Student IDs: 23198583 23205919

ASSIGNMENT TWO

PAPER NAME: Data Mining and Machine

Learning PAPER CODE: COMP809

TOTAL MARKS: 100

Students' Names: ……………………Yize(Serena) Wang…………………GuangLiang(Ricky) Yang…………………

Students' IDs: ……………………………23198583……………………………………23205919………………………

- Due date: 09 Jun 2024 midnight NZ time.
- Late penalty: maximum late submission time is 24 hours after the due date. In this case, a **5% late penalty** will be applied.
- Submit the actual code (no screenshot) separately with appropriate comments for each task.

Note: This assignment should be complemented by a group of two students and both students MUST contribute in each part.
Submission: a soft copy needs to be submitted through the canvas assessment link.

INSTRUCTIONS:

1. **The following actions may be deemed to constitute a breach of the General Academic Regulations Part 7: Academic Discipline,**
   - Communicating with or collaborating with another person regarding the Assignment
   - Copying from any other student work for your Assignment
   - Copying from any third-party websites unless it is an open book Assignment
   - Uses any other unfair means
2. **Please email DCT.EXAM@AUT.AC.NZ if you have any technical issues with your submission on Canvas immediately**
3. **Attach your code for all the datasets in.**

SAR ship Image Detection

1st Yize(Serena) Wang
Student ID: 23198583
Auckland University of Technology

2nd Guangliang(Ricky) Yang
Student ID: 23205919
Auckland University of Technology

***Abstract*—In recent years, Synthetic Aperture Radar (SAR) has become an important technology in maritime surveillance due to its ability to operate independently of weather conditions such as cloud cover and time of day. However, there are still challenges in detecting small targets and dealing with complex backgrounds in remote sensing images. Computer vision and machine learning have become effective methods for image detection by extracting meaningful features from images. This report addresses these issues using the latest object detection algorithm, You Only Look Once, version 8 (YOLOv8). YOLO is a single-stage object detection algorithm known for its high speed and accuracy. YOLOv8 consists of three main components: the backbone, the detection head, and the post-processing module. Experiments have shown that YOLOv8 achieves higher detection accuracy for SAR ship detection compared to YOLOv9. YOLOv8 is effective for detecting small targets and recognizing complex backgrounds in SAR ship detection.**

***Keywords*—Machine Learning, SAR, Deep learning, Ship recognition, YOLO, Object detection**

I. Introduction

Synthetic Aperture Radar (SAR) is a remote sensing technology that uses radar to obtain images of ground targets. In ship detection missions, it is unaffected by changing sea weather and can monitor ship targets in real time from all directions. SAR has significant research importance and application value in maritime traffic, sea rescue, and maritime military target monitoring.

Deep learning, a branch of machine learning, is widely used in image recognition, object detection, and semantic segmentation tasks. In recent years, object detection has become widely applied. Object detection mainly includes two methods. The first method is region-based two-stage detection models, such as Region-based Fast R-CNN, Region-based Fully Convolutional Network (R-FCN), and Mask R-CNN. The second method is regression-based single-stage detection. The YOLO series is a famous single-stage algorithm. Compared to the two methods, single-stage methods are faster but slightly less accurate than two-stage methods.

In the second section, I will introduce the background. In the third section, I will review related literature to explore how advanced methods address SAR data complexities. The fourth section our opinion, introduces and proposes my theoretical solution, model architecture, extend the existing work and perspective on Machine Learning and Data Mining challenges. Finally, I will summarise my work and future directions.

II. Background

Traditional feature extraction and selection for SAR ship detection typically require manually designed algorithms. With the widespread application of deep learning in SAR ship detection, these traditional methods have been replaced by deep learning due to their end-to-end characteristics and strong robustness. Deep learning methods optimize and improve image recognition accuracy by extracting features from ship images and continuously learning and training these features to build neural networks. However, some issues have been identified in the application process. SAR ship detection faces challenges in recognizing small targets and complex backgrounds, indicating significant room for improvement in terms of accuracy and detection speed.

YOLO, a classic single-stage detection algorithm, was released in 2015. It processes images by dividing them into multiple grids, with each grid predicting multiple bounding boxes and their confidence levels, then determining the target categories in the predictions. YOLO's speed and accuracy have made it a mainstream detection algorithm. With the iterations of the YOLO series, detection has become more accurate and faster. I will use YOLOv8 for SAR ship detection. YOLOv8 consists of three main parts: the backbone (BN), detection head (DH), and post-processing module (PPM). The BN extracts features from images; the DH uses these features to predict object boxes and class probabilities; and the PPM improves predictions by removing redundant boxes to enhance accuracy.

III. Literature Review

Ship detection in SAR images is of great significance for maritime surveillance tasks in various countries. Deep learning techniques have been widely applied in this field, but due to the complex background and small targets, there are still significant challenges. This report reviews some advanced technological methods and experimental results for SAR ship detection, discusses the advantages and disadvantages of these methods, and offers valuable suggestions for improvement.

To address issues such as low resolution, small targets, dense near-shore ship arrangements, and background clutter noise in SAR ship detection, the author proposes SPD-

InceptionDWConv (SIDConv). They designed lightweight SIDConv (LSIDConv) and efficient SIDConv (ESIDConv), as well as a lighter and faster detector, LFer-Net. Experiments on the SSDD and HRSID datasets showed accuracies of 98.2% and 90.6%, respectively. Overall, the SIDConv module improves feature extraction efficiency, and the ECA module indirectly reduces background clutter noise. The SIDConv model balances high accuracy, speed, and reduced complexity, making it suitable for real-time applications. But, integrating multiple new modules (SIDConv, LSIDConv, ESIDConv, ECA) may increase implementation complexity. By combining innovative convolutional techniques to enhance feature extraction while maintaining a lightweight and efficient model, this method effectively addresses key issues in SAR ship detection. The ability to handle low resolution, small targets, and background clutter sets a new standard for SAR ship detection technology. [1]

To improve small target detection and reduce near-shore interference in SAR ship detection, the author proposes improvements to the YOLOv5 model: optimizing input size and anchor frames, introducing asymmetric pyramid non-local block (APNB) and Simple Attention Mechanism (SimAM) , channel transformation in the C3 module, and Ghost Convolution. Experiments on the HRSID and SSDD datasets showed accuracies of 93.4% and 95.8%, respectively. The APNB and SimAM attention mechanisms reduce background interference and enhance target area focus, while Ghost Convolution and optimized module structure effectively reduce model complexity. Integrating multiple advanced modules (APNB, SimAM, channel transformation, Ghost Convolution) increases implementation complexity. The improved YOLOv5 method offers high accuracy, reduced computational demands, and robustness in various SAR scenarios. This method's enhancements solve key issues of small target detection and background clutter, making it suitable for real-time and large-scale applications in maritime surveillance and monitoring. [2]

In complex environments such as near-shore areas and small ships in SAR images, traditional methods face false alarms and missed detections. To address these issues, the author proposes a high-precision ship detection method based on the YOLOv7-tiny network, DBW-YOLO. Experiments on the HRSID and SSDD SAR datasets showed accuracies of 90.04% and 98.41%, respectively. The improved deformable convolutional networks (DCNets) and BiFormer attention mechanism significantly enhance performance in handling background and small targets. The WIoU loss function improves convergence speed and generalization ability. However, the model's complexity and computational requirements may limit its application in resource-constrained environments. The DBW-YOLO method is a valuable tool for handling complex scenarios and detecting small targets in SAR images. [3]

To address the complex backgrounds and diverse scales of ships in SAR images, a multi-granularity perception network (MGA-Net) is proposed. This network integrates two key components: the multi-granularity hybrid feature fusion module (MGHF2M) and the multi-granularity feature synergy enhancement module (MGFSEM). In experiments on the HRSID dataset, the accuracy reached 93.87%. MGHF2M enhances the capability of ship feature representation, while MGFSEM improves the representation of various scales of ship features. Compared to other advanced methods, MGA-Net shows superior performance in mAP and F1 scores. The use of MGHF2M and MGFSEM modules allows the model to capture detailed local information and global semantic context, enhancing its robustness to background interference. However, the integration of multiple modules increases model complexity, which may lead to longer training times and higher computational resource demands. MGA-Net solves key challenges in SAR ship detection, providing a powerful solution that performs well in complex and variable environments. It offers valuable insights for future research and development in this field. [4]

A ship detection scheme based on YOLOv7 has been proposed to address issues such as unclear contours, complex backgrounds, and strong scattering in SAR images. This method achieved an accuracy of 84.4% on the high-resolution SAR images dataset (HRSID). It aims to improve detection accuracy and recall rate without compromising detection speed. YOLOv7 maintains high accuracy while providing fast detection speed, making it very suitable for real-time applications. YOLOv7 introduces innovative feature extraction structures, employing multi-branch stacking, path aggregation network (PANet), and spatial pyramid pooling (SPP). These innovative feature extraction methods enhance the model's feature extraction capability and overall performance. Enhanced feature fusion, using Feature Pyramid Network (FPN) and Path Aggregation Network (PANet), achieves more effective feature fusion, thereby improving detection accuracy in complex backgrounds. However, despite its overall good performance, YOLOv7 may still have certain limitations when dealing with small or dense targets. Overall, the YOLOv7-based detection scheme addresses the complex background problem in SAR ship detection by improving detection speed and accuracy. It provides valuable insights and tools for enhancing ship detection in SAR images. [5]

To address model complexity and performance degradation caused by speckle noise in SAR ship detection, a lightweight multi-scale network is proposed. The main components of this method are IAS (ShuffleNetV2) and a lightweight attention-enhanced path aggregation feature pyramid network (LAE-PAFPN). In experiments on the SAR Ship Detection Dataset (SSDD), the accuracy reached 98.65%. This model combines high detection accuracy with reduced complexity, making it suitable for deployment on platforms with limited computational resources. However, further evaluation on other datasets is needed to provide a more comprehensive assessment of the method's robustness. The high mAP achieved on a single dataset suggests the possibility of overfitting. This research offers a lightweight and efficient solution to key challenges in SAR ship detection. Its reduced complexity and improved accuracy make it highly relevant for real-time applications on platforms with limited computational capabilities. The method's robustness to speckle noise and its ability to handle multi-scale ship detection provide valuable insights for future SAR image analysis research. [6]

In summary, the literature review provides insights into the latest advancements in SAR ship detection methods. The main findings include the effectiveness of single-stage detection algorithms, particularly YOLO, in improving detection speed and accuracy. Various innovative methods such as MGA-Net, SPD-InceptionDWConv, SimAM and DBW-YOLO have been proposed to address challenges like small target detection, complex backgrounds, and low-resolution images. These methods have shown impressive accuracy and robustness in different SAR datasets. Despite these advancements, there is still room for improvement in detection accuracy. The integration of multiple advanced modules often increases model complexity and computational requirements, posing challenges for deployment on resource-limited platforms. These studies provide significant help for future research in improving small target detection accuracy and reducing model complexity.

IV. Our Opinion

Based on the literature review, multiple authors have utilized the YOLO algorithm and have confirmed that YOLO is effective in addressing the issues of complex backgrounds and small targets in SAR ship detection. Currently, YOLO has evolved to YOLOv9. Through the analysis of the internal principles of YOLOv8 and a comparative demo (Fig. 1), it is evident that YOLOv8 is more suitable for SAR ship detection. YOLOv8 uses convolutional neural networks (CNN) as its foundational model, leveraging end-to-end features and robustness to replace traditional handcrafted feature extraction algorithms. YOLOv8 divides images into multiple grids and predicts object bounding boxes and class probabilities for each grid, achieving fast and accurate target detection. Its main advantage lies in its ability to effectively handle the complexity of SAR data, including small targets, high noise, and complex backgrounds. Additionally, YOLOv8 further improves detection accuracy and speed by addressing variations in object scale, complex object shapes, and the presence of multiple objects in cluttered scenes. The entire method framework consists of three main components: the backbone network, detection head, and post-processing module, responsible for extracting features from input images, manipulating features to generate predictions, and refining and optimising prediction results to improve final detection accuracy.

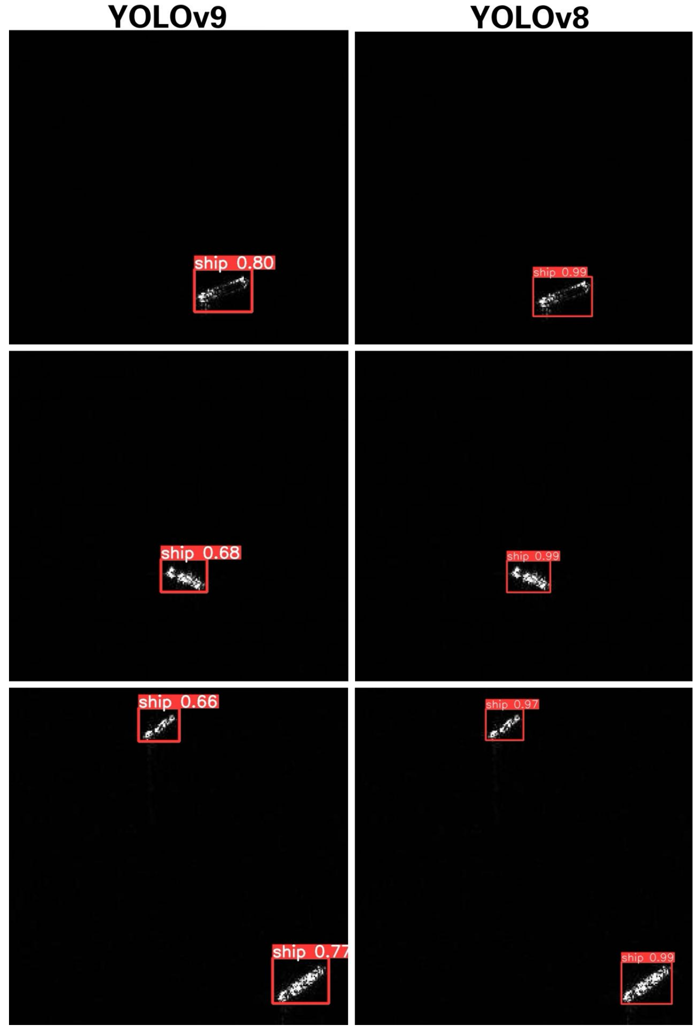

Fig. 1. Comparison of different YOLO version predictions.

A. Model Architecture

The YOLOv8 algorithm is a fast single-stage object detection method composed of a backbone network, neck, and head modules. The backbone network and neck module form the central structure of the YOLOv8 network. The backbone network is responsible for extracting features from input images and transforming the images into feature representations with rich semantic information. The neck module is used to fuse features from different levels to enhance the model's performance. The head module is then utilized to predict the position and class of the targets. The model structure is shown in Figure 2. The backbone of YOLOv8 comprises the CBS, C2f, and Spatial Pyramid Pooling—Fast (SPPF) modules. YOLOv8 processes input images through multiple Conv and C2f modules to extract features at different scales. Unlike YOLOv5, the C2f module in YOLOv8 is an improved version of the original C3 module and serves as the main residual learning module. It combines the advantages of the ELAN structure in YOLOv7 by reducing one standard con-

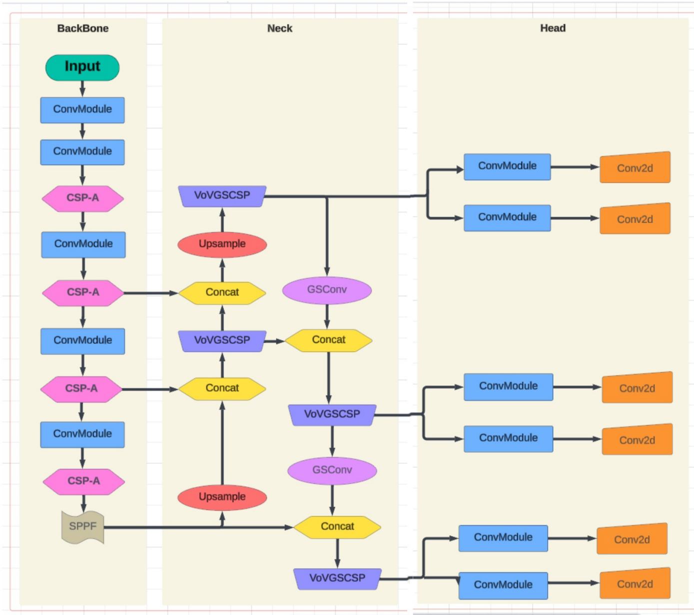

Fig. 2. The model architecture of YOLOv8.

volutional layer and making full use of bottleneck modules to enhance gradient branching. This method not only retains lightweight characteristics but also captures richer gradient flow information. YOLOv8 integrates the SPPF module from YOLOv5 to process the output feature maps. In contrast to YOLOv5, the SPPF in YOLOv8 is not aimed at obtaining a fixed-size representation layer but rather at pooling at different size scales.

The neck of YOLOv8 adopts the PAN-FAN structure, enhancing the model's feature fusion capability. In comparison to YOLOv5, YOLOv8 removes the 1×1 downsampling layer. Through this approach, the network can better fuse features of objects at different scales, thereby improving its detection performance.

The detection head of YOLOv8 utilizes a decoupled-head structure, separating the classification head from the detection head. Additionally, anchor-based is replaced with anchor-free.

Loss calculation includes the positive and negative sample assignment strategy and the loss computation. The positive and negative sample assignment strategy assigns a weight to each sample during training to prioritize important samples for the model. The loss calculation includes two parts: classification and regression. BCE loss is used as the classification loss, while DFL loss and CIoU loss are used as the regression loss.

B. *Extend the existing work*

Data Collection and Preprocessing

Through web scraping, I collected 2,000 SAR ship images, which include various complexities such as very small ships, single ships, and multiple ships in one image. Then, label the datasets accurately to ensure high-quality training. For this research project, I divided the dataset into three subsets: the training set, the validation set, and the test set. 1,701 images were randomly selected for training, 200 images for validation, and the remaining 99 images for testing. I resized the images from 256×256 to 640×640 and applied low contrast, low sharpening, low brightness, and Gaussian noise to some images. By processing the data in this manner, it can help the model generalise better.

Model Optimization

Model training is a key step in improving YOLOv8's performance. During training, we need to choose the right optimizer, learning rate, and training epochs. To prevent overfitting, we can use techniques like data augmentation and regularization. By adjusting these parameters, we can find the best training strategy for our task.

- Initial training: The YOLOv8 model will be trained for 5 epochs with input images resized to 640x640 pixels. A batch size of 16 images will be used, and the AdamW optimizer with a learning rate of 0.002 and momentum of 0.9 will help update weights efficiently. A weight decay of 0.0005 will be applied for regularization, and the learning rate will be gradually increased over the first 3 epochs (warm-up epochs). The loss function weights will be balanced, with box loss, class loss, and distribution focal loss weighted at 1.674, 0.9914, and 1.381, respectively. During inference, a non-maximum suppression threshold of 0.7 and a maximum of 300 detections per image will be used. Various augmentations will be applied, and training will be done using Ultralytics YOLOv8.0.196, Python 3.10.12, Torch 2.2.1+Cu121, and a 15102MiB Tesla T4 GPU for faster training.
- Parameter adjustment:
  - Adjust the learning rate: Too high a learning rate can make the model converge to suboptimal solutions too quickly, while too low a learning rate can make training slow and stuck in local minima. Use learning rate schedules like cosine annealing or learning rate decay.
  - Adjust batch size: Experiment with different batch sizes to find the best balance between memory usage and training speed, enhancing model stability and generalization.
  - Adjust anchor box sizes and aspect ratios: Use k-means clustering on the training dataset to determine optimal anchor boxes to improve model performance.
  - Network depth and width: Use cross-validation to find the optimal architecture. Adjusting the network's depth (number of layers) and width (number of neurons per layer) helps improve performance.

- Regularization: Implement dropout layers and L2 regularization in the model to prevent overfitting.
- Early stopping: Monitor validation loss during training and use an early stopping parameter to define how many epochs to wait before stopping when the loss starts increasing.
- Dropout: Add dropout layers to randomly drop neurons during training, helping prevent the model from over-relying on specific neurons.
- Cross-validation: Use k-fold cross-validation by dividing the dataset into k subsets, training the model k times, each time using a different subset as the validation set and the remaining as the training set, ensuring consistent model performance across different data subsets.

Incorporate Advanced Techniques

- Add SPD-Conv (Stride Pooling Deconvolution Convolution) without convolution strides or pooling: A new CNN module for low-resolution images and small objects.
- Add SimAM (Simple Attention Mechanism): An attention mechanism that reduces background interference and enhances focus on target areas.
- Add Ghost Convolution: A lightweight convolution module.

Evaluation and Comparison

After training the model, we need to validate it on a test set to evaluate its performance.

Evaluation Metrics: Use precision, recall, F1-Score, and mean Average Precision (mAP). Plot confusion matrices, PR curves, and ROC curves to evaluate the model.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

$$AP = \int_0^1 P(R)\, dR \tag{4}$$

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \tag{5}$$

Test the model on benchmark datasets (e.g., HRSID and SSDD) to ensure it generalizes well to different SAR images.

By doing these, we can understand the model's performance on different tasks and further optimize it. For example, if the model's accuracy is low, we can try increasing training epochs or adjusting the learning rate. If the recall is low, we can add more positive samples or adjust the weights of positive and negative samples.

Real-World Application and Deployment

Once the model is trained and validated, we can deploy it in real-world applications. During inference, input images are fed into the model to get detection results. To improve inference speed, we can use techniques like GPU acceleration and code optimization. Continuous monitoring and maintenance are necessary to ensure the model's stability.

Continuous Improvement

Collect feedback from practical applications and continuously improve the model. Stay updated with the latest advancements in SAR technology and deep learning to integrate new techniques.

C. Perspective on Machine Learning and Data Mining Challenges

Data Quality and Preprocessing: Data quality is crucial. Low-quality data can lead to inaccurate models and unreliable insights. Significant effort should be placed on data preprocessing, including cleaning, normalization, and handling missing values. Advanced techniques like anomaly detection and imputation methods can improve data quality.

Model Complexity and Performance: Balancing model complexity and computational efficiency is key, especially when deploying on resource-limited platforms. High-performing lightweight models are desirable but often require innovative design and optimization techniques.

Overfitting and Generalization: Overfitting is a common issue where the model performs well on training data but poorly on unseen data. Regularization techniques, cross-validation, and ensuring a sufficient amount of diverse training data are essential strategies. Models should be evaluated on different datasets to ensure good generalization.

Scalability and Efficiency: As datasets grow larger, scalability becomes crucial. Efficient algorithms and distributed computing frameworks (e.g., Hadoop and Spark) are necessary for handling large-scale data. Optimization techniques and hardware accelerators (e.g., GPUs and TPUs) can significantly improve model training and inference times.

Continuous Learning and Adaptation: The environment in which models operate can change over time, necessitating continuous learning and adaptation. Online learning algorithms and periodic retraining of models are important to maintain the relevance and accuracy of predictions.

V. Conclusion and Future Work

This report begins with an abstract, followed by an instruction section that discusses the issues and techniques for recognizing various ships in SAR images. Then, through a literature review, it introduces some of the latest methods. In the our opinion section, we present our ideas, explaining the theory behind YOLOv8 and how to extend the existing work. We also share our opinions on machine learning and data mining-related issues.

Although our YOLOv8 demo predictions are good results, further experiments and optimisations are needed in the future to address small target detection and complex background issues.

References

[1] C. Tian, D. Liu, F. Xue, Z. Lv and X. Wu, "Faster and Lighter: A Novel Ship Detector for SAR Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024, doi: 10.1109/LGRS.2024.3351132.

[2] Z. Wang, G. Hou, Z. Xin, G. Liao, P. Huang and Y. Tai, "Detection of SAR Image Multiscale Ship Targets in Complex Inshore Scenes Based on Improved YOLOv5," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 5804-5823, 2024, doi: 10.1109/JSTARS.2024.3370722.

[3] X. Tang, J. Zhang, Y. Xia and H. Xiao, "DBW-YOLO: A High-Precision SAR Ship Detection Method for Complex Environments," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 7029-7039, 2024, doi: 10.1109/JSTARS.2024.3376558.

[4] L. Ying, Y. Liu, Z. Zhang and D. Miao, "Multigranularity-Aware Network for SAR Ship Detection in Complex Backgrounds," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024, doi: 10.1109/LGRS.2024.3352633.

[5] C. Yu and Y. Shin, "A Deep Learning-Based SAR Ship Detection," *2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pp. 744-747, 2023, doi: 10.1109/ICAIIC57133.2023.10067131.

[6] H. Yu, S. Yang, Z. Liu and S. Zhou, "A Lightweight Network for Multiscale SAR Ship Detection Under Complex Imagery Backgrounds," *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena*, pp. 6406-6409, 2023, doi: 10.1109/IGARSS52108.2023.10282664.