electronics

MDPI

Article

An SAR Imaging and Detection Model of Multiple Maritime Targets Based on the Electromagnetic Approach and the Modified CBAM-YOLOv7 Neural Network

Peng Peng *, Qingkuan Wang, Weike Feng, Tong Wang and Chuangming Tong

Air Defense and Antimissile School, Air Force Engineering University, Xi'an 710054, China; qingkuan_wang@163.com (Q.W.); fengweike007@163.com (W.F.); tongwang001@163.com (T.W.); cmtong156@126.com (C.T.)
* Correspondence: pengpengxidian@163.com

Abstract: This paper proposes an Synthetic Aperture Radar (SAR) imaging and detection model of multiple targets at the maritime scene. The sea surface sample is generated according to the composite rough surface theory. The SAR imaging model is constructed based on a hybrid EM calculation approach with the fast ray tracing strategy and the modified facet Small Slope Approximation (SSA) solution. Numerical simulations calculate the EM scattering and the SAR imaging of the multiple cone targets above the sea surface, with the scattering mechanisms analyzed and discussed. The SAR imaging datasets are then set up by the SAR image simulations. A modified YOLOv7 neural network with the Spatial Pyramid Pooling Fast Connected Spatial Pyramid Convolution (SPPFCSPC) module, Convolutional Block Attention Module (CBAM), modified Feature Pyramid Network (FPN) structure and extra detection head is developed. In the training process on our constructed SAR datasets, the precision rate, recall rate, mAP@0.5 and mAP@0.5:0.95 are 97.46%, 90.08%, 92.91% and 91.98%, respectively, after 300 rounds of training. The detection results show that the modified YOLOv7 has a good performance in selecting the targets out of the complex sea surface and multipath interference background.

Keywords: SAR imaging; YOLOv7 network; electromagnetic scattering model; maritime targets

check for updates

Citation: Peng, P.; Wang, Q.; Feng, W.; Wang, T.; Tong, C. An SAR Imaging and Detection Model of Multiple Maritime Targets Based on the Electromagnetic Approach and the Modified CBAM-YOLOv7 Neural Network. *Electronics* **2023**, *12*, 4816. https://doi.org/10.3390/electronics12234816

Academic Editor: Heung-Il Suk

Received: 17 October 2023
Revised: 23 November 2023
Accepted: 24 November 2023
Published: 28 November 2023

1. Introduction

Synthetic Aperture Radar (SAR) is an all-weather and all-time remote sensing platform that can provide high-resolution microwave radar imaging features of the target. The SAR imaging and detections of maritime targets have significant applications in the aspects of sea target monitoring and recognition, sea rescuing, sea crisis management, etc. [1–5]. Sea targets usually have different sizes and various scattering mechanisms. These factors are challenges for sea target detection. In recent years, Deep Learning (DL) methods have been introduced and show excellent performance in the image recognition and detection domain [6–10]. You Only Look Once (YOLO) is one of those methods, which has been developed in various series and shows good performance in image detection [11–15]. YOLOv7 is the latest version at the time of this research. It has an optimized training process and a higher detection accuracy, as indicated in the latest studies [15–17]. In maritime SAR detections, the DL method requires plenty of SAR datasets to serve as training samples. The practical maritime environment varies with time, showing different radar features for various sea states. In addition, it can take a long time to obtain extensive SAR samples by means of practical measurements. Furthermore, SAR image samples during high sea states are also not easy to acquire. The physical mechanisms behind the SAR features and their influence on SAR detection performance are also not easy to illuminate by measurement alone.

Electromagnetic models can serve as a close-to-real tool to set up the SAR imaging model and employ the SAR data simulations with higher efficiencies owing to their low cost, high verisimilitude and convenient implementation. The electromagnetic scattering calculation model needs to be set up for the sea targets. This issue has complex composite scattering mechanisms, which include not only the individual target and sea scattering but also the target-and-target and target-and-sea interactions. At the SAR frequencies, the composite targets and real sea surface model are electrically very large with complex scattering mechanisms in the radar sights. All these factors make the issue troublesome and intriguing. The rigorous full-wave models (e.g., MoM, MLFMA, FEM, FDTD) [18–20] are accurate, but they are mostly limited to handling relatively electrically small-scale EM simulations, which are far from practical uses. Thus, the analytical models are more attractive in practical applications since they have more applicable formulations with computational efficiencies. Some models have emerged in recent years, such as the Four Path Model (FPM) [21] and Half-Space Physical Optics (HSPO) [22]. They are simple and efficient to apply in reality. However, they generally simplify the scattering mechanisms and ignore the local sea surface scattering and the complex high-order interactions in the composite target-sea scattering calculation, which is significant in the SAR imaging simulations. A "facet-based" model was developed in some previous works, which can describe the local scattering in the electrically large maritime scene accurately and efficiently. The large-scale sea surface is meshed by the facet elements with local scattering configurations. The SAR echo from these facet elements can be easily obtained. Franceschetti et al. [23,24] developed the facet model with the Kirchhoff solution and applied the model in ocean SAR echo simulations. Min Zhang et al. [25,26] modified the facet model with Two-Scale Methods (TSM) by further considering the Bragg phenomena and combining it with the four-path model for simulating the SAR image of the composite ship–ocean scene. However, deficiencies still exist. The two-scale method utilizes an ad hoc "cutoff" wavenumber as the selection criteria to determine the proportion of contributions from the multi-scale sea surface structures, which may affect the correctness of the calculation results. In addition, the facet-based four-path model is also too approximate to fully consider the local interactions between the multiple targets and the sea.

This paper aims to provide an SAR imaging and detection model with the EM approach and a modified CBAM-YOLOv7 neural network. The EM approach utilizes a facet-based sea scattering description and a hybrid high-frequency electromagnetic model. The EM approach is applied to build the SAR image simulation datasets. A modified CBAM-YOLOv7 neural network is applied for multiple target detection in the maritime scene. The rest part of the paper is organized as follows: The description of the EM approach to calculate the composite target and sea scattering is given in Section 2. Then, the SAR imagery simulator is set up, and the simulations are created. Section 3 constructs the modified YOLOv7 neural network model and accomplishes the SAR image training and detections, with the performance of the model compared and discussed. Finally, Section 4 ends with the conclusions.

2. SAR Imaging Model with the Electromagnetic Approach

2.1. The Sea Surface Scattering Calculations

The practical sea surface has composite structures, as illustrated in Figure 1. Assume the time-varied sea height map is $\xi(\boldsymbol{r}, t)$. The large-scale waves, $z(\boldsymbol{r}, t)$, can be generated at each time moment as a rough surface sample by the Monte Carlo method, according to the rough surface theory with a certain sea spectrum [27]. The Elfouhaily's sea spectrum and the spreading function [28] are used in this paper. In the facet model, the long waves are meshed by large planar facets. The short waves $\zeta_c(\boldsymbol{r})$ serve as modifications upon the planar facet. As indicated in the Bragg theory, the radar scattering contributions from the short waves are mainly from its Bragg wave components [29,30], which can be represented

by the expansion of sinusoidal waves traveling at the Bragg resonant frequency with the form:

$$\zeta_c(\boldsymbol{r}) = B(\boldsymbol{\kappa}_c)\sin(\boldsymbol{\kappa}_c \cdot \boldsymbol{r}_c + \omega_c t + \psi) \tag{1}$$

where $\boldsymbol{\kappa}_c$ is the Bragg wave vector, $B(\boldsymbol{\kappa}_c) = \sqrt{S(\boldsymbol{\kappa}_c)/\Delta S}$ is the amplitude, ψ is the random phase, $S(\boldsymbol{\kappa}_c)$ is the short wave spectrum, ω_c is the circle frequency, which is assumed to follow the dispersion relationship in the deep water with $\omega_c^2 = \kappa g$ and g as the gravity acceleration.

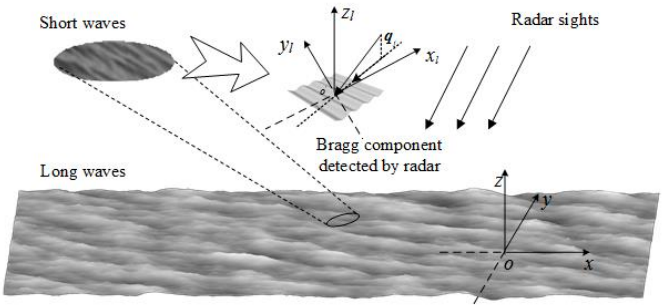


Figure 1. The geometry of multi-scale sea surface and the coordinates.

A Facet Small Slope Approximation (SSA) method is employed as the EM approach in calculating the sea surface scattering. The SSA method was proposed by Voronovich et al. [31–33] in a description of the scattering from an electrically large and complex environment. However, it only provides an average scattering coefficient of the sea surface, which is not enough to provide a high-resolution SAR imaging return. In the earlier studies, the SSA method was combined with the facet model to study the local scattering of the multi-scale structures of the complex environment, with good agreement with the measurement data of the practical environment radar scattering [34–36]. In Figure 1, the global and local coordinates are adopted to illustrate the scattering configurations. The Scattering Amplitude (SA) can be calculated by

$$S^{SSA} = \frac{2(qq_0)^{1/2}}{(q+q_0)P_{inc}^{1/2}}\boldsymbol{B}(\boldsymbol{k},\boldsymbol{k}_0)\int \frac{T(\boldsymbol{r},z(\boldsymbol{r}))}{(2\pi)^2} e^{-i(\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{r}-i(q+q_0)\xi(\boldsymbol{r})} d\boldsymbol{r} \tag{2}$$

where $\boldsymbol{B}(\boldsymbol{k}, \boldsymbol{k}_0)$ is a polarization matrix [31–33], $\boldsymbol{k}_0$ and q_0 are the horizontal and vertical projections of the incident vector $\boldsymbol{k}_i$, and $\boldsymbol{k}$ and q are the horizontal and vertical projections of the scattering vector $\boldsymbol{k}_s$. The scattering vector is $\boldsymbol{q} = k(\boldsymbol{k}_s - \boldsymbol{k}_i)$. $T(\boldsymbol{r}, z(\boldsymbol{r}))$ is the taper wave function. P_{inc} is the incident power. The taper factor is set as $g_x = L_x/4$, and $g_y = L_y$, as suggested by [37]. L_x and L_y are the lengths of the sea surface along the x- and y-directions.

In the classical SSA, the size of the meshed facets is requested to ensure that the integral upon the short-wave structure is precisely solved [31–33]. It is hard for the method to handle the electrically very large environment. The calculation also consumes a lot of computation resources. Here, the SSA is employed on the large sea facets with the Bragg wave modifications, which can save a lot of calculation time. The integral kernel upon the facet can be analytically solved as:

$$\begin{aligned} I &= \int \frac{T(\boldsymbol{r},z(\boldsymbol{r}))}{(2\pi)^2} \cdot \mathrm{e}^{-i(\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{r}_c - i(q+q_0)(z(\boldsymbol{r}_c)+\zeta_c(\boldsymbol{r}))} d\boldsymbol{r} \\ &= \frac{T(\boldsymbol{r}_c,z(\boldsymbol{r}_c))}{(2\pi)^2} \cdot \frac{\Delta S}{2n_z} \cdot \mathrm{e}^{-i(\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{r}_c - i(q+q_0)z(\boldsymbol{r}_c)} \\ &\quad \cdot \left\{ B(\boldsymbol{\kappa}_c^+) \sum_{n=-\infty}^{\infty} (-i)^n J_n[q_z \mathrm{B}(\boldsymbol{\kappa}_c^+)] I_0(\boldsymbol{\kappa}_c^+) + B(\boldsymbol{\kappa}_c^-) \sum_{n=-\infty}^{\infty} (-i)^n J_n[q_z \mathrm{B}(\boldsymbol{\kappa}_c^-)] I_0(\boldsymbol{\kappa}_c^-) \right\} \end{aligned} \tag{3}$$

where ΔS is the facet area, $\boldsymbol{r}_c$ is the center of the large sea facet, $J_n(\cdot)$ is the Bessel function, and only the dominant terms ($n = 0, \pm 1$) are reserved in the Bessel series expansion. At the facet, the local coordinate, $I_0(\boldsymbol{\kappa}_c)$, is given by

$$\begin{aligned} I_0(\boldsymbol{\kappa}_c) = {} & \mathrm{e}^{-i(1+n)\omega_c t} sinc\left\{\tfrac{\Delta x}{2}[(1+n)\kappa_{\mathrm{cx}} - q_x - q_z z_x]\right\} \cdot sinc\left\{\tfrac{\Delta y}{2}[(1+n)\kappa_{\mathrm{cy}} - q_y - q_z z_y]\right\} \\ & + \mathrm{e}^{-i(1-n)\omega_c t} sinc\left\{\tfrac{\Delta x}{2}[(1-n)\kappa_{\mathrm{cx}} + q_x + q_z z_x]\right\} \cdot sinc\left\{\tfrac{\Delta y}{2}[(1-n)\kappa_{\mathrm{cy}} + q_y + q_z z_y]\right\} \end{aligned} \tag{4}$$

The scattering coefficient is defined as

$$\sigma^{SSA} = 4\pi q q_0 \left| S^{SSA} \right|^2 \tag{5}$$

Figure 2 presents the simulated backward scattering coefficients. The simulated area of the sea is 128 m × 128 m. The sea facets are meshed into the size of 1 m × 1 m. The dielectric constant of the sea is calculated according to the Klein model [38] at 20 °C and 32.5% of salinity. The incident frequency is 13.9 GHz. The incident angle is θ_{i} = 30° and φ_{i} = 0°. The results are obtained by averaging the simulated scattering from 100 sea realizations. Comparisons are made between the simulated results and the measurement data in the reference [39], the results from the SSA-2 method proposed by Voronovich [31–33], and the two-scale method proposed by Andreas Arnold-Bos and Ali Khenchaf [40]. The sea wind speed in Elfouhaily's sea spectrum U_{10} = 5 m/s. From the results, one can see that our method and the SSA-2 method better agree with the measurement results in the real scene compared with the TSM results for different incident angles in both polarizations. Additionally, the facet SSA method has a unified formula and better adaptability in the description of the specular and diffuse scattering upon the local sea surface in different sea conditions. This can validate the accuracy of our method.

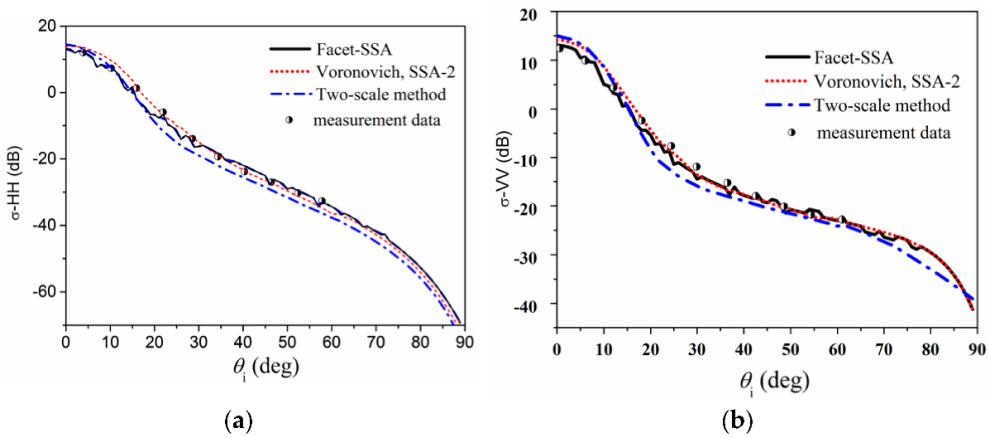


Figure 2. The backscattering coefficient of the sea. (**a**) HH–polarization. (**b**) VV–polarization.

2.2. The Calculations of the Composite Target and Sea Model Scattering

The composite target and sea model involve hybrid scattering mechanisms. At the high microwave radar frequencies, the scattering from the target and sea surface model becomes localized and uncorrelated. A facet-based hybrid high-frequency scattering calculation approach is adopted in calculating the composite EM scattering. For the illuminated target facets, the scattering field is calculated by the Physical Optics (PO) method, as given by

$$\boldsymbol{E}^{PO} = \frac{jk\mathrm{e}^{-jkR}}{4\pi R} \iint \eta \cdot \hat{k}_s \times (\hat{k}_s \times \boldsymbol{J}(\boldsymbol{r}')) \mathrm{e}^{jk\hat{k}_s \cdot \boldsymbol{r}'} ds' \tag{6}$$

where $\hat{k}_s$ is the unit scattering vector. R is the radar-to-target distance. $\boldsymbol{J}(\boldsymbol{r}')$ is the equivalent current for the illuminated facets, which is calculated as

$$\boldsymbol{J}\left(\boldsymbol{r}'\right) = 2\hat{\boldsymbol{n}} \times \boldsymbol{H}_{inc} \tag{7}$$

Here, the tapered wave is also used as the incident wave. The integral in Equation (6) can be solved by Gordon's method [41]. The coupling interactions among facets are processed by a fast ray tracing process, as shown in Figure 3.

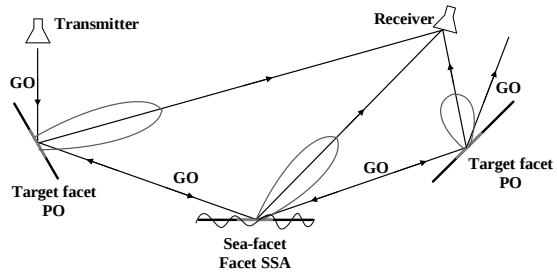


Figure 3. The coupling interactions calculation.

The incident wave is modeled by the ray tube, which traces and bounces according to the GO principles until it no longer intersects with any facet. Once the target or the sea facet is shone by the ray, the scattered field is, respectively, calculated by PO and the facet SSA method. The total scattering field is vector summed by the scattered field from all the facet elements. The ray tracing process can be further accelerated by the bi-directional ray tracing technique [42] and the KD-tree acceleration technique [43].

Apart from the surface scattering, the diffractions from the target edges also contribute. They make a specific contribution to the composite scattering. These components are evaluated by the EEC method as follows

$$\boldsymbol{E}^{\mathrm{d}} = \frac{ik_0}{4\pi}\frac{e^{-\mathrm{j}k_0 r}}{r}\int_l [\hat{s}\times(\hat{s}\times\hat{t})I_e + (\hat{s}\times\hat{t})I_m]e^{ik_0\hat{\mathbf{s}}\cdot\boldsymbol{r}'}\mathrm{d}t' \tag{8}$$

I_e and I_m have the form:

$$I_e = \frac{i2\boldsymbol{E}_{inc}\cdot\hat{t}D_e^{EEC}}{k\sin^2\beta_i} + \frac{i2\eta\boldsymbol{H}_{inc}\cdot\hat{t}D_{em}^{EEC}}{k\sin\beta_i} \tag{9}$$

$$I_m = \frac{i2\eta\boldsymbol{H}_{inc}\cdot\hat{t}}{k\sin\beta_i\sin\beta_s}D_m^{EEC} \tag{10}$$

where $\boldsymbol{E}_{inc}$ and $\boldsymbol{H}_{inc}$ are the incident fields. $\hat{t}$ is the unit vector along the edges. D_e^{EEC}, D_{em}^{EEC}, D_m^{EEC} are the EEC diffraction coefficients [44]. The scattering coefficient is defined for the composite scattering, as given by

$$\sigma = \lim_{r\to\infty} 4\pi r^2|E_{sca}|^2/|E_{inc}|^2 \tag{11}$$

The bistatic scattering coefficients are shown in Figure 4.

Figure 4a shows the different polarization scattering coefficients of the multiple targets above the sea surface, which consisted of a long cone (cone length is 5 m, base radius is 1 m) and a short cone target (cone length is 3 m, base radius is 1 m). Figure 4b calculates the bistatic scattering from a group of two long cones and two short cones. The cones are 5 m high above the sea. The incident angle wave is $\theta_i = 40°, \varphi_i = 0°$. One can see that the composite scattering has the strongest value in the specular direction, which is mainly contributed by the sea surface scattering. The reflection from the cone surface, as well as the coupling between the sea and the cone, also causes peak values in the backward directions. One can also see that the coupling effect for the four-cone group is stronger, which disperses the scattering energy into other diffuse directions.

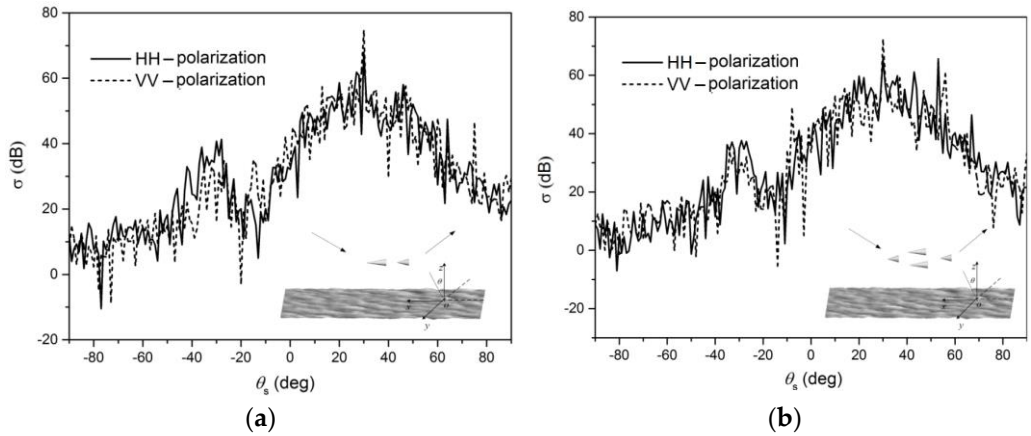


Figure 4. The scattering from multiple cone targets above the sea surface. (**a**) Two cone targets. (**b**) Four cone targets.

2.3. The SAR Imaging Model

The SAR image model is set up based on the airborne strip mode in this section, as shown in Figure 5.

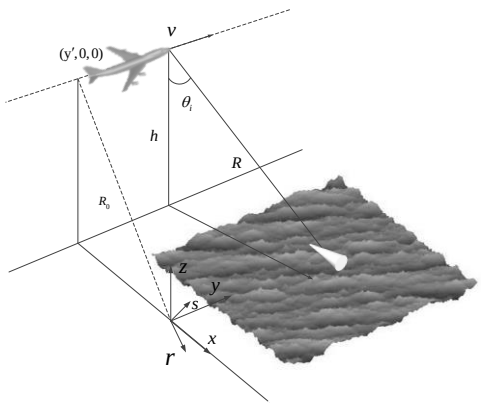


Figure 5. Sketch of the synthetic aperture radar imaging model.

The SAR carrier platform is flying at a velocity of v along the y-axis. θ_i is the radar squint angle. R is the distance from the radar to the scatter in the simulated scene. The SAR raw echo is set up at the coordinate system in Figure 5, given by

$$s(y', r', t) = \iint \sigma(y, r, t) g(r' - r, y' - y) dy dr \tag{12}$$

where

$$g(r' - r, y' - y) = rect\left[\frac{2(r' - r - \Delta R)}{c\tau}\right] \cdot \omega^2\left(\frac{y' - y}{Y}\right) \mathrm{e}^{-j\frac{4\pi}{\lambda}l} \cdot \mathrm{e}^{-j\frac{4\pi}{\lambda}\{(r'-r-\Delta R)^2+(y-y'-\Delta y)^2\}^{\frac{1}{2}}} \tag{13}$$

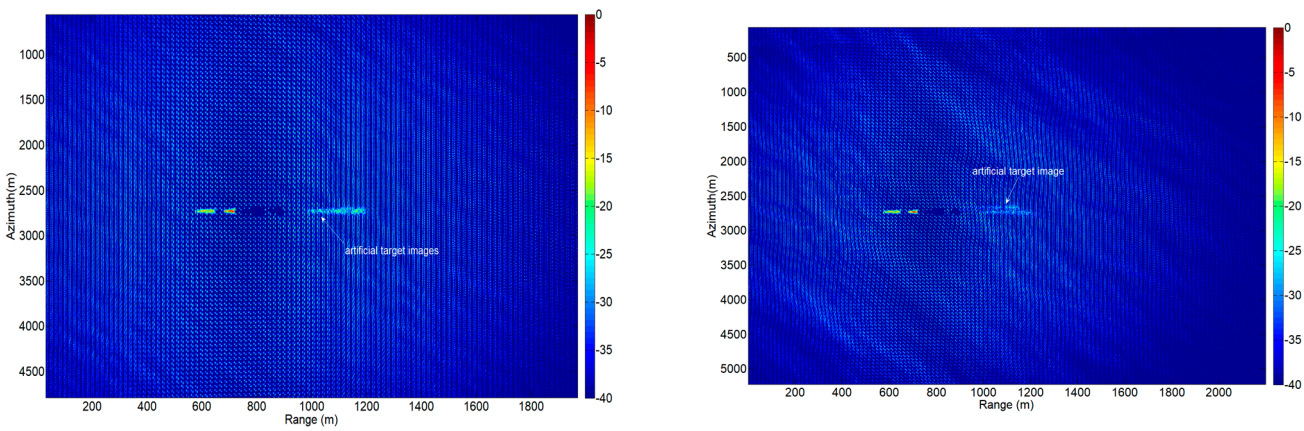
where ΔR and Δy are the range and azimuth position alterations of the scatterers. f is the radar working center frequency. Δf is the frequency bandwidth. τ is the pulse duration. λ is the wavelength. $rect(\cdot)$ is the rectangular window function. c is the propagation speed of the electromagnetic waves. $\omega(\cdot)$ is the antenna directionality function. $Y = \lambda R_0/L$ is the azimuth width of the antenna beam footprint. σ is the backscattering coefficient of the scatterer in the scene. l is the length of the radar wave traveling. The sea height maps are generated, and σ for each scatterer is calculated at each slow time moment. Through image focusing [45], the ensemble-averaged SAR image intensity is obtained by

$$I(x,y) = B\iint dx'dy'\left\langle\left|\sigma(x',y',t)\sigma^*(x',y',t)\right|\right\rangle \cdot \frac{\rho_{aN}}{\rho'_{aN}(x',y')} \cdot f_r(x-x') \cdot \mathrm{e}^{-\left(\frac{\pi}{\rho'_{aN}(x',y')}\right)^2 \cdot \left[y-y'-\frac{R}{v}\cdot u_r(x',y')\right]^2} \tag{14}$$

where $\langle\cdot\rangle$ refers to the average of the scattering coefficients in the integration time. u_r refers to the orbital velocity. $f_r(\cdot)$ refers to the range resolution function. $\rho_{aN} = N\rho_a$ is the azimuth resolution after N times of SAR incoherent illuminations. $\rho_{aN}{}'$ is the degraded azimuth resolution because of the target and sea scatter movements.

In the following simulations, the SAR image characteristics of the low-flying cone targets at the maritime scene are investigated. The targets fly along the negative x-axis direction with a constant velocity of 100 m/s. The velocity of the SAR platform is v = 100 m/s. The radar incident angle θ_i = 30°. The carrier frequency is 10 GHz. The base-band bandwidth is 1 GHz.

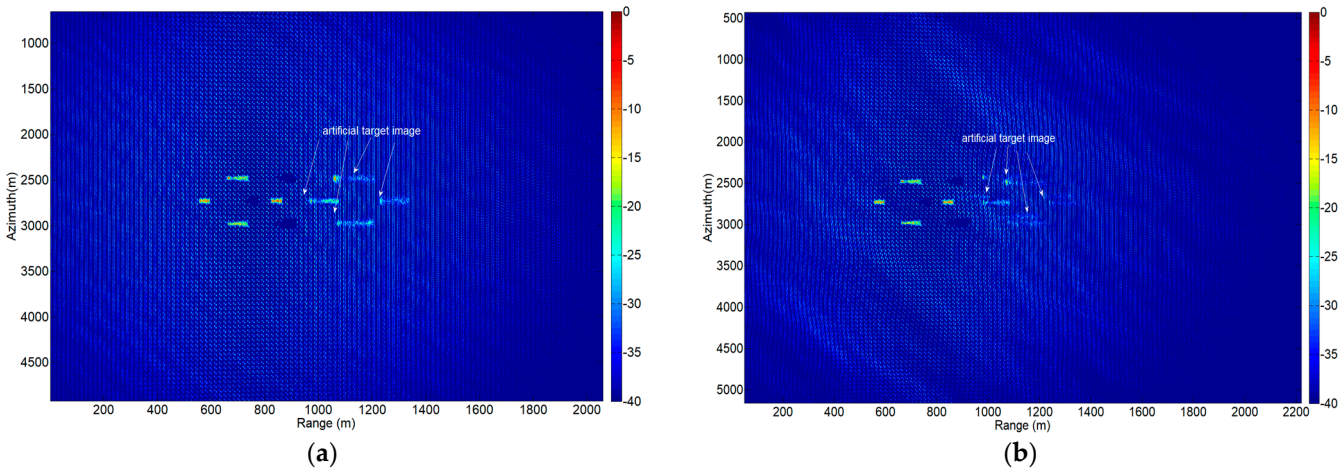
Figure 6 presents the simulated SAR images for the two-cone group targets at different sea states (U_{10} = 3 m/s and U_{10} = 10 m/s). The long and short cone targets, as well as the distribution characteristics of the wave propagations along the wind directions, can all be clearly identified. The color bar indicates the radar scattering intensities. The shadow areas of the sea images are the targets' shadows under the radar wave illuminations. One can also identify the artificial target images caused by the multipath echo responses, which can also be clearly observed in Figure 6a. These images are quite similar to the images from the direct target responses, but they have weaker intensities, shifts and distortions compared with the target images in both the range and azimuth directions. They are caused by the phase delay of the multipath interactions. These phenomena are also observed and referred to as "multipath ghosts" in other related studies [46,47]. One can also observe in Figure 6b that the intensities of the multipath response images can become even weaker at the higher sea state and are not easily identified.



(**a**) (**b**)

Figure 6. Synthetic aperture radar images of two-cone group targets above the sea at different sea states. (**a**) Low sea state (U_{10} = 3 m/s). (**b**) High sea state (U_{10} = 10 m/s).

In Figure 7, the SAR characteristics of the four-cone group targets are further investigated. It can be seen that the artificial targets are distributed in a broader area. At the high sea state in Figure 7b, more artificial images exist, but their intensities are weaker. From the above simulations, one can see that the SAR images of the targets at maritime scenes are much more complex than those of targets in the free space. In this scenario, the target detections can suffer more influence from the sea clutter and the multipath interactions between the target and the sea surface. This issue is also indicated in many references [46–49].



(**a**) (**b**)

Figure 7. Synthetic aperture radar images of four-cone group targets above the sea at different sea states. (**a**) Low sea state (U_{10} = 3 m/s). (**b**) High sea state (U_{10} = 10 m/s).

3. SAR Image Detections with the YOLOv7 Neural Network

3.1. Modified CBAM-YOLOv7 Neural Network

YOLOv7 neural network is an effective target detection method in the YOLO series [11–15]. In comparison with the earlier version of the YOLO series, the YOLOv7 network introduces a multi-scale testing technique that shows good potential in complex SAR image target detections [15–17]. To handle the difficulties in the issue of the SAR detections of the targets in the complex maritime scene, a modified YOLOv7 method is adopted, the structure of which is shown in Figure 8.

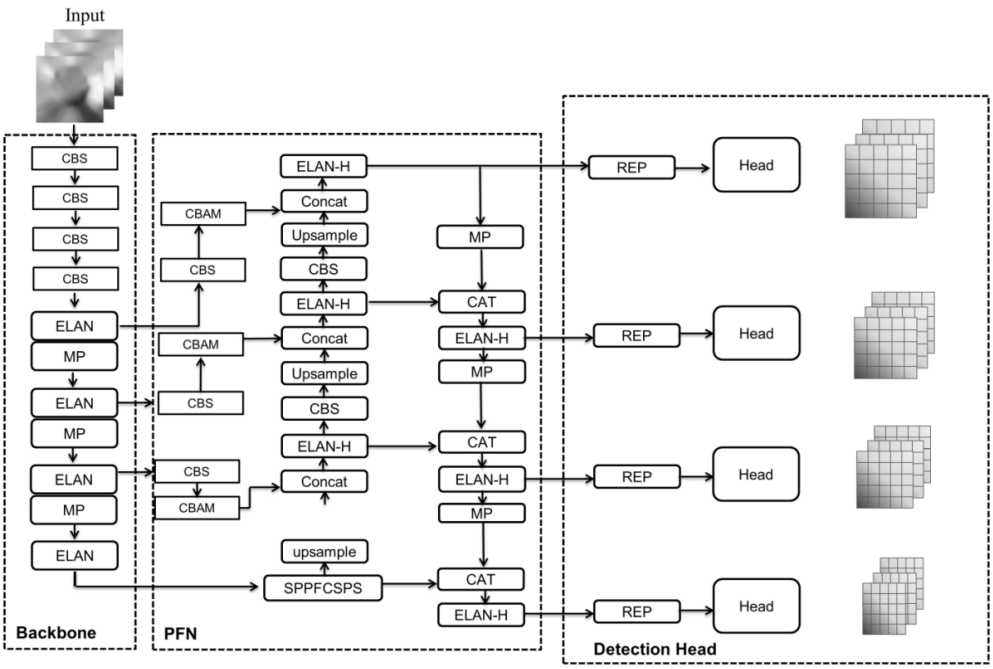


Figure 8. The modified Convolutional Block Attention Module–You Only Look Once version 7 (CBAM-YOLOv7) structure.

The YOLOv7 structure has three main parts: the backbone network, the Feature Pyramid Network (FPN) and the Detection head. The backbone network has four CBS modules. The CBS module extracts the underlying target features by means of the composite convolutional module, the batch normalization (BN) module and the SiLU module. The Max Pool

(MP) module and the Efficient Layer Aggregation Network (ELAN) are used for feature fusion [15]. Their structures are shown in Figure 9.

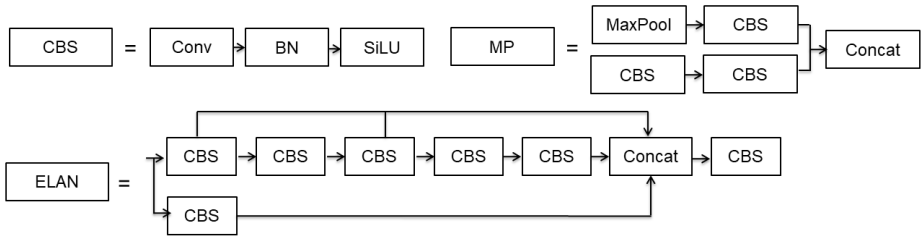


Figure 9. The structures of the CBS, Max Pool (MP) and Efficient Layer Aggregation Network (ELAN) in the You Only Look Once version 7 (YOLOv7) backbone.

The Spatial Pyramid Pooling Connected Spatial Pyramid Convolution (SPPCSPC) module is used in the original YOLOv7 [15], which performs down-sampling through parallel pooling operations, as in Figure 10. The SPPCSPC has three differently sized convolution kernels, which can greatly improve network calculations. A Spatial Pyramid Pooling Fast Connected Spatial Pyramid Convolution (SPPFCSPC) module is used in this study instead of the SPPCSPC module [50], whose structure is shown in Figure 10. In the SPPFCSPC module, the serial maximum pooling operations are adopted in the parallel maximum pooling operation, which can raise the computation efficiency while maintaining the receptive field.

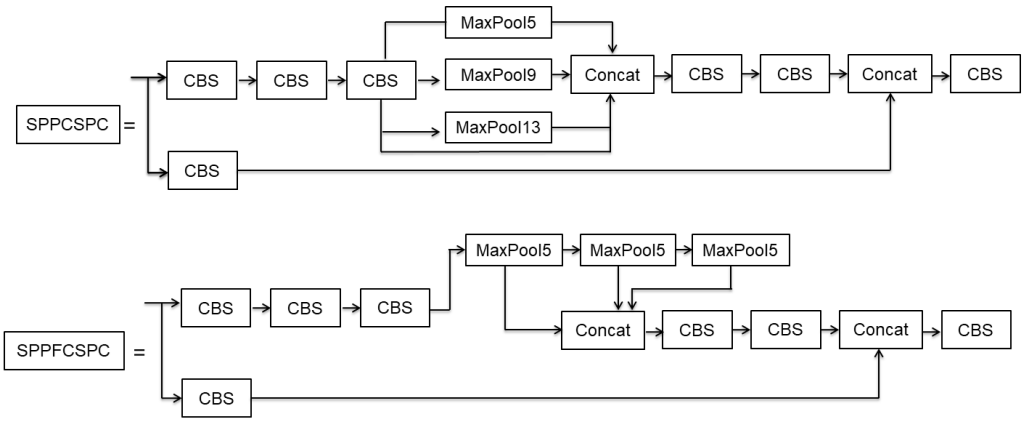


Figure 10. The structures of the Spatial Pyramid Pooling Connected Spatial Pyramid Convolution (SPPCSPC) and the Spatial Pyramid Pooling Fast Connected Spatial Pyramid Convolution (SPPFCSPC).

On the basis of the original YOLOv7 network, the Convolutional Block Attention Module (CBAM) [51–53] is employed to highlight the important target features. It can enhance the important features while suppressing non-important features. The spatial features are added together through the average pooling and the maximum pooling. The structure of the CBAM is presented in Figure 11.

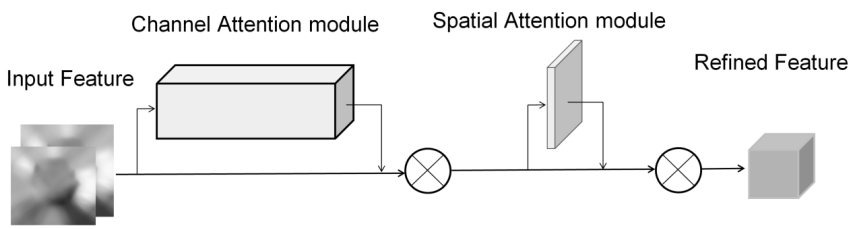


Figure 11. The Convolutional Block Attention Module (CBAM) structure.

It contains the Channel Attention Mechanism (CAM) module and the Spatial Attention Mechanism (SAM) module [51]. The channel attention concerns the meaningful information in a given image. The spatial attention concerns the location of the meaningful area in the image. The input feature **F** is transferred into the feature map of $\mathbf{F} \in \mathbb{R}^{C\times H\times W}$, where C refers to the channel number and H and W refer to the height and width of the feature map. The channel attention map $\mathbf{M_c} \in \mathbb{R}^{C\times 1\times 1}$ is calculated by the CAM module by squeezing the feature maps of the pooling layers with the activation function given by

$$\mathbf{M_c}(\mathbf{F}) = \text{Signmoid}(\text{MLP}(\text{AvgPool}(\mathbf{F})) + \text{MLP}(\text{MaxPool}(\mathbf{F}))) \tag{15}$$

where Signmoid() refers to the sigmoid activation function [53]. MLP is the parameterized multilayer perceptron function, which can integrate the channel information and output the weighted map. The spatial attention is acquired by the average pooling and max pooling operations along the channel. Then, the spatial attention map $\mathbf{M_s} \in \mathbb{R}^{1\times H\times W}$ is obtained by concatenation and convolution operations as given by

$$\mathbf{M_s}(\mathbf{F}) = \text{Signmoid}\left(f^{7\times 7}([AvgPool(\mathbf{F}); MaxPool(\mathbf{F})])\right) \tag{16}$$

where 7×7 denotes the 7×7 filter-sized convolution operation. The output feature map can be obtained by

$$\mathbf{F}_{\text{out}} = \mathbf{M}_s(\mathbf{M}_c(\mathbf{F}) \otimes \mathbf{F}) \otimes (\mathbf{M}_c(\mathbf{F}) \otimes \mathbf{F}) \tag{17}$$

The target-sea image has multi-scale image features. The target features decrease with the increase in the convolutional layers. The Feature Pyramid Network (FPN) structure is adopted to mix the shallow and high-level features [54–56]. The semantic features are transferred from the top to the bottom by up-sampling. The localization feature fusion is conducted by the down-sampling of the feature map. On the basis of the traditional FPN layers, an extra 160×160 feature layer is added to fuse the information of the smaller feature layers before down-sampling. Correspondingly, an extra detection head is introduced in the modified YOLOv7, as presented in Figure 8. The added prediction head is produced from the low-level and high-resolution feature maps to improve the capability of target feature extractions in the complex environment. The ELAN-H [57] is employed to increasingly improve the training and learning ability of the network, whose structure is given in Figure 12.

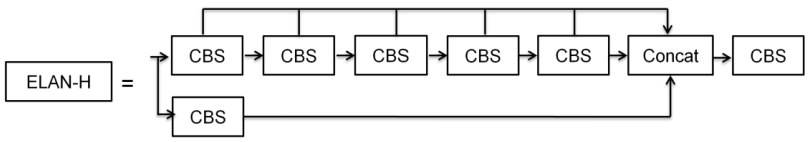


Figure 12. The ELAN-H structure.

3.2. SAR Imaging Training and Detection Process

The SAR imaging training and detection process is given in Figure 13.

Firstly, the SAR dataset is constructed. The SAR images are obtained by the SAR imaging model with the EM calculations of the scattering from the long and short cone targets, and sea surface samples are generated at different times in different sea states and radar incident conditions. Then, the SAR image data are framed and labeled by well-trained researchers with the help of the image annotation tool labelImg [58]. The SAR images are then preprocessed, resized into 640×640 and input into the backbone structure of the YOLOv7 network. The Stochastic Gradient Descent (SGD) optimizer is employed to train YOLOv7. The batch size for the optimization is set at 16. The training epoch is 300. There are 1024 SAR image samples used to establish the training datasets; 1/3 of the 1024 training samples are randomly selected to establish the validation dataset. The initial training was conducted using the pre-trained weights for YOLOv7 on the COCO dataset [59]. The

best weights can be obtained after the whole training process, which is used in the final SAR imaging detections. The numerical simulations are conducted on a WorkStation with Intel(R) Core(TM) i9-10885H CPU @ 2.40 GHz. The GPU version is NVIDIA GeForce GTX 1650Ti with Max-Q Design.

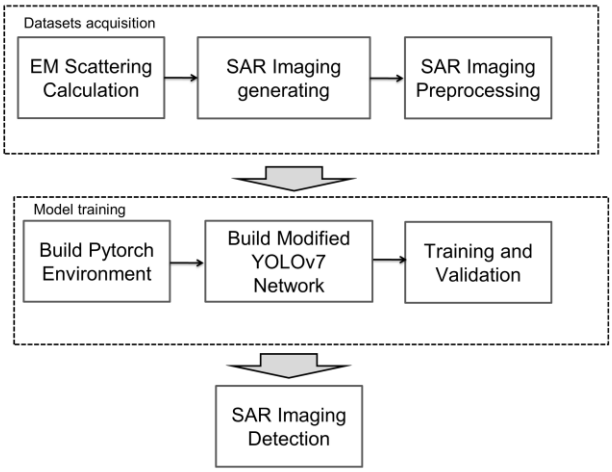


Figure 13. The synthetic aperture radar imaging training and detection process.

The precision rate, recall rate and Mean Average Precision (mAP) are employed to evaluate the training performance of the network according to the comparisons with the Intersection-Over-Union (IOU) value [60]. *TP* (True Positive) represents the number of correctly detected targets. The IOU of the predicted result is no less than 0.5 with the actual label. Otherwise, it is regarded as a False Positive (FP) detection. False Negative (FN) represents the number of missed targets. The precision rate and recall rate refer to the proportion of the correct detected targets in the total detected targets and the total number of targets, respectively, as given by

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{18}$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{19}$$

The Average Precision (AP) can be calculated by the integration of the average value of the highest precision under different recall conditions. The mean Average Precision (mAP) is the mean value of AP for each category, given by

$$\text{AP} = \int_0^1 P(r)dr \tag{20}$$

$$\text{mAP} = \frac{1}{M}\sum_{k=1}^{M} \text{AP}(k) \tag{21}$$

There are two kinds of commonly used mAPs, mAP@0.5 and mAP@0.5:0.95. The mAP@0.5 is the mAP with the IoU set at 0.5; mAP@0.5:0.95 has an average mAP at different IoU thresholds from 0.5 to 0.95 with a step size of 0.05. Figure 14 compares the evaluation metrics of the original YOLOv7 and the modified YOLOv7 when training on the SAR image datasets constructed by the numerical simulations. One can see from the comparison results that the precision and recall rates of the modified YOLOv7 are 97.46% and 90.08% after 300 rounds of training, which are 4.04% and 3.71% higher than those of the original YOLOv7 network. In addition, the value of the precision and recall for the modified YOLOv7 can stabilize with much fewer training times. The values of mAP@0.5 and mAP@0.5:0.95 of the

modified YOLOv7 model are 92.91% and 91.98%, which are 3.94% and 14.26% higher than those of the original YOLOv7 model.

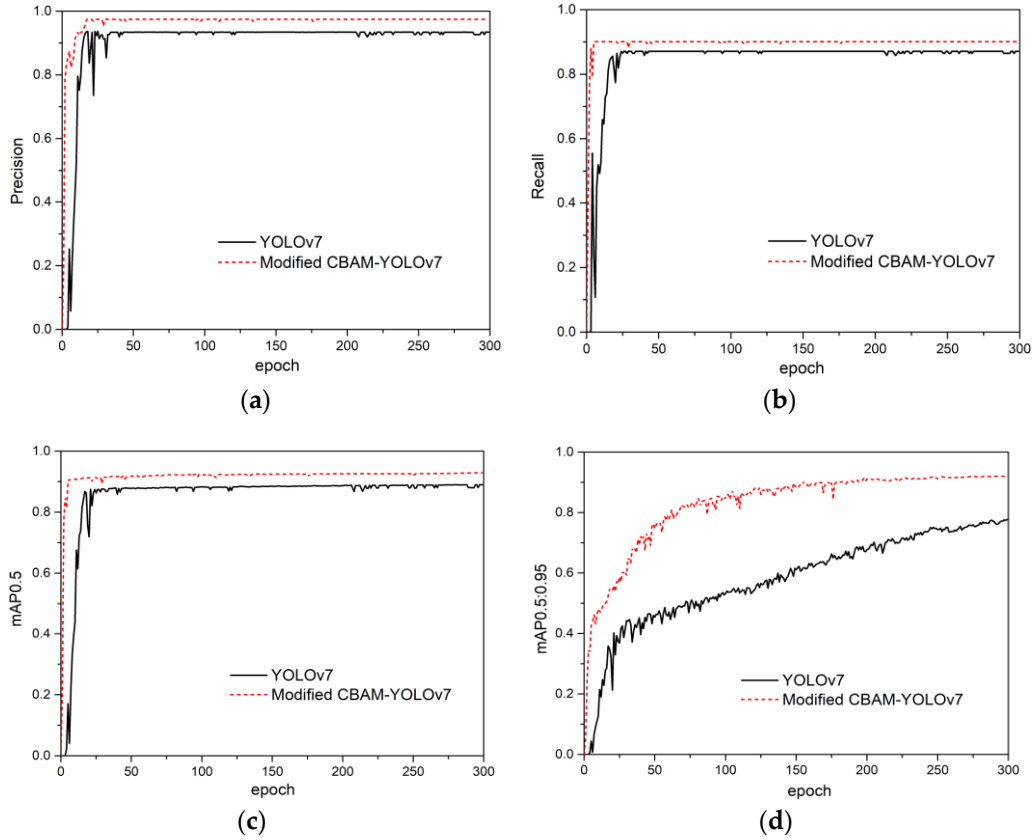


Figure 14. The comparisons of the evaluation metrics between the original You Only Look Once version 7 (YOLOv7) network and the modified YOLOv7 network. (**a**) Precision rate. (**b**) Recall rate. (**c**) mAP@0.5. (**d**) mAP@0.5:0.95.

Figure 15 shows the comparisons of the bounding box loss and object loss curves for the two networks. It can be seen that the loss value of the modified YOLOv7 network is lower and decreases more quickly than the original one.

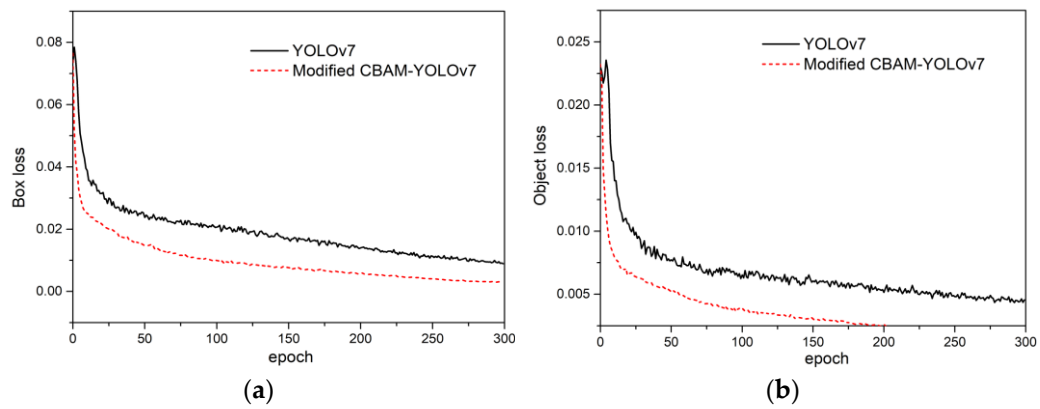
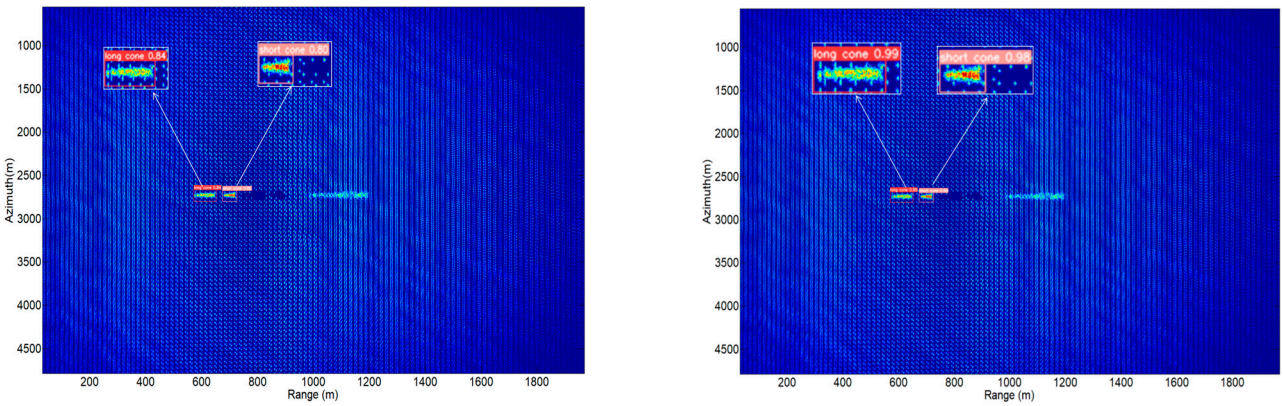


Figure 15. The loss value comparisons between the original and modified You Only Look Once version 7 (YOLOv7) networks. (**a**) Box loss (**b**) Object loss.

In the training process, the hyperparameters of the networks were fine-tuned as well. The default hyperparameters were used in the first training to establish a performance

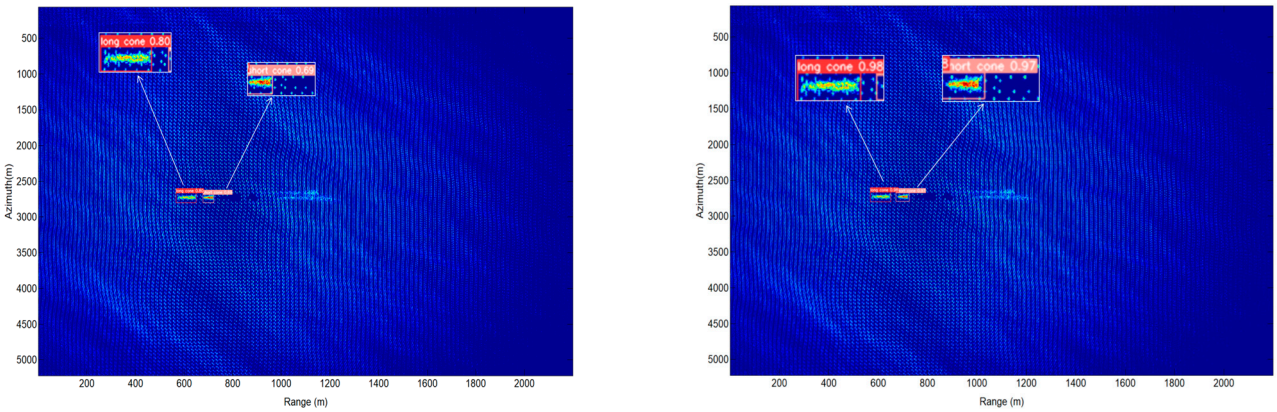
baseline. Then, the random search method was adopted to find the hyperparameters with better performance than the baseline. By doing so, the fine-tuned neural network and desired precision, recall and mAP values are achieved. The above comparisons show that the modified YOLOv7 can obtain a great performance improvement in the training. These results are consistent with the fact that modules, like CBAM and FPN in references [61–63], are effective in improving the performance of neural network models.

In the following simulations, the detections are employed on the SAR images of the multiple cone targets in the maritime environment with the two YOLOv7 networks. Figures 16 and 17 show the comparisons of the SAR detection results for a long-cone and short-cone target above the sea surface when U_{10} = 3 m/s and U_{10} = 10 m/s by means of the original YOLOv7 network and the modified YOLOv7 network. In the detected images, the targets need to be selected from the sea surface background and the multipath echo interference.



(**a**) (**b**)

Figure 16. The synthetic aperture radar detection results for the two-cone targets at the maritime scene when U_{10} = 3 m/s. (**a**) Original You Only Look Once version 7 (YOLOv7) network. (**b**) Modified YOLOv7 network.



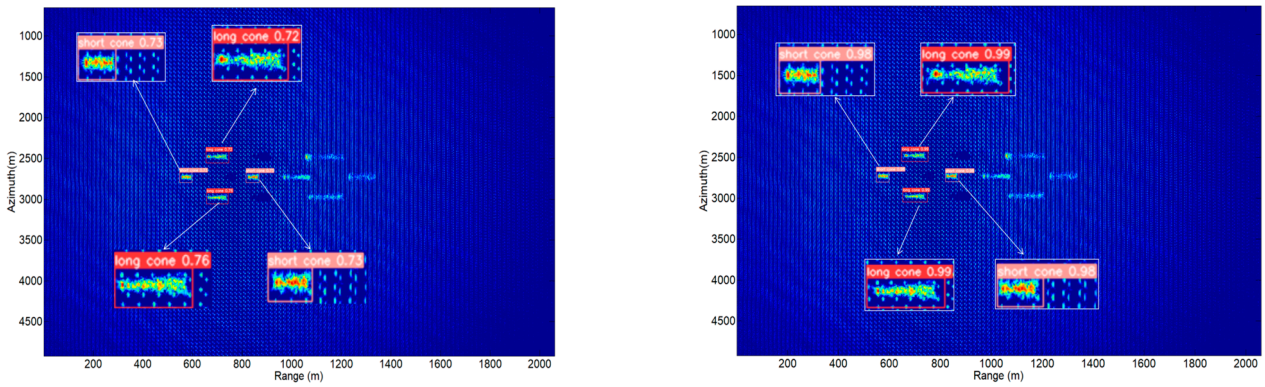
(**a**) (**b**)

Figure 17. The synthetic aperture radar detection results for the two-cone targets at the maritime scene when U_{10} = 10 m/s. (**a**) Original You Only Look Once version 7 (YOLOv7) network. (**b**) Modified YOLOv7 network.

Figures 16a and 17a show the detection results by the original YOLOv7 network. From the results, one can see that the detection performance of the original YOLOv7 network worsens in the higher sea state conditions. According to the aforementioned analysis, the backscattering intensities of the sea surface can be stronger at the higher sea state. In this condition, the target detection can be more severely affected by the sea surface

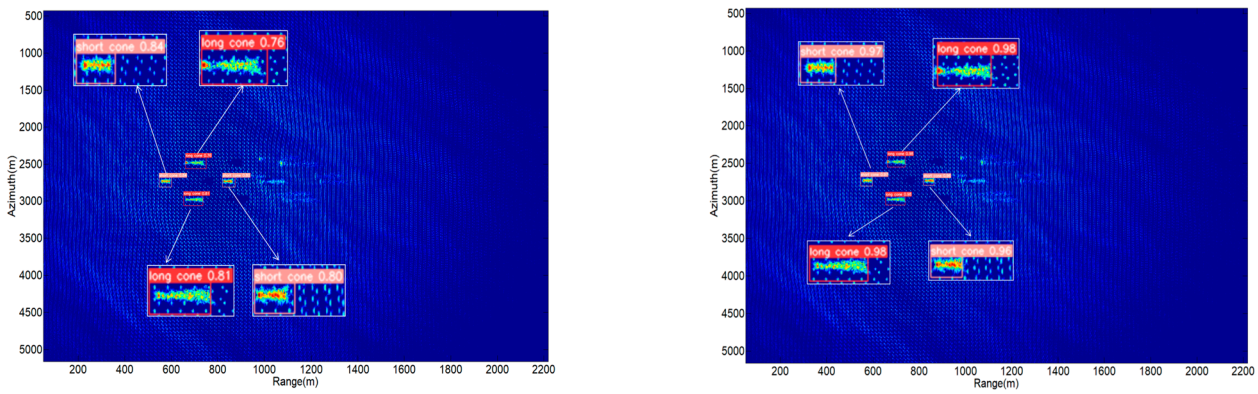
background. Additionally, it can be seen that the detections of the smaller cone target are more easily affected by the sea background. It means that smaller target detections in large maritime scenes are more difficult. Figures 16b and 17b show the detection results by the modified YOLOv7 network. One can also see that in comparison with the results of the original YOLOv7 network, the modified YOLOv7 network apparently improves the ability to extract the multi-scale features from the complex background. Thus, a better detection performance can be obtained.

Figures 18 and 19 show the detection results in a more complex environment. SAR detections are employed for the SAR images of the two long-cone targets and two short-cone targets in the maritime scene.



(**a**) (**b**)

Figure 18. The synthetic aperture radar detection results for the four cone targets at the maritime scene when U_{10} = 3 m/s. (**a**) Original You Only Look Once version 7 (YOLOv7) network, (**b**) modified YOLOv7 network.



(**a**) (**b**)

Figure 19. The synthetic aperture radar detection results for the four cone targets at the maritime scene when U_{10} = 10 m/s. (**a**) Original You Only Look Once version 7 (YOLOv7) network, (**b**) modified YOLOv7 network.

From the above results, one can see that the interference of the sea background and the artificial targets caused by the multipath interactions between the targets, as well as the target and sea surface, are more severe. The detection performance of the original YOLOv7 network suffered more loss. In Figure 18a, the target detection performance is mainly affected by the artificial target image caused by the multipath echo at the low sea state. In this condition, the backscattering intensity of the multipath echo is apparently very strong. The detections of the nearby targets can be affected. In Figure 19a, the target detection performance is mainly affected by the sea surface when the sea wind speed is strong. The reason is similar to the analysis of the results in Figures 16 and 17. In Figures 18b and 19b,

an apparent improvement in the detection performance of the modified YOLOv7 network has been seen in the same detection condition. It further proves that the modified YOLOv7 network has a stronger ability to extract accurate multi-scale features and detect the targets in the complex maritime environment.

4. Conclusions

SAR imaging and detection of maritime targets play a significant role in many fields, but a reliable and convenient way to obtain SAR datasets is still lacking. This paper proposes an SAR imaging model of the multiple targets in the maritime scene based on a hybrid electromagnetic approach. A facet SSA is derived to evaluate the local scattering with the Bragg phenomenon upon the electrically large sea environment with great accuracies and efficiencies, which is validated through a comparison with methods in previous studies and the measurement data [39,40]. The coupling interactions among the target and sea surface scatterers are calculated by a fast ray tracing process with the bidirectional tracing technique and the KD-tree acceleration method. The model is well applied in handling the SAR imagery simulations of the multiple targets in the maritime scene with different sea wind speeds. Numerical simulations show the distribution characteristics of the sea wave propagations as well as the multipath artificial images caused by the coupling interactions. The target image has a relatively distinct profile. The images caused by the multipath interference are obscure and weaker. The phenomena are also indicated in the references [46–49]. A modified YOLOv7 neural network with the SPPFCSPC module, CBAM, FPN structure and the extra detection head is developed for the SAR image target detections. The dataset is constructed by the simulated SAR images of the short-cone and long-cone targets above the sea surface in the various sea and target conditions. In the training process, the modified YOLOv7 network shows apparent better performance in the precision rate, recall rate and average precision. The modules used in the modified YOLO network were also used in other research fields and showed good performance [61–63], which further demonstrates the improvement effect of the modified YOLOv7 network. In references [61–63], the SAR image simulations provide diverse samples to train the neural network model. Thus, the overfitting to a specific dataset of the target features will have a limited influence on the generalization of the target detections in different maritime scenes and conditions. The modified YOLOv7 network also shows a stronger ability to extract the features from the multiple targets in the maritime scene at different sea states. However, it also needs to be admitted that the real sea surface structures and scattering mechanisms are very complex. It's hard for a model to wholly characterize the geometry features of various kinds of real sea surfaces. In this article, the non-Bragg scattering mechanisms from the sea surface structures in some complex sea states, such as breaking waves, whitecaps and internal waves, as indicated in references [33], are not included. In addition, the practical measurement experiments in the scenarios described in this article have not been carried out yet. These can be further researched and serve as improvements to the existing model to make it more robust in future studies.

Author Contributions: Conceptualization, P.P., C.T., W.F. and Q.W.; methodology, P.P. and Q.W.; validation, W.F.; formal analysis, P.P. and T.W.; investigation, Q.W. and P.P.; writing—original draft preparation, P.P.; writing—review and editing, C.T.; visualization, P.P., C.T. and Q.W.; funding acquisition, P.P. and W.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Nature Science Foundation of China (61901510) and the Program of The Youth Innovation Team of Shannxi University.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the datasets set up in the article are very large, which are not easy to be publicly uploaded.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Ulaby, F.T.; Fung, A.K. *Microwave Remote Sesing (Active and Passive)*; Addison Wesley: New York, NY, USA, 1982.
2. Martorella, M.; Pastina, D.; Berizzi, F.; Lombardo, P. Spaceborne Radar Imaging of Maritime Moving Targets with the Cosmo-SkyMed SAR System. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2797–2810. [CrossRef]
3. Long, Y.; Zhao, F.; Zheng, M.; Zhao, L.; He, W. An Unambiguous Imaging Method of Moving Target for Maritime Scenes With Spaceborne High-Resolution and Wide-Swath SAR. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–19. [CrossRef]
4. Cao, T.V.; Luckett, C.; Williams, J.; Cooke, T.; Yip, B.; Rajagopalan, A.; Wong, S. SARFish: Space-Based Maritime Surveillance Using Complex Synthetic Aperture Radar Imagery. In Proceedings of the 2022 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, Australia, 30 November–2 December 2022; pp. 1–8.
5. Zhang, Y.; Yin, Y.; Shao, Z. An Enhanced Target Detection Algorithm for Maritime Search and Rescue Based on Aerial Images. *Remote Sens.* **2023**, *15*, 4818. [CrossRef]
6. Lin, Z.; Ji, K.; Leng, X.; Kuang, G. Squeeze and Excitation Rank Faster R-CNN for Ship Detection in SAR Images. *IEEE Trans. Geosci. Remote Sens Lett.* **2019**, *16*, 751–755. [CrossRef]
7. Patel, K.; Bhatt, C.; Mazzeo, P.L. Deep learning-based automatic detection of ships: An experimental study using satellite images. *J. Imaging* **2022**, *8*, 182. [CrossRef]
8. Jiang, S.; Hong, Z. Unexpected Dynamic Obstacle Monocular Detection in the Driver View. *IEEE Intelli. Trans. Syst. Mag.* **2023**, *15*, 68–81. [CrossRef]
9. Jiang, S.; Tarabalka, Y.; Yao, W.; Hong, Z.; Feng, G. Space-to-speed architecture supporting acceleration on VHR image processing. *ISPRS J. Photogramm. Remote Sens.* **2023**, *198*, 30–44. [CrossRef]
10. Yu, X.; Salimpour, S.; Queralta, J.P.; Westerlund, T. General-Purpose Deep Learning Detection and Segmentation Models for Images from a Lidar-Based Camera Sensor. *Sensors* **2023**, *23*, 2936. [CrossRef]
11. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.
12. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
13. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. *CVF* **2021**, *2108*, 11539.
14. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6:A Single-Stage Object Detection Framework for Industrial Applications. *arXiv* **2022**, arXiv:2209.02976.
15. Wang, C.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.
16. Zhu, Q.; Ma KWang, Z.; Shi, P. YOLOv7-CSAW for maritime target detection. *Front. Neurorobot.* **2023**, *17*, 1210470. [CrossRef] [PubMed]
17. Zhao, H.; Zhang, H.; Zhao, Y. YOLOv7-sea: Object Detection of Maritime UAV Images based on Improved YOLOv7. In Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW), Waikoloa, HI, USA, 3–7 January 2023; pp. 233–238.
18. Guan, B.; Zhang, J.F.; Zhou, X.Y.; Cui, T.J. Electromagnetic Scattering From Objects Above a Rough Surface Using the Method of Moments With Half-Space Green's Function. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3399–3405. [CrossRef]
19. Kuang, L.; Jin, Y.Q. Bistatic scattering from a three-dimensional object over a randomly rough surface using the FDTD algorithm. *IEEE Trans. Antennas Propag.* **2007**, *55*, 2302–2312. [CrossRef]
20. Liu, P.; Jin, Y.-Q. Numerical simulation of bistatic scattering from a target at low altitude above rough sea surface under an EM-wave incidence at low grazing angle by using the finite element method. *IEEE Trans. Antennas Propag.* **2004**, *52*, 1205–1210. [CrossRef]
21. Johnson, J.T. A Numerical Study of Scattering from an Object above a Rough Surface. *IEEE Trans. Antennas Propag.* **2002**, *50*, 1361–1367. [CrossRef]
22. Li, X.F.; Xie, Y.J.; Yang, R. High-frequency method analysis on scattering from homogenous dielectric objects with electrically large size in half space. *Prog. Electromagn. Res. B* **2008**, *1*, 177–188. [CrossRef]
23. Franceschetti, G.; Migliaccio, M.; Riccio, D.; Schirinzi, G. SARAS: A synthetic aperture radar (SAR) raw signal simulator. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 110–123. [CrossRef]
24. Franceschetti, G.; Migliaccio, M.; Riccio, D. On ocean SAR raw signal simulation. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 84–100. [CrossRef]
25. Zhang, M.; Chen, H.; Yin, H.C. Facet-based investigation on EM scattering from electrically large sea surface with two-scale profiles: Theoretical model. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1967–1975. [CrossRef]
26. Zhao, Y.; Yuan, X.F.; Zhang, M.; Chen, H. Radar scattering from the composite ship-ocean scene: Facet-based asymptotical model and specular reflection weighted model. *IEEE Trans. Antennas Propag.* **2014**, *62*, 4810–4815. [CrossRef]
27. Tsang, L.; Kong, J.A.; Ding, K.-H.; Ao, C.O. *Scattering of Electromagnetic Waves, Numerical Simulations*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
28. Elfouhaily, T.; Chapron, B.; Katsaros, K.; Vandemark, D. A unified directional spectrum for long and short wind-driven waves. *J. Geophys. Res.* **1997**, *102*, 15781–15796. [CrossRef]

29. West, J.C. Correlation of Bragg scattering from the sea surface at different polarizations. *Waves Random Media* **2005**, *15*, 395–403. [CrossRef]
30. Plant, W.J. Microwave Sea Return at Moderate to High Incidence Angles. *Waves Random Media* **2003**, *13*, 339–354. [CrossRef]
31. Voronovich, G. *Wave Scattering from Rough Surfaces*; Springer: Berlin/Heidelberg, Germany, 1994.
32. Voronovich, G. Small-slope approximation for electromagnetic wave scattering at a rough interface of two dielectric half-spaces. *Waves Random Media* **1994**, *4*, 337–367. [CrossRef]
33. Voronovich, G.; Zavorotni, V.U. Theoretical model for scattering of radar signals in Ku- and C-bands from a rough sea surface with breaking waves. *Waves Random Media* **2001**, *11*, 247–269. [CrossRef]
34. Wei, P.B.; Zhang, M.; Jiang, W.Q.; Nie, D. A New Model for Sand-Ripple Scattering Based on SSA Method and Practical Ripple Profiles. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2450–2459. [CrossRef]
35. Li, J.X.; Zhang, M.; Zhao, Y.; Jiang, W.Q. Efficient Numerical Full-Polarized Facet-Based Model for EM Scattering from Rough Sea Surface within a Wide Frequency Range. *Remote. Sens.* **2019**, *11*, 75. [CrossRef]
36. Peng, P.; Guo, L.X. A Facet-Based Simulation of the Multipath Effect on the EM Scattering and Doppler Spectrum of a Low-Flying Target at Maritime Scene. *IEEE Trans. Geosci. Remote Sens Lett.* **2018**, *15*, 1495–1499. [CrossRef]
37. Thorsos, E. The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum. *J. Acoust. Soc. Am.* **1988**, *83*, 79. [CrossRef]
38. Klein, L.A. An improved model for the dielectric constant of sea water at microwave frequencies. *IEEE Trans. Antennas Propag.* **1977**, *25*, 104–111. [CrossRef]
39. Schroeder, L.; 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]
40. Arnold-Bos, A.; Khenchaf, A.; Martin, A. Bistatic radar imaging of the marine environment—Part I: Theoretical background. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3372–3383. [CrossRef]
41. Gordon, W.B. High Frequency Approximations to the Physical Optics Scattering Integral. *IEEE Trans. Antennas Propag.* **1994**, *42*, 427–432. [CrossRef]
42. Fan, T.Q.; Guo, L.X.; Lv, B.; Liu, W. An Improved Backward SBR-PO/PTD Hybrid Method for the Backward Scattering Prediction of an Electrically Large Target. *IEEE Trans. Antennas Propag. Lett.* **2015**, *15*, 512–515. [CrossRef]
43. Tao, Y.-B.; Lin, H.; Bao, H.J. KD-tree based fast ray tracing for RCS prediction. *Prog. Electromagn. Res.* **2008**, *81*, 329–341. [CrossRef]
44. Michaeli, A. Equivalent edge currents for arbitrary aspects of observation. *IEEE Trans. Antennas Propag.* **1984**, *23*, 252–258. [CrossRef]
45. Alpers, W.R.; Bruening, C. On the relative importance of motion related contributions to the SAR imaging mechanism of ocean surface waves. *IEEE Trans. Geosci. Remote Sens.* **1986**, *GE-24*, 873–885. [CrossRef]
46. Gennarelli, G.; Soldovieri, F. Multipath ghosts in radar imaging: Physical insight and mitigation strategies. *IEEE Sel. Top. Appl. Earth Obs.* **2015**, *8*, 1078–1086. [CrossRef]
47. Duan, C.W.; Hu, W.D.; Du, X.Y. Probability model of multipath delays in radar echoes of scattering centres above ocean surface. *Electron. Lett.* **2012**, *48*, 177–179. [CrossRef]
48. Huo, W.; Huang, Y.; Pei, J.; Zhang, Q.; Gu, Q.; Yang, J. Ship Detection from Ocean SAR Image Based on Local Contrast Variance Weighted Information Entropy. *Sensors* **2018**, *18*, 1196. [CrossRef] [PubMed]
49. Yang, W.; Qi, C.H. Multipath scattering simulation and SAR analysis for a composite model. *Microw. Opt. Technol. Lett.* **2015**, *57*, 1689–1695. [CrossRef]
50. Piao, Y.; Jiang, Y.; Zhang, M.; Wang, J.; Lu, H. PANet: Patch-Aware Network for Light Field Salient Object Detection. *IEEE Trans. Cybern.* **2021**, *53*, 379–391. [CrossRef] [PubMed]
51. Woo, S.; Park, J.; Lee, J.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the 15th European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 3–19.
52. Jiang, M.; Yin, S.L. Facial expression recognition based on convolutional block attention module and multi-feature fusion. *Int. J. Comput. Vis. Robot.* **2023**, *13*, 21–37. [CrossRef]
53. Xia, Y.; Xu, X.; Pu, F. PCBA-Net: Pyramidal Convolutional Block Attention Network for Synthetic Aperture Radar Image Change Detection. *Remote Sens.* **2022**, *14*, 5762. [CrossRef]
54. Zhang, T.; Zhang, X.; Ke, X. Quad-FPN: A Novel Quad Feature Pyramid Network for SAR Ship Detection. *Remote Sens.* **2021**, *13*, 2771. [CrossRef]
55. Sun, Z.; Meng, C.; Cheng, J.; Zhang, Z.; Chang, S. A Multi-Scale Feature Pyramid Network for Detection and Instance Segmentation of Marine Ships in SAR Images. *Remote Sens.* **2022**, *14*, 6312. [CrossRef]
56. Chen, Z.; Liu, C.; Filaretov, V.F.; Yukhimets, D.A. Multi-Scale Ship Detection Algorithm Based on YOLOv7 for Complex Scene SAR Images. *Remote Sens.* **2023**, *15*, 2071. [CrossRef]
57. Hu, J.; Chuang, F.; Wang, Z.; Ruan, J.; Wu, S.Y. Fruit detection and counting in apple orchards based on improved Yolov7 and multi-Object tracking methods. *Sensors* **2023**, *23*, 5903. [CrossRef]
58. Zhang, T.; Zhang, X.; Shi, J.; Wei, S. Depth wise Separable Convolution Neural Network for High-Speed SAR Ship Detection. *Remote Sens.* **2019**, *11*, 2483. [CrossRef]
59. Shunjun, W.; Xiangfeng, Z.; Qizhe, Q.; Mou, W.; Hao, S.; Jun, S. HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation. *IEEE Access* **2020**, *8*, 120234–120254.

60. Amarasingam, N.; Gonzalez, F.; Salgadoe, A.S.A.; Sandino, J.; Powell, K. Detection of White Leaf Disease in Sugarcane Crops Using UAV-Derived RGB Imagery with Existing Deep Learning Models. *Remote Sens.* **2022**, *14*, 6137. [CrossRef]
61. Yi, W.; Wang, B. Research on Underwater Small Target Detection Algorithm Based on Improved YOLOv7. *IEEE Access* **2023**, *11*, 66818–66827. [CrossRef]
62. Yang, H.; Liu, Y.; Wang, S.; Qu, H.; Li, N.; Wu, J.; Yan, Y.; Zhang, H.; Wang, J.; Qiu, J. Improved Apple Fruit Target Recognition Method Based on YOLOv7 Model. *Agriculture* **2023**, *13*, 1278. [CrossRef]
63. Jiang, K.; Xie, T.; Yan, R.; Wen, X.; Li, D.; Jiang, H.; Jiang, N.; Feng, L.; Duan, X.; Wang, J. An Attention Mechanism-Improved YOLOv7 Object Detection Algorithm for Hemp Duck Count Estimation. *Agriculture* **2022**, *12*, 1659. [CrossRef]