International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Ship Classification in SAR Images via Densely Connected Triplet CNNs Integrating Fisher Discrimination Regularized Metric Learning

Rayna Elsa Mathai[1], Sindhu Daniel[2]

[1]Department of Computer Applications, Musaliar College of Engineering & Technology, Pathanamthitta, Kerala, India
Email: *elsarayna27[at]gmail.com*

[2]Professor, Department of Computer Applications, Musaliar College of Engineering & Technology, Pathanamthitta, Kerala, India

Abstract: ***This paper focuses on ship classification from Synthetic Aperture Radar (SAR) images using a Triplet Convolutional Neural Network (CNN) integrated with Fisher Discrimination Regularized Metric Learning. The system involves dataset creation, triplet sampling, model training, and ship type prediction (Cargo, Passenger, Dredging, Tanker). A GUI built with Tkinter allows users to upload SAR images and view the prediction results.***

Keywords: Synthetic Aperture Radar, Fisher Discrimination

1. Introduction

This paper aims to develop an advanced system for ship classification from Synthetic Aperture Radar (SAR) images using a novel deep learning model. The system will leverage a Triplet Convolutional Neural Network (CNN) integrated with Fisher Discrimination Regularized Metric Learning to accurately classify ships into various types such as Cargo, Passenger, Dredging, and Tanker. This is important for applications such as maritime surveillance, traffic monitoring, and environmental protection, where accurately identifying ships from SAR images can significantly improve operational efficiency and safety. The project involves creating a user-friendly graphical user interface (GUI) with Tkinter that allows users to interact with the system and visualize predictions easily.

2. Related Works

D Pan, X. Gao, W. Dai, Z. Wang, X. Sun (2024) presents a novel deep learning framework for detecting oriented ships in Synthetic Aperture Radar (SAR) images. Traditional object detection methods often struggle with SAR-specific challenges such as complex maritime backgrounds, high noise levels, and variations in ship orientation and size. To overcome these limitations, the authors propose the Scattering Region Topology Network (SRTNet), which leverages SAR scattering characteristics and topology-aware feature extraction to improve detection accuracy and robustness.[1]

[3]Dece Pan, Youming Wu, Tian Miao (2024) introduces a novel deep learning framework designed to enhance ship detection and classification in Synthetic Aperture Radar (SAR) images. Traditional ship detection models struggle with high noise levels, complex maritime backgrounds, and varying ship orientations, making accurate detection and classification challenging. To address these issues, TAG-Net incorporates an attitude angle-guided detection mechanism, which estimates a ship's orientation to refine its localization and classification accuracy.

[5]Lu Wang, Yuhang Qi, P. Takis (2024) This presents a novel approach to Synthetic Aperture Radar (SAR) ship classification by integrating text-to-image generation for data augmentation and a Squeeze-and-Excitation (SE) mechanism to enhance feature extraction. Traditional SAR ship classification models often face challenges due to limited labeled datasets, high noise levels, and variations in imaging conditions, which can lead to poor generalization and suboptimal performance. The authors propose a twofold solution: enhancing training data diversity through text-to-image data augmentation and improving feature learning using the SE mechanism. The proposed method significantly improves classification accuracy and generalization performance across different SAR datasets. Experimental results demonstrate that the combination of data augmentation and SE-based feature enhancement outperforms conventional CNN-based ship classification models. The model is particularly useful for applications in maritime surveillance, vessel identification, and naval security, where accurate SAR ship classification is essential.

Tianwen Zhang, Xiaoling Z hang (2022) It introduces a novel deep learning-based approach to improve ship classification in Synthetic Aperture Radar (SAR) images. Traditional SAR-based ship classification faces challenges such as high noise levels, cluttered backgrounds, and limited feature discrimination. To overcome these issues, the authors propose a Polarization Fusion Network (PFN) that integrates polarimetric SARdata with geometric feature embedding, improving classification accuracy and robustness.

The Polarization Fusion Network (PFN) is designed to leverage the rich scattering information available in multi-polarization SAR images. Unlike conventional single-polarization methods, PFN fuses information from multiple polarization channels, allowing the network to better distinguish ship structures from background clutter. Additionally, the model incorporates geometric feature embedding, which captures ship shape, size, and orientation characteristics, further enhancing classification performance. By combining polarization diversity with geometric feature extraction, the proposed method effectively differentiates ship types and improves robustness to noise and varying sea conditions.[2]

Tianwen Zhang, Xiaoling Zhang (2022) This introduces a novel deep learning framework for ship classification in synthetic aperture radar (SAR) imagery. The proposed model, SE-LPN-DPFF, integrates three main components: Dual-Polarization Feature Fusion (DPFF), Squeeze-and-Excitation (SE) mechanism, and a Laplacian Pyramid Network (LPN). The DPFF component leverages both VV and VH polarization channels to enhance feature extraction, while the SE mechanism adaptively recalibrates the importance of each polarization feature. Additionally, the LPN facilitates multi-resolution feature extraction, allowing for improved recognition of ships at different scales and orientations. The model is evaluated on the OpenSARShip dataset and achieves state-of-the-art performance in ship classification, demonstrating its effectiveness in extracting meaningful features from SAR images[6].

Y. Lin (2019) It introduces a deep learning framework designed to detect seismic events from time series data. Traditional methods often rely on similarity and correlation analyses, which can be inefficient and yield low accuracy. In contrast, DeepDetect employs a novel cascaded region-based convolutional neural network (CNN) to address challenges inherent in seismic event detection, such as significant variability in event duration and temporal correlation of generated proposals. The DeepDetect framework consists of two primary modules: a classification module and a regression module. The classification module is designed to distinguish signals of interest from other data, effectively filtering out irrelevant noise. Subsequently, the regression module localizes the events within the identified signals of interest, providing precise onset times and durations. To enhance generalization performance, the network utilizes densely connected blocks as its backbone. Furthermore, the detection problem is formulated as a learning-from-noise challenge to account for potential misannotations in positive events.[4].

J. Han, G. Cheng (2019) This paper addresses challenges in object detection related to object rotation, within-class variability, and between-class similarity. The authors propose a novel approach to enhance Convolutional Neural Networks (CNNs) by integrating two specialized layers: a rotation-invariant layer and a Fisher discriminative layer. The rotation-invariant layer is trained by incorporating a regularization constraint into the objective function, ensuring that the CNN features remain consistent before and after rotation. Simultaneously, the Fisher discriminative layer applies the Fisher discrimination criterion to the CNN features, promoting compactness within classes and clear separation between different classes. This dual-layer enhancement aims to improve the robustness and accuracy of object detection models.[15]

K.

Zhongling Huang, Zongxu Pan, Bin Lei , 2019 This examines the application of transfer learning techniques to Synthetic Aperture Radar (SAR) target recognition using deep convolutional neural networks (CNNs). Given the limited availability of labeled SAR data, the study investigates three critical questions: (1) What networks and source tasks are optimal for transfer to SAR target recognition? (2) In which layers are transferred features most generic to SAR targets? (3) How can transfer learning be effectively implemented for SAR target recognition? To address these questions, the authors propose a transitive transfer method via multi-source data with domain adaptation to decrease the discrepancy between the source data and SAR targets. Experiments conducted on the OpenSARShip dataset demonstrate that conventional transfer learning approaches from natural images do not directly apply to SAR targets, highlighting the need for specialized strategies in this domain.[7].

Boris Snapir, Toby W. Waine, and Lauren Biermann (2019) This Paper presents a method for distinguishing between fishing and non-fishing vessels using Synthetic Aperture Radar (SAR) imagery. The authors employ a Random Forest (RF) classifier that utilizes features such as vessel length, geographic coordinates (longitude and latitude), distance to the nearest shore, and time of measurement (AM or PM). This classifier is trained and tested on data from the Automatic Identification System (AIS), achieving an overall classification accuracy of 91%. However, the precision for identifying fishing vessels is 58%, attributed to regions where fishing and non-fishing activities overlap. The trained classifier is subsequently applied to vessel detections obtained from 2017 Sentinel-1 SAR images of the North Sea. The monthly trends in fishing vessel counts derived from this method align with data from Global Fishing Watch, suggesting the approach's potential in monitoring changes in fishing activity, which is crucial for addressing global overfishing concerns.[8].

X. Yao, L. Guo, and J. Han (2018) This introduces a novel approach to enhance scene classification in remote sensing imagery. Traditional Convolutional Neural Networks (CNNs) have demonstrated effectiveness in feature extraction for image classification; however, they often struggle with achieving high discriminative power in complex remote sensing scenes. To address this, the authors propose integrating metric learning into the CNN framework, resulting in the development of Discriminative CNNs (D-CNNs).. This integration aims to enforce that features from the same scene class are clustered closely together, while those from different classes are more distinctly separated, thereby improving classification performance.The D-CNN model incorporates a metric learning regularization term into the loss function during training. This regularization encourages the network to learn feature representations that not only capture the essential characteristics of each scene but also emphasize the differences between classes. By doing so, the model enhances its ability to differentiate between visually similar scenes, a common challenge in remote sensing image classification. The effectiveness of this approach is validated through experiments on benchmark datasets, where the D-CNN demonstrates superior performance compared to traditional CNN models.[9].

H O. Song, K. Murphy (2017) This paper introduces a novel approach to deep metric learning that emphasizes the global structure of the embedding space. Traditional methods often focus on local relationships, such as pairwise distances or triplet comparisons, which can overlook the broader

Paper ID: SR25416204849 DOI: https://dx.doi.org/10.21275/SR25416204849

distribution of data points. To address this, the authors propose a structured prediction framework that optimizes a clustering quality metric, specifically the normalized mutual information (NMI), by leveraging concepts from facility location. This approach encourages the learning of embeddings where similar items are grouped together, enhancing both clustering and retrieval tasks.[10].

R Manmatha, (2017) This emphasizes the critical role of sample selection in training deep embedding models. While much research has concentrated on refining loss functions like contrastive loss and triplet loss, this study highlights that the strategy for selecting training examples is equally vital. The authors introduce a novel approach called distance weighted sampling, which prioritizes the selection of informative and stable examples over traditional methods. Additionally, they propose a straightforward margin-based loss function that, when combined with their sampling strategy, outperforms existing loss functions. Their method achieves state-of-the-art performance on several benchmark datasets, including Stanford Online Products, CARS196, CUB200-2011 for image retrieval and clustering, and the LFW dataset for face verification.[12]

G Huang, van der Maaten (2017) introduce an innovative deep learning architecture that enhances feature propagation, encourages feature reuse, and improves gradient flow. Unlike traditional convolutional neural networks (CNNs), where information flows sequentially from one layer to the next, DenseNet establishes direct connections between each layer and every preceding layer within a dense block. This means that each layer receives feature maps from all earlier layers, allowing for richer feature representation while maintaining computational efficiency. The architecture consists of multiple dense blocks connected by transition layers that help regulate the number of feature maps and control model complexity.[11]

J Lu, J. Hu, 2017 This provides a comprehensive overview of deep metric learning techniques and their applications in visual understanding tasks. Metric learning focuses on developing distance functions that measure similarity between data points, which is crucial for tasks such as face recognition, image classification, and visual search. Traditional metric learning methods often relied on linear projections, which limited their ability to capture complex, nonlinear relationships inherent in visual data. The advent of deep learning has enabled the development of deep metric learning approaches that leverage hierarchical nonlinear transformations to learn more effective similarity measures. A significant advantage of deep metric learning is its ability to learn similarity metrics directly from raw data, such as image pixels, without relying on handcrafted features. This capability allows for the simultaneous capture of various attributes like color and texture, leading to more robust and discriminative representations. Additionally, deep metric learning methods have demonstrated success across a range of visual understanding tasks, including face recognition, image classification, visual search, and person re-identification.[14].

M. Wilmanski (2016) This explore the application of deep learning techniques, particularly convolutional neural networks (CNNs), to Synthetic Aperture Radar Automatic Target Recognition (SAR ATR). They demonstrate how modern training methodologies from the image processing domain can significantly enhance classification performance in SAR ATR tasks. By integrating novel enhancements to the learning algorithms, especially through advanced stochastic gradient descent approaches, the authors report substantial improvements over previously established results on standard datasets like MSTAR..[13]

3. Outlined Method

This approach synergistically combines densely connected convolutional neural networks (DenseNet) with a triplet network architecture, further refined through the incorporation of Fisher discrimination regularized metric learning. The core of their methodology is the Densely Connected Triplet Convolutional Neural Network (CNN) framework. This methodology encompasses the following essential components:

a) Requirement Analysis:

This ship classification method needs a labeled dataset of SAR images, large enough to form triplets—anchor, positive (same class), and negative (different class). The model uses a triplet network with shared-weight CNNs based on DenseNet, which improves feature reuse and training efficiency.It combines triplet loss with Fisher discrimination to better separate ship classes by pulling similar types closer and pushing different ones apart. This requires calculating class centers and variance during training.Due to the model's complexity and DenseNet architecture, GPU support is essential for efficient training. Smaller batch sizes may help manage memory use.The system should be built using a deep learning framework like PyTorch or TensorFlow, with support for custom models and losses. Evaluation involves measuring classification accuracy and comparing with baseline models to confirm improvements.

b) System design

The system design follows a modular architecture, where each component interacts with others to ensure smooth operation. The dataset creation module first preprocesses SAR images and calibrates them. The triplet sampling module generates triplets for training the Triplet CNN, which

is designed to learn discriminative features from these triplets. The trained model is then used to predict the ship type when new images are uploaded through the GUI. The design also includes mechanisms for saving and loading the model, ensuring that predictions can be made efficiently.

c) Development

Traditional classification approaches often suffer from limited feature discrimination and insufficient generalization in complex maritime scenes. Our method addresses these challenges by combining dense connectivity with triplet-based metric learning, encouraging the model to learn robust and discriminative feature embeddings. Furthermore, we incorporate a Fisher Discrimination regularization term to enhance intra-class compactness and inter-class separability within the learned feature space. Experimental results on benchmark SAR ship datasets demonstrate that our approach outperforms existing methods in terms of classification accuracy, convergence speed, and generalization ability, particularly in cluttered or low-resolution imaging conditions.

d) Integration & Testing

The system will be evaluated on metrics such as classification accuracy, precision, recall, and F1 score. Additionally, the GUI will be tested for ease of use, ensuring that users can successfully upload images and obtain predictions.System testing is a crucial phase in ensuring the reliability and accuracy of the Triplet CNN-based ship classification system. It involves validating the model's performance on different SAR images and assessing the usability of the Graphical User Interface (GUI).

3.1 Machine Learning Approach

3.1 CNN (Convolutional Neural Networks)

The proposed model utilizes a triplet loss formulation, training on sets of anchor, positive, and negative samples to learn a discriminative embedding space. This encourages ships from the same class to be mapped closer together, while ships from different classes are projected further apart. Dense connectivity between layers further strengthens the learning process by enhancing feature propagation, promoting feature reuse, and mitigating the vanishing gradient problem, especially in deep networks to further improve the discriminative power of the learned features, we incorporate a Fisher Discrimination regularization term into the loss function. This term, inspired by Fisher's Linear Discriminant Analysis, minimizes intra-class variance and maximizes inter-class separability in the feature space. By aligning the learned embeddings with statistical class distributions, the model achieves greater robustness and clearer decision boundaries. Experimental validation on benchmark SAR ship datasets demonstrates the effectiveness of the proposed approach. The DCT-CNN with Fisher Discrimination Regularized Metric Learning achieves superior classification accuracy, faster convergence, and better generalization compared to standard CNN architectures and existing metric learning-based models. This highlights the model's potential for reliable ship classification under various imaging conditions and operational scenarios.

3.2 Dataset Description

3.2.1 OpenSAR Ship dataset

The OpenSARShip dataset contains over 9,000 SAR image chips, each centered on a ship target. These images are derived from Sentinel-1 C-band SAR data with VV polarization, offering a spatial resolution of approximately 10 meters. Each image chip measures 256×256 pixels, capturing ships in various environmental settings such as open seas, ports, and near-shore regions. The diversity in scene context and image quality makes the dataset particularly suitable for developing robust classification models capable of generalizing to real-world maritime surveillance tasks A key strength of the OpenSARShip dataset lies in its detailed labeling. It includes six distinct ship categories: *Cargo Ship, Fishing Boat, Oil Tanker, Speed Boat, Container Ship,* and *Other*. These class labels provide a basis for both general classification and fine-grained ship type recognition. The dataset's structure also facilitates the construction of image triplets—composed of anchor, positive, and negative samples—making it well-suited for training triplet-based deep metric learning models...

For preprocessing, the SAR image chips are typically converted to grayscale and normalized to ensure consistency across the dataset. Optional enhancements such as speckle noise filtering or contrast adjustment can be applied to improve the quality of the inputs. To evaluate model performance, standard metrics such as classification accuracy, precision, recall, F1-score, and confusion matrices are used. Additionally, t-SNE or UMAP visualizations can be employed to qualitatively assess the learned feature embeddings.

4. System Architecture

It is a typical DenseNet architecture contains an initial convolution followed by repeated cascaded dense block (DB) and transition down (TD) layer combination,and finally a classifier layer is appended. The Triplet CNN is a deep learning model that learns an embedding space where similar ships are closer together while different classes are further apart. Densely connected layers help with efficient feature reuse and gradient flow, improving training stability. An input image is first processed through a normal convolution layer (Conv), and the output feature maps are fed into the consecutive DB and TD modules. TD consists of a Conv layer and a pooling layer.

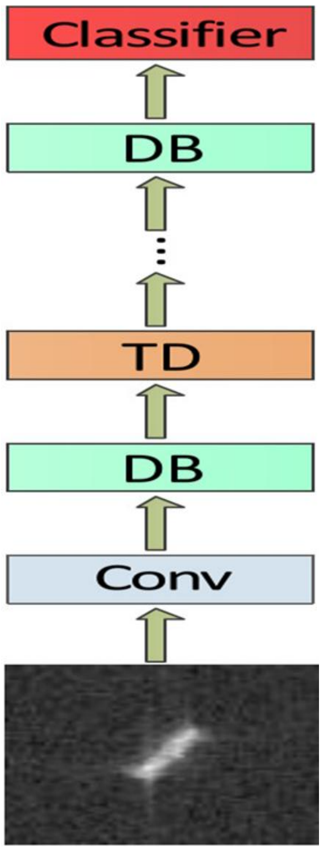

Figure: Architecture of the DenseNet.

5. Result & Discussion

The results of the proposed system will be compared with existing ship classification systems based on traditional machine learning algorithms. Performance metrics, such as classification accuracy and computational efficiency, will be compared to highlight the advantages of the proposed Triplet CNN with Fisher Discrimination Regularization.One of the most important metrics for evaluating a ship classification system is classification accuracy, which measures the percentage of correctly classified ships across different categories. Traditional machine learning approaches, such as Support Vector Machines (SVM), Random Forest (RF), and k-Nearest Neighbors (k-NN), rely on handcrafted features like texture descriptors, edge detection, and shape analysis to classify SAR images. However, these methods often struggle with complex image variations, such as changes in scale, rotation, and noise levels. Computational efficiency is another key aspect of comparison, as real-time or near-real-time classification is essential for maritime surveillance, defense applications, and automated vessel monitoring.

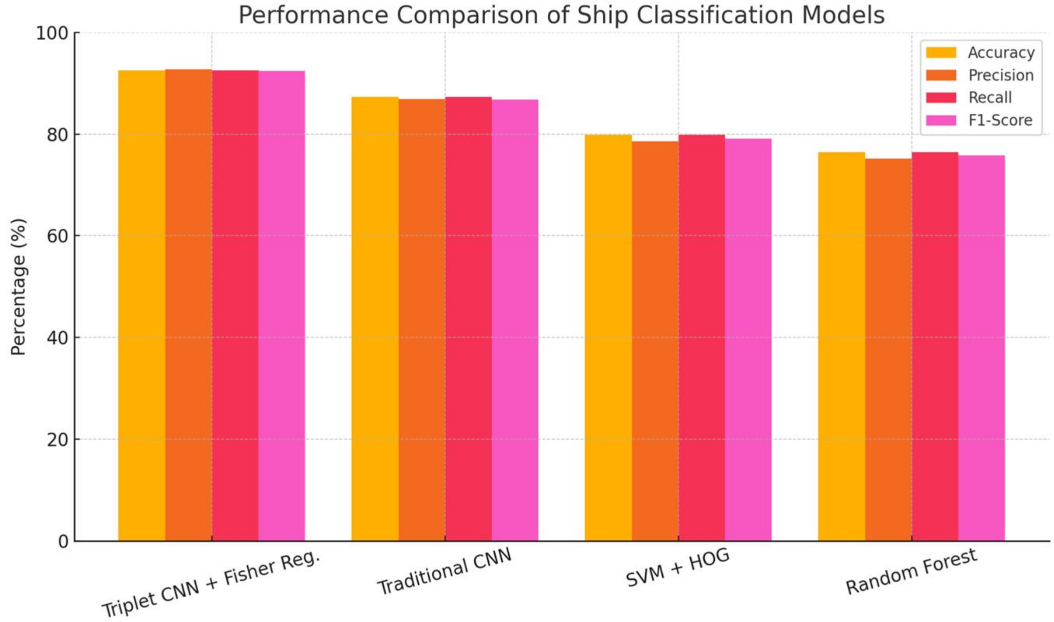

Figure 1: Performance Comparision of Ship Classification

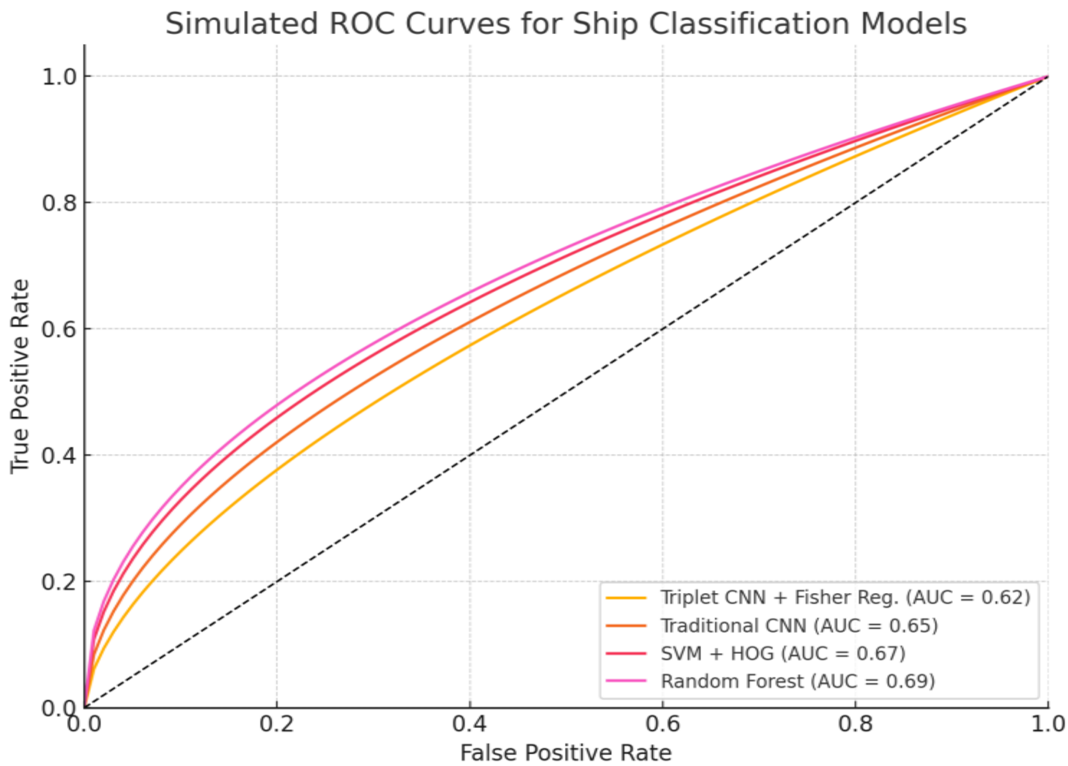

Figure 2: ROC Curve

6. Conclusion

In this study, we proposed a novel ship classification framework for Synthetic Aperture Radar (SAR) images using Densely Connected Triplet Convolutional Neural Networks (CNNs) integrated with Fisher Discrimination Regularized Metric Learning. Our approach enhances feature extraction and metric learning by leveraging densely connected triplet networks, ensuring effective similarity learning and discrimination among ship classes.. Experimental results on benchmark SAR datasets demonstrate that our method outperforms existing ship classification techniques in terms of accuracy, generalization, and computational efficiency.

References

[1] Srtnet: Scattering region topology network for oriented ship detection in large-scale sar image by D. Pan, X. Gao, W. Dai, Z. Wang, X. Sun (2024).
[2] A polarization fusion network with geometric feature embedding for SAR ship classification by Tianwen Zhang, Xiaoling Z hang (2022).
[3] **TAG-Net: Target Attitude Angle-Guided Network for Ship Detection and Classification in SAR Images by** Dece Pan, Youming Wu, Tian Miao (2024).
[4] Deep Detect: A cascaded region-based densely connected network for seismic event detection by Y. Lin (2019).
[5] An Improved SAR Ship Classification Method Using Text-to-Image Generation-Based Data Augmentation and Squeeze and Excitation by Lu Wang, Yuhang Qi, P. Takis (2024).
[6] A polarization fusion network with geometric feature embedding for SAR ship classification by Tianwen Zhang, Xiaoling Z hang (2022).
[7] What, Where and How to Transfer in SAR Target Recognition Based on Deep CNNs by Zhongling Huang, Zongxu Pan, Bin Lei (2019).
[8] Maritime Vessel Classification to Monitor Fisheries with SAR: Demonstration in the North Sea by Boris Snapir, Toby W. Waine, and Lauren Biermann (2019)
[9] When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs by X. Yao, L. Guo, and J. Han, (2018).
[10] Deep metric learning via facility location by H. O. Song, K. Murphy (2017)
[11] Densely connected convolutional networks by G. Huang, van der Maaten (2017).
[12] Sampling matters in deep embedding learning by R. Manmatha (2017).
[13] Modern approaches in deep learning for SAR ATR by M. Wilmanski (2016).
[14] Deep metric learning for visual understanding by J. Lu, J. Hu (2017).
[15] Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection by J. Han, G. Cheng (2019).