Improving data-driven estimation of significant wave height through preliminary training on synthetic X-band radar sea clutter imagery

Vadim Rezvov[1,2]*, Mikhail Krinitskiy[1,2], Alexander Gavrikov[2], Viktor Golikov[1,2], Mikhail Borisov[1,2], Alexander Suslov[2] and Natalia Tilinina[2]

[1]Machine Learning for Earth Sciences Laboratory, Moscow Institute of Physics and Technology, Dolgoprudny, Russia, [2]Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia

X-band marine radar captures the signal reflected from the sea surface. Theoretical studies indicate that the initial unfiltered signal contains meaningful information about wind wave parameters. Traditional methods of significant wave height (SWH) estimation rely on physical laws describing signal reflection from rough surfaces. However, recent studies suggest the feasibility of employing artificial neural networks (ANNs) for SWH approximation. Both classical and ANN based approaches necessitate costly *in situ* data. In this study, as a viable alternative, we propose generating synthetic radar images with specified wave parameters using Fourier-based approach and Pierson–Moskowitz wave spectrum. We generate synthetic images and use them for unsupervised learning approach to train a convolutional component of the reconstruction ANN. After that, we train the regression ANN based on the previous convolutional part to obtain SWH back from the synthetic images. Then, we apply preliminary trained weights for the regression model to train SWH approximation on the dataset of real sea clutter images. In this study, we demonstrate the increase in SWH estimation accuracy from radar images with preliminary training on synthetic data.

1 Introduction

X-band marine radars play an essential role in ship navigation and safety through obstacle detection (Huang et al., 2017). Besides their primary function, raw radar images capturing sea clutter contain substantive information (Young et al., 1985). Analysis of the spatial distribution of the reflected signal facilitates the derivation of parameters associated with wind waves and swell. Additionally, these radar images allow estimating such vital sea surface characteristics as significant wave height (SWH). Sea surface examination is strongly linked to ocean-atmosphere interaction and, consequently, to long-term climate reconstructions, such as Global Atlas of Ocean Waves[1].

Given that the intensities of sea clutter radar images do not correspond directly to the ocean surface elevation on a one-to-one scale, a problem of accurate backscatter signal modeling is of great scientific interest. Such an elaborated model would give the opportunity to study sea surface with high temporal and spatial resolution (Nieto-Borge et al., 2004).

The classical approaches of SWH estimation include Fourier analysis and a linear dispersion relationship to recognize wave signals within the temporal series of radar data. One of the examples is the signal-to-noise ratio (SNR)-analysis-based method proposed by (Nieto-Borge et al., 2008). The method necessitates the incorporation of modulation transfer functions and calibration coefficients specific to individual radar antennas (Nieto-Borge et al., 1999), thereby constraining its generalizability. Despite this limitation, the classical methodology remains extensively applied for real-time estimation of ocean wave parameters, processing the back-scatter spectrum derived from radar images (Tilinina et al., 2022).

In addition to classical approaches, radar data processing includes methodologies that are potentially faster and highly independent of radar antenna specifications. Notably, certain publications show superior SWH estimation quality through contemporary artificial intelligence (AI)-based techniques (Vicen-Bueno et al., 2012), in comparison to classical methodologies. Within machine-learning paradigm, the functional relationships, such as artificial neural networks (ANNs), between radar images and corresponding SWH values are approximated through training on extensive datasets (Park et al., 2020). For this group of methods, the efficacy of regression quality can potentially be correlated with the size and the distribution of the training dataset (Long et al., 2017).

The authors of (Ludeno and Serafino, 2019) address the challenge of limited data by employing an algorithm for generating synthetic radar images with arbitrary SWH value. This methodology uses technique of generating realistic ocean surfaces, as detailed in (Mastin et al., 1987). While the methodology in (Ludeno and Serafino, 2019) does not incorporate image synthesis within a machine-learning framework, the ANN model and synthetic datasets can potentially enhance the overall SWH estimation from real X-band radar images.

Convolutional neural networks (CNNs) show their efficacy in image recognition and feature extraction across various scientific domains, including geosciences. The notable instances are illustrated in (Choi et al., 2020; Kim et al., 2021), where CNNs are employed for real-time SWH estimation utilizing X-band radar images. The temporal convolutional network (TCN) is a variant of CNN architecture proposed in (Bai et al., 2018) for sequence processing. Based on it, (Huang et al., 2021) propose a TCN-based model to retrieve SWH from X-band marine radar images.

In summary, the recent progress in ANN and the emergence of modern machine-learning techniques have substantially diversified radar image processing methods. Nonetheless, the findings in this research domain present conflicting outcomes, and numerous questions persist. Consequently, evaluating current methodologies for radar image generation and SWH estimation remains a subject of ongoing scientific interest.

In this paper, we demonstrate the utilization of a preliminarily trained CNN for SWH estimation from real marine X-band radar images of sea clutter. Initially, we elaborate the methodology of generating a synthetic dataset of radar images. We train a model that reconstructs synthetic images in unsupervised learning approach. After that, we use preliminarily trained part of the reconstruction model to construct and train a regression model that approximates SWH from the synthetic image. Further, the pre-trained regression architecture is trained on a dataset of real radar images. We then compare the results of this pre-trained model with the quality of the simple SWH regression model.

In this study, we do not aim to develop a model that would accurately reproduce the radar signal from a given sea surface. Our goal is to increase the accuracy of the ANN-based model approximating SWH value from the radar image.

The paper is organized as follows. In Section 2, we provide the details of the real radar image dataset, the methodology of the artificial radar image synthesis. and the architectures of the applied ANNs, quality metrics and training and evaluation procedures. In Section 3, we provide the results of the elaborated model training. In Section 4, we analyze how SWH estimation quality increases with pre-training on synthetic dataset. Concluding remarks are made in Section 5.

2 Materials and methods

2.1 Initial data

For this research, we adopt the data collection methodology outlined in (Tilinina et al., 2022). Our dataset comprises samples obtained during four research expeditions conducted in the Atlantic and Arctic oceans. These expeditions were undertaken by the Shirshov Institute of Oceanology of the Russian Academy of Sciences within the governmental program of regular ocean observations.

The routes of the expeditions encompass points with SeaVision radar images and/or Spotter buoys measurements[2]. Comprehensive details about the research expeditions can be found in Table 1.

---

1 Gulev, S. (2014) *Global Atlas of Ocean Waves*. Available at: http://www.sail.msk.ru/atlas/ (Accessed: 21 December 2023).

2 Sofar Ocean (2023) *Spotter Buoy by Sofar*. Available at: https://www.sofarocean.com/products/spotter (Accessed: 22 December 2023).

TABLE 1 Summary of undertaken scientific research marine cruises equipped with SeaVision and Spotter buoy for collecting real dataset of this study and summary of the synthetic dataset.

Expedition	Departure	Arrival	Spotter buoy stations	SeaVision stations	Number of measurements
ASV50	Kaliningrad, Russia, 7 Aug 2020	Arkhangelsk, Russia, 13 Sep 2020	24	157	27 468
AI57	Kaliningrad, Russia, 25 Jun 2021	Arkhangelsk, Russia, 21 Jul 2021	12	76	15 617
AI58	Kaliningrad, Russia, 10 Aug 2021	Kaliningrad, Russia, 9 Sep 2021	16	55	21 978
AI63	Arkhangelsk, Russia, 29 Sep 2022	Arkhangelsk, Russia, 7 Dec 2022	30	209	30 304
Real data			82	497	95 367
Synthetic data			60 000		

Here, abbreviation ASV stands for the research vessel Akademik Sergey Vavilov, and AI stands for the research vessel Akademik Ioffe.

For clarity, a "station" refers to an instance where the ship is adrift, thereby creating optimal conditions for data acquisition. Table 1 outlines the departure and arrival ports, as well as the number of stations where data from both the Spotter buoy and the SeaVision radar were collected.

The selection of locations for sea wave observation was based on local weather conditions and temporal constraints. In total, there around 95 000 SeaVision images and 62 hours of simultaneous Spotter buoy *in situ* measurements were collected. It is noteworthy (Table 1) that the number of Spotter buoy stations differs from the number of SeaVision stations. This discrepancy arises because the engineering support required for operations at the buoy stations is significantly more complex compared to the SeaVision operations, resulting in some SeaVision stations lacking corresponding buoy measurements. In this study, however, we focus exclusively on the subset of data comprising synchronous observations from both the Spotter buoy and the SeaVision radar, as summarized in Table 1. The summary of the station data is in Table 2.

Spotter wave buoy measures vertical and horizontal displacements of the ocean surface with 2.5 Hz sampling frequency providing highly accurate measurements of wind wave characteristics (Raghukumar et al., 2019) and collecting the training dataset of SWH.

Random sea-surface elevation can be modeled as a stationary Gaussian process, wherein the statistical properties are defined by spectral moments. For instance, the surface elevation variance corresponds to the zeroth moment. To determine the wave spectrum, Welch's method is employed, utilizing a 600-second Hann window. From the zeroth spectral moment, SWH is calculated.

To compute the current value of SWH, the ten minutes of buoy data are analyzed, comprising approximately 1546 records, so that every time interval corresponds in time to the radar image coinciding with the center of the interval. The spectrum is calculated using a Hann window of equivalent width. This window is then incrementally shifted forward by one record at a time until the analysis encompasses the most recent ten minutes of data.

SeaVision radar creates one sea clutter image every two seconds. Research vessels ASV and AI are equipped with JRC JMA-9110-6XA and JMA-9122-6XA X-band radars, respectively. These radars have the same principal characteristics, such as the frequency of 9.41 GHz, a 6-foot antenna with a horizontal directional resolution of 1.2°, and almost identical shortest pulse lengths of 0.08 μs and 0.07 μs, respectively. Images are stored externally along with the GPS coordinates, speed over the ground (SOG), and course over the ground (COG). These files are later combined and converted into NetCDF format for post-processing. Radar antenna characteristics with the detailed description of the image collection process can be found in (Tilinina et al., 2022).

The center of the resulting sea clutter image coincides with the location of the radar antenna. Raw images with the spatial resolution of 1.875 m cover >7 km radius around the ship. The image exhibits a signal with diverse structures and intensities attributed to local wind direction, vessel rotation, and reflections of electromagnetic signal from the rough ocean surface (Lyzenga and Walker, 2015).

There is a "blind" zone near the center due to signal reflection from the vessel. To mitigate this effect, we exclude the part of the image within a 300 m radius around the ship. After visual examination of radar images, we observe that the reflected signal becomes less distinct with increasing distance from the radar antenna. As a result, we choose a radius of ≈2000 m to ensure the highest possible significant variability observed in radar data. Consequently, first, we restrict our analysis to the 300–2000 m range for further processing.

The important step of the data pre-processing is a choice of the optimal 180° sector containing the most informative wave signal. The technique used in this research extracts the most contrasting area from the entire radar image to provide the most clearly identified wind waves. For every station, the sector with the largest temporal

TABLE 2 Summary of stations.

Station	Number of measurements	Expedition
2763	571	ASV50
2771	1336	ASV50
2777	1007	ASV50
2782	733	ASV50
2787	596	ASV50
2792	686	ASV50
2797	711	ASV50
2803	853	ASV50
2809	799	ASV50
2821	866	ASV50
2833	832	ASV50
2841	833	ASV50
2849	859	ASV50
2856	912	ASV50
2863	988	ASV50
2868	1594	ASV50
2881	1691	ASV50
2885	1664	ASV50
2899	1774	ASV50
2901	1753	ASV50
2903	1497	ASV50
2913	1549	ASV50
2928	1632	ASV50
2937	1732	ASV50
Total ASV50	27 468	
3831	1095	AI57
3836	829	AI57
3841	3195	AI57
3847	1144	AI57
3853	1045	AI57
3858	1168	AI57
3863	1145	AI57
3870	616	AI57
3875	1271	AI57
3880	1344	AI57
3884	1374	AI57
3899	1391	AI57
Total AI57	15 617	

(Continued)

TABLE 2 Continued

Station	Number of measurements	Expedition
3911	1362	AI58
3929	1484	AI58
3930	1001	AI58
3939	1335	AI58
3946	1042	AI58
3956	1439	AI58
3972	1463	AI58
3982	1438	AI58
3990	1510	AI58
3997	1267	AI58
4013	1436	AI58
4020	1412	AI58
4025	1442	AI58
4029	1439	AI58
4031	1469	AI58
4040	1439	AI58
Total AI58	21 978	
4338	1198	AI63
4339	1100	AI63
4340	1319	AI63
4341	1487	AI63
4342	1441	AI63
4343	1420	AI63
4344	1374	AI63
4345	1387	AI63
4346	1371	AI63
4347	1414	AI63
4348	745	AI63
4349	737	AI63
4350	769	AI63
4351	730	AI63
4352	774	AI63
4353	783	AI63
4354	733	AI63
4355	736	AI63
4356	733	AI63
4357	762	AI63
4358	859	AI63

(Continued)

TABLE 2 Continued

Station	Number of measurements	Expedition
4361	937	AI63
4364	832	AI63
4366	744	AI63
4368	762	AI63
4371	774	AI63
4374	755	AI63
4378	730	AI63
4392	2137	AI63
4410	761	AI63
Total AI63	30 304	
Total	95 367	

standard deviation is the optimal sector. In this study, considering spatial resolution of 1.875 m, we choose the outer radius of the meaningful area as ≈1920 m (1024 pixels). Hence, we work with pre-processed 1024×2048 pixels images in contradistinction to 384×384 pixels images in (Tilinina et al., 2022).

Then, we linearly transform the values of the back-scattered radar signal image so that for every specific image, the minimum value is 0 and the maximum is 255. For further purposes, we use only masked real 1024×2048 pixels images. Namely, we replace the pixels outside 300–1920 m radius area with zeros. The example of the pre-processed radar image is shown in Figure 1.

2.2 Synthesis of realistic sea surface

We develop the methodology of generating synthetic radar images (Ludeno and Serafino, 2019) based on realistic ocean scenes (Mastin et al., 1987). The authors of (Mastin et al., 1987) elaborated a technique that presents fully developed sea with an empirical modified Pierson–Moskowitz sea power spectrum (Pierson and Moskowitz, 1964).

As proposed in (Mastin et al., 1987), we first generate a 2088×2088 pixels array with the noise values uniformly distributed from −0.5 to 0.5. We choose the image size that is bigger than 2048 pixels to further exclude possible edge effects. The result of this step is a white-noise image (see Figure 2A).

The spatial resolution is 1.875 m, equal to that of the real radar images. This parameter allows us to perform a two-dimensional forward Fourier transform of the white-noise image to generate an array of complex numbers. The magnitude of these complex Fourier components is shown in Figure 2B. The coordinate axes in Figure 2B are orthogonal components k_x and k_y [m^{-1}] of spatial frequency vector.

The methodology of creating a synthetic sea surface is based on the following theory. In this research, we consider a simple case of fully developed wind waves, characterized by a wave power spectrum constant in time. Under these assumption, W. Pierson and L. Moskowitz (Pierson and Moskowitz, 1964) empirically approximated the mathematical form of the downwind power spectrum $F_{PM}(f)$:

$$F_{PM}(f) = \frac{\alpha g^2}{(2\pi)^4 f^5} \exp\left[-\frac{5}{4}\left(\frac{f_m}{f}\right)^4\right], \quad (1)$$

where f is the temporal frequency [Hz], f_m is the peak temporal frequency [Hz], $\alpha = 0.0081$ is the Phillips constant, and g is the mean gravitational acceleration [m/s^2].

The form of the fully-developed wind spectrum (1) depends solely on a single parameter f_m that indicates the frequency of the spectrum maximum. It was also discovered in (Pierson and Moskowitz, 1964) that f_m is a function of 10 m surface wind speed U_{10} with the constant mean gravitational acceleration g:

$$f_m \approx 0.13 \frac{g}{U_{10}}. \quad (2)$$

SWH is usually defined as an average measurement of the largest 33% of waves[3]: $SWH = \frac{1}{N/3}\sum_{j=1}^{N/3} H_j$, where N is the total number of measured waves, H_j is a height of the j-th wave from the largest 33%. For a given power spectrum $S(\omega)$, we can equivalently calculate significant wave height[4]: $SWH = 4\sqrt{\int_0^{+\infty} S(\omega)\ d\omega}$. Thus, from (2), for Pierson–Moskowitz spectrum (1), we obtain:

$$SWH \approx 0.22 \frac{U_{10}^2}{g}. \quad (3)$$

Hence, we completely determine the shape of the one-dimensional downwind fully developed wave spectrum (1) with either U_{10} or SWH (3).

The two-dimensional wave power spectrum $F(f, \varphi)$ was proposed in (Hasselmann et al., 1980) as an extension of (1) taking into account wind direction:

$$F(f, \varphi) = F_{PM}(f)D(f, \varphi), \quad (4)$$

$$D(f, \varphi) = \frac{\Gamma^2(p+1)}{2^{1-2p}\pi\Gamma(2p+1)}\left[\cos\left(\frac{\varphi}{2}\right)\right]^{2p}, \quad (5)$$

where $F_{PM}(f)$ is one-dimensional Pierson-Moskowitz spectrum from (1), $D(f,\varphi)$ is a normalized directional multiplier at angle φ from the downwind direction, and Γ is the gamma function. The empirical parameters in (5) are defined by (Mastin et al., 1987): $p = 9.77\left(\frac{f}{f_m}\right)^\mu$ and μ is equal to 4.06 for $f \le f_m$ and –2.34 for $f > f_m$.

---

3 National Weather Service (2023) *Significant Wave Height*. Available at: https://www.weather.gov/mfl/waves (Accessed: 22 December 2023).

4 National Data Buoy Center (2023) *How are significant wave height, dominant period, average period, and wave steepness calculated?* Available at: https://www.ndbc.noaa.gov/faq/wavecalc.shtml (Accessed: 22 December 2023).
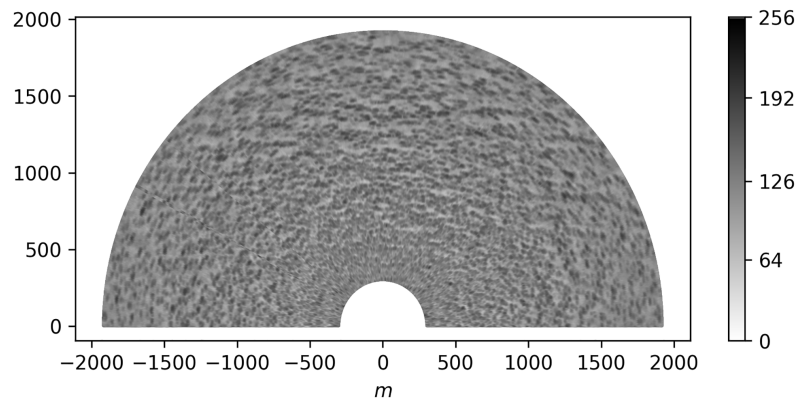
FIGURE 1
The example of the pre-processed radar image from AI57 marine research mission of Shirshov Institute of Oceanology of Russian Academy of Sciences. The position of a radar antenna is in (0,0) point.

The example of the normalized filter $F_{PM}(f)D(f,\varphi)$ for $U_{10} = 15$ m/s for the frequency domain of Figure 2B is shown in Figure 3A. To transform the white-noise spectrum into the two-dimensional Pierson–Moskowitz spectrum (4), we multiply the magnitudes of the Fourier components of the initial white-noise image by (1) and (5).

The resulting spectrum creates a narrower profile near $f_m(2)$ in the downwind direction, forms a bimodal spectrum shape for $\varphi \approx 90°$ from the downwind direction, and suppresses the long-crested peak frequency components, while retaining non-peak frequencies (Mastin et al., 1987). The filtered magnitudes of the white-noise Fourier components from Figure 2B are shown in Figure 3B.

Subsequently, we combine the filtered magnitudes with the original phase of the white-noise complex Fourier components, and apply the inverse Fourier transform, as described in (Mastin et al., 1987). The result of this procedure is a 2088×2088 pixels array of complex numbers. The desired synthetic sea surface is a real part of this array. We illustrate the full-domain realistic sea surface and its central part in Figure 4.

2.3 Synthesis of radar images

Here we describe transformation of the synthetic sea surfaces obtained in 2.2 into synthetic X-band radar images.

As (Nieto-Borge et al., 2004) claim, sea clutter primarily arises due to backscatter through Bragg resonance with ocean waves that have wavelengths similar to the transmitted electromagnetic waves. Longer waves become visible in radar images because they modulate the sea clutter signals with three effects: hydrodynamic modulation, tilt modulation, and shadowing (Nieto-Borge et al., 2004). Hydrodynamic modulation transforms the energy of the smaller ripples due to their interaction with longer waves. Tilt modulation results from changes in the effective incidence angle along the slope of the long waves. Shadowing involves the partial shadowing of the sea surface by higher waves (Plant and Keller, 1990; Wetzel, 1990; Lee et al., 1995; Nieto-Borge et al., 2004).

The numerical simulations show that when the wave field is homogeneous in space and stationary in time, hydrodynamic modulation is negligible (Nieto-Borge et al., 2004). Moreover,
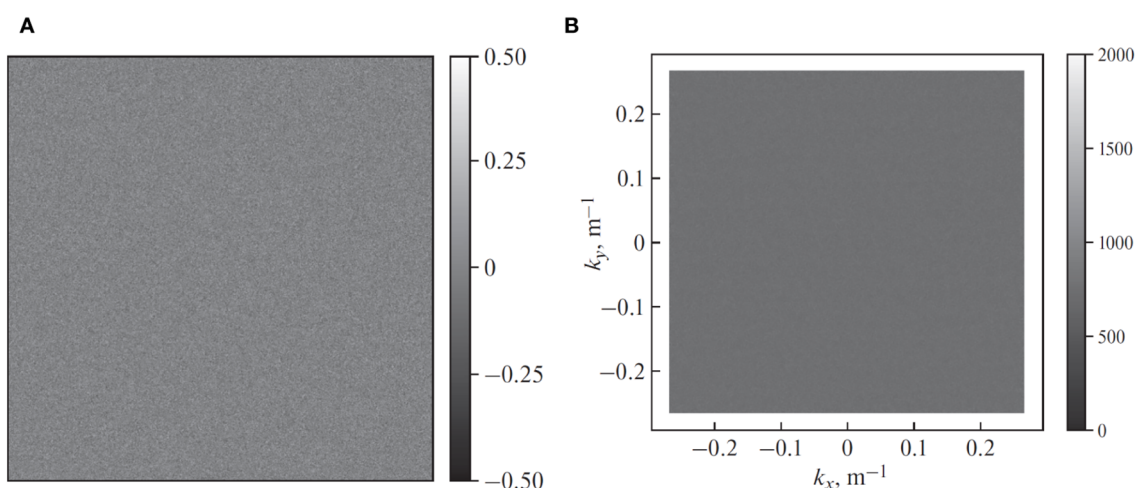

FIGURE 2
The first stage of the sea surface synthesis. **(A)** White noise. **(B)** Fourier-transformed white noise.
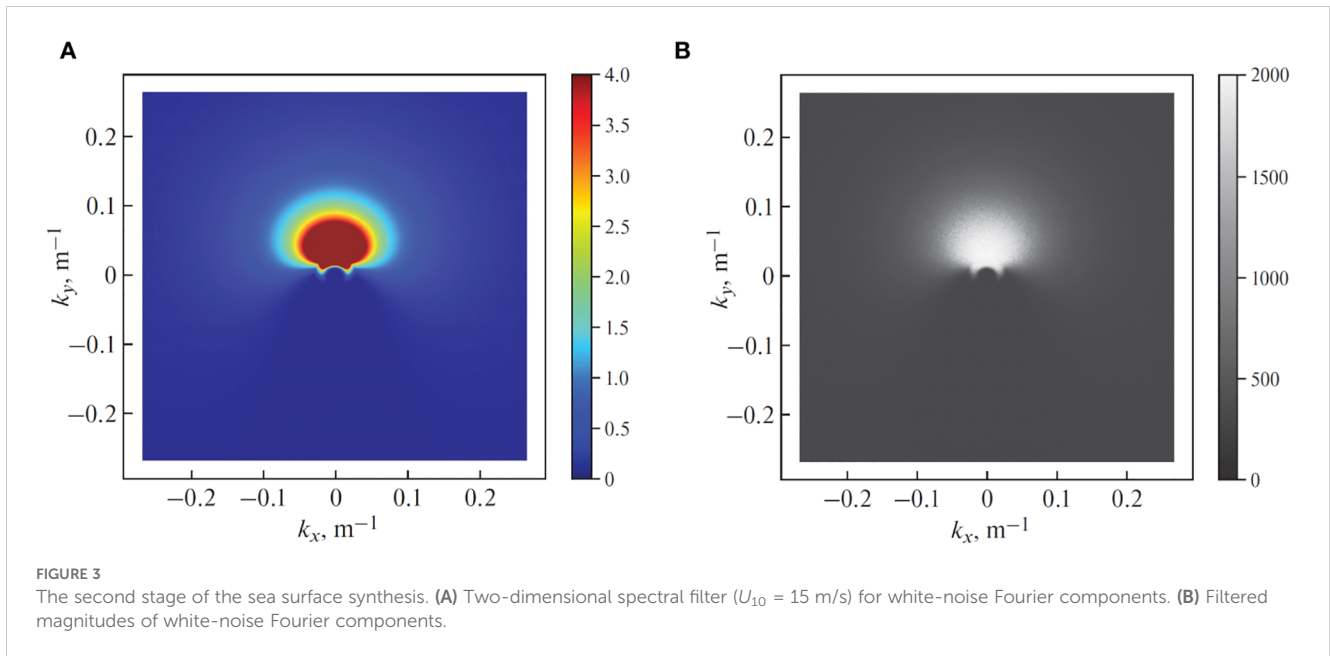
FIGURE 3
The second stage of the sea surface synthesis. **(A)** Two-dimensional spectral filter (U_{10} = 15 m/s) for white-noise Fourier components. **(B)** Filtered magnitudes of white-noise Fourier components.

local wind velocity modulation is pre-filtered from the received radar signal (Ludeno and Serafino, 2019).

Although it is proved that geometrical shadowing serves as a first-order approximation of the backscattering phenomenon, while tilt modulation have a relatively minor impact on the imaging mechanism compared to geometric shadowing at grazing incidence (Nieto-Borge et al., 2004; Wijaya and van Groesen, 2016), in this study we consider both geometrical effects omitting only hydrodynamic and local wind velocity modulations.

Thus, for simplification purposes, we propose that two main effects that influence the X-band signal reflected from the sea surface are geometrical shadowing (Gangeskar, 2014; Wijaya and van Groesen, 2016; Ludeno and Serafino, 2019) and simplified tilt modulation (Ludeno and Serafino, 2019).

As in (Wijaya and van Groesen, 2016), we believe it is important to note that the following algorithm that transforms the synthetic sea surface into the X-band radar image presented in this paper pertain to "idealized" scenarios, specifically, linear seas that are not influenced by local or global wind effects.

Simplifications mentioned above are based mainly on the fact that we use synthetic radar images solely for pre-training but not for the final ANN-based regression model. Hence, we do not claim complete realism for the generated radar images, in accordance with the elaborated papers (Nieto-Borge et al., 2004; Wijaya and van Groesen, 2016; Ludeno and Serafino, 2019).

We further follow the transformation technique proposed by (Nieto-Borge et al., 2004) with the subsequent modifications
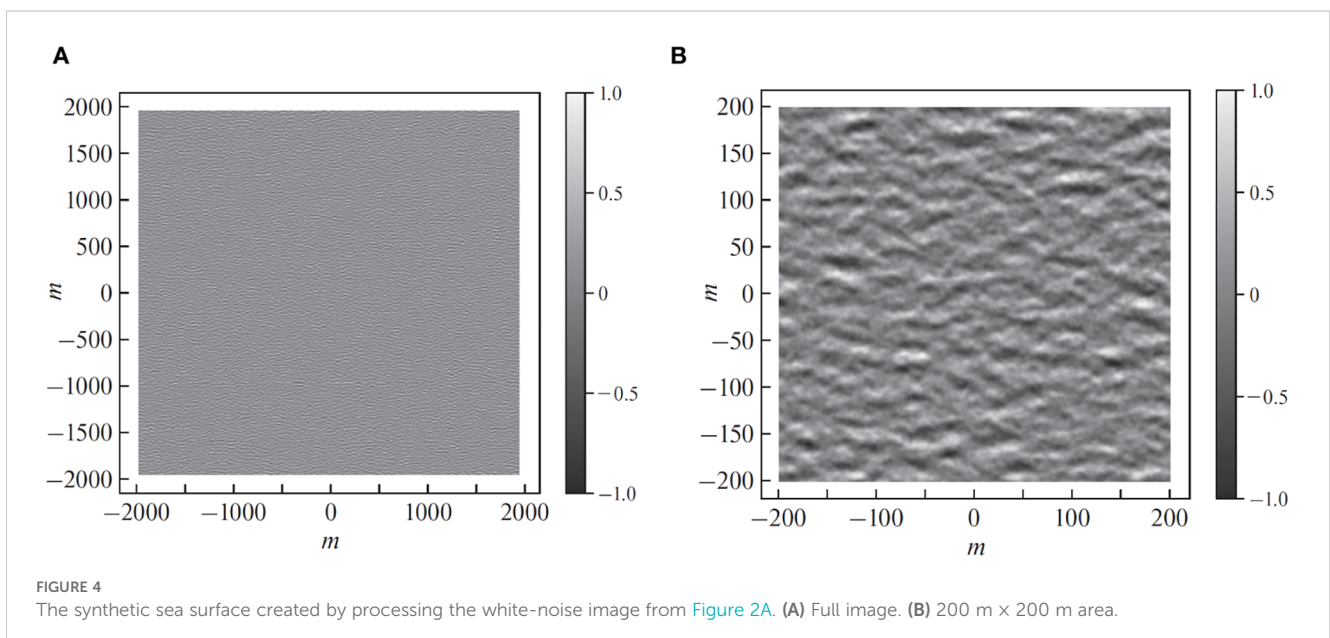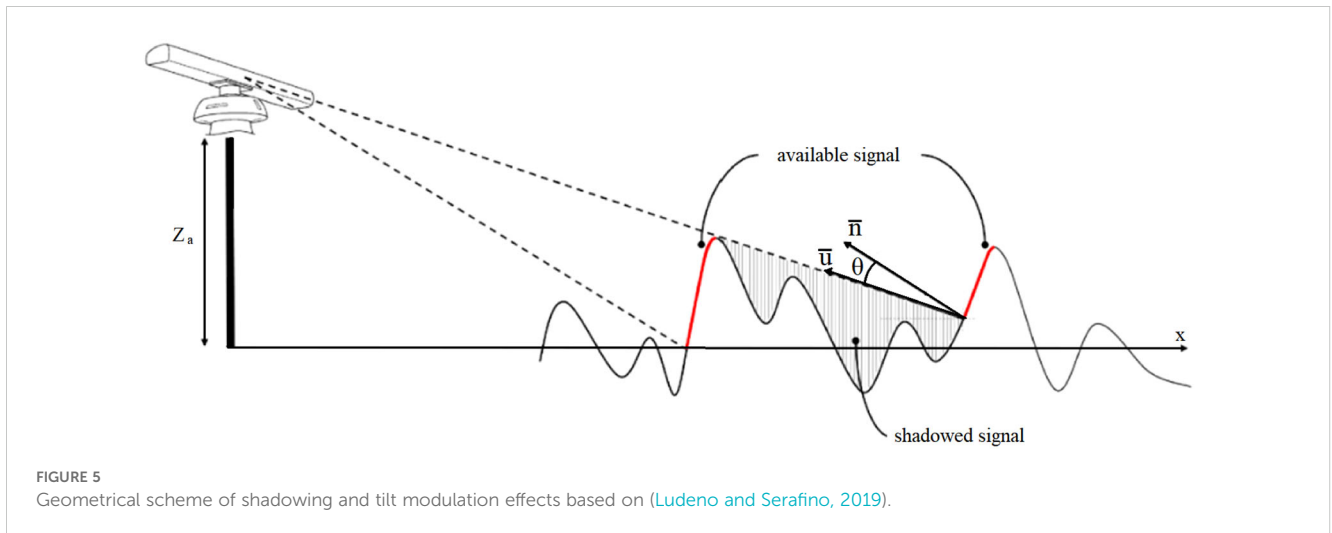

FIGURE 4
The synthetic sea surface created by processing the white-noise image from Figure 2A. **(A)** Full image. **(B)** 200 m × 200 m area.

FIGURE 5
Geometrical scheme of shadowing and tilt modulation effects based on (Ludeno and Serafino, 2019).

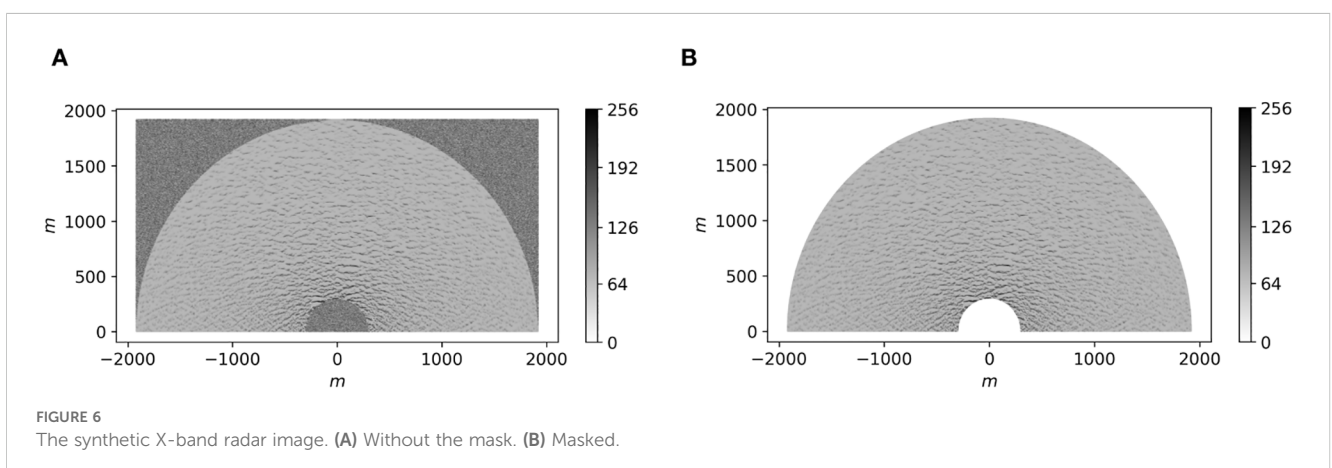(Gangeskar, 2014; Wijaya and van Groesen, 2016; Ludeno and Serafino, 2019).

Briefly, in this study, the shadowing effect refers to the geometrical optics approximation. In certain areas, the sea surface obstructs the reflection of radar rays from adjacent areas, leading to the shadowing of nearby waves. Consequently, the radar antenna receives no meaningful signal from the shadowed parts of the sea surface (Ludeno and Serafino, 2019). Obviously, this phenomenon depends on the grazing angle that is determined by the relationship between the radar antenna height Z_a and the distance x to the antenna (Figure 5).
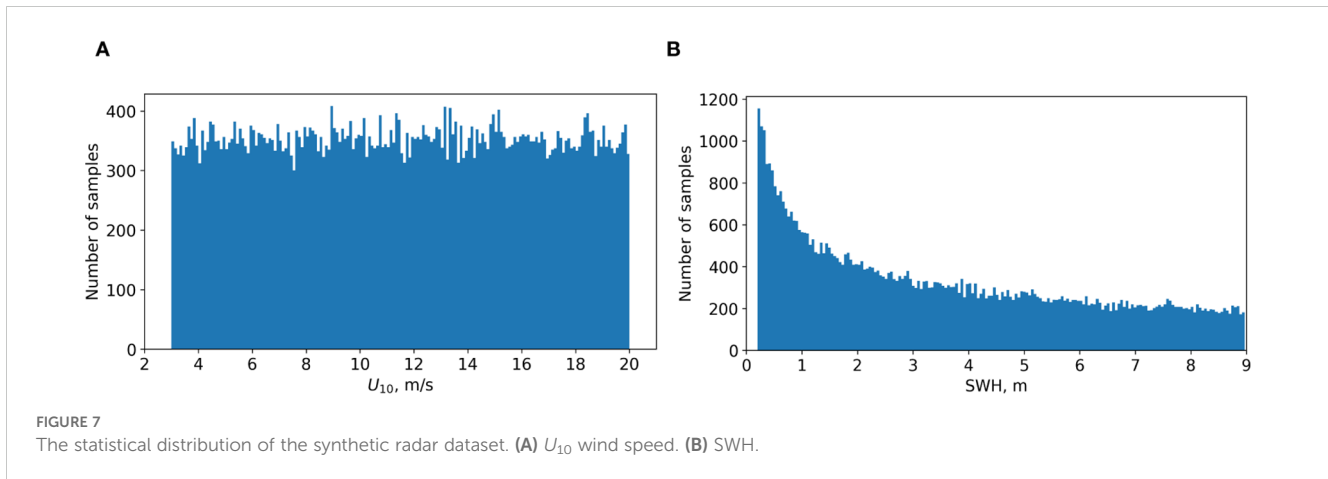
In the real radar images, the shadowed area is not equal to zero, so the data in these areas are modelled as Gaussian white noise (Ludeno and Serafino, 2019).

For tilt modulation, the steepness of the observed surface slope affects the power amplitude received by the radar antenna (Ludeno and Serafino, 2019). Thus, for non-shadowed areas, the received back-scattered signal is proportional to $\cos\theta$, where θ is the angle between the radar ray $\vec{u}$ and the vector $\vec{n}$ normal to the wave surface (Figure 5).

Summarizing, we present the following algorithm that transforms the synthetic sea surface into the X-band radar image. For every synthetic sea surface:

1. Distance mask: we determine the masked points – the points with the distance to the radar more than 1920 m or less than 300 m.
2. Shadowing: we determine the points that are not available for a radar ray.
3. Tilt modulation: for every non-shadowed point we determine the angle θ between the radar ray and the vector normal to the wave surface.
4. For every non-shadowed point the amplitude of the back-scattered signal is $\cos\theta$.
5. We compute the minimum and the maximum values of the back-scattered amplitude among the non-shadowed points.
6. Normalization: we linearly transform the amplitude values so that the new minimum value is 0 and the new maximum value is 255.
7. Noise: for shadowed and masked points we set the amplitude values as random numbers distributed uniformly from 0 to 255.
8. We choose the half of the array that corresponds to the downwind direction.

FIGURE 6
The synthetic X-band radar image. **(A)** Without the mask. **(B)** Masked.

FIGURE 7
The statistical distribution of the synthetic radar dataset. **(A)** U_{10} wind speed. **(B)** SWH.

2.4 Dataset of synthetic images

To generate a set of synthetic images, we first generate 60 000 values of U_{10} speed distributed uniformly in the range from 3 m/s to 20 m/s. We include extremely low and extremely high wind speed values to provide proper learning quality in comparison with the real wind wave conditions. For fully-developed sea, SWH value is a quadratic function of U_{10} (3) leading to a bigger share of lower SWH values. The distributions of U_{10} and SWH are shown in Figure 7.

After that, for every wind speed value we generate a corresponding sea surface, as described in 2.2, and then convert it to a synthetic radar image, according to 2.3. We compute the grazing angles with the radar antenna height Z_a= 20 m. This procedure results in 60 000 samples of synthetic radar images (Table 1).

2.5 Data pre-processing

In this research, we apply convolutional neural networks (CNNs). Consistent with standard practice, we linearly normalize both input data and the target SWH values. Namely, for every pixel px of the real and synthetic radar images, the normalized radar image pixel $px_{norm} = \frac{px - 127.5}{255}$. This procedure limits pixel values between –0.5 and 0.5. We also adjust SWH to approximately have a zero mean and a variance equal to one: $SWH_{norm} = \frac{SWH - 4.6}{8.8}$ After normalization, for the masked areas of the image (white areas in Figure 6B), we set the zero value.

CNNs lack inherent rotation invariance, necessitating diverse training data with varied spatial feature orientations (Chidester et al., 2018). In this study, the orientation of spatial features may not exhibit high diversity. Consequently, we employ two-dimensional data augmentation for the real SeaVision and synthetic radar semicircles, namely, random rotation with an angle ranging from –5° to 5°. This approach encourages the CNN-based model to acquire rotation invariance through training on augmented data, thereby enhancing the generalization ability of the CNN.

2.6 ANN models

In this study, we apply convolutional neural networks (CNNs), i.e., parametric mappings where the model parameters are optimized through sequential application of a fixed-size convolutional kernel to two-dimensional input data (Rezvov et al., 2022).

The high depth of CNNs is anticipated to enhance the predictive output quality. However, the large number of layers introduces training instability in the back-propagation algorithm, leading to learning inefficiency due to the vanishing gradients. This effect results from the accumulation of excessively small gradients for model parameters. Consequently, the product of the gradient vector and the learning rate coefficient tends toward zero, causing the parameters to remain constant during each optimization step.

An effective approach to address the issue of learning instability involves the incorporation of connections that bypass the intermediate layers of the model. These skip connections serve to diminish the likelihood of accumulating small gradients. A notable example is the implementation of residual connections, wherein the output of an intermediate layer is added to the output of a subsequent level.

In this paper, we build two CNN architectures based on the model for processing SeaVision radar images as it is described in (Krinitskiy et al., 2023). The basic architecture is the modified ResNet50 combining the advantages of deep CNNs and residual connections.

The convolutional core of the modification adheres to the original ResNet architecture, with the addition of sinusoidal positional encoding of various wavelengths to enable the CNN to capture the wavelengths peculiar to real SeaVision images (Krinitskiy et al., 2023). Two-dimensional positional encoding introduces additional channels with generated harmonic maps featuring various wavelengths and directions. Specifically, there are cosine- and sine-based positional encoding channels that vary in both horizontal and vertical

directions. These maps are concatenated with the output of the ResNet blocks so that the subsequent ResNet block processes activation maps from the previous layers along with the positional encoding channels. It is worth noting that for modified ResNet50, positional encoding maps are injected into the activation maps after each ResNet building block.

2.6.1 Unsupervised reconstruction model

The first architecture elaborated for this research is the reconstruction CNN model with the structure similar to U-Net (Ronneberger et al., 2015). U-Net is based on a typical contracting network augmented with consecutive upsampling blocks. Hence, U-Net architecture consists of two parts: the contracting path from the input to the bottleneck, and the expansive path from the bottleneck to the output. The upsampling blocks of the model contain a large number of feature channels to effectively propagate context information to higher resolution layers.

The main distinction and the principal concept of U-Net is the presence of the specific skip connections that transmit activation maps from the intermediate layers of the contracting path to the intermediate layers of the expansive path. In other words, the high-resolution features from the contracting path are added to the upsampled output on the expansive path to facilitate localization. Subsequently, a consecutive convolution layer is capable of learning to generate a more precise output based on this information (Ronneberger et al., 2015).

For this research, we develop a U-Net architecture with the contracting path based on the convolutional core of modified
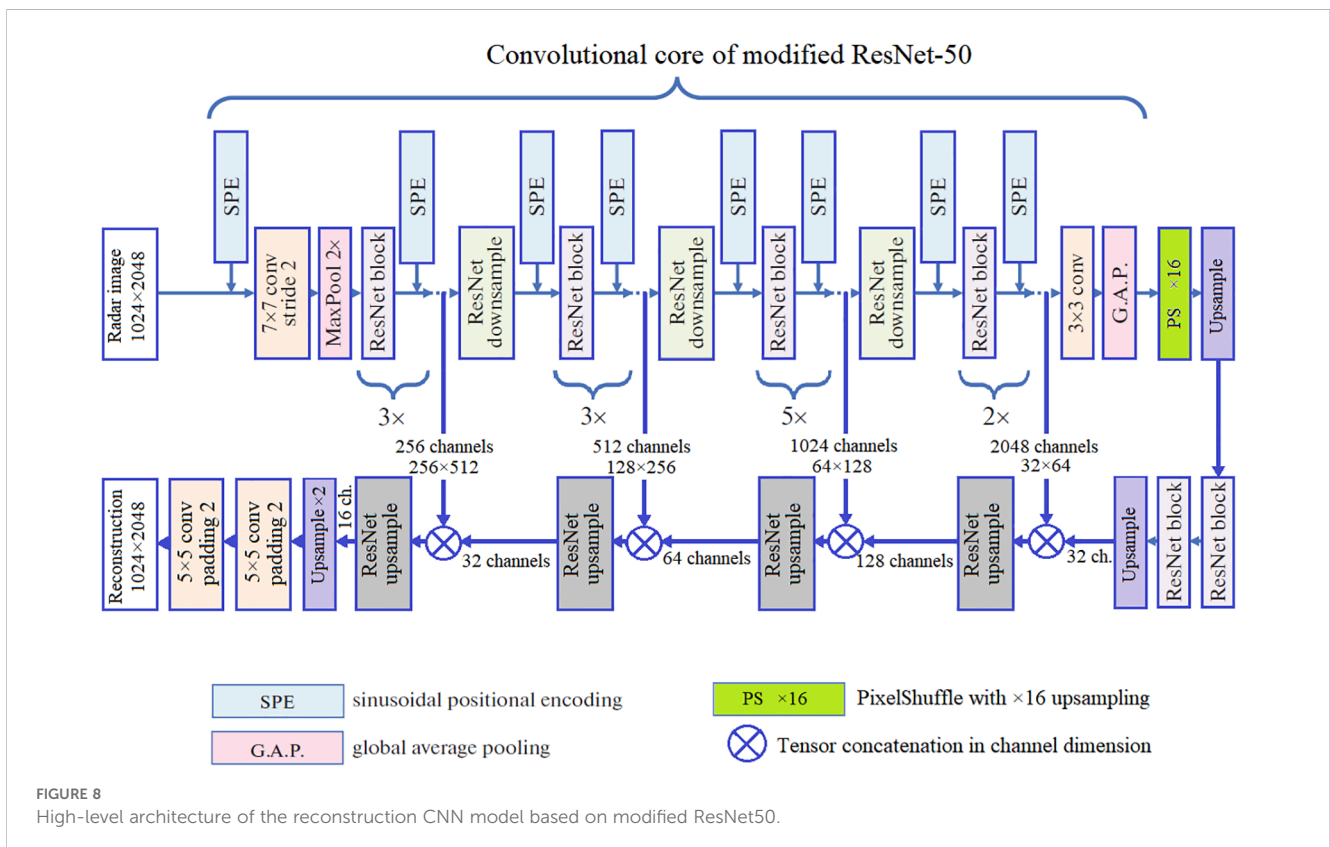
ResNet50 (the upper half of Figure 8). The input of the model is supposed to be a two-dimensional array with 1024×2048 size. Four skip connections and a bottleneck pass the multilevel features into the expansive path of the model (the lower half of Figure 8). It leads to the output with the input size. With the equal size of the input and the output, the model architecture is aimed at unsupervised reconstruction of the input array.
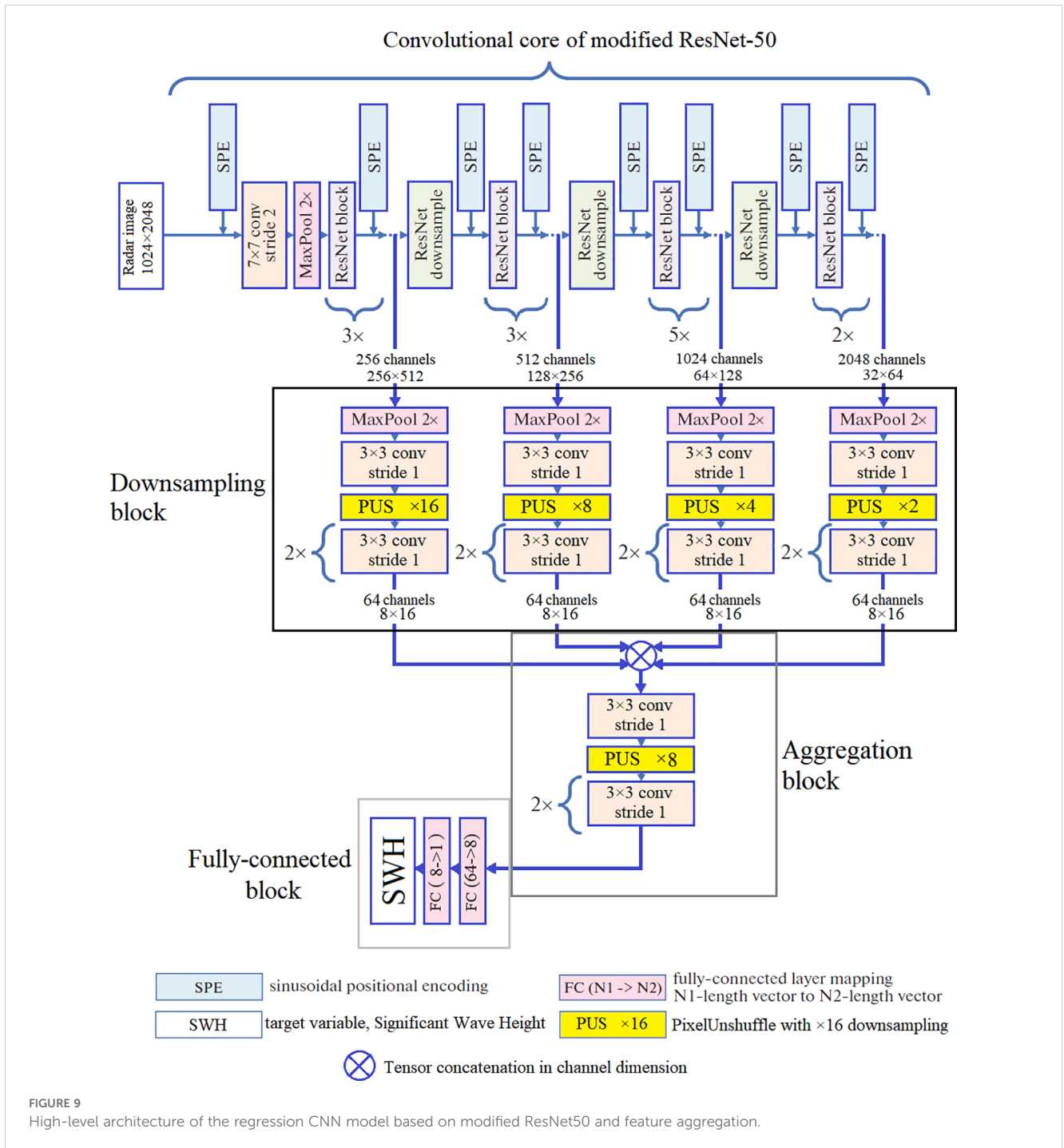
As Figure 8 shows, the expansive part of the model consists of upsampling ResNet blocks with the additional upsampling layers if necessary. On the expansive path, we coordinate the spatial size of the block outputs with the outputs of the skip connections to properly concatenate the tensors before the subsequent ResNet block or layer. The number of channels for the concatenated activation maps and the output size of the skip connections are shown in Figure 8.

2.6.2 Regression model

The second architecture built for this study is the regression CNN model designed to approximate the scalar target SWH value through processing 1024×2048 radar images. Unlike the regression model from (Krinitskiy et al., 2023) with the same input and output objects, we base our model on the reconstruction architecture described in 2.6.1 including the skip connections (see Figure 9).

As Figure 9 shows, our regression model consists of four fundamental parts. The first part is the convolutional core of modified ResNet50. We preserve four contracting paths to pass the outputs of the intermediate layers to the downsampling blocks. Every parallel downsampling block consists of a pooling layer, a

FIGURE 8
High-level architecture of the reconstruction CNN model based on modified ResNet50.

FIGURE 9
High-level architecture of the regression CNN model based on modified ResNet50 and feature aggregation.

convolutional layer, a pixel unshuffling layer and two consecutive convolutions. The task of the downscaling part of the model is to bring four activation maps of the skip connections to the same size. In the third part of the architecture, the features from the previous block are aggregated through channel concatenation. The aggregated activation map is passed to the downsampling block resulting in vector of size 64. The fully-connected subnet following the aggregation part contains two sequential fully-connected layers of the widths 64 and 8. The terminating layer is of the width 1 since in this study, we approximate SWH scalar.

2.7 Quality metrics

Various quality metrics are employed to assess and monitor the learning process of the models. For the unsupervised reconstruction model, the simplest quality metric utilized is the root mean-squared error, denoted as $RMSE_{rec}$. This form of reconstruction error quantifies the disparity between the input data and the data obtained after the U-Net compresses and decompresses the input. The $RMSE_{rec}$ metric provides an evaluation of the noise introduced to the input as it passes through the bottleneck and the skip connections.

If we denote our reconstruction model as R and a batch of input images as y, then an output batch is $y^* = R(y)$. We compute $RMSE_{rec}$ summarizing only by unmasked points to exclude non-meaningful areas of radar images. The number of points in the input batch is N, and the result is:

$$RMSE_{rec} = \sqrt{\frac{1}{N} \sum_{i,j,k} (y^*_{i,j,k} - y_{i,j,k})^2} \,. \qquad (6)$$

In (6), $y_{i,j,k}$ and $y^*_{i,j,k}$ are the identical elements of the input and the output tensors. We summarize the difference between the input and the output point-by-point. The indices i, j and k denote points in ranges of respective spatial dimensions and a batch size, respectively, taking into account distance mask.

The inadequacy of RMSE in detecting visually altered defective regions in images with consistent intensity values has been demonstrated in (Bergmann et al., 2019). In this research, we adopt a perceptual quality metric grounded in structural similarity, which assesses the inter-dependencies among local image regions. Unlike RMSE, which compares pixel values, structural similarity takes into account contrast and structural information. The computation of the structural similarity index measure (SSIM) is outlined as follows:

$$SSIM(y, y^*) = \frac{(2\mu\mu^* + C_1)(2\sigma_{yy^*} + C_2)}{(\mu^2 + \mu^{*2} + C_1)(\sigma^2 + \sigma^{*2} + C_2)},$$

where μ is the pixel sample mean of y, μ^* is the pixel sample mean of y^*, σ^2 is the variance of y, σ^{*2} is the variance of y^*, $\sigma_{yy}*$ is the cross-correlation of y and y^*. C_1 and C_2 are constant values equal to 0.01^2 and 0.03^2, respectively.

In this paper, we compute the Structural Similarity Index (SSIM) between two windows, each sized 128×128 pixels, applied to both the input and the output images. The SSIM values range between −1 and 1, where a higher SSIM signifies greater similarity.

In case of the regression model, only root mean-squared error $RMSE_{SWH}$ is utilized since the output of this CNN is a vector containing significant wave height values with a length equal to the batch size BS:

$$RMSE_{SWH} = \sqrt{\frac{1}{BS} \sum_{k=1}^{BS} (SWH^*_k - SWH_k)^2} \,.$$

where SWH is a target value measured by Spotter buoy (see 2.1) or determined by the synthetic ocean surface (see 2.2), SWH^* is the output value of the regression model, and k denotes a dataset element of a batch with the size BS.

2.8 Training and evaluation

The training process of artificial neural networks is known for its sensitivity to various details, and the selection of a training algorithm and hyperparameters plays a critical role in determining the quality of the resulting model. Currently, the Adam optimizer (Kingma and Ba, 2017) stands out as the most stable and widely employed training algorithm, utilizing a momentum approach to estimate lower-order moments of the loss function gradients. In our study, we leverage the Adam optimization procedure.

In the optimization algorithms, the batch size and the learning rate hold particular significance. Due to the substantial dimensions of SeaVision radar images, substantial variations in batch size were impractical. Instead, we selected the largest feasible batch size for our computer hardware (batch_size = 2) to mitigate noise in CNN gradient estimates.

Adhering to best practices, we fine-tuned the learning rate schedule not only to attain high quality in SWH regression but also to enhance robust generalization skills. Generalization is evaluated by scrutinizing the disparity between the quality estimated on the training and validation subsets, where a small gap signifies good generalization and a large gap indicates poor generalization.

Aligned with recent research findings, we implemented the specialized learning rate schedule (Ding, 2021; Loshchilov and Hutter, 2017). This cyclical schedule encompasses a cosine-shaped decrease in the learning rate throughout the training process. We incorporate a multiplicative form of increase in the simulated annealing period with each cosine cycle. We also employ exponential decay of the simulated annealing magnitude with each cosine cycle, utilizing the multiplicative form.

In this research, training consists of three principal consecutive steps. For all the steps, we use normalized values as it was described in 2.3.

On Step 1, we train the reconstruction model using the dataset of synthetic images described in 2.4. The dataset comprising 60 000 synthetic radar images is randomly divided into two segments: a training dataset (50 000 images) and a validation dataset (10 000 images).

In machine learning, it is customary to assess the performance of a model by computing quality metrics on a validation subset obtained through random sampling from the original set of labeled examples. This methodology presupposes that the examples are independent and identically distributed (i.i.d.). The random partition ensures a uniform distribution of modeled surface wind wave conditions across the datasets. Moreover, the distribution of SWH values is similar for both datasets.

We train the reconstruction model using the mean-squared error between an input and an output images summing only over unmasked points. The length of training is 80 epochs. We change the first cosine cycle from learning_rate = $7.5{\times}10^{-5}$ to learning_rate = $1{\times}10^{-8}$.

On Step 2, we train the regression model using the same dataset of synthetic images described in 2.4. The number of images for training and evaluation is the same as for Step 1. Nevertheless, we change the random split of the dataset to prevent overfitting.

Step 2 consists of two stages: 2^1 and 2^2. For both stages of Step 2, the input of the model is the masked radar image. We train the regression model using the mean-squared error between an output value and a target SWH.

For Stage 2^1, we initialize the weights of the convolutional modified ResNet50 core and the skip connections with the weights from the pre-trained reconstruction model as described for Step 1. After that, we train only downsampling block, feature aggregation and fully-connected layers. The convolutional core weights are frozen.

After 20 epochs, we start Stage 2^2, training the whole regression model. Epoch 1 of Stage 2^2 is ininitialized with the weights of the full regression model after epoch 20 of Stage 2^1. The length of Stage 2^2

training is 40 epochs. We change the cosine cycles from learning_rate = $7.5×10^{-5}$ to learning_rate = $1×10^{-8}$.

On Step 3, we train the regression model using the dataset of real SeaVision synthetic images described in 2.1.

In studies concerning the application of statistical models, particularly ANNs, for the analysis of remote sensing data, it is crucial to account for the auto-correlation inherent in the observational time series dataset. Owing to the inherent evolution of underlying physical phenomena, consecutive observations may demonstrate substantial auto-correlation, thereby influencing the accuracy of the model (Krinitskiy et al., 2023).

It is imperative to refrain from systematically incorporating consecutive examples into the training and testing sets. In this research, we tackle the challenge of strongly correlated successive examples by adopting station-wise random sampling, a strategy from (Krinitskiy et al., 2023). This approach avoids the systematic sampling of successive examples into the training and validation subsets, ensuring a reliable assessment of our model's quality.

In other words, to prevent overfitting and potential data leakage, we train, validate and test the model distinct stations. If the station is included in the training dataset, it is excluded from the validation and testing datasets.

Another significant challenge in statistical modeling is the problem of covariate shift, which occurs when there is a degradation in model performance during evaluation due to a discrepancy between the distribution of the testing data and that of the training data. For an accurate assessment of model quality, it is crucial that the testing and validation datasets adhere to the same distribution as the training dataset. In this study, we use a straightforward sampling strategy designed to promote, though not ensure, alignment of the target value distributions across the training, validation, and testing subsets. This approach involves station-wise sampling that is stratified according to mean SWH (Krinitskiy et al., 2023).

Step 3 consists of three stages: 3^1, 3^2 and 3^3. The input of the model is the masked radar image. We train the regression model using the mean-squared error between an output value and a target Spotter buoy SWH.

For Stage 3^1, we initialize the weights of the full model with the weights from the pre-trained regression model after Stage 2^2. Then, we train only fully-connected block for 10 epochs. After that, on Stage 3^2 we start training fully-connected block, feature aggregation and downsampling block while the covolutional core weights are frozen. After 40 epochs of Stage 3^2, we start Stage 3^3, training the full regression model. The length of Stage 3^3 training is 20 epochs. We change the cosine cycles from learning_rate= $7.5 × 10^{-5}$ to learning_rate = $1 × 10^{-8}$ for stages 3^1 and 3^2, and from learning_rate = $3.75×10^{-5}$ to learning_rate = $2×10^{-5}$ for stage 3^3.

3 Results

In this section, we present the results of training the SWH estimation model on SeaVision radar imagery including pre-training steps on synthetic radar images.

First, the unsupervised reconstruction training shows relatively high (i.e., relatively good) Structural Similarity Index value $SSIM =$

0.503 on validation dataset. This result means that the reconstruction model is able to extract the multi-level features from the synthetic radar images. The convolutional core passes the extracted and processed features through the skip connections. This detail of the architecture allows subsequent use of Step 1 weights for Step 2.

The regression model trained on the synthetic dataset (Step 2) demonstrates root mean-squared values $RMSE_{SWH}$= 0.047 m after Stage 2^1 and $RMSE_{SWH}$= 0.035 m after Stage 2^2 on validation dataset. The monotonous decrease in RMSE value proves the necessity of gradual unfreezing of pre-trained weights for smooth transition from the previous step. These values of $RMSE_{SWH}$ are too good to be true in case of real ocean data, however, one may consider them as the measure of success in case of synthetic dataset. These low (i.e., good) values of $RMSE_{SWH}$ indicate that the neural network is capable of handling the variety of imagery similar to radar-captured sea clutter images.

Finally, the regression model trained on the real dataset (Step 3) shows root mean-squared error values $RMSE_{SWH}$= 0.442 m after Stage 3^1, $RMSE_{SWH}$= 0.188 m after Stage 3^2 and $RMSE_{SWH}$= 0.204 m after Stage 3^3 on validation dataset.

The results are summarized in Table 3.

We compare the final results in our study with the simple modified ResNet50 without pre-training (Krinitskiy et al., 2023). As one may see, pre-training with synthetic data and architecture modifications proposed in our study significantly improve the quality of SWH approximation in terms of root mean-squared error from $RMSE_{SWH}$= 0.48 m reported in (Krinitskiy et al., 2023) to $RMSE_{SWH}$= 0.204 m.

4 Discussion

This study initially proposes the efficacy of neural-network regression methods for determining SWH from sea clutter imagery acquired by a navigation X-band radar. However, *in situ* data collection is both costly and limited in its ability to offer a diverse range of ocean wave conditions.

The synthesis of radar images effectively addresses the issue of insufficient data. Moreover, the generation of diverse synthetic sea surface conditions allows for simulating radar images with specific

TABLE 3 Results.

Training stage	Quality metric
Step 1	SSIM=0.503
Step 2	
Stage 2^1	RMSE=0.047 m
Stage 2^2	RMSE=0.035 m
Step 3	
Stage 3^1	RMSE=0.442 m
Stage 3^2	RMSE=0.188 m
Stage 3^3	RMSE=0.204 m

hyperparameters essential for the research. Synthetic images contribute to the expansion of the dataset, rendering it in a more uniform manner.

We present an enhanced approach for generating a synthetic sea surface and synthetic radar images under the condition of a fully-developed sea. We suppose this method promising as it allows for the utilization of various power spectra of wind waves. The resulting images exhibit realism for visual perception.

Recent studies demonstrate proficiency in the task of SWH approximation, particularly in cases of low wind speed. To enhance CNN generalization capabilities, we developed two different architectures capable of proper preliminary training to conduct experiments involving the unsupervised reconstruction and the regression model on an elaborated synthetic dataset.

The results of the unsupervised reconstruction illustrate that the U-Net-like model with skip connections and ResNet blocks is capable of capturing the harmonic structure of wind waves.

Our findings indicate that the reconstruction model reproduces the fine structure of ocean wind waves. Despite the fact that the MSE training loss, which is sensitive to point-by-point distortions, tends to smooth images and eliminate small-scale features, SSIM quality metric is relatively high. At the same time, the pre-trained regression model shows good results for both synthetic and real datasets.

It is precisely the small-scale structure of a radar image that encompasses crucial information about wave lengths and heights. Consequently, we suppose that the convolutional core of the reconstruction and regression models does not lose wave information and successfully extracts multi-scale features from the radar images. As a result, the pre-trained model shows better SWH approximation quality than the simple modified ResNet50 without pre-training (Krinitskiy et al., 2023).

5 Conclusions and outlook

In this study, we demonstrate the capability of a convolutional neural network (CNN) proposed in this paper of approximating significant wave height (SWH) with the quality superior to the one reported in recent studies exploiting CNNs for this task (Krinitskiy et al., 2023). The superiority is achieved through the preliminary training of our CNN within the subsequent stages of (a) unsupervised autoencoder-like training and (b) further supervised pre-training with a synthetic target similar to the real SWH. We demonstrate significant improvement of the SWH estimation quality in terms of root mean-squared error, compared to the results reported recently in a similar CNN-involving setting (Krinitskiy et al., 2023).

Synthetic dataset simulating radar-acquired sea clutter imagery with corresponding SWH values that we used in our pre-training may provide additional improvements for an utilized ANN due to the control over the distribution of both wind speed and corresponding SWH values.

Nevertheless, the application of synthetic images has a number of restrictions for SWH regression problem. The synthetic dataset based on fully developed sea assumption is not able to reproduce a variety of different real conditions, such as rain or obstacles on sea surface. Thus, the synthetic images do not exhibit the same statistical and physics-reflecting properties as those in the real

dataset. Moreover, the proposed algorithm is not applicable fundamentally for generating sea surface under low wind conditions. Such conditions require taking into account additional physical phenomena that distort the wave spectrum.

In the future, we plan to add more real radar images collected from different locations into the ANN to further improve the regression model quality. Different advanced methods of synthetic dataset generation are also to be explored. Furthermore, we should examine various ANN architectures for SWH regression problem.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: the PANGAEA repository https://doi.org/10.1594/PANGAEA.939620.

Author contributions

VR: Software, Visualization, Writing – original draft. MK: Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – review & editing. AG: Data curation, Formal analysis, Validation, Writing – review & editing. VG: Methodology, Software, Visualization, Writing – review & editing. MB: Software, Validation, Writing – review & editing. AS: Data curation, Visualization, Writing – review & editing. NT: Conceptualization, Funding acquisition, Project administration, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bai, S., Kolter, J. Z., and Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [online] [Preprint]. doi: 10.48550/arXiv.1803.01271. (Accessed September 17, 2024)

Bergmann, P., Löwe, S., Fauser, M., Sattlegger, D., and Steger, C. (2019). "Improving unsupervised defect segmentation by applying structural similarity to autoencoders," *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*. Prague, Czech Republic, February 25-27, 2019. Setúbal, Portugal: SciTePress, 372–380. doi: 10.48550/arXiv.1807.02011

Chidester, B., Do, M. N., and Ma, J. (2018). Rotation equivariance and invariance in convolutional neural networks [online] [Preprint]. doi: 10.48550/arXiv.1805.12301. (Accessed September 17, 2024)

Choi, H., Park, M., Son, G., Jeong, J., Park, J., Mo, K., et al. (2020). Real-time significant wave height estimation from raw ocean images based on 2D and 3D deep neural networks. *Ocean Eng.* 201, 107129. doi: 10.1016/j.oceaneng.2020.107129

Ding, Y. (2021). "The impact of learning rate decay and periodical learning rate restart on artificial neural network," in *Proceedings of the 2021 2nd International Conference on Artificial Intelligence in Electronics Engineering*, New York, NY, USA. 6–14 (Thailand. New York, N.Y.:Association for Computing Machinery). doi: 10.1145/3460268.3460270

Gangeskar, R. (2014). An algorithm for estimation of wave height from shadowing in X-band radar sea surface images. *IEEE Trans. Geosci. Remote Sens.* 52, 3373–3381. doi: 10.1109/TGRS.2013.2272701

Hasselmann, D. E., Dunckel, M., and Ewing, J. A. (1980). Directional Wave Spectra Observed during JONSWAP 1973. *J. Phys. Oceanogr.* 10, 1264–1280. doi: 10.1175/1520-0485(1980)010<1264:DWSODJ>2.0.CO;2

Huang, W., Liu, X., and Gill, E. W. (2017). Ocean wind and wave measurements using x-band marine radar: A comprehensive review. *Remote Sens.* 9 (12), 1261. doi: 10.3390/rs9121261

Huang, W., Yang, Z., and Chen, X. (2021). Wave Height Estimation From X-Band Nautical Radar Images Using Temporal Convolutional Network. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 14, 11395–11405. doi: 10.1109/JSTARS.2021.3124969

Kim, H., Ahn, K., and Oh, C. (2021). Estimation of Significant Wave Heights from X-Band Radar Based on ANN Using CNN Rainfall Classifier. *J. Korean Soc. Coast. Ocean Eng.* 33, 101–109. doi: 10.9765/KSCOE.2021.33.3.101

Kingma, D. P., and Ba, J. (2017). Adam: A method for stochastic optimization. doi: 10.48550/arXiv.1412.6980. (Accessed September 17, 2024)

Krinitskiy, M., Golikov, V., Anikin, N., Suslov, A., Gavrikov, A., and Tilinina, N. (2023). Estimating Significant Wave Height from X-Band Navigation Radar Using Convolutional Neural Networks. *Moscow Univ. Phys. Bull.* 78, S128–S137. doi: 10.3103/S0027134923070159

Lee, P. H. Y., Barter, J. D., Beach, K. L., Hindman, C. L., Lake, B. M., Rungaldier, H., et al. (1995). X-band microwave backscattering from ocean waves. *J. Geophysical Res.: Oceans* 100, 2591–2611. doi: 10.1029/94JC02741

Long, Y., Gong, Y., Xiao, Z., and Liu, Q. (2017). Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 55, 2486–2498. doi: 10.1109/TGRS.2016.2645610

Loshchilov, I., and Hutter, F. (2017). SGDR: Stochastic Gradient Descent with Warm Restarts [online] [Preprint]. doi: 10.48550/arXiv.1608.03983. (Accessed September 17, 2024)

Ludeno, G., and Serafino, F. (2019). Estimation of the significant wave height from marine radar images without external reference. *J. Mar. Sci. Eng.* 7 (12), 432. doi: 10.3390/jmse7120432

Lyzenga, D. R., and Walker, D. T. (2015). A simple model for marine radar images of the ocean surface. *IEEE Geosci. Remote Sens. Lett.* 12, 2389–2392. doi: 10.1109/LGRS.2015.2478390

Mastin, G. A., Watterberg, P. A., and Mareda, J. F. (1987). Fourier synthesis of ocean scenes. *IEEE Comput. Graphics Appl.* 7, 16–23. doi: 10.1109/MCG.1987.276961

Nieto-Borge, J. C., Hessner, K., Jarabo-Amores, P., and de la Mata-Moya, D. (2008). Signal-to-noise ratio analysis to estimate ocean wave heights from X-band marine radar image time series. *IET Radar Sonar Navigation* 2, 35–41(6). doi: 10.1049/iet-rsn:20070027

Nieto-Borge, J. C., Hessner, K., and Reichert, K. (1999). "Estimation of the significant wave height with X-band nautical radars," *18th International Conference on Offshore Mechanics and Arctic Engineering*. St. John's, Newfoundland, Canada, July 11-16, 1999. New York, N.Y.: American Society of Mechanical Engineers.

Nieto-Borge, J. C., Rodriguez, G. R., Hessner, K., and González, P. I. (2004). Inversion of marine radar images for surface wave analysis. *J. Atmospheric Oceanic Technol.* 21, 1291–1300. doi: 10.1175/1520-0426(2004)021<1291:IOMRIF>2.0.CO;2

Park, J., Ahn, K., Oh, C., and Chang, Y. S. (2020). Estimation of Significant Wave Heights from X-Band Radar Using Artificial Neural Network. *J. Korean Soc. Coast. Ocean Eng.* 32, 561–568. doi: 10.9765/KSCOE.2020.32.6.561

Pierson, W. J., and Moskowitz, L. (1964). A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii. *J. Geophysical Res.* 69, 5181–5190. doi: 10.1029/JZ069i024p05181

Plant, W. J., and Keller, W. C. (1990). Evidence of Bragg scattering in microwave Doppler spectra of sea return. *J. Geophysical Res.: Oceans* 95, 16299–16310. doi: 10.1029/JC095iC09p16299

Raghukumar, K., Chang, G., Spada, F., Jones, C., Janssen, T., and Gans, A. (2019). Performance Characteristics of "Spotter," a Newly Developed Real-Time Wave Measurement Buoy. *J. Atmospheric Oceanic Technol.* 36, 1127–1141. doi: 10.1175/JTECH-D-18-0151.1

Rezvov, V., Krinitskiy, M., and Gulev, S. (2022). Approximation of high-resolution surface wind speed in the North Atlantic using discriminative and generative neural models based on RAS-NAAD 40-year hindcast. *Proceedings of Science.* 2022, 23 doi: 10.22323/1.429.0023

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, eds. N. Navab, J. Hornegger, W. Wells and A. Frangi (Cham: Springer), 234–241. doi: 10.1007/978-3-319-24574-4_28

Tilinina, N., Ivonin, D., Gavrikov, A., Sharmar, V., Gulev, S., Suslov, A., et al. (2022). Wind waves in the North Atlantic from ship navigational radar: SeaVision development and its validation with the Spotter wave buoy and WaveWatch III. *Earth System Sci. Data* 14, 3615–3633. doi: 10.5194/essd-14-3615-2022

Vicen-Bueno, R., Lido-Muela, C., and Nieto-Borge, J. (2012). Estimate of significant wave height from non-coherent marine radar images by multilayer perceptrons. *EURASIP J. Adv. Signal Process.* 84. doi: 10.1186/1687-6180-2012-84

Wetzel, L. B. (1990). "Electromagnetic Scattering from the Sea at Low Grazing Angles," in *Surface Waves and Fluxes. Environmental Fluid Mechanics*, vol 8., eds. G. L. Geernaert and W. L. Plant (Dordrecht: Springer), 109–171. doi: 10.1007/978-94-009-0627-3_3

Wijaya, A., and van Groesen, E. (2016). Determination of the significant wave height from shadowing in synthetic radar images. *Ocean Eng.* 114, 204–215. doi: 10.1016/j.oceaneng.2016.01.011

Young, I. R., Rosenthal, W., and Ziemer, F. (1985). A three-dimensional analysis of marine radar images for the determination of ocean wave directionality and surface currents. *J. Geophysical Res.: Oceans* 90, 1049–1059. doi: 10.1029/JC090iC01p01049