Training Data Selection Strategy for CFAR Ship Detection in Range-Compressed Radar Data

Sushil Kumar Joshi
Microwaves and Radar Institute
German Aerospace Center (DLR)
Wessling, Germany
Sushil.Joshi@dlr.de

Stefan V. Baumgartner
Microwaves and Radar Institute
German Aerospace Center (DLR)
Wessling, Germany
Stefan.Baumgartner@dlr.de

***Abstract*—In the paper a novel ship detection algorithm using range-compressed airborne radar data is proposed. Ships are detected in range-Doppler domain using CFAR (constant false alarm rate) based sea clutter models. The advantage of this domain is that even ships with low radar cross sections (RCS) can be detected when they move with a certain line-of-sight (LOS) velocity so that they are shifted to the exo-clutter region. For estimating accurately the parameters of the sea clutter models and for computing a true CFAR threshold, the training data has to be selected carefully and has to be free from any targets for avoiding estimation biases. Therefore, an automatic ocean training data extraction procedure is proposed. A novel pre-detection algorithm cancels unwanted targets from the training data. The applicability and performance of the proposed algorithm are demonstrated with experimental radar data acquired with DLR's airborne radar sensor F-SAR. Since the algorithm requires comparatively low processing effort, it is principally suitable for real time applications.**

Keywords—Radar clutter, oceanography, radar detection, synthetic aperture radar (SAR).

I. INTRODUCTION

Ship detection and monitoring from a real time perspective shall ensure maritime traffic safety. AIS (Automatic identification systems) and ground based systems have been extensively used to serve this purpose [1]. However, many ships, especially the smaller ones, aren't equipped with (range limited) AIS transceivers. SAR (synthetic aperture radar) remote sensing due to its day night acquisition and all weather independent capabilities is an effective way to overcome the difficulties in aforementioned systems. Ship detection using SAR images is an established method and the existing literature is enormous [2]. One of the approaches to detect ships is to model the backscatter from the ocean, known as *sea clutter,* using suitable clutter models and applying a computed CFAR detection threshold [3]. Considering a SAR image, pixels exceeding the detection threshold are declared as potential targets.

So far most of the investigations on clutter models have been carried out over fully focused SAR images at shallow incidence angle. However, range-compressed (RC) radar data can also be considered as a suitable choice for real time ship monitoring since no comprehensive and time consuming SAR processing is required. In addition, a transformation of the data to range-Doppler domain may improve the detection capability of the ships, especially if they move with a certain line-of-sight (LOS) velocity so that they are shifted in Doppler away from the clutter region [4]–[6]. To obtain a valid CFAR detection threshold in Doppler domain, ocean statistics have to be described accurately. Bright targets and high clutter peaks contaminate the training data and may cause a significant bias in the estimated ocean statistics, especially for low CFARs.

In this paper, an approach for automatic extraction of ocean training data for CFAR ship detection is proposed. Additionally various sea clutter models are compared in range-Doppler domain in terms of false alarm rate mismatch and CFAR threshold error. The proposed training data selection strategy and the clutter models are validated using real single-channel RC radar data acquired with DLR's F-SAR sensor [7].

II. PRINCIPLE OF RANGE-DOPPLER SHIP DETECTION

In Fig. 1 the overall principle of our proposed ship detection method is illustrated. It starts in time-domain (top) with the successive extraction of range-compressed data patches, which have only a small extension in azimuth direction. Thus, the coherent processing interval (CPI) for the followed fast Fourier transform (FFT) along azimuth (denoted as *Azimuth FFT*) is kept very short. Due to the short CPI the linear phase of the azimuth signals is dominant so that potential target peaks appear well focused (but not in high-resolution) in range-Doppler domain. For applying later on a single range and Doppler independent CFAR threshold, the data in range-Doppler has to be normalized (cf. Fig. 1 right). This leads to a flat clutter spectrum. If prior pre-detection was carried out and if potential target peaks were removed, the flat Doppler spectrum can also be used as training data for estimating the parameters of suitable clutter models and for computing the CFAR detection threshold. The detected target pixels are clustered to physical objects by using a DBSCAN algorithm (cf. Fig. 1 bottom right). Afterwards tracking of the cluster centroids can be carried out for reconstructing the ship track over longer observation times (Fig. 1 bottom left), which for circular airborne radar flight tracks can be in the order of several minutes (cf. Fig. 8 and [5]).

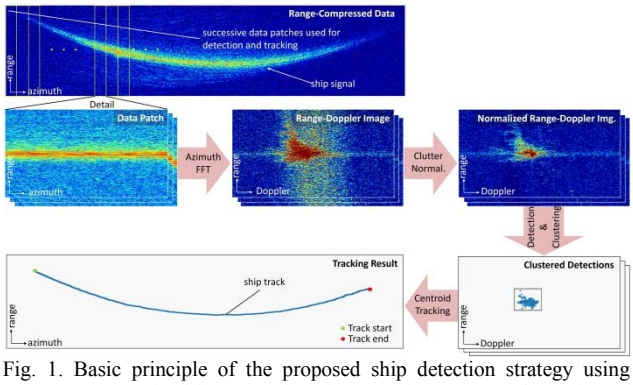

Fig. 1. Basic principle of the proposed ship detection strategy using range-compressed airborne radar data as input. For the sake of completeness, clustering and tracking results are added.

978-1-7281-2660-9/19/$31.00©2019 European Union

Further details on the proposed method is found in [8].

Outliers in the form of bright target peaks and high clutter peaks degrade the performance of the chosen clutter models and may lead to erroneous detection threshold which do not lead to a CFAR anymore. Therefore, proper training data selection, pre-detection and cancellation of such disturbing peaks is a very crucial step before performing clutter normalization and clutter model parameter estimation.

III. TRAINING DATA SELECTION

A. *Target pre-detection*

A valid and bias-free CFAR detection threshold can only be estimated if a suitable clutter model is available and if the considered ocean data patch is free of ship targets, which in most cases appear brighter than the clutter background in the radar data [9], [10], as depicted exemplarily in Fig. 2.

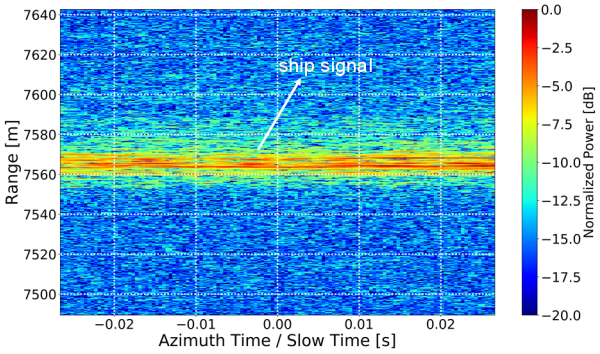

Fig. 2. Presence of a bright ship signal in an X-band HH polarized single channel RC data acquired with DLR's airborne sensor F-SAR. The shown image patch is composed of 512 range and 128 azimuth samples.

To better illustrate the negative influence of a bright ship target, the estimated probability density functions (PDFs) of the ocean only and the ocean including a ship target signal are shown in Fig. 3.

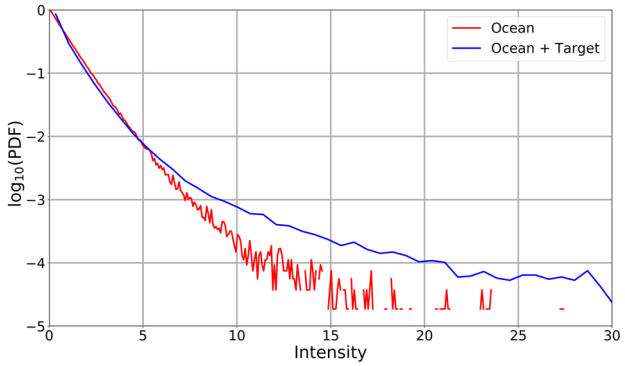

Fig. 3. PDFs of a RC data patch containing a ship target (blue) and the patch containing only the ocean (red). The PDFs were derived from the data shown in Fig. 2. For visualization purposes, the intensity axis is truncated as the maximum intensity due to the ship is around 600.

Comparing both the PDFs from Fig. 3, it is clear how the presence of a target distorts the shape of the histogram (blue). It is skewed towards higher intensity values. This causes a significant bias to the background ocean statistics. Therefore, it has to be ensured that the data used for clutter statistics estimation and clutter model fit, the so called *Training Data*, is free of targets. Potential targets have firstly to be pre-detected and removed from the data.

Several methods of CFAR detectors were proposed in the past to eliminate such potential bright peaks [11]. Very recently proposed methods based on truncated statistics or iterative censoring, are also discussed in the literature. However, they are computationally inefficient [10].

We propose a rather simple but effective method to exclude potential target peaks. This method operates in time-domain (i.e., before performing the *Azimuth FFT*) and estimates an adaptive threshold which varies along range. Deriving such a range dependent threshold is important to take into account the range and incidence angle dependency of the data as depicted in Fig. 4. The basic steps for performing target pre-detection can be summarized as:

1. Data patch extraction in time domain consisting of certain number of azimuth and range samples (cf. Fig. 1 top).
2. Incoherent summation over azimuth to obtain an average amplitude profile (cf. blue line in Fig. 4).
3. Computation of an adaptive pre-detection threshold (cf. red line in Fig. 4).
4. Cancellation of the target signals along azimuth (i.e., target azimuth lines) by additionally considering certain range guard zones.

In the following it is explained how the pre-detection threshold is computed. If r is the range vector and $A(r)$ is the average amplitude profile (cf. step 2 of above list), the equation to estimate the target pre-detection threshold as a function of range is given as

$$\eta_{pre}(r) = \tilde{A}(r) + f.\mathrm{SG}\big(k.\mathrm{MAD}(r)\big), \qquad (1)$$

where $\eta_{pre}(r)$ is the pre-detection threshold, $\tilde{A}(r)$ is the median of $A(r)$, $f > 1$ is multiplicative factor, SG is the Savitzky Golay filter, k is a scaling factor with $k \approx 1.4826$ representing 0.75 quantile of the standard Gaussian distribution, and $\mathrm{MAD}(r)$ is called the median absolute deviation which is given as

$$\mathrm{MAD}(r) = \mathrm{median}\big(\left|A(r) - \tilde{A}(r)\right|\big). \qquad (2)$$

To illustrate the adaptive range dependent behavior of the pre-detection threshold, an amplitude over range profile obtained from range-compressed X-band F-SAR data is shown in Fig. 4.

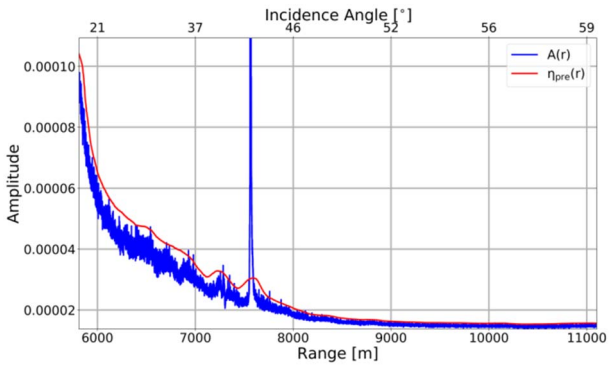

Fig. 4. Average amplitude range profile $A(r)$, with range dependent pre-detection threshold $\eta_{pre}(r)$ for a RC X-band F-SAR data patch. A high target peak (= ship) is present at a range of approximately 7500 m.

In Fig. 4 it clearly can be seen how the computed threshold follows the range profile. The bright target peak at a range of approx. 7500 m as well as other high clutter peaks between 6000 and 7500 m are effectively detected by the proposed algorithm.

B. *Normalization over Doppler*

After target pre-detection and cancellation in time domain, the remaining data are transformed into range-Doppler domain. Clutter statistics estimation and CFAR detection threshold computation is performed in this domain. The principle challenge associated to compute a single, Doppler frequency independent detection threshold in range-Doppler domain is illustrated in Fig. 5.

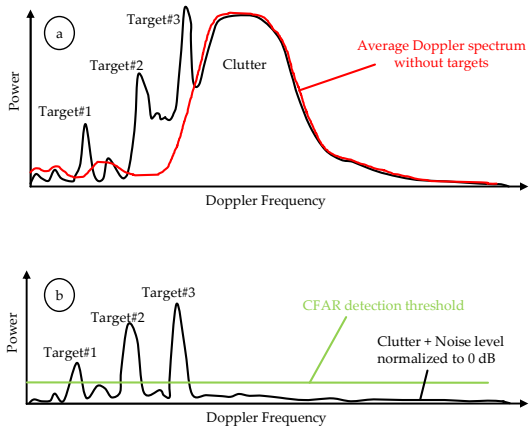

Fig. 5. Illustration of the principle of the normalization over Doppler. (a) Average power of clutter plus targets (black curve) and clutter only (red curve) without performing clutter normalization. (b) Average power of clutter plus targets after clutter normalization. In this case a constant CFAR threshold (green) can be applied.

For performing normalization over Doppler, it is important to firstly estimate an average Doppler spectrum without targets (cf. red curve in Fig. 5(a)). The estimated average power profile $\hat{R}(f_\mathrm{a})$ can be written as [12]

$$\hat{R}(f_\mathrm{a}) = \frac{1}{K}\sum_{k=1}^{K} |Z(r_\mathrm{k}, f_\mathrm{a})|^2, \qquad (3)$$

where K is the number of range bins used for averaging, r_k is the range to a certain range bin, f_a is the Doppler frequency, and $Z(r_\mathrm{k}, f_\mathrm{a})$ is the range-Doppler domain representation of the RC radar data. For simplicity we write $|Z(r_\mathrm{k}, f_\mathrm{a})|^2 = P_\mathrm{k}(f_\mathrm{a})$, which is known as power spectral density or Doppler spectrum $Z(r_\mathrm{k}, f_\mathrm{a})$. The target free and clutter normalized data $Z_\mathrm{CN}(r_\mathrm{k}, f_\mathrm{a})$ can then be written as

$$|Z_\mathrm{CN}(r_\mathrm{k}, f_\mathrm{a})|^2 = P_\mathrm{k}(f_\mathrm{a})/\hat{R}(f_\mathrm{a}). \qquad (4)$$

After using (4), the clutter and noise power are scaled to an average value of 0 dB. The term $|Z_\mathrm{CN}(r_\mathrm{k}, f_\mathrm{a})|^2$ can be considered as training data which can be used for estimating the ocean clutter statistics.

IV. Ocean Statistics Estimation and Threshold Computation

One of the widely used distribution models for characterizing heterogeneous sea clutter in high resolution radar systems is the K-distribution [13]. Due to its compound nature, it is able to capture the fast varying Rayleigh distributed speckle modulated by the Gamma distributed underlying radar cross section (RCS) variations. An example for K-distribution based ship detection using real single-channel X-band HH polarized F-SAR data for desired CFAR of 10^{-6} is shown in Fig. 6.

For generating the results shown in Fig. 6 and in the following figures, a moving window was used for extracting the required range-compressed data patches in time domain (cf. principle of the algorithm depicted in Fig. 1). The moving window was composed of 10 coherent processing intervals (CPIs) in azimuth direction and 512 range samples, where a single CPI is composed of 128 azimuth bins.

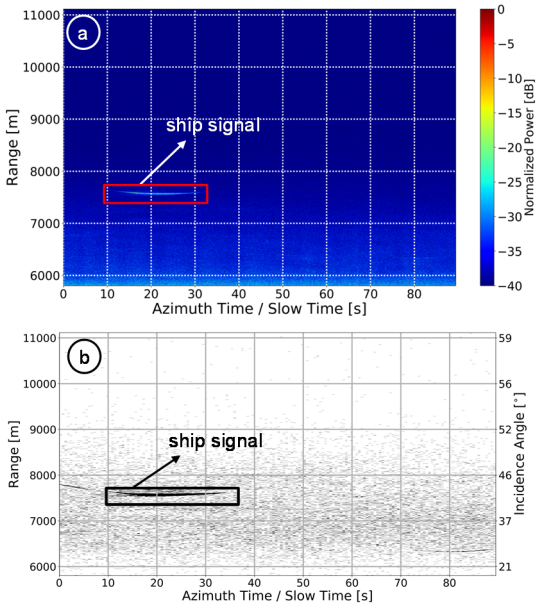

Fig. 6. (a) Real single-channel RC X-band HH polarized F-SAR data including a ship signal. (b) Binary detection map backprojected to time-domain after applying K-distribution based CFAR detection in range-Doppler domain. The set false alarm rate was 10^{-6}.

The radar data shown in Fig. 6(a) and Fig. 8(a) were acquired during a dedicated experiment carried out with a controlled German federal police ship during an F-SAR campaign in the North Sea in 2016 [5].

In Fig. 6(b) it can be clearly observed that the number of false detections is significantly higher at lower incidence angles. This is due to the presence of discrete sea spikes whose intensity matches with the target. In such cases the K-distribution fails to capture these spikes. Therefore, we use the K-Rayleigh distribution function which models these spikes as an extra Rayleigh component [14]. In previous studies, it was found that the K-Rayleigh distribution performs very well for high clutter-to-noise ratios (CNRs). In region with extremely low CNR, as this for the shown example is especially the case for far range, a different model should be used. In our case, we chose the tri-modal discrete texture (3MD) model introduced in [10].

To investigate the models behavior, the data shown in Fig. 6(a) were divided into three regions: near (15°-30° incidence angle), mid (30°-50°) and far range (> 50°). The obtained false alarm rate error was estimated for each of these regions. The estimated false alarm rate is then compared with the set (desired) false alarm rate. The results are shown in Table I. In the optimum case, where there are no clutter spikes and no other target peaks, the ratio should be 1.

TABLE I. ESTIMATED FALSE ALARM RATE ERROR FOR NEAR, MID AND FAR RANGES. SET FALSE ALARM RATE = 10^{-6}

Slant range	K-distribution	K-Rayleigh distribution	3MD model
Near	80.5	1.31	149.2
Mid	112.1	1.68	135.9
Far	3.08	-	1.56

From Table I it can be recognized that the false alarm rate errors are much higher in near and mid ranges for the K-distribution and 3MD model compared to K-Rayleigh. Only in far range the estimated false alarm rate is comparable to the desired value. On the contrary, the K-Rayleigh distribution gives the least error in near and mid range, but is found to be not applicable in far range due to too low CNR [14].

Additionally, we also computed the threshold errors for each of these models using the data shown in Fig. 6(a). The threshold error is the absolute difference between the thresholds estimated from the data complementary cumulative distribution function (CCDF) and the CCDF computed from the models. The obtained threshold errors (in log scale) for a CCDF of 10^{-4} are shown in Table II.

TABLE II ESTIMATED THRESHOLD ERROR FOR DIFFERENT CLUTTER MODELS AT CCDF = 10^{-4}

Slant range	K-distribution	K-Rayleigh distribution	3MD model
Near	6.89	-6.68	8.94
Mid	6.02	2.27	6.94
Far	-5.86	-	-10.86

From Table II it can be seen that the threshold error is minimum in case of K-Rayleigh distribution in near and mid range. In far range, 3MD model gives the minimum error.

Based on the previous analyses, we applied the K-Rayleigh distribution in near and mid range of the data and the 3MD model in far range. The detection results are shown in Fig. 7. Compared to Fig. 6(b), where only the K-distribution was used, there are much less false detections in near and mid range. Discrete sea spikes are no longer detected.

The same methodology of using two different clutter models for CFAR detection in near/mid and far range was also applied on circularly acquired RC radar data. The detection result is shown in Fig. 8(b). The results show clearly, that the proposed methodology also works well for circular SAR data.

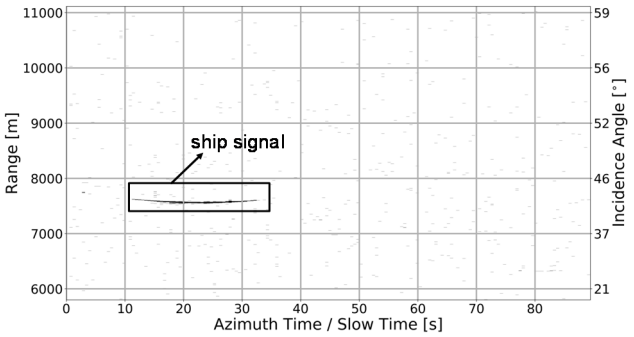

Fig. 7. Binary detection map obtained from real X-band HH polarized RC F-SAR radar data, where the CFAR threshold was computed in range-Doppler domain. The K-Rayleigh distribution function was used in near and mid range and the 3MD model was used in the far range for threshold computation. The desired false alarm rate was set to 10^{-6}.

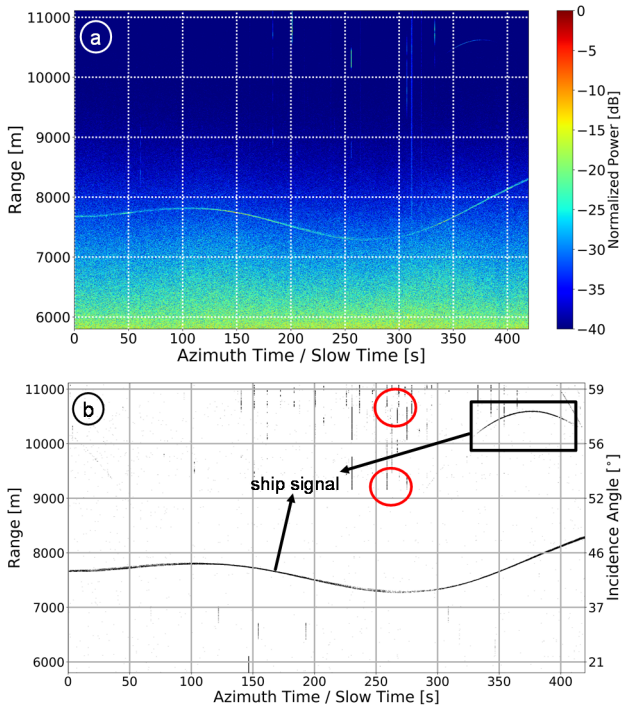

Fig. 8. (a) Circularly acquired real single-channel HH polarized RC L-band F-SAR radar data. (b) Binary detection map shown backprojected to time-domain after applying CFAR detection in range-Doppler domain. The K-Rayleigh distribution was used in near and mid range and the 3MD model was used in the far range. The desired false alarm rate was set to 10^{-6}. The detections marked by the red circles are due to interfering signals from a ground surveillance radar located close to the test site (for visualization purposes not all of the interfering signals are marked).

V. CONCLUSION

In this paper a novel CFAR based ship detection algorithm using range-compressed airborne radar data, was presented. Since this algorithm requires comparatively low processing effort, it is principally suitable for future real time applications. The paper emphasized on automatic ocean training data extraction, which is imperative for accurate CFAR target detection. Considering the strong sea spikes, present especially in the near and mid range, and the low CNR in the far range of the investigated X- and L-band data, we recommend to use two different clutter

models for CFAR detection: the K-Rayleigh distribution in near and mid range and the 3MD model in far range. This recommendation is at least valid for the limited amount of RC F-SAR data we had available for our investigations, and the current sea state conditions during the data acquisitions. It may happen that for different sea states other models are better suited. Under this viewpoint more radar data acquired during different sea states are needed for more sophisticated investigations to be carried out in future.

REFERENCES

[1] F. Te Hennepe *et al.*, "Space-based detection of AIS signals: Results of a feasibility study into an operational space-based AIS system," in *ASMS/SPSC 2010: 2010 5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop*, 2010, pp. 17–24.

[2] D. J. Crisp, "The State-of-the-Art in Ship Detection in Synthetic Aperture Radar Imagery," *Inf. Sci. (Ny).*, no. DSTO-RR-0272, p. 115, 2004.

[3] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. Van Schie, "Ship surveillance with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 1092–1103, 2011.

[4] S. K. Joshi and S. V Baumgartner, "Sea clutter model comparison for ship detection using single channel airborne raw SAR data," in *EUSAR 2018; 12th European Conference on Synthetic Aperture Radar*, 2018, no. 1, pp. 731–735.

[5] S. V Baumgartner, "Linear and Circular ISAR Imaging of Ships Using DLR ' s Airborne Sensor F-SAR," in *IET Radar*, 2017.

[6] S. Watts and L. Rosenberg, "A comparison of coherent and non-coherent radar detection performance in radar sea clutter," in *International Conference on Radar Systems (Radar 2017)*, 2017, pp. 1–6.

[7] A. Reigber *et al.*, "Current status of DLR's new F-SAR sensor," in *EUSAR*, 2010, pp. 1078–1081.

[8] S. K. Joshi, S. V Baumgartner, A. B. C. Silva, and G. Krieger, "Range-Doppler Based CFAR Ship Detection with Automatic Training Data Selection," *Remote Sens.*, vol. 11, no. 11, p. 1270/1-1270/39, 2019.

[9] X. Leng, K. Ji, S. Zhou, X. Xing, and H. Zou, "An adaptive ship detection scheme for spaceborne SAR imagery," *Sensors (Switzerland)*, vol. 16, no. 9, 2016.

[10] C. H. Gierull and I. Sikaneta, "A Compound-Plus-Noise Model for Improved Vessel Detection in Non-Gaussian SAR Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1444–1453, 2018.

[11] El_Mashade Mohamed B., "Monopulse detection analysis of the trimmed mean level CFAR processor in nonhomogeneous situations," *IEE Radar, Sonar Navig.*, vol. Vol.143, no. No.2, p. 87–94., 1996.

[12] L. Rosenberg and S. Watts, "Model based coherent detection in medium grazing angle sea-clutter," in *2016 IEEE Radar Conference, RadarConf 2016*, 2016.

[13] K. D. Ward, "Compound representation of high resolution sea clutter," *Electron. Lett.*, vol. 17, no. 16, pp. 561–563, 1981.

[14] L. Rosenberg, S. Watts, and S. Bocquet, "Application of the K+Rayleigh distribution to high grazing angle sea-clutter," in *2014 International Radar Conference, Radar 2014*, 2014.