




## Article

# Use of Neural Networks and Computer Vision for Spill and Waste Detection in Port Waters: An Application in the Port of Palma (MaJorca, Spain)

Mariano Morell <sup>1,2,\*</sup> , Pedro Portau <sup>2</sup>, Antoni Perelló <sup>2</sup>, Manuel Espino <sup>1</sup> , Manel Grifoll <sup>1</sup>  and Carlos Garau <sup>2</sup><sup>1</sup> LIM Laboratori d'Enginyeria Marítima, UPC-BarcelonaTech, 08034 Barcelona, Spain<sup>2</sup> Garau Ingenieros, SLU, 07012 Palma de Mallorca, Spain

\* Correspondence: mariano.nicolas.morell@upc.edu; Tel.: +34-657243912

**Featured Application:** Port Environmental Management systems; automated spill and waste detection in port waters.

**Abstract:** Water quality and pollution is the main environmental concern for ports and adjacent coastal waters. Therefore, the development of Port Environmental Management systems often relies on water pollution monitoring. Computer vision is a powerful and versatile tool for an exhaustive and systematic monitoring task. An investigation has been conducted at the Port of Palma de Mallorca (Spain) to assess the feasibility and evaluate the main opportunities and difficulties of the implementation of water pollution monitoring based on computer vision. Experiments on surface slicks and marine litter identification based on random image sets have been conducted. The reliability and development requirements of the method have been evaluated, concluding that computer vision is suitable for these monitoring tasks. Several computer vision techniques based on convolutional neural networks were assessed, finding that Image Classification is the most adequate for marine pollution monitoring tasks due to its high accuracy rates and low training requirements. Image set size for initial training and the possibility to improve accuracy through retraining with increased image sets were considered due to the difficulty in obtaining port spill images. Thus, we have found that progressive implementation can not only offer functional monitoring systems in a shorter time frame but also reduce the total development cost for a system with the same accuracy level.

**Keywords:** computer vision; marine litter; marine pollution; monitoring technologies; port water quality

**Citation:** Morell, M.; Portau, P.; Perelló, A.; Espino, M.; Grifoll, M.; Garau, C. Use of Neural Networks and Computer Vision for Spill and Waste Detection in Port Waters: An Application in the Port of Palma (MaJorca, Spain). *Appl. Sci.* **2023**, *13*, 80. <https://doi.org/10.3390/app13010080>

Academic Editors: Andrei Doncescu, Wen-June Wang and Chung-Hsun Sun

Received: 20 September 2022

Revised: 18 November 2022

Accepted: 21 November 2022

Published: 21 December 2022



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Ports and surrounding areas of the coast are zones in which a multitude of human activities are concentrated in a limited space with usually low water renewal rates. In consequence, ports and adjacent waters are very sensitive to pollution and accumulation of solid waste and their impact on the aquatic environment and, in turn, socioeconomic impact [1]. A relevant mechanism of water pollution in port areas is waste discharge and accumulation caused by non-continuous discharge events either intentionally or accidentally. This means solid or liquid pollutant waste is discharged into the water instantaneously or during a short period of time. These events constitute one of the most significant aspects to be considered in port and coastal environmental management; thus, economic and robust monitoring techniques are paramount to achieve adequate port water quality [2,3]. This issue is especially sensitive in city ports where there is a close relation between port operation and city activity, and where city waste and pollution can easily get into port water [4]. Currently, the most common approach for marine pollution monitoring in ports relies on conventional methods of collecting in situ water and waste samples for subsequent analysis in a laboratory. Such methods are time-consuming, expensive and do not provide a real-time picture of water quality in port waters. Thus, in practice they tend to be

implemented at minimum levels in order to comply with regulations, especially in ports with scarce resources. The consequences of this limited monitoring at environmental management level are in many cases significant [5,6]. Additionally, real-time or near real-time measurement and monitoring methods for marine pollutants and waste are necessary for managing their environmental impacts and understanding the processes governing their spatial distribution [7]. These techniques offer a complementary perspective on marine pollution to hydrodynamics-based environmental management techniques [8–10]. Thus, real-time pollution monitoring techniques can be linked with hydrodynamic models to obtain improved environmental management systems [11].

Given the nature and frequency of these discharges, management systems will usually consider the statistical parameters of the spatial and temporal distribution of the frequency of discharges instead of individual events. Therefore, these systems do not require very high levels of accuracy in monitoring as opposed to critical systems like biomedical applications, but rather enough to offer statistically significant distributions. Monitoring systems that offer 80% or higher accuracy are considered admissible based on the usual values required in these types of applications [12].

In this context, it is important to note that pollution discharge events in ports are, in most cases, visually perceivable. Consequently, it seems feasible to investigate the possibility of establishing automated monitoring systems for these discharges using cameras installed at strategic points in the port. Associated with automatic image analysis systems, computer vision techniques seem an excellent complement according to previous experiences in other fields for detecting and recognition [12–14]. Computer vision techniques have recently experienced a quick evolution, being implemented in a wide range of different applications with high efficiency and performance [15–17]. Deep learning on convolutional neural networks is proven to achieve very high performance on computer vision tasks [18]. In fact, remote sensing technology is proven to provide spatially synoptic and near real-time measurements that can be effectively used to detect and manage pollutants such as suspended sediments, oil and chemical spills, algal blooms and high suspended solids [7,19]. Additionally, recent contributions in waste and pollutant detection used Image Classification based on deep convolutional networks [20,21]. Such approaches have been successful at addressing pollution detection in large surface areas. In the case of port waters, satellite images cannot be used due to poor image resolution, and a monitoring system tailored to smaller scales has to be generated. Specifically, a computer vision system, supported by “in situ” mounted camera images would be a robust alternative for water pollution monitoring at ports. This system would allow continuous and low-cost monitoring of surface water pollution, addressing the limitations of traditional observational techniques. In addition, it would constitute a leap forward in the digitalization of ports through the practical application of artificial intelligence technology in coastal infrastructures at limited cost. It is important to note that the aim of this novel monitoring system is not only to give warnings for each discharge so that immediate action can be taken, especially in particularly relevant episodes that generate a significant risk for health or navigation, but also to obtain knowledge about the discharges that threaten the port waters where and when they happen or if they are related to specific operations. In consequence, computer vision, combined with traditional or Artificial Intelligence based analysis, may provide operational knowledge in specific port areas and facilities, thus allowing development of adequate environmental management strategies.

Computer vision techniques can be classified according to the problem considered [22]. There are several classifications and the set of problems considered has grown in recent times, but Image Classification is one of the most common applications and, in consequence, is very promising in port environmental management [23]. Image Classification involves assigning a label to an entire image; the labels (i.e., the categories in which images were classified) that should be considered in the context of port environmental management systems are three: clean water, pollution (spill) or floating waste (waste). One of the most important requirements for the implementation of computer vision systems is the

generation of a database of tagged images that can be used to train the algorithm. In this respect, it is important to take into account that gathering a significant database of images of spills can be time consuming, as they can only be achieved by installing cameras in the port to record images of eventual spills. Thus, images will be incorporated to the database progressively, and the question arises in terms of how many images—and image types—are required to train the algorithm to achieve an adequate level of confidence on the system. Specifically, it is important to determine whether it is preferable to train the algorithm with all images available even when the number of images in each category is different, or whether optimal results will be obtained only when there is an equal number of images in each category. In the first scenario, the least common class will be underrepresented, potentially affecting proper system performance, and in the second, the number of images to be gathered increases, and consequently so does the time required to achieve a working system.

In addition to image requirements, computer vision systems are evaluated according to specific performance metrics. Four of the most common metrics are Accuracy, Precision, Recall and F1-score [24,25]. However, the Accuracy metric does not provide a relevant metric for a port environmental management system because clean water images will be significantly more abundant than waste and spill ones; here Accuracy will provide mostly a measure of how many times clean water is correctly predicted. However, preliminary designs of computer vision systems for port environmental management suggest the need to generate correct alarms on spill and waste instances. Thus, an alternative metric needs to be put forward in order to compare trained algorithms with a set of images that are not evenly distributed between categories, as will be the case in the current application.

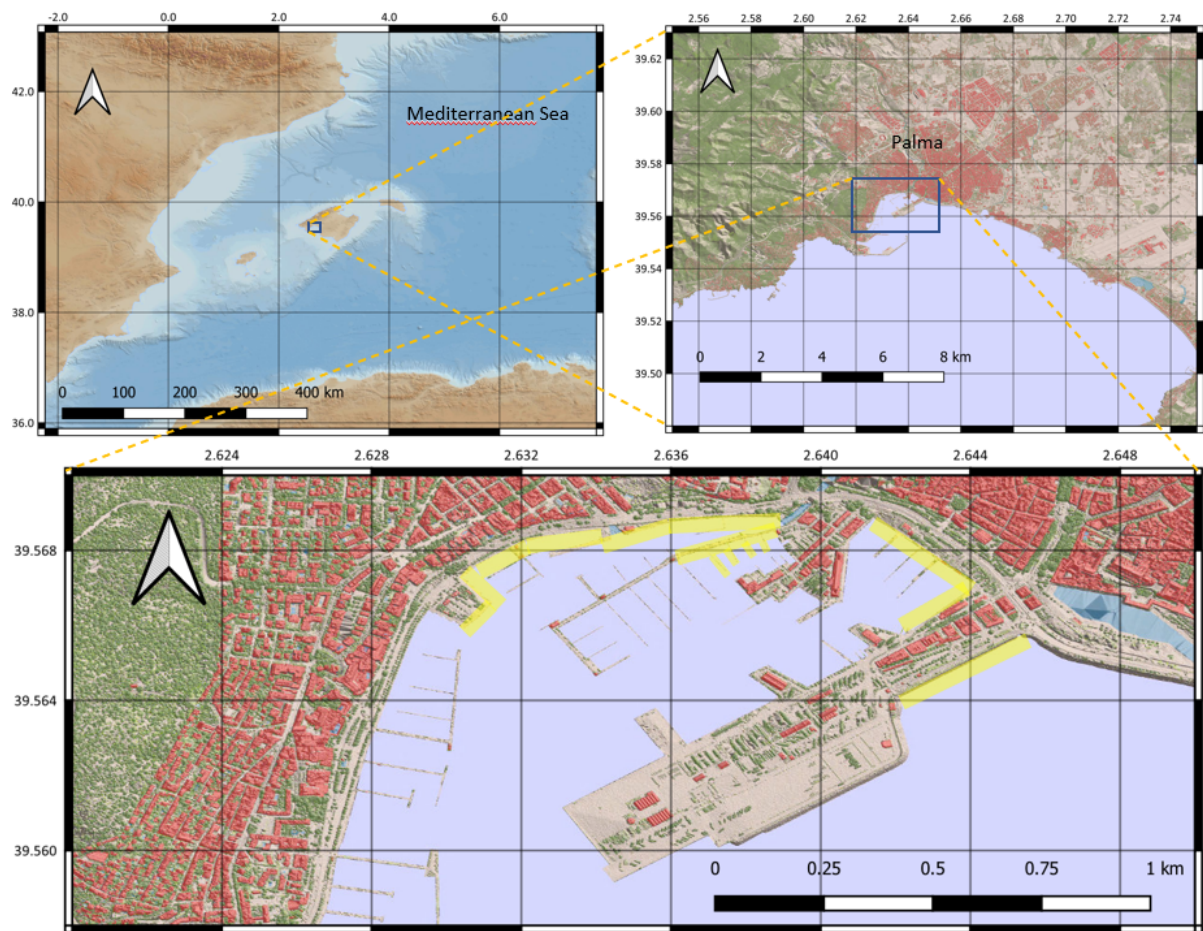
The present paper evaluates the results of a set of experiments on surface spills and floating marine waste identification based on random images as an initial stage of the development of a system for port water quality monitoring. After the methodological process (i.e., post-process) has been implemented, image sets have been obtained and analyzed to determine the amount and proportion of each image class that is required. In this sense, several computer vision techniques have been assessed, including Image Classification as the most promising one identified preliminarily. In order to evaluate the performance of the algorithm specifically for port environmental management applications, a novel performance index (the error index) has been proposed. The set of images has been conducted in the port of Palma de Majorca, which suffers important events of water quality degradation.

The paper is organized as follows. Section 2 introduces the study area, the computer vision technique used, the spill and waste classification, the system layout, the images used, the algorithm training and the statistical reliability of the algorithms. Section 3 shows the results of the training processes and a comparison for different amounts of data available in terms of image set sizes and distribution. Section 4 presents a discussion on the design criteria for the system set-up and its further development. Finally, in Section 5, the conclusions of the study are summarized.

## 2. Materials and Methods

### 2.1. Study Area

The port of Palma de Mallorca is located in the city of Palma, on the island of Majorca (Balearic Islands, Spain; see location in Figure 1) in the Western Mediterranean Sea, with approximate coordinates of: 2°38.4' E, 39°33.7' N. The management resides at the Port Authority of the Balearic Islands in a landlord governance model. From the impact on water quality degradation and environmental management, the port has the following characteristics: (i) Strong Port–City relation. (ii) Development of several different port activities (i.e., recreational boating, transport of passengers and goods, fishing, repair and maintenance of boats and restoration and services on land). (iii) Sporadic discharges of rainwater through four gullies and several collectors of stormwater drainage networks, in some cases with risk of discharge of mixed rainwater and wastewater.



**Figure 1.** Study area location. Zones where images were obtained are highlighted in yellow in the lower panel.

## 2.2. Computer Vision Technique and Application to Pollution in Port Waters

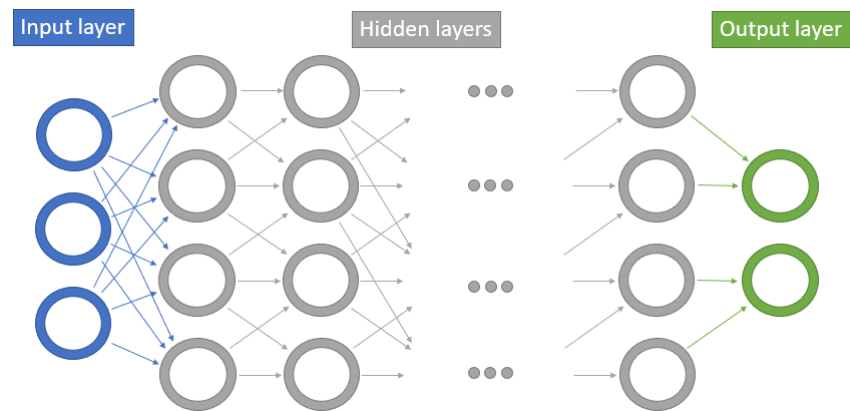
Computer vision is a field where applications are developed using convolutional neural networks that are trained using deep learning techniques. Specifically, it can be defined as a set of techniques to automatically obtain descriptions or significant parameters from the images of physical objects; these descriptions can be useful for decision making. This is the case of the current investigation included in the field of marine waste and litter detection. Due to the numerous potential applications of Computer vision, it has experienced an important development in the recent years.

An artificial neural network is a collection of connected nodes which loosely model the neurons in a biological brain [26]. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. When an artificial neuron receives a signal, it processes it and, as a result, sends outputs (real numbers) to neurons connected to it. In turn, the output signals of each neuron are computed by some non-linear function of the sum of its inputs. Typically, neurons are aggregated into layers; different layers may perform different transformations on their inputs coming from the one before. Signals travel from the first layer (or input layer) to the last one (or output layer). Figure 2 depicts schematically how neurons in different layers interact to provide meaningful results.

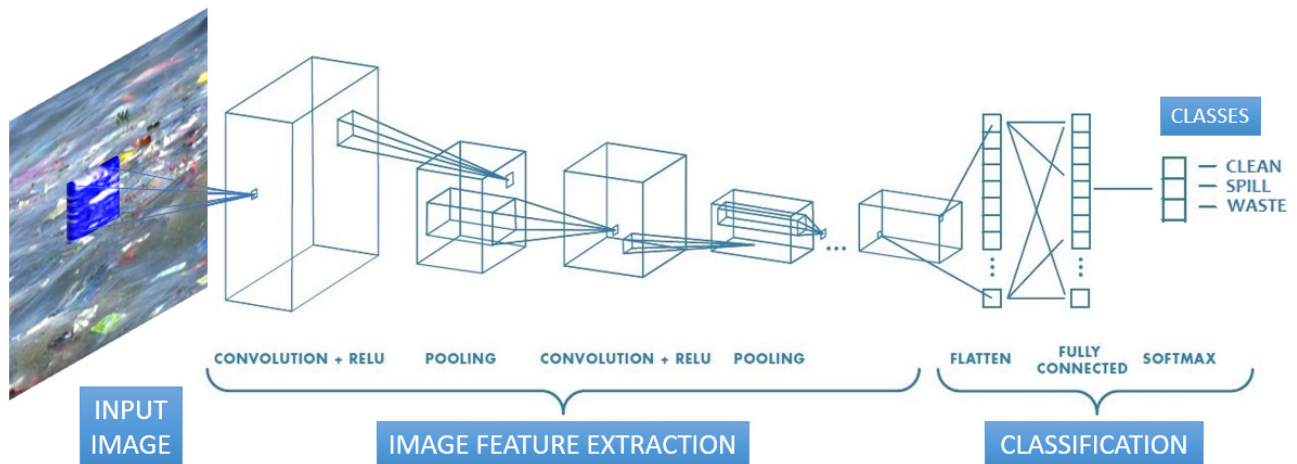
A Convolutional Neural Network (CNN) is a type of artificial neural network most commonly applied to analyze visual imagery because they are shift invariant (or space invariant), meaning that the position of a feature in an image is not important. This is due to the CNN having a shared-weight architecture of the convolution kernels or filters that slide along input features and provide translation equivariant responses known as feature maps.



Figure 3 shows how the CNN architecture works towards generating relevant information from an input image.



**Figure 2.** Neurons, layers and signal transmissions.



**Figure 3.** Basic CNN architecture (modified from [26]).

The most important computer vision techniques are Image Classification, Object detection, Object tracking, Semantic segmentation and Semantic instance segmentation. Although all these techniques have a potential application in port water quality monitoring, the most appropriate technique according to the input–output information desired is Image Classification. This is due to the fact that the amount of information needed to train the system is lower compared to other techniques, and it allows the classification of images into simple classes that can be used to build temporal and spatial distributions of pollution events [27].

Thus, the aim of this investigation is to evaluate the efficiency of training Image Classification algorithms that, when taking as input the images of port water provided by a camera monitoring system and operating in real-time, provide as output the class to which each image belongs with the highest probability, according to a classification that is relevant for proper environmental management of port water.

### 2.3. Computer Vision Classes Considered

The selection of the clean, spill and waste classes has been carried out after careful consideration of the nature of pollution in ports as well as the level of detail that is useful in port environmental management activities. Specifically, spills in ports have four main origins (although in specific ports or terminals there may be others): users on land, users on boats, discharges of mixed drainage networks and port operations. Considering their

physical and chemical nature, there is an enormous variety in waste and contaminants that can reach port waters, including suspended matter, hydrocarbons or eutrophication (not a spill in itself, but a consequence of a nutrient spill) (see Figure 4). Identification of both the origins and chemical nature of spills could be pursued, but the applicability of such information is very limited; all these contamination events are managed in a similar fashion and thus their identification would not provide any relevant input in port environmental management. In contrast, a type of pollution that follows a different type of action from a port environmental management perspective is floating waste (see Figure 4). Consequently, for the computer vision system designed, two categories of pollution have been considered, namely spill and waste. The spill class (class 1 in this study) refers to liquids mixing and/or diluting in the water, or to clouds of suspended solids. The waste class refers to large individual solids floating on the water or near the surface (class 2 in this study). Finally, clean water has been labelled as class 0. These three classes provide sufficient information for a port environmental management system to take relevant decisions on time and cost.



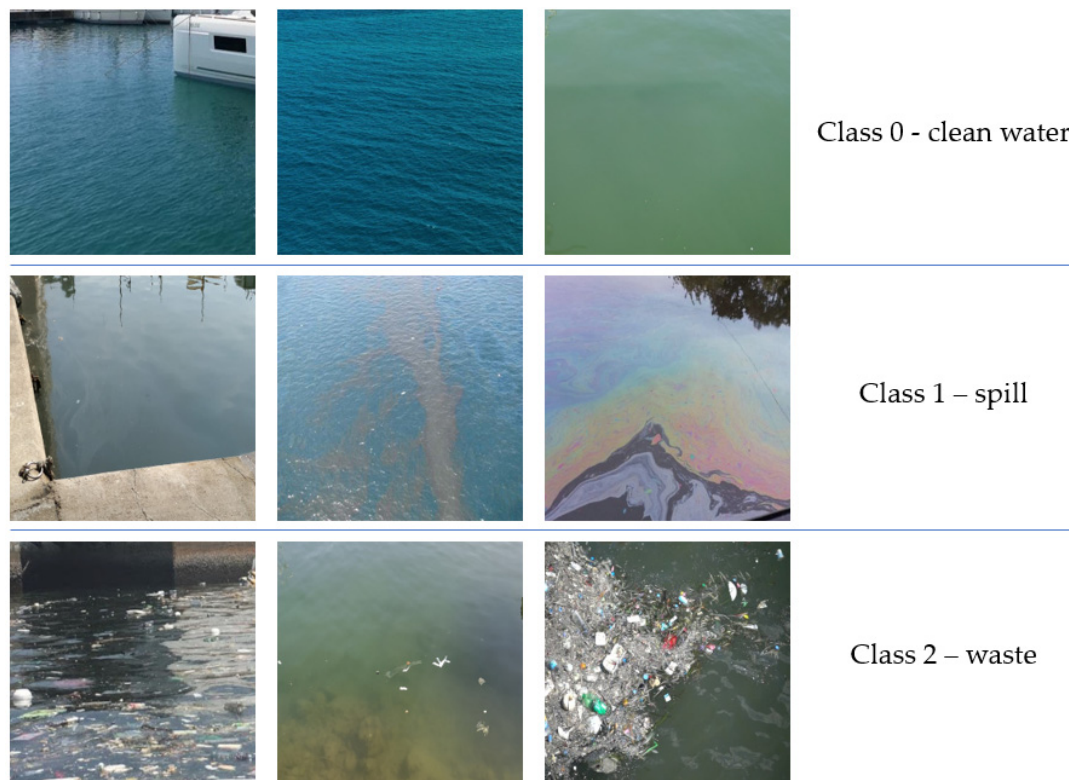
**Figure 4.** Right image: Spill class example. Center image: Waste class example. Left image: Mixed Spill and Waste example.

The Image Classification technique does not consider the possibility of one image belonging to two or more classes; it simply returns the most likely class. This may constitute a limitation of the method since spill and waste could theoretically appear simultaneously in an image. To overcome this limitation, an additional class should be defined including images with the presence of both (see Figure 4). However, this situation is very infrequent in ports, and, in fact, it did not occur in any of the images obtained in this investigation. The most common cases in which we could theoretically find spill and waste together are: (i) pollution originating from two or more independent incidents ending up in an accumulation zone due to the hydrodynamic characteristics of the port; or (ii) mixed pollution released by rivers or collector systems that discharge into the port. In the context of the system proposed in this work, the first case is irrelevant because the main objective is to monitor discharge episodes rather than the persistence of discharges within the port. The second case is limited to specific areas and its processing constitutes a particularity that is to be faced in future research. Therefore, although this limitation exists, it does not seem to be an import limitation at this stage due to the infrequency of the combined (i.e., spill and waste) event. The segregated monitoring system proposed represents the reality of most existing ports and thus is easily scalable to other infrastructures.

#### 2.4. Dataset Used

The dataset used in the current study consists of images obtained through manual sampling in several different locations in the Port of Palma. About 3400 images were obtained, of which only 1379 were actually used; 660 were selected as instances of clean water class, 389 of spill class and 330 of waste class. Discarded images were too similar to other images that were used or constituted excess clean water class and spill class images. The number of images obtained in spill and waste classes were the main limitation as actual pollution events are required to happen in the port during the fieldwork visit in order to obtain them.

In this study, different amounts of spill/waste and water images have been used, as detailed in Section 2.5, in order to investigate the practical applicability of the developed system. The images were gathered using different digital cameras, in 4:3 format and different image resolutions (1 Megapixel and higher). Nevertheless, when using the images for the training and validation of algorithms, they were transformed into square pictures and their resolution was reduced (see Section 2.5). Figure 5 shows three images from each class with square shape and reduced resolution.



**Figure 5.** Example images for each class.

### 2.5. Experiments Description: Algorithm Training and Validation

In order to evaluate the feasibility of implementing a computer vision water quality monitoring system in ports, three experiments have been carried out in the present study using a CNN type system. The experiments intend to evaluate the feasibility of a computer vision system in port environmental management and the performance impact of the results on image set size and distribution. The main characteristics of each experiment are shown in Table 1, including the research objectives.

**Table 1.** Summary of computer vision experiments in the current study.

Experiment Number	Research Objective	Number of Images (Spill/Waste/Clean Water)	Image Resolution (Pixels)
Experiment 1	Screening of computer vision system overall performance and feasibility for port environmental management	389/330/660	300 × 300
Experiment 2	Investigating the performance impact of image set distribution	1320 images in different proportions	256 × 256
Experiment 3	Investigating the performance impact of image set size	Different numbers in equal proportions	256 × 256

Keras open-source software library for Python (version: 2.4.3) on Tensorflow Google developed open-source software library (version: 2.3.0) backend framework based on the Anaconda3 platform was used in these experiments. Python 3.8.10 programming language was used for training and validation process programming. The computer used was equipped with an Intel Core i7-6700HQ CPU with 16 GB RAM and a NVIDIA GeForce GTX 960M graphics card. The computer operating system was the 64-bit Windows 10 home edition. In the three experiments, a neural network InceptionV3, with “imagenet” weights and a 3-channel resolution was deployed. InceptionV3 was chosen between Keras available models, after discarding models designed for mobile devices considering the compromise between accuracy and speed according to Keras documentation [28] and CNN research [29,30]. An additional GlobalAveragePooling2D layer was added with 1024 additional neurons with ReLU activation (0.2 dropout), as well as another layer with 3 neurons with softmax activation. The latter layer is the one bearing the spill/waste/clean water class information. In order to feed the neural network, two image generators were used. For the training images, a series of transformations were applied (rotation, horizontal and vertical shifts, crop, zoom and horizontal reflex) including a standard normalization. In addition, data augmentation techniques were used on the image set [31]. For image validation purposes, only normalization was applied. Data ingestion was carried out in batches of 8 images. The training set images consisted of 80% of the set and the remaining 20% were used for validation purposes. Firstly, a training of additional layers was conducted and subsequently a fine-tuning was simultaneously carried out of both final inception blocks and additional layers. The cost function used was CategoricalCrossentropy (logit) and Adam was deployed as the optimization algorithm (learning rate of 0.001 and 0.00001 was each of the training phases described previously).

For Experiment 2, 14 algorithms were trained, two for each image set distribution tested. The distributions of images considered in these experiments are the ones shown in Table 2.

**Table 2.** Distribution of images considered in Experiment 2.

Image Ratio	Number of Images of Class 0	Number of Images of Class 1	Number of Images of Class 2
1/1	330	330	330
1/2	660	330	330
2/5	660	264	264
3/1	660	196	196
1/4	660	165	165
1/8	656	82	82
1/16	656	41	41

For Experiment 3, 82 trainings based on image sets formed randomly of different sizes (ranging between 18 and 990 total images). Here, one third of the total number of images corresponds to each class.

In experiments 2 and 3, each algorithm training was started from the initial model and not from the previously trained algorithm in order to prevent the propagation of errors or beneficial traits from one algorithm to the next.

## 2.6. Algorithm Performance Evaluation

Some of the metrics used in this study are the ones commonly reported in the literature and applied investigations when evaluating the performance of computer vision systems (8). These are the following:

*Accuracy:* Commonly defined as the ratio of true positives and true negatives to all positive and negative observations. That is, how often we can expect the computer vision system to correctly predict an outcome out of the total number of times it made predictions.



Mathematically, it is formulated as the ratio of the sum of true positives and true negatives out of all the predictions, namely:

$$Accuracy = \frac{TP + TN}{(TP + FN + TN + FP)} \quad (1)$$

where  $TP$  = true positives;  $TN$  = true negatives;  $FN$  = false negatives; and  $FP$  = false positives.

*Precision*: It represents the proportion of labels that were correctly predicted to be positive. That is, it is a performance metric that is most useful when trying to control false positives. As well as for *Accuracy*, *Precision* is also affected by class distribution; if there are more images for a class that does not happen frequently, precision becomes lower.

Mathematically, it is formulated as the ratio of true positive to the sum of true positives and false positives, namely:

$$Precision = \frac{TP}{(FP + TP)} \quad (2)$$

*Recall*: It represents the system's capacity to correctly predict the positives from the set of actual positives. *Recall* is most useful when identifying positives as critical. Mathematically, it represents the ratio of true positive to the sum of true positives and false negatives.

$$Recall = \frac{TP}{(FN + TP)} \quad (3)$$

*F1 score*: It is obtained as a harmonic mean of the *Precision* and *Recall* scores, giving each of them an equal weight. It is often used as a single value that provides high-level information about the model's output quality and *Precision/Recall* balance.

Mathematically, it is formulated as a harmonic mean of the *Precision* and *Recall* scores.

$$F1\ Score = \frac{2 * Precision * Recall}{(Precision + Recall)} \quad (4)$$

In the case of experiment 1, where the objective is to validate the algorithm generated for its application in port environmental management, the prior metrics are relevant and sufficient. However, in experiment 2, as well as in realistic system application, we would need an additional index that evaluates the performance of the system as an alternative to the common *Accuracy* metric. This is due to the fact that the *Accuracy* metric is not the most reliable in computer vision models trained on datasets where one event (in this case clean water) is much more frequent than the rest of the events (in this case spill or waste). In this case, *Accuracy* will mostly determine that clean water is detected correctly most of the time but will not provide decisive information on the spill and waste detection performance. As the latter are the actual events (alarms) to be detected by a computer vision system applied in a port setting, *Accuracy* is not a parameter that becomes useful in the present study or in real-life applications of the system. *Precision*, *Recall* and *F1-Score* indexes are also not suitable for experiment 2 because they are class specific and for comparison purposes an all-class synthetic index is needed. Consequently, a novel index has been defined for the purpose of this application (as well as others that might face similar issues as the one presented): the *Error index*. This index is defined as the ratio of the sum of errors made in providing warnings (either false alarms or alarms that are incorrectly not provided) to the sum of total alarms provided by the system. Adapting for the current application with three classes (i.e., 1, 2 and 3), the *Error index* is defined as:

$$Error\ index = \frac{(FP0 + FP1 + FP2)}{(TP1 + FP1 + TP2 + FP2)} \quad (5)$$

where  $TP_i$  = true positives for class,  $i$ ;  $TN_i$  = true negatives for class,  $i$ ;  $FN_i$  = false negatives for class  $i$  and  $FP_i$  = false positives.

The definition corresponds to a parameter that is more meaningful than *Accuracy* for port water quality monitoring applications, as it eliminates the issue of the unequal distribution of images during the application of the system. However, two limitations have been detected: (i) *Error index* is not a normalized parameter and (ii) it overestimates the errors made overall by the system because it eliminates a set of prediction successes. Nevertheless, it is a conservative and meaningful index useful for port managers because of its comprehensiveness.

### 3. Results

#### 3.1. Experiment 1

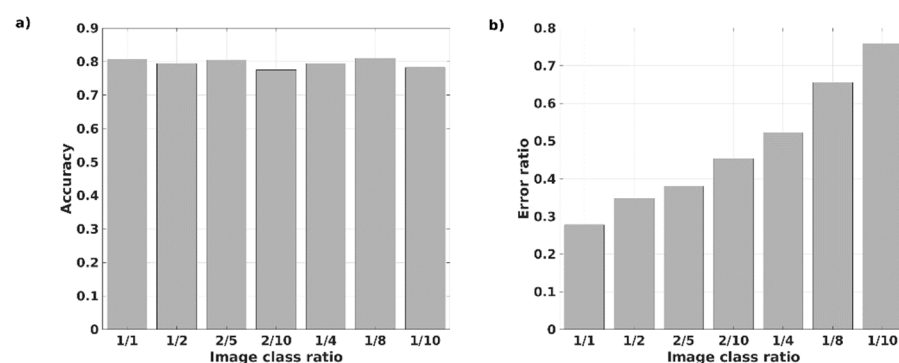
The Image Classification algorithm has been trained and validated and an *Accuracy* of 0.91 has been obtained with an image evaluation time of about one second. Table 3 presents the performance metrics for the identification of each of the classes. In general, adequate performance metrics with the Image Classification technique have been achieved, proving that the system is promising. As a shortcoming to be addressed with the validation dataset, a proportion (>10%) of the cases classified as clean water are really contaminated water. This aspect will be improved in the upcoming experiments, where image set distribution and image set size are investigated in order to generate a more applied monitoring technique.

**Table 3.** Performance results for Experiment 1.

Class	Precision	Recall	F1-Score
0—clean water	0.89	0.94	0.91
1—spill	0.95	0.86	0.90
2—waste	0.93	0.93	0.93

#### 3.2. Experiment 2: Impact of Image Set Distribution

Figure 6a shows the *Accuracy* (y-axis) versus the image class ratio (x-axis) in the different simulations carried out. In this figure, *Accuracy* remains relatively stable with changing proportions of images in the training and validation dataset. On the other hand, the *Error index*, formulated in the present study to be able to capture how adequate the system is in correctly detecting contamination alarms, shows that the performance of the system decreases significantly with an increasing disproportion of image classes (see Figure 6b).

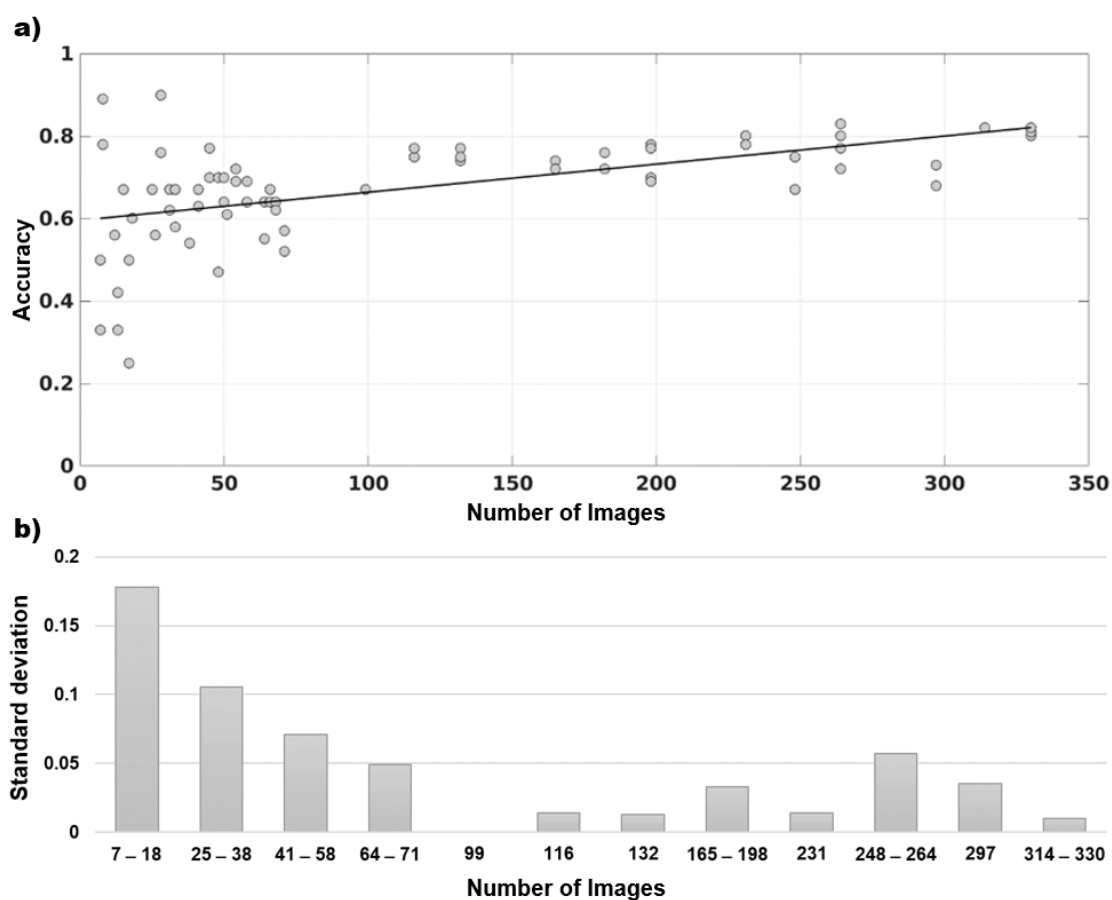


**Figure 6.** (a) Accuracy vs. Image class ratio. (b) Error ratio vs. Image class ratio.

#### 3.3. Experiment 3: Impact of Image Set Size

Figure 7a shows *Accuracy* in different simulations where image set sizes vary. *Accuracy* has been used in this experiment since it is a normalized parameter and thus easier to interpret graphically, but similar conclusions have been reached with the *Error index* in this case. Figure 7b shows how simulations with image set sizes lower than 297 images (99 images per class) generate a significant dispersion in performance. Dispersion shows a decreasing trend up to 99 images per class and from that point on there is no clear trend,

remaining at moderate values. The number of 99 images per class is also the closest among those used to the benchmark of 100 images per class usually recommended for training Image Classification algorithms [27]. With datasets that have a number of images over this amount, Accuracy becomes stable and increases in a linear manner with increasing images provided to train and validate the algorithm. When carrying out a regression in datasets with over 99 images per class (297 total images), both linear and quadratics fits have been considered. Finally, a linear regression, shown in Figure 7a, was selected because the quadratic fit is only marginally better than the linear one and because the linear fit showed significantly more robustness. Robustness was here evaluated as the change in fit parameters when random datapoints are removed from the set of results. Thus, in the range of image set sizes considered in the study, also the relevant range for the application at hand, the *Accuracy* presents a linear tendency with increasing image set size after a certain number of images have been achieved.



**Figure 7.** (a) Dots: Accuracy vs. Number of Images per class. Line: Regression. (b) Standard deviation of accuracy measures vs. Number of Images per class (when two numbers are displayed it means a number range).

#### 4. Discussion

##### 4.1. Results Discussion System Set-Up

Results showed in Section 3 demonstrate that Image Classification is adequate for marine pollution monitoring tasks due to its high *Accuracy* rates and low training requirements (Table 2). The system obtained a 91% accuracy rating, which can be considered a sufficient value for the requirements of a discharge management system in which the use of pollution event data is statistical and, if action is required, it will entail necessary direct validation. The time required by the trained algorithm to classify an image is approximately one second, which is compatible with the needs of real-time monitoring. The best performance was proven to be achieved when image set sizes for all classes are similar (Figure 5),

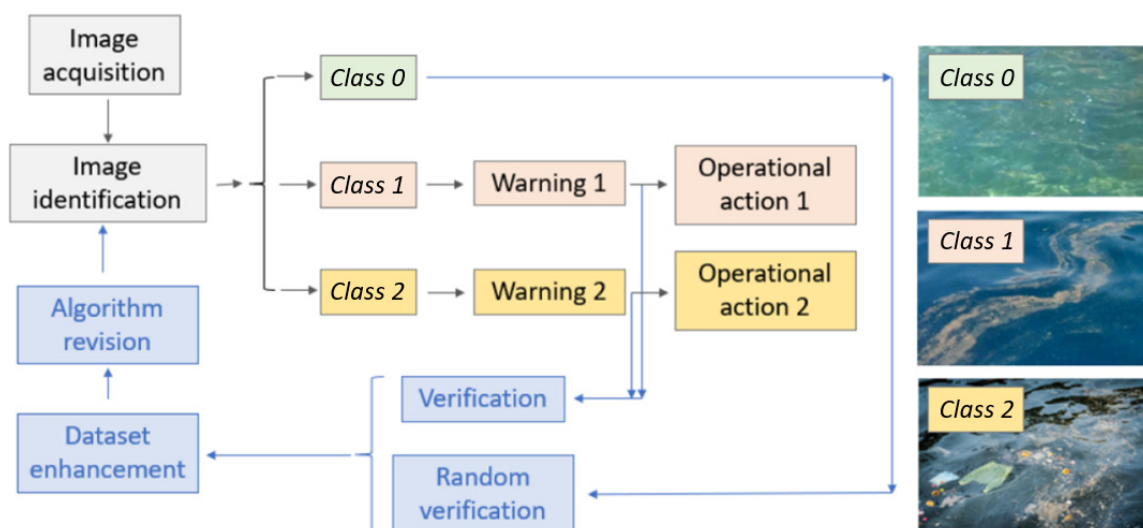
providing the first insight into the requirements for adequate system implementation. In practice, spill images are difficult to obtain in great numbers and commonly clean water images will be the most dominant class. Thus, in order to achieve an algorithm that holds optimal performance, spill and waste images have to be obtained to achieve a total number between the three classes which is higher than 297 (Figure 6). In this sense, the results of experiments 2 and 3 are consistent with other documented application cases based on computer vision (e.g., [27]).

The most appropriate performance metric to evaluate these systems in operation is the proposed *Error index*, since in operating conditions it is foreseeable to find a much higher number of class 0 images than those of classes 1 and 2.

Our work suggests that the most appropriate way for the monitoring system to be implemented is through progressive implementation. In this sense, datasets would be ever increasing when additional spill and waste images were attained. At these points, the algorithm would be retrained with new datasets in order to generate higher *Accuracy* and lower *Error Index* rates, improving the information provided by the system in a gradual manner. After a total image data set of 297 (between the three classes considered) has been reached, retraining would be less frequent due to the fact that performance only increases gradually after that point. In this type of progressive implementation, functional monitoring systems would be provided to port decision makers in a shorter time frame while also reducing the total development cost for a system with the same accuracy level.

Considering that the training time of the algorithm is in the order of minutes and that retraining of the algorithm will be carried out very infrequently (due to the difficulty of obtaining pollution images), it is preferable that each training is carried out from the initial model and not from the previously used algorithm in order to prevent error propagation.

Figure 8 depicts how the proposed implementation would be carried out in practice. In addition to the image acquisition and identification of the three classes (with alarms generated when spill or waste were detected), with increasing waste and spill images a verification and dataset enhancement step would be prompted in the system. With the enhanced dataset, the algorithm would be revised in order to achieve gradually better *Accuracy* and *Error index* performance.



**Figure 8.** Progressive implementation of port water monitoring systems. The operational action refers to eventual anti-pollution measures planned by the port authority.

#### 4.2. Future Applications

The most critical part of an applied computer vision system for the detection of pollution in ports is the availability of images with spills or floating waste. Therefore, future implementations would include tools and developments that improve upon the



speed and cost of image obtention. Specifically, when a spill or floating waste image is detected, considering information on the duration of the pollution event would be critical to accumulate as much images as possible from the same event. This could be achieved either manually when a spill or waste event system was detected by the computer vision system, or a hybrid hydrodynamic computer vision system could be generated. In this sense, a hydrodynamic model in the framework of operational oceanography systems [32] would automatically provide information on the duration of the pollution event, and the dataset would be increased also in an automatic manner. However, hybrid systems can be complex; thus, proper investigation of the actual practicality of developing such a system should be further investigated. Future implementation in operational mode (with a large amount of images acquired) may entail an increase in the number of classes considered either by subdivision of some of the current classes or even by incorporation of a new class to codify the simultaneous presence of spill and waste as explained in Section 2.3. Additionally, pre-filtering and preparation of images could provide better image sets that would increase the performance without relying on algorithm retraining. This would include—for instance—filtering to avoid classification interference by passing boats and port infrastructures and detecting of waste and discharge events located far from the camera location with less loss of image resolution. In addition, pre-filtering may avoid or reduce the effect of sunlight reflections, and other transformation of the images may yield better detail of the contamination event and reduce interference of other less relevant details also contained in the images.

## 5. Conclusions

Experiments on port water quality identification based on random image sets have been conducted. The reliability and development requirements of the method have been evaluated showing that computer vision tools are suitable for these monitoring tasks. Several computer vision techniques were considered for use in real-time marine pollution monitoring, with the decision that Image Classification was the most adequate for such tasks due to its high accuracy rates and low training requirements. These requirements and the possibility to improve accuracy through retraining with increased image sets were considered due to the difficulty in obtaining port spill images, finding that progressive implementation can not only offer functional monitoring systems in a shorter time frame, but also reduce the total development cost for a system with the same accuracy level. A novel performance metric for the case of computer vision systems in the port environmental management application was put forward and tested, providing meaningful conclusions.

Future lines of research include the development of additional methods that improve the time taken to obtain spill and waste images, ultimately increasing the applicability and speed in which it provides meaningful information to port decision makers. In addition, future works include the consideration of a new class for combined spill and waste for those ports that receive mixed (i.e., waste and spills) discharges from waterways or water collection infrastructure. In addition, image preparation and pre-filtering could also yield algorithms with higher performance metrics and help overcome limitations for monitoring systems where camera location is not optimal or where reflected sunlight makes images hard to classify.

**Author Contributions:** Conceptualization, M.E., M.G. and C.G.; Data curation, M.M. and A.P.; Formal analysis, M.M. and P.P.; Funding acquisition, C.G.; Investigation, M.M. and A.P.; Methodology, P.P., M.E. and M.G.; Project administration, M.M.; Software, A.P.; Supervision, M.E., M.G. and C.G.; Writing—original draft, M.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by ECOBAYS project (MCIN/AEI/10.13039/501100011033) funded from the Agencia Española de Investigación (Spanish Research Agency).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Project datasets are not publicly available as images used are property of Garau Ingenieros, SLU.

**Acknowledgments:** The authors are grateful for the help provided by Ports de Balears for providing the site; to Vicepresidència i Conselleria d’Innovació, Recerca i Turisme del Govern de les Illes Balears through Direcció General d’Innovació i Recerca for support to SPILLSURVEY project; and to Secretaria d’Universitats i Recerca del Dpt. d’Economia i Coneixement de la Generalitat de Catalunya (ref. 2014SGR1253) for supporting the research group.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ng, A.K.; Song, S. The environmental impacts of pollutants generated by routine shipping operations on ports. *Ocean Coast. Manag.* **2010**, *53*, 301–311. [\[CrossRef\]](#)
- Puig, M.; Azarkamand, S.; Wooldridge, C.; Selén, V.; Darbra, R. Insights on the environmental management system of the European port sector. *Sci. Total Environ.* **2021**, *806*, 150550. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hossain, T.; Adams, M.; Walker, T.R. Role of sustainability in global seaports. *Ocean Coast. Manag.* **2020**, *202*, 105435. [\[CrossRef\]](#)
- Li, Y.; Zhang, X.; Lin, K.; Huang, Q. The Analysis of a Simulation of a Port-City Green Cooperative Development, Based on System Dynamics: A Case Study of Shanghai Port, China. *Sustainability* **2019**, *11*, 5948. [\[CrossRef\]](#)
- Wooldridge, C.F.; McMullen, C.; Howe, V. Environmental management of ports and harbours—Implementation of policy through scientific monitoring. *Mar. Policy* **1999**, *23*, 413–425. [\[CrossRef\]](#)
- Puig, M.; Wooldridge, C.; Michail, A.; Darbra, R.M. Current status and trends of the environmental performance in European ports. *Environ. Sci. Policy* **2015**, *48*, 57–66. [\[CrossRef\]](#)
- Hafeez, S.; Wong, M.S.; Abbas, S.; Kwok, C.Y.T.; Nichol, J.; Lee, K.H.; Tang, D.; Pun, L. Detection and Monitoring of Marine Pollution Using Remote Sensing Technologies. In *Monitoring of Marine Pollution*; IntechOpen: London, UK, 2019. [\[CrossRef\]](#)
- Grifoll, M.; Jordà, G.; Espino, M.; Romo, J.; García-Sotillo, M. A management system for accidental water pollution risk in a harbour: The Barcelona case study. *J. Mar. Syst.* **2011**, *88*, 60–73. [\[CrossRef\]](#)
- Mali, M.; Malcangio, D.; Dell’Anna, M.M.; Damiani, L.; Mastroianni, P. Influence of hydrodynamic features in the transport and fate of hazard contaminants within touristic ports: Case study: Torre a Mare (Italy). *Heliyon* **2018**, *4*, e00494. [\[CrossRef\]](#)
- Villalonga, M.M.; Infantes, M.E.; Colls, M.G.; Ridge, M.M. Environmental Management System for the Analysis of Oil Spill Risk Using Probabilistic Simulations. Application at Tarragona Monobuoy. *J. Mar. Sci. Eng.* **2020**, *8*, 277. [\[CrossRef\]](#)
- La Loggia, G.; Capodici, F.; Ciraolo, G.; Drago, A.; Maltese, A. Monitoring Mediterranean marine pollution using remote sensing and hydrodynamic modelling. *Proc. SPIE Int. Soc. Opt. Eng.* **2011**, *8174*, 817416. [\[CrossRef\]](#)
- Arribas, J.I.; Sánchez-Ferrero, G.V.; Ruiz-Ruiz, G.; Gómez-Gil, J. Leaf classification in sunflower crops by computer vision and neural networks. *Comput. Electron. Agric.* **2011**, *78*, 9–18. [\[CrossRef\]](#)
- Eskandari, R.; Mahdianpari, M.; Mohammadimanesh, F.; Salehi, B.; Brisco, B.; Homayouni, S. Meta-Analysis of Unmanned Aerial Vehicle (UAV) Imagery for Agro-Environmental Monitoring Using Machine Learning and Statistical Models. *Remote. Sens.* **2020**, *12*, 3511. [\[CrossRef\]](#)
- Storbeck, F.; Daan, B. Fish species recognition using computer vision and a neural network. *Fish. Res.* **2001**, *51*, 11–15. [\[CrossRef\]](#)
- Chen, H.-C.; Li, Z.-T. Automated Ground Truth Generation for Learning-Based Crack Detection on Concrete Surfaces. *Appl. Sci.* **2021**, *11*, 10966. [\[CrossRef\]](#)
- Dong, R.; Na, X. Quantitative Retrieval of Soil Salinity Using Landsat 8 OLI Imagery. *Appl. Sci.* **2021**, *11*, 11145. [\[CrossRef\]](#)
- Ngeljaratan, L.; Moustafa, M.A. Underexposed Vision-Based Sensors’ Image Enhancement for Feature Identification in Close-Range Photogrammetry and Structural Health Monitoring. *Appl. Sci.* **2021**, *11*, 11086. [\[CrossRef\]](#)
- Leonard, J.K. Image Classification and Object Detection Algorithm Based on Convolutional Neural Network. *Sci. Insights* **2019**, *31*, 85–100. [\[CrossRef\]](#)
- Ciappa, A.C. Marine Litter Detection by Sentinel-2: A Case Study in North Adriatic (Summer 2020). *Remote Sens.* **2022**, *14*, 2409. [\[CrossRef\]](#)
- Panwar, H.; Gupta, P.; Siddiqui, M.K.; Morales-Menendez, R.; Bhardwaj, P.; Sharma, S.; Sarker, I.H. AquaVision: Automating the detection of waste in water bodies using deep transfer learning. *Case Stud. Chem. Environ. Eng.* **2020**, *2*, 100026. [\[CrossRef\]](#)
- Jiao, Z.; Jia, C.G.; Cai, C.Y. A new approach to oil spill detection that combines deep learning with unmanned aerial vehicles. *Comput. Ind. Eng.* **2018**, *135*, 1300–1311. [\[CrossRef\]](#)
- Khan, A.I.; Al-Habsi, S. Machine Learning in Computer Vision. *Procedia Comput. Sci.* **2020**, *167*, 1444–1451. [\[CrossRef\]](#)
- Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote. Sens.* **2007**, *28*, 823–870. [\[CrossRef\]](#)
- Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In *Advances in Information Retrieval, Lecture Notes in Computer Science*; Losada, D.E., Fernández-Luna, J.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 345–359. [\[CrossRef\]](#)
- Powers, D.M.W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv* **2020**. [\[CrossRef\]](#)

26. Introduction to Deep Learning: What Are Convolutional Neural Networks? Video [WWW Document], n.d. URL. Available online: <https://es.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html> (accessed on 18 September 2022).
27. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [[CrossRef](#)]
28. Keras Documentation: Keras Applications. [WWW Document]. 2022. Available online: <https://keras.io/api/applications/> (accessed on 18 September 2022).
29. He, T.; Zhi, Z.; Hang, Z.; Zhongyue, Z.; Junyuan, X.; Mu, L. Bag of Tricks for Image Classification with Convolutional Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 558–567. [[CrossRef](#)]
30. Hussain, M.; Bird, J.; Faria, D. A Study on CNN Transfer Learning for Image Classification. In *Advances in Computational Intelligence Systems Proceedings of 18th Annual UK Workshop on Computational Intelligence*. Nottingham, Great Britain; Lofti, A., Bouchachia, H., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 191–202. [[CrossRef](#)]
31. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
32. Sotillo, M.G.; Cerralbo, P.; Lorente, P.; Grifoll, M.; Espino, M.; Sanchez-Arcilla, A.; Álvarez-Fanjul, E. Coastal ocean forecasting in Spanish ports: The SAMOA operational service. *J. Oper. Oceanogr.* **2019**, *13*, 37–54. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.