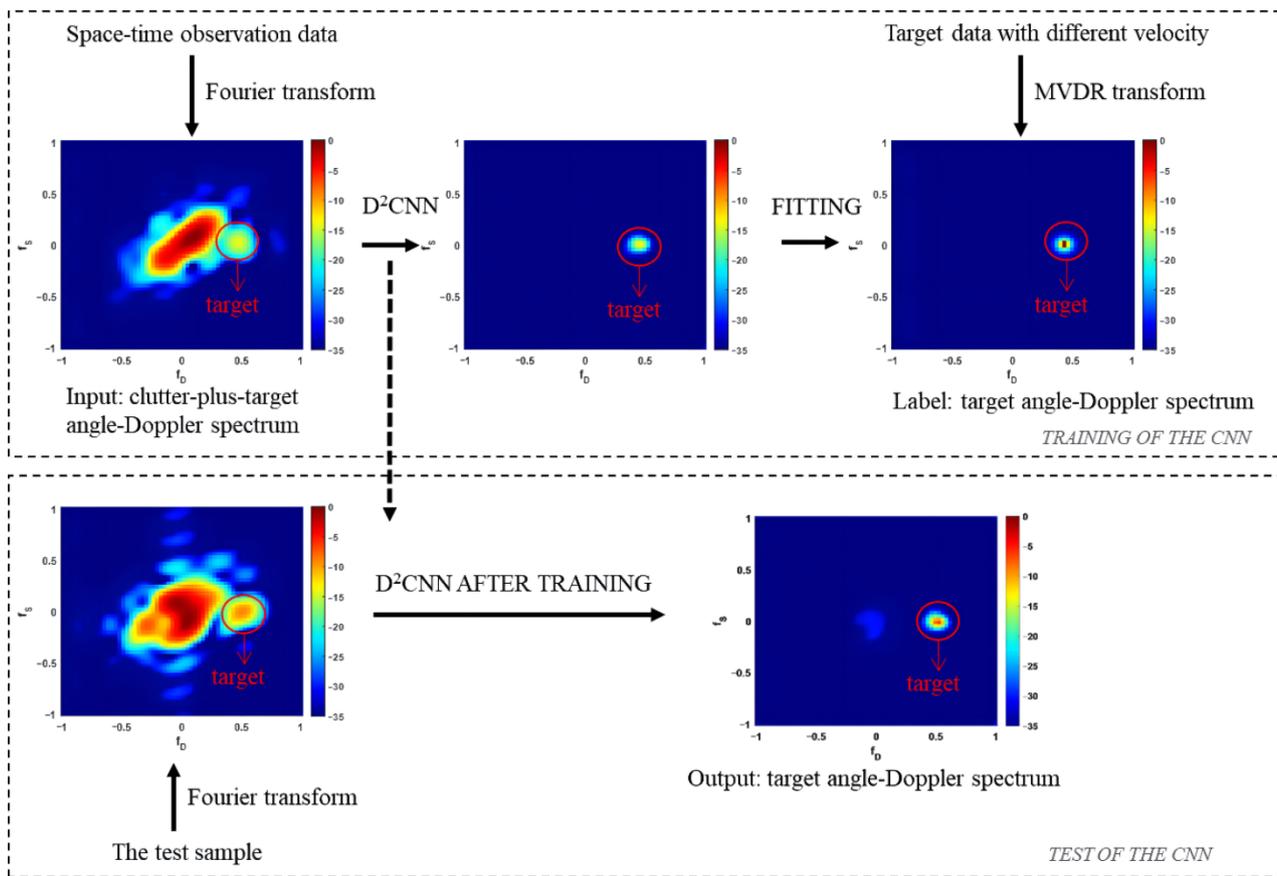

Figure 2. Deep learning framework for ETE-MTI.

Firstly, we can discretize the angle-Doppler plane into $N_s = \rho_s N$ and $N_d = \rho_d M(\rho_s, \rho_d \gg 1)$ cells, where ρ_s and ρ_d are the angle and Doppler frequency discretization factors, respectively. Then, the collection of all steering vectors in the two-dimensional space–time plane is given by:

$$\mathbf{V} = \left[\mathbf{v}(f_{s,1}, f_{d,1}), \ldots, \mathbf{v}(f_{s,N_s}, f_{d,1}), \ldots, \mathbf{v}\left(f_{s,N_s}, f_{d,N_d}\right)\right], \tag{12}$$

where $f_{s,i}, 1 \leq i \leq N_s$ and $f_{d,k}, 1 \leq k \leq N_d$ denote the normalized spatial and Doppler frequencies, respectively.

The power spectrum estimation is performed on the training sample of the STO data $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_L] \in C^{NM \times L}$. $P(f_{s,i}, f_{d,k})$ is the spectrum intensity of the corresponding grid. Therefore, the Fourier spectrum transform can be defined as:

$$P(f_{s,i}, f_{d,k}) = \left(\mathbf{v}(f_{s,i}, f_{d,k})^H \mathbf{X}\right)\left(\mathbf{v}(f_{s,i}, f_{d,k})^H \mathbf{X}\right)^H = \mathbf{v}(f_{s,i}, f_{d,k})^H \widehat{\mathbf{R}}_{\text{MLE}} \mathbf{v}(f_{s,i}, f_{d,k}). \tag{13}$$

The Fourier spectrum transform plays an important role in the network's input. According to Figure 2, it converts STO data into the form of the angle-Doppler spectrum as the network's input.

Similarly, the Minimum Variance Distortionless Response (MVDR) spectrum transform is represented by:

$$P(f_{s,i}, f_{d,k}) = \frac{1}{\mathbf{v}(f_{s,i}, f_{d,k})^H \widehat{\mathbf{R}}_{\mathrm{MLE}}^{-1} \mathbf{v}(f_{s,i}, f_{d,k})}. \tag{14}$$

The MVDR spectrum transform converts the target data with the exact angle and Doppler into the form of the angle-Doppler spectrum as the network's label.

In this paper, the angle-Doppler spectrum is obtained by superimposing each grid's spectrum intensity. Therefore, the angle-Doppler spectrum can be represented by:

$$\mathbf{P}[\mathbf{X}] = \sum_{i=1}^{N_s} \sum_{k=1}^{N_d} P(f_{s,i}, f_{d,k}). \tag{15}$$

The Rayleigh resolution limits the Fourier spectrum transform. However, the MVDR spectrum transform has a high resolution due to its ability to break the Rayleigh limit. Based on these properties, CNN-STAP [32] and SR-CNN [33], we use the low-resolution clutter-plus-target angle-Doppler spectrum as the network's input. The D^{2}CNN is a specific neural network for reconstructing and filtering the input so that we can obtain the expected high-resolution target angle-Doppler spectrum of the output. The task of achieving the high-resolution target angle-Doppler spectrum can be formulated as a supervised deep learning problem. The whole mathematical model process is given by:

$$\mathbf{Z} = F[\mathbf{P}[\mathbf{X}]], \tag{16}$$

where $\mathbf{Z} \in R^{N_s \times N_d}$ is the expected target high-resolution space–time spectrum at the network's output. $F : R^{N_s \times N_d} \to R^{N_s \times N_d}$ characterizes the D^{2}CNN operator.

Consequently, there are two stages in the deep learning from Figure 2. In the proposed D^{2}CNN, the input was constructed by the clutter-plus-target angle-Doppler spectrum with a low-resolution estimated by a few samples according to Equation (13). The label was constructed by the target angle-Doppler spectrum with a high-resolution obtained by the exact spatial and Doppler frequency according to Equation (14). In the training stage, the training data set was used for the D^{2}CNN parameter optimization and fitting. Once trained, the D^{2}CNN can be used to predict the target high-resolution angle-Doppler spectrum using a few samples in near-real-time in the test stage.

3.2. Construction of D^{2}CNN

Based on the CNN-STAP [32], we constructed the convolutional neural network structure, as shown in Figure 3. The network consisted of five convolution layers. The input was the low-resolution clutter-plus-target spectrum estimated by Fourier spectrum transform, and the output was the high-resolution target spectrum estimated by MVDR spectrum transform after filtering out the clutter and noise.

The low-resolution angle-Doppler spectrum contains the clutter-plus-target rough information of the actual position and energy distribution. Its characteristics are more intuitive and effective. Therefore, the characteristics of the training samples can be extracted in the first layer:

$$F_1 = \max(0, \mathbf{W}_1 \times \mathbf{Y} + \mathbf{b}_1), \tag{17}$$

where $\mathbf{Y} = \mathbf{P}[\mathbf{X}]$; $\mathbf{W}_1$ and $\mathbf{b}_1$ denote the convolution kernel and bias, respectively; $\mathbf{W}_1$ is of a size $c \times f_1 \times f_1 \times n_1$, where c, f_1 and n_1 denote the number of input image channels, the size of the kernels, and the number of convolution kernels, respectively.
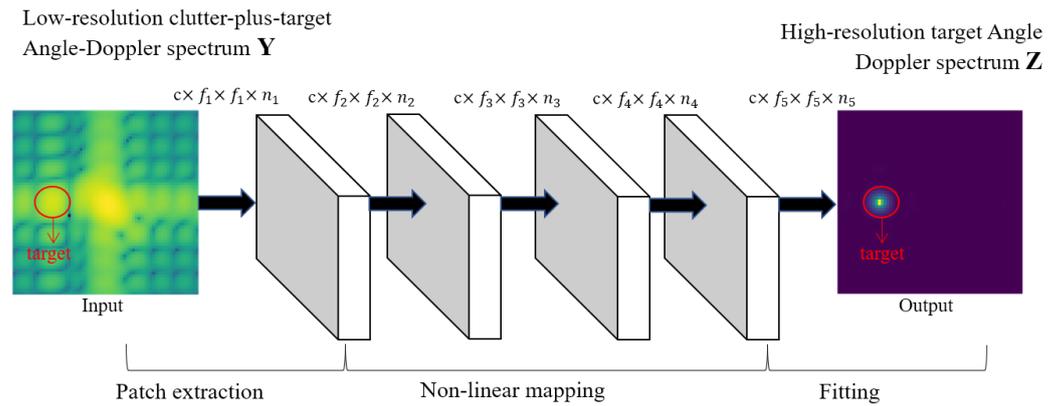
Figure 3. Architecture of the proposed convolutional neural network.

The five convolutional layers all utilize the ReLU activation function, which acts as feature extraction and high-dimensional mapping. The edge-complementary zero operation ensures that each layer's input and output images are the same sizes. The second to fourth layers are all features nonlinear mapping where the extracted feature is mapped nonlinearly into the transformed high-dimensional:

$$F_i = \max(0, \mathbf{W}_i \times F_{i-1} + \mathbf{b}_i) \quad i = 2, 3, 4, \tag{18}$$

where $\mathbf{W}_i$ denotes a size of $n_{i-1} \times f_i \times f_i \times n_i$ and $\mathbf{b}_i$ is an n_i-dimensional vector. The fifth layer is the image reconstruction layer, which generates the high-resolution output image:

$$\mathbf{Z} = \mathbf{W}_5 \times F_4 + \mathbf{b}_5, \tag{19}$$

where $\mathbf{W}_5$ is of a size $n_4 \times f_5 \times f_5 \times c$. $\mathbf{b}_5$ is a c-dimensional vector.

Assume that the low-resolution clutter-plus-target angle-Doppler spectrum is the input $(\mathbf{Y}_t)_{t=1}^{T}$, and the high-resolution target angle-Doppler spectrum is the label $(\hat{\mathbf{Z}}_t)_{t=1}^{T}$. $(\mathbf{Y}_t)_{t=1}^{T}$ and $(\hat{\mathbf{Z}}_t)_{t=1}^{T}$ are passed through the minimization model mean squared error (MSE), resulting in a nonlinear mapping relationship between the label and output:

$$\text{Loss}(\boldsymbol{\Theta}) = \frac{1}{T} \sum_{t=1}^{T} \left\| F(\mathbf{Y}_t; \boldsymbol{\Theta}) - \hat{\mathbf{Z}}_t \right\|_F^2, \tag{20}$$

where T is the number of the training data. $\boldsymbol{\Theta} = \{\mathbf{W}_i, \mathbf{b}_i\}, i = 1, 2, \cdots, 5$ are the network parameters, while the stochastic gradient descent method is used to update the parameters.

3.3. Construction of Training Dataset

The inputs $(\mathbf{Y}_t)_{t=1}^{T}$ and the labels $(\hat{\mathbf{Z}}_t)_{t=1}^{T}$ should be included in the training dataset, which is defined as:

$$\Gamma = (\mathbf{Y}_t, \hat{\mathbf{Z}}_t)_{t=1}^{T}. \tag{21}$$

In the proposed method, we first apply a beamforming procedure, using $\mathbf{V}$ in Equation (12), to the clutter-plus-target echo data $\mathbf{X}$, which constructs an initial clutter-plus-target angle-Doppler spectrum $\mathbf{Y}$. Consequently, the input $\mathbf{Y}$ clutter-plus-target angle-Doppler spectrum is constructed by the Fourier transform in Equation (13). In the CNN, the label $\hat{\mathbf{Z}}$ target angle-Doppler spectrum is constructed by the MVDR transform in Equation (14), which has a high-resolution performance. Therefore, the input uses the clutter-plus-target covariance matrix, and the label uses the target covariance matrix in Equation (10). As a result, to improve the target detection and suppress the clutter, we apply the D^{2}CNN to the intermediate reconstruction and filter, which outputs a high-resolution

target angle-Doppler spectrum **Z** according to Equation (16). In a word, this process can be viewed as a supervised deep learning problem.

In the following simulation experiments, we artificially generated sufficient training dataset Γ using samples from four range cells adjacent to the CUT. For simplicity, the NSF of the expected target was known, and the NDF varied between $[-1, 1]$. The experiments used two datasets corresponding to STO's ideal and non-ideal cases to fully validate the ETE-MTI performance. For the ideal case, the dataset was generated for the simulation, of which 80% was used as the training dataset and the remaining 20% was used as the validation dataset to verify the performance of the network. In the airborne radar system, aircraft crabbing, array errors and ICM will affect the clutter distribution on the angle-Doppler spectrum, thereby affecting the target indication. Therefore, the values of each non-ideal factor parameter, such as the array errors $\sigma_e^2 \in [0, 0.2]$, the ICM $\sigma_v^2 \in [0, 0.2]$ and the aircraft crabbing angles $\theta_{crab} \in [0, 5°]$ can be randomly selected to generate the clutter space–time snapshot vector to construct the dataset. Additionally, the SNR was set to between 20 dB and 60 dB in order to verify that the method can obtain good detection performance even at low SCR environments. Similarly, 80% of the dataset was used for training and 20% was used for validation.

4. Results and Discussion

In this section, simulation experiments were used to verify the effectiveness of the proposed method. The simulation parameters are listed in Table 1. The number of used training samples was 4. The angle frequency discretization factor ρ_s was 6 and the Doppler frequency discretization factor ρ_d was 6. The network parameters were given as: the number of channels c is 1 and $f_i \times f_i \times n_i, i = 1, 2, \ldots, 5$ are set to $11 \times 11 \times 16, 9 \times 9 \times 8, 7 \times 7 \times 4, 5 \times 5 \times 2, 3 \times 3 \times 1$, respectively. Meanwhile, the learning rate was set to 10^{-2}. Moreover, the pairs dataset was used for training with a batch size of 64. Furthermore, we conducted the experiment using an AMD Ryzen 7 5700 G with Radeon Graphics CPU.

Table 1. Simulation parameters of the radar system.

Parameter	Value
Platform height	7 km
Platform velocity	150 m/s
Element number	8
Pulses in one CPI	8
Element spacing	0.15 m
Pulse repetition frequency	4k Hz
CNR	50 dB
Noise power	1 W

4.1. Convergence Analysis

This subsection analyzes each network's overall training and validation MSEs concerning the number of iterations. Figure 4 presents the variation of the training and validation MSEs with the training iterations in the ideal and non-ideal cases. Two networks were trained for 350 and 400 iterations, respectively. The training MSE in both the ideal and non-ideal cases decreases rapidly in the early training period and essentially reaches convergence at the 300th training iteration with only minor changes in the subsequent training iterations. In addition, the network converges faster in the ideal case than in the non-ideal case, since the training dataset in the ideal case does not contain other non-ideal factors. The clutter distribution is relatively single. Therefore, ETE-MTI can quickly learn the distribution characteristics between the clutter and the target. In contrast, in the non-ideal case, the clutter-plus-target contains various non-ideal factors. So the clutter spectrum distribution is complicated to affect the target indication, which makes ETE-MTI need a longer period to learn. Moreover, the validation curves level off after about 150 iterations

and remain roughly constant thereafter. The result confirms that there is no overfitting in the two networks.
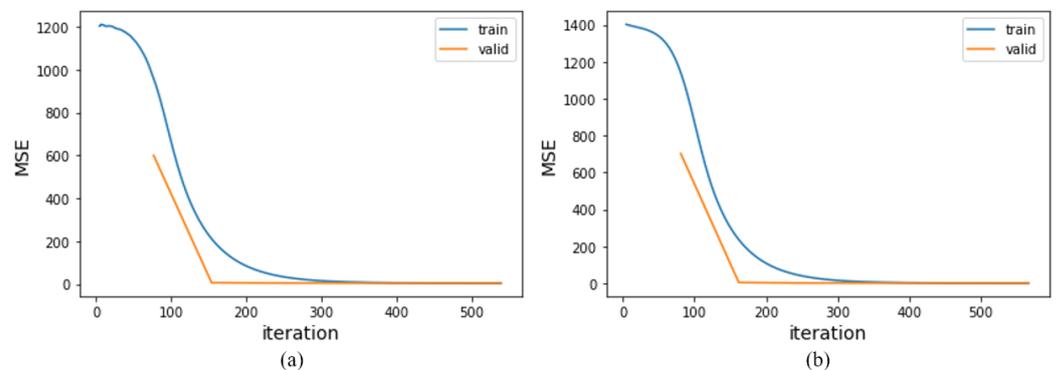

Figure 4. Training and validation MSE versus the number of iterations. (**a**) ideal case; (**b**) non-ideal case.

4.2. Visualization of Prediction Results

This subsection analyzes the prediction performance of ETE-MTI. For simplicity, the NSF was made to be 0.

If the clutter and the target were easily distinguishable on the space–time spectrum, the target's NDF was set to 0.556. Figure 5 shows the predicted target angle-Doppler results. Figure 5a,b show the clutter-plus-target and the target spectrum in the ideal case. The target was estimated by the proposed method. ETE-MTI can predict the target position well without the clutter remaining, realizing end-to-end target indication. The prediction performance in the case of the aircraft crabbing angle $\theta_{crab} = 5°$ is shown in Figure 5c,d. The clutter spectrum is bent due to the influence of the aircraft crabbing and is mixed with a part of the target in Figure 5c. Nonetheless, It can be seen that, from Figure 5d, the expected target can be detected after the CNN, but there is a bit of residual clutter at the zero Doppler position. As shown in Figure 5e, in the presence of array errors, the energy of the clutter spectrum leaks along the angle direction and undergoes spectral broadening. The predicted result in the case of crabbing is shown in Figure 5f. Although the target can be indicated, there is relatively more clutter remaining at zero Doppler along the angle direction. Figure 5g,h show the clutter-plus-target and the target Fourier spectrum in the case of ICM. The target was estimated by the proposed method. As shown in Figure 5g, the clutter spectrum is broadened due to the wind speed. The predicted result is shown in Figure 5h, that the target can still be indicated with NSF = 0.556 after the deep learning network.

In the following, we discuss the performance when the target is close to the mainlobe of the clutter. The target's NDF was set to 0.1429. Figure 6 shows the predicted target angle-Doppler results. Figure 6a,b show the clutter-plus-target Fourier spectrum and the predicted target spectrum in the ideal case. The target was buried in the clutter with the high power; ETE-MTI could still predict the target after the trained network, but the target's power was weakened at this time. In the non-ideal case, the factor parameters were set to the array error $\sigma_e^2 = 0.1$, the ICM $\sigma_v^2 = 0.2$, and the aircraft crabbing angle $\theta_{crab} = 5°$. Figure 6c,d show the clutter-plus-target Fourier spectrum and the predicted target in the non-ideal case. As is shown in Figure 6c, although the clutter spectrum is completely mixed with the target due to the bending, energy leakage and spectral broadening because of the aircraft crabbing, array errors and ICM, the expected target can be indicated after the deep learning from Figure 6d.
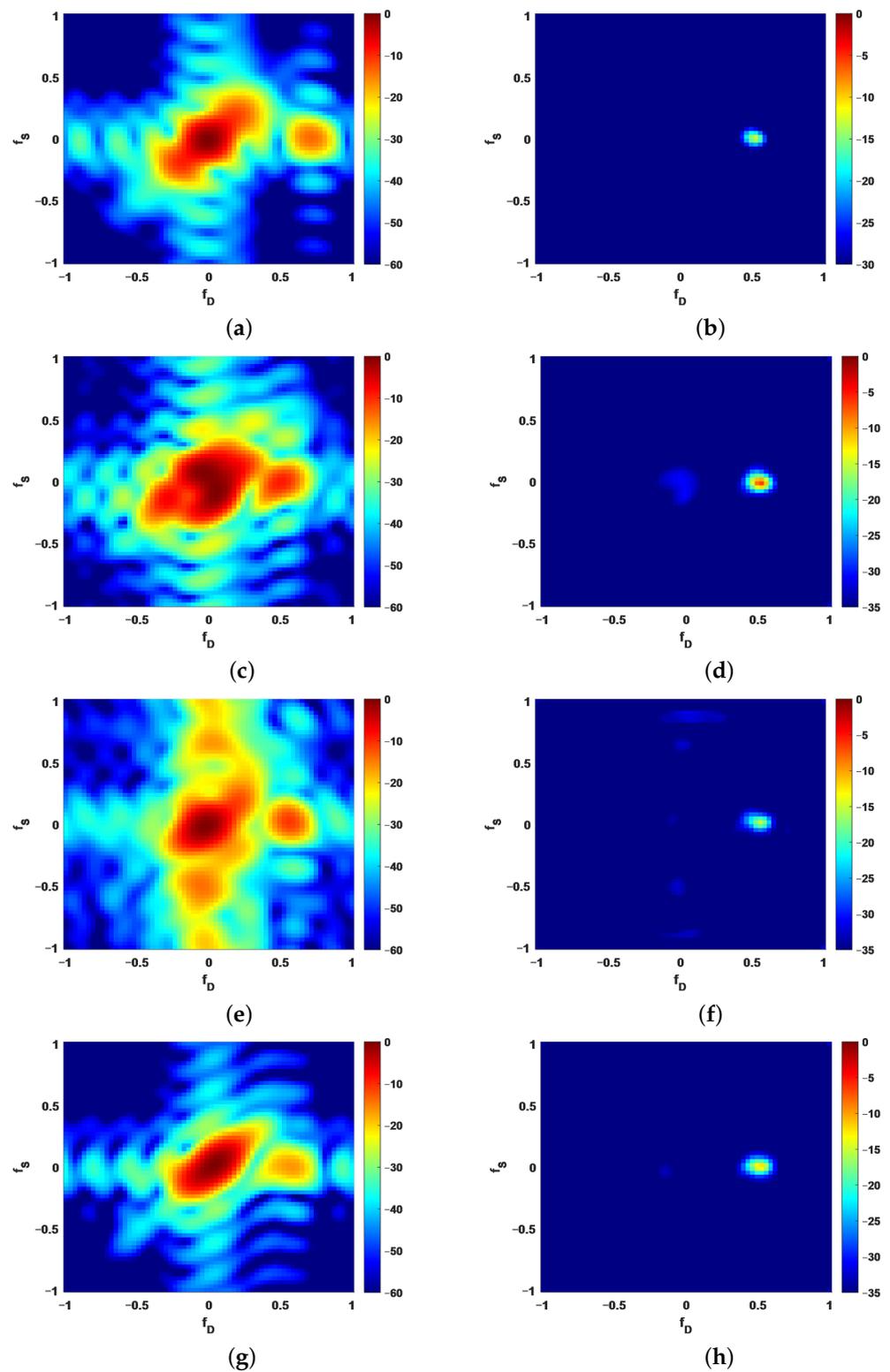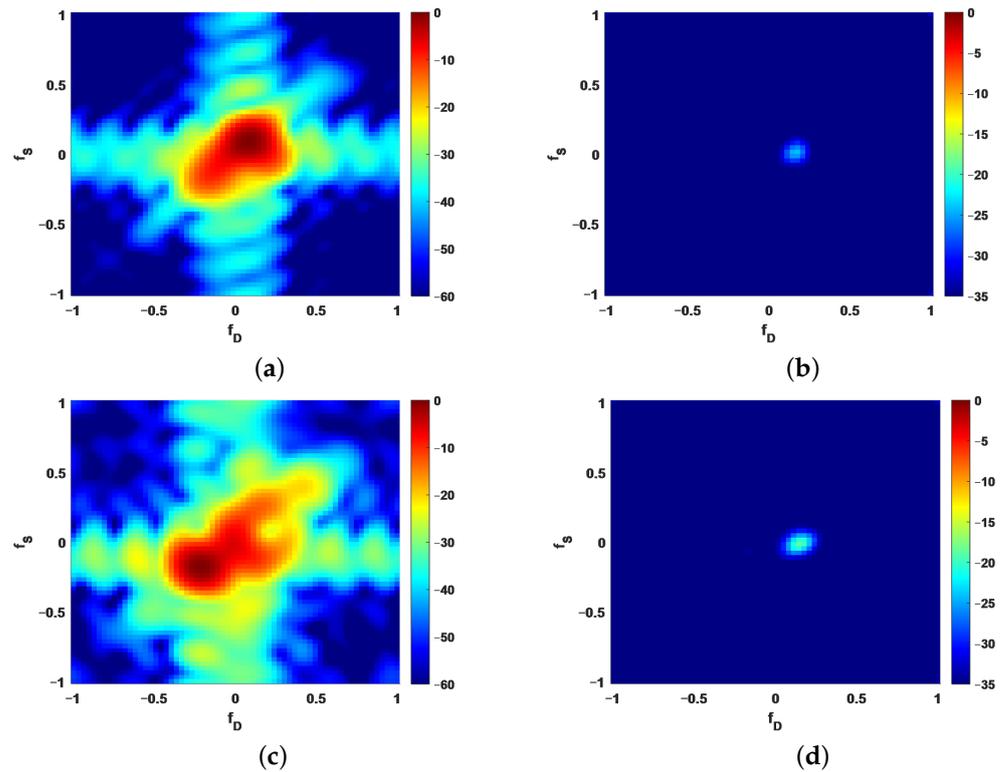
Figure 5. Processing results of different clutter environments with SCR = −15 dB in the case where the target and clutter are distinguishable. (**a**,**b**) ideal case; (**c**,**d**) in the presence of aircraft crabbing; (**e**,**f**) in the presence of spatial error; (**g**,**h**) in the presence of ICM.

Figure 6. Processing results of different clutter environments with SCR = −15 dB in the case where the target and clutter are indistinguishable. (**a**,**b**) ideal case; (**c**,**d**) non-ideal case.

As a result, when the target is buried and covered by the clutter with high power or the target is at low speed, ETE-MTI can quickly learn the spatial–temporal distribution characteristics of the clutter and the target through the neural network to extract the target information, realizing the end-to-end target indication.

4.3. Detection of Probability under Different SCR Scenarios

In this subsection, we evaluate the target detection performance of different NDFs by the probability of detection (PD) versus SNR curves. There are 31 artificially generated test datasets with different SCRs, which are produced by the different target powers under the same clutter power of 50 dB. In different test datasets, the targets' powers varied from 20 dB to 60 dB with equal intervals. In each dataset, 1000 test samples were generated by adding the target signals with the same power and candidate NDFs to the clutter. The samples from each test dataset were fed into the trained D^2CNN. The detection performance was evaluated by PD which were obtained by using the adaptive matched filter (AMF) detector. PD is the average percentage of correctly classified test samples for each target in the test dataset. Figure 7 shows the effect of non-homogeneous clutter on detection performance. Two cases are also considered in Figure 7. In the non-ideal case, the non-ideal factors were set to the array error $\sigma_e^2 = 0.1$, the ICM $\sigma_v^2 = 0.1$, and the aircraft crabbing angle $\theta_{crab} = 5°$. The target's NSF was fixed to 0, while the NDF considers three values; 0.167, 0.367 and 0.5, respectively.

As depicted in Figure 7a,b, with the increase of SCR, the detection performance of the ETE-MTI method has improved. The three curves indicate that the ETE-MTI method have superior target detection performance whether in the mainlobe region ($f_{dt} = 0.1667$) or in the sidelobe region ($f_{dt} = 0.367$ or $f_{dt} = 0.5$) at the high SCR conditions. The PD approximately approaches 100% in the sidelobe region ($f_{dt} = 0.367$ or $f_{dt} = 0.5$) with the SCR of −15 dB. As the target's NDF increases, the proposed method's detection performance improves. It can be seen that the detection performance of the proposed method in the sidelobe region

is better than that in the mainlobe region. It will degrade the target detection performance when the clutter exists with non-ideal factors. From Figure 7a,b, compared with the PD curves in the ideal case, the PD in the non-ideal case is slightly decreasing, although in non-homogeneous clutter environments, the PD can remain above 100% in the sidelobe region ($f_{dt} = 0.367$ or $f_{dt} = 0.5$) at -10 dB SCR. Thus, the results demonstrate that the ETE-MTI method has a good detection performance in the non-homogeneous clutter environments and low SCR conditions.
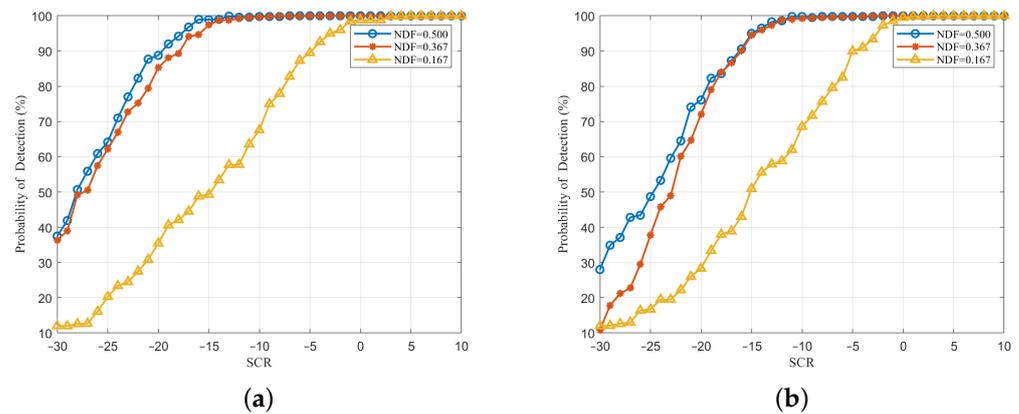

Figure 7. PD versus SCR curves of different NDFs. (**a**) ideal case; (**b**) non-ideal case.

4.4. Comparison of Computation Complexity

The calculation burden mainly comes from convolution operations during the D^{2}CNN's training and test. For the mentioned D^{2}CNN, the component complexity formula is as follows [34]:

$$O\left(\sum_{l=1}^{C} n_{l-1} \cdot s_l^2 \cdot n_l \cdot m_l^2\right), \tag{22}$$

where l is the index of a convolutional layer, and C is the depth. n_l is the number of filters in the l-th layer. n_{l-1} represents the number of input channels of the l-th layer. s_l is the spatial size (length) of the filter. m_l is the spatial size of the output feature map. The calculation complexity of the ETE-MTI method is obtained by substituting the network parameters set in this paper into Equation (22). According to Table 1 and Figure 3, the computation complexity of the proposed ETE-MTI is in the order of $O(10^5 MN)$. However, the computation complexity of the CNN-MTI method is $O(10^6 MN)$, which is one order of magnitude more than the method proposed in this paper.

4.5. Comparison of Detection of Probability

In this subsection, PD verifies the detection performance of different methods. First, we evaluated the proposed ETE-MTI method's detection performance under different doppler channels' PD compared with other methods. Other conditions were the same; the target's power was set to 30 dB and the target's velocity in different test datasets varied from -150 m/s to 150 m/s. Fifteen test datasets with different target velocities were generated, which corresponded to the 15 Doppler channels. The results of the traditional optimal method (OPT-STAP-MTI) and the CNN-MTI method in [31] for comparison were used to verify the ETE-MTI method's accuracy and effectiveness. ETE-MTI used four IID data range cells, CNN-MTI used 105 IID data range cells around the CUT, and OPT-STAP-MTI used all range cells. The PD of the three methods in different Doppler channels were compared as shown in Figure 8. ETE-MTI had the lowest PD of 53% in the zero Doppler channel since the clutter entirely buried the target. As the target velocity increased, the target was further and further away from the main lobe of the clutter spectrum. Therefore, the distinguishability between the target and the clutter increased and ETE-MTI could detect the target more accurately with PD of up to 100%.
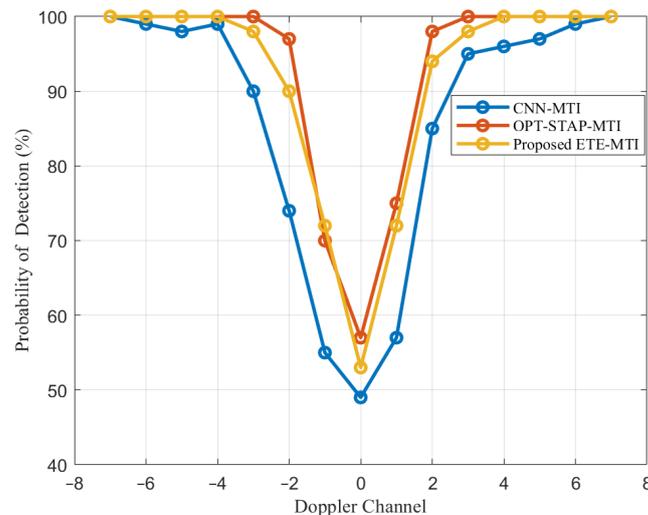
Figure 8. Performance comparison of the proposed ETE-MTI, the OPT-STAP-MTI and the CNN-MTI method [31].

It can be observed that the detection performance of CNN-MTI is poor, and its PD is lower than that of the other two methods in the zero Doppler channel. The detection performance of all three methods improves as the Doppler channel increases, and ETE-MTI and STAP-MTI can detect the target with the PD of 100% in multiple Doppler channels. Moreover, ETE-MTI and STAP-MTI can detect the target when the target's velocity is low. The reason for the improved detection performance of the three methods is that, as the Doppler channel increases, the target velocity increases relative to the stationary clutter. Hence, the clutter and the target can be distinguished in the spectrum, making it easier to detect the target.

The comparison shows that the average PD of ETE-MTI exceeds that of CNN-MTI. Moreover, the ETE-MTI 's PD curve is very close to that of OPT-STAP-MTI. Thus, the results demonstrate that ETE-MTI can achieve an excellent performance under different Doppler channels and excels in detecting low-speed targets. Furthermore, the ETE-MTI method will outperform the traditional STAP method when the training sample is limited.

In addition, we compared the detection performance of the proposed method ETE-MTI and CNN-MTI with different SCRs. The target's NDF was set as 0.367 and the number of test samples was 1000. The test samples' generation was the same as in Section 4.3. The performance comparison is shown in Figure 9. The PD of both methods gradually improves with the increase of the SCR. The highest PD of ETE-MTI can reach 100%, while the highest PD of CNN-MTI is close to 92%. The detection performance of the proposed ETE-MTI is better than that of CNN-MTI at low SCRs. Therefore, the result demonstrates that the proposed ETE-MTI has a much lower computational load and a higher detection accuracy than the existing CNN-MTI method with a few samples in the non-homogeneous and low SCR environments.

Consequently, the two methods—ETE-MTI and CNN-MTI—differ in the form of the data entered. The proposed method's input is the power spectrum amplitude data of the clutter-plus-target. In CNN-MTI, the input is the space–time observation data. Furthermore, the five-layer D^2CNN built in this paper considers the target's high resolution for target indication, allowing our method to detect the target more easily in the non-homogeneous clutter and low SCR environments. From the results, the proposed simpler D^2CNN with less computation is more efficient in learning the power spectrum amplitude data and therefore has a better detection performance.
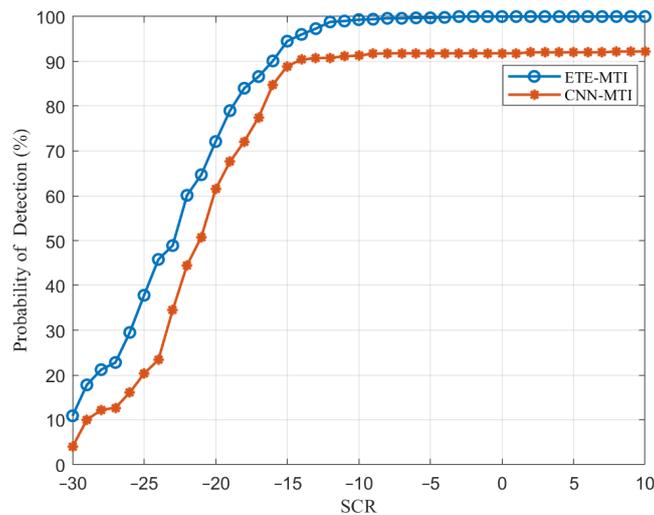
Figure 9. PD versus SCR curves of the proposed ETE-MTI and the CNN-MTI method [31] .

5. Conclusions

This paper proposes an end-to-end moving target indication method for airborne radar based on deep learning. First, we constructed the training dataset including non-ideal factors in non-homogeneous clutter environments. In the dataset, the low-resolution clutter-plus-target spectrum was considered as the D^2CNN's input, which was estimated by a few samples to solve the problem of insufficient samples. Then, the high-resolution target spectrum is taken as the D^2CNN's label. Secondly, the proposed five-layer D^2CNN is established to extract the input's feature. Finally, once the clutter and target distribution characteristics are learned, the D^2CNN can predict the target space–time information from the output's high-resolution spectrum, realizing the end-to-end moving target indication. The D^2CNN with five layers is in consideration of the high-resolution requirements, which can improve the target detection. Furthermore, unlike other traditional STAP technologies, the proposed method mainly uses the D^2CNN's mapping characteristics to complete clutter filtering to realize the target indication directly. The results demonstrate that the proposed ETE-MTI with a few samples has a much lower computational load and a higher detection accuracy in non-homogeneous and low-SCR environments than the existing CNN-MTI [31] method.

The limitation of the proposed method is that it has studied the target indication performance in the non-homogeneous environments for the time being. Target indication in the heterogeneous environments is the next research goal. In our future research, the more realistic physical effects, such as heterogeneous clutter environments, should also be considered to validate the robustness of our method.

Author Contributions: Conceptualization, Y.G. and J.W.; methodology, L.Z.; software, Y.G.; validation, Y.G., J.W. and Y.F.; formal analysis, Y.G.; investigation, Y.G.; resources, J.W.; data curation, L.Z.; writing—original draft preparation, Y.G.; writing—review and editing, Y.G., J.W. and Y.F.; visualization, Y.G. and Y.F.; supervision, J.W.; project administration, J.W.; funding acquisition, Q.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

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

References

1. Wang, H.; Yao, Z.; Fan, Z.; Yang, J.; Liu, G. A robust STAP beamforming algorithm for GNSS receivers in high dynamic environment. *Signal Process.* **2020**, *172*, 107532.1–107532.10. [CrossRef]
2. Zhao, Y.; Wu, J.; Suo, Z.; Liu, X.; Liang, Y. Robust low-range-sidelobe target synthesis for airborne FDMA–MIMO STAP radar. *IET Radar Sonar Navig.* **2020**. [CrossRef]
3. Xu, J.; Zhu, S.; Liao, G. Space-Time-Range Adaptive Processing for Airborne Radar Systems. *IEEE Sens. J.* **2015**, *15*, 1602–1610.
4. Reed., I.S.; Mallett, J.D.; Brennan, L.E. Rapid Convergence Rate in Adaptive Arrays. *IEEE Trans. Aerosp. Electron. Syst.* **1974**, *AES-10*, 853–863. [CrossRef]
5. Kreyenkamp, O.; Klemm, R. Doppler compensation in forward-looking STAP radar. *IEE Proc. Radar Sonar Navig.* **2001**, *148*, 253–258. [CrossRef]
6. Himed, B.; Zhang, Y.H.; Hajjari, A. STAP with angle-Doppler compensation for bistatic airborne radars. In Proceedings of the IEEE Radar Conference, Long Beach, CA, USA, 25 April 2002.
7. Jaffer, A.G.; Ho, P.T.; Himed, B. Adaptive compensation for conformal array STAP by configuration parameter estimation. In Proceedings of the IEEE Conference on Radar, Verona, NY, USA, 24–27 April 2006.
8. Sun, G.; He, Z.; Tong, J.; Zhang, X. Knowledge-Aided Covariance Matrix Estimation via Kronecker Product Expansions for Airborne STAP. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 527–531. [CrossRef]
9. Riedl, M.; Potter, L.C. Knowledge-Aided Bayesian Space-Time Adaptive Processing. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1850–1861. [CrossRef]
10. Riedl, M.; Potter, L.C. Multi-Model Shrinkage for Knowledge-Aided Space-Time Adaptive Processing. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 2601–2610. [CrossRef]
11. Chen, H.; Liu, J.; Sun, H.; Yi, X.; Mu, H.; Lu, Y. Knowledge-aided Space Time Adaptive Processing for Airborne Radar in Heterogeneous Environments. In Proceedings of the 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 11–13 December 2019; pp. 1–5. [CrossRef]
12. Duan, K.; Wang, Z.; Xie, W.; Chen, H.; Wang, Y. Sparsity-based STAP algorithm with multiple measurement vectors via sparse Bayesian learning strategy for airborne radar. *Iet Signal Process.* **2017**, *11*, 544–553. [CrossRef]
13. Wu, Q.; Zhang, Y.D.; Amin, M.G.; Himed, B. Space–Time Adaptive Processing and Motion Parameter Estimation in Multistatic Passive Radar Using Sparse Bayesian Learning. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 944–957. [CrossRef]
14. Sun, K.; Meng, H.; Wang, Y.; Wang, X. Direct data domain STAP using sparse representation of clutter spectrum. *Signal Process.* **2011**, *91*, 2222–2236. [CrossRef]
15. Yang, X.; Sun, Y.; Zeng, T.; Long, T.; Sarkar, T. Fast STAP Method Based on PAST with Sparse Constraint for Airborne Phased Array Radar. *IEEE Trans. Signal Process.* **2016**, *64*, 4550–4561. [CrossRef]
16. Su, Y.; Wang, T.; Tao, F.; Li, Z. A Grid-Less Total Variation Minimization-Based Space-Time Adaptive Processing for Airborne Radar. *IEEE Access* **2020**, *8*, 29334–29343. [CrossRef]
17. Eikvil, L.; Aurdal, L.; Koren, H. Classification-based vehicle detection in high-resolution satellite images. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 65–72. [CrossRef]
18. Chen, Y.; Nasrabadi, N.M.; Tran, T.D. Simultaneous Joint Sparsity Model for Target Detection in Hyperspectral Imagery. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 676–680. [CrossRef]
19. Xiong, G.; Wang, F.; Yu, W.; Truong, T.K. Spatial Singularity-Exponent-Domain Multiresolution Imaging-Based SAR Ship Target Detection Method. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [CrossRef]
20. Khatib, A.E.; Assaleh, K.; Mir, H. Learning-based space–time adaptive processing. In Proceedings of the International Conference on Communications, Budapest, Hungary, 9–13 June 2013.
21. Khatib, A.E.; Assaleh, K.; Mir, H. Space-Time Adaptive Processing Using Pattern Classification. *IEEE Trans. Signal Process.* **2015**, *63*, 766–779. [CrossRef]
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
23. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
24. Jian, Y.; Ni, J.; Yang, Y. Deep Learning Hierarchical Representations for Image Steganalysis. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 2545–2557.
25. Dai, W.; Mao, Y.; Yuan, R.; Liu, Y.; Pu, X.; Li, C. A Novel Detector Based on Convolution Neural Networks for Multiscale SAR Ship Detection in Complex Background. *Sensors* **2020**, *20*, 2547. [CrossRef]
26. Scarpa, G.; Gargiulo, M.; Mazza, A.; Gaetano, R. A CNN-Based Fusion Method for Feature Extraction from Sentinel Data. *Remote Sens.* **2018**, *10*, 236. [CrossRef]
27. Wang, P.; Zhang, H.; Patel, V.M. SAR Image Despeckling Using a Convolutional Neural Network. *IEEE Signal Process. Lett.* **2017**, *24*, 1763–1767. [CrossRef]
28. Bentes, C.; Velotto, D.; Tings, B. Ship Classification in TerraSAR-X Images with Convolutional Neural Networks. *IEEE J. Ocean. Eng.* **2017**, *43*, 258–266. [CrossRef]
29. Wang, C.; Zheng, J.; Jiu, B.; Liu, H. Model-and-Data-Driven Method for Radar Highly Maneuvering Target Detection. *IEEE Trans. Aerosp. Electron. Syst.* **2021**, *57*, 2201–2217. [CrossRef]

30. Liu, Z.M.; Zhang, C.; Yu, P.S. Direction-of-Arrival Estimation Based on Deep Neural Networks with Robustness to Array Imperfections. *IEEE Trans. Antennas Propag.* **2018**, *66*, 7315–7327. [CrossRef]
31. Liu, Z.; Ho, D.K.C.; Xu, X.; Yang, J. Moving Target Indication Using Deep Convolutional Neural Network. *IEEE Access* **2018**, *6*, 65651–65660. [CrossRef]
32. Duan, K.; Chen, H.; Xie, W.; Wang, Y. Deep learning for high-resolution estimation of clutter angle-Doppler spectrum in STAP. *IET Radar Sonar Navig.* **2022**, *16*, 193–207. [CrossRef]
33. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [CrossRef]
34. He, K.; Jian, S. Convolutional neural networks at constrained time cost. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.