







## Deep Learning for Marine Vehicles Parking Availability: A ResNet50-Based Deep Feature Engineering Model



Mert Gurturk<sup>1</sup>, Veysel Yusuf Cambay<sup>2</sup>, Abdul Hafeez-Baig<sup>3</sup>, Rena Hajjiyeva<sup>4</sup>, Sengul Dogan<sup>5\*</sup>, Turker Tuncer<sup>5</sup>

<sup>1</sup> Civil Engineering Department, Engineering Faculty, Adiyaman University, Adiyaman 02040, Turkey

<sup>2</sup> Department of Electrical and Electronics Engineering, Mus Alparslan University, Mus 49001, Turkey

<sup>3</sup> School of Business, University of Southern Queensland, Toowoomba 4350, Australia

<sup>4</sup> Department of Information Technologies, Western Caspian University, Baku AZ1001, Azerbaijan

<sup>5</sup> Department of Digital Forensics Engineering, College of Technology, Firat University, Elazig 23119, Turkey

Corresponding Author Email: [sdogan@firat.edu.tr](mailto:sdogan@firat.edu.tr)

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420206>

### ABSTRACT

**Received:** 24 July 2024

**Revised:** 23 October 2024

**Accepted:** 6 November 2024

**Available online:** 30 April 2025

#### Keywords:

*deep feature engineering, feature selection with INCA, marine engineering, parking availability detection for ships, ResNet50, semi overlapped patch division*

In this research, our essential objective is to evaluate the availability of parking spaces within port/marine/fisher shelters employing a novel computer vision-based approach. Therefore, we collected a new dataset and developed a ResNet50-based image classification model to detect parking status. Initially, we collected a new image dataset using an unmanned aerial vehicle (UAV) from over 200 fisher shelters and the collected image dataset contains two classes which are parking available or not (full). To automatically detect parking available fisher shelters, a new ResNet50-based deep feature engineering (DFE) approach has been recommended. In the recommended DFE approach, we introduced a novel semi-overlapped patch division strategy to extract local features like transformers. To implement this model, we first trained the ResNet50 approach on our collected training dataset and a trained ResNet50 model has been obtained. Subsequently, deep features have been derived using the proposed semi-overlapped patch division approach and the global average pooling (GAP) layer of the trained ResNet50. Nine feature vectors have been generated using patches and a feature vectors has been extracted from the whole image. By using this strategy, we have generated both local and global features and these features have been merged to create the ultimate feature vector. To select informative features from the generated ultimate feature vector, iterative neighborhood component analysis (INCA) feature selector has been applied. The chosen features by INCA were employed as input of the support vector machine (SVM), is a shallow classifier, classifier to create classification results. The used ResNet50 convolutional neural network (CNN) attained 100% training accuracy and 94.23% validation accuracy. Subsequently, the recommended DFE model was assessed on test images, achieving a test classification accuracy of 97.27%. Furthermore, we utilized Grad-CAM and feature analysis to provide interpretable results for the presented model. The achieved classification performance and the explanatory outcomes demonstrably illustrate the capability of the proposed model for automatic detection of parking availability in fisher shelters. These findings support the utility of computer vision as a viable solution for this application.

## 1. INTRODUCTION

The maritime industry is a sector that is constantly expanding and actively trading. Ports, marinas and fisher shelters are needed for the continuity and effective management of this sector. Especially the efficient use of these areas will ensure the continuity of trade and transportation. One of the issues of efficient use is the correct identification of parking areas and planning according to these parking areas [1, 2]. The management methods used in this field (the shipping/maritime industry) are generally divided into two main branches: human control/command-based models and sensor network models. Even though systems under human

control are old and simple, human errors are frequently encountered. Sensor network systems provide automatic control using multiple sensors. Although this is a more modern and efficient method, it is more expensive and complex to install and maintain [3]. Along with developing deep learning (DL) models, computer vision-based automatic systems have the potential to provide important solutions for the usability of parking areas [4].

DL models/networks are very effective for image classification and segmentation. Moreover, DL networks have changed the basics of computer vision due to their high performance and findings [5-7]. There's a missing piece in the puzzle when it comes to applying DL to specific tasks, like

figuring out parking spots for large ships [8]. Current/existing methods for detecting parking for marine vessels have their downsides/limitations. To overcome these limitations, most research focuses on Internet of Things (IoT) systems [9, 10]. However, the lack of publicly available data on maritime/shipping industry and the nature limitations of IoT systems (complex and costly) make it difficult to develop the next generation of advanced models [11].

We have aimed to address this limitation by proposing a DL model. Our proposal is built on a curated image dataset and a novel deep feature engineering (DFE) approach. This study demonstrates the potential of DL for maritime parking

optimization. Moreover, we have presented a new patch-division technique for DFE and this patch-division technique is named semi-overlapped patch division. We have presented an innovative approach to contribute computer vision.

### 1.1 Related works

Machine learning (ML) techniques are used in literature for different disciplines [12-14]. In this section, we have surveyed literature about to maritime industry and ML and the surveyed papers have been outlined in Table 1.

**Table 1.** ML-based studies using image dataset for ship detection classification and marine parking availability

Study	Method	Classifier	Data	Number of Class	Data Augmentation	The Results
Kim et al. [15]	Faster region-based CNN	Softmax	7000 (75:25)	7	No	The mean average precision: 93.92 Accuracy: 98.75
Escorcia-Gutierrez et al. [16]	Mask regional CNN, Colliding body's optimization	Weighted extreme learning machine	2000 (80:20)	2	Yes	F1: 98.75 Precision: 98.60 Recall: 98.90 1. Overall accuracy: 95.20
Ma et al. [17]	CNN	Softmax	1. 3210 (79:21) 2. 1727 (32:30:38)	1. 8 2. 6	1. Yes 2. No	2. The mean average precision: 93.92
Feng et al. [18]	CNN	Softmax	1. 1061 images in the HRSC2016 dataset (41:17:42) 2. 934 MWC (80:20)	1. 19 2. 11	Yes	1. Overall accuracy: 84.50 2. Average precision: 87.10
Huang et al. [19]	The regressive deep CNN based on YOLOv2/v3	Softmax	4200 (80:20)	7	No	The mean average precision: 92.09 Recall: 98.18
Shi et al. [20]	CNN	Softmax	1. 800 images in BCCT200-resize dataset 2. 2865 images in VAIS dataset	1. 4 2. 6	No	Accuracy 1. 98.33 2. 88.00
Chang et al. [21]	YOLOv2	YOLOv2	1. 1160 images in SSDD (70:20:10) 2. 1174 images in DSSDD (70:20:10)	1. 2.	No	Accuracy 1. 90.05 2. 89.13
Zhang et al. [22]	YOLOv5	YOLOv5	7000 (60:20:20)	6	No	Accuracy: 71.60
Wang et al. [23]	Faster R-CNN	Softmax	12650 (Unspecified)	3	No	The mean average precision: 74.26
Leclerc et al. [24]	Deep CNN	Softmax	9216 (89:11) images in Marvel dataset	26	No	Accuracy: 78.73
Shao et al. [25]	CNN	Softmax	31455 (split ratio: unspecified)	7	Yes	The mean average precision: 92.40
Zhang et al. [26]	Regional-based convolutional neural network	Softmax	2152 (80:10:10)	2	Yes	Precision: 95.79 Rec: 96.46
Pan et al. [27]	Target attitude angle-guided network	Softmax	116 (81:19)	11	No	The mean average precision: 73.91
Yu and Shin [28]	YOLOv5	YOLOv5	1.1160 2. 5604 (65:35)	1. 2.	No	The mean average precision: 1. 95.02 2. 85.11

Table 1 section showcases existing ship detection models, which mainly focus on general ship classification using CNNs and object detection approaches like Faster R-CNN and YOLO. Although these models perform well in identifying various ship types, they do not specifically address the challenge of parking availability detection. In addition, IoT-based solutions, though effective, are costly and complex, making them less practical for widespread use.

The introduced ResNet50-based DFE model overcomes these limitations by specifically targeting parking availability detection. By using the semi-overlapping patch splitting technique, it captures both local and global features that are important for identifying parking spaces.

The introduced approach provides a novel and effective solution to the specific challenge of parking availability detection.

Table 1 has summarized a wide range of methods that combine marine vessels with ML. Table 1 has presented the characteristics of the datasets utilized by these methods, as well as the classification/detection performances achieved. Table 1 demonstrates the diversity of current research and the performance of various approaches in the field of ship detection and classification. Most of the studies listed achieved high accuracy rates by utilizing various DL architectures. For example, the study by Kim et al. [15] employed a faster region-based CNN to attain a mean average precision of 93.92%. Another study by Escorcia-Gutierrez et al. [16], which used a mask regional CNN and Weighted extreme learning machine, achieved an accuracy of 98.75%. These studies illustrate the effectiveness of DL techniques in applications within the maritime domain. However, a research by Zhang et al. [22], which utilized YOLOv5, listed a relatively lower performance with an accuracy of 71.60%. This indicates significant performance disparities across research due to the variety of methods and datasets used. This table also shows that some studies applied data augmentation, suggesting that enhancing the model's generalization capability can be achieved through this technique. Overall, Table 1 demonstrates the successful application of ML and DL techniques in ship detection and classification. The results of these studies highlight the potential of these technologies for significant applications in the maritime industry, including determining the availability of parking spaces. These contributions provide insights that could serve as a foundation for future research.

## 1.2 Literature gaps

The following gaps have been identified according to the literature review:

- There is a notable scarcity of research papers specifically focused on ship parking detection. Most existing studies rely on IoT-based systems, which are complex and expensive to implement, limiting their practicality in large-scale marine environments.
- Current methods emphasize IoT systems, but there is a pressing need for more mobile and cost-effective solutions to detect parking availability for ships. Such systems could provide greater flexibility and scalability in diverse marine environments.
- The lack of open-access datasets for ship parking detection presents a significant barrier to developing and benchmarking new models, particularly those using advanced computer vision and DL techniques. This hinders the progress of next-generation solutions in the field.

## 1.3 Motivation and study outline

Ship parking detection is a critical task in marine environments such as ports, marinas, and fishing ports to ensure smooth operations and minimize congestion. However, existing solutions in this area are mostly based on IoT systems, which are often complex and costly to implement. These systems are typically expensive. Moreover, there is a significant lack of research specifically focused on ship parking detection, which further limits the development of innovative, automated solutions. Additionally, the scarcity of open-access datasets poses a major obstacle to the advancement of computer vision and DL models that could address these challenges more efficiently. Given the urgent need for cost-effective alternatives, this research aims to fill

the gaps in the literature by proposing a novel DL-based approach for ship parking detection.

The literature review highlights a scarcity of automated systems for detecting ship parking availability despite the potential of DL models to address various computer vision challenges. Our primary goal is to address the parking availability issue in marines, ports, and fisher shelters, among other areas. To this end, we have compiled a new dataset using an unmanned aerial vehicle (UAV) divided into two categories: 'full' and 'parking available.'

The proposed ResNet50-based DFE model is inspired by a popular image transformers (ViT) [29], which demonstrate the effectiveness of fixed-size patch division for extracting features from local areas. However, the patch-based feature extraction approach is a complex approach. For example, ViT uses patch sizes such as  $14 \times 14$ ,  $16 \times 16$  and  $32 \times 32$ , and using these sizes, 256, 196 and 49 fixed-size patches are obtained from a  $224 \times 224$  image, respectively. To obtain a more linear method and use the effectiveness of fixed-size patch division, we proposed a new patch division approach, the semi-overlapped patch division technique. In this work, we used  $112 \times 112$  patches with a stride of 56 and the created 9 patches have been employed to extract local features. Additionally, we used the entire image to extract global features. At this stage, we used trained ResNet50 as feature extractor. We generated the ultimate feature vector by combining the obtained local and global features. Then, we employed iterative neighborhood component analysis (INCA) [30] to choose the most meaningful/valuable features to solve the problem in the ultimate feature vector.

Classification results have been obtained using a shallow classifier, specifically a support vector machine (SVM) [31]. Additionally, a Bayesian optimizer [32] has been employed to tune the optimal hyperparameters for the utilized SVM classifier to attain higher classification performance.

## 1.4 Novelties and contributions

We have presented a parking availability detection model for marine vehicles and as our surveyed literature, we are the first team to present a solving for parking availability detection for marine vessels using computer vision.

*Novelties:*

- We have gathered a novel image dataset using an UAV to create a testbed for parking availability detection.
- The semi-overlapped patch division technique has been presented to extract local features from the local areas using less complexity in this research.

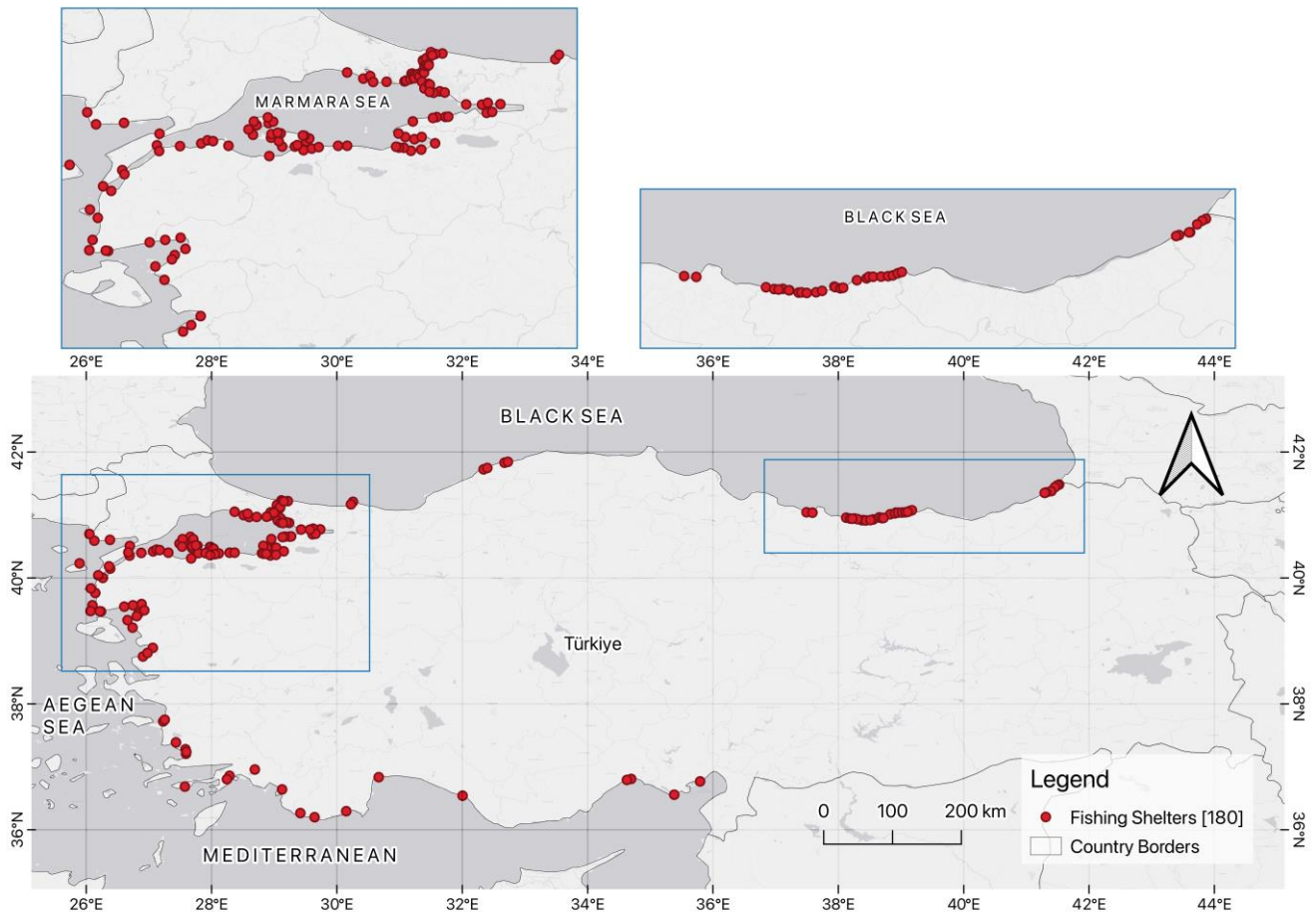
*Contributions:*

- Addressing the parking availability issue for ships is a prevalent and economically significant problem, especially considering its implications for port congestion. By compiling a new dataset from fisher shelters and marines specifically for parking availability detection, we contribute to refining the methodology for parking availability analysis.
- We propose a novel DFE model inspired by ViT. This model employs semi-overlapped blocks to extract local and global deep features effectively. The DFE model has demonstrated a test classification accuracy of 97.27% on the marine image dataset, underscoring its capability to detect parking availability for ships accurately. This result shows that DL can be useful for real-world problems in marine logistics. It also supports the use of ML for detecting ship parking availability.

## 2. THE COLLECTED IMAGE DATASET

In this study, images were collected from over 200 fisher shelters with varying vessel capacities along the coasts of the

Mediterranean, Black Sea, Marmara, and Aegean Seas in Turkey. Figure 1 provides a visual overview of the data collection locations.



**Figure 1.** The data collection locations

The images were captured using a DJI Mavic 3 Enterprise UAV equipped with a 56x zoom camera, a 4/3 CMOS 20MP sensor, and a mechanical shutter. The data were recorded in 4K resolution video format, and the UAV was capable of capturing images with a 0.7-second interval between shots. To ensure the robustness of the dataset, images were collected under varying environmental conditions, including different lighting (morning, afternoon, and evening) and weather conditions (clear, cloudy, and rainy days). This diversity helps to capture a wide range of visual scenarios, testing the model's ability to generalize across different situations.

The collected dataset was separated into training and test subsets. The details are presented in Table 2. The dataset consists of two classes: "Full" (indicating no parking available) and "Parking available." In total, 3,687 images were collected, with 2,772 allocated for training and 915 for testing, maintaining a roughly 75:25 split. However, the dataset is imbalanced, with more images of "Parking available" than "Full." The gathered dataset is partitioned into train and test folders and the distribution of the collected image dataset is demonstrated in Table 2.

As can be highlighted from Table 2, we have used an imbalanced dataset and training and test split ration is approximated 75:25.

The dataset's variations in lighting, weather, and time of day

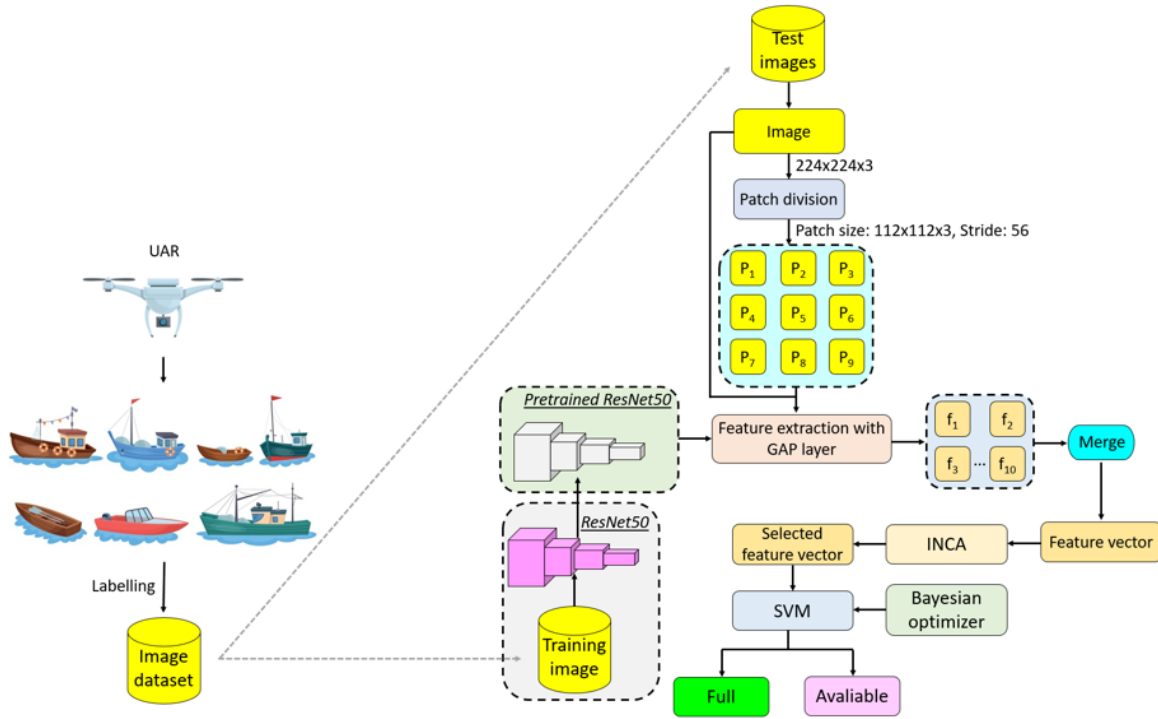
contribute to its robustness and ensure that the approach is exposed to a wide variety of conditions. This diversity allows the model to better generalize to real-world situations.

**Table 2.** ML-based studies using image dataset for ship detection classification and marine parking availability

No.	Class	Train	Test	Total
1	Full	1063	350	1413
2	Parking available	1709	565	2274
	Total	2772	915	3687

## 3. THE INTRODUCED RESNET50-BASED DEEP FEATURE ENGINEERING APPROACH

We introduce a novel DFE approach based on ResNet50, employing one of the most prominent CNNs in our research. ResNet50 is renowned for utilizing residual blocks to mitigate the vanishing gradient problem, a key feature that enhances its performance for DL tasks. The proposed model aims to detect ship parking availability. To achieve this goal, we gathered a new dataset and designed a ResNet50-based exemplar DFE approach. The schematic overview of the recommended approach is demonstrated in Figure 2.



**Figure 2.** The general block diagram of the recommended approach

The dataset was collected using a UAV to capture images of various ship parking areas, and labeling was performed after the dataset was collected. ResNet50, a widely used CNN, was selected as the base feature extractor for the approach. Leveraging the pre-trained version of ResNet50 allowed us to benefit from transfer learning, reduce the need for large amounts of training data, and increase model efficiency. The ResNet50 architecture includes residual connections to avoid the vanishing gradient problem, which enables the approach to perform well even with deeper networks.

A novel semi-overlapping patch splitting strategy was used to improve feature extraction. The images were split into  $112 \times 112 \times 3$  patches with a stride of 56 steps. After patch splitting, both local features from the patches and global features from the entire image were extracted using the GAP (Global Average Pooling) layer of the pre-trained ResNet50 approach. This step reduced the complexity of the feature maps while preserving important information. In total, nine feature vectors were created from the patches and one global feature vector from the entire image, and these were combined to form a comprehensive feature representation of the image.

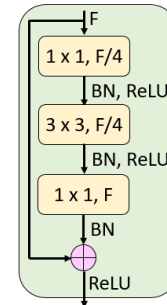
INCA was then implemented to the combined feature vector to choose the most informative features. The chosen features were passed through the SVM. Bayesian optimization was used to fine-tune the hyperparameters of the SVM, improving the classification accuracy by minimizing the misclassification rate.

In this research, the used CNN is ResNet50. Therefore, the core architecture of ResNet50, along with the mathematical formulation of the recommended approach, is detailed in Figure 3.

We present a schematic representation of the main block of ResNet50 in Figure 3. The mathematical framework underpinning ResNet50 is detailed in the following sections.

$$F: (256,512,1024,2048), R: (3,4,6,3) \quad (1)$$

Herein, F: number of filters and R: number of repetitions.



**Figure 3.** The block design of the ResNet50

Utilizing ResNet50, we trained the collected image dataset and created a pretrained ResNet50 CNN alongside a DFE model. This DFE model comprises three main phases:

**Feature extraction:** In this phase, semi-overlapped patch division and the GAP layer of the pretrained ResNet50 were used to generate local and global features.

**Feature selection:** The INCA method was employed to obtain the most informative results.

**Classification:** The SVM classifier was implemented in this phase. Furthermore, a Bayesian optimizer was applied to the SVM classifier to determine the optimal hyperparameters.

Through these phases, we implemented the introduced ResNet50-based DFE approach. A graphical overview of the presented ResNet50-based approach is outlined in Figure 4.

The details of the introduced ResNet50-based DFE approach are given as below step by step.

Step 1: Train the ResNet50 by implementing the training image dataset.

Step 2: Employ semi-overlapped patch division to each test image.

$$P_k = Im(i:i + 111, j:j + 111, :), \quad (2)$$

$$i \in \{1,57, \dots, 113\},$$

$$k \in \{1,2, \dots, 9\}, j \in \{1,57, \dots, 113\}$$

Herein, P: fixed-size patch, Im: image. Herein, the size of the used patch is equal to  $112 \times 112 \times 3$  and stride is 56. In this step, we have generated nine fixed-size patches.

The semi-overlapping patch splitting technique was used to overcome the limitations of traditional fixed-size patch splitting methods, such as those used in Vision Transformers (ViT). Traditional patch splitting usually relies on non-overlapping patches, which can result in the loss of important contextual information between adjacent patches. Additionally, to minimize this issue, the complexity of the model is often increased by using too many patches. In contrast, the semi-overlapping approach allows partial overlap between patches, preserving more local information, ensuring continuity between regions, and enabling more meaningful information to be extracted with larger patch sizes (fewer patches). This approach is especially important in marine parking spot detection, where small details such as ship edges or parking boundaries need to be captured accurately.

Step 3: Generate global features from all images by implementing the GAP layer of the pretrained ResNet50.

$$f_1 = ResNet50(Im, GAP) \quad (3)$$

Herein,  $f_1$  defines the created feature vector as having a length of 2048, and this feature vector represents the global feature vector.

Step 4: Extract local features from the generated patches implementing the GAP layer of the pretrained ResNet50.

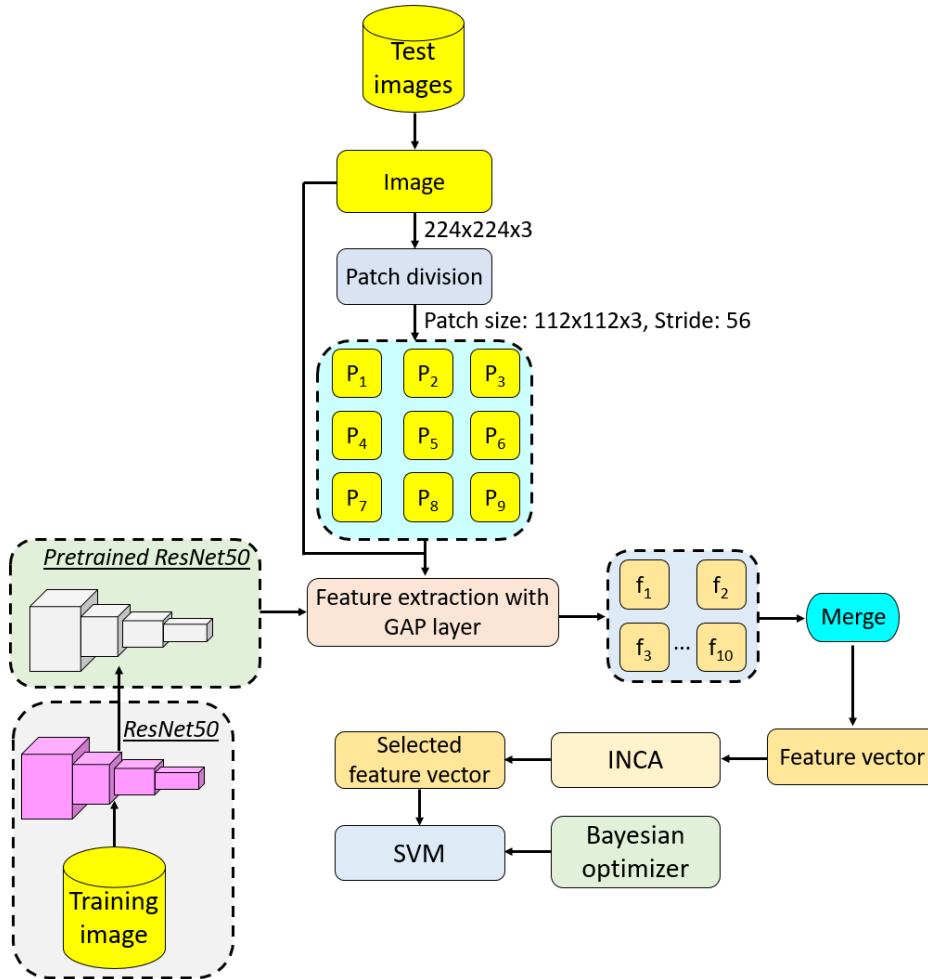
$$f_{k+1} = ResNet50(P_k, GAP) \quad (4)$$

where,  $f_{(2-10)}$ : the local features and the length of each of them is 2048.

Step 5: Concatenate the derived feature vectors.

$$F = merge(f_1, f_2, \dots, f_{10}) \quad (5)$$

where, F: the derived feature vector with a length of 20,480 (=  $2048 \times 10$ ).



**Figure 4.** The introduced DFE approach based on the ResNet50

Step 6: Implement the INCA to the merged feature vector.

$$idx = NCA(F, y) \quad (6)$$

$$S^{r-99}(d, i) = F(d, idx(i)), i \in \{1, 2, \dots, r\}, r \in \{100, 101, \dots, 1000\} \quad (7)$$

where, F: the derived feature vector with a length of 20,480 (=  $2048 \times 10$ ),  $idx$ : the qualified indices,  $y$ : actual output,  $S$ : selected feature vector, and  $r$ : range of the iteration. Above (see Eqs. (6) and (7)), the iterative feature selection process has mathematically been explained, and INCA chooses 901 (=  $1000 - 100 + 1$ ) feature vectors. Among these 901 feature

vectors, the best one has been selected using a greedy algorithm according to the loss values. We have used the SVM classifier to compute loss values. The best feature vector selection process is specified below.

$$\text{loss}(r - 99) = \text{SVM}(S^{r-99}, y) \quad (8)$$

$$[\text{mini}, \text{inx}] = \min(\text{loss}) \quad (9)$$

$$\text{SelFeat} = S^{\text{inx}} \quad (10)$$

Herein, loss: misclassification rate, mini: minimum loss rate, inx: index of the minimum loss rate, and SelFeat: the ultimate chosen feature vector by INCA. Herein, the length of the optimal selected feature vector is 111.

Step 7: Optimize the hyperparameters of the SVM classifier by implementing a Bayesian optimizer.

Step 8: Classify the choose features by deploying the fine-tuned SVM classifier.

$$\text{Pred} = \text{SVM}(\text{SelFeat}, y) \quad (11)$$

where, Pred: predicted vectors and we have used this predicted vector to compute the classification results.

#### 4. EXPERIMENTAL RESULTS

This section explains the results derived from the introduced ResNet50-based DFE approach. The MATLAB programming environment was utilized for the implementation of this model. Specifically, the MATLAB Deep Network Designer was employed to deploy ResNet50, which was pre-trained using the ImageNet1K dataset. Subsequently, our collected dataset was trained using ResNet50 on a personal computer (PC) equipped with 32 gigabytes of RAM, a 3.1 gigahertz CPU, and a GeForce RTX 4090 GPU. The training of ResNet50 was conducted with the following parameters:

Solver: Stochastic Gradient Descent with Momentum (SGDM),

Learning Rate: 0.01,

L2 Regularization: 0.0001,

Epochs: 30,

Mini-batch Size: 32,

Training and Validation Split: 75:25.

The training and validation curves obtained with the aforementioned parameters are illustrated in Figure 5.

The ResNet50 approach attained a training accuracy of 100%, a validation accuracy of 94.23%, and a validation loss value of 0.3877. Utilizing this trained ResNet50 model, we evaluated the test classification performance using several metrics, including classification accuracy, sensitivity, specificity, F1-score and the geometric mean. These performance metrics were calculated based on the confusion matrix derived from the test results depicted in Figure 6.

Based on the confusion matrix presented in Figure 6, the calculated test classification accuracies of the ResNet50 approach are outlined in Table 3.

Table 3 reveals that the ResNet50 approach attained a test accuracy of 96.39% and a test geometric mean of 96.59%. To enhance these results with a reduced feature set, we introduced a ResNet50-based DFE approach with the following initial configurations:

- Feature Extraction:

Patch Division Method: Patches sized  $112 \times 112 \times 3$  with a stride of 56,

Number of Patches: 9,

Feature Extraction Function: GAP layer of the trained ResNet50.

- Feature Selection:

Range of INCA: [100,1000],

Loss Value Calculation: SVM with 10-fold cross-validation, Length of Selected Feature Vector: 111.

- Classification:

Bayesian Optimizer: 100 iterations, focusing on minimizing the misclassification rate.

SVM Configuration: Kernel function is Gaussian, Kernel scale is set to 1.4484, Box constraint is 3.1192, Standardization is not applied, Coding strategy is One-vs-One, and Validation is conducted through 10-fold cross-validation.

Utilizing these settings, the ResNet50-based DFE model was implemented. The calculated test confusion matrix is depicted in Figure 7.

Figure 7 outlines the test results obtained from the implemented DFE model in Table 4.

Table 4 shows that the proposed DFE model achieved a classification accuracy of 97.27% and a geometric mean of 97.40% on the test image dataset. Furthermore, the DFE model demonstrated enhanced test classification performance using only 111 features in conjunction with a shallow classifier.

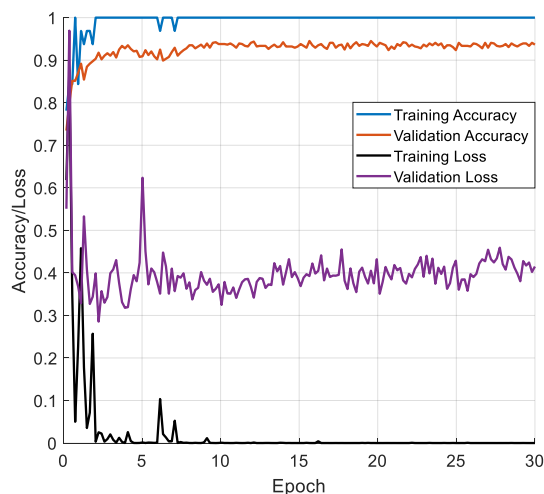
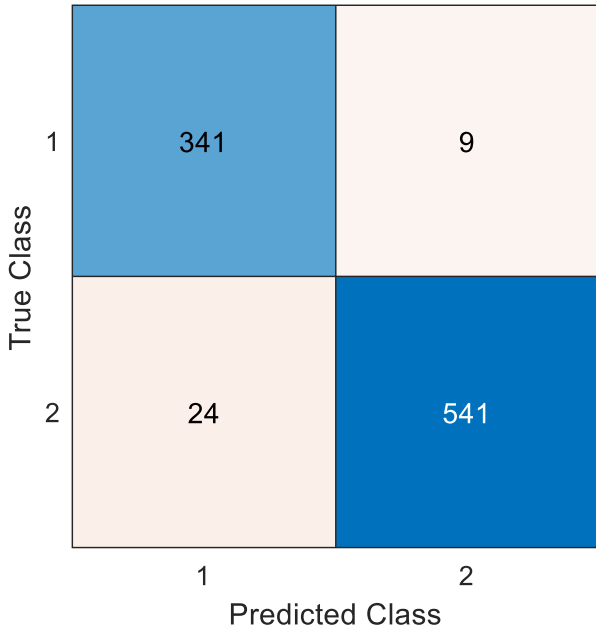


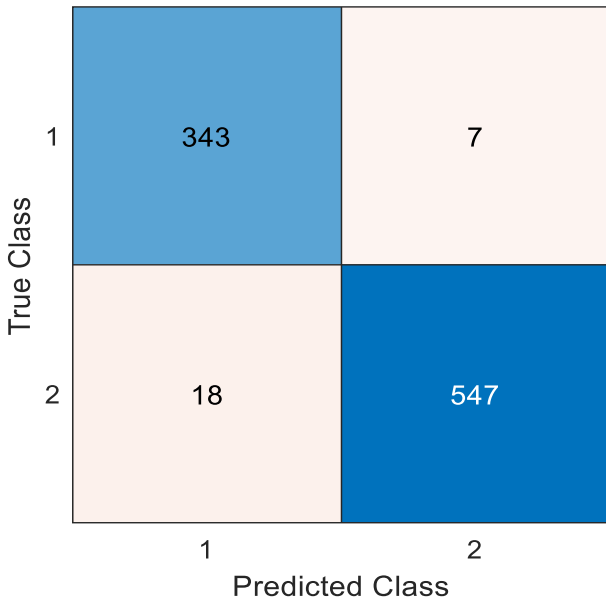
Figure 5. Training and validation curve of the ResNet50 for the gathered image dataset

Table 3. The classification performance metrics

Performance Metric	Class	Result (%)
Classification accuracy	Full	-
	Parking available	-
	Overall	96.39
Sensitivity	Full	97.43
	Parking available	95.75
	Overall	96.59
Specificity	Full	95.75
	Parking available	97.43
	Overall	96.59
F1-score	Full	95.43
	Parking available	97.06
	Overall	96.25
Geometric mean	Full	-
	Parking available	-
	Overall	96.59



**Figure 6.** Test confusion matrix of ResNet50 for ship parking available dataset. 1: Full, 2: Parking available



**Figure 7.** The confusion matrix of the proposed DFE

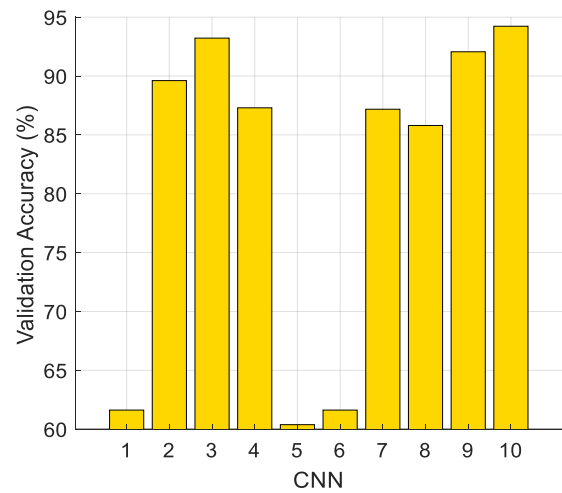
**Table 4.** Results of the presented DFE model

Performance Metric	Class	Result (%)
Classification accuracy	Full	-
	Parking available	-
	Overall	97.27
Sensitivity	Full	98.00
	Parking available	96.81
	Overall	97.41
Specificity	Full	96.81
	Parking available	98.00
	Overall	97.41
F1-score	Full	96.51
	Parking available	97.78
	Overall	97.15
Geometric mean	Full	-
	Parking available	-
	Overall	97.40

## 5. DISCUSSION

In this study, a novel DFE model based on ResNet50 was introduced, and when implemented to the collected image dataset, it achieved a test accuracy of over 97% (97.27% and see Table 4 for more details) utilizing this ResNet50-based DFE model.

To design the proposed model, we tested the well-known CNNs and these are (1) SqueezeNet [33], (2) EfficientNetB0 [34], (3) DenseNet201 [35], (4) DarkNet53 [36], (5) AlexNet [37], (6) VGG19 [38], (7) GoogLeNet [39], (8) ConvNeXt [40] (a CNN designed by us), (9) MobileNetV2 [41], and (10) ResNet50 [42]. These CNNs were trained using our collected image dataset to obtain validation accuracies. Herein, we selected the training and validation separation ration as 70:30. Herein, we aimed to use the best CNN to extract features. Therefore, we used a greedy-based the most suitable CNN selection strategy and the computed validation results of these 10 CNN architectures have been illustrated in Figure 8.



**Figure 8.** Comparison of the CNNs

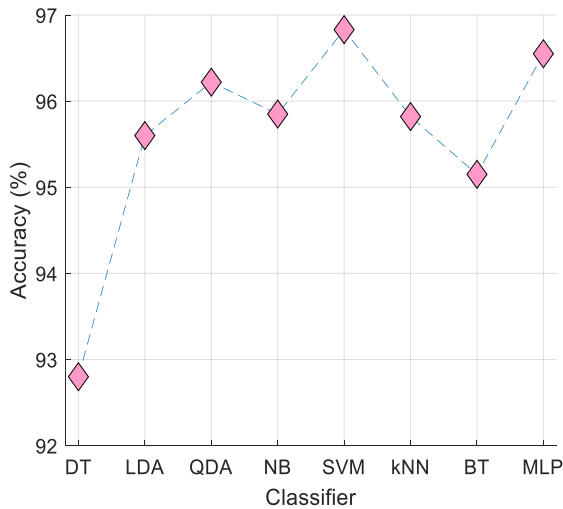
Figure 8 has demonstrated that the best-performing CNN among these 10 CNNs is the 10<sup>th</sup> CNN, which is ResNet50, since this CNN attained a 94.23% validation accuracy. The second CNN, according to validation accuracy, is the 3<sup>rd</sup> CNN (DenseNet201), which reached a 93.22% validation accuracy. Therefore, we selected ResNet50 (ResNet50 has been employed as a feature extractor in our DFE) to create our presented DFE model.

By using ResNet50 and a semi-overlapped patch division technique, we designed the feature extraction method for the suggested DFE model.

To choose the best features from the generated feature vector, INCA (it is an effective iterative feature selector) feature selector has been used.

The last phase of the presented approach is the classification. To choose the most effective shallow classifier for the recommended DFE model, we tested eight shallow classifiers and these are: Decision Tree (DT) [43], Linear Discriminant Analysis (LDA) [44], Quadratic Discriminant Analysis (QDA) [45], Naïve Bayes (NB) [46], SVM [31], k-Nearest Neighbors (kNN) [47], Bagged Tree (BT) [48], and Multilayer Perceptron (MLP) [49]. We selected the best accurate classifier for our proposed DFE model and the computed classification accuracies by using these classifiers have been depicted in Figure 9.



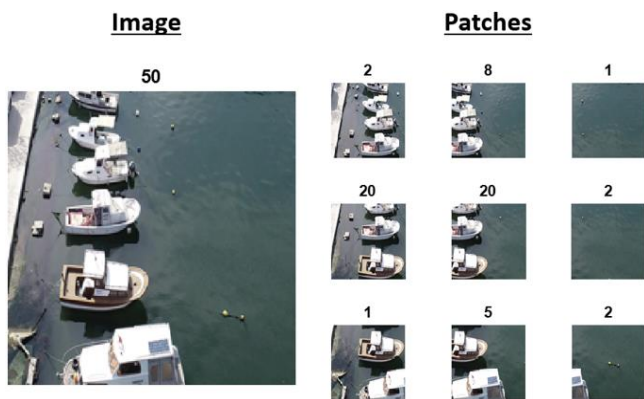


**Figure 9.** The classification accuracies of the employed classifiers

According to Figure 9, the SVM (accuracy: 96.83%) is the most accurate classifier for the parking space availability detection problem. To set the parameters of the SVM, we applied Bayesian optimization. With the fine-tuning obtained through the Bayesian optimizer, the test classification accuracy increased from 96.83% to 97.27%.

Above, we have summarized the story of the presented DFE model and highlighted that a cognitive approximation was utilized to create the proposed model. In the second step of the discussion, we explained the explainable results of our model. To present XAI results, we used feature vector analysis and activations.

The presented DFE model has used a semi-overlapped patch division model, and nine patches have been generated by deploying the presented patch division technique. By deploying patches, the local features have been extracted, and the whole image has been used to extract global features. To evaluate the effectiveness of each created patch and the local and global features, we have employed the findings of the INCA feature selector. The INCA has selected the best 111 features. Here, we have analyzed the source of these features by using a sample image, and this analysis has been illustrated in Figure 10.

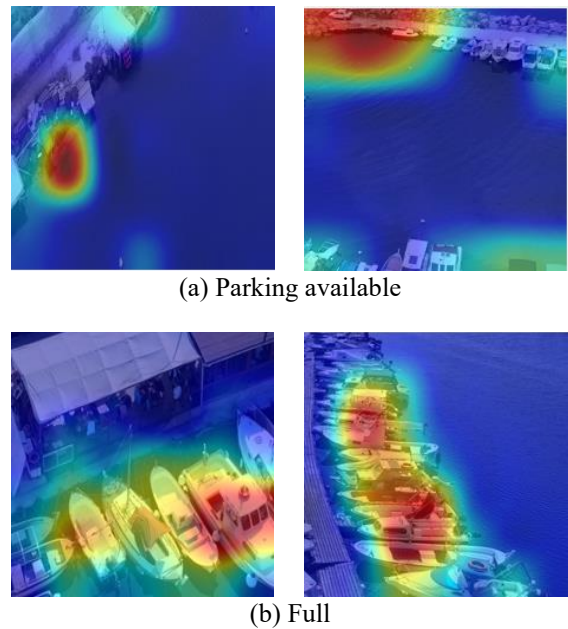


**Figure 10.** The number of the chosen features according to the employed input

Figure 10 has illustrated that 50 out of the selected 111 features were generated from the whole image. Therefore,

these 50 features are termed global features, and the remaining 61 features were generated from the patches. In this aspect, these features are considered local features. The most valuable patches are the 4<sup>th</sup> and 5<sup>th</sup>, since 20 features were generated using the 4<sup>th</sup> patch, and 20 features were selected from those generated by the 5<sup>th</sup> patch.

Furthermore, Gradient-weighted Class Activation Mapping (Grad-CAM) has been employed to provide more explainable results about the proposed model, and we have used sample images to give visually explainable results. These visualizations are shown in Figure 11. In this figure, we have demonstrated heatmaps of the sample images and selected sample images from both classes. The used Grad-CAM model shows how ResNet50 focuses on the region of interest (ROI).



**Figure 11.** Heatmap images by generating Grad-CAM

Figure 11 illustrates that the proposed ResNet50-based model can easily focus on the ROI.

The findings obtained using Grad-CAM allow the model to highlight specific areas of interest, such as parking areas or ship locations, during decision-making. This capability is critical for practical deployment in the maritime industry, as it helps port managers trust the model and easily integrate it into their workflows.

Grad-CAM also ensures the reliability of the model by verifying that the model focuses on relevant features, avoiding noise or irrelevant parts of the image. It helps troubleshoot and improve the model, identify areas for improvement, and reduce misclassifications. The explainability offered by Grad-CAM increases user confidence, making it more feasible to deploy in real-world applications such as port congestion management and ship tracking. Grad-CAM findings improve the transparency and robustness of the model, facilitating deployment in dynamic environments and enabling data-driven, interpretable decisions.

These results (see Figure 9 and Figure 10) clearly highlight the explainability of the presented ResNet50-based model.

By using these results and findings, the findings, advantages, limitations and future directions of the presented research are discussed below.

Findings:

- The best CNN among the tested CNNs is ResNet50 (see

Figure 7). Therefore, we have used it as a feature extractor in the proposed model).

- SVM is the best classifier among the tested classifiers. Thus, we tuned the parameters of this classifier.
- The length of the chosen feature vector is 111, and this feature vector contains both global and local features. According to the feature analysis, 50 of them are global, while 61 are local features.
- Figure 10 demonstrates that the presented model can focus on the ROI.

#### Advantages:

- We have introduced a novel solution for detecting ship parking space availability, employing computer vision as our proposed solution.
- This study unveils a new patch division model termed semi-overlapped patch division.
- A cognitive model has been introduced to attain high classification performance.
- The proposed DFE approach attained over 97% test accuracy on the collected image dataset for parking availability detection.
- In this research, we have provided both classification results and explainable outcomes.

#### Limitations:

- The dataset includes images from over 200 fisher shelters in Turkey and focuses on parking space availability. However, the dataset may have regional or seasonal biases that could affect generalizability.
- The dataset may also contain class imbalances, potentially skewing model performance in underrepresented scenarios.
- UAV-collected data could be influenced by external factors such as weather or lighting conditions, impacting model accuracy.

#### Future directions:

- We will collect a larger, more diverse dataset using UAVs and internet images. This will address biases related to region, season, and lighting for better generalization.
- We plan to extend the model to detect ship congestion, dock occupancy, and vessel types, beyond parking availability.
- Grad-CAM will be used to improve model explainability, providing better visual insights into predictions.
- Future work will explore EfficientNet and lightweight custom CNN models to boost performance and reduce computational costs.
- We will partner with maritime authorities and stakeholders to test the model in real-world environments like ports and fisher shelters.
- The model will be adapted for parking and occupancy detection in other areas like urban car parks, airports, and logistics hubs.
- A real-time monitoring system using UAV feeds and AI-based alerts will be developed for ongoing parking management.
- We plan to release part of our dataset publicly to promote research and collaboration in maritime computer vision.
- Future studies will combine radar, thermal imaging, and visual data to enhance detection accuracy in challenging weather or lighting conditions.

- We plan to introduce an explainable DFE model employing Grad-CAM.

## 6. CONCLUSIONS

The main purpose of this research is to solve the parking space availability problem for the maritime industry with a cost-effective and autonomous method. For this purpose, a novel image dataset was collected from approximately 200 fisher shelters using a UAV. In the second stage, the ResNet50-based DFE, which was designed cognitively and whose design process is detailed in the discussion section, was applied to the collected dataset. The introduced approach attained a 97.27% test accuracy and a 97.40% test geometric mean, effectively solving the parking space availability problem for marine vessels using computer vision. The reliability of the method has been demonstrated by presenting explainable results in the study. The obtained classification results, explainable outcomes, and findings clearly illustrate that the presented model is effective for parking space availability detection for marine vessels.

## AUTHOR CONTRIBUTIONS

Conceptualization, MG, VYC, AHB, RH, SD, TT; methodology, MG, VYC, AHB, RH, SD, TT; software, SD, TT; validation, MG, VYC, AHB, RH, SD, TT; formal analysis, MG, VYC, AHB, RH, SD, TT; investigation, MG, VYC, AHB; resources, MG; data curation, MG, VYC, AHB, RH, SD, TT; writing—original draft preparation, MG, VYC, AHB, SD, TT; writing—review and editing, MG, VYC, AHB, RH, SD, TT; visualization, MG, VYC, AHB, RH; supervision, TT; project administration, TT. All authors have read and agreed to the published version of the manuscript.

## REFERENCES

- [1] Jugović, A., Mezak, V., Lončar, S. (2006). Organization of maritime passenger ports. *Pomorski Zbornik*, 44(1): 93-104.
- [2] OECD (2011), *Environmental Impacts of International Shipping: The Role of Ports*, OECD Publishing. <http://doi.org/10.1787/9789264097339-en>
- [3] Khusainov, R., Azzi, D., Achumba, I.E., Bersch, S.D. (2013). Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations. *Sensors*, 13(10): 12852-12902. <https://doi.org/10.3390/s131012852>
- [4] Yuan, J., Zhang, L., Kim, C.S. (2023). Multimodal interaction of MU plant landscape design in marine urban based on computer vision technology. *Plants*, 12(7): 1431. <https://doi.org/10.3390/plants12071431>
- [5] Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, 2018(1): 7068349. <https://doi.org/10.1155/2018/7068349>
- [6] O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Walsh, J. (2020). Deep learning vs. traditional computer vision. In *Advances in Computer Vision: Proceedings of*

- the 2019 Computer Vision Conference (CVC), Las Vegas, USA, pp. 128-144. [https://doi.org/10.1007/978-3-030-17795-9\\_10](https://doi.org/10.1007/978-3-030-17795-9_10)
- [7] Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., Pietikäinen, M. (2020). Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, 128: 261-318. <https://doi.org/10.1007/s11263-019-01247-4>
- [8] Aslam, S., Michaelides, M.P., Herodotou, H. (2020). Internet of ships: A survey on architectures, emerging applications, and challenges. *IEEE Internet of Things Journal*, 7(10): 9714-9727. <https://doi.org/10.1109/JIOT.2020.2993411>
- [9] Fraga-Lamas, P., Fernández-Caramés, T.M., Suárez-Albela, M., Castedo, L., González-López, M. (2016). A review on internet of things for defense and public safety. *Sensors*, 16(10): 1644. <https://doi.org/10.3390/s16101644>
- [10] Bandyopadhyay, D., Sen, J. (2011). Internet of things: Applications and challenges in technology and standardization. *Wireless Personal Communications*, 58: 49-69. <https://doi.org/10.1007/s11277-011-0288-5>
- [11] Nizetić, S., Šolić, P., Gonzalez-De, D.L.D.I., Patrono, L. (2020). Internet of Things (IoT): Opportunities, issues and challenges towards a smart and sustainable future. *Journal of Cleaner Production*, 274: 122877. <https://doi.org/10.1016/j.jclepro.2020.122877>
- [12] Abdalla, G., Özyurt, F. (2021). Sentiment analysis of fast food companies with deep learning models. *The Computer Journal*, 64(3): 383-390. <https://doi.org/10.1093/comjnl/bxaa131>
- [13] Özyurt, F. (2021). Automatic detection of COVID-19 disease by using transfer learning of light weight deep learning model. *Traitement du Signal*, 38(1): 147-153. <https://doi.org/10.18280/ts.380115>
- [14] Tuncer, T., Aydemir, E., Ozyurt, F., Dogan, S. (2022). A deep feature warehouse and iterative MRMR based handwritten signature verification method. *Multimedia Tools and Applications*, 81: 3899-3913. <https://doi.org/10.1007/s11042-021-11726-x>
- [15] Kim, K., Hong, S., Choi, B., Kim, E. (2018). Probabilistic ship detection and classification using deep learning. *Applied Sciences*, 8(6): 936. <https://doi.org/10.3390/app8060936>
- [16] Escorcía-Gutierrez, J., Gamarra, M., Beleño, K., Soto, C., Mansour, R.F. (2022). Intelligent deep learning-enabled autonomous small ship detection and classification model. *Computers and Electrical Engineering*, 100: 107871. <https://doi.org/10.1016/j.compeleceng.2022.107871>
- [17] Ma, M., Chen, J., Liu, W., Yang, W. (2018). Ship classification and detection based on CNN using GF-3 SAR images. *Remote Sensing*, 10(12): 2043. <https://doi.org/10.3390/rs10122043>
- [18] Feng, Y., Diao, W., Sun, X., Yan, M., Gao, X. (2019). Towards automated ship detection and category recognition from high-resolution aerial images. *Remote Sensing*, 11(16): 1901. <https://doi.org/10.3390/rs11161901>
- [19] Huang, Z., Sui, B., Wen, J., Jiang, G. (2020). An intelligent ship image/video detection and classification method with improved regressive deep convolutional neural network. *Complexity*, 2020(1): 1520872. <https://doi.org/10.1155/2020/1520872>
- [20] Shi, Q., Li, W., Zhang, F., Hu, W., Sun, X., Gao, L. (2018). Deep CNN with multi-scale rotation invariance features for ship classification. *IEEE Access*, 6: 38656-38668. <https://doi.org/10.1109/ACCESS.2018.2853620>
- [21] Chang, Y.L., Anagaw, A., Chang, L., Wang, Y.C., Hsiao, C.Y., Lee, W.H. (2019). Ship detection based on YOLOv2 for SAR imagery. *Remote Sensing*, 11(7): 786. <https://doi.org/10.3390/rs11070786>
- [22] Zhang, X., Yan, M., Zhu, D., Guan, Y. (2022). Marine ship detection and classification based on YOLOv5 model. In *Journal of Physics: Conference Series*, Huzhou, China, p. 012025. <https://doi.org/10.1088/1742-6596/2181/1/012025>
- [23] Wang, R., You, Y., Zhang, Y., Zhou, W., Liu, J. (2018). Ship detection in foggy remote sensing image via scene classification R-CNN. In *2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, Guiyang, China, pp. 81-85. <https://doi.org/10.1109/ICNIDC.2018.8525532>
- [24] Leclerc, M., Tharmarasa, R., Florea, M.C., Boury-Brisset, A.C., Kirubarajan, T., Duclos-Hindíé, N. (2018). Ship classification using deep learning techniques for maritime target tracking. In *2018 21st International Conference on Information Fusion (FUSION)*, Cambridge, UK, pp. 737-744. <https://doi.org/10.23919/ICIF.2018.8455679>
- [25] Shao, Z., Wu, W., Wang, Z., Du, W., Li, C. (2018). Seaships: A large-scale precisely annotated dataset for ship detection. *IEEE Transactions on Multimedia*, 20(10): 2593-2604. <https://doi.org/10.1109/TMM.2018.2865686>
- [26] Zhang, S., Wu, R., Xu, K., Wang, J., Sun, W. (2019). R-CNN-based ship detection from high resolution remote sensing imagery. *Remote Sensing*, 11(6): 631. <https://doi.org/10.3390/rs11060631>
- [27] Pan, D., Wu, Y., Dai, W., Miao, T., Zhao, W., Gao, X., Sun, X. (2024). TAG-Net: Target attitude angle-guided network for ship detection and classification in SAR Images. *Remote Sensing*, 16(6): 944. <https://doi.org/10.3390/rs16060944>
- [28] Yu, C., Shin, Y. (2024). SAR ship detection based on improved YOLOv5 and BiFPN. *ICT Express*, 10(1): 28-33. <https://doi.org/10.1016/j.icte.2023.03.009>
- [29] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Houshby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- [30] Tuncer, T., Dogan, S., Özyurt, F., Belhaouari, S.B., Bensmail, H. (2020). Novel multi center and threshold ternary pattern based method for disease detection method using voice. *IEEE Access*, 8: 84532-84540. <https://doi.org/10.1109/ACCESS.2020.2992641>
- [31] Vapnik, V. (1998). *The support vector method of function estimation*. In *Nonlinear Modeling: Advanced Black-Box Techniques*, Boston, MA, USA, pp. 55-85. [https://doi.org/10.1007/978-1-4615-5703-6\\_3](https://doi.org/10.1007/978-1-4615-5703-6_3)
- [32] Frazier, P.I. (2018). *A tutorial on Bayesian optimization*. *arXiv preprint arXiv:1807.02811*. <https://doi.org/10.48550/arXiv.1807.02811>
- [33] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. *arXiv preprint arXiv:1602.07360*. <https://doi.org/10.48550/arXiv.1602.07360>
- [34] Tan, M., Le, Q. (2019). *Efficientnet: Rethinking model*

- scaling for convolutional neural networks. In International Conference on Machine Learning, PMLR, pp. 6105-6114.
- [35] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. (2017). Densely connected convolutional networks. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 4700-4708. <https://doi.org/10.1109/CVPR.2017.243>
- [36] Redmon, J., Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 7263-7271. <https://doi.org/10.1109/CVPR.2017.690>
- [37] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25: 1097-1105.
- [38] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>
- [39] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., Fergus, R. (2013). Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199. <https://doi.org/10.48550/arXiv.1312.6199>
- [40] Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S. (2022). A convnet for the 2020s. In 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Orleans, LA, USA, pp. 11976-11986. <https://doi.org/10.1109/CVPR52688.2022.01167>
- [41] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [42] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [43] Safavian, S.R., Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3): 660-674. <https://doi.org/10.1109/21.97458>
- [44] Zhang, Y., Zhou, X., Witt, R.M., Sabatini, B.L., Adjeroh, D., Wong, S.T. (2007). Dendritic spine detection using curvilinear structure detector and LDA classifier. *Neuroimage*, 36(2): 346-360. <https://doi.org/10.1016/j.neuroimage.2007.02.044>
- [45] Bhattacharyya, S., Khasnobish, A., Chatterjee, S., Konar, A., Tibarewala, D.N. (2010). Performance analysis of LDA, QDA and KNN algorithms in left-right limb movement classification from EEG data. In 2010 International Conference on Systems in Medicine and Biology, Kharagpur, India, pp. 126-131. <https://doi.org/10.1109/ICSMB.2010.5735358>
- [46] Ng, A., Jordan, M. (2001). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in Neural Information Processing Systems*, 14: 841-848.
- [47] Maillo, J., Ramírez, S., Triguero, I., Herrera, F. (2017). kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data. *Knowledge-Based Systems*, 117: 3-15. <https://doi.org/10.1016/j.knosys.2016.06.012>
- [48] Hothorn, T., Lausen, B. (2003). Bagging tree classifiers for laser scanning images: A data-and simulation-based strategy. *Artificial Intelligence in Medicine*, 27(1): 65-79. [https://doi.org/10.1016/S0933-3657\(02\)00085-4](https://doi.org/10.1016/S0933-3657(02)00085-4)
- [49] Biswas, S.K., Mia, M.M.A. (2015). Image reconstruction using multi layer perceptron (MLP) and support vector machine (SVM) classifier and study of classification accuracy. *International Journal of Scientific & Technology Research*, 4(2): 226-231.