Ocean Systems Engineering, *Vol. 15, No. 2 (2025) 219-239*
https://doi.org/10.12989/ose.2025.15.2.219

An improved method of detecting the horizon for tracking maritime obstacles

Yun-Sik Kim[1a], Myung-Il Roh[*2], Ha-Yun Kim[1b] and In-Chang Yeo[1c]

[1]*Department of Naval Architecture and Ocean Engineering, Seoul National University, Republic of Korea*
[2]*Department of Naval Architecture and Ocean Engineering, and Research Institute of Marine Systems Engineering, Seoul National University, Republic of Korea*

(Received May 17, 2025, Revised June 20, 2025, Accepted June 21, 2025)

Abstract. For safe navigation of ships, it is essential to accurately detect and continuously track surrounding obstacles. The horizon, serving as the boundary between the sky and the sea, plays a crucial role in enabling ships to maintain precise routes and effectively assess the positions of obstacles. Reliable horizon detection is, therefore, highly significant. Conventional horizon detection methods, such as the Hough transform and edge-based approaches, have shown good performance in relatively simple environments. However, their accuracy significantly decreases in realistic maritime environments, which involve complex factors such as terrain, obstacles, and reflections. To address these limitations, this study proposes a novel horizon detection method that fine-tunes the SAM (Segment Anything Model), a deep learning model specifically designed for maritime images. An efficient adapter-based fine-tuning technique was implemented on the SAM's mask decoder, enabling the model to effectively learn the distinct visual and structural characteristics of maritime environments. Experimental evaluations demonstrated that the method combining the SAM fine-tuning with the vertical edge response approach achieved superior performance, significantly reducing height error and slope error by an average of 75.58% and 70.17%, respectively, even in highly complex environments. These findings highlight the superior accuracy and robustness of the proposed method, indicating its substantial potential for practical applications in autonomous navigation systems and enhanced maritime safety.

Keywords: autonomous ship; fine-tuning; horizon detection; maritime obstacle detection; SAM (Segment Anything Model)

1. Introduction

1.1 Research background

For ships to navigate safely, it is crucial to accurately perceive their surroundings and swiftly detect and track potential obstacles. In particular, the capability to identify other ships or fixed

*Corresponding author, Professor, E-mail: miroh@snu.ac.kr
[a] M.Sc. Student, E-mail: kyus99@snu.ac.kr
[b]Ph.D. Student, E-mail: hykim1274@snu.ac.kr
[c]Ph.D. Student, E-mail: iys02191@snu.ac.kr

http://www.techno-press.org/?journal=ose&subpage=7 ISSN: 2093-6702 (Print), 2093-677X (Online)

structures at an early stage and continuously track their positions is essential for collision avoidance at sea (Zhang *et al.* 2021). Since maritime collisions can result in significant loss of life and property, technologies that enable rapid identification and immediate response to potential hazards are urgently required. To address these requirements, significant advancements have been made in smart ship technologies, particularly in autonomous navigation systems, leading to increased interest in USVs (Unmanned Surface Vessels). Autonomous navigation systems continuously monitor their surroundings, utilizing various sensors to detect risks proactively and respond swiftly (Lee *et al.* 2023, Kim *et al.* 2024). Specifically, recent studies have extensively explored camera-based obstacle detection methods using onboard cameras (Lee *et al.* 2021, *et al.* 2024). Cameras are favored for maritime monitoring and obstacle detection due to their ease of installation, low maintenance, and cost-effective performance. Moreover, real-time video analysis offers high accessibility and rapid information acquisition in diverse maritime environments. Consequently, camera-based obstacle detection methods effectively complement the advantages and disadvantages of existing sensors, significantly enhancing maritime navigation safety.

For camera-based obstacle detection methods to operate effectively, precise horizon detection, which serves as a critical reference point in maritime imagery, is essential. In maritime images, the horizon serves as the boundary between the sky and the sea, playing a crucial role in localizing the positions of obstacles and other ships. Conventional horizon detection methods, typically employing traditional computer vision algorithms such as edge detection and the Hough transform, perform well in simple maritime environments. However, real-world maritime environments often feature terrain or obstacles that partially overlap the horizon, posing challenges to these methods and increasing the likelihood of false positives and missed detections (Zardoua *et al.* 2024). Therefore, it is essential to develop more advanced and robust methods for horizon detection.

Recently, with the rapid advancement of deep learning in image processing, methods based on deep learning that deliver superior performance across diverse maritime environments have been actively investigated. Notably, SAM (Segment Anything Model) has emerged as a state-of-the-art deep learning model in segmentation, demonstrating exceptional performance and robustness. It is recognized as a promising tool for effectively processing images under complex and various maritime environments (Kirillov *et al.* 2023). This study proposes an enhanced horizon detection method based on the SAM, aiming to address the limitations of conventional horizon detection methods. In particular, the goal is to achieve accurate and reliable horizon detection even in realistic and complex environments where terrain or obstacles overlap the horizon. This method is anticipated to substantially enhance maritime navigation safety by improving the localization and tracking accuracy of both obstacles and ships, thereby contributing to safer maritime traffic environments

1.2 Related works

Horizon detection is a fundamental and essential task for ensuring navigation safety and effective obstacle localization in maritime environments. Extensive research has been conducted in this area for many years. Previous studies have primarily focused on edge-based methods for extracting or classifying horizon edges. A summary of these studies is presented in Table 1 below.

Aggarwal and Karl (2006) proposed an effective horizon detection method utilizing a regularized Hough transform. Gershikov (2013) detected the horizon by analyzing strong vertical edge responses using a vertical edge response method. These conventional methods offer simplicity and intuitive computations but exhibit significant limitations in detection accuracy

Table 1 Comparison of the related works with this study

Related works	Segmentation	Extraction method of horizon edges	Classification method of horizon edges
Aggarwal and Karl (2006)	No	No	Hough transform
Gershikov (2013)	No	No	Vertical edge response
Fefilatyev *et al*. (2006)	No	Hough transform	Machine learning approach
Ahmad *et al.* (2013)	No	Machine learning (SVM)	Machine learning approach
Jeong *et al*. (2018)	No	Deep learning (NN)	Regional approach
Kuanar *et al*. (2022)	No	Deep learning (CNN)	Hough transform
This study	**Yes**	**Deep learning (SAM fine-tuning)**	**Vertical edge response**

under complex maritime environments. To address these limitations, Fefilatyev *et al.* (2006) combined machine learning methods with the Hough transform, improving overall detection accuracy. Ahmad *et al.* (2013) introduced a local feature-based horizon detection method using SVM (Support Vector Machines), achieving superior detection accuracy compared to conventional methods. However, this method encountered issues such as complicated parameter settings and high computational costs during feature extraction, making real-time detection infeasible. To further enhance robustness and reduce complexity, recent studies have actively investigated deep learning-based horizon detection methods that can automatically extract complex features and offer superior generalization performance. Jeong *et al.* (2018) proposed a neural network-based method utilizing a regional approach for horizon detection. Additionally, Kuanar *et al*. (2022) effectively detected horizons in nighttime images using a multi-path dilated CNN (Convolutional Neural Network). Nonetheless, further improvements in detection accuracy remain necessary, particularly for complex maritime environments.

In summary, previous studies have demonstrated limited performance in realistic maritime environments involving fog, waves, and obstacles, primarily due to their reliance on image processing and deep learning models tailored for specific environments. To overcome these limitations, this study proposes an advanced horizon detection method based on the recently developed SAM, renowned for its superior performance and robustness. The proposed method achieved highly accurate horizon detection even under complex maritime environments, offering practical applicability in real-world maritime navigation systems

1.3 Overall process of the proposed method

The method proposed in this study comprises three main stages: data preprocessing, deep learning-based horizon edge detection, and horizon edge classification. The overall process of these stages is summarized in Fig. 1, with detailed descriptions provided below.

The first stage involves data preprocessing. In this step, labeled data is used to calculate horizon points that precisely indicate the location of the horizon. These computed horizon points are subsequently employed to generate corresponding mask images automatically. This process is uniformly applied across the entire dataset, efficiently producing mask images for all input images. This preprocessing stage ensures optimal input data quality for training the deep learning model. The second stage focuses on deep learning-based horizon edge detection. In this study, SAM

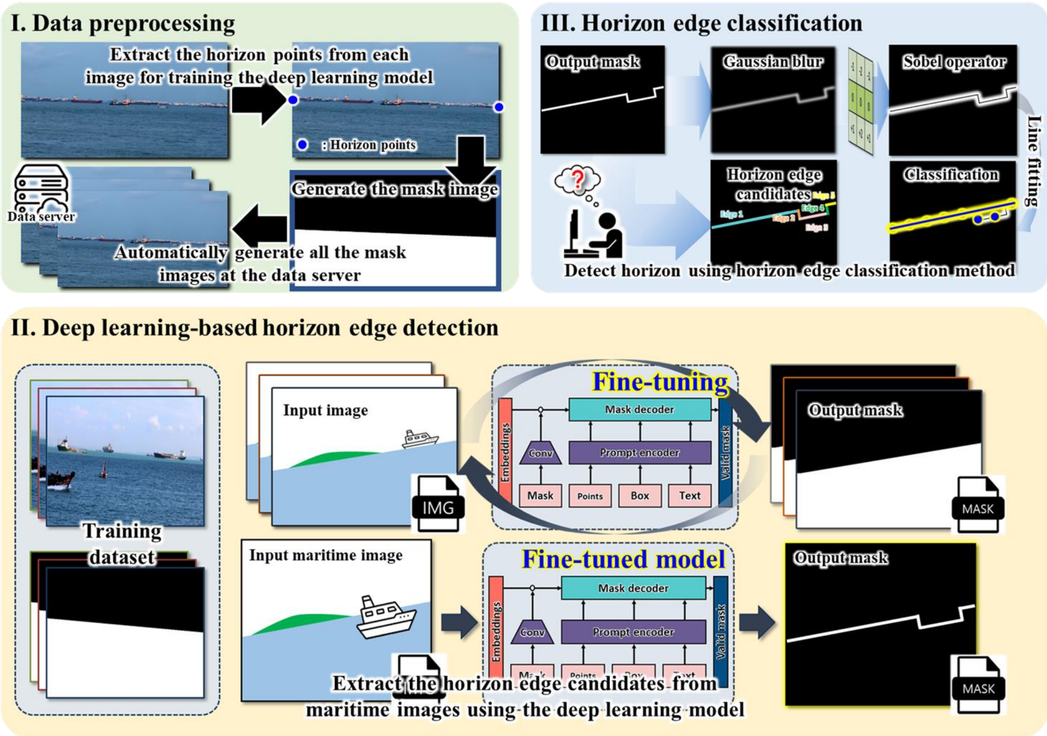

Fig. 1 Overall process of this study

undergoes fine-tuning using images captured in realistic maritime environments, along with binary masks that distinguish the sky and sea regions. Through this fine-tuning technique, the SAM effectively learns maritime-specific visual features, enabling accurate and stable detection of horizon edge information even under challenging environments, such as complex backgrounds and varying illumination. The third stage involves horizon edge classification. Based on the mask images generated by SAM fine-tuning, this stage clearly identifies and classifies horizon edges, ultimately determining the precise location of the horizon. Here, mathematical analysis methods are employed to evaluate structural information in mask images quantitatively, systematically deriving the horizon through a structured, step-by-step algorithm.

2. Data preprocessing

This study aims to accurately detect horizons in maritime images and effectively localize the positions of other ships using deep learning-based methods. Deep learning-based image processing typically requires large-scale datasets for model training. However, collecting datasets related to ship navigation in real maritime environments is a practical challenge. Therefore, researchers often utilize publicly available datasets for model training and performance evaluation.

In this study, we utilized SMD (Singapore Maritime Dataset) published by Prasad *et al*. (2017). The SMD comprises 81 videos recorded from July 2015 to May 2016 across various locations,

Fig. 2 Example image of the SMD (Prasad *et al*. 2017)

routes, and periods in Singapore. All videos were captured using a Canon 70D camera with a resolution of 1080 x 1920 pixels. The dataset is categorized into two groups: videos recorded from shore (onshore) and videos recorded onboard ships (onboard). For this research, we specifically selected onboard videos, as these more closely resemble actual maritime navigation environments. Frames extracted from these onboard videos were subsequently used as training data. An example of the utilized SMD is illustrated in Fig. 2.

SMD includes images captured across diverse maritime environments along with corresponding labeling information. This dataset encompasses a diverse range of weather conditions, illumination changes, wave heights, and ships of varying sizes and shapes. Consequently, it is particularly suitable for image analysis and training deep learning models specialized for maritime environments. SMD comprises a total of 51 maritime videos containing over 33,440 individual images. Each image provides labeled information in the form of bounding boxes indicating the pixel locations of other ships, as well as ground-truth horizon data marking the boundary between the sky and the sea. Notably, for the specific horizon detection problem addressed in this study, these precisely labeled horizon positions serve as a reliable reference for model training and performance evaluation. For each frame, the dataset provides two horizon points in XML format (structXML), clearly representing the location of the horizon and the direction vector. This direction vector is defined as a unit vector that describes the slope of the horizon and its orientation through cosine and sine values. In this study, labeled information was utilized to generate mask images accurately reflecting the location of the horizon. Specifically, for each frame, the two horizon points where the horizon intersects the left and right image boundaries were calculated using the provided center coordinates (X, Y) and the direction unit vector ($\cos\alpha$, $\sin\alpha$). Mathematically derived intersection points between the horizon direction vector and the left and right image boundaries were computed for precise horizon positioning. An example of the calculated horizon points is illustrated in Fig. 3.

Subsequently, binary mask images were generated based on the straight line connecting the two calculated horizon points. In these masks, pixels corresponding to the sky region above the horizon were assigned a value of 0. In contrast, pixels corresponding to the sea region below the horizon were assigned a value of 1. In this study, an automated algorithm was implemented to generate

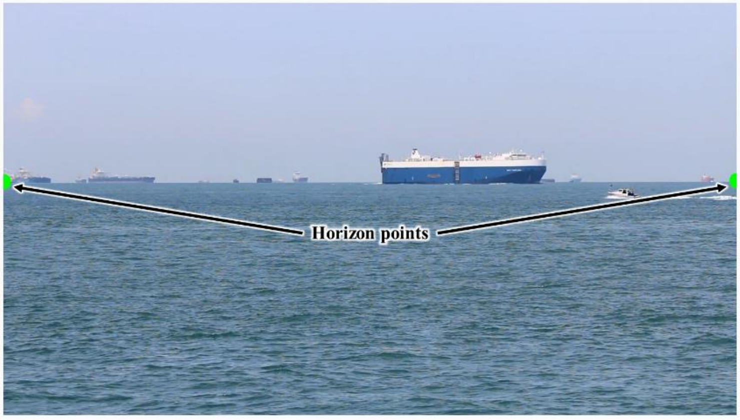

Fig. 3 Computation of horizon points from the labeled data

Fig. 4 Generated binary mask image for horizon detection

masks efficiently, applying the procedure uniformly across all images in the dataset. This allowed mask data to be collected systematically and effectively. An example of the generated binary mask images is presented in Fig. 4.

3. Deep learning-based horizon edge detection

3.1 Overview of the SAM (Segment Anything Model)

SAM is a general-purpose image segmentation model developed by Meta AI Research (Kirillov *et al.* 2023). Utilizing user-provided prompts in the form of points, boxes, or masks, SAM accurately segments a wide variety of images without requiring additional training. The model is pretrained on the SA-1B (Segment Anything 1 Billion) dataset, comprising 1.1 billion masks and

11 million images, enabling it to perform robust zero-shot segmentation even on novel data not encountered during training. While conventional image segmentation models are typically trained on specific tasks or datasets, limiting their adaptability, SAM demonstrates strong generalization capabilities driven by extensive pretraining on a large-scale image dataset.

SAM consists of three primary components: an image encoder, a prompt encoder, and a mask decoder. The image encoder, based on a pretrained ViT (Vision Transformer), efficiently extracts both global image context and detailed features, enabling it to effectively distinguish objects from the background (Vaswani *et al*. 2017, Dosovitskiy *et al.* 2021). ViT partitions images into small patches, treating each patch as an independent token, and learns relationships between the tokens using a self-attention mechanism. This enables SAM to accurately comprehend both the global context and the detailed characteristics of objects, including their shape and position. The prompt encoder converts user-provided prompts into spatial embeddings indicating object locations within the image. Prompts can be given in various forms, such as points, boxes, or text, and are translated into spatial vector representations that are passed to the mask decoder. SAM can effectively interact with a wide range of user interfaces. Finally, the mask decoder combines visual features extracted by the image encoder with spatial embeddings from the prompt encoder to generate the final segmentation masks. The decoder progressively refines masks through multiple decoding stages, clearly delineating boundaries between objects and the background.

Recently, SAM has demonstrated outstanding performance in diverse application domains, including medical image analysis (Zhang *et al*. 2024), autonomous driving systems (Zhang *et al*. 2023), and drone imagery analysis (Shahraki *et al.* 2025). In maritime image analysis, SAM is also evaluated as an effective tool. The prompt-based approach of SAM enables precise and robust horizon edge detection, overcoming limitations of conventional methods in complex maritime environments, such as varying illumination and complex backgrounds. Using minimal prompts, SAM produces highly accurate horizon segmentation masks, significantly enhancing the precision of maritime obstacle localization.

3.2 Fine-tuning of the SAM for horizon detection

3.2.1 Limitations of the pretrained SAM for maritime horizon detection

The pretrained SAM, designed as a general-purpose segmentation model, exhibits excellent performance on diverse images. However, it presents several fundamental limitations when directly applied to maritime environments. Specifically, horizons in maritime images have distinctive structural and visual characteristics. Consequently, using the pretrained SAM without modification often results in reduced accuracy and unstable masks when detecting horizon edges. These limitations of the pretrained SAM are illustrated in Fig. 5.

First, horizons in maritime images typically exhibit simple and consistent structures, often appearing as straight or gently curved lines. In contrast, the images primarily learned by SAM include multi-dimensional features, such as complex shapes, irregular contours, and diverse textures and colors associated with people and various objects. Due to these inherent differences, the simple horizon structure may not align well with the complex, detail-oriented feature extraction approach of SAM. As a result, SAM frequently demonstrates lower accuracy when segmenting horizons in maritime images, significantly reducing the reliability of horizon detection. Second, maritime environments frequently present various visual challenges, such as fog, waves, and intense sunlight reflections, making horizon boundaries ambiguous or irregular. Consequently, horizons often lack clear, distinct edges. SAM, trained primarily on objects with clearly defined

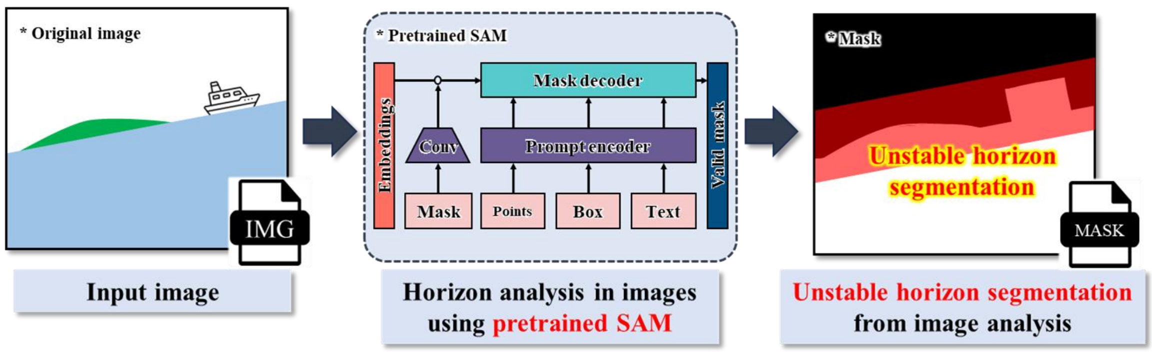

Fig. 5 Limitations of the pretrained SAM for maritime horizon detection

boundaries, struggles to accurately detect horizons characterized by blurred or indistinct edges. Third, horizons represent global boundaries separating the sky from the sea within the entire image. Here, global information refers to the overall structure or contextual meaning of the image as a whole, whereas local information pertains to detailed features or textures within specific regions. Since SAM primarily focuses on detailed, local features during its training, it may not effectively utilize contextual, global structural information. Consequently, its performance becomes limited in segmentation tasks such as horizon detection, where understanding the global structure and overall context is crucial.

3.2.2 Fine-tuning and application to maritime horizon detection

In this study, we determined that domain adaptation of SAM, specifically tailored for horizon detection, is essential given the distinct characteristics of maritime images. Previous studies, such as those by Jebari *et al*. (2025), have also highlighted significant improvements achieved by employing fine-tuning techniques specifically tailored to maritime applications. They demonstrated that fine-tuning models effectively address domain-specific challenges, resulting in enhanced performance and robustness in maritime tasks. To achieve this, we employed the fine-tuning technique that maintains the SAM's original architecture while providing additional training focused on reliably distinguishing the characteristic visual information, boundaries, and positional patterns of maritime horizons. This fine-tuning process is illustrated in Fig. 6.

The fine-tuning technique for SAM employed in this study was designed based on a comprehensive comparative study conducted previously (Gu et al., 2025). In that research, various fine-tuning techniques were evaluated using medical imaging data, among which the PEFT (Parameter-Efficient Fine-Tuning) technique demonstrated superior efficiency and performance. Based on these findings, this study conducted fine-tuning exclusively on the mask decoder of SAM by employing an adapter-based fine-tuning technique.

As described in Section 3.1, SAM comprises three main components: an image encoder, a prompt encoder, and a mask decoder. To facilitate automatic horizon detection, we modified the SAM's structure by removing the prompt encoder, enabling the SAM to generate segmentation masks automatically without user-supplied prompts. Generally, the prompt encoder identifies object locations based on spatial hints such as user-specified points, boxes, or masks. However, as the goal of this study was to achieve fully automated horizon detection without manual

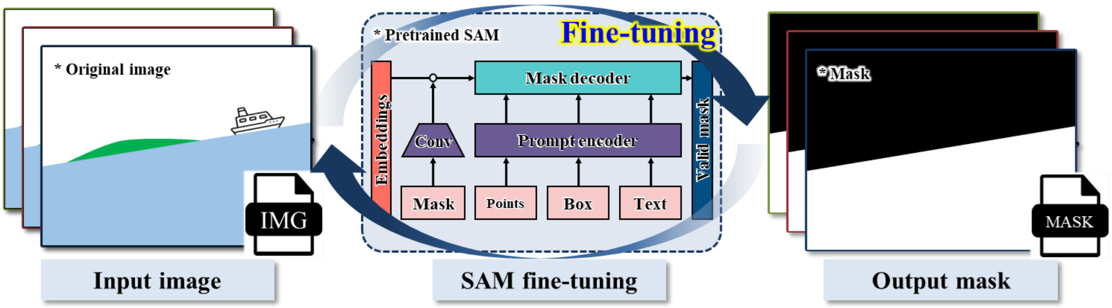

Fig. 6 Fine-tuning process of the SAM for maritime horizon detection

intervention, we substituted these spatial hints with dummy inputs (None), thereby allowing the SAM to generate masks directly from the input images autonomously. For the training data, we utilized the SMD, as previously described in Section 2, which provides paired sets of original maritime images and corresponding binary masks indicating horizon regions. In total, 33,400 images from the SMD dataset were randomly divided into training, validation, and test sets at an 8:1:1 ratio. The random splitting was performed to ensure that no data leakage occurred between these subsets. The input images were resized to a resolution of 1024 x 1024 pixels, as per the SAM's input specifications, and single-channel grayscale images were converted to RGB format. Output masks were downsampled to a resolution of 256 x 256 pixels to enhance computational efficiency.

In this study, the image encoder of SAM, which had already effectively learned visual feature extraction capabilities from general image datasets, remained frozen during the fine-tuning process. Fine-tuning was exclusively applied to the mask decoder, utilizing the ViT-B (Vision Transformer-Base) backbone due to its balanced performance and computational efficiency, making it well-suited for maritime horizon detection tasks. The adapter-based fine-tuning technique involves adding two lightweight MLP (Multi-Layer Perceptron) modules within each Transformer block, thereby limiting additional parameter updates to only a small fraction of the overall model. Specifically, the adapter modules were inserted after attention and MLP blocks within the mask decoder, as illustrated in Fig. 7. Consequently, only approximately 0.14% of the total parameters required additional training. The adapter-based fine-tuning technique was chosen over alternatives such as full model fine-tuning and LoRA (Low-Rank Adaptation; Hu *et al*. 2022) due to its optimal balance of computational efficiency and stability. Unlike full model fine-tuning, it minimizes computational overhead and the risk of catastrophic forgetting. Compared to LoRA, adapters integrate more naturally into Transformer structures, effectively capturing maritime-specific features with fewer parameters. This technique effectively prevents common problems encountered during fine-tuning, such as model collapse, where the model repeatedly converges to limited outputs, and catastrophic forgetting, where previously learned essential information is lost. Thus, the proposed method enables efficient and stable domain adaptation. During the training process, a combination of cross-entropy loss and dice loss was used as the objective function. The AdamW optimizer was employed, with a learning rate of 0.001 and a weight decay of 0.1. The maximum number of training epochs was set at 200, and model performance on the validation dataset was evaluated every 100 iterations. Early stopping was applied when validation performance ceased to improve, effectively preventing overfitting.

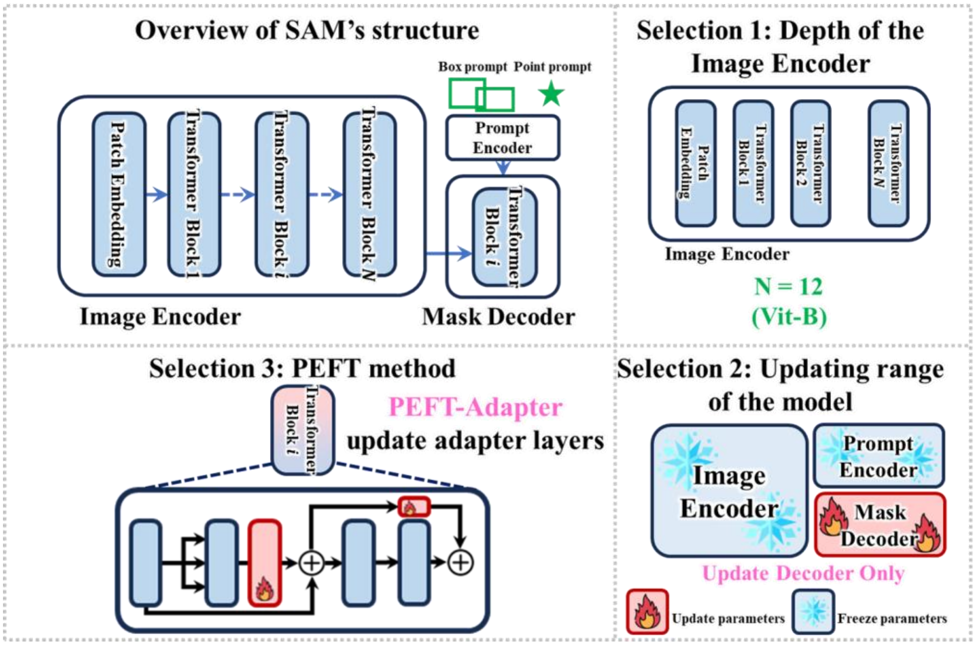

Fig. 7 Overview of the SAM's structure (Gu *et al*. 2025)

Additionally, model states were periodically saved throughout the training process to facilitate further evaluation and subsequent experimentation.

The fine-tuning technique proposed in this study enabled SAM to effectively learn the global and continuous boundary patterns characteristic of maritime horizons, significantly enhancing detection accuracy and stability, which were previously difficult to achieve with the original model. Consequently, this technique is anticipated to provide a robust foundation for practical applications such as horizon edge classification and maritime obstacle tracking.

4. Horizon edge classification

The binary mask produced by SAM fine-tuning effectively distinguishes the sky and the sea regions within an image. However, the horizon is not merely a set of pixels but must be clearly defined as a continuous and structurally consistent boundary line. Consequently, further analysis and processing are necessary to identify the precise location and shape of the horizon accurately. In this study, three post-processing methods were applied to quantitatively determine the position of the horizon from the binary masks generated by SAM fine-tuning. Firstly, a Hough transform method was employed to detect the strongest straight line across the entire image, identifying potential horizon candidates. Secondly, a regional approach method utilizing texture statistics differences between the sky and the sea was applied, partitioning the mask vertically to estimate the horizon. Lastly, a vertical edge response method was employed, analyzing the strongest vertical gradient response in each column and approximating the horizon using the RANSAC (RANdom SAmple Consensus) regression.

These three methods evaluate the horizon edge candidates using different image-processing principles, structurally organizing the extracted edge information into a linear horizon model. Variations in method performance were observed depending on environmental factors such as image brightness variation, edge clarity, and horizon curvature. By comparatively analyzing these methods, this study aimed to propose the most effective horizon detection method applicable to diverse maritime environments.

4.1 Hough transform method

The Hough transform method is a widely used image-processing method that detects specific shapes, notably straight lines, within images (Duda and Hart 1972). In this study, the Hough transform method was specifically utilized to identify and extract the straight-line structure representing the horizon effectively. The fundamental principle involves initially converting the edge pixel information from the image into a polar coordinate space. Within this coordinate space, lines are represented by two parameters: ρ, which indicates the perpendicular distance from the origin to the line, and θ, the angle between the x-axis and the line. In this polar coordinate space, each edge pixel contributes to multiple potential line candidates, forming an accumulator space. A voting mechanism accumulates these contributions, with each edge pixel adding votes to potential line candidates. The line with the highest accumulated votes is considered the most prominent and, therefore, the strongest candidate for the horizon. Subsequently, two points (x_1, y_1) and (x_2, y_2), representing the identified horizon at the left and right image boundaries, are calculated using the parameters (ρ, θ). These points are then connected to visually represent the horizon within the image clearly. This method is highly effective in simple environments where the horizon is distinctly visible and unobstructed. However, in more complex maritime environments with multiple linear structures or less defined edges, the method may lead to false detections or inaccuracies.

4.2 Regional approach method

The regional approach method detects the horizon by analyzing the texture and brightness characteristics of specific regions within an image (Jeong *et al.* 2018a). This method is particularly effective in complex environments where the horizon appears indistinct or curved. The regional approach involves several sequential steps. Initially, the image is vertically partitioned into multiple subregions. Texture analysis, which quantitatively evaluates visual attributes such as surface textures and patterns, is conducted for each subregion to identify boundaries separating the sky and the sea. Typically, the sea region exhibits various textures due to structural complexities like waves or ships, whereas the sky region generally displays uniform and smoother textures. Determining the boundary between the sky and the sea in each subregion involves calculating statistical texture attributes, including contrast and brightness. These attributes enable the establishment of thresholds that clearly distinguish the sky and sea regions. Initially, a Canny edge detection algorithm is applied to the mask image to identify edge structures effectively (Ding and Goshtasby 2001). Canny edge detection is an image processing technique that detects edges by identifying regions with significant changes in brightness. Subsequently, the detected edges are segmented vertically. For each segment, GLCM (Gray-Level Co-occurrence Matrix) is computed to analyze the spatial correlation of pixel intensity values, thus extracting detailed texture characteristics. From these computations, texture contrast values are derived, and texture weights are calculated based on mean brightness and structural complexity. These derived weights effectively distinguish between the sky and the sea regions, clearly defining their boundary positions within each subregion. Finally, the boundary positions from each subregion are interconnected to form a continuous and coherent horizon throughout the entire image. Additional continuity-focused post-processing steps refine the horizon further, addressing any indistinct or distorted boundaries. The regional approach method incrementally detects the horizon, considering the entire image structure, thus showcasing different strengths compared to global methods such as the Hough transform method. Nevertheless, the performance of the method may decline in maritime environments where texture differences are subtle or where illumination conditions vary significantly within the image.

4.3 Vertical edge response method

The vertical edge response method detects the horizon by analyzing the magnitude of the vertical gradient within the image. This method is particularly effective because the horizon typically represents a clear boundary between the sky and the sea, characterized by significant changes in pixel intensities. Consequently, locations exhibiting high vertical gradient magnitudes in the binary mask image can be candidates for the horizon. The procedure of this method includes several steps. Initially, preprocessing is applied to the input image using a Gaussian blur filter. The Gaussian blur smooths out intensity variations between pixels, effectively reducing noise and clarifying edge structures. This preprocessing step minimizes the likelihood of falsely detecting edges that are not associated with the horizon due to random noise. Following preprocessing, the vertical gradient magnitude is computed using the Sobel operator. The Sobel operator is a prominent edge-detection technique that measures changes in pixel intensity to detect edges. The vertical gradient identifies locations where pixel intensities sharply change vertically, which are highly indicative of horizon positions. Subsequently, the pixel position with the highest vertical gradient magnitude in each column is selected as a potential horizon candidate. Thus, horizon candidate points are determined for every column, providing critical information about the approximate location of the horizon across the image. Finally, these horizon candidate points are refined into an accurate horizon using the RANSAC regression. The RANSAC regression is a robust statistical method capable of identifying the most appropriate linear model, even when the dataset contains outliers. The method iteratively selects random subsets of candidate points, validates each subset to exclude outliers, and eventually determines the most accurate linear model. This robust approach ensures accurate horizon detection, even in environments with noise or obstructions. The vertical edge response method effectively utilizes local pixel-level information, demonstrating superior performance in realistic maritime environments where the horizon may not be perfectly straight and could include minor curvatures or indistinct boundaries.

5. Applications

In this study, the proposed horizon detection method was applied to a variety of maritime images obtained from real maritime environments to evaluate its performance. The evaluation procedure is systematically divided into the following subsections. Section 5.1 presents a qualitative comparative analysis between the pretrained SAM and SAM fine-tuning specifically for maritime environments. This analysis aims to highlight the impact of fine-tuning on the

Fig. 8 Results of the SAM and SAM fine-tuning for horizon segmentation in simple maritime environments

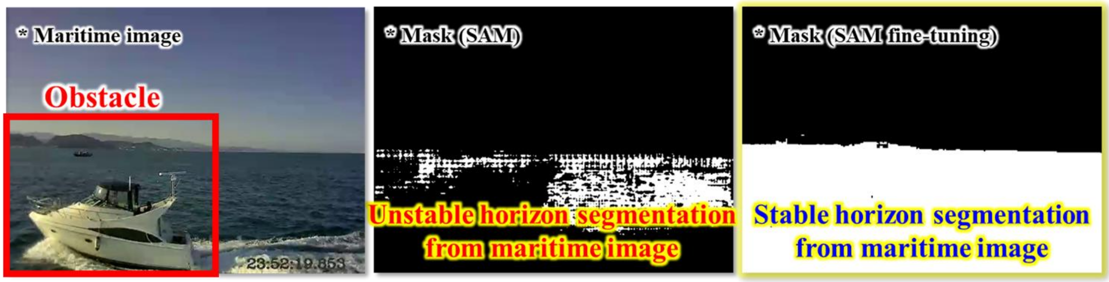

Fig. 9 Results of the SAM and SAM fine-tuning for horizon segmentation in complex maritime environments with obstacles

accuracy and robustness of horizon detection. In Section 5.2, both quantitative and qualitative evaluations of the three horizon edge classification methods (Hough transform, regional approach, and vertical edge response) are conducted. The distinct characteristics, strengths, and potential limitations of each method are thoroughly analyzed and discussed. Section 5.3 comprehensively assesses the combined effectiveness of SAM fine-tuning and horizon edge classification methods. Three distinct methods (Method 1, Method 2, and Method 3) are comparatively evaluated, utilizing both quantitative and qualitative criteria, to identify the most effective method for horizon detection. Through these systematic evaluations, this research verifies that the proposed horizon detection method is both practical and effective, offering reliable performance in diverse maritime environments.

5.1 Comparison of the SAM and SAM fine-tuning

This section presents a comparative analysis of the pretrained SAM and SAM fine-tuning specifically for maritime environments, highlighting their performance differences. In simple maritime environments, primarily composed of the sky and the sea, both models exhibited minimal differences in segmentation performance. As illustrated in Fig. 8, both the pretrained SAM and SAM fine-tuning produced similar horizon segmentation from maritime images, effectively delineating stable horizon boundaries.

Fig. 10 Results of three methods for horizon edge classification

However, as depicted in Fig. 9, significant differences emerged in complex maritime environments that contain obstacles, such as ships, structures, and floating debris. The pretrained SAM exhibited distorted and unstable horizon segmentation from maritime images due to interference from visual elements near the horizon, negatively affecting the accuracy of subsequent horizon classification processes. Conversely, SAM fine-tuning consistently demonstrated clear and stable horizon boundaries, even near obstacles, producing structurally coherent binary masks. This stability significantly enhanced the precision of the subsequent horizon edge classification stage, underscoring the effectiveness of fine-tuning in adapting SAM to the unique characteristics of maritime imagery.

5.2 Comparison of the horizon edge classification methods

This section compares the horizon detection performance of the three methods described in Section 4 using original maritime images unrelated to SAM fine-tuning. Each method extracts horizon candidates using distinct image-processing principles and approximates these candidates into a straight line. Two quantitative metrics were selected to evaluate the performance of each method: horizon height error (in pixels) and horizon slope error (in degrees). The performance evaluation results are summarized in Fig. 10 and Table 2.

The Hough transform method identifies the horizon as the strongest straight line across the entire image, showing relatively accurate detection in simple environments without obstacles. However, errors occurred due to under-detection caused by sunlight near the horizon or interference from other vertical structures. In this experiment, the Hough transform method recorded a horizon height error of 139 pixels and a slope error of 0.95°. The regional approach method vertically divides the image into several subregions. It estimates the boundary between the sky and the sea based on texture contrast and average brightness within each subregion. While this texture-based analysis can detect less distinct horizon boundaries, it experiences over-detection errors in cases with subtle texture differences or low illumination conditions, such as terrain near the horizon. In this experiment, the regional approach method exhibited a horizon height error of 93 pixels and a slope error of 2.85°, with slope accuracy being lower than that of the Hough transform method. The vertical edge response method selects the pixel with the highest vertical gradient response in each column as horizon candidate points, subsequently approximating the

Table 2 Comparison of the horizon edge classification methods

Method	Horizon height error (pixel)	Horizon slope error (°)
Hough transform	139	0.95
Regional approach	93 (33.09% reduction)	2.85 (200.00% increase)
Vertical edge response	**55** **(60.43% reduction)**	**0.50** **(47.36% reduction)**

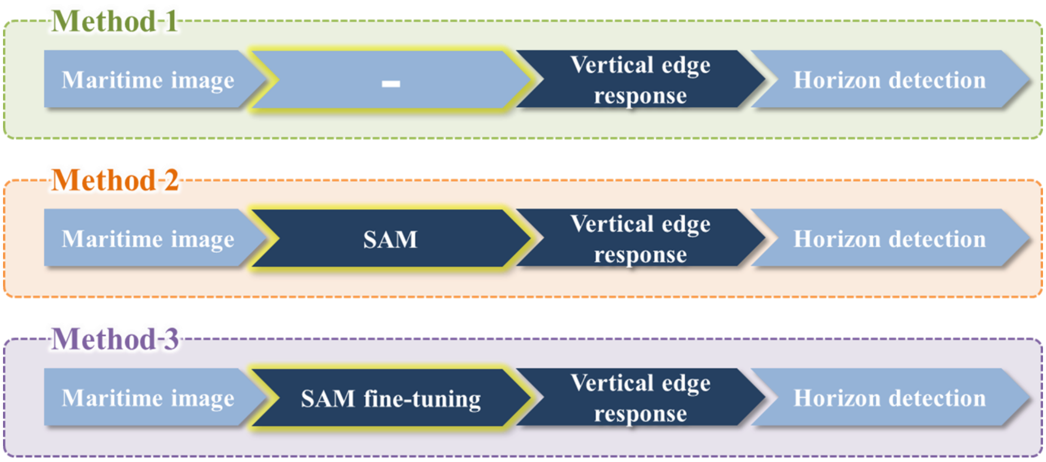

Fig. 11 Definition of three methods for horizon detection

horizon using the RANSAC regression. This method robustly models the horizon, effectively handling obstacles, distortions, and noise, thereby demonstrating the lowest average error. Specifically, the horizon height error was 55 pixels, representing a 60.43% reduction compared to the Hough transform method. Additionally, the slope error was the lowest at 0.50°. Based on these numerical results, the vertical edge response method was ultimately selected as the optimal method for horizon edge classification.

5.3 Horizon detection results

Horizon detection performance is influenced by a combination of data preprocessing methods, horizon edge detection methods, and horizon edge classification methods rather than relying on a single method alone. Therefore, this study developed and comparatively analyzed three distinct horizon detection methods, as defined in Fig. 11.

Method 1 directly applies the vertical edge response method to maritime images without generating any separate mask. Method 2 utilizes the vertical edge response method based on binary masks generated by the pretrained SAM. Method 3 incorporates SAM fine-tuning specifically adapted for maritime environments to create more precise binary masks before applying the vertical edge response method.

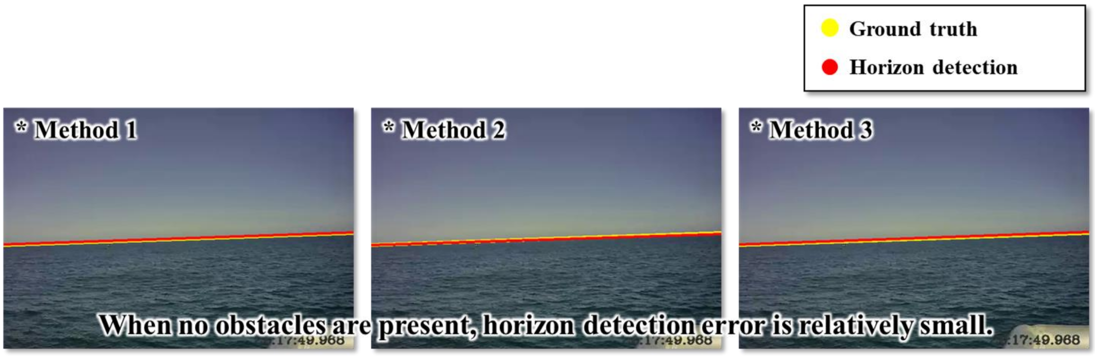

Fig. 12 Results of the three methods for horizon detection in simple maritime environments

Table 3 Comparison of the horizon detection methods in simple maritime environments

Method	Horizon height error (pixel)	Horizon slope error (°)
Method 1	2	0.03
Method 2	2	0.13
Method 3	3	0.07

In simple maritime environments without obstacles near the horizon, all three methods exhibited similar performance, as illustrated in Fig. 12. Methods 1 and 2 each showed an average horizon height error of approximately 2 pixels and minor slope errors of 0.03° and 0.13°, respectively. Method 3 yielded comparable outcomes with a height error of 3 pixels and a slope error of 0.07°. These results suggest that in clear environments devoid of obstacles, the vertical edge response method consistently and reliably detects the horizon irrespective of mask generation. Detailed performance results for these scenarios are summarized in Table 3.

However, notable differences among the methods were observed in complex maritime environments characterized by various visual disturbances, including obstacles, reflections of sunlight, and terrain near the horizon. Fig. 13 illustrates a challenging environment involving terrain above the horizon and intense sunlight reflections, which significantly impact detection accuracy. In this environment, Method 1 inaccurately identified the horizon substantially above its actual position, resulting in a height error of 55 pixels and a slope error of 0.50°. Method 2, which leveraged masks generated from the pretrained SAM, reduced the height error to 19 pixels but exhibited increased instability in slope estimation, with a slope error of 1.53°. Conversely, Method 3, which utilizes SAM fine-tuning, effectively mitigates interference from obstacles and reflections, achieving a significantly lower height error of 12 pixels and a minimal slope error of 0.21°. Compared to Method 1, Method 3 achieved a height error reduction of 78.18% and a slope error improvement of 58.00%. These comparative results are presented in Table 4.

Additionally, Fig. 14 illustrates an even more challenging maritime environment characterized by distinct obstacles, such as ships and structures near the horizon. Method 1 was overly sensitive to obstacle edges and notably over-detected the horizon position, resulting in a height error of 37 pixels and a slope error of 3.34°. Method 2 improved slightly due to the enhanced mask quality

Fig. 13 Results of three methods for horizon detection in complex maritime environments with terrain and reflection

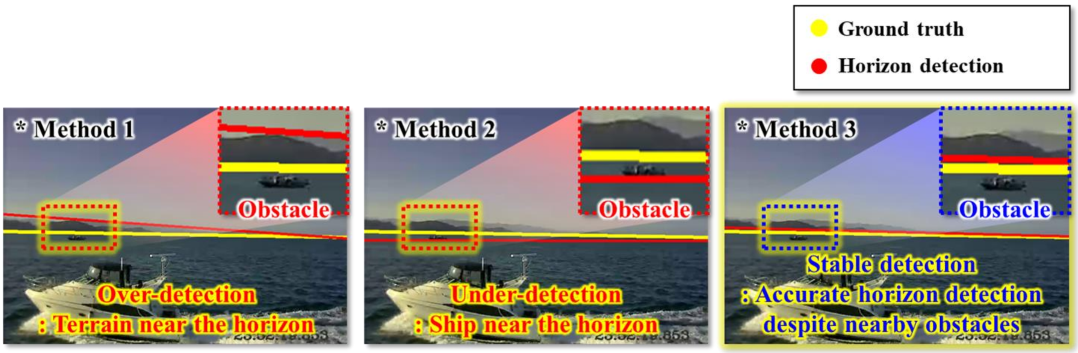

Fig. 14 Results of three methods for horizon detection in complex maritime environments with terrain and ships

Table 4 Comparison of the horizon detection methods in complex maritime environments with terrain and reflection

Method	Horizon height error (pixel)	Horizon slope error (°)
Method 1	55	0.50
Method 2	19 (65.45% reduction)	1.53 (206.00% increase)
Method 3	**12** **(78.18% reduction)**	**0.21** **(58.00% reduction)**

provided by the pretrained SAM, achieving errors of 25 pixels in height and 1.01° in slope. In contrast, Method 3, utilizing SAM fine-tuning, exhibited superior robustness to visual disturbances, significantly reducing errors to 10 pixels in height and 0.59° in slope. Consequently, Method 3 demonstrated relative improvements of 72.97% in height accuracy and 82.33% in slope accuracy compared to Method 1. The detailed performance metrics are summarized in Table 5.

Table 5 Comparison of the horizon detection methods in complex maritime environments with terrain and ships

Method	Horizon Height Error (pixel)	Horizon Slope Error (°)
Method 1	37	3.34
Method 2	25 (32.43% reduction)	1.01 (69.76% reduction)
Method 3	**10** **(72.97% reduction)**	**0.59** **(82.33% reduction)**

Furthermore, the average inference speed of each method was evaluated to assess its practical applicability in real-time maritime navigation systems. Method 1, utilizing only the vertical edge response method, achieved the highest average inference speed of 16.22 fps (frames per second) due to its relatively simple computational requirements. Method 2 demonstrated the lowest average inference speed of 2.15 fps, primarily due to the additional computational load associated with generating masks using the pretrained SAM without domain-specific fine-tuning. Method 3, which incorporates the SAM fine-tuning specifically tailored for maritime environments, demonstrated an intermediate average inference speed of 4.54 fps, providing a balanced trade-off between enhanced accuracy and computational efficiency. These results clearly demonstrate the importance of selecting appropriate horizon detection methods that balance both accuracy requirements and computational constraints in practical, real-time maritime applications.

6. Conclusions

This study proposes a novel deep learning-based horizon detection method designed to enhance maritime navigation safety and improve the accuracy of obstacle tracking. Conventional image-processing methods have shown limited performance due to the complexity and variability inherent in maritime environments. To address these limitations, the general-purpose image segmentation model, SAM, was fine-tuned specifically for maritime applications. Particularly, the efficient adapter-based fine-tuning technique was applied to the SAM mask decoder, significantly enhancing the model's ability to learn the visual characteristics of the horizon and consequently improving horizon detection performance.

The key conclusions of this study are summarized as follows:

- The SAM fine-tuning successfully addressed the limitations encountered by conventional methods, effectively capturing distinct maritime visual features and improving accuracy in detecting the horizon, particularly in complex maritime environments that involve obstacles, various terrains, and intense reflections.
- The vertical edge response method demonstrated superior performance compared to the horizon edge classification methods analyzed. It significantly reduced horizon height error by 60.43%, decreasing from 139 pixels to 55 pixels, and slope error by 47.36%, decreasing from 0.95° to 0.50°, when compared with the Hough transform method. Additionally, it improved accuracy relative to the regional approach method, reducing height error from 93 pixels to 55 pixels and slope error from 2.85° to 0.50°.
- Method 3, integrating SAM fine-tuning with the vertical edge response method, delivered

the most accurate horizon detection in complex maritime environments. It substantially decreased height error by 78.18%, from 55 pixels to 12 pixels, and reduced slope error by 58.00%, from 0.50° to 0.21°, compared to Method 1. Furthermore, relative to Method 2, Method 3 improved height accuracy from 19 pixels to 12 pixels and reduced slope error significantly from 1.53° to 0.21°.

- The generation of high-quality masks through the SAM fine-tuning was crucial for accurate horizon detection. This approach significantly minimized interference from maritime obstacles, reduced false detections, and ensured a clear, continuous, and structurally consistent horizon, thus enhancing the robustness and reliability of detection results under challenging maritime environments.

Future research will involve the real-time application and evaluation of this method in actual maritime environments, utilizing expanded datasets that encompass diverse environments, including weather variations, different resolutions, and ship movements. In these diverse environments, horizon detection performance may be impacted by factors such as reduced visibility, reflections, and camera instability, underscoring the need for advanced post-processing image methods to ensure reliability. Additionally, exploring the application of this method to related domains, such as rivers, lakes, and aerial maritime environments, is expected to improve safety, navigation accuracy, and environmental monitoring capabilities. Lastly, the integration of supplementary sensor data from AIS (Automatic Identification System) and RADAR will be pursued to develop a more robust multi-sensor-based obstacle detection and tracking system, significantly enhancing maritime navigational safety.

Acknowledgments

This work was partially supported by (a) Research Institute of Marine Systems Engineering of Seoul National University, Republic of Korea, (b) Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (RS-2025-02263945, HRD Program for Industrial Innovation), and (c) the project titled ‘Fostering Talent in Advanced Ship Blue Tech (RS-2025-02221147)’, funded by the Ministry of Oceans and Fisheries, Republic of Korea.

References

Aggarwal, N. and Karl, W.C. (2006), “Line detection in images through regularized hough transform”, *IEEE T. Image Process*., **15**(3). https://doi.org/10.1109/TIP.2005.863021

Ahmad, T., Bebis, G., Regentova, E.E. and Nefian, A. (2013), “A machine learning approach to horizon line detection using local features”, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8033 LNCS (PART 1). https://doi.org/10.1007/978-3-642-41914-0_19.

Ding, L. and Goshtasby, A. (2001), “On the canny edge detector”, *Pattern Recogn*., **34**(3). https://doi.org/10.1016/S0031-3203(00)00023-6.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houlsby, N. (2021), “An image is worth 16x16 words: Transformers for image recognition at scale”, ICLR 2021 – *Proceedings of the 9th International Conference on Learning Representations*.

Duda, R.O. and Hart, P.E. (1972), “Use of the hough transformation to detect lines and curves in pictures”, *Commun. ACM*, **15**(1). https://doi.org/10.1145/361237.361242.

Fefilatyev, S., Smarodzinava, V., Hall, L.O. and Goldgof, D.B. (2006), "Horizon detection using machine learning techniques", *Proceedings of the 5th International Conference on Machine Learning and Applications*, ICMLA 2006. https://doi.org/10.1109/ICMLA.2006.25.

Gershikov, E. (2013), "Horizon line detection in marine images: Which method to choose?", *Int. J. Adv. Intell. Syst.*, **6**(1-2).

Gu, H., Dong, H., Yang, J. and Mazurowski, M.A. (2025), "How to build the best medical image segmentation algorithm using foundation models: A comprehensive empirical study with Segment anything model", *Mach. Learn. Biomed. Imaging*, **3**, 88-120. https://doi.org/10.59275/j.melba.2025-86a6.

Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L. and Chen, W. (2022), "Lora: Low-rank adaptation of large language models, ICLR 2022 - *Proceedings of the 10th International Conference on Learning Representations*.

Jebari, O., Hong, S.H., Hunsucker, T., Kim, D.K. and Jin, C. (2025), "Wave-field reconstruction using stereo cameras on a floating platform in a synthetic environment", *Ocean Eng.*, **327**, 120958. https://doi.org/10.1016/J.OCEANENG.2025.120958.

Jeong, C.Y., Yang, H.S. and Moon, K.D. (2018a), "Fast horizon detection in maritime images using region-of-interest", *Int. J. Distrib. Sens. N.*, **14**(7). https://doi.org/10.1177/1550147718790753.

Jeong, C.Y., Yang, H.S. and Moon, K.D. (2018b), "Horizon detection in maritime images using scene parsing network", *Electron. Lett.*, **54**(12). https://doi.org/10.1049/el.2018.0989.

Kim, H.Y., Roh, M.I., Lee, H.W., Yeo, I.C., Jo, Y.M., Ha, J.S. and Son, N.S. (2024), "A hybrid tracking method for maritime obstacles using sensor data", *Ocean Eng.*, **312**, 119242. https://doi.org/10.1016/J.OCEANENG.2024.119242.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P. and Girshick, R. (2023), "Segment anything", *Proceedings of the IEEE International Conference on Computer Vision*, 3992-4003. https://doi.org/10.1109/ICCV51070.2023.00371.

Kuanar, S., Mahapatra, D., Bilas, M. and Rao, K.R. (2022), "Multi-path dilated convolution network for haze and glow removal in nighttime images", *Visual Computer*, **38**(3). https://doi.org/10.1007/s00371-021-02071-z.

Lee, H.W., Roh, M.I., Cho, Y.M. and Park, J.H. (2023), "Data association for autonomous ships based on virtual simulation environment", *Ocean Eng.*, **281**. https://doi.org/10.1016/j.oceaneng.2023.114646.

Lee, W.J., Roh, M.I., Lee, H.W., Ha, J.S., Cho, Y.M., Lee, S.J. and Son, N.S. (2021), "Detection and tracking for the awareness of surroundings of a ship based on deep learning", *J. Comput. Des. Eng.*, **8**(5). https://doi.org/10.1093/jcde/qwab053

Park, J.H., Roh, M.I., Lee, H.W., Jo, Y.M., Ha, J.S. and Son, N.S. (2024), "Multi-vessel target tracking with camera fusion for unmanned surface vehicles", *Int. J. Naval Architect. Ocean Eng.*, **16**, 100608. https://doi.org/10.1016/J.IJNAOE.2024.100608.

Prasad, D.K., Rajan, D., Rachmawati, L., Rajabally, E. and Quek, C. (2017), "Video processing from electro-optical sensors for object detection and tracking in a maritime environment: A survey", *IEEE T. Intell. Transp.*, **18**(8). https://doi.org/10.1109/TITS.2016.2634580.

Shahraki, M., Elamin, A. and El-Rabbany, A. (2025), "SAMNet++: A segment anything model for supervised 3D point cloud semantic segmentation", *Remote Sens.*, **17**(7). https://doi.org/10.3390/rs17071256.

Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I. (2017). Attention Is All You Need. In Advances in Neural Information Processing Systems (Vol. 30).

Zardoua, Y., Boulaala, M., El Mrabet, M. and Astito, A. (2024), "A fast horizon detector and a new annotated dataset for maritime video processing", *J. Image Graphics (United Kingdom)*, **12**(4), 345-361. https://doi.org/10.18178/joig.12.4.345-361.

Zhang, C., Liu, L., Cui, Y., Huang, G., Lin, W., Yang, Y. and Hu, Y. (2023), "A comprehensive survey on segment anything model for vision and beyond. http://arxiv.org/abs/2305.08196.

Zhang, R., Li, S., Ji, G., Zhao, X., Li, J. and Pan, M. (2021), "Survey on deep learning-based marine object

detection", *J. Adv. Transport*., https://doi.org/10.1155/2021/5808206.
Zhang, Y., Shen, Z. and Jiao, R. (2024), Segment anything model for medical image segmentation: Current applications and future directions. http://arxiv.org/abs/2401.03495.

MK