International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

Maritime Surveillance in SAR Data Using Multiscale Attention Models

Suresh Kumar M[1], *Giridharan L*[2], *Hemanth K R*[3]

[1]*Associate Professor - Department of Computer Science and Engineering, Sri Venkateswara College of Engineering, Chennai, Tamil Nadu, India.*

[2]*PG - Department of Computer Science and Engineering, Sri Venkateswara College of Engineering, Chennai, Tamil Nadu, India.*

[3]*UG - Department of Computer Science and Engineering [Artificial Intelligence and Data Science], Sri Venkateswara College of Engineering, Chennai, Tamil Nadu, India.*

Emails: *sureshm@svce.ac.in*[1] , *giridharan6369@gmail.com*[2], *ktrhemanthh@gmail.com*[3]

Abstract

A novel hybrid deep learning framework for ship detection in Synthetic Aperture Radar (SAR) satellite imagery is proposed, integrating MobileNetV3 for lightweight and efficient feature extraction, Feature Pyramid Networks (FPNs) for multi-scale feature fusion, and an optimized YOLOv8n object detection module for precise and real-time ship localization. This framework addresses key challenges in SAR-based maritime detection, including varying ship scales, low-contrast targets, and complex noisy sea backgrounds. To enhance model generalization under limited annotated data, the system leverages transfer learning using pre-trained MobileNetV3 weights and applies robust data augmentation techniques, such as geometric transformations and noise simulations. Experimental evaluations conducted on a dedicated SAR ship detection dataset demonstrate that the proposed model achieves an average precision of approximately 90.4% at an IoU threshold of 0.5, while significantly reducing false alarm rates compared to conventional methods. The model's real-time inference ability and minimal computing overhead make it ideal for deployment on platforms with limited resources, like UAVs and onboard satellite systems. The findings demonstrate the framework's promise as a dependable and effective solution for automated vessel monitoring, coastal security, and marine surveillance applications.

Keywords: *Convolutional Neural Networks (CNN); Maritime Surveillance; Object Detection; Ship Detection; Synthetic Aperture Radar (SAR) Images; Yolo.*

1. Introduction

Maritime surveillance has become an indispensable tool for ensuring national security, safeguarding commercial shipping, and protecting marine ecosystems. With the rapid proliferation of satellite technology, Synthetic Aperture Radar (SAR) images have emerged as a valuable resource for continuous and all-weather monitoring of oceanic regions. However, the effective utilization of SAR imagery for ship detection remains a significant challenge due to the inherent complexities associated with these images. Factors such as low resolution, speckle noise, and varying environmental conditions complicate the accurate identification and localization of ships in SAR data. Traditional methods for ship detection in SAR imagery relied heavily on classical image processing techniques such as thresholding, edge detection, and morphological filtering to isolate potential targets. Although these methods provided an initial framework for ship detection, they often lacked robustness, particularly when dealing with cluttered backgrounds or varying ship sizes. Moreover, manual interpretation of SAR images is both time-consuming and error-prone, which underscores the need for automated solutions. This work presents an advanced multiscale attention mechanism for precise ship detection in SAR images.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

Our proposed framework integrates advanced techniques to address the limitations of traditional methods. At its core, we utilize a CNN-based backbone for initial feature extraction. To effectively handle ships appearing at different scales, we incorporate a Feature Pyramid Network (FPN), which enhances multiscale representations by merging low- and high-level features. This improves the model's ability to detect small, occluded, or low-contrast ships, ensuring robust performance in challenging conditions. Additionally, to address the scarcity of annotated training data a common issue in SAR image analysis comprehensive data augmentation strategies are employed. Techniques such as rotation, flipping, and scaling are applied to artificially expand the dataset, thereby enhancing the generalizability of the model and reducing the risk of overfitting. The proposed hybrid framework marks a significant advancement in ship detection for SAR imagery. By combining the strengths of CNNs, FPNs, and advanced attention mechanisms, our approach enhances detection accuracy and robustness across various maritime scenarios. Experimental results on a curated SAR ship dataset highlight its potential as a practical solution for real-time maritime surveillance and security applications.

2. Related Work

In this section the review of existing works is done for SAR ship detection. Peng Wang et.al [1] Multiscale ship recognition in SAR images is severely hampered by small-scale ships' poor contrast and limited pixel count when compared to large-scale ships. A universal multiscale pyramid attention module (MPAM) has been developed as a lightweight, plug-and-play part that can be incorporated into different ship detection networks in order to overcome these problems. The deep feature extraction submodule of MPAM initially divides the feature map into discrete levels by utilizing a multiscale pyramid structure. Rich characteristics that capture both high-resolution details and semantic information necessary for ship detection at various scales can be extracted thanks to this architecture. The module then presents two submodules of attention fusion, one concentrating on spatial multilayer attention and the other on channel multilayer attention. Together, these submodules merge attention blocks from various feature map levels, enabling the network to learn more dependent characteristics in both spatial and channel dimensions. Improved detection results are obtained by feeding the improved feature representation that this fusion provides into already-existing ship detection networks. Results from experiments on datasets with multiscale ships show that adding a MPAM significantly increases detection accuracy, especially in difficult situations with small or low-contrast ships. Chen Shiqi et.al [2] Although real-time applications still face many obstacles, automatic ship detection in SAR imagery has long been a crucial field of research for marine monitoring. Many current detectors struggle to deliver high-quality localization, especially for small objects amidst complicated backgrounds, despite substantial breakthroughs in deep learning approaches. A single-stage detector built on an attention mechanism has been suggested as a solution to these drawbacks. This method significantly reduces noise and suppresses background interference by converting pixel-level into box-level segmentation. The detector's focus is guided by an attention map that is created by utilizing weak segmentation to automatically identify and roughly pinpoint regions of interest. Additionally, the model integrates a multi-branch fusion module with a top-down feature pyramid structure. By combining high-level semantic information with specific low-level characteristics, this architecture enhances the feature representation and makes it possible to recognize multi-scale and multi-directional targets more precisely. Promising results from experimental evaluations on a SAR ship dataset demonstrate the detector's improved capacity to precisely identify ships in difficult image scenarios. Rong Yang et.al [3] Deep learning methods have advanced SAR ship detection, yet challenges remain particularly mismatches between feature scales and target sizes, conflicting learning tasks, and imbalanced positive sample distributions that hinder the detection of small ships. To address these issues, a novel one-stage detection framework, R-RetinaNet, is proposed, which builds on RetinaNet with rotatable bounding boxes (RBox). Three major innovations are

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

introduced in the framework: a task-wise attention FPN that decouples feature optimization across tasks; and a scale calibration method that aligns backbone feature maps with target scales. Tests conducted on a publicly available SAR ship detection dataset show notable gains in detection accuracy, with average precision gains of up to 13.26% when compared to the most advanced RBox-based techniques. This succinct architecture provides a solid solution for real-time maritime surveillance by successfully addressing important issues in SAR ship detection. Rui Qin et.al [4] Speckle noise reduces SAR target recognition accuracy, and conventional reduction techniques that depend on transform domain and spatial filtering necessitate expertly adjusted thresholds, which frequently impair performance. A multilayer wavelet speckle reduction network (Wavelet-SRNet) is suggested as a solution to this problem. This method embeds a CNN framework with a trainable wavelet soft threshold denoising module. A CNN-based classifier is then cascaded to create a unified SAR target identification network with an automatically learned denoising threshold after a two-level wavelet denoising branch has been built and fused with noisy image. Wavelet-SRNet provides high test accuracy at noise augmentation studies and surpasses state-of-the-art approaches in classification accuracy, according to experiments conducted on a target acquisition and recognition database. Cong Wang et.al [5] Data dispersion causes the traditional fuzzy C-means (FCM) technique to converge slowly and be susceptible to noise. We suggest a more reliable and computationally effective version of the FCM algorithm for image segmentation. To minimize noise and preserve details, pictures are first filtered using a morphological grayscale reconstruction (MGR) process. Feature spaces are obtained from observed and filtered images, we use a three-step iterative approach that uses Lagrangian multipliers, a hard-threshold operator, and normalization to solve the FCM objective after adding a sparse regularization term. MGR filters the clustering labels to further improve segmentation. Experiments on color, synthetic, and medical pictures show that our system performs faster and more accurately in segmentation than current FCM variations. Jiahuan Jiang et.al [6] Synthetic Aperture Radar (SAR) is crucial for marine surveillance because of its capacity to operate in any weather conditions, day or night. SAR ship detection is essential for law enforcement and maritime security. Deep learning techniques have replaced traditional methods, however many of them rely on intricate, GPU-intensive tactics that are inappropriate for frontline deployment and frequently ignore SAR's signal-to-noise ratio. We suggest a multi-channel SAR processing technique based on YOLO-V3 to address these problems. Our customized YOLO-V3-light network optimizes three-channel image processing while lowering model size, load, and memory consumption. Our approach demonstrated its potential for emergency rescue and real-time maritime monitoring when tested on portable computers, achieving an Average Precision of 90.37% on the SAR Ship Dataset. Delphine Cerutti-Maori et.al [7] This study uses the scan-MTI mode of the aerial radar sensor PAMIR to examine wide-area traffic monitoring in real-world scenarios. the versatile GMTI mode quickly scans wide areas and finds moving things from a variety of aspect angles. Important performance metrics are measured and contrasted with the theoretical GMTI expectations. The outcomes validate the scan-MTI mode's appropriateness for real-time, practical applications by showing that it is especially successful for effective wide-area traffic monitoring. Peng Wang et.al [8] Due to unfavorable imaging conditions, spectral imagery frequently has poor resolution and a large number of mixed pixels, which causes errors in mapping land cover. Conventional super-resolution mapping (SRM) methods ignore nonlinear imaging effects and suffer from spectral unmixing errors since they mostly rely on linear spatial correlations in their attempt to retrieve subpixel information. A unique spatial-spectral correlation (SSC)-based SRM technique is put forth in order to get beyond these restrictions. SSC uses a spectral correlation measure based on the nonlinear KLD to more precisely extract spectral features and a MSAM that uses distance to capture spatial correlation. the SSC method successfully addresses both linear and nonlinear imaging difficulties by integrating both spatial and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

spectral correlations, leading to increased mapping accuracy. Colin P. Schwegmann et.al [9] It is intrinsically difficult to locate ships in SAR pictures, particularly when dealing with unfriendly vessels that do not transmit transponder signals. CA-CFAR prescreening technique, which uses a set scalar threshold to categorize pixels as ships, is frequently used in traditional methods. Although this threshold indirectly regulates the quantity of false alarms and establishes the brightness needed for detection, its static nature may result in less-than-ideal performance. A innovative method transforms the scalar threshold into a dynamic threshold manifold in order to get beyond these restrictions. A simulated annealing (SA) approach is used to adaptively alter this manifold under the guidance of ship dispersion data obtained from transponder data. The technique reduces the typical computational inefficiencies of SA by carefully selecting threshold boundaries. This method demonstrated a remarkable low false alarm rate of 1.01×10^{-7} and a detection accuracy of 85.2% when tested on six ASAR images against five alternative approaches. Ruijin Jin et.al [10] A level set technique for segmenting high-resolution PolSAR images incorporate a heterogeneous clutter model called the distribution. Conventional level set techniques for PolSAR data depend on the intricate Wishart distribution, which frequently falls short in modeling diverse locations like cities and woods. For multilook PolSAR data, on the other hand, the £ distribution proves to be a more adaptable and reliable model. According to experimental results, this approach outperforms Wishart, Kummer-U, and Markov random field-based approaches in terms of accurately segmenting and characterizing high-resolution PolSAR data.

Table 1 Merits and Demerits of Related work

Authors	Merits	Demerits
Peng Wang et al.	Enhances multiscale detection using lightweight MPAM with rich feature extraction.	Slightly increases model complexity and training time.
Chen Shiqi et al.	Boosts real-time detection accuracy with attention-guided weak segmentation.	May underperform in clutter-heavy scenes or fine object boundaries.
Rong Yang et al.	Yolov framework improves small ship detection using scale calibration and adaptive IoU.	Involves complex training strategies and decoupled task optimization.
Rui Qin et al.	Learns denoising thresholds automatically via Wavelet-SRNet for robust classification.	Training complexity increases and may limit generalization.
Cong Wang et al.	Improved FCM algorithm enhances segmentation accuracy and speed using MGR and sparsity.	Relies heavily on preprocessing and less focus on real-time deployment.
Jiahuan Jiang et al.	YOLO-V3-light enables efficient real-time ship detection with high AP.	Sacrifices some detection accuracy for reduced computational load.
Delphine Cerutti-Maori et al.	ScanMTI mode enables wide-area moving target detection with high revisit rates.	Not suitable for static ship detection and limited image detail.
Peng Wang et al. (SRM)	SSC method improves SR mapping using spatial-spectral correlation under nonlinear conditions.	Designed for land-cover mapping, not directly applicable to ships.
Colin P. Schwegmann et al.	Dynamic threshold manifold via SA achieves high accuracy and low false alarms.	High computational cost and dependence on transponder data.
Ruijin Jin et al.	£-distribution-based level set method yields superior segmentation on PolSAR data.	Requires complex parameter tuning and modeling effort.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

3. Proposed Work

The proposed work aims to develop an efficient and robust ship detection system for Synthetic Aperture Radar (SAR) imagery, addressing the inherent challenges of low contrast, noise, and the small size of targets, while ensuring real-time performance on portable devices. The proposed framework follows an optimized multi-stage pipeline, integrating deep learning techniques across key components: data preprocessing and augmentation for noise reduction and contrast enhancement, multi-scale feature extraction using a CNN and SE-Net with an FPN to improve ship detection at various scales, and a YOLO-based object detection module for fast and accurate localization. Table 1 shows Merits and Demerits of Related work. By eliminating redundant classification steps and leveraging efficient backbone architectures, the framework enhances both detection accuracy and real-time performance, making it well-suited for maritime surveillance and security applications. In the initial stage, raw SAR images undergo preprocessing to enhance their suitability for deep learning models. Each image is normalized to scale pixel values between 0 and 1, ensuring consistent input for the network. Given the limited availability of annotated SAR datasets, extensive data augmentation techniques, including multi-angle rotations and both horizontal and vertical flips, are applied to generate additional training samples. This process increases dataset diversity and improves the model's ability to generalize. The preprocessing pipeline plays a crucial role in optimizing the input data while mitigating overfitting through augmentation.

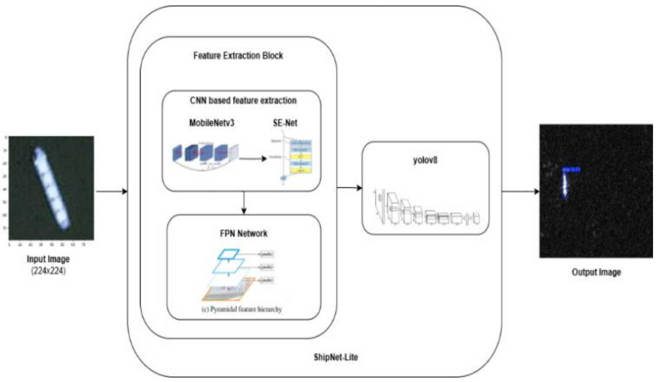

Figure 1 Architecture of The Proposed System

The final stage of the system integrates the FPN-enhanced features into a modified YOLO detection module as shown in Figure 1, optimized for real-time performance. By adapting the YOLO-V8n architecture to incorporate the multi-scale features from the FPN, the detection head is able to generate accurate bounding boxes and confidence scores for ships in SAR imagery. This tailored YOLO model, referred to as YOLO-V8n, reduces the overall computational complexity by minimizing model size, decreasing detection time, and lowering memory consumption, making it suitable for deployment on portable computing devices. The entire detection framework is trained end-to-end, ensuring that feature extraction, classification, and object localization are optimized simultaneously. Overall, the proposed work offers a comprehensive solution to the challenges of ship detection in SAR images. By integrating robust data augmentation, CNN-based classification, multi-scale feature extraction through FPN, and an optimized YOLO detection framework, the system achieves high detection accuracy and fast inference speeds. Experimental evaluations on a dedicated SAR ship detection dataset demonstrate that this integrated approach not only outperforms conventional methods but also meets the practical requirements of real-time maritime surveillance and emergency response.

4. Result and Discussion

Experimental evaluations were conducted on a dedicated SAR ship detection dataset to assess the performance of the proposed framework. The integrated system, incorporating extensive data augmentation, CNN-based feature extraction with MobileNetV3, multi-scale feature extraction via FPN, and a YOLOv8n-based detection module, achieved high accuracy in SAR ship detection. Specifically, the system reached an average precision (AP) of approximately 90.4% under an IoU threshold of 0.5. In addition to accuracy, the model demonstrated a significant reduction in false alarm rates, maintaining a low false alarm rate (FAR) compared to baseline methods. Ablation studies further underscore the importance of the phonetic conversion and phonological similarity modules. When these modules were disabled, the FID

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

increased by over 15%, and the generated images exhibited noticeable artefacts and reduced visual coherence. This finding confirms that incorporating auditory cues into the embedding process is critical for capturing nuanced visual characteristics such as the smoothness associated with soft sounds or the angularity linked to sharp, staccato phonetics. Figure 2 shows Input Image with Its Pixel Intensity Distribution.

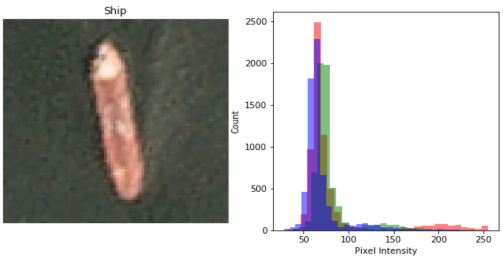

Figure 2 **Input Image with Its Pixel Intensity Distribution**

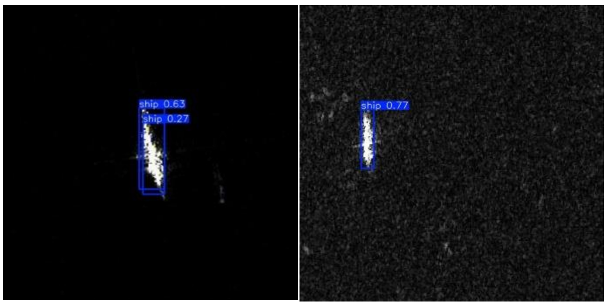

Figure 3 **Output Produced by Shipnet-Lite Mode**

Figure 3 illustrates the model's ability to distinguish a ship from background noise in a SAR (Synthetic Aperture Radar) image. The detected ship is highlighted with blue bounding boxes, each labeled with a confidence score. The model assigns a probability of 0.77 and 0.63 to the detected regions, reinforcing its ability to localize and classify the ship while filtering out irrelevant noise in the surrounding sea. Our proposed architecture enabled adaptive predictions across various ship sizes while maintaining low inference time, making real-time applications feasible. Despite these successes, certain limitations were observed. In extremely noisy environments or cases with very low contrast, the model occasionally misclassified targets. Future research may focus on integrating advanced attention mechanisms Beyond accuracy and robustness, the system maintained a low inference time, benefiting from the lightweight architecture of YOLOv8n and the computational efficiency of MobileNetV3. This supports real-time application feasibility, a crucial requirement for deployment in maritime monitoring systems aboard UAVs, coastal radar stations, or satellite-based platforms. The integration of multi-scale feature extraction through FPN contributed significantly to the detection of ships of varying sizes and orientations, particularly improving performance in detecting small or partially occluded vessels. This is crucial in real-world maritime environments where ships appear at different resolutions due to varying altitudes and imaging angles.

Conclusion

In summary, this work presents an efficient and effective SAR ship detection framework designed to address the core challenges of maritime object detection, such as background noise, low target contrast, and varying ship sizes. The proposed approach integrates several key components: MobileNetV3 for lightweight and high-performance feature extraction, Feature Pyramid Network (FPN) for multi-scale feature fusion, and YOLOv8n as a fast and optimized detection head tailored for real-time inference on resource-constrained devices. Comprehensive experiments conducted on a dedicated SAR ship detection dataset demonstrate that the system achieves an impressive average precision of approximately 90.4% at an IoU threshold of 0.5, while also maintaining a low false alarm rate a critical factor in operational maritime surveillance scenarios. The use of MobileNetV3 enables fast and energy-efficient processing without compromising on detection quality, making the framework suitable for deployment on UAVs, edge devices, or satellite-based monitoring platforms. The integration of FPN allowed the network to effectively detect ships across multiple scales, especially smaller vessels often missed in traditional detection pipelines. Additionally, the adoption of the YOLOv8n detection architecture ensured efficient bounding box regression and confidence scoring, resulting in accurate and rapid object localization. This architecture not only demonstrated competitive

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2455 - 2461
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0365

performance but also showed resilience to typical SAR-specific challenges such as speckle noise and target occlusion. Its ability to distinguish ships from complex backgrounds further supports its real-world applicability. Looking forward, future enhancements may include incorporating attention mechanisms (e.g., Transformer blocks or channel-spatial attention modules) to further refine feature maps, as well as adaptive thresholding techniques to better handle uncertain detections in highly noisy scenes. There is also potential for leveraging temporal SAR sequences or multi-modal data fusion to boost detection stability in dynamic maritime conditions. Overall, the promising results achieved in this study validate the strength of combining lightweight CNNs, multi-scale feature processing, and optimized detection heads. This sets the stage for deploying real-time, low-power ship detection systems in large-scale maritime surveillance and emergency response operations.

Reference

[1]. Peng Wang, Yongkang Chen, Yi Yang, Ping Chen, Gong Zhang, Daiyin Zhu, Yongshi Jie, Cheng Jiang, and Henry Leung, (2024), "A General Multiscale Pyramid Attention Module for Ship Detection in SAR Images", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 17, pp. 2815-2827.

[2]. Chen Shiqi, Zhan Ronghui, Zhang Jun, (2019), "Regional attention-based single shot detector for SAR ship detection", The Journal of Engineering, Vol. 21, pp. 7381 – 7384.

[3]. Rong Yang, Zhenru Pan, Xiaoxue Jia, (2021), "A Novel CNN-Based Detector for Ship Detection Based on Rotatable Bounding Box in SAR Images", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 14, pp. 1938-1958.

[4]. Rui Qin, Xiongjun Fu, (2022), "Multilevel Wavelet-SRNet for SAR Target Recognition", IEEE Geoscience and Remote Sending Letters, Vol. 19, pp. 1-5.

[5]. Cong Wang, Witold Pedrycz, (2021), "Sparse Regularization-Based Fuzzy C-Means Clustering Incorporating Morphological Grayscale Reconstruction and Wavelet Frames", IEEE Transactions of Fuzzy Systems, Vol. 29, pp. 1826-1840.

[6]. Jiahuan Jiang, Xiongjun Fu, Rui Qin, Xiaoyan Wang and Zhifeng Ma, (2021), "High Speed Lightweight Ship Detection Algorithm Based on YOLO-V4 for Three-Channels RGB SAR Image", Remote Sensing ,13, pp. 1909.

[7]. Delphine Cerutti-Maori, Jens Klare, Andreas R. Brenner, and Joachim H. G. Ender, (2008), "Wide-Area Traffic Monitoring With the SAR/GMTI System PAMIR", IEEE Transactions on Geoscience and Remote Sensing, Vol. 46, No. 10, pp. 3019-3030.

[8]. Peng Wang, Liguo wang, Henry Leung and Gong Zhang, (2021), "Super-Resolution Mapping Based on Spatial–Spectral Correlation for Spectral Imagery", IEEE Transactions on Geoscience and Remote Sensing, Vol. 59, No. 3, pp. 2256-2268.

[9]. Colin P. Schwegmann, Waldo Kleynhans, and Brian P. Salmon, (2015), "Manifold Adaptation for Constant False Alarm Rate Ship Detection in South African Oceans", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 8, pp. 3329-3337.

[10]. Ruijin Jin, Junjun yin, Wei Zhou, and Jian Yang, (2017), "Level Set Segmentation Algorithm for High-Resolution Polarimetric SAR Images Based on a Heterogeneous Clutter Model", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 10, pp. 4565-4579.