

# A Classification-Aware Super-Resolution Framework for Ship Targets in SAR Imagery

Ch Muhammad Awais\*<sup>†‡</sup>, Marco Reggiannini<sup>†‡</sup>, Davide Moroni<sup>†</sup>, and Oktay Karakus<sup>§</sup>

\*PhD School in Computer Science, University of Pisa, 56126 Pisa, Italy

<sup>†</sup>Institute of Information Science and Technologies, National Research Council of Italy, 56124 Pisa, Italy,

<sup>‡</sup>National Biodiversity Future Center - NBFC, Palermo, Italy, <sup>§</sup>School of Computer Science and Informatics, Cardiff University, Cardiff, U.K.

**Abstract**—High-resolution imagery plays a critical role in improving the performance of visual recognition tasks such as classification, detection, and segmentation. In many domains, including remote sensing and surveillance, low-resolution images can limit the accuracy of automated analysis. To address this, super-resolution (SR) techniques have been widely adopted to attempt to reconstruct high-resolution images from low-resolution inputs. Related traditional approaches focus solely on enhancing image quality based on pixel-level metrics, leaving the relationship between super-resolved image fidelity and downstream classification performance largely underexplored. This raises a key question: can integrating classification objectives directly into the super-resolution process further improve classification accuracy? In this paper, we try to respond to this question by investigating the relationship between super-resolution and classification through the deployment of a specialised algorithmic strategy. We propose a novel methodology that increases the resolution of synthetic aperture radar imagery by optimising loss functions that account for both image quality and classification performance. Our approach improves image quality, as measured by scientifically ascertained image quality indicators, while also enhancing classification accuracy.

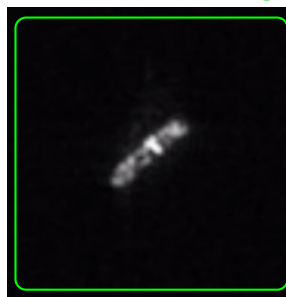
**Index Terms**—SAR ship classification, Deep learning, Synthetic Aperture Radar, Super-resolution

## I. INTRODUCTION

Ship classification algorithms help identify vessels, a critical task given that approximately 80% [1] of the world’s trade is carried out via maritime transport, making maritime traffic monitoring crucial. Ship classification can be performed using images captured by synthetic aperture radar (SAR) [2], [3]; however, SAR data processing and its effectiveness are frequently hindered by two significant challenges: data scarcity and inherent low resolution of publically available data collections [4]. In addition to these factors, SAR images are affected by sidelobe interference, speckle, and angle-dependent scattering behaviour. Strong sidelobes generated by bright scatterers can mask weak targets and

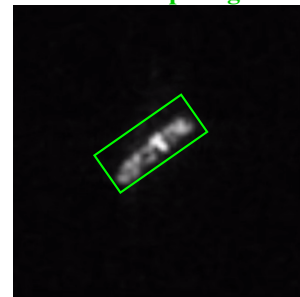
distort local structure [5], while variations in azimuth can introduce nonlinear changes in scattering patterns [6].

Focus on Whole Image



(a) Traditional SR (Global Attention)

Focus on ship target



(b) Proposed (Localized Attention)

Fig. 1: Comparison of attention strategies in super-resolution. Traditional methods apply global attention, while the proposed classification-aware approach uses class label supervision to guide attention toward task-relevant regions, enhancing super-resolution for downstream classification.

Deep learning supervised networks are powerful algorithms that effectively learn from large amounts of annotated data. Such methods can also be applied to solve super resolution (SR) problems [7]. SR models are typically pretrained on large-scale optical image datasets, where they have demonstrated considerable success. Following this pretraining, these models are often repurposed for specialized applications, such as enhancing SAR imagery. However, SR methods developed for optical images do not directly transfer to SAR, as SAR data contain speckle noise, sidelobes, and sea-cluttered backgrounds that alter feature distributions. In this context, the efficacy of SR techniques is typically quantified using pixel-based Image Quality (IQ) metrics, including the Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) [8]–[11].

While these metrics may indicate an improvement in perceptual image quality, it is crucial to consider their relevance for subsequent analytical tasks. This raises a pivotal question: does an enhancement in SR-driven image quality, as measured by conventional metrics like SSIM and PSNR, necessarily translate to improved performance in downstream applications such as image classification? This disconnect is particularly relevant in SAR, where polarimetric studies show that preserving scattering entropy and backscatter contrast is critical for discrimination tasks [12]. When super-resolution models are trained solely on pixel-level reconstruction losses (e.g., L1, L2), they optimize for pixel level accuracy but often produce over-smoothed results. In SAR ship classification, where targets are small and the scene is dominated by background clutter, this smoothing can inadvertently degrade the fine-grained scattering features crucial for accurate classification.

To the best of our knowledge, no prior work has introduced a classification-aware super-resolution strategy specifically for SAR ship classification. While previous studies have optimized SR using pixel-level image quality metrics, our framework combines super-resolution and classification objectives to maintain scattering integrity and semantic consistency, addressing noise and imbalance challenges inherent to SAR ship classification. This approach encourages preservation of discriminative scattering characteristics that are critical for ship-class discrimination, producing super-resolved imagery that maintains both visual fidelity and classification accuracy (see Fig. 1). We validate this hypothesis through a systematic experimental progression: applying off-the-shelf ImageNet SR models (SR-I), training SR models on SAR data with standard reconstruction loss (SR-PT), and finally incorporating classification awareness into the SR training process (SR-FT). To rigorously assess the effectiveness of the proposed methodology, we conducted experiments using three SR models and five classification networks. Evaluations on the widely used OpenSARShip dataset [13], which is known for its low-resolution imagery and severe class imbalance, revealed that image quality enhancements contribute positively to classification accuracy.

Our key contributions are summarised as follows:

- We introduce two novel loss functions specifically designed to guide the training of SR models.
- We propose a unified framework that integrates classification feedback into the SR training process, enabling task-aware image enhancement.
- We demonstrate that the proposed approach simultaneously improves both image super-resolution quality and downstream classification performance.

Next sections are organized as follows: Sec. II provides a brief overview of the background and motivation

for this research, Sec. III details the proposed methodology, Sec. IV presents the results, Sec. V analyzes the findings, and we conclude our study in Sec. VI.

## II. BACKGROUND

High-resolution images play a critical role in deep learning algorithms by preserving fine-grained spatial details that are essential for accurate feature extraction and robust model performance across visual recognition tasks. In SAR ship classification, where objects are small, high-resolution images provide richer visual details [14]–[18] that can be crucial for maritime applications, including traffic monitoring, maritime security, and environmental protection. However, SAR images often suffer from speckle noise and moderate resolution, leading to performance deficiencies in these applications. This underscores the need for techniques that enhance the resolution of SAR data while preserving critical details to improve ship classification accuracy.

Several examples exist in previous literature of how super-resolution techniques can be leveraged to achieve higher resolution in SAR imagery [15], [18], [19]. Most researchers have focused on improving resolution by optimising image quality metrics such as PSNR [10], [20]–[23] and SSIM [9], [10], [20]–[22]. Higher scores in these metrics provide evidence of the ability of SR models to improve image quality.

Typically, supervised SR machine learning algorithms are guided by the minimization of a loss function. This function estimates the difference between the super-resolved image generated by the model and its corresponding ground-truth high-resolution counterpart; the computed deviation is used to iteratively optimise the model's parameters. Also, it should be noted that the performance of an SR model, as assessed by PSNR and SSIM scores, depends on the choice of the loss function [24], [25]. Pixel-wise losses, such as L1 and MSE, are commonly used to minimise reconstruction errors.

Despite the significance of SR in predicting high-resolution imagery and its applications in SAR data, its use in SAR ship classification remains scarcely explored. While PSNR and SSIM are widely used for image quality evaluation, they have not been incorporated as part of a loss function. The ultimate goal of generating high-resolution images is to enhance the performance of downstream tasks such as classification, detection, and segmentation. However, the majority of SR research remains focused on improving visual or perceptual image quality, often overlooking its practical impact on task-specific performance. In this study, we address this gap by integrating SR into the SAR ship classification pipeline and explicitly incorporating classification feedback into the SR training process. This enables the SR

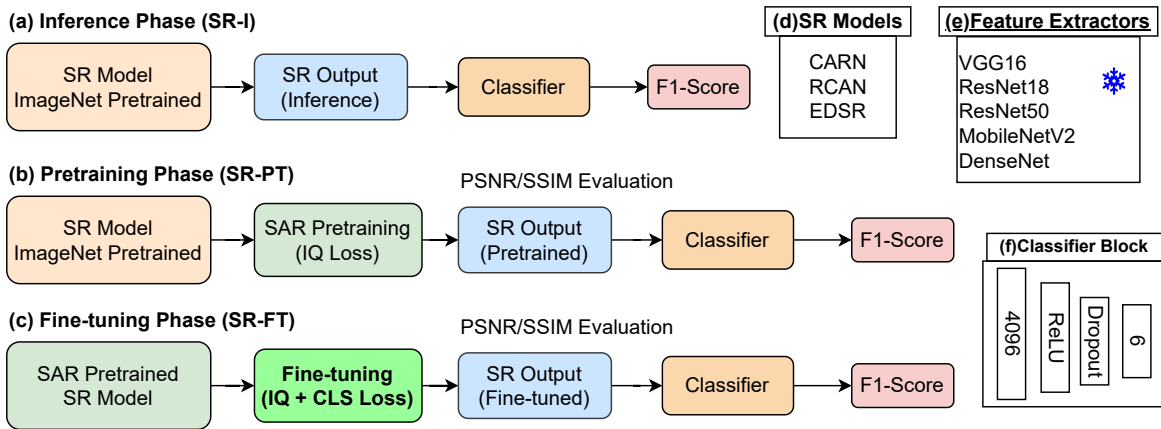


Fig. 2: **Overview of the proposed classification-aware super-resolution pipeline** illustrating the (a) **SR-I**: inference using ImageNet pretrained SR models, (b) **SR-PT**: pretraining the SR models on SAR data using Image Quality (IQ) loss functions, and (c) **SR-FT**: finetuning on SAR data using IQ and classification focused (CLS) loss functions. Evaluation includes PSNR, SSIM, and F1-score. (d) **SR Models**: SR models used in the pipeline. (e) **Feature Extractors**: feature extractors used with frozen weights. (f) **Classifier Block**: classifier layers used for ship classification, where 6 is the number of OpenSARShip classes.

model to be optimised not only for visual quality but also for improving downstream classification accuracy.

### III. METHODOLOGY

We divide our methodology into three stages, as illustrated in Fig. 2. In the first stage (SR-I), we apply ImageNet pretrained SR models to infer high-resolution images and evaluate their performance using classification scores. The second stage (SR-PT) involves training the same SR models inferred in SR-I block on SAR data using loss functions driven by image quality metrics. In the third stage (SR-FT), we perform fine-tuning of the SR models pretrained in the SR-PT stage by incorporating a classification-guided loss function, enabling the model to optimise for both visual quality and classification performance. Each component of this methodology is described in detail below.

#### A. Dataset and Models

We utilized the publicly available OpenSARShip dataset [13], characterized by low-resolution synthetic aperture radar imagery. All classes within OpenSARShip were employed for training and evaluating the SR models; however, for the classification training and evaluation, we specifically selected six classes (Cargo, Tanker, Fishing, Dredging, Passenger, and Tug). Typically, prior studies have limited their analyses to three to five classes due to significant class imbalance issues. We expanded the analysis to six classes to robustly demonstrate the effectiveness of our proposed multi-stage methodology.

The dataset was preprocessed to create two subsets: one containing high-resolution (HR) original images and another consisting of low-resolution (LR) images generated by downsampling the original images by a factor of two. Three pretrained deep-learning-based SR models [26], [27] were selected to perform initial resolution enhancement: Enhanced Deep Super-Resolution (EDSR), Cascading Residual Network (CARN), and Residual Channel Attention Network (RCAN) (Fig. 2(d)). Subsequently, classification performance was evaluated using five widely recognized architectures (Fig. 2(d)): ResNet50, ResNet18 [28], VGG16 [29], MobileNetV2 [30], and DenseNet121 [31]. The classification layers (Fig. 2(e)) of these architectures were modified appropriately, while their feature extraction layers were fine-tuned without freezing to optimize both visual quality and classification accuracy.

#### B. SR-I

To test the effectiveness of the SR models, in the inference module (Fig. 3), we utilized the SR-I models to generate SR images from LR images. Then, to evaluate the quality of images in terms of downstream tasks, we used the obtained SR images to train classification networks. The results are evaluated in terms of F1-scores.

#### C. SR-PT

In the SR-PT block (Fig. 4), the SR models considered in previous block SR-I are trained on all classes of our dataset. As before, the SR images produced by these

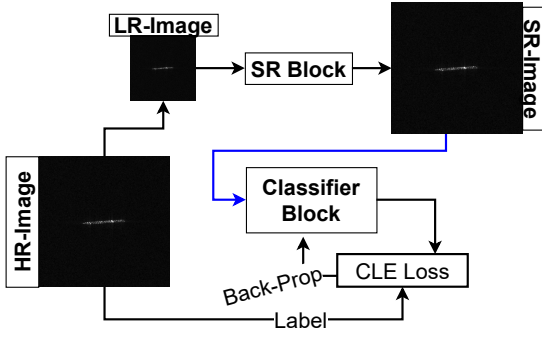


Fig. 3: Overview of the SR-I stage

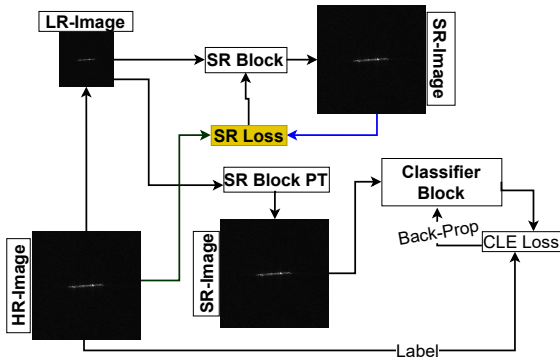


Fig. 4: Overview of the SR-PT stage

trained networks are then used to train the cascaded classification network. Three different loss functions, which are discussed in detail as follows, are used to train the SR models:

1) *L1-Loss*: The first adopted loss function is the L1 [32], which is defined in Eq. (1):

$$L_1 = \frac{1}{N} \sum_{p=1}^N |sr^{(p)} - hr^{(p)}| \quad (1)$$

where  $hr^{(p)}$  is the  $p$ -th original image,  $sr^{(p)}$  is the corresponding super resolved image, and  $N$  is the total number of images.

2) *Combo-Loss*: As previously noted, PSNR and SSIM are commonly used to assess the quality of super-resolved images, but they are rarely leveraged to directly guide the training of super-resolution models. To address this, we propose the ‘‘Combo-Loss’’ function (Eq. (2)) that integrates both metrics to explicitly steer the training process toward generating perceptually and quantitatively improved images. The loss is defined as a weighted sum of a PSNR- and an SSIM-based terms:

$$L_{\text{combo}} = \alpha L_{\text{PSNR}} + \beta L_{\text{SSIM}} \quad (2)$$

PSNR is taken from [33] and is used to define the loss function as follows:

$$L_{\text{PSNR}} = \frac{\text{PSNR}_{\text{max}} - \text{PSNR}}{\text{PSNR}_{\text{max}}} \quad (3)$$

$\text{PSNR}_{\text{max}}$  is the maximum PSNR value and is given by

$$\text{PSNR}_{\text{max}} = 10 \log_{10} \left( \frac{M^2}{\epsilon} \right),$$

where  $M$  is the maximum possible intensity value in the image and  $\epsilon$  is a small constant ( $10^{-8}$  was used in this work) to prevent division by zero.

Concerning the SSIM-based term, it is worth considering that the SSIM between any two images is a real number ranging from 0 to 1 (respectively indicating null to perfect similarity). Thus, given a set of  $N$  couples of super-resolved and high-resolution images, the SSIM [33] term in the loss function (which penalises lower similarity) is defined as:

$$L_{\text{SSIM}} = \frac{1}{N} \sum_{p=1}^N \left( 1 - \text{SSIM}(sr^{(p)}, hr^{(p)}) \right) \quad (4)$$

The resulting combination loss is formulated as a weighted average of the PSNR and SSIM losses, as shown in Equation 2. In the experiments conducted in this study, the weighting parameters were set to  $\alpha = 0.5$  and  $\beta = 0.5$ , assigning equal importance to both components. We determined the internal weights of the Combo loss through quantitative evaluation of PSNR, SSIM, and classification F1-scores. Equal weights of 0.5 and 0.5 provided the best balance between structural fidelity and signal preservation, and therefore this configuration was adopted in our experiments.

3) *Hybrid-Loss*: To investigate whether integrating image quality metrics with a traditional pixel-level loss can enhance model performance, a hybrid loss function was developed. This function is formulated to strike a balance between pixel-wise accuracy, structural integrity, and perceptual quality. The proposed hybrid loss is a weighted sum of three components, respectively related to L1, SSIM and PSNR loss terms.

We determined the internal weights of the Hybrid loss by evaluating six different weight permutations of the (0.7, 0.2, 0.1) configuration across L1, SSIM, and PSNR. The arrangement that assigned the highest weight to L1, a moderate weight to SSIM, and the lowest weight to PSNR was found to yield the most stable performance in both image quality and classification accuracy.

L1 loss is assigned the highest weight (0.7) to preserve fine-grained structural integrity, SSIM is given a moderate weight (0.2) to enhance overall visual quality, and PSNR is assigned the lowest weight (0.1) to avoid overemphasis on pixel-wise accuracy, which may not

always align with SAR-specific features. The precise formulation of hybrid loss is presented in Eq. (5):

$$L_{\text{hybrid}} = 0.7L_1 + 0.2L_{\text{SSIM}} + 0.1L_{\text{PSNR}} \quad (5)$$

#### D. SR-FT

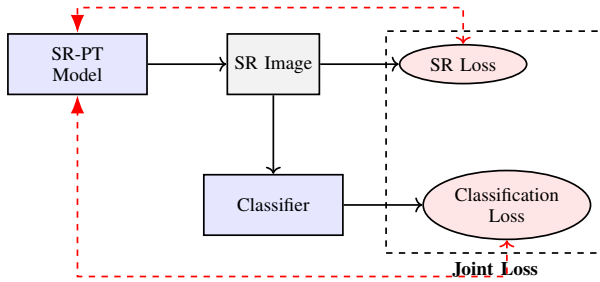


Fig. 5: Architecture of the fine-tuning (SR-FT) stage. The SR model, pre-trained for perceptual quality during SR-PT stage, is fine-tuned using a joint loss function that combines super-resolution loss and task-specific classification loss.

In the fine-tuning block (Fig. 5) resolution is increased through transfer learning by leveraging SR models initially trained using image quality-focused losses in the SR-PT block. We fine-tune the SR models to further refine their ability to reconstruct high-resolution details. A key aspect of this fine-tuning process is the integration of a task-specific objective. Alongside the SR refinement, we compute a classification loss by passing both the generated SR images and the ground-truth HR images through the five classification architectures mentioned in section III.A. The feature extraction layers of these classifiers remain trainable (i.e., unfrozen) to allow them to adapt to the specific image domain, while their final classification layers have been modified for the specific task. Mainly (Fig. 2(f)), we have used a two-layer feedforward head comprising a linear transformation to 4096 units, followed by a ReLU activation and dropout regularization, and a final linear layer projecting to 6 output values, i.e. the chosen number of target classes. This parallel computation of classification loss allows the SR model's fine-tuning to be guided not only by image reconstruction fidelity but also by the ultimate goal of improving downstream classification performance.

One specific novelty of this work concerns the definition of the loss function employed to train the models in this block (see Algorithm 1). To force the SR models to take into consideration the peculiar information intrinsic in the categorical label of the training data, a novel factor is introduced in the loss, which exploits this information to constrain the classifier prediction. In particular, the SR images and HR images are fed into the five different

---

**Algorithm 1** Pseudo-code describing the calculation of the loss in the SR-FT block. The resulting loss function incorporates super-resolution and classification penalty terms

---

- 1: **Input:**  $hr, lr, sr\_model, cls\_model, criteria$
  - 2: **Output:**  $merged\_loss$
  - 3: Criteria  $sr\_criterion, cls\_criterion \leftarrow criteria$
  - 4: Compute SR image:  $sr \leftarrow sr\_model(lr)$
  - 5: HR prediction:  $out\_hr \leftarrow cls\_model(hr)$
  - 6: SR prediction:  $out\_sr \leftarrow cls\_model(sr)$
  - 7: CLS loss:  $cl\_loss \leftarrow cls\_criterion(out\_sr, out\_hr)$
  - 8: SR loss:  $sr\_loss \leftarrow sr\_criterion(sr, hr)$
  - 9:  $merged\_loss \leftarrow cl\_loss + sr\_loss$
- 

classifiers, and the loss is computed by comparing the predicted labels of SR images with those of HR images: (Eq. (6)):

$$L_{merged} = L_{SR} + L_{CLS} \quad (6)$$

where  $L_{SR}$  represents the loss from the SR training block, and  $L_{CLS}$  is the mean-squared error (MSE) [33] loss between the predicted labels.

Once more, following the fine-tuning process, the SR images produced by this SR-FT block are used to train and evaluate the suite of classification models. Performance is then quantified by calculating the final F1-scores for classification and the corresponding IQ metrics, ensuring consistency with the evaluation protocol of the previous stages.

#### E. Training Settings

For our experiment we use the OpenSARShip dataset, where LR images are generated by downscaling the original images to  $32 \times 32$ , while HR images have a resolution of  $64 \times 64$ . The learning rate for both model types (classification and SR) is set to 0.0001, and training is conducted for 10 epochs, with 64 batch size. Adam is used as the optimiser, and random seeds are set to ensure reproducibility. The implementation details, including access to the code, are available online [github.com/cm-awais/sar\\_classification\\_informed\\_sr](https://github.com/cm-awais/sar_classification_informed_sr).

## IV. RESULTS

This section focuses on the impact of fine-tuning SR models using classification-aware loss functions and is structured into three main parts. First, we assess the performance of various SR algorithms using image quality metrics and downstream classification accuracy. Second, we compare different loss functions to evaluate their effectiveness in improving image fidelity and classification outcomes. Finally, we analyse the performance of multiple classification architectures to determine the most suitable models for integration with SR frameworks.

TABLE I: Main evaluation of SR methods (EDSR, CARN, RCAN) across training stages (baseline, pretraining, fine-tuning) using PSNR (dB), SSIM, and F1-score. F1-scores are averaged across five classification models. **Bold** indicates the best per column (training phase); green marks overall best performance for each metric. Averages per model are shown at the bottom, and best average is highlighted using **bold** font.

SR Method	Loss	PSNR ( $\uparrow$ )			SSIM ( $\uparrow$ )			F1-score ( $\uparrow$ )		
		SR-I	SR-PT	SR-FT	SR-I	SR-PT	SR-FT	SR-I	SR-PT	SR-FT
	Baseline LR	-	-	-	-	-	-	61.664	-	-
	Baseline HR	-	-	-	-	-	-	62.614	-	-
EDSR	Baseline	42.47	-	-	0.97	-	-	48.718	-	-
	<i>L1</i>	-	43.95	37.60	-	0.9762	0.9030	-	60.122	58.754
	<i>Combo</i>	-	43.82	37.33	-	0.9708	0.8970	-	59.99	60.914
	<i>Hybrid</i>	-	40.98	35.91	-	0.9334	0.8290	-	60.902	59.558
	<b>Average</b>	42.47	42.92	36.95	0.97	0.9601	0.8763	48.718	60.338	59.742
CARN	Baseline	<b>42.71</b>	-	-	<b>0.98</b>	-	-	57.938	-	-
	<i>L1</i>	-	42.87	37.34	-	0.9757	0.9070	-	60.324	60.844
	<i>Combo</i>	-	43.23	<b>38.14</b>	-	0.9734	<b>0.9260</b>	-	59.73	60.47
	<i>Hybrid</i>	-	43.01	37.95	-	0.9696	0.9220	-	60.668	61.166
	<b>Average</b>	42.71	43.04	37.81	<b>0.98</b>	0.9729	0.9183	57.938	60.241	60.827
RCAN	Baseline	42.66	-	-	0.97	-	-	<b>58.826</b>	-	-
	<i>L1</i>	-	43.75	37.48	-	0.9760	0.9190	-	60.594	62.04
	<i>Combo</i>	-	<b>44.31</b>	37.75	-	0.9756	0.9050	-	<b>61.18</b>	<b>63.41</b>
	<i>Hybrid</i>	-	43.42	37.75	-	0.9678	0.9100	-	60.022	63.15
	<b>Average</b>	42.66	<b>43.83</b>	37.66	0.97	0.9731	0.9113	58.826	60.599	<b>62.867</b>

#### A. SR Models

Throughout the first experimental analysis, the baseline represents the scores calculated without training the SR models on the SAR data (SR-I), whilst pretrained (SR-PT) and fine-tuned (SR-FT) refer to the scores calculated using the proposed methodology described in Fig. 2.

The evaluation of SR methods is conducted from two complementary perspectives: image quality assessment and classification performance. As summarised in Tab. I, SR-PT models generally achieved higher PSNR and SSIM values than their fine-tuned counterparts. RCAN obtained the highest PSNR across all loss functions in SR-PT (Fig. 2(b)), whereas EDSR consistently produced the lowest PSNR in both SR-PT and SR-FT (Fig. 2(c)). A similar trend was observed for SSIM, with RCAN outperforming the remaining models and EDSR ranking lowest across configurations. These results indicate that SAR-specific pretraining substantially enhances perceptual fidelity, even though fine-tuning prioritises classification relevance rather than pure image quality.

In terms of PSNR (Tab. I), the RCAN model optimised with Combo-Loss in the SR-PT block achieved the highest score. Overall, SR-PT models outperformed SR-FT and SR-I models in PSNR. For SSIM, the best score was obtained by the ImageNet-pretrained baseline (SR-I) with a value of 0.98. However, when averaged across all methods, SSIM values were generally similar, indicating consistent structural fidelity across different

training strategies.

To assess whether super-resolved images lead to better classification performance, we evaluated the F1-score of the classification networks introduced in Section III-A, using images generated by the SR models described in Sections III-B, III-C and III-D. The results indicate that RCAN with Combo-Loss, when fine-tuned in the SR-FT block, achieved the highest F1-score. With the exception of EDSR (which may be limited by the absence of adaptive mechanisms), all other SR models showed improved performance during SR-FT compared to SR-PT, demonstrating the broad applicability of our proposed classification-aware strategy.

#### B. Loss Functions

The performance of image quality loss functions followed a similar trend to that of SR models in terms of PSNR and SSIM scores with SR-PT gaining better scores with respect to SR-FT. Consistently, SR-FT led to higher F1-score values across all loss functions (see Tab. II), and both Combo and Hybrid beat the L1 loss function for image quality.

While SR-PT optimises for perceptual fidelity, SR-FT shifts the optimisation toward discriminative feature recovery. This explains why the gain in F1-score is more pronounced for Combo and Hybrid losses: both incorporate structural and perceptual cues that translate more effectively into class-relevant features once classification gradients are introduced. In contrast, L1

TABLE II: The table presents a comparison of F1-scores (averaged across the employed classification models), demonstrating the impact of SR-FT on both classification performance and image quality, LR is low resolution data, HR is high resolution data, SRHR is the data which is super-resolved using SR-I models without any training on current dataset. The best scores for each loss functions are highlighted in bold.

	LR	HR	SRHR	L1	Combo	Hybrid
No Train	61.66	62.614	55.16	–	–	–
SR-PT	–	–	–	60.35	60.30	60.53
SR-FT	–	–	–	<b>60.56</b>	<b>61.60</b>	<b>61.29</b>

focuses primarily on pixel-wise accuracy, which does not necessarily capture the scattering patterns or high-level semantics required for reliable ship discrimination. This behaviour is consistent with the trends observed across the SR models and highlights that loss functions incorporating perceptual structure, provide a stronger foundation for classification-aware fine-tuning.

### C. Classification Models

To evaluate the impact of super resolution on classification, five different classification architectures were tested. VGG16 achieved the highest average F1-score (Fig. 6), while Resnet18 had the lowest. Although the performance difference among lower-performing models was minimal, VGG consistently outperformed others, making it the best-performing classification model for this task.

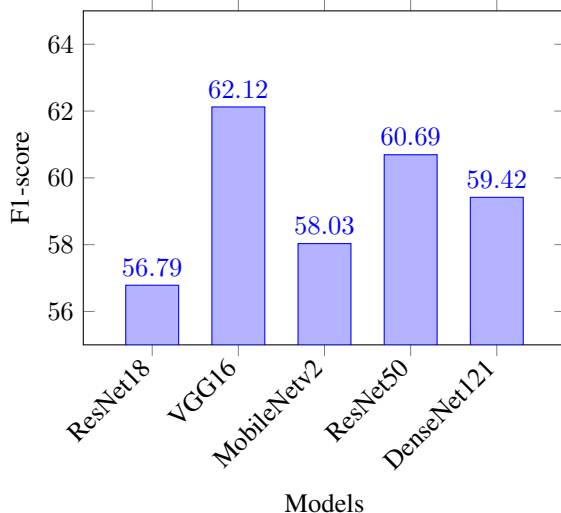


Fig. 6: Average F1-scores of different models clearly indicating **VGG16** as the best-performing model.

To further illustrate how the three SR stages affect the classifier’s decision-making process, we provide Grad-

CAM visualizations in Fig. 7. These show how SR-I and SR-PT tend to focus on background or noisy regions, whereas SR-FT yields compact, target-centered activations aligned with the ship structure.

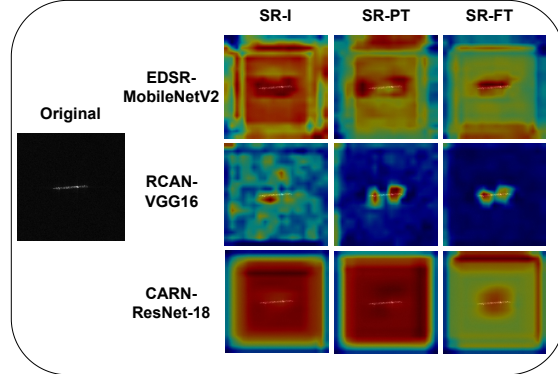


Fig. 7: Grad-CAM visualizations illustrating classifier attention for SR-I, SR-PT, and SR-FT across different SR-classifier combinations. Warmer colors indicate stronger contributions to the classifier’s decision. The SR-FT stage produces compact, target-centered activations focused on the ship, demonstrating improved task relevance.

## V. DISCUSSION

This section discusses the relationship between image quality, classification accuracy, and the behavior of different SR models under the proposed training pipeline. We first examine the effect of resolution scale on both image quality and classification behavior.

The 32×32 to 64×64 configuration represents a low-information regime that is typical in publicly available SAR ship datasets, where targets occupy a very small number of pixels and fine details are inherently limited. Although visual improvements at this scale are subtle, the classification-aware fine-tuning stage consistently enhanced the discriminative features most relevant to the classifier. Similar trends were observed when scaling from 64×64 to 128×128, demonstrating that the proposed approach remains effective across different resolution levels. This further highlights that the primary objective of the framework is to improve task-relevant feature representations rather than purely perceptual quality.

### A. SR Models

Although one might expect classification outcomes to mirror image quality trends, our results show a clear divergence. The only exception was EDSR trained using L1 and Hybrid losses, where a marginal drop in F1-score was observed. All other models showed consistent

TABLE III: **Best Performing combinations:** Classification F1-scores for the best-performing SR-classifier configurations. LR reports performance on low resolution inputs; HR denotes performance on the original high resolution images. SR-PT and SR-FT show F1 scores after applying 2× super resolution to HR images using the pretrained and fine-tuned SR models, respectively. Improvement gives the absolute increase in F1-score achieved by fine-tuning (SR-FT - SR-PT). Boldface highlights the highest fine-tuned score for each classifier.

Model (Loss)	LR	HR	SR-PT	SR-FT	Improvement
VGG16 (CARN-Combo)	61.95	65.03	63.12	<b>65.40</b>	+2.28
MobileNetv2 (RCAN-Combo)	58.92	63.73	60.84	<b>62.40</b>	+1.56
MobileNetv2 (RCAN-L1)	58.92	63.73	60.56	<b>61.35</b>	+0.79
ResNet50 (EDSR-Combo)	61.22	60.40	62.34	<b>62.85</b>	+0.51

improvements in classification performance after fine-tuning. For example, RCAN with Combo Loss improved from an F1-score of 61.18% to 63.41% (Tab. I). These findings confirm our earlier hypothesis [34] that higher image quality, as measured by PSNR and SSIM, does not necessarily translate into better classification performance.

A possible explanation of this behavior can refer to the nature of SAR imagery. SAR images often contain speckle and structural noise. SR methods can unintentionally enhance these noisy regions, especially if the model lacks mechanisms to distinguish signal from noise. EDSR is effective at preserving global image structure due to its deep residual architecture. However, it lacks adaptive mechanisms, such as attention modules or multi-scale feature fusion, which are found in RCAN and CARN. As a result, EDSR struggles to adapt to the semantic characteristics of SAR data during SR-FT (finetuning). This leads to suboptimal performance on the classification task.

In contrast, RCAN and CARN include architectural components (such as residual cascades and channel attention) that may help them focus on class-relevant features (e.g. ship contours and high-frequency textures). They achieve this even if it comes at the cost of traditional image quality metrics. This suggests a trade-off, where preserving semantic features becomes more important than maximizing pixel-level accuracy. These findings highlight the need to evaluate SR models based not only on reconstruction metrics but also on their performance in downstream tasks. For SAR ship classification, where there is a lot of noise, fine-tuning is essential to optimize class-discriminative features, even if it results in lower PSNR or SSIM values.

### B. Loss Functions

Also, in the context of loss functions (Tab. II), it is critical to identify which configuration yields the best performance for SAR ship classification. Since the trends observed in image quality metrics align with those seen across different SR models, our analysis focuses on their

influence on classification performance, specifically the F1-score.

To evaluate this, we used both SR-PT (Fig. 2(b)) and SR-FT (Fig. 2(c)) models to generate SR images, which were subsequently used to train classification networks. The resulting F1-scores were analyzed to assess the relative effectiveness of different SR loss functions. Eventually, it was observed that SR images coming from models trained with Combo and Hybrid loss functions (see section III-C) were consistently classified better than those coming from L1-based SR models. Combo Loss with VGG16 and CARN emerged as the best-performing configuration (see Tab. III, which reports classification on HR images after 2× super-resolution).

An important trend was observed following fine-tuning with classification loss (SR-FT): all models exhibited improved F1-scores compared to their pre-trained counterparts. This indicates that while all three loss functions performed comparably during the pretraining phase, Combo and Hybrid losses provided a significant advantage during fine-tuning, confirming their suitability for enhancing classification performance in SAR applications.

To demonstrate the effectiveness of the proposed loss functions, additional comparative experiments have been performed taking into consideration several different training datasets FUSAR [35], MSTAR [36], and HRSID [37]. For FUSAR and MSTAR, performance was averaged across both SR-PT and ST-FT phases. For HRSID, only the SR-PT stage results were considered, as this dataset is designed for object detection rather than classification. The models were evaluated on each dataset using standard image quality metrics, including PSNR and SSIM.

### C. Classification Models

The classification model exhibited lower scores when trained solely during SR-PT with SR loss functions (Tab. I), as shown in previous results. We attribute this to OpenSARShip being the lowest-resolution dataset, leaving the model with minimal information for training.

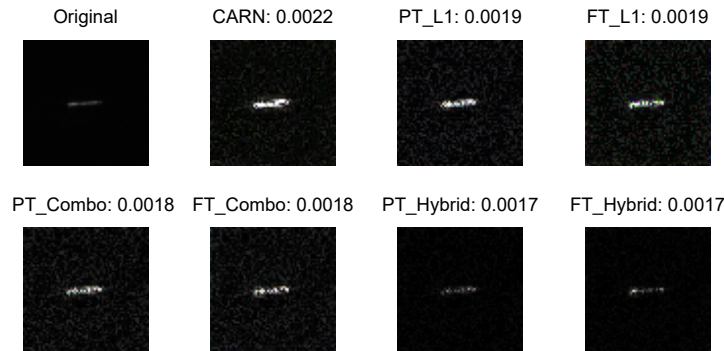


Fig. 8: Error visualisation of super-resolved images using CARN. The images represent the pixel-wise difference between the original image and the SR outputs, with brighter regions indicating areas where the SR models failed to reconstruct details accurately. PT corresponds to the SR image generated by the model from the SR training block, while FT represents the SR image produced by the Fine-Tuned model. Each image includes a numerical error score, where higher values indicate greater reconstruction errors and poorer SR performance.

Additionally, lower resolution causes fewer distinguishable features in images. To address this, classification loss was introduced during fine-tuning, which effectively improved performance in both target quality and classification scores.

The divergence between image quality metrics and classification performance underscores a fundamental limitation in current evaluation methodologies for SAR super-resolution. Traditional metrics like PSNR and SSIM, while valuable for natural image assessment, may not adequately capture the preservation of discriminative features essential for SAR ship classification. Our error analysis (Fig. 8) supports this hypothesis, revealing that reconstruction errors are predominantly concentrated in ship target regions, precisely where accurate feature preservation is most critical for classification performance. This spatial distribution of errors suggests that conventional super-resolution approaches may struggle to maintain the subtle textural and structural characteristics that distinguish different ship classes in SAR imagery.

Fine-tuning enabled the SR models to better adapt to the underlying data distribution, resulting in consistent classification gains compared to pretraining alone. By integrating classification supervision into the SR training process, our framework allows the SR model to optimize for task-relevant features

## VI. CONCLUSION

This paper presented a classification-aware super-resolution framework for SAR ship classification, combining image-quality losses with a classification-guided objective. The proposed methodology improves both perceptual image quality and downstream classification

accuracy. Experiments conducted using three SR models and five classifiers demonstrated that incorporating classification information into the SR training process leads to consistent performance gains.

The main contributions of this work are as follows:

- Two loss functions were introduced to guide SR models toward improved perceptual quality while preserving task-relevant structure.
- A comparative analysis identified CARN as the most effective SR model and VGG16 as the strongest classifier for SAR ship classification.
- It was shown that aligning SR optimization with classification metrics yields higher classification accuracy than traditional pixel-focused training.

Future work will explore extending this classification-aware SR strategy to additional SAR domains and tasks, such as detection and segmentation, by adapting the framework to task-specific architectures.

## ACKNOWLEDGMENTS

This work was supported by National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union – NextGenerationEU. Award Number: Project code CN\_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP D33C22000960007, Project title “National Biodiversity Future Center - NBFC”.

## REFERENCES

- [1] A. B. Antony, "Emerging significance and challenges of maritime industry," *Journal of Research & Development: A Multidisciplinary International Level Referred and Peer Reviewed Journal*, vol. 13, no. 4, pp. 93–96, 2022.
- [2] S. K. Chaturvedi, C.-S. Yang, K. Ouchi, and P. Shanmugam, "Ship recognition by integration of sar and ais," *The Journal of Navigation*, vol. 65, no. 2, pp. 323–337, 2012.
- [3] H. Lang, S. Wu, and Y. Xu, "Ship classification in sar images improved by ais knowledge transfer," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 439–443, 2018.
- [4] H. Lang, J. Zhang, X. Zhang, and J. Meng, "Ship classification in sar image by joint feature and classifier selection," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 212–216, 2016.
- [5] D. Xiang, W. Li, X. Sun, H. Wang, and Y. Su, "Sidelobe suppression for high-resolution sar imagery based on spectral reshaping and feature statistical difference," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–14, 2024.
- [6] D. Xiang, Y. Liu, J. Cheng, X. Lu, Y. Xie, and D. Guan, "Sar target recognition with image generation and azimuth angle feature constraints," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025.
- [7] S. Anwar, S. Khan, and N. Barnes, "A deep journey into super-resolution: A survey," *ACM computing surveys (CSUR)*, vol. 53, no. 3, pp. 1–34, 2020.
- [8] L. Yanshan, Z. Li, X. Fan, and C. Shifu, "Ogsmr: Optical-guided super-resolution network for sar image," *Chinese Journal of Aeronautics*, vol. 35, no. 5, pp. 204–219, 2022.
- [9] L. Wang, M. Zheng, W. Du, M. Wei, and L. Li, "Super-resolution sar image reconstruction via generative adversarial network," in *2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, 2018, pp. 1–4.
- [10] C. He, L. Liu, L. Xu, M. Liu, and M. Liao, "Learning based compressed sensing for sar image super-resolution," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1272–1281, 2012.
- [11] C. Zhang, Z. Zhang, Y. Deng, Y. Zhang, M. Chong, Y. Tan, and P. Liu, "Blind super-resolution for sar images with speckle noise based on deep learning probabilistic degradation model and sar priors," *Remote Sensing*, vol. 15, no. 2, p. 330, 2023.
- [12] D. Xiang, Y. Lu, D. Guan, G. Li, J. Cheng, and B. Li, "Oil spill detection in polsar imagery using composite scattering power entropy and multi-scale hybrid feature fusion network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025.
- [13] B. Li, B. Liu, L. Huang, W. Guo, Z. Zhang, and W. Yu, "Opensarship 2.0: A large-volume dataset for deeper interpretation of ship targets in sentinel-1 imagery," in *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)*, 2017, pp. 1–5.
- [14] Y. Wang, C. Wang, and H. Zhang, "Ship classification in high-resolution sar images using deep learning of small datasets," *Sensors*, vol. 18, no. 9, p. 2929, 2018.
- [15] T. Li, Z. Liu, R. Xie, and L. Ran, "An improved superpixel-level cfar detection method for ship targets in high-resolution sar images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 1, pp. 184–194, 2018.
- [16] C. Lu and W. Li, "Ship classification in high-resolution sar images via transfer learning with small training dataset," *Sensors*, vol. 19, no. 1, p. 63, 2018.
- [17] Y. Dong, H. Zhang, C. Wang, and Y. Wang, "Fine-grained ship classification based on deep residual learning for high-resolution sar images," *Remote Sensing Letters*, vol. 10, no. 11, pp. 1095–1104, 2019.
- [18] Z. Sun, M. Dai, X. Leng, Y. Lei, B. Xiong, K. Ji, and G. Kuang, "An anchor-free detection method for ship targets in high-resolution sar images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7799–7816, 2021.
- [19] O. Karakuş and A. Achim, "On solving SAR imaging inverse problems using nonconvex regularization with a Cauchy-based penalty," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 7, pp. 5828–5840, 2020.
- [20] S.-J. Lee and S.-G. Lee, "Efficient super-resolution method for targets observed by satellite sar," *Sensors*, vol. 23, no. 13, p. 5893, 2023.
- [21] N. Jiang, W. Zhao, H. Wang, H. Luo, Z. Chen, and J. Zhu, "Lightweight super-resolution generative adversarial network for sar images," *Remote Sensing*, vol. 16, no. 10, p. 1788, 2024.
- [22] F. Gu, H. Zhang, C. Wang, and F. Wu, "Sar image super-resolution based on noise-free generative adversarial network," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 2575–2578.
- [23] H. Shen, L. Lin, J. Li, Q. Yuan, and L. Zhao, "A residual convolutional neural network for polarimetric sar image super-resolution," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, pp. 90–108, 2020.
- [24] Y. Jo, S. Yang, and S. J. Kim, "Investigating loss functions for extreme super-resolution," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 424–425.
- [25] S. Shojaei and A. Mahmoudi-Aznavah, "Analyzing different loss functions for single image super-resolution," in *2024 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP)*, 2024, pp. 1–6.
- [26] G. Gouvine, "Super-resolution networks for pytorch," <https://github.com/Coloquinte/torchSR>, 2021.
- [27] G. Gabriel, "Ninasr: Efficient small and large convnets for super-resolution," <https://github.com/Coloquinte/torchSR/blob/main/doc/NinaSR.md>, 2021.
- [28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [29] K. Simonyan, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.
- [31] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
- [32] X. He and J. Cheng, "Revisiting l1 loss in super-resolution: a probabilistic view and beyond," *arXiv preprint arXiv:2201.10084*, 2022.
- [33] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.
- [34] C. M. Awais, M. Reggiannini, and D. Moroni, "Image quality vs performance in super-resolution for sar ship classification," in *2025 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2025, pp. 1–5.
- [35] X. Hou, W. Ao, Q. Song, J. Lai, H. Wang, and F. Xu, "Fusar-ship: Building a high-resolution sar-ais matchup dataset of gaofen-3 for ship detection and recognition," *Science China Information Sciences*, vol. 63, pp. 1–19, 2020.
- [36] E. R. Keydel, S. W. Lee, and J. T. Moore, "Mstar extended operating conditions: A tutorial," *Algorithms for Synthetic Aperture Radar Imagery III*, vol. 2757, pp. 228–242, 1996.
- [37] S. Wei, X. Zeng, Q. Qu, M. Wang, H. Su, and J. Shi, "Hrsid: A high-resolution sar images dataset for ship detection and instance segmentation," *IEEE Access*, vol. 8, pp. 120 234–120 254, 2020.

## VII. BIOGRAPHY SECTION



**Ch Muhammad Awais** is a Ph.D. student at the University of Pisa, specializing in SAR ship classification. Hailing from Pakistan, he has been engaged in deep learning research for the past six years and has worked on SAR ship classification for the last two years.



**Marco Reggiannini** received his M.Sc. in Physics and a Ph.D. in Automation Engineering from the University of Pisa. Currently, he is a researcher at the Signals and Imaging Lab at ISTI-CNR, with more than ten years of experience devoted to multisensor imagery analysis.



**Davide Moroni** received his M.Sc. in Mathematics (Hons.) from the University of Pisa and a Ph.D. in Mathematics from the University of Rome La Sapienza. He is a Senior Researcher and Head of the Signals and Images Lab at ISTI-CNR, specialising in image processing and computer vision.



**Oktay Karakus** (Member, IEEE) received his B.Sc. (Hons.) degree in Electronics Engineering from Istanbul Kültür University, Turkey, in 2009, and his M.Sc. and Ph.D. degrees in Electronics and Communication Engineering from İzmir Institute of Technology (IZTECH), Turkey, in 2012 and 2018, respectively. From 2009 to 2018, he held research assistant positions at several universities in Turkey. In 2017, he was a Visiting Scholar at the Institute of Information Science and Technologies (ISTI-CNR), Pisa, Italy. From 2018 to 2021, Dr. Karakus was a Research Associate in image processing at the Visual Information Laboratory, University of Bristol, UK. He is currently a Lecturer at the School of Computer Science and Informatics, Cardiff University, where he also serves as Deputy Director of the Cardiff Data Science Academy and leads the Remote Sensing Image and Data Analysis (ReSIDA) research group. His research interests include remote sensing image analysis, environmental data science for marine and coastal sciences, ecology and wildlife, machine learning and AI, and football analytics. He is on the editorial board of multiple reputable journals and has published in high-impact venues in remote sensing, environmental monitoring, and artificial intelligence.