applied sciences

MDPI

Communication

Proposal of Innovative Methods for Computer Vision Techniques in Maritime Sector

Bo Jiang [1], Xuan Wu [2], Xuecheng Tian [2,*], Yong Jin [3] and Shuaian Wang [2]

1. Division of Logistics and Transportation, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China; jb22@mails.tsinghua.edu.cn
2. Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Kowloon, Hong Kong; isabella125.wu@connect.polyu.hk (X.W.); hans.wang@polyu.edu.hk (S.W.)
3. Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Hong Kong; jimmy.jin@polyu.edu.hk

* Correspondence: xuecheng-simon.tian@connect.polyu.hk

Abstract: Computer vision (CV) techniques have been widely studied and applied in the shipping industry and maritime research. The existing literature has primarily focused on enhancing image recognition accuracy and precision for water surface targets by refining CV models themselves. This paper introduces innovative methods to further improve the accuracy of detection and recognition using CV models, including using ensemble learning and integrating shipping domain knowledge. Additionally, we present a novel application of CV techniques in the maritime domain, expanding the research perspective beyond the traditional focus on the accurate detection and recognition of water surface targets. Specifically, a novel solution integrating a CV model and the transfer learning method is proposed in this paper to address the challenge of relatively low-speed and high-charge internet services on ocean-going vessels, aiming to improve the online video viewing experience while conserving network resources. This paper is of importance for advancing further research and application of CV techniques in the shipping industry.

Keywords: shipping industry; maritime research; computer vision; ensemble learning; transfer learning

Citation: Jiang, B.; Wu, X.; Tian, X.; Jin, Y.; Wang, S. Proposal of Innovative Methods for Computer Vision Techniques in Maritime Sector. *Appl. Sci.* **2024**, *14*, 7126. https://doi.org/10.3390/app14167126

Academic Editors: José A. Orosa and Atsushi Mase

Received: 3 July 2024
Revised: 4 August 2024
Accepted: 13 August 2024
Published: 14 August 2024

1. Introduction

With the popularity and development of artificial intelligence (AI), big data, and the Internet of Things (IoT), computer vision (CV) techniques have been extensively utilized in the shipping industry and maritime research, including maritime surveillance [1], maritime safety [2], maritime transportation [3], marine environmental protection [4], and international maritime policies and laws [5]. Table 1 shows the SWOT analysis for the applications of CV techniques in the maritime domain. CV techniques present distinct advantages in the shipping industry, facilitating the real-time monitoring and automatic detection of vessels, cargo, and marine environments. This capability not only reduces operational costs but also enhances the efficiency and safety of maritime operations. With an increasing global demand for automation and intelligence in the shipping industry, the application and proliferation of CV techniques are facing unprecedented opportunities. Governments in many countries and regions are also actively supporting the development of smart shipping and maritime safety technologies. However, the accuracy and stability of CV models are critically dependent on the availability of extensive and high-quality training data in the maritime domain and may be influenced by complex maritime environments. On the other hand, the application of CV techniques may encounter legal requirements concerning data privacy and security, which could pose challenges. Furthermore, their widespread adoption may also raise social and ethical issues such as employment impacts on the traditional shipping industry. Consequently, the widespread implementation of CV techniques still encounters numerous challenges and issues.

Table 1. SWOT analysis for the applications of CV techniques in the maritime domain.

Strengths	Weaknesses
• Real-time monitoring and automatic detection of vessels, cargo, and marine environments; • Reduce operational costs; • Enhance the efficiency and safety of maritime operations.	• Dependent on the availability of extensive and high-quality training data in the maritime domain; • Influenced by complex maritime environments.
Opportunities	**Threats**
• An increasing global demand for automation and intelligence in the shipping industry; • Government support.	• Legal requirements concerning data privacy and security; • Social and ethical issues.

One of the most important issues of CV techniques in the shipping industry is the timely, automatic, and accurate detection and recognition of water surface targets. Many traditional CV models and algorithms have been employed to detect and recognize ships and objects of interest, including mean shift [6], deformable part-based models (DPMs) [7], support vector machines (SVMs) [8], and sparse representation [9]. Prasad et al. [10] provide a comprehensive overview of various CV models for target detection and tracking in the maritime environment, and the performance of several maritime and computer vision techniques is evaluated on the Singapore Maritime Dataset (SMD), a dataset including on-shore videos, on-board videos, and near-infrared (NIR) videos, which are acquired at various locations. Kontopoulos et al. [11] introduce an innovative approach that merges CV and trajectory classification to provide a highly accurate classification of vessel mobility patterns. Varga et al. [12] present a large-scale water surface object detection and tracking benchmark, which aims to bridge the gap between land-based vision systems and sea-based ones. Prasad et al. [13] discuss and define the assessment metrics suitable for maritime CV, indicating that existing CV approaches are indeed promising for maritime research.

In the past several years, maritime research has witnessed and benefited from the applications of deep learning methods with superior performance in target detection and recognition. Qiao et al. [14] make a survey on four aspects of current research progress: full scene parsing of an image, target vessel re-identification, target vessel tracking, and multimodal data fusion with data from visual sensors. Deep learning methods, such as regions with convolutional neural network features (R-CNN) [15], Fast R-CNN [16], and Faster R-CNN [17], are proposed and adopted, which can obtain accurate detection results at the expense of a high computational cost. For a balance between detection accuracy and computational efficiency, the method called "you only look once" (YOLO) [18], which is based on image global information, is widely employed in practical applications. Ergasheva et al. [2] develop an accurate ship fire detection model based on YOLO for discerning the presence of fires on vessels.

Due to the complicated water surface environments, such as light reflection and poor weather conditions (e.g., fog and rain), the quality of captured images is often inevitably degraded, leading to unsatisfactory and unreliable target detection results. To make the vision techniques more feasible and practicable, many intelligent visibility enhancement methods have been developed, such as low-visibility enhancement networks (LVENets) [19], learned parameter sharing-based versatile visibility enhancement networks (LPSNets) [20], efficient and effective multi-deep neural networks (EEMNNs) [21], and multiple-feature fusion-guided low-visibility enhancement networks (MFF-Nets) [22]. Guo et al. [23] propose a neural network-empowered water surface target detection framework with a data augmentation strategy, which illustrates its effectiveness in terms of detection accuracy. However, it should be noted that although these visibility enhancement methods can theoretically improve image visual quality, it is impossible to generate restored images without the loss of any detail [24]. *Therefore, the accuracy of detection and recognition still needs to be improved.*

Based on existing studies, this paper proposes some innovative methods to further improve the accuracy of the detection and recognition of CV models in maritime research, such as the application of ensemble learning and the integration of maritime domain knowledge. Furthermore, considering that current CV studies and practical applications in the shipping industry primarily focus on the timely, automatic, and accurate detection and recognition of water surface targets, we innovatively expand the research perspective and put forward a novel application of CV techniques in the maritime domain.

The remainder of this study is organized as follows. Section 2 introduces our two proposed innovative methods for improving the recognition accuracy of CV models. Section 3 presents a novel application of CV techniques in the maritime domain. Section 4 concludes this study.

2. Innovative Methods to Further Improve the Recognition Accuracy of CV Models

In this section, we propose two innovative methods to further improve the detection and recognition accuracy of CV models in the shipping industry and maritime research.

2.1. Enhancing Recognition Accuracy through Ensemble Learning

Illegal shipping activities not only increase the occurrence of maritime traffic accidents and impact regular maritime transportation but also potentially harm the marine ecological environment, through activities such as illegal river vessels entering the sea, illegal fishing, and the illegal excavation/dumping of sand. To enhance the surveillance of maritime vessels, cameras are installed for surveillance in key areas such as river estuaries and sea entrances, where some illegal shipping activities may happen. Then, CV models are utilized to recognize the specific type of ships in the captured images of these cameras. For example, when a river vessel is identified as preparing to sail into the open sea, the recognition system will trigger an alarm and alert surveillance officers to take enforcement actions. However, due to the complex conditions and environments, such as heavy fog and water surface light reflection, the accuracy of ship recognition by the CV model is relatively low; for instance, mistakenly identifying a sea vessel as a river vessel is possible, resulting in inefficient management.

Existing research has primarily focused on optimizing the accuracy and precision of image recognition by improving the CV model itself [23]. However, the improvement in accuracy is limited due to the complicated environment. Ensemble learning [25] is a machine learning method that completes learning tasks by constructing and combining multiple learners. The core idea is to integrate multiple learners (weak learners) to achieve better performance than a single learner. By combining the prediction results of multiple learners, ensemble learning can improve overall generalization ability and stability, thereby achieving better performance in complex tasks. In other words, even if one weak learner obtains an incorrect prediction, other weak learners may correct the error. Based on the concept of ensemble learning, we propose an innovative approach to addressing the issue by installing multiple cameras at different orientations in key areas to capture images of ships from various angles. Specifically, the captured images from each orientation are independently processed using a CV model for recognition, and the majority (more than half) of recognition results are taken as the final recognition output. Therefore, even if the recognition result of the image captured from a certain angle is incorrect, as long as the majority of the recognition results are accurate, the output result is correct. The proposed method, thereby, can enhance the accuracy of detection and recognition, with more details illustrated in Example 1.

Example 1. *At an estuary, nine cameras are installed at different orientations to identify illegal entries of river ships into the sea. Assume each camera has a recognition accuracy rate of 70%, and the recognition results from photos taken at different angles are independent of each other. Suppose at a given moment, the system captures nine photos from different angles, and as long as five or more (more than half) photos are correctly recognized, the system's output is correct, with the accuracy rate R calculated as follows:*

$$R = \sum_{k=5}^{9} C_9^k 0.7^k (1-0.7)^{9-k} \approx 90.12\%,$$

where C_9^k represents the number of combinations of k elements taken from nine different elements. The accuracy rate of the system with nine cameras increases by 28.74% compared to the recognition accuracy rate of 70% of a single camera.

2.2. Integrating Shipping Domain Knowledge: A Tugboat Example

In the realm of detecting and recognizing water surface targets, conventional methods typically focus on static training data and the CV model itself. Nevertheless, the presence of model errors is an inescapable reality, leading to the possibility of recognition inaccuracies. Consequently, rather than exclusively enhancing the CV model itself, the integration of shipping domain knowledge should be considered as a supplementary component in the process of water surface target detection and recognition. Yang et al. [26] propose a novel perspective of integrating shipping domain knowledge, such as the length and width of ships and the speed of navigation, into CV models for auxiliary ship recognition in maritime surveillance, and they give two specific examples in practice. Based on that, we propose another example of using shipping domain knowledge for recognizing tugboats to further improve the accuracy of ship recognition, as illustrated in Example 2.

Example 2. *A tugboat refers to a vessel used for towing barges and other ships, characterized by a small hull, good stability, high power, and strong traction. Tugboats are primarily utilized in maritime transport and rescue operations, specifically for towing barges carrying cargoes and various workboats, as well as for towing giant vessels and other large water structures such as offshore platforms and floating docks. It is important to note that tugboats themselves generally lack cargo handling capabilities. In practice, when a tugboat operates without a load, it typically travels at a faster speed, whereas when towing other vessels through ropes, it usually travels at a slower speed. This practical phenomenon is not only due to the increased load but also to avoid potential collisions between vessels resulting from the sudden decelerations. Based on the above discussion, suppose we have two CV models, one for recognizing ship types and the other for detecting ship speeds. Then, if the two CV models identify a slow-moving tugboat without towing other vessels behind, or a fast-moving tugboat towing other large vessels, at least one of the models' output results is likely to be incorrect. That is because such scenarios are rare in practice, and, thus, it is necessary to re-identify or intervene manually.*

3. A Novel Application of CV in the Maritime Domain

In this section, we innovatively put forward a novel application of CV techniques in the maritime domain; this is to enhance the online video viewing experience on ocean-going vessels while saving network resources.

Network communication services on ocean-going vessels are typically provided by communication satellites, as there are no communication fiber optic cables in the open sea. Consequently, the network speed on ocean-going vessels is relatively slow, and the charges are comparatively high. However, ocean-going vessels often sail on the open sea for weeks or even months, and the crew members on board have a practical demand for Internet access to watch online videos such as news programs and movies, which require fast and cost-effective network communication services. One of the existing solutions is to reduce the network transmission load by compressing and decompressing video files, aiming to save network resources; however, this method is not timely and convenient enough because the compression and decompression processes are time-consuming and may potentially decrease the resolution of online videos. Therefore, the online video viewing experience is not good enough.

We propose a novel solution that uses a CV model and the concept of transfer learning to achieve a better online video viewing experience while conserving network resources. Transfer learning [27] is a machine learning method that applies knowledge and models learned from old domains to new related ones based on the similarity between problems. The core idea is that the knowledge contained in a model trained on one task can be partially or fully transferred to another similar task.

Specifically, we pre-train a CV model based on a large amount of video sample data to convert low-resolution videos into high-resolution ones. When crew members on ocean-going vessels watch videos online, a software tool (embedded with the pre-trained CV model), first loads the high-resolution frames of the beginning portion of the video and the low-resolution frames of the remaining portion. For example, it loads the first 10% of the video in high resolution and the remaining 90% in low resolution. The first 10% high-resolution part is used to fine-tune the built-in pre-trained CV model through transfer learning, as the specific features of different types of videos, such as news programs and movies, may vary. The optimized CV model is better able to adapt to the characteristics of the particular video after transfer learning, and then it converts the remaining low-resolution part into a high-resolution one; the converted remainder is then combined with the first 10% part to result in the final video for viewing by the crew members. The entire process, including video loading and CV model fine-tuning, is accomplished automatically in the software. While the resulting video may differ from the original high-resolution video, this method can enhance the online video viewing experience to some extent while conserving network resources. For example, a high-resolution movie requires 2 GB of data usage, but using this method may only require 0.5 GB of data while providing a similar viewing experience. The flowchart of our proposed novel solution utilizing a CV model is shown in Figure 1.

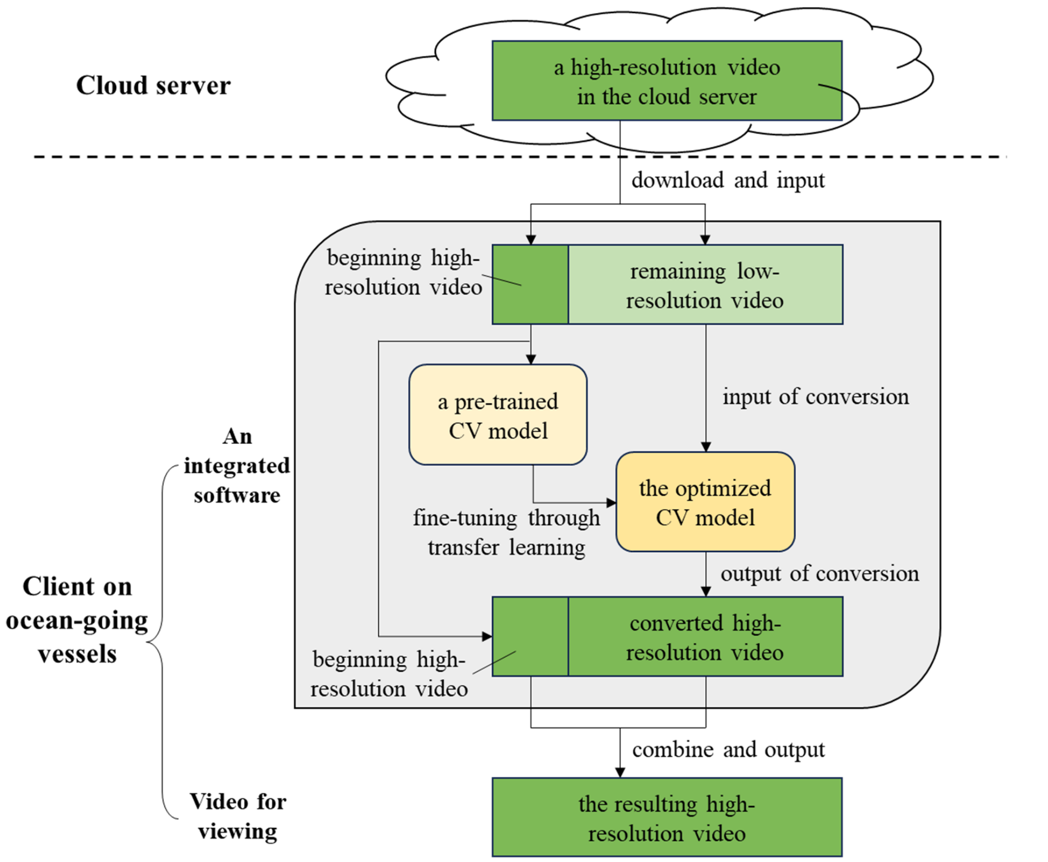

Figure 1. The flowchart of our proposed novel solution utilizing a CV model.

4. Conclusions

CV techniques have seen widespread application in the shipping industry and maritime research. The challenging water surface environment, characterized by factors such as light reflection and adverse weather conditions (e.g., fog and rain), often results in degraded image quality, thereby compromising the reliability of target recognition outcomes. The existing literature has predominantly concentrated on enhancing the accuracy and precision of image recognition through improvements in the underlying CV models.

This study introduces novel approaches aimed at further enhancing the detection and recognition accuracy of CV models in maritime research. Specifically, we propose the utilization of ensemble learning methods and the incorporation of shipping domain knowledge to augment the performance of CV models. In addition, while the current landscape of CV research and applications in the shipping industry primarily emphasizes the prompt, automated, and precise identification of water surface targets, we present a pioneering shift in research perspective by introducing a novel application of CV techniques within the maritime domain. Specifically, we leverage a CV model and incorporate the concept of transfer learning to address the challenge posed by the provision of relatively low-speed and high-charge network services on ocean-going vessels. The proposed novel approach is designed to enhance the online video viewing experience while conserving network resources, thereby offering a promising response to the demands of the shipping industry.

Future research can be expanded from the following two aspects. First, it is imperative to propose more innovative methods to further improve the accuracy of CV models for detecting and recognizing water surface targets of interest, such as integrating more systematic and effective shipping domain knowledge into CV models. Meanwhile, quantitative methods and extensive numerical experiments could be conducted to more accurately demonstrate the effectiveness of CV models. With higher accuracy and stability, future CV techniques would greatly enhance the efficiency of automatic monitoring and detection in the shipping industry, thereby promoting the development of maritime transportation and marine safety technologies. Second, feasible CV models and software tools are worth developing for our proposed solutions to address the challenge of online video viewing on ocean-going vessels in Section 3, with extensive numerical experiments analyzing and evaluating model performance. It is of great significance to enhance the online video viewing experience for crew members on ocean-going vessels while saving network resources.

Author Contributions: Conceptualization, B.J. and S.W.; methodology, B.J., X.W., X.T., Y.J. and S.W.; investigation, B.J.; resources, Y.J. and S.W.; writing—original draft preparation, B.J. and X.T.; writing—review and editing, X.W., X.T., Y.J. and S.W.; visualization, B.J.; supervision, S.W.; project administration, Y.J.; funding acquisition, Y.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Prayudi, A.; Sulistijono, I.A.; Risnumawan, A.; Darojah, Z. Surveillance system for illegal fishing prevention on uav imagery using computer vision. In Proceedings of the 2020 IEEE International Electronics Symposium (IES), Surabaya, Indonesia, 29–30 September 2020; pp. 385–391.
2. Ergasheva, A.; Akhmedov, F.; Abdusalomov, A.; Kim, W. Advancing maritime safety: Early detection of ship fires through computer vision, deep learning approaches, and histogram equalization techniques. *Fire* **2024**, *7*, 84. [CrossRef]
3. Yu, M.; Han, S.; Wang, T.; Wang, H. An approach to accurate ship image recognition in a complex maritime transportation environment. *J. Mar. Sci. Eng.* **2022**, *10*, 1903. [CrossRef]
4. Goudemant, T.; Francesconi, B.; Aubrun, M.; Kervennic, E.; Grenet, I.; Bobichon, Y.; Bellizzi, M. Onboard anomaly detection for marine environmental protection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 7918–7931. [CrossRef]
5. Emerick De Magalhães, M.; Barbosa, C.E.; Cordeiro, K.D.F.; Isidorio, D.K.M.; Souza, J.M.D. Improving maritime domain awareness in Brazil through computer vision technology. *J. Mar. Sci. Eng.* **2023**, *11*, 1272. [CrossRef]

6. Zhou, H.; Yuan, Y.; Shi, C. Object tracking using sift features and mean shift. *Comput. Vis. Image Underst.* **2009**, *113*, 345–352. [CrossRef]
7. Felzenszwalb, P.F.; Girshick, R.B.; McAllester, D.; Ramanan, D. Object detection with discriminatively trained part-based models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *32*, 1627–1645. [CrossRef] [PubMed]
8. Juang, C.F.; Chen, G.C. A TS fuzzy system learned through a support vector machine in principal component space for real-time object detection. *IEEE Trans. Ind. Electron.* **2012**, *59*, 3309–3320. [CrossRef]
9. Chen, X.; Wang, S.; Shi, C.; Wu, H.; Zhao, J.; Fu, J. Robust ship tracking via multi-view learning and sparse representation. *J. Navig.* **2019**, *72*, 176–192. [CrossRef]
10. Prasad, D.K.; Rajan, D.; Rachmawati, L.; Rajabally, E.; Quek, C. Video processing from electro-optical sensors for object detection and tracking in a maritime environment: A survey. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 1993–2016. [CrossRef]
11. Kontopoulos, I.; Makris, A.; Zissis, D.; Tserpes, K. A computer vision approach for trajectory classification. In Proceedings of the 2021 22nd IEEE International Conference on Mobile Data Management (MDM), Toronto, ON, Canada, 15–18 June 2021; pp. 163–168.
12. Varga, L.A.; Kiefer, B.; Messmer, M.; Zell, A. Seadronessee: A maritime benchmark for detecting humans in open water. In Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2022; pp. 3686–3696.
13. Prasad, D.K.; Dong, H.; Rajan, D.; Quek, C. Are object detection assessment criteria ready for maritime computer vision? *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 5295–5304. [CrossRef]
14. Qiao, D.; Liu, G.; Lv, T.; Li, W.; Zhang, J. Marine vision-based situational awareness using discriminative deep learning: A survey. *J. Mar. Sci. Eng.* **2021**, *9*, 397. [CrossRef]
15. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
16. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
17. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef] [PubMed]
18. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
19. Guo, Y.; Lu, Y.; Liu, R.W. Lightweight deep network-enabled real-time low-visibility enhancement for promoting vessel detection in maritime video surveillance. *J. Navig.* **2022**, *75*, 230–250. [CrossRef]
20. Qu, J.; Gao, Y.; Lu, Y.; Xu, W.; Liu, R.W. Deep learning-driven surveillance quality enhancement for maritime management promotion under low-visibility weathers. *Ocean Coast. Manag.* **2023**, *235*, 106478. [CrossRef]
21. Lu, Y.; Guo, Y.; Zhu, F.; Liu, R.W. Towards low-visibility enhancement in maritime video surveillance: An efficient and effective multi-deep neural network. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; pp. 2869–2874.
22. Zhou, W.; Li, B.; Luo, G. Multi-feature fusion-guided low-visibility image enhancement for maritime surveillance. *J. Mar. Sci. Eng.* **2023**, *11*, 1625. [CrossRef]
23. Guo, Y.; Lu, Y.; Guo, Y.; Liu, R.W.; Chui, K.T. Intelligent vision-enabled detection of water-surface targets for video surveillance in maritime transportation. *J. Adv. Transp.* **2021**, *2021*, 9470895. [CrossRef]
24. Lu, W.; Duan, J.; Qiu, Z.; Pan, Z.; Liu, R.W.; Bai, L. Implementation of high-order variational models made easy for image processing. *Math. Methods Appl. Sci.* **2016**, *39*, 4208–4233. [CrossRef]
25. Dong, X.; Yu, Z.; Cao, W.; Shi, Y.; Ma, Q. A survey on ensemble learning. *Front. Comput. Sci.* **2020**, *14*, 241–258. [CrossRef]
26. Yang, Y.; Yan, R.; Wang, S. Integrating shipping domain knowledge into computer vision models for maritime transportation. *J. Mar. Sci. Eng.* **2022**, *10*, 1885. [CrossRef]
27. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76. [CrossRef]