Mediterranean Smart Cities Conference (MSCC 2024)
2-4 May, 2024, Martil, Morocco

Improving Coastal and Port Management in Smart Cities with UAVs and Deep Learning

Murat Bakirci
Unmanned/Intelligent Systems Lab
Faculty of Aeronautics and Astronautics
Tarsus University
Mersin, Türkiye
muratbakirci@tarsus.edu.tr

Irem Bayraktar
Unmanned/Intelligent Systems Lab
Faculty of Aeronautics and Astronautics
Tarsus University
Mersin, Türkiye
irem_bayraktar@tarsus.edu.tr

Abstract—**Efficient coast and harbor management is integral to the vitality, sustainability, and resilience of smart cities. With bustling harbors serving as vital hubs of commerce, trade, and tourism, optimizing port operations is paramount for economic growth and prosperity. Smart technologies play a pivotal role in this optimization, leveraging advanced sensor networks, real-time monitoring systems, and predictive analytics to enhance safety, mitigate environmental risks, and improve overall efficiency. Additionally, smart coastal management strategies focus on preserving ecosystems, mitigating climate change impacts, and safeguarding against natural disasters. Aerial imagery, facilitated by Unmanned Aerial Vehicles (UAVs) equipped with high-resolution cameras and sensors, provides comprehensive insights into coastal dynamics, harbor operations, and environmental conditions. These images enable efficient monitoring of coastal areas, ports, and harbors, capturing crucial information for informed decision-making in coastal management and port operations. Object detection, particularly in ship detection, stands as a transformative technology for enhancing coastal and harbor management within smart cities. Leveraging advanced algorithms and high-resolution aerial imagery, ship detection systems offer real-time monitoring crucial for optimizing maritime operations and ensuring port security. Object detection algorithms, particularly Faster R-CNN, have shown promise in accurately detecting ships in aerial imagery, offering valuable insights for harbor planning and infrastructure development. This study focuses on utilizing the Faster R-CNN detection algorithm for ship detection in coastal and harbor environments, highlighting its potential to bolster security applications and contribute to the resilience of smart city infrastructure. Through rigorous evaluation and optimization, this research aims to enhance the effectiveness of ship detection systems in safeguarding coastal and harbor environments within smart cities.**

Keywords—object detection; coastal management; smart cities; unmanned aerial vehicle; faster R-CNN

I. INTRODUCTION

In smart cities, efficient coast and harbor management represents a pivotal component intertwining economic vitality, environmental sustainability, and urban resilience. Coastal cities with bustling harbors often serve as hubs of commerce, trade, and tourism, making efficient port management critical for economic growth and prosperity. Smart technologies play a crucial role in optimizing harbor operations, encompassing vessel traffic management, cargo handling, and port logistics [1, 2]. Advanced sensor networks, real-time monitoring systems, and predictive analytics enable harbor authorities to enhance safety, mitigate environmental risks, and improve overall efficiency [3]. Moreover, smart coastal management strategies focus on preserving and restoring coastal ecosystems, mitigating the impacts of climate change, and safeguarding against natural disasters such as storms and sea-level rise.

Aerial imagery serves as a powerful tool in addressing coast and harbor management issues within smart cities by providing comprehensive insights into coastal dynamics, harbor operations, and environmental conditions [4, 5]. Through the deployment of Unmanned Aerial Vehicles (UAVs) equipped with high-resolution cameras and sensors, aerial imagery enables efficient and cost-effective monitoring of coastal areas, ports, and harbors [6]. These images capture crucial information such as shoreline changes, sediment transport patterns, and infrastructure deterioration, facilitating informed decision-making for coastal management and port operations. Additionally, aerial imagery can aid in identifying potential hazards such as pollution, oil spills, or illegal activities, enabling prompt intervention and mitigation measures [7, 8]. Moreover, by employing advanced image processing techniques and machine learning algorithms, aerial imagery can automate the analysis of vast amounts of data, extracting actionable insights for harbor authorities, environmental agencies, and urban planners [9].

Object detection, particularly in the context of ships, stands as a transformative technology for enhancing coastal and harbor management within smart cities. By leveraging advanced algorithms and high-resolution aerial imagery, ship detection systems offer real-time monitoring capabilities crucial for optimizing maritime operations and ensuring port security [10]. These systems can detect and track ships entering and leaving harbors, facilitating efficient vessel traffic management and berth allocation. Additionally, ship detection aids in identifying illegal activities such as smuggling or unauthorized anchorage, enabling prompt intervention by harbor authorities [11]. Moreover, by analyzing ship movement patterns and traffic density, these systems provide valuable insights for harbor planning and infrastructure development, helping to optimize navigational routes and mitigate environmental impacts.

Leveraging cutting-edge technologies such as deep learning and computer vision, object detection algorithms are adept at detecting ships in aerial imagery captured by satellites or UAVs with high accuracy and speed. One prominent approach involves the utilization of convolutional neural networks (CNNs), which excel at identifying patterns and features within images, allowing for the precise detection and classification of ships amidst complex maritime environments [12, 13]. Furthermore, object detection algorithms enable the tracking of ship movements over time, providing valuable insights into vessel traffic patterns, port activities, and maritime anomalies. These insights empower harbor authorities and maritime agencies to optimize port operations, enhance safety measures, and respond effectively to emergencies such as oil spills or illegal fishing activities [14]. By harnessing object detection algorithms, smart cities can bolster their coastal and harbor management capabilities, fostering sustainable maritime development and ensuring the resilience of coastal communities in the face of evolving challenges.

Aerial harbor monitoring through the utilization of UAVs represents an alternative approach to enhancing maritime management within smart cities [15]. Equipped with high-resolution cameras and advanced sensors, these vehicles provide a bird's-eye view of harbor activities, offering real-time insights into vessel traffic, port operations, and environmental conditions. By conducting systematic aerial surveys, UAVa enable efficient coverage of expansive harbor areas, facilitating the collection of rich and diverse datasets crucial for informed decision-making [16, 17]. These aerial images serve as invaluable inputs for object detection algorithms, which are adept at identifying and tracking ships, cargo containers, and other maritime assets with precision and speed [18-20]. By leveraging object detection algorithms, harbor authorities gain the ability to monitor vessel movements, detect anomalies such as unauthorized entry or hazardous cargo, and optimize port logistics in real-time. Moreover, the integration of UAV-based aerial monitoring and object detection technology enhances port security, safety, and operational efficiency, contributing to the resilience and sustainability of smart cities' maritime infrastructure [21].

This study is centered on the detection of ships along coastlines and within harbors utilizing the Faster R-CNN detection algorithm. Focused within the scope of smart city management, the research delves into the intricate process of analyzing images acquired through aerial surveillance via UAVs, emphasizing ship detection for coastal and port security purposes. Through detailed examination, the algorithm's efficacy, trained on datasets compiled from UAV-acquired images, is thoroughly assessed on a novel and unprecedented test set. Insights gleaned from this evaluation shed light on the potential of harnessing the algorithm as a powerful tool in bolstering security applications within the ambit of smart cities. This holistic investigation not only underscores the algorithm's performance in ship detection but also outlines avenues for its optimization and integration into real-world security frameworks, contributing to the advancement of smart city infrastructure and safeguarding coastal and harbor environments.

II. METHODOLOGY

Convolutional Neural Networks (CNNs) represent a specialized form of linear processing within the broader field of deep learning, primarily employed for visual data analysis. In image processing, the CNN algorithm receives images as input and employs mathematical operations, namely neural network retrieval, to extract and classify features present within the images. This algorithm plays a pivotal role in imbuing machines with vision capabilities. Structurally, CNNs are composed of five layers: the input layer, convolutional layer, pooling layer, activation layer, flattening layer, and fully connected layer. Upon entering the algorithm, the image undergoes convolutional processing in the primary convolutional layer, where filters, smaller than the image size, traverse the image's pixels, filtering out specific values as they move across. The subsequent pooling layer serves to eliminate extraneous features, enhancing processing efficiency, and comprises two primary pooling methods: maximum pooling and average pooling. In the activation layer, the Droupout function, previously utilized to deactivate select neurons, has been supplanted by the Rectifier (ReLu) function, expediting neural network training. The primary objective of this layer is to confer autonomy upon neurons, thereby enhancing system efficiency. Following this, the flattening layer prepares the input data for the pivotal fully connected layer by converting multidimensional matrices into one-dimensional counterparts. In the final and most crucial layer, the fully connected layer, the one-dimensional matrix representing the image is transformed into a flat vector, initiating the learning process. In essence, CNNs adeptly extract and preserve image features while reducing the image size, thereby facilitating efficient visual data analysis.

The development of the R-CNN (Region-Based Convolutional Neural Network) algorithm was motivated by the need to improve the efficiency of object detection and address the complexity of multi-region selection tasks. Unlike traditional methods that entail classifying numerous regions individually, R-CNN adopts a more streamlined approach by segmenting the image into approximately two thousand region proposals. These proposals serve as potential locations where objects may be present, allowing the algorithm to focus its computational resources on relevant areas. Moreover, R-CNN goes beyond mere classification by incorporating offset value estimation, aiming to refine the precision of bounding boxes surrounding detected objects. Despite its innovative approach, R-CNN faces a significant challenge in terms of training duration. Each image requires the classification of two thousand region proposals, leading to

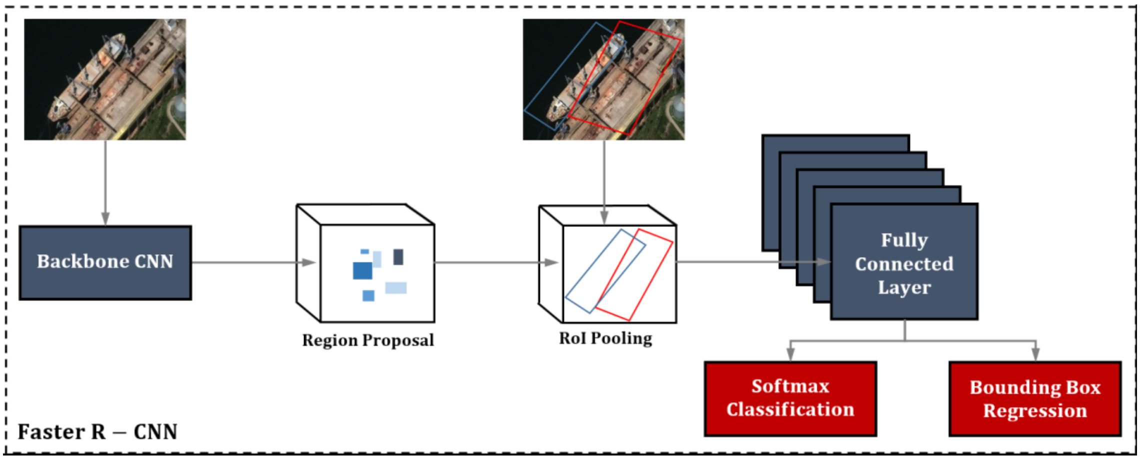

Fig. 1. Faster R-CNN network architecture.

considerable computational overhead. Consequently, the training process for R-CNN is time-intensive, with an average duration of approximately 47 seconds per test image and an overall training time spanning around 84 hours. As a result, while R-CNN demonstrates promising capabilities in object detection, its impractical training requirements render it unsuitable for real-time applications, highlighting the need for more efficient algorithms in the field of computer vision.

To address the limitations of R-CNN, a more efficient object detection algorithm, known as Faster R-CNN, was developed. In the Faster R-CNN algorithm, the image is directly fed into the CNN without the need for prior segmentation into region proposals. Instead, a convolutional feature map is generated, serving as the foundation for subsequent processing. Once the convolutional feature map generates region proposals, they are mapped back to pixels and then transformed into a one-dimensional matrix through the Region of Interest (RoI) pooling layer, facilitating integration into the fully connected layer. Employing a softmax layer, the algorithm estimates both the class of the region proposal obtained from the RoI feature vector and the offset values of the bounding box. The efficiency of Faster R-CNN is underscored by its ability to create a feature map through a single convolution process for each image, contrasting with R-CNN's iterative segmentation of two thousand region proposals. Training the Faster R-CNN algorithm typically requires approximately 8.75 hours, striking a balance between speed and accuracy crucial for real-time object detection. Its superior performance has positioned Faster R-CNN as the preferred choice, delivering more precise results within a comparatively shorter timeframe than its predecessors.

The architecture of Faster R-CNN is depicted in Fig. 1. The algorithm takes the entire image and a collection of object proposals as input. The initial step involves processing the image through a series of convolutional and max-pooling layers to generate a convolutional feature map. Subsequently, the RoI pooling layer operates on this feature map to extract a fixed-length feature vector for each object proposal. These feature vectors are then passed through a set of fully connected layers. Notably, the output layers are bifurcated into two parallel branches. One branch computes softmax probability estimates for K object classes, while the other outputs real-valued numbers corresponding to each class. Within this latter branch, every fourth value encodes adjustments for bounding box positions, facilitating refinement for each of the K classes. The RoI pooling mechanism employs max pooling to transform features within designated regions of interest into a compact feature map with fixed spatial dimensions, with parameters such as image height and width being layer-specific hyperparameters, independent of any RoI.

As part of this study, training and test datasets were carefully curated through extensive aerial monitoring conducted using a UAV, encompassing multiple reconnaissance flights conducted along coastal regions. Special emphasis was placed on ensuring the comprehensiveness of the dataset, incorporating a diverse array of ship types, varying scales, and backgrounds within the images. Sample images from the training dataset are illustrated in Fig. 2, providing a glimpse into the breadth and depth of the dataset used for algorithmic training. The training dataset comprises a total of 3655 images, collectively featuring 8761 ships. Within this dataset, 2833 images were allocated for training purposes, while 244 images were reserved for validation, and 578 images were designated for inclusion in the test dataset.

III. Results

The algorithm, trained on a comprehensive training dataset and refined through validation, underwent rigorous detection testing using a carefully crafted test dataset. A subset of detection results is illustrated in Fig. 3, providing insights into the algorithm's performance. Notably, images 1-5 showcase seamless object and ship detections, underscoring the algorithm's robust capabilities under optimal conditions. The clarity of these detections can be attributed to the absence of waves on the sea surface, rendering ships distinctly visible against a serene backdrop. Furthermore, the relatively large size of ships and their distinct coloration contribute to the efficacy of detection. The uniformity of the background further facilitates the algorithm's task, enabling ships to be readily distinguished. Detailed scrutiny of the bounding boxes reveals precise framing around the ships, affirming the algorithm's accuracy in delineating detected objects. This exemplary performance underscores the algorithm's

Fig. 2. Examples of images from the data set used to train the algorithm.

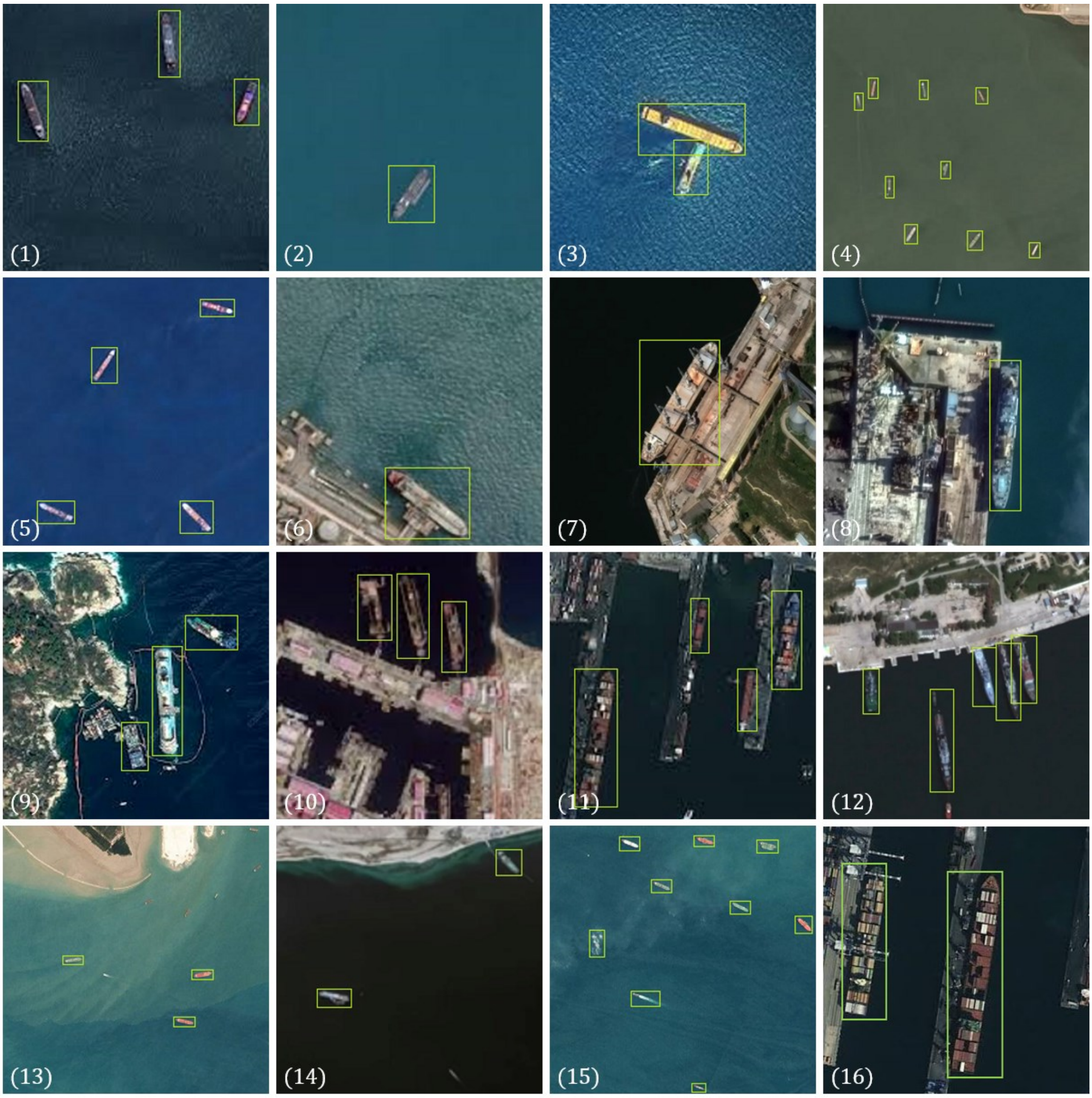

Fig. 3. Examples of detection tests performed on UAV-based images with Faster R-CNN. While images 1-8 are examples of error-free detections, images 9-16 show examples that undetected or contain incorrect detections.

effectiveness in scenarios characterized by favorable environmental conditions, where clear distinctions between objects and backgrounds enhance detection accuracy.

In examples 6-8, despite the proximity of ships to the coastline, the algorithm executed ship detection flawlessly. Detecting ships in coastal regions presents inherent challenges due to the diverse array of objects and backgrounds scattered across the land. Unlike the open sea, where ships stand out against a uniform backdrop, coastal areas feature a myriad of structures and terrain variations, making it arduous for the algorithm to discern ships with the same clarity. However, in the specific instances highlighted (6-8), the algorithm demonstrated remarkable efficacy in detecting ships amidst coastal complexities.

In the ninth image, despite the object on the far left not being a ship, the algorithm erroneously identified it as one. This misclassification can be attributed to the algorithm comparing the object with the ship on the far right, leading to a false positive detection. As mentioned earlier, the algorithm is prone to detection errors, especially in terrestrial areas characterized by a multitude of textures and diverse features. Similarly, in image number 10, the leftmost object was inaccurately identified as a ship, despite lacking the characteristic attributes of a vessel. These instances underscore the challenges faced by the algorithm in accurately distinguishing between objects in complex and varied environments, where the presence of diverse textures and structures can confound detection algorithms.

As depicted in the eleventh image, the algorithm encountered difficulty in detecting two dark ships positioned behind the central red vessel. This challenge likely stemmed from the indistinct outlines of the ships and the closely matched color tones between the ships and the surrounding sea. The lack of clear delineation and the similarity in color between the ships and their aquatic backdrop posed obstacles to accurate detection. Conversely, the other two red ships, which were approximately the same size, were effortlessly detected due to their comparatively greater visibility. Their vivid coloration and distinct outlines enabled the algorithm to discern them with ease, underscoring the pivotal role of visual clarity and color contrast in facilitating accurate object detection.

Regarding the twelfth image, despite clear visibility and the absence of obstructions in its vicinity, a small ship positioned behind the larger vessel in the central area remained undetected by the algorithm. This unexpected outcome highlights a limitation in the algorithm's ability to discern objects of varying sizes and spatial relationships, particularly in complex visual environments. Similarly, in image number 13, the algorithm failed to detect six small ships, a predicament attributed to their diminutive size relative to the surrounding context. The inherent challenges associated with detecting small objects amidst dynamic backgrounds, compounded by the undulating structure of the area and the presence of diverse background colors, further complicated detection efforts. These factors collectively contribute to the algorithm's difficulty in accurately identifying small ships, underscoring the need for ongoing algorithmic refinement to enhance detection performance in scenarios characterized by spatial intricacies and varied visual elements.

As portrayed in image 14, the ship positioned at the bottom right, relatively smaller in size compared to its counterparts, eluded detection by the algorithm. This observation underscores a recurring challenge wherein the algorithm struggles to discern smaller-sized ships within the same image. While the algorithm managed to detect all other objects in image number 15, it erroneously identified a condensed cloud cluster in the middle left section as a ship. Furthermore, the algorithm included a small ship located in the lower right diagonal of the cloud cluster within the same bounding box, along with its trace. While this reflects accurate object detection, the inclusion of the ship's trajectory in the bounding box could potentially complicate critical scenarios such as rescue operations. Finally, in the last image, a partially visible ship situated on the far right remained undetected by the algorithm, illustrating a limitation in its ability to identify objects with obscured visibility.

IV. Conclusion

Based on the results obtained from the rigorous testing and evaluation of the Faster R-CNN's performance in ship detection scenarios, several key observations and insights have been gleaned. The algorithm, trained and validated on comprehensive datasets, has demonstrated varying degrees of effectiveness across different environmental conditions and object characteristics. Firstly, the algorithm exhibited exemplary performance under optimal conditions. The absence of waves on the sea surface, coupled with the relatively large size and distinct coloration of ships against a uniform backdrop, facilitated accurate detection. However, challenges arose when detecting ships in coastal regions, where the algorithm encountered complexities inherent to varied landscapes and structures. Despite these challenges, the algorithm demonstrated some degree of efficacy in detecting ships amidst coastal complexities, showcasing its adaptability to diverse environmental settings. Nevertheless, instances of misclassification and undetected objects were observed, particularly in scenarios characterized by visual intricacies and spatial complexities.

Misclassifications occurred when the algorithm erroneously identified non-ship objects as ships, highlighting the algorithm's susceptibility to detection errors in environments with diverse textures and features. Additionally, challenges were encountered in accurately detecting ships with indistinct outlines or obscured visibility. The algorithm's limitations in discerning objects of varying sizes and spatial relationships, particularly in dynamic visual environments, were also evident in instances where small ships remained undetected. While the algorithm has demonstrated promising capabilities in ship detection, there remains room for improvement, particularly in addressing challenges associated with diverse environmental conditions, object characteristics, and spatial intricacies. Continued refinement and optimization of the algorithm, coupled with advancements in deep learning and computer vision technologies, hold the potential to enhance its performance and applicability in real-world scenarios, thereby contributing to the advancement of coastal and port management within smart cities.

References

[1] D'Amico, g., Szopik-Depczyńska, K., Dembińska, I., Ioppolo, G., "Smart and sustainable logistics of Port cities: A framework for comprehending enabling factors, domains and goals," *Sustai. Cit. Soc.*, vol. 69, no. 102801, 2021.

[2] Zhang, Z., Song, C., Zhang, J., Chen, Z., Liu, M., et al., "Digitalization and innovation in green ports: A review of current issues, contributions and the way forward in promoting sustainable ports and maritime logistics," *Sci. Tot. Env.*, vol. 912, no. 169075, 2024.

[3] Zaman, I., Pazouki, K., Norman, R., Younessi, S., Coleman, S., "Challenges and opportunities of big data analytics for upcoming regulations and future transformation of the shipping industry," *Proced. Eng.*, vol. 194, pp. 537-544, 2017.

[4] Rumson, A.G., Hallett, S.H., "Opening up the coast," *Ocean Coast. Manag.*, vol. 160, pp. 133-145, 2018.

[5] Rumson, A.G., Hallett, S.H., Brewer, T.R., "Coastal risk adaptation: the potential role of accessible geospatial big data," *Marine Pol.*, vol. 83, pp. 100-110, 2017.
[6] Yang, Z., Yu, X., Dedman, S., Rosso, M., Zhu, J., et al., "UAV remote sensing applications in marine monitoring: Knowledge visualization and review," *Sci. Tot. Env.*, vol. 838, no. 155939, 2022.
[7] Messinger, M., Silman, M., "Unmanned aerial vehicles for the assessment and monitoring of environmental contamination: An example from coal ash spills," *Env. Poll.*, vol. 218, pp. 889-894, 2016.
[8] Bakirci, M., Ozer, M.M., "Post-disaster area monitoring with swarm UAV systems for effective search and rescue," *2023 10th International Conference on Recent Advances in Air and Space Technologies*, Istanbul, Turkiye, 2023, pp. 1-6.
[9] Yaseen, Z.M., "An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions," Chemosphere, vol. 277, no. 130126, 2021.
[10] Okafor-Yarwood, I., Eastwood, O., Chikowore, N., Paes, L.O., "Technology and maritime security in Africa: Opportunities and challenges in Gulf of Guinea," *Marine Pol.*, vol. 160, no. 105976, 2024.
[11] Shanthi, T.S., Dheepanbalaji, L., Priya, R., Kumar, V.D.A., Kumar, A., "Illegal fishing, anomalous vessel behavior detection through automatic identification system," *Mater. Today: Proceed.*, vol. 62, pp. 4685-4690, 2022.
[12] Jiang, B., Ma, X., Lu, Y., Li, Y., Feng, L., Shi, L., "Ship detection in spaceborne infrared images based on Convolutional Neural Networks and synthetic targets," *Infr. Phys. Tech.*, vol. 97, pp. 229-234, 2019.
[13] Deng, Z., Sun, H., Zhou, S., Zhao, J., Lei, L., Zou, H., "Multi-scale object detection in remote sensing imagery with convolutional neural networks," ISPRS J. *Photog. Rem. Sens.*, vol. 145, pp. 3-22, 2018.
[14] Alpers, W., Holt, B., Zeng, K., "Oil spill detection by imaging radars: Challenges and pitfalls," Remot. Sens. Env., vol. 201, pp. 133-147, 2017.
[15] Konar, B., Iken, K., "The use of unmanned aerial vehicle imagery in intertidal monitoring," *Deep Sea Res.: Topic. Stud. Ocean.*, vol. 147, pp. 79-86, 2018.
[16] Cheng, K.H., Chan, S.N., Lee, J.H.W., "Remote sensing of coastal algal blooms using unmanned aerial vehicles (UAVs)," *Mari. Poll. Bull.*, vol. 152, no. 110889, 2020.
[17] Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M.F., et al., "Use of unmanned aerial vehicles for efficient beach litter monitoring," *Mari. Poll. Bull.*, vol. 131, pp. 662-673, 2018.
[18] Zhao, C., Liu, R.W., Qu, J., Gao, R., "Deep learning-based object detection in maritime unmanned aerial vehicle imagery: Review and experimental comparisons," *Eng. Appl. Artif. Intel.*, vol. 128, no. 107513, 2024.
[19] Yuan, H., Xiao, C., Wang, Y., Peng, X., Wen, Y., et al., "Maritime vessel emission monitoring by an UAV gas sensor system," *Ocean Eng.*, vol. 218, no. 108206, 2020.
[20] Kumar, S.V., Jayaparvathy, R., Priyanka, B.N., "Efficient path planning of AUVs for container ship oil spill detection in coastal areas," *Ocean Eng.*, vol. 217, no. 107932, 2020.
[21] Meng, Y., Wang, W., Han, H., Ban, J., "A visual/inertial integrated landing guidance method for UAV landing on the ship," *Aerosp. Sci. Tech.*, vol. 85, pp. 474-480, 2019.