SUPPLEMENTARY MATERIAL: A NOVEL CONTEXT-ADAPTIVE FUSION OF SHADOW AND HIGHLIGHT REGIONS FOR EFFICIENT SONAR IMAGE CLASSIFICATION

Author(s) Name(s)

Author Affiliation(s)

This supplementary material provides additional insights into the experimental results presented in the main paper. We extend our analysis with visualizations, qualitative comparisons, and further evaluations of the proposed Context-adaptive sonar image classification framework.

1. SONAR IMAGE CONTEXT-ADAPTIVE FUSION OF SHADOW AND HIGHLIGHT

The intuitive visualization of attention weights for shadow and shadow + highlight, as illustrated in Fig. 1, demonstrates how the model prioritizes features for classification. In Fig. 1(a), the image classified as "Ship" exhibits a dominant alpha (Shadow + Highlight) weight (0.514) over beta (Shadow) (0.486), indicating reliance on combined feature representations. In contrast, Fig. 1(b), where the predicted class is "Human," the beta weight (0.538) surpasses the alpha weight (0.462), suggesting that shadow-based features play a crucial role in this classification. Similarly, Fig. 1(c) shows another "Human" prediction with a balanced attention distribution (Alpha: 0.392, Beta: 0.608), implying the model's adaptability to diverse feature dependencies. Finally, in Fig. 1(d), the "Plane" classification is guided primarily by the alpha weight (0.887) over beta (0.113), reinforcing the significance of highlight features in distinguishing this category. These visualizations emphasize the model's ability to dynamically adjust its attention distribution based on contextual variations, enhancing classification accuracy in real-world scenarios.

2. SHADOW FEATURE PREPROCESSING BLOCK

To effectively combine shadow and highlight regions for improved sonar image classification, the first step is to extract shadow regions from the sonar images. Figure 2 illustrates shadow regions for different objects, including planes, ships, humans, and mines. This extraction is performed using a shadow preprocessing block, as described in Section 3.1.1 of the main paper.

3. BENCHMARK DATASET CREATION

The S3SIMULATOR+ benchmark dataset consists of naval mine-like objects, which are challenging to collect due to high acquisition costs and limited availability of real-world data. To support AI applications and generative AI models, high-quality datasets are essential, and S3SIMULATOR+ provides a solution, as illustrated in Figure 3.

As shown in Figure 4 the dataset creation process begins with real-world observations of naval mines, revealing three primary shapes: cylindrical, truncated cone, and spherical. Based on these observations, 3D models were developed using AutoCAD Fusion. These models were then deployed in a Gazebo simulation environment to render objects and their corresponding shadows from various orientations. The simulation enabled the recording of videos at different ranges, positions, and seabed conditions, ensuring diversity in the dataset.

Real-world naval mine datasets are scarce, particularly at short ranges, where mines often appear as small, indistinct objects. To bridge this gap, computational imaging techniques were applied to enhance realism by incorporating grayscale sonar effects, noise patterns, color mapping, and nadir zone artifacts, ensuring that the synthetic dataset closely resembles real sonar imagery.

4. QUANTITATIVE EVALUATION OF REGION-AWARE DENOISING MODEL

The performance comparison in Table 1 highlights the effectiveness of the proposed Region-Aware Denoising model in preserving image features while denoising. Across all datasets and classes, our model achieves consistently higher SSIM values, indicating superior structural preservation compared to traditional methods like Mean, Median, and Wiener filters. For backscatter noise, our model achieves the highest SSIM for both the *Plane* (0.48) and *Ship* (0.52) classes, demonstrating its robustness in retaining features. Under Gaussian noise, our model maintains its lead, achieving the highest SSIM for the *Plane* (0.45) and *Ship* (0.52) classes. In real-world noisy scenarios, our model outperforms all other

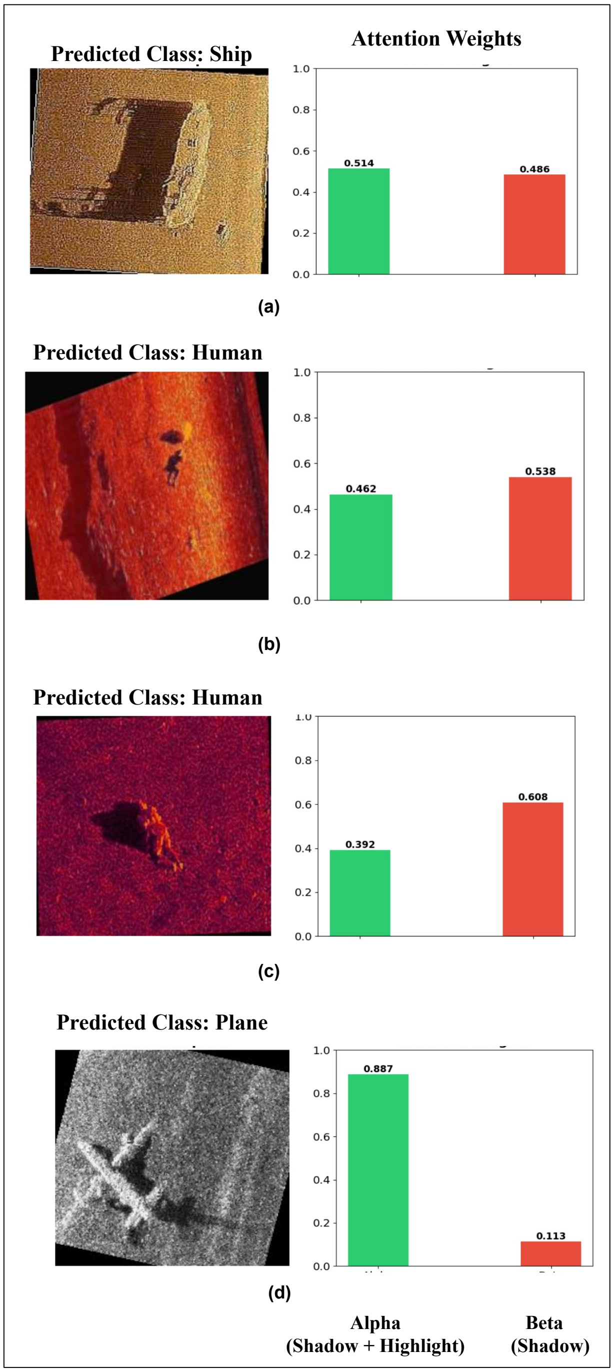


Fig. 1: Visualization of attention weight distributions for different predicted classes. Alpha represents the combined influence of shadow and highlight features, while Beta corresponds to shadow-only contributions.

methods with SSIM values of 0.63 and 0.72 for the *Plane* and *Ship* classes, respectively, validating its practical applicability. While Mean and Wiener filters occasionally report lower MSE values, their lower SSIM suggests over-smoothing and loss of critical features. our model strikes a strong balance across PSNR, SSIM, and MSE metrics, making it highly effective in denoising without compromising essential image details.

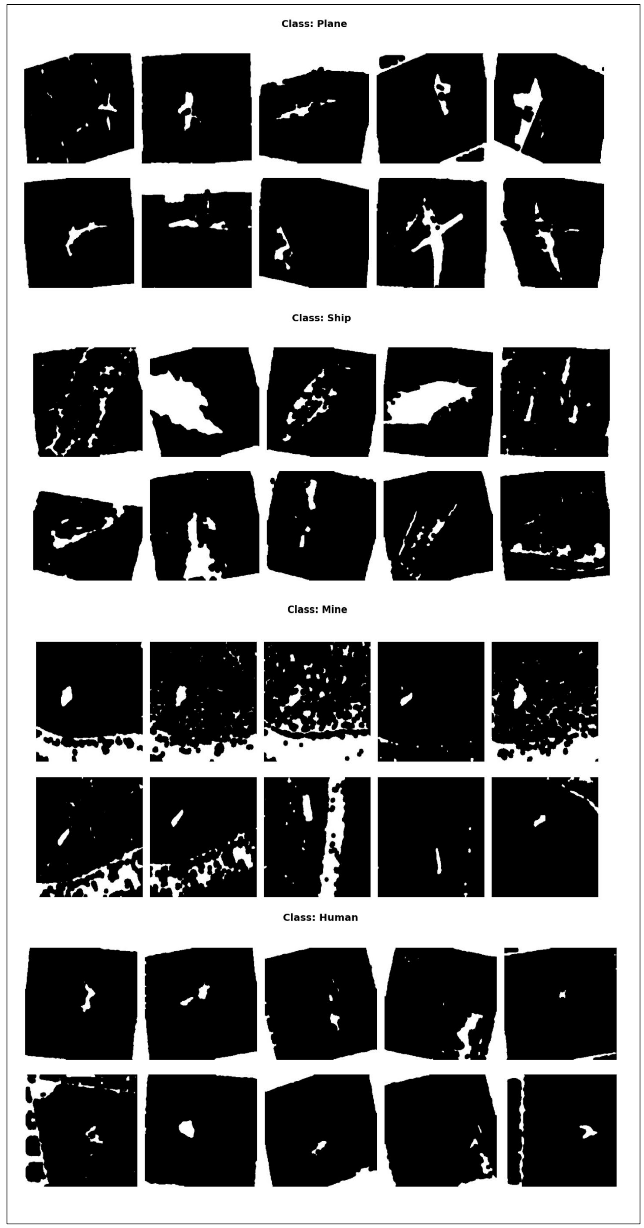


Fig. 2: Sample Images from Shadow-Feature Preprocessing block of plane, ship, mine, human class.

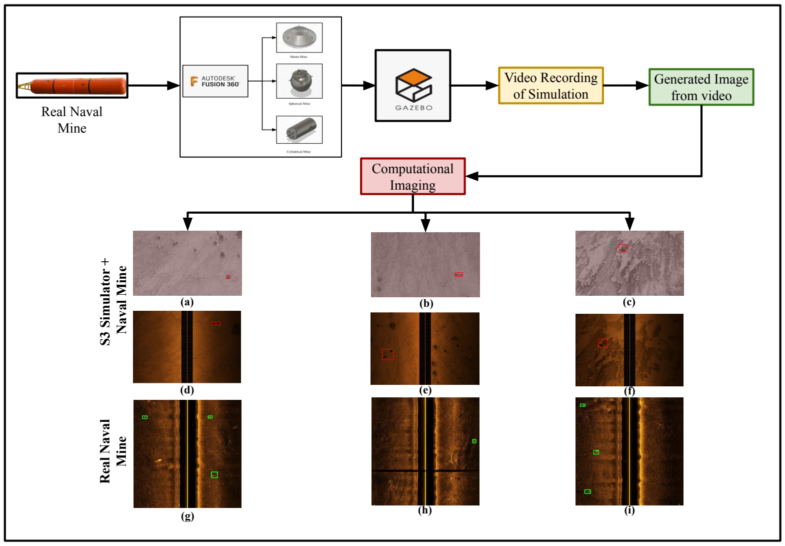


Fig. 3: Illustration of the S3SIMULATOR+ framework. The diagram highlights the synthetic dataset generation process and its integration into the proposed classification pipeline.

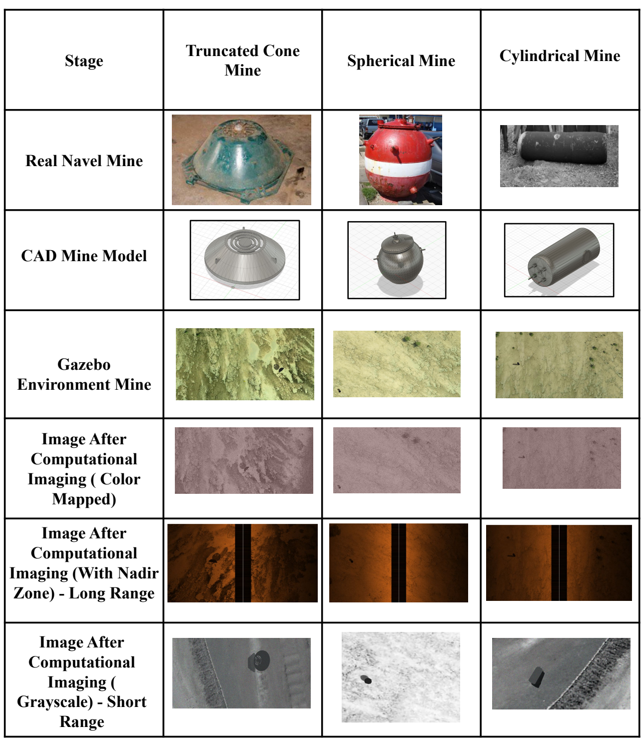
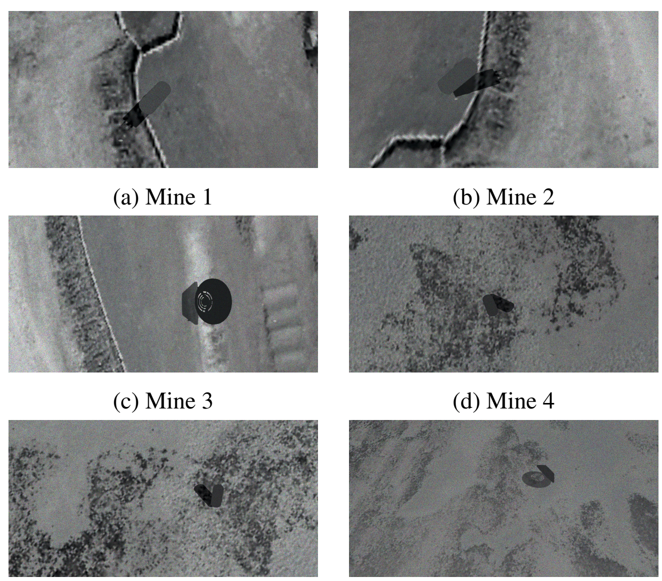


Fig. 4: Qualitative representation of mine classes across various stages and imaging conditions. Columns correspond to different mine types (Cylindrical, Truncated Cone, Spherical), and rows depict the stages, from original references to processed sonar images, including close-range and long-range variations.



(a) Mine 1 (b) Mine 2

(c) Mine 3 (d) Mine 4

(e) Mine 5 (f) Mine 6

Fig. 5: Sample Images of Mine Like Objects(MLO) from S3Simulator+ dataset.

Table 1: Performance comparison of various denoising methods. Bold values indicate the best-performing metric.

Dataset	Class	Method	PSNR	SSIM	MSE
Backscatter	Plane	Region aware denoising (Ours)	18.63	**0.48**	1409.34
		Mean	**18.99**	0.36	**1167.98**
		Median	18.79	0.36	1239.94
		Wiener	**18.99**	0.36	**1167.98**
	Ship	Region aware denoising (Ours)	19.29	**0.52**	1290.99
		Mean	**19.52**	0.41	**1093.44**
		Median	19.38	0.42	1156.77
		Wiener	**19.52**	0.41	**1093.44**
Gaussian	Plane	Region aware denoising (Ours)	**20.22**	**0.45**	1040.21
		Mean	20.08	0.30	**768.21**
		Median	20.11	0.31	770.88
		Wiener	20.08	0.30	**768.21**
	Ship	Region aware denoising (Ours)	**21.61**	**0.52**	627.80
		Mean	21.19	0.37	**540.85**
		Median	21.29	0.38	531.64
		Wiener	21.19	0.37	**540.85**
Real	Plane	Region aware denoising (Ours)	22.38	**0.63**	1107.71
		Mean	24.04	0.54	**512.39**
		Median	**24.48**	0.55	506.83
		Wiener	24.04	0.54	**512.39**
	Ship	Region aware denoising (Ours)	24.35	**0.72**	601.52
		Mean	25.54	0.65	**266.82**
		Median	**26.11**	0.67	249.73
		Wiener	25.54	0.65	**266.82**