



Article

Use of a Residual Neural Network to Demonstrate Feasibility of Ship Detection Based on Synthetic Aperture Radar Raw Data

Giorgio Cascelli ¹, Cataldo Guaragnella ², Raffaele Nutricato ^{1,*}, Khalid Tijani ¹, Alberto Morea ¹, Nicolò Ricciardi ¹ and Davide Oscar Nitti ¹

¹ Geophysical Applications Processing G.A.P. s.r.l., 70126 Bari, Italy; khalid.tijani@gapsrl.eu (K.T.); alberto.morea@gapsrl.eu (A.M.); nicolo.ricciardi@gapsrl.eu (N.R.); davide.nitti@gapsrl.eu (D.O.N.)

² Department of Electrics and Information Engineering, Politecnico di Bari, 70126 Bari, Italy; cataldo.guaragnella@poliba.it

* Correspondence: raffaele.nutricato@gapsrl.eu

Abstract: Synthetic Aperture Radar (SAR) is a well-established 2D imaging technique employed as a consolidated practice in several oil spill monitoring services. In this scenario, onboard detection undoubtedly represents an interesting solution to reduce the latency of these services, also enabling transmission to the ground segment of alert signals with a notable reduction in the required downlink bandwidth. However, the reduced computational capabilities available onboard require alternative approaches with respect to the standard processing flows. In this work, we propose a feasibility study of oil spill detection applied directly to raw data, which is a solution not sufficiently addressed in the literature that has the advantage of not requiring the execution of the focusing step. The study is concentrated only on the accuracy of detection, while computational cost analysis is not within the scope of this work. More specifically, we propose a complete framework based on the use of a Residual Neural Network (ResNet), including a simple and automatic simulation method for generating the training data set. The final tests with ERS real data demonstrate the feasibility of the proposed approach showing that the trained ResNet correctly detects ships with a Signal-to-Clutter Ratio (SCR) > 10.3 dB.

Keywords: SAR raw data processing; ship detection; convolutional neural networks



Citation: Cascelli, G.; Guaragnella, C.; Nutricato, R.; Tijani, K.; Morea, A.; Ricciardi, N.; Nitti, D.O. Use of a Residual Neural Network to Demonstrate Feasibility of Ship Detection Based on Synthetic Aperture Radar Raw Data. *Technologies* **2023**, *11*, 178. <https://doi.org/10.3390/technologies11060178>

Academic Editors: Pietro Zanuttigh, Yangquan Chen, Nunzio Cennamo, M. Jamal Deen, Junseop Lee, Subhas Mukhopadhyay and Simone Morais

Received: 30 August 2023
Revised: 27 November 2023
Accepted: 5 December 2023
Published: 11 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the deployment of the first satellite with a synthetic aperture remote sensing system into orbit in 1978 [1], the use of SAR imagery has been a vital part of several scientific domains, including environmental monitoring, early warning systems, and public safety [2].

The SAR functioning principle can be described as follows [3]: microwave signals are emitted from a radar mounted on a moving object, such as a satellite, and later, the back-scattered signals (echoes) are collected to generate raw data. The motion of the sensor is opportunistically used to synthesize a very long antenna and obtain a high-resolution image of the scene being observed by properly integrating the collected raw data. SAR is distinguished from other mapping techniques by its active operational mode. In fact, the sensor has its own energy source for illuminating the local region. Thanks to this, SAR permits service regardless of solar radiation or time of day. In addition, it is possible to select the transmitted wavelength in a specific range wherein the attenuation of electromagnetic waves generated by the atmosphere can be neglected. As a result, the SAR sensor can work in the day, night, and nearly all weather conditions and can even monitor a cloud-covered region. SAR could be described as “non-literal imaging” since the raw data do not resemble an optical image and are incomprehensible to humans (see Figure 1). For this reason, SAR raw data are typically processed by means of a focusing algorithm that, in a nutshell, implements space-variant two-dimensional filtering of raw data. More specifically, the focusing step uses the motion of the sensor to synthesize a very long antenna and

obtain a higher-resolution image of the scene being observed. The result of this process is a complex image, known as a Single Look Complex (SLC) image. The generic pixel of the SLC image corresponds to a specific cell on the ground. The phase of the image is typically used to extract information about the sensor pixel distance in interferometric applications related to Digital Elevation Model (DEM) generation [4] and ground motion monitoring [5]. The amplitude of the image is related to the backscattering coefficient of the cells on the ground and is typically used in segmentation and classification applications, such as oil and ship detection (see Figure 2).

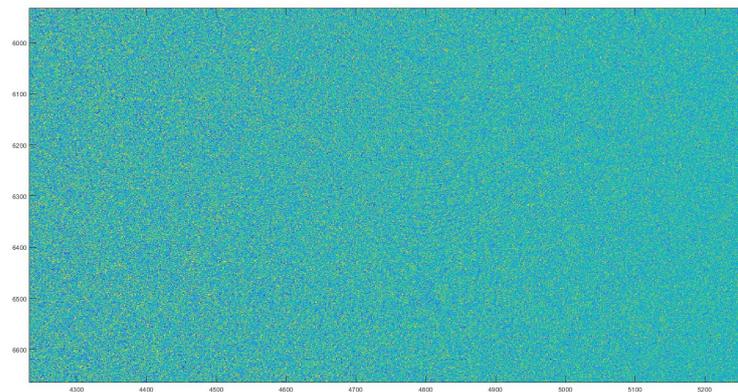


Figure 1. Example of the real part of an SAR raw image taken in a ship and an oil spill scene.

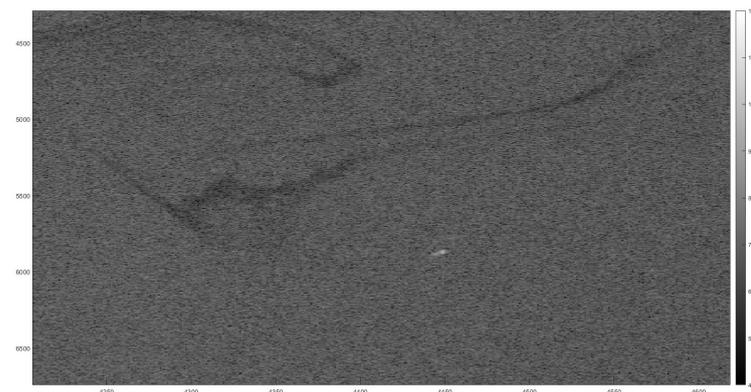


Figure 2. Example of focused SAR image taken in a ship and an oil spill scene.

Although several efficient focusing approaches are available in the scientific literature [6–9], the processing of raw data requires a significant amount of computational power. Moreover, focusing represents only the first step of the detection process since it may be necessary to introduce further processes, such as calibration [10] and despeckle filtering [11], before proceeding with the actual detection phase. As a result, it is almost never practical to perform it onboard, and consequently, the data are transmitted back to Earth to be processed. This introduces a non-negligible latency, which is the main limitation to the implementation of SAR-based real-time services. This is the case, for example, for oil and ship detection services. Although the potential of SAR data has been well demonstrated and consolidated for maritime surveillance [12–14], the latency due to the transmission of radar data to the ground segment currently represents a strong limitation.

In this scenario, onboard detection would represent an interesting solution to the problem thanks to the possibility of performing detection as soon as data are taken, thus also enabling transmission to the ground segment of alert signals only, with a notable reduction in both latency and downlink bandwidth. However, as already mentioned, the reduced computational capabilities onboard prevent the application of algorithms adopted on ground computing centers and even the simple focusing step can represent a complex operation from a computational point of view. The objective of next-generation

studies [15] is then to optimize Earth Observation (EO) data processing in order to deliver EO products to the end user with very low latency using a combination of advancements in onboard processing. As a result, in the field of the EO, the demand for real-time decision-making capabilities jointly with the need to optimize data acquisition and transmission has led to a significant interest in onboard intelligence, particularly in scenarios such as the early detection of disasters, extreme events, maritime situation awareness, and automatic target recognition, where swift and informed decisions can mitigate potential risks and minimize impacts and damages [16].

In this scenario, Artificial Intelligence (AI) can provide the needed boost, thus allowing new applications to be realized [17], by focusing directly on the information contained in the data, autonomously extracting relevant features for a given application. In particular, as highlighted in [18], Convolutional Neural Networks (CNNs) have demonstrated remarkable results in several space applications, such as scene classification, object recognition, pose estimation, change detection, and others; while the initial paradigm was to have these applications run using a server-hosted processor, recent advances in microelectronics provide efficient hardware accelerators enabling the implementation of AI algorithms “at the edge” [19]. Furthermore, according to [18], mission lifetime extensions, as well as improvement by means of delta training, could also be immediately feasible for AI solutions via dynamic reconfigurability of the CNN.

However, as pointed out in [16], the research on Edge Computing (EC) to reduce the amount of data transmission and energy consumption in EO applications is still in its infancy. As a response to this demand, preliminary research activities have been directed toward the exploration of onboard intelligence for EO applications based on optical sensors. In particular, the ability to run Machine Learning (ML) algorithms with EO-oriented platforms has been demonstrated by SmallSat missions with optical payloads such as ϕ sat missions, using off-the-shelf ultra-low-power AI-optimized chips [18]. As of today, it is noteworthy that fewer efforts have been dedicated to SAR sensors in comparison to optical sensors. This is also well recognized by the European Space Agency (ESA), which has recently launched several initiatives (e.g., the onboard intelligence for SAR missions invitation to tender) in which it is recognized that, despite the immense potential of SAR, it remains an under-explored domain in terms of onboard intelligence implementation. One of the main reasons identified by ESA is the complexities arising from onboard SAR data’s inherent characteristics with unfocused SAR raw data, including its complex interactions with the Earth’s surface, that present challenges (but also opportunities) for integrating intelligent decision-making processes directly into SAR-equipped spacecraft. For these reasons, approaches based on detection applied directly to raw data may represent a possible solution that requires more investigation. More specifically, the accuracy and computational cost of these approaches are both of extreme interest as they are not sufficiently addressed in the literature. In this regard, the present work analyzes the ship detection accuracy that can be achieved by using a Convolutional Neural Network directly applied to SAR raw data.

Recent studies have indicated that CNN might outperform human performance in a variety of fields [20,21]. As described in a dedicated recent survey [22], CNNs have also received increasing attention from researchers in target detection with SAR images and have evolved rapidly, achieving interesting results in many application fields. More specifically, in [22] traditional detection algorithms are classified into three categories (based on structural features, gray features, and texture features) and are compared with CNN-based algorithms, highlighting the interesting advantages of deep learning, mainly consisting in the capability to (i) achieve high classification accuracy, (ii) accurately extract high-level features, and (iii) provide strong robustness and adaptability to complex environments. However, all the traditional and CNN-based approaches analyzed in [22] and the references therein introduce a preliminary pre-processing stage that involves, at least, the focusing step and possibly despeckle filtering. The (traditional or CNN-based) detection stage is then applied after the pre-processing stage. It is also interesting to note

that both focusing and despeckling are effectively spatial filtering operations (space-variant or space-invariant, depending on the chosen algorithm) and, therefore, can ultimately be mathematically described using the convolution operation. Considering that CNN is a typical deep learning model that uses convolution operations and non-linear mapping to effectively extract target features, there exists a need to further investigate approaches that could integrate pre-processing and the detection stage into a single CNN algorithm, thus working directly on raw data. More specifically, our study attempts to eliminate any pre-processing by training a CNN to directly recognize ships on raw data. We only focused on the feasibility of detection, while computational cost is not within the scope of this work, even though the possibility to skip some pre-processing steps intrinsically opens new scenarios to substantially shorten processing and delivery times.

In this analysis, it must also be considered that, in recent years, the visual attention mechanism used by humans to deal with complex visual signals (selecting relevant information and filtering out irrelevant information) has gained attention in the fields of object detection and image quality assessment [23,24]. In [25], a novel SAR image target detection method based on a visual attention model is proposed with performance that, in simulation tests, shows good capability and robustness compared to traditional algorithms such as Constant False Alarm Rate (CFAR). However, these encouraging results are obtained on focused SAR images where the target response is concentrated in a well-confined pixel area. On the contrary, in raw images, the radar target response is spread over several thousands of pixels, and very often, it is not detectable by the human eye (as illustrated later in this paper), thus leading us to avoid testing this approach in this work and leaving it for future research activities.

Another important aspect of this work concerns training data sets that, as well known, are critical to the development of machine learning algorithms [26,27]. In this regard, the efficacy of the final machine-learning-powered solution for a specific application is ultimately determined by the quality and amount of the training data. In particular, as detailed in [28], a major challenge in using CNNs in the SAR domain is the availability of large (and realistic) data sets with annotated SAR data, thus introducing overfitting risk. This is particularly true for training data sets based on raw data. As a result of the above challenges, the scientific literature proposes transfer learning techniques focused on the transfer of knowledge from a secondary related domain, where labeled data are easy and inexpensive to obtain. For example, refs. [29,30] give details concerning transfer learning based on optical images. Alternatively, the use of SAR simulators can also be a valid approach, as detailed in [31], where it is shown that a simulated SAR data set can be transferred successfully to a real SAR data scenario.

Four possible approaches are then available to produce, in the present work, a suitable training data set for raw-data-based ship detection: (i) the use of real data extracted from SAR images, (ii) the use of optical-based transfer learning, (iii) the use of an SAR data simulator, and (iv) the use of a hybrid approach (real and synthesized data combined together). All of these approaches have advantages and disadvantages. The approach based on the use of real data allows the training of the CNN in conditions similar to those of the operational service, but data generation is, in general, a time-consuming activity. The use of transfer learning and SAR simulators allows the quick production of a very large number of training data but may produce unrealistic data. The hybrid approach can represent a compromise between the two previous approaches, but it can also represent a complication related to the lack of harmonization between real and simulated data.

In our work, we mainly wanted to give priority to the ability to have physical control of data set generation at the expense of the complexity of the SAR scenes; therefore, we adopted a simplified simulation approach. This choice will provide a quantitative analysis of the applicability of a simple simulation tool in real application scenarios. The proposed simulator produces a set of complex matrices simulating SAR raw data acquired in a marine environment with the presence of ships. The synthetic data set was then used to train and evaluate a state-of-the-art CNN. More specifically, the simulator developed is

based on the parameters of the ERS SAR mission, which is a well-documented system [32] and, therefore, lends itself to the simulation of data sets in realistic conditions. As already stated, the simulation approach is based on a low level of complexity aimed at developing a fast and automatic simulator. The adequacy of this choice was subsequently verified by evaluating the CNN performance with real data extracted from ERS images. The scientific literature provides many applications of ERS data to the maritime context [33,34], and this allows the easy identification of case studies to be used for the test of the proposed CNN in real scenarios.

2. Materials and Methods

As described in the previous section, this study is aimed at the implementation of a CNN-based approach capable of detecting ships directly on SAR raw data. This required two main steps:

- The implementation of an SAR simulator of raw data in a maritime environment;
- The configuration and training of the selected CNN by using the training data sets generated with the SAR simulator.

As already stated, the implementation of the simulator followed an approach based on a low level of complexity aimed at developing a fast and automatic simulator. More accurate approaches based on, for example, advanced physical models have been discarded due to the large number of external parameters to consider and their complex interaction. This would implicate a challenging activity with the implementation of a complex simulator that may lead to the generation of training data sets based on data extraction from real images, which is not the strategy that we wished to adopt in our work. As a result, we simplified the simulation of the sea and ship raw data set according to the block scheme presented in Figure 3. Details for each block are addressed in the next subsections.

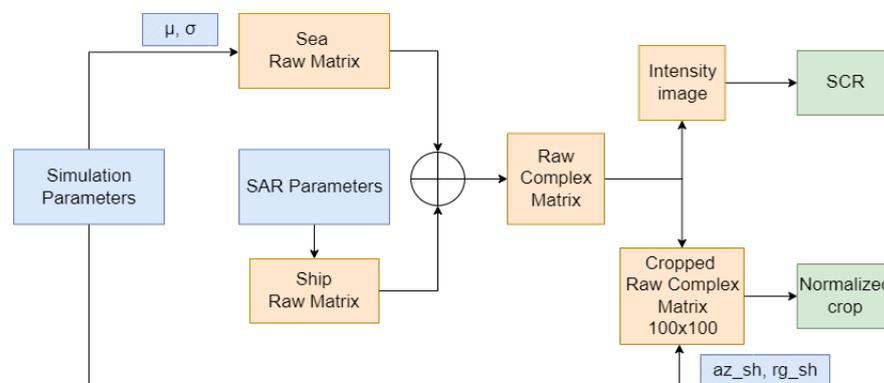


Figure 3. Simulator scheme.

2.1. SAR Parameters

As already mentioned, the real and simulated data are based on the ERS-1 and ERS-2 satellites. Although these satellites no longer operate, they together represent, even still today, one of the most successful satellite SAR missions. The ERS-1 satellite was launched in 1991, followed by ERS-2 in 1995. The two ESA satellites, at the time, represented the most technologically advanced Earth-observing spacecraft that had ever been manufactured in Europe and were able to collect data for more than fifteen years, thus generating a very wide data set suitable for both land and sea applications with a very extensive bibliography [35].

The main ERS parameters are listed in Table 1. The values are derived from parameters defined in the ERS handbook [32]. The ERS operating frequency falls in the C band that, according to IEEE designation, spans from 4.0 to 8.0 GHz, which is a typical operating band for Earth Observation radar systems. Additional parameters that depend on the specific acquisition time, such as the radar's altitude and velocity, can be extracted from the

auxiliary files associated with each image. For our simulator, we used the average values indicated in Table 2.

Table 1. ERS parameters.

Name	Symbol	Value
Carrier frequency	f_0	5.3 GHz
Signal bandwidth	BW	15.55 MHz
Pulse width	T_R	37.10 μ s
Chirp rate	f_R	2.0889×10^{11} Hz/s
Pulse repetition frequency	PRF	1.68 kHz
Sampling frequency	f_s	18.96 GHz
Radar aperture	W_A	10 m
Look angle	θ	23°
Transmitter peak power	P_{peak}	4.8 kW

Table 2. ERS flight parameters.

Name	Symbol	Value
Velocity	v_{sat}	7.12 km/s
Radar altitude	H_{sat}	785 km

2.2. Ship Simulation

The SAR response of a ship depends on many factors, such as the material, shape, dimensions, and orientation of the ship toward the radar acquisition system. For this reason, we propose an approach aimed at implementing a minimal level of complexity in the simulation of a ship's response, choosing the single-point model. The SAR signal model used for the implementation of the simulated data set is based on the acquisition geometry depicted in Figure 4.

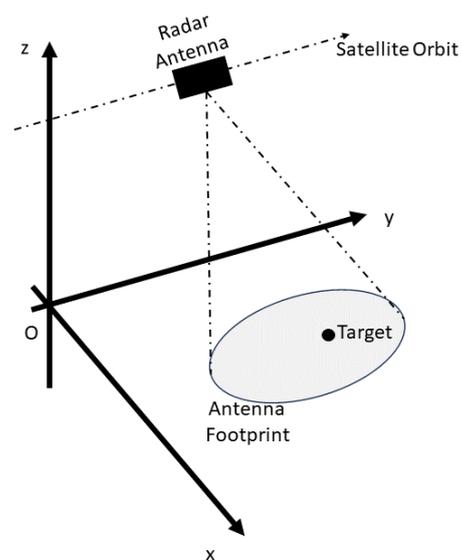


Figure 4. SAR acquisition geometry.

The SAR sensor travels along a flight path (known as the azimuth direction) such that the antenna phase center, at a given instant τ , has a three-dimensional spatial location denoted by:

$$r_A = [x_A(\tau), y_A(\tau), z_A(\tau)] \quad (1)$$

In this acquisition geometry, a generic stationary (or slowly moving) point target located in the scene has the following coordinates:

$$r_T = [x_T, y_T, z_T] \quad (2)$$

The distance from the antenna phase center to the target, known as range, is denoted by:

$$d_{AT}(\tau) = \sqrt{(x_A(\tau) - x_T)^2 + (y_A(\tau) - y_T)^2 + (z_A(\tau) - z_T)^2} \quad (3)$$

At periodic intervals (given by the inverse of the pulse repetition frequency), the radar antenna transmits a pulse, known as chirp and defined by the following equation:

$$s_T(t) = A_T \cdot \text{rect}\left(\frac{t}{T_R}\right) \cdot \exp(j(2\pi f_0 t + 2\pi f_R t^2)) \quad (4)$$

where A_T is the amplitude of the transmitted signal, T_R is the duration of the chirp signal, f_0 is the carrier frequency, f_R is the chirp rate, and t is the time axis centered in the generic instant τ_P when the transmission occurs.

The point target on the ground receives this signal and produces a so-called echo, which is a delayed and rescaled version of the transmitted signal, as defined by the following equation:

$$s_R(t) = A_R \cdot \exp(j\phi) \cdot s_T\left(t - \frac{2d_{AT}(\tau)}{c}\right) \quad (5)$$

where c is the speed of light, A_R is the magnitude of the received signal and is different from A_T since it takes into account the energy loss during propagation and reflection, and ϕ represents the phase change that may result from the target scattering processing. The time delay of $2d_{AT}(\tau)/c$ takes into account the two-way propagation from antenna to target.

The received signal is demodulated in order to remove the carrier frequency:

$$s_{Rdem}(t) = s_R(t) \cdot \exp(-j2\pi f_0 t) \quad (6)$$

Finally, the received demodulated signal is sampled (at $t_n = n/f_s$) and quantized in order to obtain a discrete-domain signal. The sampling frequency used in the ERS mission produces $N_{rg} = 704$ range samples for the chirp signal. It can be observed that the antenna footprint allows the collection of several echoes from the same target in different τ_P instants. In particular, for the ERS case, we used a realistic value of $N_{az} = 1600$ echoes. As a result, a single-point generates an overall ERS impulse response that takes about 700 samples in the range direction and about 1600 samples in the azimuth direction:

$$I_{impulse}(m, n); m = 1, \dots, N_{az}; n = 1, \dots, N_{rg} \quad (7)$$

where m and n are the matrix coordinates of the azimuth and range directions, respectively.

As illustrated in the simulation of Figure 5, despite the single-point simplification, it must be noted that the relevant raw data generated are a complex matrix, wherein rows represent azimuth samples and columns represent range samples, with azimuth being the flight direction of the satellite and range being the look direction (distance) of the SAR sensor. Although this is a considerable simplification, it has proven to be extremely useful for verifying whether the training of the selected CNN in such conditions is still valid enough to subsequently operate with real data.

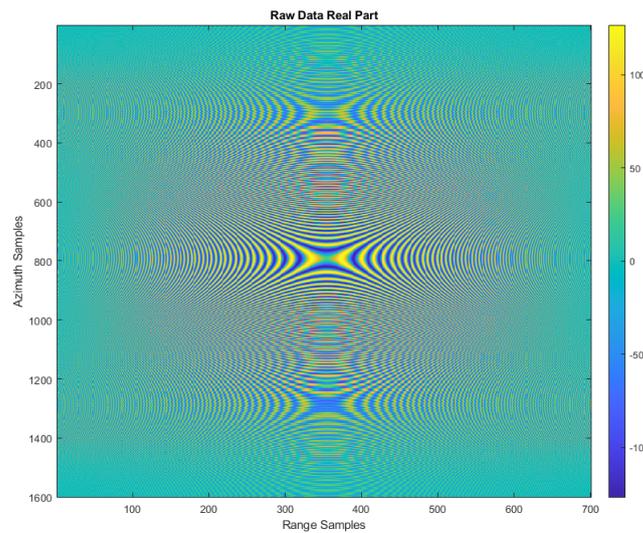


Figure 5. Real part of the ERS impulse response of a one-point modeled ship centered in the image.

Once the single-point impulse response is available, $I_{impulse}(m,n)$, it is possible to generate raw data images with a generic position of the ship by simply adding a range and an azimuth shift by means of the following equation:

$$I_{ship}(m, n) = I_{impulse}(m - az_sh, n - rg_sh) \quad (8)$$

where the parameters az_sh and rg_sh are the applied shifts according to the block diagram of Figure 3.

It is interesting to note that the typical SAR focusing step is in charge of collecting all the echoes and ideally compressing their energy in a single pixel located at coordinates az_sh and rg_sh . This is obtained by means of several techniques proposed in the literature [6–9] that efficiently implement a 2D convolution with a matched filter $H_{matched}$, thus obtaining a focused SLC image:

$$I_{SLC}(m, n) = I_{ship}(m, n) \otimes H_{matched}(m, n) \quad (9)$$

2.3. Sea Simulation

A realistic marine simulation is a difficult goal to attain. Many factors must be considered, including weather, wind, sea state, marine currents, biological spots, and oil slicks. For this work, we propose a simplified approach based on a statistical model of the sea directly derived from the distribution of the values of the raw data.

As outlined in [36], either in the high-resolution or low-resolution case, with the ideal hypothesis of a sea background having a constant radar cross section and the central limit theorem, both the real and imaginary parts are Gaussian-distributed. Although more complex modeling can be proposed, the analysis carried out on a number of different crops of ERS-1/2 raw complex data (see Figure 6) shows that it is acceptable to model both the real and the imaginary parts with a truncated Gaussian distribution according to the following equation:

$$s_{sea}(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \cdot \exp\left(-\frac{1}{2} \cdot \left(\frac{x - \mu}{\sigma}\right)^2\right); x \in [X_{min}, X_{max}] \quad (10)$$

The four variables used to characterize the truncated Gaussian distribution are the mean (μ), the standard deviation (σ), and the truncation limits (X_{min} and X_{max}). In particular, for our simulation, the following were employed:

- μ was set to 0;
- σ was varied in order to generate different rough sea conditions, i.e., different levels of the sea signal with respect to the ship signal, thus obtaining different Signal-to-Clutter Ratio (SCR) values;
- The truncation limits were set according to the quantization of the real and imaginary parts performed by the 5-bit analog-to-digital converter of the ERS acquisition system. This allows the quantized data to take integer values in the range of -16 ($X_{min} = -2^4$) to $+15$ ($X_{max} = +2^4 - 1$).

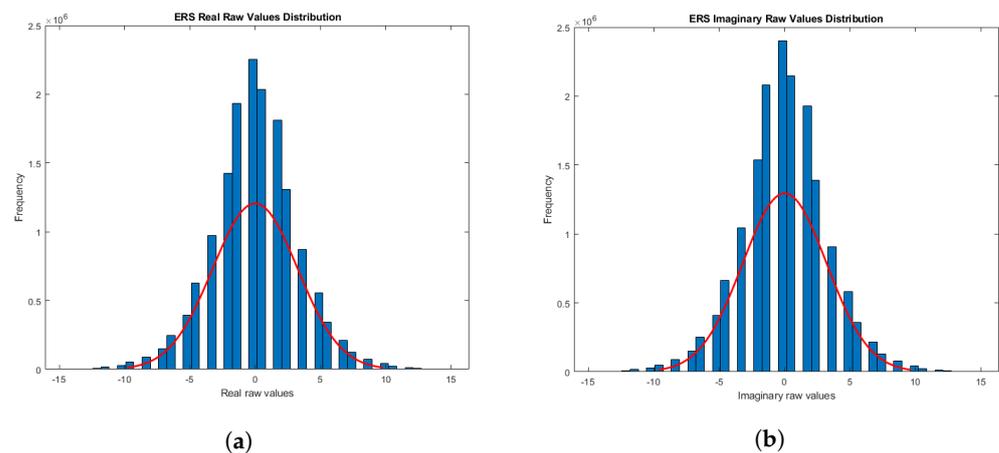


Figure 6. Gaussian-fitted real part (a) and imaginary part (b). The histogram shows the distribution of values extracted from crops of ERS raw data, and the red line is the fitted distribution.

The simulation is simply implemented by initializing an empty matrix with complex values by using random values generated via the truncated Gaussian distribution (here indicated with X_{sea}). This matrix represents the raw data generated from the sea return:

$$I_{sea}(m, n) = X_{sea_real}(m, n) + i \cdot X_{sea_imag}(m, n) \quad (11)$$

where i is the imaginary operator and $X_{sea_real}(m, n)$ and $X_{sea_imag}(m, n)$ are real matrices randomly generated with the truncated Gaussian distribution.

2.4. Training Data Set Generation

In accordance with the block scheme illustrated in Figure 3, once a pair of sea and ship simulations are produced, these two contributions are summed into a single complex matrix that represents the simulated noisy raw data image:

$$I_{raw}(m, n) = I_{ship}(m, n) + I_{sea}(m, n) \quad (12)$$

In this way, we introduce full control of both the reference image (the ship radar response, I_{ship}) and clutter (the sea radar response, I_{sea}), with the possibility of conducting parameter sensitivity tests in terms of noise intensity and ship extension and position, as detailed later in this paper. This approach allowed us to perform an exhaustive sensitivity analysis, thus avoiding (at least in this study) working with blind quality assessment approaches that are very common in the literature [37–39].

After this stage, the simulator activates two branches aimed at the generation of the training data set and computing image quality. It must be noted that high-quality training data are important for the successful use of intelligent systems. Various survey works are available in the literature regarding the estimation of the quality of an image, known as Image Quality Assessment (IQA) [40,41]. In the specific case of this work, it was decided to take as a reference the Signal-to-Clutter Ratio (SCR), which is a specific parameter used in the SAR field, largely employed in the field of ship detection [42]. It is useful in our work

to verify whether the raw data generated in the simulation are capable of producing an SAR response with the presence of an intensity peak in a noisy background.

To this aim, it was inserted in the block diagram of Figure 3, a branch dedicated to focusing the simulated raw data and the subsequent calculation of the SCR. More specifically, the first step generates the intensity image, obtained by focusing the raw data and computing the magnitude of the result:

$$I_{int}(m, n) = |I_{raw}(m, n) \otimes H_{matched}(m, n)| \quad (13)$$

Subsequently, the SCR is computed according to Formula (14).

$$SCR = 10 \log_{10} \left(\frac{P_{ship}}{N} \right) \quad (14)$$

where P_{ship} is the peak intensity of the ship return, while N is the average intensity of the sea around the ship. These two parameters are computed on the intensity image I_{int} .

There is an alternative to this expression, where P_{ship} is calculated as the average power of all the pixels belonging to the ship. This second version, however, is susceptible to estimation errors on real data since it is not possible to know precisely the extent of the ship in the image. This second expression is very common in the scientific literature; we will refer to it with the notation SCR_{AV} in this paper.

In parallel, a second branch of the block diagram extracts a crop of the simulated raw data image that contributes to populating the training data set. For this study, we looked at the effects of different cropping dimensions on raw data. Following several tests, we determined that the best performance dimension was 100×100 samples, which is significantly shorter than the impulse response. A possible explanation is that, in this way, we can concentrate the crop on the peak response, which typically contains significant energy of ship radar return, and guide the selected CNN to learn to identify this part of the signal, which is less disturbed by the background noise generated by the sea response.

Furthermore, the selection of a small crop introduces two other advantages. First, it must be considered that the CNN we will examine in the next sections of the paper is a classification network; this means that it does not specify where the ship is in the scene, but whether it is present or not. The use of small crops is, therefore, recommendable in order to reduce uncertainty about the ship's position. Second, a smaller crop requires less training and classification times. To make the simulation closer to realistic conditions, the ship position is not centered in the crop; it is moved from the center with a random shift of ± 12.5 pixels in the range and azimuth directions (az_sh and rg_sh parameters are shown in Figure 3). The goal of this shift is to improve ship detection performance even when the ship peak is not exactly centered in the crop.

2.5. Convolutional Neural Network

The goal of the Neural Network is to categorize the raw data crops into two classes: Ship and NoShip. Due to the fact that input raw data are a complex matrix, the selected CNN was configured to work with two separate real channels representing the real and imaginary parts of the input data, respectively (see Figure 7). In a preliminary analysis, we also attempted to work with only a single real channel selected among the real part, imaginary part, amplitude, and phase of the input complex data, but none of them provided satisfying results. Therefore, the use of both real and imaginary channels was the final choice for the present work, and only the results concerning this choice will be presented.

Since the main objective of the present work is to prove the feasibility of ship detection on SAR raw data by means of a CNN approach and not to test different machine learning algorithms, it was necessary to choose an algorithm with a high probability of success in the proposed experiment. To this aim, we evaluated the performance of a deep neural residual

network known as “ResNet” [43] that can be considered a state-of-the-art Convolutional Neural Network. CNNs are a class of deep learning models designed to process and extract features from visual data, such as images and videos. These networks employ convolutional layers to detect local patterns and hierarchical structures, followed by pooling layers to reduce spatial dimensions and increase computational efficiency. The fully connected layers at the end of the CNN interpret the learned features to perform classification or regression tasks. Deep convolutional neural networks, during the learning phase, use kernels that are normally randomly initialized and, in the subsequent phases of convolutions and max pooling, the kernels able to extract the information necessary to justify the class of belonging of the input crop are retained, while the others are discarded.

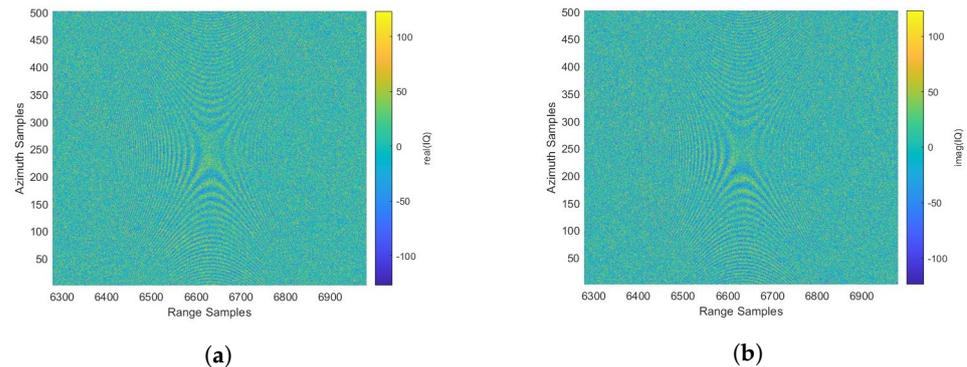


Figure 7. Real part (a) and imaginary part (b) of a simulated ship and sea raw data crop.

The network working on a generic case of the absence of the ship (null hypothesis) works only on sea images. Such images are obtained as described in [11]. No pattern is recurrent in the learning phase, so the network chooses the most similar patterns from the background noise for each crop in the training set corresponding to the null hypothesis; such kernels are not present on all training image sets, so the decision process is chosen on this basis for the null hypothesis. If a ship is present in the image, a specific pattern appears to be present in the input data and is recurrent on all the crops in the training set of the detected hypothesis; thus, max pooling extracts the kernels able to detect this kind of pattern. The convolutional neural network is useful in our application as many crops can contain patterns only in a small part of the image while the rest may be affected by different patterns that do not correspond to the class the network is learning to recognize. Such other patterns are discarded in the max pooling phase. This process is able to perform pattern recognition, enhancing the searched patterns while reducing the effects of background noise that may shade the presence of the pattern to a human eye.

ResNets use so-called residual blocks that implement shortcut connections in the network architecture [43]. The stack of convolution layers within each residual block only needs to learn a residual term that refines the input of the residual block toward the desired output. This makes the ResNet easier to train because the shortcut connections enable the direct propagation of information and gradients across multiple layers of the network, leading to better gradient flow and the convergence properties of the network during calibration. This ensures high flexibility for this CNN and increases the potential to understand more complex features. The ResNet used in this work was created and trained on Matlab. The command used to create its structure is `resnetLayers(inputSize, numClasses)`, which creates a 2D residual network with an image input size specified by `inputSize` and a number of classes specified by `numClasses`. More specifically:

- `inputSize` is a 3-element vector in the form (height, width, depth), where depth is the number of channels. In our case, `inputSize` is set to [100, 100, 2], considering that the input image crops have a size of 100×100 pixels in the form of real and imaginary parts.

- *numClasses* is an integer greater than 1. In our case, *numClasses* is set to 2, considering the need to detect Ship and NoShip classes.

The architecture parameters of the ResNet are presented in Table 3.

Table 3. ResNet architecture parameters.

Name	Description	Value
Initial filter size	Filter size in first convolutional layer.	(Height, weight) = (7, 7)
Initial num filters	Number of filters in the first convolutional layer. The number of initial filters determines the number of channels (feature maps) in the output of the first convolutional layer in the residual network.	64
Initial stride	Stride in the first convolutional layer.	(Height, weight) = (2, 2)
Initial pooling layer	First pooling layer before the initial residual block.	Max pooling
Residual block type	Specifies the location of the batch normalization layer in the standard and downsampling residual blocks.	Before the addition layer in the residual blocks.
Bottleneck type	Uses bottleneck residual blocks that perform downsampling in the first convolutional layer of the downsampling residual blocks, using a stride of 2. A bottleneck residual block consists of three convolutional layers: a 1-by-1 layer for downsampling the channel dimension, a 3-by-3 convolutional layer, and a 1-by-1 layer for upsampling the channel dimension.	Enabled
Stack depth	Number of residual blocks in each stack, specified as a vector of positive integers. Specifies the number of filters in the convolutional layers of each stack using the Num Filters argument.	(3, 4, 6, 3)
Num filters	Number of filters in convolutional layers of each stack.	(64, 128, 256, 512)
Normalization	Subtracts the mean. The mean is calculated at training time.	Enabled

2.6. Data Set and Training

The training parameters of the ResNet are presented in Table 4.

Table 4. ResNet training parameters.

Parameter	Value
Number of hidden units	177
Epochs	30
Iterations per epoch	600
Solver	Adam
Learning rate	2×10^{-4}
Mini batch size	16

The data set is composed of 12,000 simulations. Half of the data set's 12,000 total crops include ships, whereas the other half only includes a sea simulation. We generated a total of 6000 ship crops by flipping 2000 ship simulations vertically and horizontally. These crops have an SCR between 42 and 56 dB. In this first phase, we chose to train the ResNet with a high SCR to assist it in learning the essential features of a bright object's radar impulse response. A total of 80% percent of the data set is utilized for training, while the remaining crops are equally divided between validation, which is performed periodically during training, and testing, which is performed at the conclusion of the training. The data set specifications are summarized in Table 5.

Table 5. ResNet data set parameters

Parameter	Value
Ship simulations	2000 crops
Ship augmented data simulations	4000 crops
Total ship simulations	6000 crops
Sea simulations	6000 crops
Total simulations	12,000 crops
Simulations used for training	80%
Simulations used for validation	10%
Simulations used for test	10%
SCR range	42–56 dB

Under these conditions, the ResNet managed to achieve high accuracy during the training phase with the simulated raw data, with several iterations exceeding 95% and up to 100% accuracy, as illustrated in Figure 8.

After the training phase, the test data set composed of 1200 crops (600 ships and 600 sea crops) was used to check the final performance of the ResNet. This test phase was aimed to assess the performance of the ResNet with an independent set of simulated data not used in the training phase. All crops were correctly classified, reaching an overall accuracy of 100%. However, to provide a realistic and quantitative performance of the ResNet, further test phases were needed that also involved real data, as detailed in the following section.

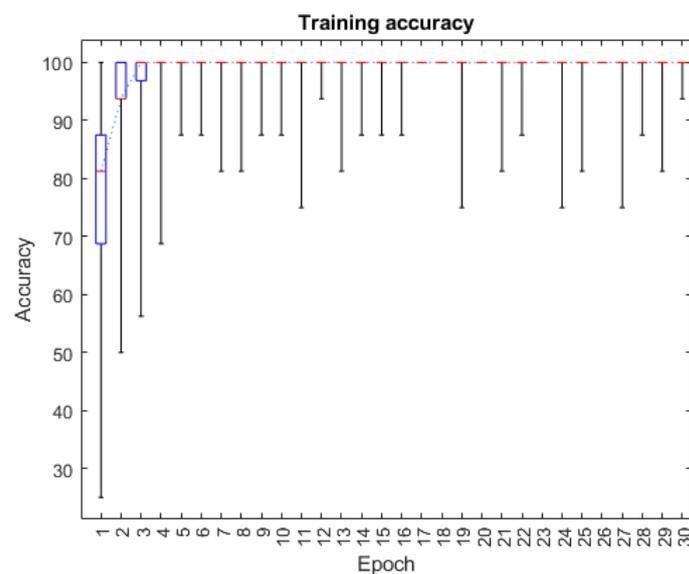


Figure 8. Training accuracy boxplot for each epoch.

3. Results

Our goal was to assess the ResNet performance by using both real and simulated ERS raw data. In this way, we were able to simultaneously evaluate the quality of our simulation model, as well as the performance of the trained ResNet in real scenarios.

3.1. Test with Single-Point Ship Simulated Data

This test phase was aimed to assess the ResNet performance with another independent set of simulated data not used in the previous training and test phases. A single-point model for ship simulation was adopted with different conditions of SCR values and range/azimuth shifts. This gave us important insights into ResNet's robustness and sensitivity to a wide range of conditions. More specifically, we investigated the detection performance as a function of SCR and shift values by further simulating 29 ship crops with an SCR varying from 15 to 45 dB. Starting from these simulations, we generated 2349 crops by applying a combination of 81 progressive azimuth and range shifts to the ship position with respect to the center of the crop. The size of the simulated crops is 100×100 samples in accordance with the training data set, while the range of ship positions was extended to ± 80 samples and then more than 6 times higher than the ± 12.5 range used in the training phase. This was conducted as a preliminary attempt to test the performance of the network beyond the limits it was trained with.

Under these conditions, we evaluated the ResNet's classification result for each crop. The classification accuracy can be analyzed by means of the scatter plot presented in Figure 9. It can be observed that the ResNet begins to detect the ship class with SCR > 33 dB. With higher SCR values, it is also able to tolerate a shift distance of <50 pixels with respect to the center of the crop.

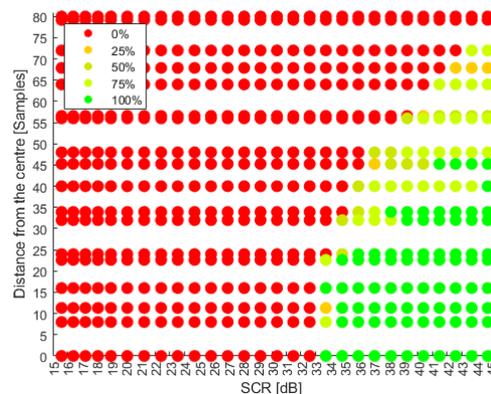


Figure 9. Accuracy of the ResNet with one-point ship simulated crops with different combinations of SCR and shifts from the crop center.

We also investigated the impact of the shift direction. Due to the different shapes of the impulse response along the azimuth and range directions, a different performance is expected when a ship shift is applied separately along these two axes. To this aim, scatter plots in Figures 10 and 11 illustrate the classification accuracy when the shift to the ship position is applied in one direction only, with a zero shift in the other direction. The analysis of both plots shows that the proposed ResNet is more robust to azimuth shifts. This different performance will be further investigated in future works.

Finally, the performance related to false ship detection was evaluated. To this aim, 1000 crops containing only the SAR response of the sea were simulated with the same size as ship crops. The ResNet was able to correctly classify all the simulated crops, thus obtaining 100% accuracy in the detection of the sea class.

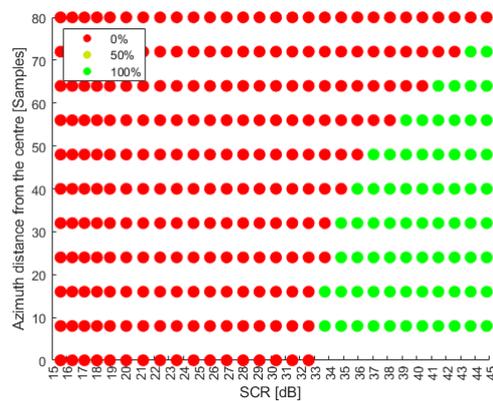


Figure 10. Classification accuracy as a function of SCR and shifts applied to simulated data along the azimuth direction only.

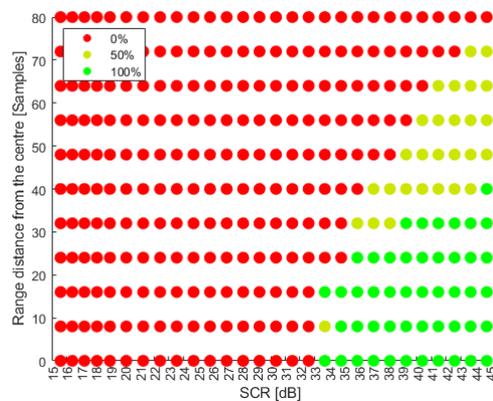


Figure 11. Classification accuracy as a function of SCR and shifts applied to simulated data along the range direction only.

3.2. Test with Multi-Point Ship Simulated Data

The results obtained using the ResNet, in the case of single-point simulations, allowed us to confirm the ResNet's ability to detect the presence of a ship using raw data. At first glance, the performance obtained in terms of the SCR may not appear particularly challenging. However, it should be considered that the resolution of the ERS data is equal to $5 \text{ m} \times 20 \text{ m}$, which means that the single-point response refers to very small vessels for which the literature is still exploring various solutions. Among the most promising approaches, we mention the use of multi-polarization sensors, which are not the subject of this analysis since the ERS mission operated exclusively in VV polarization. In this regard, in [42], it is noted that in the case of small ships, the combined use of multiple polarizations leads to a detection capacity in the range between 38% (with the GNF method) and 85% (with the Λ_M method). In particular, the Λ_M method achieves the best performance with an SCR of 18.3 dB. However, the performances achieved by it do not correspond to single-point ships but always involve a limited extension of pixels, which can favor the application of SCR enhancement before detection.

As observed, an analysis was, therefore, necessary to evaluate the impact of the extension of the ship on the performance of the detection algorithm. There are two contrasting aspects concerning this observation. First of all, it can be immediately observed that in the presence of several pixels associated with a ship, interference between the various responses is expected. This is confirmed via the graph illustrated in Figure 12, in which it is evident that by increasing the number of pixels associated with a ship, there is a consequent drop in the SCR value.

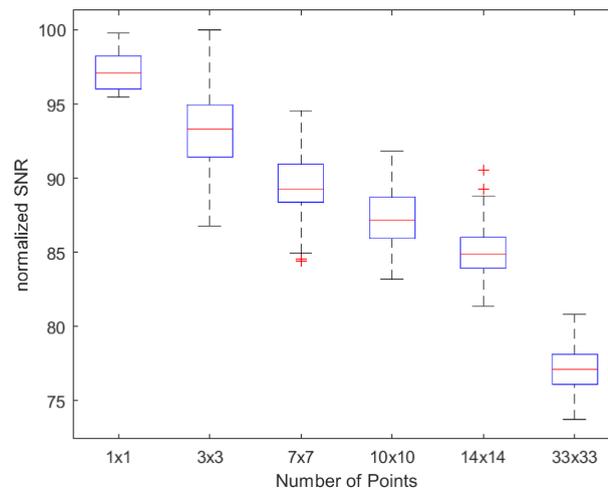
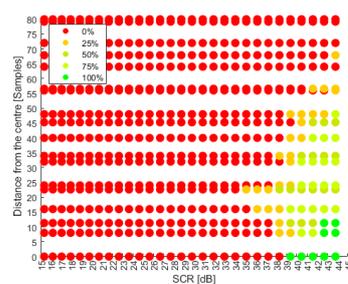


Figure 12. SCR measured with an increasing number of points used to simulate radar response of ships with different extensions. Each box shows the statistics of 100 tests with the same ship extension but with a random point position in each resolution cell.

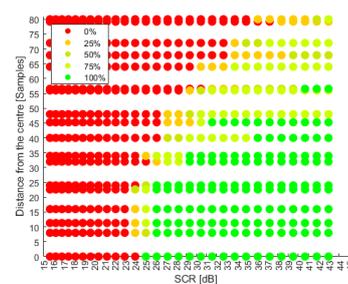
However, the availability of more pixels per ship enables the generation of multiple echoes returned by the ship and, consequently, a greater overall energy received by the radar system.

In order to investigate how these two contrasting aspects impact the performance of the ResNet, several simulations were performed with ships having dimensions of 3×3 , 7×7 , 10×10 , 14×14 , and 33×33 pixels, corresponding to ground extensions of 15×60 , 35×140 , 50×200 , 70×280 , and 165×660 m². From an implementation point of view, it was sufficient to iterate the single-point simulator in different positions and sum the results in order to generate a multi-point ship radar response. Not all dimensions are realistic for a ship, but they have been taken into account to provide a sufficiently comprehensive picture of the ResNet's performance in the case of a multi-point model.

It is observed that although the ResNet was trained with a single-point model, it was able to produce optimal performance when simulated ships have a size of 7×7 pixels. Under these conditions, it can be observed that the ResNet begins to correctly detect the ship class with $\text{SCR} > 24$ dB. With higher SCR values, it is also able to tolerate a shift distance of < 60 pixels with respect to the center of the crop. This performance is well beyond the limits measured with a single-point ship simulation (resulting in an $\text{SCR} > 33$ dB and a maximum shift distance of < 50 pixels). According to Figure 13, it can be inferred that the 7×7 -pixel size is a sort of sweet spot for ERS-based simulations. Beyond this size, the ResNet performance begins to decrease, probably due to the presence of pixels with a considerable inter-distance that introduces a strong deviation from the expected single-point ship impulse response used to train the ResNet.



(a) 3×3 -pixel simulated ship



(b) 7×7 -pixel simulated ship

Figure 13. Cont.

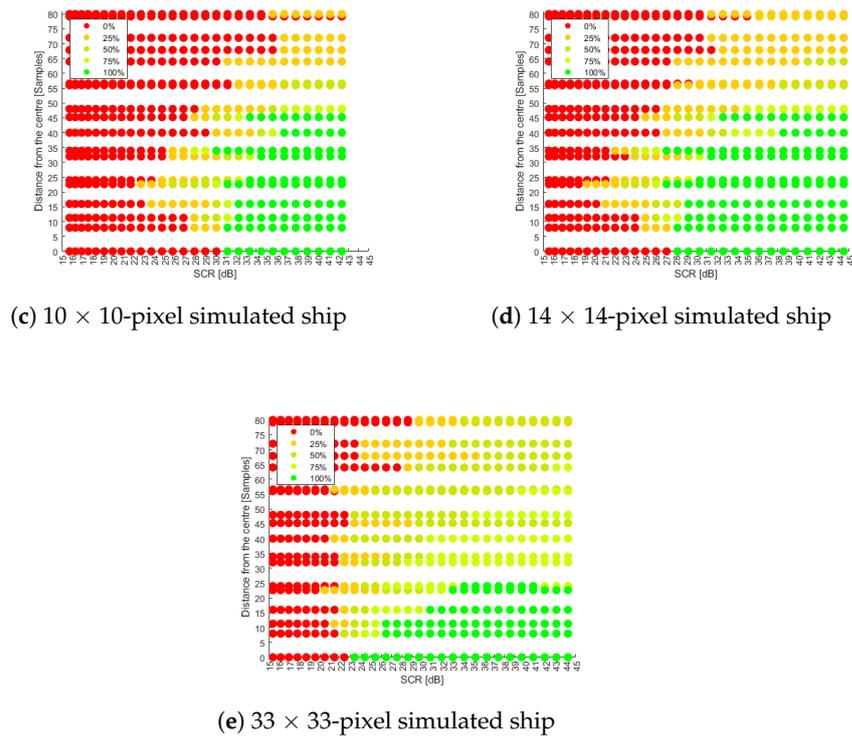


Figure 13. Classification accuracy as a function of SCR and shift in the ship position (multi-point ship simulated data).

3.3. Test with ERS Data

Finally, we conducted a thorough test of ResNet’s accuracy in ship classification using real ERS data. The first analysis consisted of verifying whether the performance measured on simulated data can be replicated on real data. To this aim, seven crops were extracted from ERS images corresponding to ships exhibiting an SCR between 24.8 dB and 40.1 dB. We applied to each crop progressive shifts to ship positions spanning ±80 samples along the azimuth and range directions. The results are illustrated in the scatter plot of Figure 14, where it is also indicated the size in pixels of each ship in order to analyze the impact of this parameter on the ResNet detection accuracy. In particular, it is observed that the ResNet begins to operate correctly starting from an SCR equal to 26 dB.

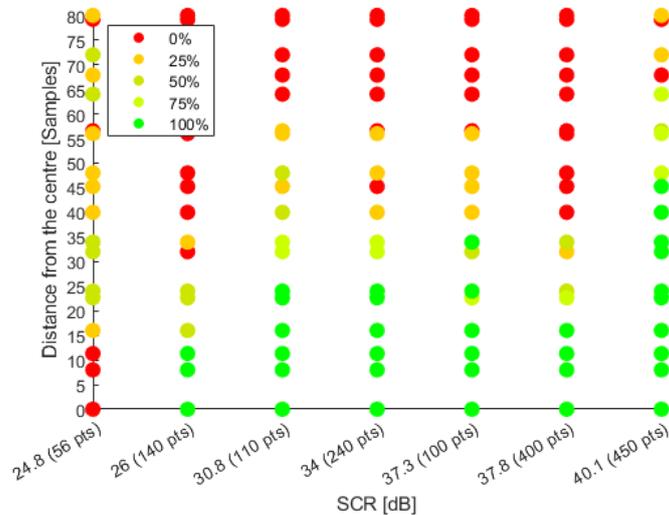


Figure 14. Classification accuracy as a function of SCR and shift in the ship position (real data). Ship size is indicated in parentheses next to the SCR value.

It is possible to compare these performances with previous works available in the literature. To this aim, it was necessary to estimate the SCR values by computing the P_{ship} parameter as the average power of all the pixels belonging to the ship, obtaining the correspondences indicated in the Table 6.

Table 6. SCR measured on real cases by computing the two alternative expressions of (14).

SCR (dB)	SCR _{AV} (dB)
24.8	8.2
26.0	10.3
30.8	23.6
34.0	23.4
37.3	32.3
37.8	20.3
40.1	37.6

It is observed that the use of the average power in Expression (14) significantly reduces the estimated SCR values. In particular, it is observed that the network begins to work correctly starting from an SCR_{AV} of 10.3 dB. This is in agreement with the results reported in [44], where it is observed that the CFAR algorithm reaches a detection capability of 29.6% with an SCR_{AV} of 5.7 dB and of 81.6% with an SCR_{AV} of 10.9 dB.

The performances of the ResNet with real data also exhibit several points of good agreement with those measured in the simulation phase by means of the scatter plots of Figure 13. More specifically, we observed the following:

- The ship with an SCR equal to 24.8 dB (having a size of 56 pixels) was not correctly detected via the ResNet. This is in good agreement with the performance measured with the 7×7 -pixel ship simulation.
- The ship with an SCR equal to 30.8 dB (having a size of 110 pixels) was correctly detected via the ResNet with shifts lower than 40–45 pixels. This is in good agreement with the performance measured with the 10×10 -pixel ship simulation.
- The ship with an SCR equal to 34.0 dB (having a size of 240 pixels) was correctly detected via the ResNet with shifts lower than 40 pixels. This is in acceptable agreement with the performance measured with the 14×14 -pixel ship simulation.
- Even with the real data, there is some evidence of a potential sweet spot in the optimal ship size. In the case of the simulated data set, a 7×7 -pixel sweet spot was identified. The reduced set of real data did not allow us to perform an exact estimation of the sweet spot; however, there are some practical considerations that lead us to infer that it should be positioned close to the one identified in the simulation phase. In particular, the following can be observed: (1) the ship with a size of 110 pixels and an SCR equal to 30.8 dB achieves the ResNet classification performances of the ship with an SCR equal to 34 dB, which has a larger size (240 pixels); (2) the two ships with similar SCR values (37.3 dB and 37.8 dB) exhibit very different performances, showing that the ship with a smaller size (100 pixels) has better performance than the ship with a 400-pixel size.

This confirms that our multi-point ship simulation methodology accurately represents the main characteristics of real ships and their interactions with clutter. Furthermore, this successful test with real data also confirms the validity of our single-point ship simulation approach for generating the simulated data, and it proves suitable for making the ResNet operative on real data.

However, the test with real data reveals some deviations in ResNet performance compared to that obtained with simulated data. In particular, it is not well understood why the real ship with an SCR equal to 26 dB exhibits better performances than those measured with simulated data. Future works will allow us to identify these anomalies.

Further analyses were instead conducted regarding the behavior of the ResNet with real data as a function of the direction of the shifts applied to the position of the ship with respect to the center of the crop. More specifically, Figures 15–17 show crops of ERS images extracted from three ships with different SCR values. For each crop, a graphic indication is provided on the shifts applied in the azimuth and range directions and the outcome of the classification corresponding to each shift:

- The detection accuracy associated with shifts applied along the azimuth direction is better than that relating to shifts along the range direction, and this is in agreement with what has already been observed in the simulation;
- The impulse response of the ship is visually recognizable only in the raw data associated with the third case (SCR, 40.1 dB), while with lower SCR values, this response is not visible, although the ResNet still demonstrates the capability to recognize the presence of the ship. Furthermore, this motivated why visual attention techniques were not considered in the first feasibility analysis.

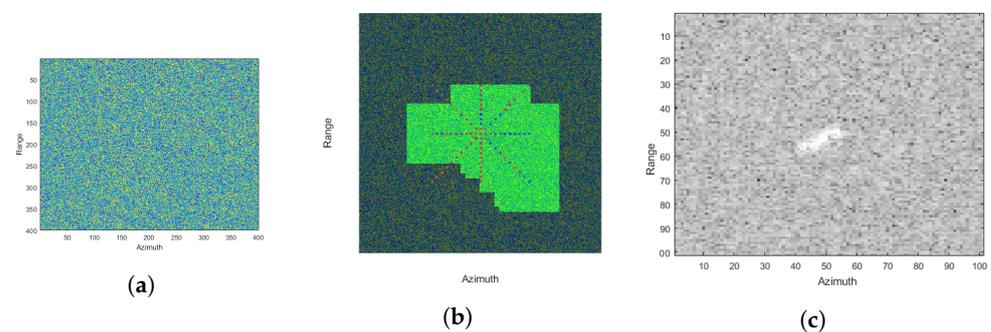


Figure 15. Ship with SCR = 24.8 dB. (a) The real part of raw data; (b) each blue point represents the center of the crop that is correctly detected via the ResNet (the relevant crop is colored green); (c) the SLC amplitude.

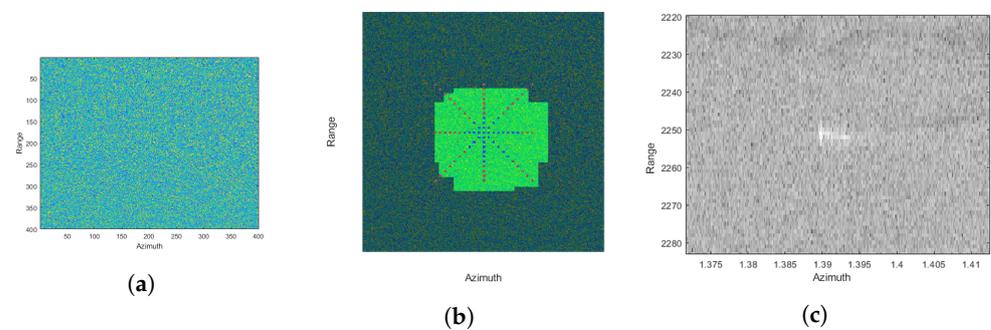


Figure 16. Ship with SCR = 30.8 dB. (a) The real part of raw data; (b) each blue point represents the center of the crop that is correctly detected via the ResNet (the relevant crop is colored green); (c) the SLC amplitude.

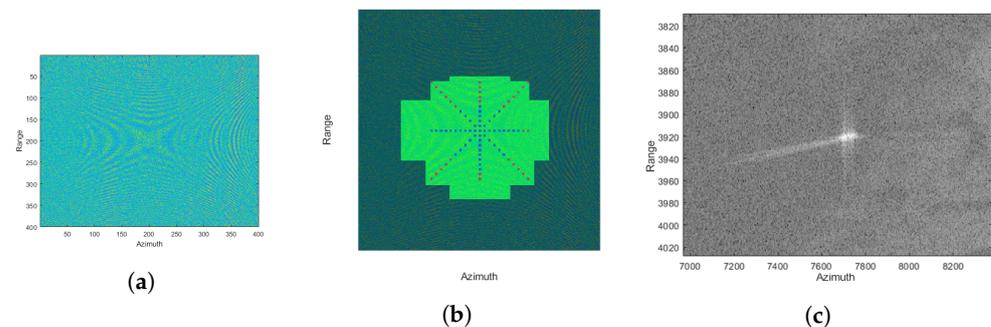


Figure 17. Ship with SCR = 40.1 dB. (a) The real part of raw data; (b) each blue point represents the center of the crop that is correctly detected via the ResNet (the relevant crop is colored green); (c) the SLC amplitude.

Finally, similarly to what was performed in the simulation, the performance related to false detection was evaluated. To this aim, 3712 crops containing only the SAR response of the sea were extracted from ERS images. The ResNet returned only 40 false positives for an overall accuracy of 98.9%. A visual analysis of the input data associated with these wrongly classified crops did not allow us to identify any interesting clues that could motivate the wrong choices made via the ResNet (see Figure 18). This prevented us from making changes in our tool for the simulation of sea crop data sets in order to improve the sea classification performance of the ResNet.

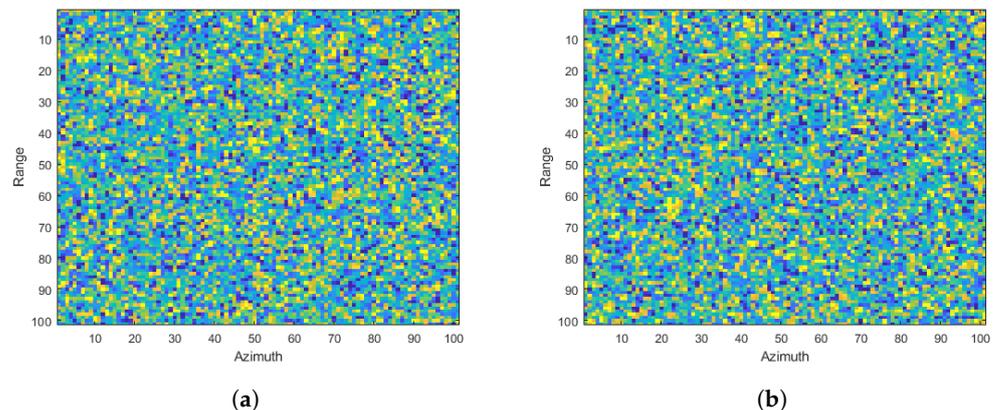


Figure 18. Real (a) and imaginary (b) parts of a sea raw data crop incorrectly classified as a ship via the ResNet.

Furthermore, this highlights the limitations of present simulation methodologies in mimicking specific natural situations and the necessity for additional research and breakthroughs in simulating complicated marine settings for superior ResNet training and testing in ship detection applications. At the same time, this leads us to reconsider, in future works, the possibility of adopting a hybrid data set, including, within the simulated data set, real cases that the ResNet has not correctly classified.

4. Conclusions

In this study, we have proposed a complete framework aimed at the detection of ships applied directly to SAR raw data by means of a ResNet. The following is a summary of our findings:

- We have proposed a simple and automatic simulation method for generating the training data set.
- The simulator produces sea/ship simulations of radar echoes integrated into complex matrices that represent the simulated raw data set.
- Ships are simulated by means of a single-point model that can be easily extended to a multi-point model.
- The sea is simulated by means of a statistical model of the sea directly derived from the distribution of raw data values extracted from ERS mission.
- A state-of-the-art ResNet CNN was configured with the aim of categorizing 100×100 -pixel crops into two classes (Ship and NoShip).
- During the training phase, the ship position was moved randomly up to the 12.5 range and azimuth samples with respect to the center of the crop. Under these simulation conditions, the ResNet was able to achieve 100% overall accuracy with an SCR in the range of 42–56 dB.
- More realistic conditions were investigated by (i) using a multi-point model for ship simulation and (ii) ship shifts of up to 80 range and azimuth samples with respect to the center of the crop. Under these simulation conditions, ships with a 7×7 -pixel size were identified as a sweet spot for ERS-based simulations. Under these conditions, it can be observed that the ResNet begins to correctly detect the ship class with

SCR > 24 dB with small shifts of the ship position. With higher SCR values, it is also able to tolerate a shift distance of <60 pixels with respect to the center of the crop. This performance is well beyond the limits measured with a single-point ship simulation (resulting in an SCR >33 dB and a maximum shift distance of <50 pixels).

- We also investigated the impact of the shift direction. Due to the different shapes of the impulse response along the azimuth and range directions, different performances are expected for shifts applied to the azimuth and range directions. In particular, the proposed ResNet exhibits more robustness to azimuth shifts.
- Final analyses were carried out with ERS real data. In particular, the ResNet begins to work correctly with an SCR_{AV} equal to 10.3 dB, which is in good agreement with the performance reported in the scientific literature. There is also a good agreement between the performance with simulated and real data. This confirms that our ship simulation methodology accurately represents the main characteristics of real ships and their interactions with clutter.

Starting from the results obtained, the following future investigations are foreseen:

- A further analysis of different ResNet performances in shifts applied along the azimuth and range directions.
- A further analysis of the ResNet performance as a function of the number of pixels (and their distribution) associated with the ship.
- Check whether the combined use of the amplitude and phase channels is a possible better alternative to the use of the real and imaginary channels.
- Check if the configured network is operating under overfitting conditions, given the higher performance obtained in the training phase.
- An analysis of the false positive classification of sea crops extracted from real ERS data.
- An analysis of the computational cost of the proposed ResNet.
- As proposed by one of the reviewers, it is also important to improve the breadth of experiments by providing a well-explored experimental setting based on performance comparisons with classic ML methods as well as VGG Very Deep Convolutional Networks and Densely Connected Convolutional Networks.

Author Contributions: Conceptualization, A.M., C.G. and R.N.; methodology, A.M., G.C. and K.T.; software, A.M., G.C., D.O.N., N.R. and K.T.; validation, C.G., G.C. and R.N.; resources, D.O.N.; data curation, G.C., N.R. and K.T.; writing—original draft preparation, G.C.; writing—review and editing, G.C. and R.N.; supervision, C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Advanced Payload Data Processing for Autonomy & Decision (APP4AD) project: contract with the Italian Space Agency (n. 2021-6-E.0), Codice Unico di Progetto (CUP) (F95F21000020005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data set used to train the ResNet can be downloaded at <http://www.gapsrl.eu/gapsrl/ricprog.php> (accessed on 30 August 2023).

Conflicts of Interest: Giorgio Cascelli, Raffaele Nutricato, Khalid Tijani, Alberto Morea, Nicolò Ricciardi, Davide Oscar Nitti are employed by the company Geophysical Applications Processing G.A.P. s.r.l. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CNN	Convolutional Neural Network
EC	Edge Computing
EO	Earth Observation
ERS	European Remote Sensing
ESA	European Space Agency
MDPI	Multidisciplinary Digital Publishing Institute
ML	Machine Learning
ResNet	Residual Neural Network
SAR	Synthetic Aperture Radar
SCR	Signal-to-Clutter Ratio
SLC	Single Look Complex

References

- Dunne, J. The SEASAT-A Project: An Overview. In Proceedings of the OCEANS '76, Washington, DC, USA, 13–15 September 1976; pp. 217–221. [\[CrossRef\]](#)
- Amitrano, D.; Di Martino, G.; Guida, R.; Iervolino, P.; Iodice, A.; Papa, M.N.; Riccio, D.; Ruello, G. Earth Environmental Monitoring Using Multi-Temporal Synthetic Aperture Radar: A Critical Review of Selected Applications. *Remote Sens.* **2021**, *13*, 604. [\[CrossRef\]](#)
- Bamler, R. Earth Environmental Monitoring Using Multi-Temporal Synthetic Aperture Radar: A Critical Review of Selected Applications. *Surv. Geophys.* **2000**, *21*, 147–157. [\[CrossRef\]](#)
- Gens, R.; van Genderen, J. SAR interferometry—Issues, techniques, applications. *Int. J. Remote Sens.* **1996**, *17*, 1803–1835. [\[CrossRef\]](#)
- Goldstein, R.M.; Zebker, H.A. Interferometric radar measurement of ocean surface currents. *Nature* **1987**, *329*, 707–709. [\[CrossRef\]](#)
- Guaragnella, C.; D.T. A Data-Driven Approach to SAR Data-Focusing. *Sensors* **2019**, *19*, 1649. [\[CrossRef\]](#)
- Bamler, R. A comparison of range-Doppler and wavenumber domain SAR focusing algorithms. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 706–713. [\[CrossRef\]](#)
- Cafforio, C.; Prati, C.; Rocca, F. SAR data focusing using seismic migration techniques. *IEEE Trans. Aerosp. Electron. Syst.* **1991**, *27*, 194–207. [\[CrossRef\]](#)
- Amein, A.S.; Soraghan, J.J. Fractional Chirp Scaling Algorithm: Mathematical Model. *IEEE Trans. Signal Process.* **2007**, *55*, 4162–4172. [\[CrossRef\]](#)
- Freeman, A. SAR calibration: An overview. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 1107–1121. [\[CrossRef\]](#)
- Lee, J.S. Speckle analysis and smoothing of synthetic aperture radar images. *Comput. Graph. Image Process.* **1981**, *17*, 24–32. [\[CrossRef\]](#)
- Chang, Y.L.; Anagaw, A.; Chang, L.; Wang, Y.C.; Hsiao, C.Y.; Lee, W.H. Ship Detection Based on YOLOv2 for SAR Imagery. *Remote Sens.* **2019**, *11*, 786. [\[CrossRef\]](#)
- Zhang, G.; Li, Z.; Li, X.; Yin, C.; Shi, Z. A Novel Salient Feature Fusion Method for Ship Detection in Synthetic Aperture Radar Images. *IEEE Access* **2020**, *8*, 215904–215914. [\[CrossRef\]](#)
- Topouzelis, K.N. Oil Spill Detection by SAR Images: Dark Formation Detection, Feature Extraction and Classification Algorithms. *Sensors* **2008**, *8*, 6642–6659. [\[CrossRef\]](#)
- Kerr, M.; Tonetti, S.; Cornara, S.; Bravo, J.-I.; Hinz, R.; Latorre, A.; Membibre, F.; Solimini, C.; Wiehle, S.; Breit, H.; et al. EO-ALERT: A Novel Architecture for the Next Generation of Earth Observation Satellites Supporting Rapid Civil Alerts. In Proceedings of the 71st International Astronautical Congress (IAC 2020), Bremen, Germany, 12–14 October 2020.
- Leyva-Mayorga, I.; Martinez-Gost, M.; Moretti, M.; Pérez-Neira, A.; Vázquez, M.Á.; Popovski, P.; Soret, B. Satellite Edge Computing for Real-Time and Very-High Resolution Earth Observation. *IEEE Trans. Commun.* **2023**, *71*, 6180–6194. [\[CrossRef\]](#)
- Bianco, S.; Cadene, R.; Celona, L.; Napoletano, P. Benchmark Analysis of Representative Deep Neural Network Architectures. *IEEE Access* **2018**, *6*, 64270–64277. [\[CrossRef\]](#)
- Giuffrida, G.; Diana, L.; de Gioia, F.; Benelli, G.; Meoni, G.; Donati, M.; Fanucci, L. CloudScout: A Deep Neural Network for On-Board Cloud Detection on Hyperspectral Images. *Remote Sens.* **2020**, *12*, 2205. [\[CrossRef\]](#)
- Denby, B.; Lucia, B. Orbital Edge Computing: Nanosatellite Constellations as a New Class of Computer System. In Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems, Lausanne, Switzerland, 16–20 March 2020.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *arXiv* **2015**, arXiv:cs.CV/1502.01852.
- Liu, X.; He, P.; Chen, W.; Gao, J. Multi-Task Deep Neural Networks for Natural Language Understanding. *arXiv* **2019**, arXiv:cs.CL/1901.11504.

22. Zhang, Y.; Hao, Y. A Survey of SAR Image Target Detection Based on Convolutional Neural Networks. *Remote Sens.* **2022**, *14*, 6240. [[CrossRef](#)]
23. Min, X.; Zhai, G.; Gu, K.; Zhu, Y.; Zhou, J.; Guo, G.; Yang, X.; Guan, X.; Zhang, W. Quality Evaluation of Image Dehazing Methods Using Synthetic Hazy Images. *IEEE Trans. Multimed.* **2019**, *21*, 2319–2333. [[CrossRef](#)]
24. Min, X.; Zhai, G.; Zhou, J.; Zhang, X.P.; Yang, X.; Guan, X. A Multimodal Saliency Model for Videos with High Audio-Visual Correspondence. *IEEE Trans. Image Process.* **2020**, *29*, 3805–3819. [[CrossRef](#)] [[PubMed](#)]
25. Gao, F.; Liu, A.; Liu, K.; Yang, E.; Hussain, A. A novel visual attention method for target detection from SAR images. *Chin. J. Aeronaut.* **2019**, *32*, 1946–1958. [[CrossRef](#)]
26. Breck, E.; Polyzotis, N.; Roy, S.; Whang, S.; Zinkevich, M. Data Validation for Machine Learning. In Proceedings of the MLSys, Stanford, CA, USA, 31 March–2 April 2019.
27. Jain, A.; Patel, H.; Nagalapatti, L.; Gupta, N.; Mehta, S.; Guttula, S.; Mujumdar, S.; Afzal, S.; Sharma Mittal, R.; Munigala, V. Overview and Importance of Data Quality for Machine Learning Tasks. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 6–10 July 2020; pp. 3561–3562. [[CrossRef](#)]
28. Rostami, M.; Kolouri, S.; Eaton, E.; Kim, K. Deep Transfer Learning for Few-Shot SAR Image Classification. *Remote Sens.* **2019**, *11*, 1374. [[CrossRef](#)]
29. Zhong, C.; Mu, X.; He, X.; Wang, J.; Zhu, M. SAR Target Image Classification Based on Transfer Learning and Model Compression. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 412–416. [[CrossRef](#)]
30. Huang, Z.; Dumitru, C.O.; Pan, Z.; Lei, B.; Datcu, M. Classification of Large-Scale High-Resolution SAR Images With Deep Transfer Learning. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 107–111. [[CrossRef](#)]
31. Malmgren-Hansen, D.; Kusk, A.; Dall, J.; Nielsen, A.A.; Engholm, R.; Skriver, H. Improving SAR Automatic Target Recognition Models With Transfer Learning From Simulated Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1484–1488. [[CrossRef](#)]
32. ESA. Available online: <https://earth.esa.int/eogateway/instruments/sar-ers/description> (accessed on 30 August 2023).
33. Eldhuset, K. An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 1010–1019. [[CrossRef](#)]
34. Solberg, A.; Storvik, G.; Solberg, R.; Volden, E. Automatic detection of oil spills in ERS SAR images. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1916–1924. [[CrossRef](#)]
35. SAR (ERS) on ESA Website. Available online: <https://earth.esa.int/eogateway/instruments/sar-ers> (accessed on 31 July 2023).
36. Gao, G. Statistical modeling of SAR images: A survey. *Sensors* **2010**, *10*, 775. [[CrossRef](#)]
37. Min, X.; Gu, K.; Zhai, G.; Liu, J.; Yang, X.; Chen, C.W. Blind Quality Assessment Based on Pseudo-Reference Image. *IEEE Trans. Multimed.* **2018**, *20*, 2049–2062. [[CrossRef](#)]
38. Min, X.; Zhai, G.; Gu, K.; Liu, Y.; Yang, X. Blind Image Quality Estimation via Distortion Aggravation. *IEEE Trans. Broadcast.* **2018**, *64*, 508–517. [[CrossRef](#)]
39. Min, X.; Ma, K.; Gu, K.; Zhai, G.; Wang, Z.; Lin, W. Unified Blind Quality Assessment of Compressed Natural, Graphic, and Screen Content Images. *IEEE Trans. Image Process.* **2017**, *26*, 5462–5474. [[CrossRef](#)] [[PubMed](#)]
40. Zhai Guangtao, M.X. Perceptual image quality assessment: A survey. *Sci. China Inf. Sci.* **2020**, *63*, 211301. [[CrossRef](#)]
41. Min, X.; Zhai, G.; Zhou, J.; Farias, M.C.Q.; Bovik, A.C. Study of Subjective and Objective Quality Assessment of Audio-Visual Signals. *IEEE Trans. Image Process.* **2020**, *29*, 6054–6068. [[CrossRef](#)]
42. Liu, G.; Zhang, X.; Meng, J. A Small Ship Target Detection Method Based on Polarimetric SAR. *Remote Sens.* **2019**, *11*, 2938. [[CrossRef](#)]
43. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
44. Lang, H.; Shi, H.; Tao, Y.; Ma, L. A comparative study on methods of improving SCR for ship detection in SAR image. *Proc. Spie* **2017**, 10432, 104320M. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.