



## Article

# Direction-of-Arrival Estimation over Sea Surface from Radar Scattering Based on Convolutional Neural Network

Xiuyi Zhao <sup>1,2</sup>, Ying Yang <sup>3,\*</sup> and Kun-Shan Chen <sup>3</sup>

<sup>1</sup> State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China; zhaoxiuyi19@mailsucas.ac.cn

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup> College of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, China; chenks@radi.ac.cn

\* Correspondence: yangying@glut.edu.cn

**Abstract:** Conventional direction-of-arrival (DOA) estimation methods are primarily used in point source scenarios and based on array signal processing. However, due to the local scattering caused by sea surface, signals observed from radar antenna cannot be regarded as a point source but rather as a spatially dispersed source. Besides, with the advantages of flexibility and comparably low cost, synthetic aperture radar (SAR) is the present and future trend of space-based systems. This paper proposes a novel DOA estimation approach for SAR systems using the simulated radar measurement of the sea surface at different operating frequencies and wind speeds. This article's forward model is an advanced integral equation model (AIEM) to calculate the electromagnetic scattered from the sea surface. To solve the DOA estimation problem, we introduce a convolutional neural network (CNN) framework to estimate the transmitter's incident angle and incident azimuth angle. Results demonstrate that the CNN can achieve a good performance in DOA estimation at a wide range of frequencies and sea wind speeds.

**Keywords:** direction-of-arrival (DOA) estimation; convolutional neural network (CNN); sea surface scattering; radar remote sensing



**Citation:** Zhao, X.; Yang, Y.; Chen, K.-S. Direction-of-Arrival Estimation over Sea Surface from Radar Scattering Based on Convolutional Neural Network. *Remote Sens.* **2021**, *13*, 2681. <https://doi.org/10.3390/rs13142681>

Academic Editor: Yukiharu Hisaki

Received: 15 May 2021

Accepted: 5 July 2021

Published: 7 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Determining the direction of arrival (DOA) of the radar signal is a fundamental problem for sea surveillance. The task of DOA estimation is to identify the signal source directions in which the signal is transmitted. Conventional DOA estimation methods, including beamforming techniques [1–3] and subspace-based methods [4–8], are primarily used in point source scenarios. With the development of machine learning and artificial intelligence, neural network (NN) has been applied in the DOA estimation domain [9–14]. This method establishes training datasets with DOA labels first, and then derives a mapping from antenna outputs to signal directions with existing methods. The derived mapping is then used on test datasets to estimate signal directions. An NN-based DOA estimation is data driven and does not rely on pre-assumptions about array geometries and whether the antenna outputs are calibrated or not. They have been demonstrated to be computationally more efficient than subspace-based methods in simulations [14].

However, sea-surface-induced local scattering causes signals from a single source to arrive via different paths and at different angles. Therefore, the signals observed from radar antenna can then be regarded as a superposition of all contributions from different propagating paths. In this case, the source is no longer perceived as a point source but instead as a spatially dispersed source with mean DOA and spatial extent. Previous works have shown that the performance of point source DOA estimation methods is degraded [15–17]. Most approaches for DOA estimation of scattered sources utilize a model of the data covariance matrix, for example, maximum likelihood estimators [18–22].

The price to pay for this increased efficiency and accuracy is that those algorithms require a multidimensional search to find the estimates and converge to a local minimum. Another commonly used approach, the spatial smoothing algorithm [23,24], obtains the robustness at the expense of a reduced effective aperture.

Additionally, methods as mentioned above require many antenna elements to achieve a high degree of accuracy. As the number of antenna elements increases, the power consumption, size, and the cost of the system increase. In [25], artificial neural networks are used to decrease the size and weight and reduce the costs of components production and whole application systems in forming a beam from radiating complex microstrip antenna. Previous studies on DOA estimation of SAR system were realized by filtering the rough-surface induced clutter and then retrieving the clutter-surpassed target signals [26,27]. In this paper, we attempt to develop a new DOA estimation technique with low computational complexity and high estimation accuracy from a radar observation.

This paper proposes a two-dimensional DOA estimation method for a radar system under a special case where the incident signal is fully scattered from the sea surface. At this point, the received signal contains sea clutter. We utilize a convolutional neural network (CNN) to establish the relationship of radar measurements with their incident directions. The advanced integral equation model (AIEM) is applied as the forward model to simulate the sea surface electromagnetic scattering under different operating frequencies, observation geometry, and winds. The input of CNN are model-generated radar measurements added with speckle noise and the output is the related incidence location of the input scattering pattern.

The contributions of this article are briefed as follows:

We explore the DOA estimation problem under a special case where the incident signal is randomly scattered by sea surface. The effect of radar frequencies and wind speeds is taken into consideration as well.

We propose a CNN framework that makes use of underlying statistical characteristics of speckle noise-contaminated radar measurements to predict their related incidence location. By experiment, the CNN shows good performance in the DOA estimation problem, and the performance of CNN is not affected by the radar frequencies and wind speeds.

This paper is organized as follows: Section 2 gives the relevant background on the random rough surface scattering models, especially the AIEM model, and illustrates the sensitivity of the received signal to influential parameters. Section 3 formulates the DOA estimation by sea surface scattering measurement, and presents an algorithm to simulate the sea surface measurement dataset and basic notations used for illustration. Besides, we present the CNN structure and interpret how it fits our requirements in Section 3. In Section 4, implementation details and experimental results are presented. Section 5 discusses the DOA estimation results with state-of-the-art algorithms and limitations in the current work. Finally, the conclusions are summarized to close the paper.

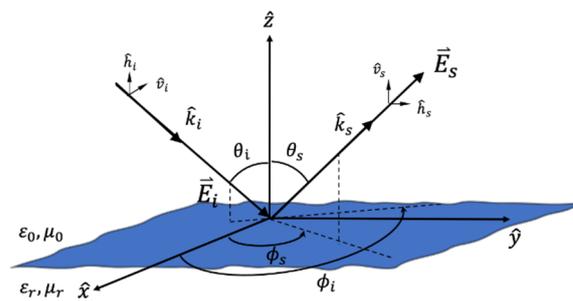
## 2. Sea Surface Scattering Datasets Generated by AIEM

We first give the AIEM bistatic scattering model, followed by a sensitivity analysis of radar measurement to incidence location and receiver location under certain sea states.

### 2.1. Bistatic Scattering Model

Figure 1 depicts a geometry of bistatic radar scattering from the sea surface. Sea surface is assumed to be a randomly rough surface with known statistical properties;  $\theta_i$  and  $\phi_i$  are the incident angle and incident azimuth angle, respectively;  $\theta_s$  and  $\phi_s$  are the scattering angle and scattering azimuth angle of the receiver, respectively;  $\epsilon_0$  and  $\mu_0$  are the permittivity and permeability of the half free space;  $\epsilon_r$  and  $\mu_r$  are relative permittivity and of the sea;  $\hat{k}_i$  is the incident wave vector and  $\hat{k}_s$  is the scattered wave vector.  $\hat{k}_i$  and  $\hat{k}_s$  are defined as:

$$\begin{aligned} k_x &= k \sin \theta_i \cos \phi_i, k_{sx} = k \sin \theta_s \cos \phi_s, \\ k_y &= k \sin \theta_i \sin \phi_i, k_{sy} = k \sin \theta_s \sin \phi_s, \\ k_z &= -k \cos \theta_i, k_{sz} = k \cos \theta_s \end{aligned} \quad (1)$$



**Figure 1.** Geometry of the bistatic scattering from a sea surface.

For the purpose of this paper, we utilize the AIEM as a forward model to simulate the scattering process of the sea surface under different observation configurations and sea states. The AIEM model is an analytical model based on the integral equation method (IEM) [28–30]. Both the IEM and AIEM models have been used on sea surface microwave scattering and are in excellent agreement with radar measurements [31–34].

In original AIEM model [29],  $\sigma_{pq}^0$  can be written as:

$$\sigma_{pq}^0 = \frac{k^2}{2} \exp[-\sigma^2(k_{iz}^2 + k_{sz}^2)] \times \sum_{n=1}^{\infty} \frac{\sigma^{2n}}{n!} |I_{pq}^n| W^{(n)}(k_{sx} - k_{ix}, k_{sy} - k_{iy}) \quad (2)$$

where the polarization indices  $p$  and  $q$  represent the  $p$ -polarized ( $v$  or  $h$ ) transmitted power and the  $q$ -polarized ( $v$  or  $h$ ) received power;  $k$  is the wave number of incident wave in free space;  $k_{ix}$ ,  $k_{iy}$ ,  $k_{iz}$ ,  $k_{sx}$ ,  $k_{sy}$ , and  $k_{sz}$  are the coordinate projection of  $\hat{k}_i$  and  $\hat{k}_s$ ;  $\sigma$  is the root-mean-squared (RMS) height of rough surface; the explicit expression of  $I_{pq}^n$  is given in [30];  $W^{(n)}$  is the Fourier transform of the  $n$ -th power of the surface correlation function  $\rho_s$ . For sea surface scattering,  $W^{(n)}$  is relevant to the relative azimuth angle  $\chi$  between the direction of sea wind  $\phi_w$  and  $\phi_i$  [35]:

$$W^{(n)}(k, \chi) = \int_0^{2\pi} \int_0^{\infty} \rho_s^n(r, \phi) e^{-jkr \cos(\chi - \phi)} r dr d\phi \quad (3)$$

An approximate representation of the sea surface correlation function is given by [36,37]:

$$\rho_s(r, \phi) = s^2 e^{-\frac{r}{L_t}} \quad (4)$$

where  $L_t = L_u \cos^2 \phi_w + L_c \sin^2 \phi_w$  is the correlation length along  $\phi_w$ .  $L_u$  is the correlation length in the upwind direction (along  $\phi_w = 0^\circ$ );  $L_c$  is the correlation length in the crosswind direction (along  $\phi_w = 90^\circ$ ). An approximate form for  $W^{(n)}(k, \chi)$  is given by [37]:

$$W^{(n)}(k, \chi) = \left(\frac{sL_t}{n}\right)^2 \left[1 + \left(\frac{kL_t}{n}\right)^2\right]^{-\frac{3}{2}} \quad (5)$$

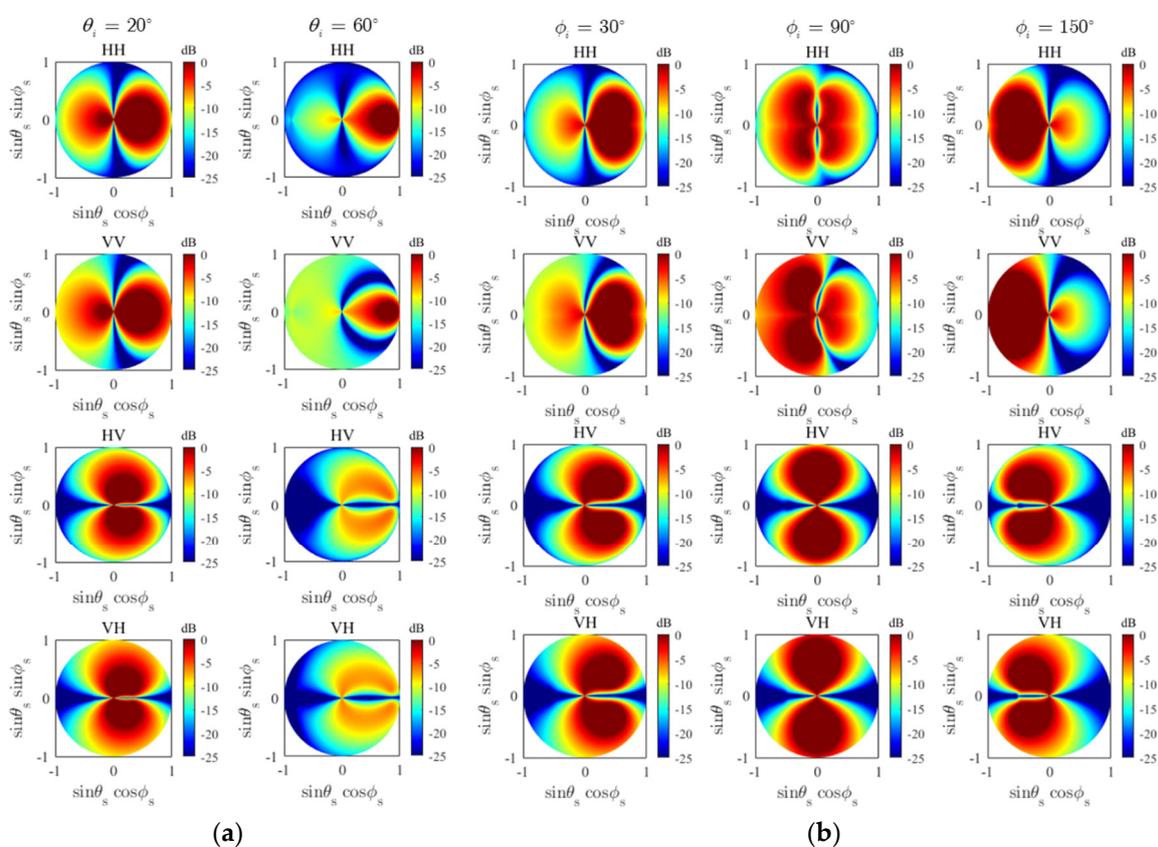
It has been successfully used in sea surface scattering computations.

## 2.2. Sensitivity of Received Signals to Incidence Location

Before devising an effective approach to DOA estimation, it is essential to conduct a sensitivity analysis of scattering responses to incidence location under specific surface parameters. According to Equations (1), (2), and (5), besides the incidence and observation location, the three rough surface parameters, including RMS height, correlation length, and dielectric constant, are related to bistatic scattering coefficients. For sea surface scattering, these parameters are influenced by sea wind speed and radar operating frequency. The values of  $L_u$ ,  $L_c$ , and  $\sigma$  under different wind speeds and frequencies can be found in [37]. They are empirical values derived by fitting the rough surface scattering model with

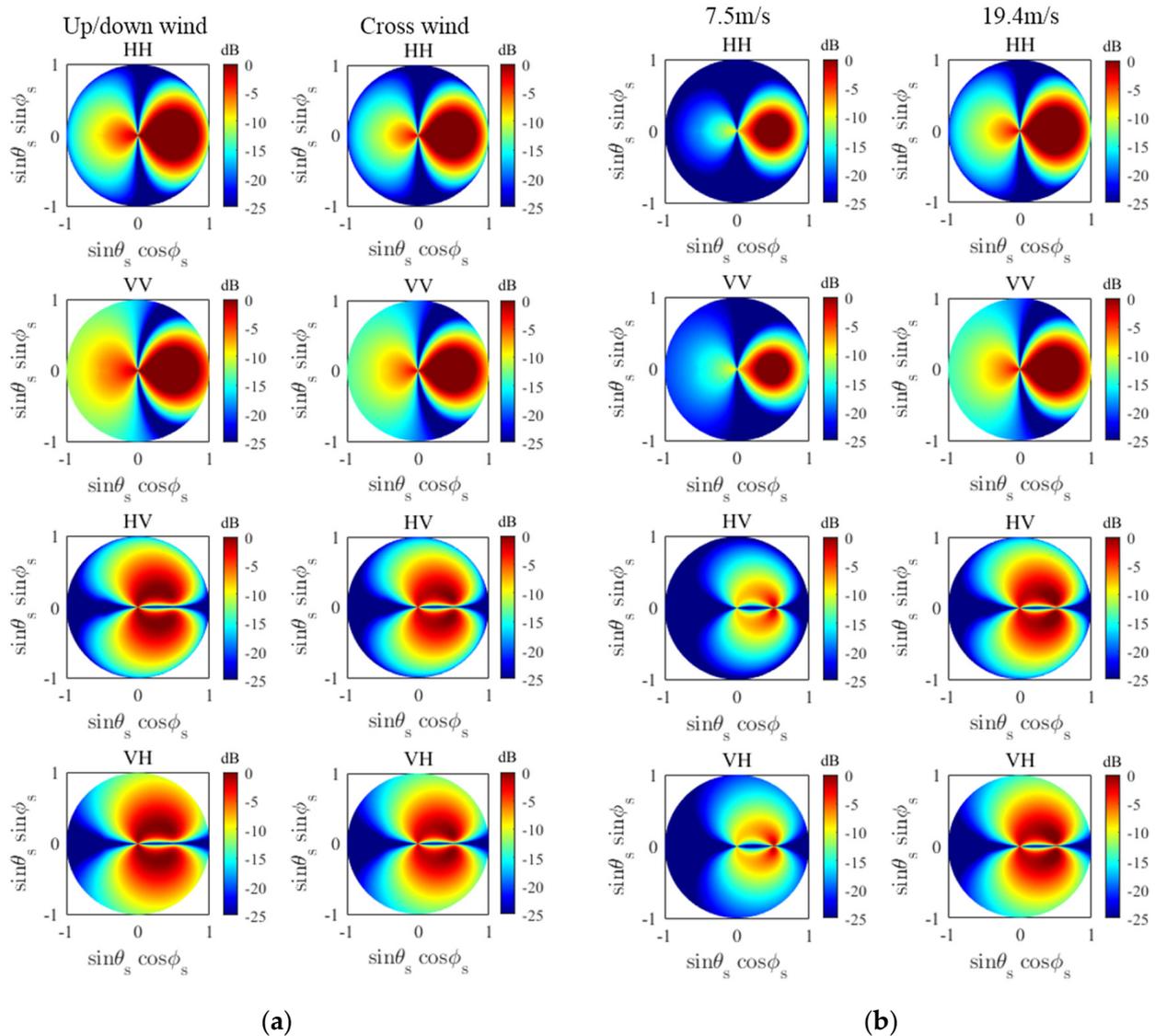
measured data. In what follows, we illustrate the bistatic hemispherical plots for the dependences of incidence location and wind direction and wind speed.

In Figure 2a,b, we examine the angular dependence of polarized bistatic scattering coefficients. Figure 2a shows the effect of the incident angle with  $f = 5.3$  GHz,  $U = 25$  m/s,  $\phi_i = 0^\circ$ ,  $\theta_i = 20^\circ$  (the left column), and  $\theta_i = 60^\circ$  (the right column). Similarly, Figure 2b shows the effect of the incident azimuth angle with  $f = 5.3$  GHz,  $U = 25$  m/s,  $\theta_i = 40^\circ$ ,  $\phi_i = 30^\circ$  (the left column),  $\phi_i = 90^\circ$  (the middle column), and  $\phi_i = 120^\circ$  (the right column). For the same set of sea surface parameters, both the incident angle and incident azimuth angle have a strong influence on the scattering patterns and strength for all polarizations. In Figure 2a, the left-hand side of the half-sphere and right-hand side of the half-sphere represent backward and forward scattering regions. In contrast, the horizontal axis and vertical axis represent the incident plane and cross-incident plane. In the forward scattering region of Figure 2a, the strong scattering is due to the specular scattering, and it moves to the right hemisphere with the increase of the incident angle. For a smaller incident angle, the small azimuth angular region shows a stronger scattering strength at HH polarization. As the incident angle increases, the region with strong scattering starts to narrow down. In the backward region, the intensity of HH polarization reduces as the incident angle increases, but a subtle increment appears at a large scattering angle and scattering azimuth angle. Similar trends can be observed in VV polarization. Meanwhile, the strength of cross-polarization, HV and VH, quickly weakened on the whole upper hemisphere with an increasing incident angle. As for Figure 2b, the forward and backward scattering regions change with the incident azimuth angle. Relatively speaking, for all polarizations, the strong scattering area moves from the right-hand hemisphere to the left-hand hemisphere with the increase of the incident azimuth angle.



**Figure 2.** Hemispherical plots of bistatic scattering coefficients  $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0$  with  $f = 5.3$  GHz,  $U = 25$  m/s,  $\phi_w = 30^\circ$ : (a)  $\phi_i = 0^\circ$  for all subplots;  $\theta_i = 20^\circ$  for subplots in the left column and  $\theta_i = 60^\circ$  for subplots in the right column; (b)  $\theta_i = 40^\circ$  for all subplots;  $\phi_i$  for subplots in the left, middle, and right column are  $30^\circ, 90^\circ, \text{ and } 150^\circ$ , respectively.

Figure 3a shows the hemispherical plot of co-and cross-polarized bistatic scattering coefficients with  $f = 5.3$  GHz,  $U = 25$  m/s,  $\theta_i = 30^\circ$ , and  $\phi_i = 0^\circ$ , with relative wind direction varying from up/downwind ( $0^\circ/180^\circ$ ) to crosswind ( $90^\circ$ ). For co-polarization, we see that the strong scattering area in the backward region at up/downwind is more pronounced than at crosswind. In addition, the cross-polarization is impacted by the relative wind direction to a lesser extent than co-polarization is.



**Figure 3.** Hemispherical plots of bistatic scattering coefficients  $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0$  with  $\theta_i = 30^\circ, \phi_i = 0^\circ$ : (a) at 5.3 GHz operating frequency,  $U = 25$  m/s; wind direction for subplots in the left and right columns is up/downwind ( $0^\circ/180^\circ$ ) and crosswind ( $90^\circ$ ), respectively (b) at 13.9 GHz,  $\phi_w = 30^\circ$  and  $U = 7.5$  m/s for left-row subplots.

Finally, we examine the sea surface roughness dependence in Figure 3b, which is the hemispherical plot of co-and cross-polarized bistatic scattering coefficients with  $f = 13.9$  GHz,  $\theta_i = 30^\circ, \phi_i = 0^\circ$ , and  $\phi_w = 30^\circ$ , but sea wind speed varying from 7.5 m/s to 19.4 m/s. With the increase of sea wind speeds, sea surface roughness is increased as well. However, the change of scattering coefficients is coupled with three roughness parameters. In general, the scattering strength under higher wind speed in both forward and backward regions is greater than the scattering strength under lower wind speed. Especially for cross-polarization under lower wind speed, there is a poorer feature in the backward scattering region than that in the forward region.

### 3. Direction-of-Arrival Estimation

#### 3.1. Problem Formulation

To better estimate the angles of the incident source, a proper data model is essential to enable a mapping of the measurement domain to the feature domain that is of interest. We detail how we came up with a good scheme to improve DOA estimation accuracy and efficiency. The radar scattering from a rough surface can be modeled as [30]:

$$\mathbf{b} = \mathbf{W}\mathbf{x} + \mathbf{u} \quad (6)$$

where  $\mathbf{x}$  is a vector including the surface parameters and radar parameters; matrix  $\mathbf{W}$  relates the parameters vector  $\mathbf{x}$  to radar scattering coefficients  $\mathbf{b}$ ; and  $\mathbf{u}$  represents the error vector induced by system and calibration errors, and speckle noise, among other factors. In this paper, the effect of speckle noise is considered only.

In statistical sense,  $\mathbf{x}$  constitutes a random variable due to spatially and temporally varying properties, such that:

$$\mathbf{x} = \mathbf{x}_t + \mathbf{x}_n \quad (7)$$

where  $\mathbf{x}_t$  is the true value and  $\mathbf{x}_n$  is the noise term. In practice, the “truth” is never obtainable and always vague. Statistically,  $\mathbf{x}_t$  and  $\mathbf{x}_n$  may be assumed to be uncorrelated, such that  $\mathbf{x}$  is an unbiased estimation of  $\mathbf{x}_t$ , i.e.,

$$E(\mathbf{x}) = E(\mathbf{x}_t), \forall \mathbf{x}_n \sim N(0, \sigma_{\mathbf{x}_n}^2) \quad (8)$$

where  $E$  denotes the statistical mean, and  $\sigma_{\mathbf{x}_n}^2$  is a variance of  $\mathbf{x}_n$ .

In general, the radar response is formed by the scattering matrix. For the purpose of this paper, we assume the radar measurement vector  $\mathbf{b}$  is formed by multi-polarized scattering coefficients and the location of the scattering source:

$$\mathbf{b} = [\sigma_{hh}^0, \sigma_{vv}^0, \sigma_{hv}^0, \sigma_{vh}^0, \theta_s, \phi_s]^t \quad (9)$$

The location of the incident source is determined by the incident angle  $\theta_i$  and incident azimuth angle  $\phi_i$ . They are denoted by:

$$\mathbf{x} = [\theta_i, \phi_i]^t \quad (10)$$

Thus, the cost function of DOA estimation problem can be written as:

$$\hat{\mathbf{x}} = \operatorname{argmin} \left\{ \left\| \mathbf{b} - \mathbf{W}\mathbf{x}_{est} \right\|_2^2 \right\} \quad (11)$$

where  $\|\cdot\|_2$  denotes  $l_2$ -norm and  $\mathbf{x}_{est}$  is the estimated incident angles.

Based on what we presented in last section, the scattering behavior is completely intricate. Therefore, it hard to get an analytical but practical solution of Equation (6). In this paper, we adapt a convolutional neural network approach for searching the cost function minima in Equation (6).

#### 3.2. Data Input-Output and Preprocessing

In this part, we introduce the method to simulate the sea surface scattering by the AIEM model. Two sets of radar operating frequencies are considered in this paper. The frequencies in set one are 5.3 GHz (C-band) and 13.99 GHz (Ku-band). For 5.3 GHz, we simulate the electromagnetic wave scattered from the sea surface under the sea wind speeds from 25 m/s to 55 m/s, and for 13.99 GHz, the sea wind speeds are from 5 m/s to 30 m/s. The frequencies in another set are 9.4 GHz (X-band), 13.9 GHz (Ku-band), and 14.6 GHz (Ku-band). The sea wind speeds of this set are from 5.5 m/s to 19.4 m/s. The values of  $L_u$ ,  $L_c$ , and  $\sigma$  under different wind speeds and frequencies are listed in [36]. The

ranges of the incidence and received angle, operating frequencies, and sea wind speeds in the DOA estimation experiment are listed in Table 1.

**Table 1.** The ranges of the incidence and received angle, operating frequencies, and sea wind speeds in the DOA estimation experiment.

Parameters	Description	Range	Step Size
$\theta_i$	Incident angle	20°–70°	5°
$\phi_i$	Incident azimuth angle	10°–170°	5°
$\theta_s$	Scattering angle	10°–70°	5°
$\phi_s$	Scattering azimuth angle	10°–170°	5°
$\phi_w$	Wind direction	0°–180°	5°
$f$	Frequency	5.3 GHz/9.4 GHz/ 13.9 GHz/13.99 GHz/ 14.6 GHz	
$U$	Wind speed	25 m/s–55 m/s (5.3 GHz) 5 m/s–30 m/s (13.99 GHz) 5.5 m/s–19.4 m/s (9.4 GHz, 13.9 GHz,14.6 Hz)	10 m/s (5.3 GHz) 5 m/s (13.99 GHz)

In regard to the DOA estimation problem, as illustrated in the previous section, the input of CNN is bistatic scattering coefficients ( $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0$ ) and related scattering angles ( $\theta_s$  and  $\phi_s$ ). The output of CNN is the corresponding incident source ( $\theta_i$  and  $\phi_i$ ). Under a given frequency and sea speed wind, the Algorithm 1 to generate datasets of sea surface scattering patterns is a quintuple iteration, i.e.,

---

**Algorithm 1** Generation of Training Data by the AIEM Model

---

```

For  $\theta_i = 20^\circ$  to  $70^\circ$  step  $5^\circ$ 
  For  $\phi_i = 10^\circ$  to  $170^\circ$  step  $5^\circ$ 
    For  $\theta_s = 10^\circ$  to  $70^\circ$  step  $5^\circ$ 
      \Define a matrix input
      input = []
      \Define a vector output
      output = []
      For  $\phi_s = 10^\circ$  to  $170^\circ$  step  $5^\circ$ 
        For  $\phi_w = 0^\circ$  to  $180^\circ$  step  $5^\circ$ 
          [ $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0$ ] = AIEM ( $\theta_i, \phi_i, \theta_s, \phi_s, \phi_w$ )
          \  $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0$  are generated by function AIEM
          input = [input;  $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0, \theta_s, \phi_s$ ]
        End
      End
      Save generated input
      output = [ $\theta_i, \phi_i$ ]
      Save generated output
    End
  End
End
End
End

```

---

The iterations for  $\phi_w$  and  $\phi_s$  generate one single input sample. Hence, the length of a single input is the product of  $N_{\phi_w}$  (the number of  $\phi_w$ ) and  $N_{\phi_s}$  (the number of  $\phi_s$ ). In this paper, this value is 1221. The width of inputs is six due to six attributes:  $\sigma_{hh}^0, \sigma_{hv}^0, \sigma_{vh}^0, \sigma_{vv}^0, \theta_s$ , and  $\phi_s$ . The iterations for  $\theta_i, \phi_i$ , and  $\theta_s$  decide the number of inputs, which is the product of  $N_{\theta_i}, N_{\phi_i}$ , and  $N_{\theta_s}$ . This value is equal to 4719 in this study.

However, as illustrated above, in the actual radar image, it always exhibits large pixel-to-pixel intensity variations, referred to as speckle. To better characterize realistic scattering coefficients in SAR images, we assume that the actual radar measured data follows the noise model as:

$$\sigma_{pq}^n = \sigma_{pq}^0 + 10 \log_{10} r \quad (12)$$

where  $\sigma_{pq}^n$  and  $\sigma_{pq}^0$  represent the measured and noise-free radar signals (in decibels), respectively;  $r$  is a random number of the K-distribution, which models the speckle noise in radar measurements. In this paper, we use a kind of  $\nu$ -model to determine the values of the K-distribution shape parameter under different wind speeds. The shape parameter  $\nu$  can be expressed as [38]:

$$\nu = \nu_0 e^{-U/U_0} \quad (13)$$

where  $\nu_0$  and  $U_0$  are parameters that are dependent on the radar incident angle and wind direction. The scale parameter of the K-distribution is set to be 1. By adding speckle noise to model-generated bistatic scattering coefficients, the parameters of CNN input become  $\sigma_{hh}^n, \sigma_{vv}^n, \sigma_{hv}^n, \sigma_{vh}^n, \theta_s$ , and  $\phi_s$ . At this point, we have finished the whole process of generating sea surface scattering datasets.

### 3.3. CNN Configurations for DOA Estimation

We aim to find a proper CNN structure trained to estimate from sea surface scattering datasets under different wind speeds and radar operating frequencies. A CNN topology typically consists of multiple convolutional layers followed by the fully connected layers. In CNN architectures, the convolutional layers are pairs of convolution and pooling operations. Especially for DOA estimation, we should perform multi-output regression. We share the same convolutional layer structure to predict the incident angle and incident azimuth angle using one fully connected layer at the final stage.

The main focus is to achieve a trade-off between the training speed and accuracy of the CNN. Therefore, as outlined in Table 2, we design five different CNN models inspired by VGG-16 [39] to be evaluated. In Table 2, the convolutional layer is abbreviated as “Conv” and the layer parameters are denoted as “length  $\times$  width  $\times$  number of filters, stride”. The Average-pooling layer is abbreviated as “Avg- $p$ ” and pooling size is denoted as “length  $\times$  width, stride.” “FC” and “Regrs” represent the fully connected layer and regression layer, respectively. In Model 1, the CNN comprises eight convolutional layers, three average pooling layers, and a fully connected layer. The pooling layer follows the fourth, sixth, and eighth convolutional layers. A stride of two is fixed for the average-pooling layer of this CNN framework. When an input-output pair feeds into this CNN, two transverse 1-D convolutional layers follow to extract the features crosswise of the input sample. Then, three or four groups of vertical 1-D convolutional layers and vertical average pooling layers aim to utilize the lengthwise information, which is concerned with the location of the transmitter and wind direction. Therefore, the scattering pattern and strength under the different directions of sea wind and observation geometry can be exploited by this model at large.

Previous studies have shown that shallow networks require exponentially more neurons than deep networks to achieve accuracy for function approximation [40]. Hence, we enlarge the length of the convolutional layer rather than adopting a typical short-length kernel as the VGG-16 model (i.e., in the size of  $3 \times 3$ ). The kernel size of both convolutional and pooling layers is sequentially increased from  $5 \times 1$  to  $10 \times 1$ . At the end of Model 1, a fully connected layer with a regression layer with 1024 nodes is applied to generate the final outputs for  $\theta_i$  and  $\phi_i$  prediction.

The depth of the configurations increases from the left (Model 1) to the right (Model 2) as more layers are added in front of the fully connected layer. The amount of filters in each convolutional layer is decreased from Model 2 to Model 3. Besides, we add models “FC-512” and “FC-2048” to compare the performance between models with different numbers of hidden units in a fully connected layer; “FC-512” denotes the number of hidden units being 512, and “FC-2048” being 2048 hidden units. The various numbers of hidden units of each model in Table 2 are shown in bold.

**Table 2.** CNN configurations (shown in columns). The depth of the configurations increases from the left (Model 1) to the right (Model 2), as more layers are added. Besides, the number of filters decreases from Model 2 to Model 3. Changes are shown in bold. The aim of model “FC-512” and “FC-2048” is to compare the performance between models with different numbers of hidden units in the fully connected layer. The convolutional layer is abbreviated as “Conv” and layer parameters are denoted as “length × width × number of filters, stride”. The average-pooling layer is abbreviated as “Avg-p” and layer size is denoted as “length × width, stride”. “FC” and “Regrs” represent a fully connected layer and a regression layer, respectively.

Model 1		Model 2		Model 3		FC-512		FC-2048	
Layer Type	Size	Layer Type	Size	Layer Type	Size	Layer Type	Size	Layer Type	Size
Input	1221 × 6 × 1	Input	1221 × 6 × 1	Input	1221 × 6 × 1	Input	1221 × 6 × 1	Input	1221 × 6 × 1
Conv	1 × 2 × 32, 2	Conv	1 × 2 × 32, 2	Conv	1 × 2 × 16, 2	Conv	1 × 2 × 32, 2	Conv	1 × 2 × 32, 2
Conv	1 × 3 × 32	Conv	1 × 3 × 32	Conv	1 × 3 × <b>16</b>	Conv	1 × 3 × 32	Conv	1 × 3 × 32
Conv	5 × 1 × 64	Conv	5 × 1 × 64	Conv	5 × 1 × <b>32</b>	Conv	5 × 1 × 64	Conv	5 × 1 × 64
Conv	5 × 1 × 64	Conv	5 × 1 × 64	Conv	5 × 1 × <b>32</b>	Conv	5 × 1 × 64	Conv	5 × 1 × 64
Avg-p	6 × 1, 2	Avg-p	6 × 1, 2	Avg-p	6 × 1, 2	Avg-p	6 × 1, 2	Avg-p	6 × 1, 2
Conv	7 × 1 × 128	Conv	7 × 1 × 128	Conv	7 × 1 × <b>64</b>	Conv	7 × 1 × 128	Conv	7 × 1 × 128
Conv	7 × 1 × 128	Conv	7 × 1 × 128	Conv	7 × 1 × <b>64</b>	Conv	7 × 1 × 128	Conv	7 × 1 × 128
Avg-p	8 × 1, 2	Avg-p	8 × 1, 2	Avg-p	8 × 1, 2	Avg-p	8 × 1, 2	Avg-p	8 × 1, 2
Conv	9 × 1 × 256	Conv	9 × 1 × 256	Conv	9 × 1 × <b>128</b>	Conv	9 × 1 × 256	Conv	9 × 1 × 256
Conv	9 × 1 × 256	Conv	9 × 1 × 256	Conv	9 × 1 × <b>128</b>	Conv	9 × 1 × 256	Conv	9 × 1 × 256
Avg-p	10 × 1, 2	Avg-p	10 × 1, 2	Avg-p	10 × 1, 2	Avg-p	10 × 1, 2	Avg-p	10 × 1, 2
FC	1024	Conv	<b>11 × 1 × 512</b>	Conv	11 × 1 × <b>256</b>	Conv	11 × 1 × 512	Conv	11 × 1 × 512
Regrs	2	Conv	<b>11 × 1 × 512</b>	Conv	11 × 1 × <b>256</b>	Conv	11 × 1 × 512	Conv	11 × 1 × 512
		Avg-p	<b>12 × 1, 2</b>	Avg-p	12 × 1, 2	Avg-p	12 × 1, 2	Avg-p	12 × 1, 2
		FC	1024	FC	1024	FC	<b>512</b>	FC	<b>2048</b>
		Regrs	2	Regrs	2	Regrs	2	Regrs	2

## 4. Results

### 4.1. Comparison of CNN Configurations

In the previous section, we presented the details of five CNN configurations to be evaluated. We aimed to find a CNN structure with the best performance to estimate the direction of the incident source from simulated radar measurement data. We implemented the CNN models as mentioned earlier with MATLAB® Deep Learning Toolbox. We used a Dell T7810 Series desktop with two Inter Xeon processors and an NVIDIA Quadro K620 GPU for the training and test. The dataset used in the comparison experiment was the simulated sea surface scattering dataset at 13.9 GHz, 5.5 m/s wind speed.

We trained and tested the proposed CNN models with each dataset 100 times to get a robust and stable result. The dataset was randomly divided into three parts: training, validation, and testing set with a 0.7:0.15:0.15 ratio in every realization. The dropout strategy was used in the neural network, and the dropout rate was 0.05. The learning rate was set to be 0.01 initially and decay 0.1 for every 10 epochs. Besides, we applied the batch normalization technique for the training set to reduce time consumption. The size of the mini-batch was set to be 50. The validation of the current CNN occurs at the end of each mini-training batch. The early stopping strategy was also taken into consideration. The loss on the validation set can be more significant or equal to the previous loss in five epochs before network training stops.

#### 4.1.1. Training

In deep learning, the training algorithm is based on backpropagation using gradient descent algorithms. Therefore, a loss function must be defined that computes the difference between the network output and the truth value. In this work, the half-mean-squared error loss was used as the loss function  $L$  to optimize the loss between the true values of incident angles and the model predictions.  $L$  is expressed as follows:

$$L(y, \hat{y}) = \sum_{i=1}^N \frac{(y_i^1 - \hat{y}_i^1)^2 + (y_i^2 - \hat{y}_i^2)^2}{2N} \quad (14)$$

where  $y^1$  and  $y^2$  denote the truth value of  $\theta_i$  and  $\phi_i$ ;  $\hat{y}^1$  and  $\hat{y}^2$  are the predicted values of  $\theta_i$  and  $\phi_i$ ;  $N$  represents the size of the mini batch.

To minimize the loss function, we used the stochastic gradient descent with momentum (SGDM) as an optimizer of CNN. The standard stochastic gradient descent algorithm can oscillate along the path of steepest descent towards the optimum. Adding a momentum term to the network parameter (weights in the convolutional kernel) update is one way to reduce this oscillation. The SGDM update is:

$$w_{l+1} = w_l - \alpha \nabla E(w_l) + \gamma(w_l - w_{l-1}) \quad (15)$$

where  $w$  represents the network parameter vector,  $l$  is the iteration number,  $\alpha$  is a learning rate,  $E(w)$  is the loss function, and  $\gamma$  determines the contribution of the previous gradient step to the current iteration. In this experiment, we set  $\gamma$  to 0.9.

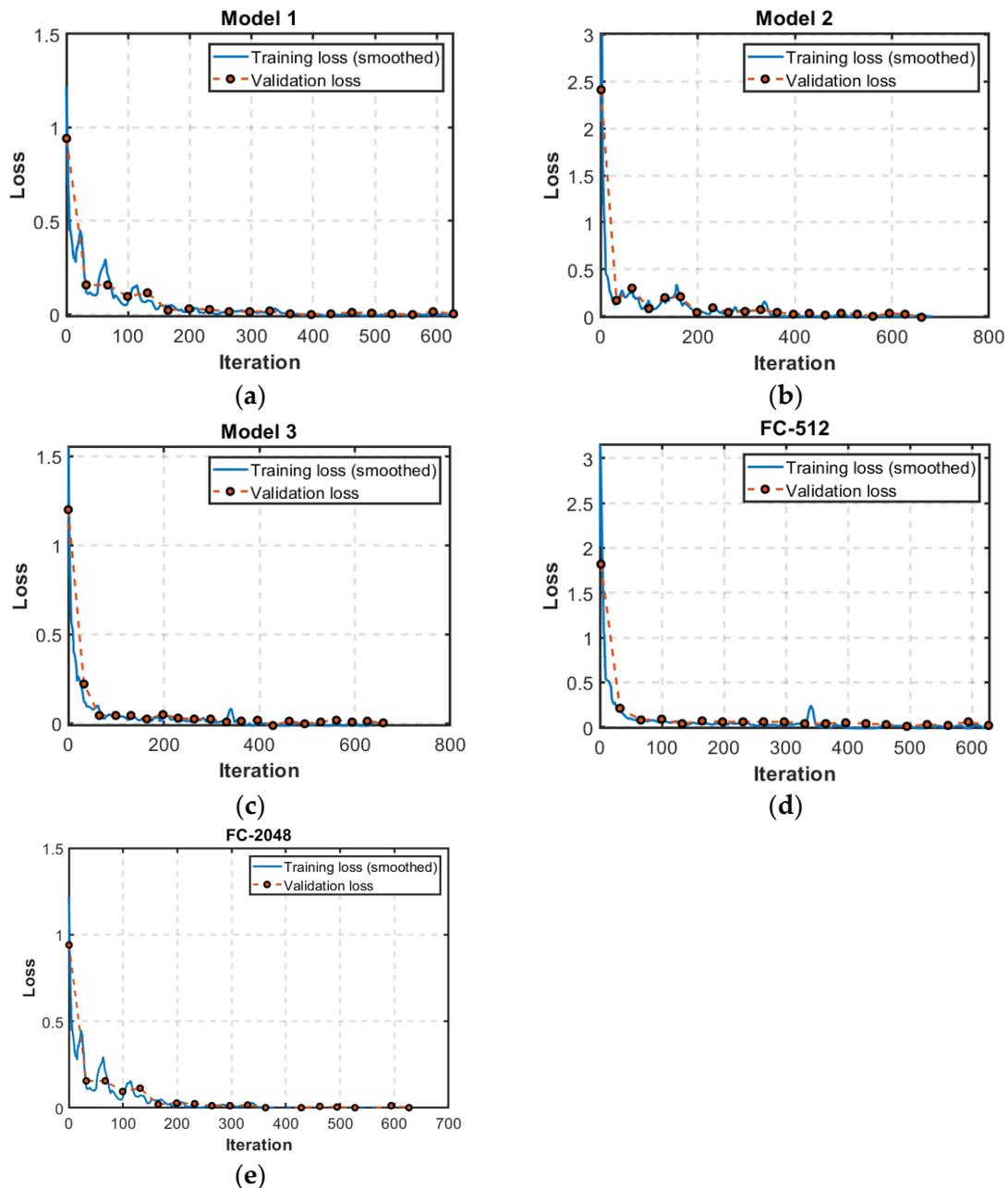
Figure 4 illustrates the loss function minimization process of CNN configurations to be evaluated. The solid line represents the smoothed training loss, and the red dots represent the loss of the validation set. The training process will be stopped if the validation loss satisfies the early stopping condition or the number of epochs reaches the maximum. The model will achieve better performance if the loss function finds a good local minimum or a global minimum. From Figure 4, we can see that the training loss functions of all models are convergent to minimal loss. Similar results were obtained for every realization but are not included here for brevity.

#### 4.1.2. Testing

After the training process, we tested CNNs with corresponding test sets. In this paper, the root-mean-square error was chosen to measure the error of CNN in DOA estimation. The results are shown in Table 3. Note that the CNN configuration with fewer stacked convolutional layers (Model 1) and fewer filters in the convolutional layer (Model 3) will lead to less time consumption but performs worse compared with Model 2. By comparing model FC-512, Model 2, and FC-2048, we found that the number of hidden units in a fully-connected layer has an effect on CNN performance. According to our experiment, more hidden units in a fully connected layer improves the prediction accuracy, but the computation time increases. In comparison, Model 2 can balance both estimation accuracy and computation time in the DOA estimation task. Therefore, in the following section, the CNN architecture of Model 2 was employed.

**Table 3.** Performance of 100 realizations of the dataset at 13.9 GHz, 5.5 m/s using the CNN methods illustrated in Table 2. The numbers reported in this table are the average of all realizations.

	Model 1	Model 2	Model 3	FC-512	FC-2048
<b>Time Consuming (sec)</b>	315	425	175	389	621
<b>RMSE of Incident Azimuth Angle</b>	3.89°	3.31°	4.39°	4.55°	3.13°
<b>RMSE of Incident Angle</b>	1.04°	1.05°	1.14°	1.11°	1.01°



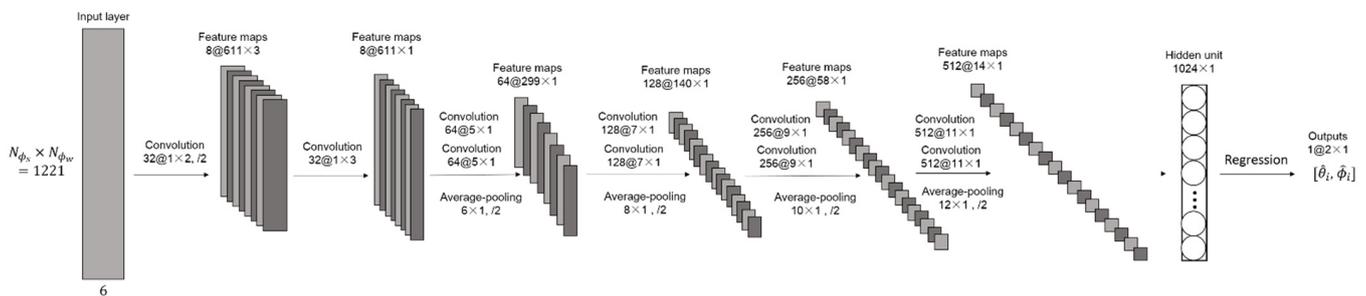
**Figure 4.** Loss function minimization process of five different CNN configurations: (a) Model 1; (b) Model 2; (c) Model 3; (d) FC-512; (e) FC-2048. The experiment is based on the simulated sea surface scattering dataset at 13.9 GHz, 5.5 m/s.

## 4.2. DOA Estimation Based on CNN

### 4.2.1. Model Training

After deciding on a certain CNN model, we utilized the simulated sea surface scattering datasets illustrated in the previous section to estimate DOA at different frequencies and wind speeds. Figure 5 shows the CNN configuration used for DOA estimation. Feature maps in Figure 5 represent the output of each average-pooling layer. The operating frequencies of the datasets are 5.3 GHz (at a wind speed of 25 m/s, 35 m/s, 45 m/s, and 55 m/s), 9.4 GHz (at wind speed of 5.5 m/s, 7.5 m/s, 12 m/s, 15 m/s, and 19.4 m/s), 13.9 GHz (same wind speeds as 9.4 GHz), and 13.99 GHz (at wind speed of 5 m/s, 10 m/s, 15 m/s, 20 m/s, 25 m/s, and 30 m/s). The implementation details are mentioned at the start of Section 4.1 and are not summarized here. The loss function is defined in Equation (14), and the optimizer to minimize the loss function is SGDM. Other training options were

the same as an illustrated in Section 4.1. Similarly, we trained and tested the proposed CNN models with each dataset 100 times for a robust and stable result.



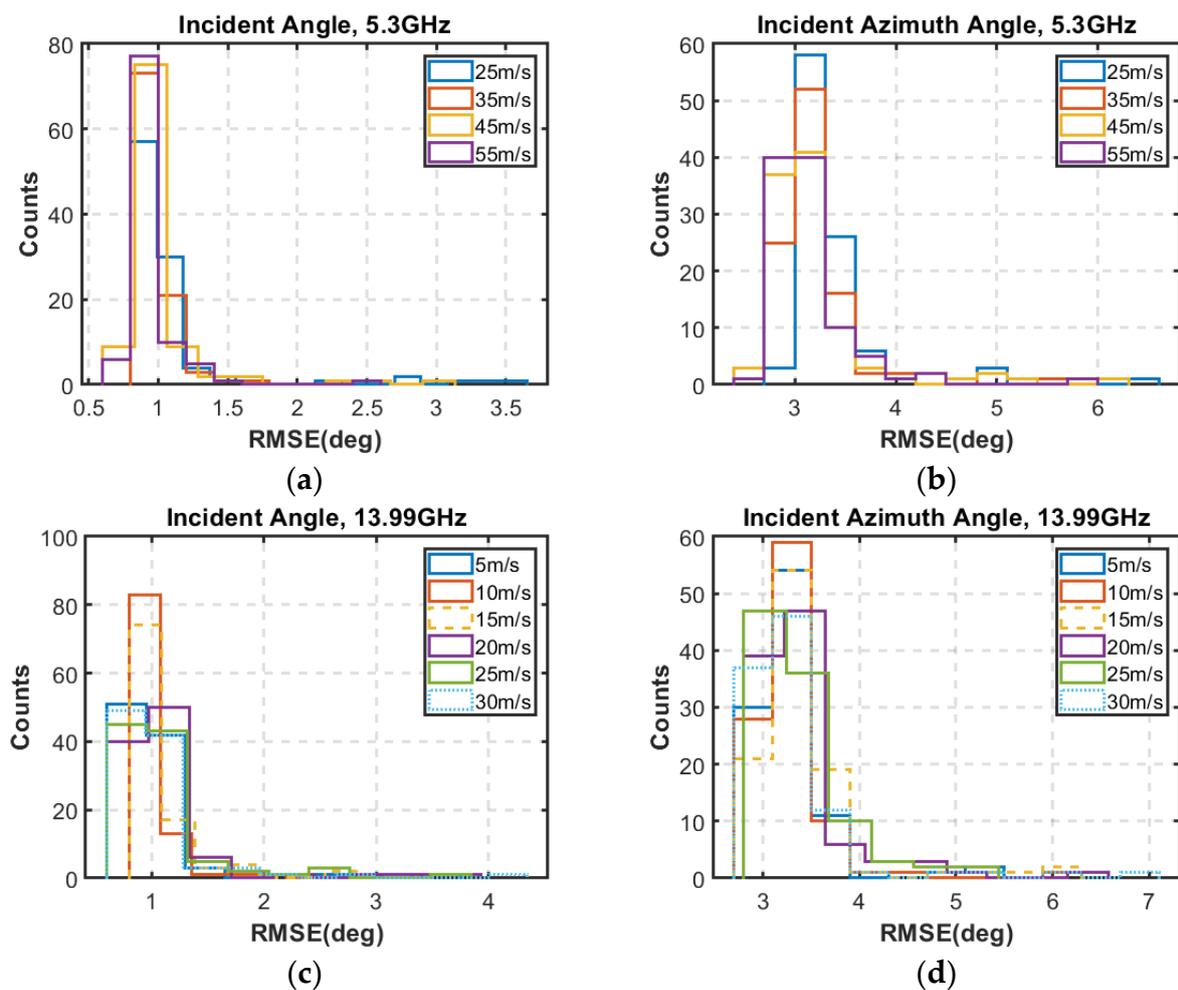
**Figure 5.** The CNN configuration employed for DOA estimation. The network is consisted of 10 convolutional layers, 4 average pooling layers, and a fully connected layer with 1024 hidden units.

#### 4.2.2. DOA Estimation Result

In this section, we evaluated the performance of Model 2 described in Figure 4. Table 2 summarizes the testing accuracy of DOA estimation-based sea surface scattering datasets at 5.3 GHz and 13.99 GHz. The values reported in this table are the average of 100 realizations. For incident angle estimation, the average RMSE of all datasets is about  $1^\circ$ , and for incident azimuth angle estimation, the average RMSE of all datasets is between  $3^\circ$  to  $3.5^\circ$ . For datasets at the 5.3 GHz operating frequency, we found that RMSE values of both the incident angle and incident azimuth angle are slightly decreased with increasing sea wind speed. However, for datasets under the 13.99 GHz operating frequency, there is no distinct relationship between RMSE and wind speeds. Figure 6 shows the distribution of RMSE between the truth and predicted incident source in 100 realizations. (a) and (b) are the RMSE distribution of  $\theta_i$  and  $\phi_i$  at 5.3 GHz and all wind speeds; (c) and (d) are the RMSE distribution of  $\theta_i$  and  $\phi_i$  at 13.99 GHz. The minimum RMSE of  $\theta_i$  and  $\phi_i$  in 100 realizations on 5.3 GHz, 25 m/s are  $0.80^\circ$  and  $2.97^\circ$ , on 35 m/s are  $0.80^\circ$  and  $2.70^\circ$ , on 45 m/s are  $0.73^\circ$  and  $2.68^\circ$ , and on 55 m/s are  $0.76^\circ$  and  $2.67^\circ$ . For 13.99 GHz, the minimum RMSE of  $\theta_i$  and  $\phi_i$  in 100 realizations on 5 m/s are  $0.80^\circ$  and  $2.87^\circ$ , on 10 m/s are  $0.82^\circ$  and  $2.77^\circ$ , on 15 m/s are  $0.81^\circ$  and  $2.93^\circ$ , on 20 m/s are  $0.82^\circ$  and  $2.90^\circ$ , on 25 m/s are  $0.79^\circ$  and  $2.94^\circ$ , and on 30 m/s are  $0.78^\circ$  and  $2.74^\circ$ . Besides, we added several samples consisting of measured data (Aquarius) and numerical model simulated data, and the network inversion are presented in Table 4.

**Table 4.** Results of 100 realizations of the 5.3 GHz and 13.99 GHz datasets using the CNN method illustrated in Figure 5. The numbers reported in this table are the average of all realizations.

$f$	$U$	RMSE of $\theta_i$	RMSE of $\phi_i$
5.3 GHz	25 m/s	$1.13^\circ$	$3.37^\circ$
	35 m/s	$0.98^\circ$	$3.20^\circ$
	45 m/s	$0.96^\circ$	$3.18^\circ$
	55 m/s	$0.94^\circ$	$3.16^\circ$
13.99 GHz	5 m/s	$1.05^\circ$	$3.31^\circ$
	10 m/s	$1.03^\circ$	$3.27^\circ$
	15 m/s	$1.09^\circ$	$3.40^\circ$
	20 m/s	$1.11^\circ$	$3.37^\circ$
	25 m/s	$1.11^\circ$	$3.42^\circ$
	30 m/s	$1.05^\circ$	$3.30^\circ$



**Figure 6.** The distribution of RMSE between the truth and predicted incident source for datasets at 5.3 GHz and 13.99 GHz in 100 realizations: (a) 5.3 GHz, incident angle; (b) 5.3 GHz, incident azimuth angle; (c) 13.99 GHz, incident angle; (d) 13.99 GHz, incident azimuth angle.

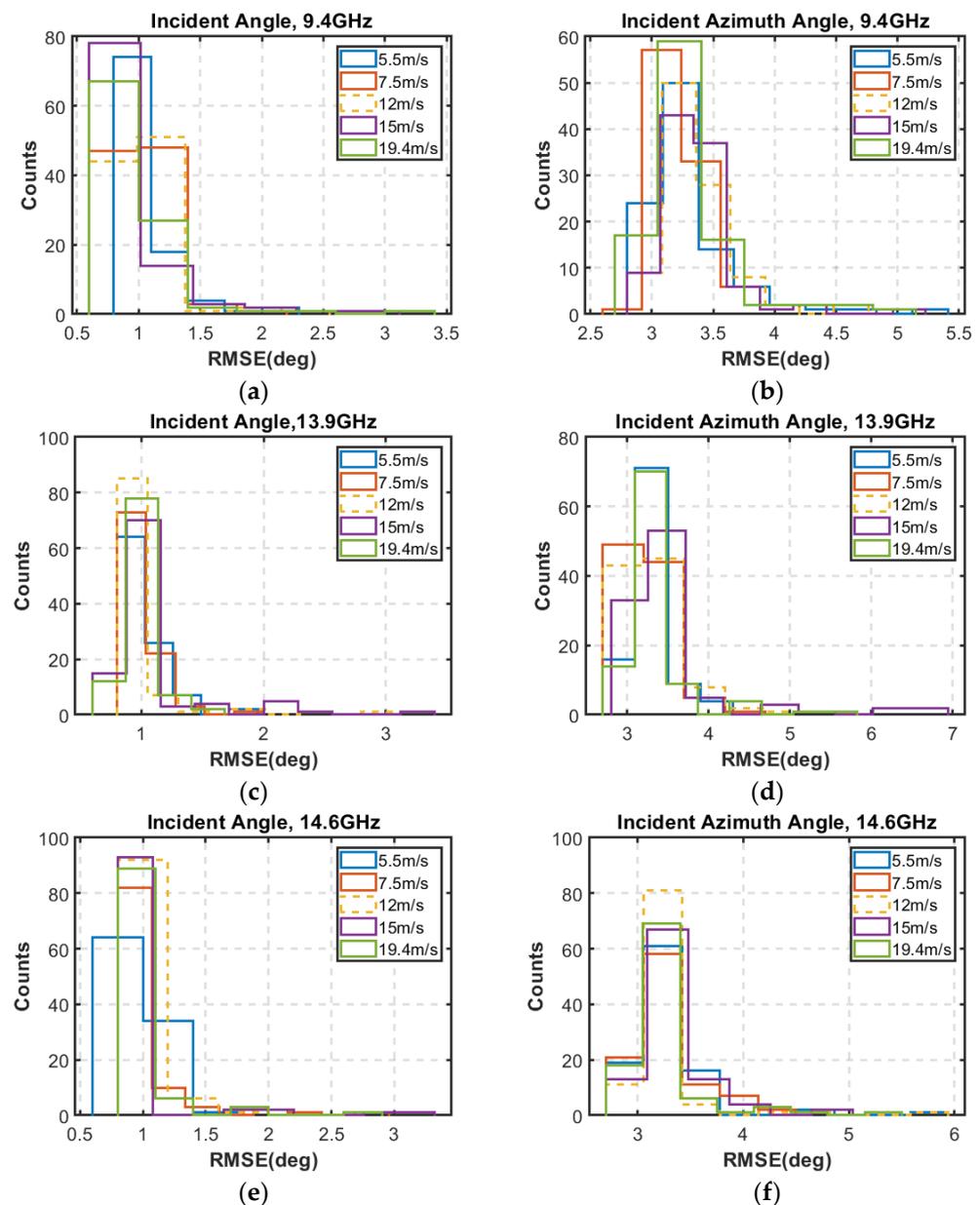
Furthermore, we explored the CNN performance for datasets at the same wind speeds but different radar frequencies. Table 5 shows the average results of 100 realizations of the 9.4 GHz, 13.9 GHz, and 14.6 GHz datasets. For incident angle estimation, the average RMSE of all frequencies and wind speeds can approximately reach  $1^\circ$ . As for the incident azimuth angle, this value is dispersed from  $3.24^\circ$  to  $3.52^\circ$ . Figure 7 displays the distribution of RMSE of the 9.4 GHz, 13.9 GHz, and 14.6 GHz datasets. For 9.4 GHz, the minimum RMSE of  $\theta_i$  and  $\phi_i$  in 100 realizations on 5.5 m/s are  $0.81^\circ$  and  $2.81^\circ$ , on 7.5 m/s are  $0.86^\circ$  and  $2.78^\circ$ , on 12 m/s are  $0.81^\circ$  and  $2.93^\circ$ , on 15 m/s are  $0.82^\circ$  and  $2.95^\circ$ , and on 19.4 m/s are  $0.82^\circ$  and  $2.85^\circ$ . For 13.9 GHz, the minimum RMSE of  $\theta_i$  and  $\phi_i$  in 100 realizations on 5.5 m/s are  $0.86^\circ$  and  $2.91^\circ$ , on 7.5 m/s are  $0.84^\circ$  and  $2.87^\circ$ , on 12 m/s are  $0.84^\circ$  and  $2.90^\circ$ , on 15 m/s are  $0.77^\circ$  and  $2.92^\circ$ , and on 19.4 m/s are  $0.80^\circ$  and  $2.94^\circ$ . For 14.6 GHz, the minimum RMSE of  $\theta_i$  and  $\phi_i$  in 100 realizations on 5.5 m/s are  $0.78^\circ$  and  $2.83^\circ$ , on 7.5 m/s are  $0.81^\circ$  and  $2.87^\circ$ , on 12 m/s are  $0.83^\circ$  and  $2.82^\circ$ , on 15 m/s are  $0.82^\circ$  and  $2.97^\circ$ , and on 19.4 m/s are  $0.81^\circ$  and  $2.88^\circ$ .

**Table 5.** Results of 100 realizations of the 9.4 GHz, 13.9 GHz, and 14.6 GHz datasets using the CNN method illustrated in Figure 5. The numbers reported in this table are the average of all realizations.

$f$	$U$	RMSE of $\theta_i$	RMSE of $\phi_i$
9.4 GHz	5.5 m/s	1.06°	3.30°
	7.5 m/s	1.04°	3.25°
	12 m/s	1.05°	3.37°
	15 m/s	1.03°	3.36°
	19.4 m/s	1.04°	3.30°
13.9 GHz	5.5 m/s	1.04°	3.33°
	7.5 m/s	1.01°	3.26°
	12 m/s	1.03°	3.32°
	15 m/s	1.08°	3.52°
	19.4 m/s	1.00°	3.36°
14.6 GHz	5.5 m/s	0.98°	3.27°
	7.5 m/s	1.04°	3.28°
	12 m/s	0.98°	3.24°
	15 m/s	1.01°	3.35°
	19.4 m/s	1.00°	3.24°

#### 4.2.3. Validation

In this paper, four types of validation data at the L-, C-, X-, and Ku-bands were adopted to examine the generalization ability of the proposed CNN structure. The L-band validation data for different wind directions come from the Aquarius scatterometer operating at 1.26 GHz, which has three incident angles from 28.7° to 45.6°. The backscattering coefficients from the sea surface are under 3, 5, 8, 10, 12, and 15 m/s wind speed [33,34]. The validation data for C- (5.3 GHz), X- (9.6 GHz), and Ku-bands (14 GHz) are generated by empirical CMOD7 [41], XMOD2 [42], and NSCAT-4 [43] models, respectively. Besides, other C- (5.3 GHz) and Ku-bands (13.9 GHz) data are radar measurements from [37] and [44], respectively. Due to the lack of bistatic scattering data, the scattering coefficients in a single validation sample are backscattered and mixed up with different frequencies and wind speeds. The incident angle and incident azimuth angle of the validation sample are 40° and 0°, respectively. Based on the above-mentioned validation data, the best prediction realized by the pre-trained CNN is 37.80° for the incident angle and 3.94° for the incident azimuth angle. As we can see, the estimation results of the mixed sample can achieve a satisfactory result.



**Figure 7.** The distribution of RMSE between the truth and predicted incident source for datasets at 13.9 GHz and 14.6 GHz in 100 realizations: (a) 9.4 GHz, incident angle; (b) 9.4 GHz, incident azimuth angle; (c) 13.9 GHz, incident angle; (d) 13.9 GHz, incident azimuth angle; (e) 14.6 GHz, incident angle; (f) 14.6 GHz, incident azimuth angle.

## 5. Discussion

### 5.1. Results Interpretation

In our experiments, by comparing the average RMSE of all datasets listed in Tables 4 and 5, we found that the performance of the CNN-based DOA estimation does not rely on radar frequencies and sea wind speeds. For almost all datasets, the average RMSE of incident angle is approaching  $1^\circ$ , and the average RMSE of the incident azimuth angle is between  $3.0^\circ$  and  $3.5^\circ$ . One possible reason is that the chosen CNN configuration is robust to the speckle noise appearing in coherent SAR measurements. Further investigation is required to generalize this claim.

### 5.2. Comparison with Other Algorithms

In this subsection, we compared the CNN-based method with the state-of-the-art DOA estimation algorithms, including machine learning or deep learning methods [13,45] and conventional parametric methods [46–48]. Those methods are proposed for local-scattered signal sources, which is quite similar to the sea surface scattering. The algorithms cited in this subsection process the signal sequence from the scattered sources, but in our case, the intensity of received signals was taken as the model input. The accuracy of the methods is summarized in Table 6. Using one-dimensional (1-D) DOA estimation means that only the incident angle was considered, while using the two-dimensional (2-D) DOA estimation, we predicted both the incident angle and incident azimuth angle of the signal sources. For 1-D DOA estimation, a machine learning-based approach, such as [45] utilizing the support vector machine (SVM), the RMSE of the incident angle was between  $1^\circ$  and  $1.5^\circ$  under the 7 dB of the signal to noise ratio (SNR). The authors of [46] used a multilayer perceptron neural network, and the RMSE of the DOA classifier was between  $0.8^\circ$  and  $3^\circ$  when 12.5% simulated training records were used (about 5895 samples in the dataset).

**Table 6.** The accuracy of other state-of-the-art signal processing-based DOA estimation algorithms.

	Algorithms	RMSE of $\theta_i$	RMSE of $\phi_i$
1-D DOA estimation	SVM [45]	$1^\circ$ – $1.5^\circ$	
	Annihilating filter with structured low rank approximation technique [46]	$0.8^\circ$ – $3^\circ$	—
	DNN [47]	$0.2^\circ$ – $0.4^\circ$	
2-D DOA estimation	ESPRIT-based method [48]	$0.4^\circ$ – $0.6^\circ$	$3^\circ$ – $4^\circ$
	Beamspace transformation [49]	$0.7^\circ$	$1^\circ$

In [13], the authors proposed a deep neural network configuration for two signals in the presence of array imperfections, and the RMSEs were between  $0.2^\circ$  and  $0.4^\circ$ . Note that the local scattering angle was not considered in [13]. For 2-D DOA estimation, [47] used the ESPRIT-based method. The RMSE of the incident angle and incident azimuth angle in [47] was between  $0.4^\circ$  and  $0.6^\circ$ ,  $3^\circ$  and  $4^\circ$  when the signal angular spreading degree was larger than  $8^\circ$ . In [48], the authors used the beamspace transformation method to estimate 2-D DOA. The RMSEs of the incident angle and incident azimuth angle were  $0.7^\circ$  and  $1^\circ$ , respectively, under 7 dB of SNR.

The accuracy in this paper is very close to the state-of-the-art signal processing-based works. Therefore, we can confirm that these results are quite satisfactory. We consider that using different deep learning techniques could lead to better results. For example, in [49], a Kalman filter-trained dynamic learning neural network (DLNN) was used to retrieve surface parameters from bistatic scattering data. Different from CNN, the configuration of DLNN is modified from multiple layer perceptron and each updated estimate of the DLNN weight is computed from the previous estimate and the new input data. The objective of [49] was to retrieve the surface parameters, while this paper was concerned with estimating the direction of the incident sources. Computationally, the sizes of input samples in [49] and in this work are also quite different.

### 5.3. Limitation

Perhaps, the biggest limitation of deep learning is its difficulty in generalization. Training a CNN to suit a variety of sea conditions and radar observations will lead to lower accuracy. Even though the observation in this article is omnidirectional, each dataset is still at a certain frequency and wind speed. Further investigation should ascertain the details of this trade-off between accuracy and generalization.

Another limitation is uncertainty. The existing uncertainty in these results is associated with the inaccuracy of speckle-interfered radar measurement modeling, sea surface

roughness assumptions, and retrieval errors. We plan to compare the simulated data with real radar measurement data and use other deep learning models as a baseline algorithm in future work to reduce this error.

## 6. Conclusions

In this study, we present a novel method of two-dimensional DOA estimation for SAR systems based on sea surface scattering and convolutional neural networks. To simulate the bistatic radar measurement data of sea surface scattering under a certain operating frequency and sea wind speed, we utilized the AIEM model to calculate the bistatic scattering coefficients and then added a K-distributed bias as speckle noise in the real radar measurements. The radar operating frequencies used in this paper included C-band, X-band, and Ku-band, and the range of the sea wind speed was from 5 m/s to 55 m/s. We proposed a CNN structure that learns the characteristics of scattering pattern and strength both lengthwise and widthwise. In this way, CNN can make full use of the radar measurement data under different sea wind conditions to improve the estimation accuracy of the incident source. For each dataset, we trained and tested it 100 times. The experimental results showed that the CNN structure in Figure 5 could be applied on DOA estimation with satisfactory accuracy at a wide range of frequencies and sea wind speeds. The average RMSE of the incident angle is about  $1^\circ$ , and the average RMSE of the incident azimuth angle is between  $3^\circ$  and  $3.5^\circ$ .

**Author Contributions:** Conceptualization, K.-S.C. and Y.Y.; methodology, X.Z. and Y.Y.; software, X.Z.; validation, X.Z., Y.Y. and K.-S.C.; formal analysis, X.Z.; investigation, X.Z.; resources, X.Z.; data curation, X.Z.; writing—original draft preparation, X.Z.; writing—review and editing, K.-S.C. and Y.Y.; visualization, X.Z.; supervision, K.-S.C.; project administration, K.-S.C.; funding acquisition, K.-S.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Guangxi Natural Science Youth Fund under Grant 2020GXNSFBFA297105, and Guangxi Natural Science Fund for Innovation Research Team under Grant 2019GXNSFGA245001.

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

## References

1. Bartlett, M.S. Smoothing Periodograms from Time-Series with Continuous Spectra. *Nature* **1948**, *161*, 686–687. [[CrossRef](#)]
2. Capon, J. High-Resolution Frequency-Wavenumber Spectrum Analysis. In *Proceedings of the IEEE*; IEEE: New York, NY, USA, 1969; pp. 1408–1418.
3. Lacoss, R.T. Data Adaptive Spectral Analysis Methods. *Geophysics* **1971**, *36*, 661–675. [[CrossRef](#)]
4. Schmidt, R.O. A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation. Ph.D. Thesis, Stanford University, Ann Arbor, MI, USA, 1982.
5. Kumaresan, R.; Tufts, D.W. Estimating the Angles of Arrival of Multiple Plane Waves. *IEEE Trans. Aerosp. Electron. Syst.* **1983**, *AES-19*, 134–139.
6. Reddi, S.S. Multiple Source Location-A Digital Approach. *IEEE Trans. Aerosp. Electron. Syst.* **1979**, *AES-15*, 95–105. [[CrossRef](#)]
7. Barabell, A. Improving the Resolution Performance of Eigenstructure-based Direction-Finding Algorithms. In Proceedings of the ICASSP '83-IEEE International Conference on Acoustics, Speech, and Signal Processing, Boston, MA, USA, 14–16 April 1983; pp. 336–339.
8. DeGroat, R.D.; Dowling, E.M.; Linebarger, D.A. The Constrained MUSIC Problem. *IEEE Trans. Signal Process.* **1993**, *41*, 1445–1449. [[CrossRef](#)]
9. Jha, S.; Durrani, T. Direction of Arrival Estimation Using Artificial Neural Networks. *IEEE Trans. Syst. Man Cybern. Syst.* **1991**, *21*, 1192–1201. [[CrossRef](#)]
10. Southall, H.L.; Simmers, J.A.; Donnell, T.H.O. Direction Finding in Phased Arrays with a Neural Network Beamformer. *IEEE Trans. Antennas Propag.* **1995**, *43*, 1369–1374. [[CrossRef](#)]
11. Rawat, A.; Yadav, R.N.; Shrivastava, S.C. Neural Network Applications in Smart Antenna Arrays: A Review. *Int. J. Electron. Commun.* **2012**, *66*, 903–912. [[CrossRef](#)]
12. Terabayashi, K.; Natsuaki, R.; Hirose, A. Ultrawideband Direction-of-Arrival Estimation Using Complex-Valued Spatiotemporal Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 1727–1732. [[CrossRef](#)]
13. Liu, Z.; 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](#)]

14. Zooghy, A.H.E.; Christodoulou, C.G.; Georgiopoulos, M.A. Neural Network-based Smart Antenna for Multiple Source Tracking. *IEEE Trans. Antennas Propag.* **2000**, *48*, 768–776. [[CrossRef](#)]
15. Valaee, S.; Champagne, B.; Kabal, P. Parametric Localization of Distributed Sources. *IEEE Trans. Signal Process.* **1995**, *43*, 2144–2153. [[CrossRef](#)]
16. Lee, Y.U.; Lee, S.R.; Kim, H.M.; Song, I. Estimation of Direction of Arrival for Angle-Perturbed Sources. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **1997**, *E80A*, 109–117.
17. Astely, D.; Ottersten, B. The Effects of Local Scattering on Direction of Arrival Estimation with MUSIC. *IEEE Trans. Signal Process.* **1999**, *47*, 3220–3234. [[CrossRef](#)]
18. Stoica, P.; Sharman, K.C. Maximum Likelihood Methods for Direction-of-Arrival Estimation. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1132–1143. [[CrossRef](#)]
19. Stoica, P.; Nehorai, A. MUSIC, Maximum Likelihood and Cramer-Rao Bound. In Proceedings of the ICASSP '88-International Conference on Acoustics, Speech, and Signal Processing, New York, NY, USA, 11–14 April 1988; Volume 4, pp. 2296–2299.
20. Stoica, P.; Nehorai, A. Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1783–1795. [[CrossRef](#)]
21. Tsakalides, P.; Nikias, C.L. Maximum Likelihood Localization of Sources in Noise Modeled as a Stable Process. *IEEE Trans. Signal Process.* **1995**, *43*, 2700–2713. [[CrossRef](#)]
22. Ottersten, B.; Viberg, M.; Stoica, P.; Nehorai, A. Exact and Large Sample Maximum Likelihood Techniques for Parameter Estimation and Detection in Array Processing. In *Radar Array Processing*; Springer: Berlin/Heidelberg, Germany, 1993; pp. 99–151.
23. Shan, T.J.; Wax, M.; Kailath, T. On Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Signals. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *33*, 806–811. [[CrossRef](#)]
24. Cai, B.; Li, Y.M.; Wang, H.Y. Forward/Backward Spatial Reconstruction Method for Directions of Arrival Estimation of Uncorrelated and Coherent Signals. *IET Microw. Antennas Propag.* **2012**, *6*, 1498–1505. [[CrossRef](#)]
25. Dudczyk, J.; Kawalec, A. Adaptive Forming of the Beam Pattern of Microstrip Antenna with the Use of an Artificial Neural Network. *Int. J. Antennas Propag.* **2012**, *2012*. [[CrossRef](#)]
26. Gierull, C.H. Ground Moving Target Parameter Estimation for Two-Channel SAR. *IEEE Proc.-Radar Sonar Navig.* **2006**, *153*, 224–233. [[CrossRef](#)]
27. Gierull, C.H. Azimuth Positioning of Moving Targets in Two-Channel SAR by Direction-of-Arrival Estimation. *Electron. Lett.* **2004**, *40*, 1380–1381. [[CrossRef](#)]
28. Fung, A.K.; Li, Z.; Chen, K.S. Backscattering from a Randomly Rough Dielectric Surface. *IEEE Trans. Geosci. Remote. Sens.* **1992**, *30*, 356–369. [[CrossRef](#)]
29. Chen, K.S.; Wu, T.D.; Tsang, L.; Li, Q.; Shi, J.; Fung, A.K. Emission of Rough Surfaces Calculated by the Integral Equation Method with Comparison to Three-Dimensional Moment Method Simulations. *IEEE Trans. Geosci. Remote. Sens.* **2003**, *41*, 90–101. [[CrossRef](#)]
30. Chen, K.S. *Radar Scattering and Imaging of Rough Surfaces: Modeling and Applications with MATLAB®*; CRC Press: Boca Raton, FL, USA, 2020.
31. Chen, K.S.; Fung, A.K.; Weissman, D.A. A Backscattering Model for Ocean Surface. *IEEE Trans. Geosci. Remote. Sens.* **1992**, *30*, 811–817. [[CrossRef](#)]
32. Xu, F.; Li, X.; Wang, P.; Yang, J.; Pichel, W.G.; Jin, Y. A Backscattering Model of Rainfall Over Rough Sea Surface for Synthetic Aperture Radar. *IEEE Trans. Geosci. Remote. Sens.* **2015**, *53*, 3042–3054. [[CrossRef](#)]
33. Xie, D.F.; Chen, K.S.; Zeng, J.F. The frequency selective effect of radar backscattering from multiscale sea surface. *Remote. Sens.* **2019**, *11*, 160. [[CrossRef](#)]
34. Xie, D.F.; Chen, K.S.; Yang, X.F. Effects of Wind Wave Spectra on Radar Backscatter from Sea Surface at Different Microwave Bands: A Numerical Study. *IEEE Trans. Geosci. Remote. Sens.* **2019**, *57*, 6325–6334. [[CrossRef](#)]
35. Fung, A.K.; Chen, K.S. *Microwave Scattering and Emission Models for Users*; Artech House: Boston, MA, USA, 2010.
36. Ulaby, F. *Microwave Radar and Radiometric Remote Sensing*; The University of Michigan Press: Chicago, IL, USA; Ann Arbor, MI, USA, 2014.
37. Fung, A.K. *Backscattering from Multiscale Rough Surfaces with Application to Wind Scatterometry*; Artech House: Boston, MA, USA, 2015.
38. Migliaccio, M.; Huang, L.; Buono, A. SAR Speckle Dependence on Ocean Surface Wind Field. *IEEE Trans. Geosci. Remote. Sens.* **2019**, *57*, 5447–5455. [[CrossRef](#)]
39. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
40. Liang, S.; Srikant, R. Why Deep Neural Networks for Function Approximation? In Proceedings of the International Conference on Learning Representations, Toulon, France, 24–26 April 2017.
41. Stoffelen, A.; Verspeek, J.A.; Vogelzang, J.; Verhoef, A. The CMOD7 Geophysical Model Function for ASCAT and ERS Wind Retrievals. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2017**, *10*, 2123–2134. [[CrossRef](#)]
42. Li, X.M.; Lehner, S. Algorithm for Sea Surface Wind Retrieval From TerraSAR-X and TanDEM-X Data. *IEEE Trans. Geosci. Remote. Sens.* **2014**, *52*, 2928–2939. [[CrossRef](#)]

43. Royal Netherlands Meteorological Institute. NSCAT-4 Geophysical Model Function. Available online: [http://knmi.nl/scatterometer/nscat\\_gmf](http://knmi.nl/scatterometer/nscat_gmf) (accessed on 1 May 2016).
44. Schroeder, L.; Schaffner, P.; Mitchell, J.; Jones, W. AAFE RADSCAT 13.9-GHz Measurements and Analysis: Wind-Speed Signature of the Ocean. *IEEE J. Ocean. Eng.* **1985**, *10*, 346–357. [[CrossRef](#)]
45. Pastorino, M.; Randazzo, A. A Smart Antenna System for Direction of Arrival Estimation Based on a Support Vector Regression. *IEEE Trans. Antennas Propag.* **2005**, *53*, 2161–2168. [[CrossRef](#)]
46. Wen, C.; Shi, G.M.; Xie, X.M. Estimation of directions of arrival of multiple distributed sources for nested array. *Signal Process.* **2017**, *130*, 315–322. [[CrossRef](#)]
47. Zheng, Z.; Li, G.J.; Teng, Y.L. Simplified Estimation of 2D DOA for Coherently Distributed Sources. *Wirel. Pers. Commun.* **2012**, *62*, 907–922. [[CrossRef](#)]
48. Zheng, Z.; Wang, W.Q.; Meng, H.P.; So, H.C.; Zhang, H.B. Efficient BeamSpace-Based Algorithm for Two-Dimensional DOA Estimation of Incoherently Distributed Sources in Massive MIMO Systems. *IEEE Trans. Veh. Technol.* **2018**, *67*, 11776–11789. [[CrossRef](#)]
49. Yang, Y.; Chen, K.S.; Shang, G.F. Surface Parameters Retrieval from Fully Bistatic Radar Scattering Data. *Remote. Sens.* **2019**, *11*, 596. [[CrossRef](#)]