





Article

A Study on the Monitoring of Floating Marine Macro-Litter Using a Multi-Spectral Sensor and Classification Based on Deep Learning

Youchul Jeong ¹, Jisun Shin ¹, Jong-Seok Lee ², Ji-Yeon Baek ^{2,3}, Daniel Schläpfer ⁴, Sin-Young Kim ², Jin-Yong Jeong ⁵ and Young-Heon Jo ^{1,6,*}

¹ Marine Research Institute, Pusan National University, Busan 46241, Republic of Korea; youchul1821@pusan.ac.kr (Y.J.); sjs1008@pusan.ac.kr (J.S.)

² BK21 School of Earth and Environmental Systems, Pusan National University, Busan 46241, Republic of Korea; hot4027@pusan.ac.kr (J.-S.L.); jiyeon@pusan.ac.kr or jy_baek@kiost.ac.kr (J.-Y.B.); tmzkt35@naver.com (S.-Y.K.)

³ Korea Ocean Satellite Center, Korea Institute of Ocean Science and Technology, Busan 49111, Republic of Korea

⁴ ReSe Applications LLC, 9500 Wil, Switzerland; daniel@rese-apps.com

⁵ Marine Disaster Research Department, Korea Institute of Ocean Science and Technology, Busan 49111, Republic of Korea; jyjeong@kiost.ac.kr

⁶ Department of Oceanography, Pusan National University, Busan 46241, Republic of Korea

* Correspondence: joyoung@pusan.ac.kr; Tel.: +82-51-510-3372

Abstract: Increasing global plastic usage has raised critical concerns regarding marine pollution. This study addresses the pressing issue of floating marine macro-litter (FMML) by developing a novel monitoring system using a multi-spectral sensor and drones along the southern coast of South Korea. Subsequently, a convolutional neural network (CNN) model was utilized to classify four distinct marine litter materials: film, fiber, fragment, and foam. Automatic atmospheric correction with the drone data atmospheric correction (DROACOR) method, which is specifically designed for currently available drone-based sensors, ensured consistent reflectance across altitudes in the FMML dataset. The CNN models exhibited promising performance, with precision, recall, and F1 score values of 0.9, 0.88, and 0.89, respectively. Furthermore, gradient-weighted class activation mapping (Grad-CAM), an object recognition technique, allowed us to interpret the classification performance. Overall, this study will shed light on successful FMML identification using multi-spectral observations for broader applications in diverse marine environments.

Keywords: floating marine macro-litter; unmanned aerial vehicle; multi-spectral sensor; atmospheric correction; reflectance retrieval; convolutional neural network



Citation: Jeong, Y.; Shin, J.; Lee, J.-S.; Baek, J.-Y.; Schläpfer, D.; Kim, S.-Y.; Jeong, J.-Y.; Jo, Y.-H. A Study on the Monitoring of Floating Marine Macro-Litter Using a Multi-Spectral Sensor and Classification Based on Deep Learning. *Remote Sens.* **2024**, *16*, 4347. <https://doi.org/10.3390/rs16234347>

Academic Editors: Yingying Dong, Chenghai Yang, Giovanni Laneve and Wenjiang Huang

Received: 29 August 2024

Revised: 15 November 2024

Accepted: 19 November 2024

Published: 21 November 2024



Copyright: © 2024 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

The escalating global use of plastic has led to a significant increase in marine pollution caused by marine litter, which is a pressing global concern. Marine litter encompasses persistent, human-made, and processed solid substances discarded or lost in marine and coastal environments [1]. FMML is defined as any floating item over 2.5 cm in length on the surface or in the surface layer of a water column [2–4]. FMML poses a significant threat to various marine life groups [5–11] and can affect beaches and the seafloor, where it has been reported to be a source of secondary micro-litter. To reduce such damage, it is essential to collect information on the distribution and trend of FMML systematically. Therefore, it is necessary to efficiently monitor the amount of FMML as well as its spatial and temporal changes, sources, and movement paths.

Many problems have occurred with marine litter in South Korea, and monitoring studies on beach litter have been conducted to solve these problems. Several studies have estimated the sources of marine plastic debris and investigated the distribution of small

pieces of plastic debris found on beaches. This can help reduce and evaluate the types and amount of beach litter [12–14]. However, due to the lack of effective monitoring and evaluation methods for FMML, more efficient monitoring is necessary to inform policies aimed at its reduction. Therefore, there is a need for remote sensing methods capable of providing a high spatial resolution and near-real-time data for the efficient monitoring and classification of macro-litter, particularly focusing on FMML.

Remote sensing methods for FMML, such as satellites [15–18], aircraft [15,16,19,20], and unmanned aerial vehicles (UAVs, also known as drones) [21–24], have emerged as efficient and cost-effective monitoring strategies. However, satellite and aircraft observations still have insufficient spatial resolutions for detecting and monitoring the small size of FMML at high flight altitudes, and obtaining permission for aircraft surveys is difficult. To address these limitations, drone observations, which offer a high spatiotemporal resolution and ease of access, have emerged as viable alternatives for monitoring FMML [20,22]. For example, Garcia-Garin et al. [20] distinguished between images with and without FMML captured by drones and aircraft in the northwest Mediterranean region. These surveys were performed using an RGB camera and analyzed using deep learning algorithms. Nonetheless, it is difficult to recognize marine litter in RGB images, as RGB bands allow only object chromaticity to be described [25]. Furthermore, using multi-spectral images to acquire multi-spectral characteristics to classify macro-litter types and materials by the spectral angle mapping (SAM) technique has been conducted on a beach dune [24]. Therefore, multi-spectral sensors can provide more detailed spectral differences in marine litter.

Machine learning approaches, such as support vector machines (SVM) and random forest (RF), have been widely applied to RGB and multi-spectral data analysis [23,26,27]. However, deep learning [28] has provided the efficient extraction of valuable information from aerial imagery [17,29,30], significantly reducing the time required for data processing. Particularly, CNNs have demonstrated a superior performance in image recognition tasks due to their capacity to automatically learn and extract complex features across multiple neural network layers [31,32]. CNN models utilizing RGB imagery have been specifically applied to the detection and classification of FMML. For instance, RGB camera-to-CNN approaches have successfully classified FMML into three distinct categories, with studies such as those conducted in Cambodia using high-resolution RGB aerial imagery to develop and train CNN-based detection and quantification systems [26]. Despite these advancements, RGB data alone imposes limitations due to its restricted spectral range, which may hinder accurate classification. Recent progress in drone-based multi-spectral data analysis now enables CNNs to leverage the extensive spectral information inherent in multi-spectral data, allowing for a more nuanced and automated feature extraction process. Building upon these advances, this study aims to address the limitations of RGB imagery by employing CNN models trained on high-resolution multi-spectral data, thereby enhancing FMML detection and classification. Therefore, while previous studies have primarily relied on CNNs trained with RGB data from drone imagery [20], this study seeks to advance FMML classification and monitoring by utilizing multi-spectral imagery, marking a significant step forward in remote sensing-based marine litter detection [33].

The aim of this study is to develop a deep learning model using drone-acquired multi-spectral images to monitor and classify FMML. First, we calculated the reflectance of FMML at different flight altitudes after correcting for atmospheric effects. Second, we analyzed the reflectance characteristics of FMML, which showed a wide range of colors and material types. Third, we trained and tested three CNN models to accurately classify FMML. Finally, we used a visualization technique called Grad-CAM to better understand how the models identified the FMML based on different types and colors, providing insight into the models' performance.

Therefore, for the first time, this study presents a comprehensive approach for classifying and detecting FMML using drone-based multi-spectral images. The remainder of this paper is organized as follows: Section 2 details the study area, data acquisition process, and methods used. In Section 3.1, we outline the methodology for acquiring and

preparing UAV-based multi-spectral images, including the use of atmospheric correction techniques for reflectance across varying altitudes. Section 3.2 describes the analysis of FMML reflectance characteristics, considering the wide range of sizes, colors, and materials. Section 3.3 presents the development and training of three CNN models optimized for FMML detection and classification. Section 3.4 applies Grad-CAM to visualize model performance, providing insights into how the models identify FMML based on different types and colors. Section 4 discusses the results, comparing the performance of different models and analyzing the effectiveness of our approach. Finally, Section 5 presents the conclusions of this study and outlines potential applications for marine litter monitoring and management.

2. Data and Methods

Figure 1 shows the overall workflow for classifying FMML using drone-acquired data and deep learning models. In this study, we selected the Dong-Sun seawall, Gadeok Island, as the study area, which is the location of the National Coastal Litter Monitoring Program conducted by the Ministry of Oceans and Fisheries of South Korea [34], according to different geographical features and human activities. Drone surveys were also conducted in the study area. The drone was equipped with a multi-spectral sensor to acquire multi-spectral images of FMML. We manually selected and used all acquired images without interfering with the FMML detection. The FMML images were divided into four categories for model training: film, fiber, fragment, and foam. The DROACOR (version 2.0.1, ReSe Application LLC, Wil, Switzerland) method [35] was used to obtain consistent reflectance from the drone's multi-spectral images obtained at various flight altitudes. The entire FMML dataset obtained through the drone surveys increased the number of images through image augmentation to improve the effective classification and accuracy before using deep learning algorithms. The images were then resized for the utilization a CNN model. Finally, three different models were constructed based on the number of training images. The completed models were classified using the CNN algorithm, and Grad-CAM was employed to visualize the classification results, helping to identify and improve model performance in the challenging areas of FMML detection.

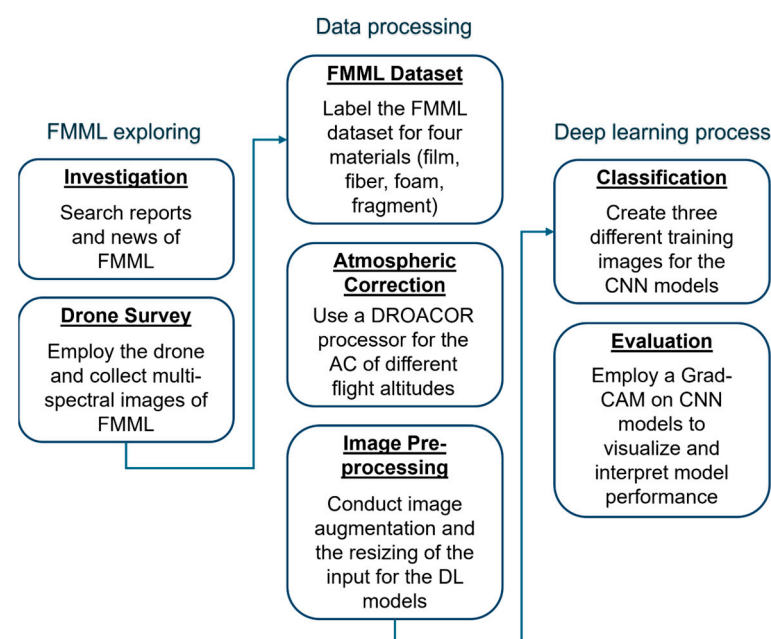


Figure 1. The overall workflow shows the processes that led to the classification of FMML using drone-acquired data and deep learning models. We performed three steps: (1) FMML exploration; (2) data processing for the deep learning models; and (3) deep learning to process FMML classification and visualization.

2.1. Study Area

This study was conducted in the Dong-Sun seawall area of Gadeok Island in South Korea, which was chosen as one of the sites for the National Coastal Litter Monitoring Program, supported by the MOF and organized by the Korea Marine Environment Management Corporation [34]. This program, which has been ongoing since 2018, conducts regular surveys every two months at 40 selected coastal sites along South Korea's shorelines. The selection criteria for these sites include a minimum coastline length of 100 m, characterized by sandy and pebbly beaches. These terrain-imposed mobility restrictions during the two-month monitoring period and were infrequently subjected to cleaning activities, allowing litter to accumulate during the monitoring process. Based on these criteria, we conducted three pre-site field trips, identifying the study area as the Dong-Sun seawall on the southern coast of South Korea's South Sea (Figure 2). This location is near the mouth of the Nakdong River, one of the largest river mouths in South Korea, with geographical features that facilitate the accumulation of marine litter. The region is popular for tourism activities, such as camping and fishing, contributing to the generation of marine litter due to anthropogenic activities. Furthermore, because of ocean currents and weather patterns, a significant amount of marine litter has been found near the study area, including on Jinwoo Island and Shinja Island. Therefore, we chose this study area based on National Coastal Litter Monitoring Program criteria and our pre-site field trips.

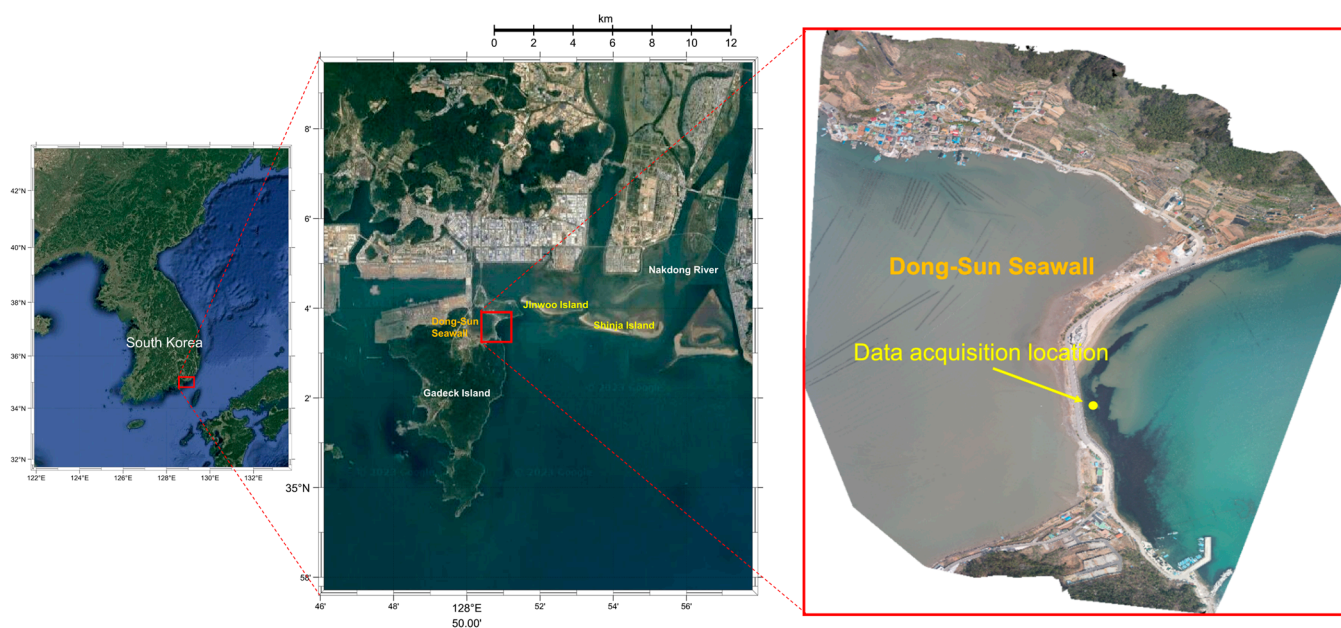


Figure 2. The study location on Gadeok Island in South Korea and the data acquisition location of the drone surveys in the study area in drone-based imagery (red rectangle). Maps of the study area and a Pix4Dmapper image were used to illustrate the data acquisition.

2.2. Drone Surveys

This study used the DJI Inspire 2 (DJI, Shenzhen, China), a drone with a maximum payload capacity of 4 kg. The flight time varies according to the battery and payload, providing approximately 27 min of operation. It features a dual-battery system for extended flight times and can reach speeds of up to 94 km/h in sport mode. The drone was equipped with a 5-channel Micasense RedEdge-MX sensor (Micasense Inc., Seattle, WA, USA) for data collection. Table 1 provides detailed information regarding the specification of the Micasense RedEdge-MX sensor. This sensor captures images in five spectral bands with central wavelengths of 475, 560, 668, 717, and 840 nm, corresponding to the blue, green, red, red edge, and near-infrared (NIR) bands, respectively. The blue and green bandwidths were 20 nm, while the red and red edge bandwidths were 10 nm, and the NIR bandwidth

was 40 nm. Using drone surveys to acquire images for object detection and classification was important for image resolutions, expressed as the ground sampling distance (GSD). The GSD is determined by flight altitudes and sensor properties. The Micasense sensor had width of 4.8 mm and a focal length of 5.4 mm. It achieved a GSD of 8 cm per pixel at a flight altitude of 120 m [36,37].

Table 1. Detail of Micasense RedEdge-MX sensor Specification.

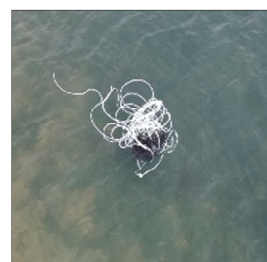
| Band Name | Blue | Green | Red | Red Edge | NIR |
|-------------------------|------|-------|-----|----------|-----|
| Central wavelength (nm) | 475 | 560 | 668 | 717 | 840 |
| Bandwidth (nm) | 20 | 20 | 10 | 10 | 40 |

Drone surveys were conducted on 11 January, 29 March, 20 July, and 7–8 September 2023 to represent the seasonal seawater conditions and sunlight of this study area. These surveys were conducted over seawater approximately 3 m from the coastline of the Dong-Sun seawall on Gadeok Island. Observations were conducted between 11:00 a.m. and 4:30 p.m. on cloudy and clear days to minimize the potential influence of wind on object detection and calm sea conditions (i.e., a Beaufort sea state < 3). This precaution aimed to mitigate the effects of sun glint, turbidity, and cloud shadows, as these factors can significantly affect the detection of marine litter. However, it is worth noting that we manually selected images after the surveys, ensuring that these parameters did not affect the results. Furthermore, the flight altitude of the drone used for the image acquisition was carefully controlled. All images captured using the Micasense sensor fell within the range of 10 to 170 m, ensuring a GSD that ranged from 0.69 to 11.8 cm per pixel. The sensor angles were adjusted to the nadir (90° to the ground) to maintain proper alignment, and a shooting interval of 2–3 s was set to enable automated capturing and facilitate the acquisition of multi-spectral images.

2.3. Data Preparation

2.3.1. FMML Dataset

We collected a variety of colors, sizes, and types of litter to represent the FMML accurately, including plastic bags, buoys, fishing ropes, baskets, and plastic bottles. Figure 3 shows sample imagery collected around the study area and represents the most common marine litter found on South Korean coastlines. The marine litter was labeled by the source of its type (material), following the study in [34] for the National Coastal Litter Monitoring Program. Table 2 lists the different size ranges, colors, and types of litter included in the dataset. The size ranges were suggested for the floating litter surveys based on the report and were determined by the size of the largest dimensions of the FMML. The size range was from 2.5 cm to over 50 cm [38]. Overall, 43% of the FMML dataset was composed of film, and 29% was composed of fragments; both fiber and foam were 14% of the dataset. Different colors were balanced among the types of marine litter, and we put a white color in each type to compare the reflectance of the types.



White - fishing rope - fiber



White - buoy - foam

Figure 3. Cont.

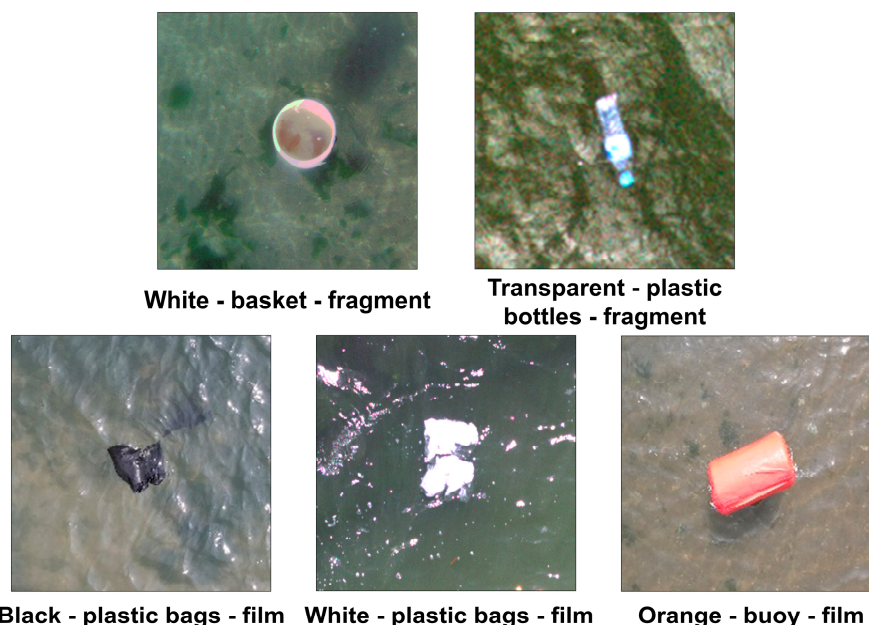


Figure 3. FMML dataset of images captured by the drone in the study area.

Table 2. The information of the FMML dataset, including the colors, types, marine litter, and sizes of the FMML, measured by its largest dimensions (A: $5 \leq 10$ cm; B: $10 \leq 20$ cm; C: $20 \leq 30$ cm; D: $30 \leq 50$ cm; E: > 50 cm).

| Color | Type | Marine Litter | Size |
|-------------|----------|-----------------|------|
| White | Film | Plastic bags | A |
| Black | | Plastic bags | A |
| Orange | | Buoy | E |
| White | Fiber | Fishing rope | D |
| White | Fragment | Basket | C |
| Transparent | | Plastic bottles | B |
| White | Foam | Buoy | E |

2.3.2. Atmospheric Correction for Multi-Spectral Measurements

The DROACOR method retrieves reflectance values from sensor radiance. It dynamically adjusts to specific devices and prevalent atmospheric conditions, offering a versatile solution for correcting multi-spectral and hyperspectral images [35]. Its function is to meticulously remove atmospheric influences from data taken by drone-based sensors, accurately retrieving reflectance. The processing methodology for the DROACOR method involves a systematic sequence of steps to ensure accurate data analysis. The process encompasses data preparation, inflight radiometric calibration, libRadtran, look-up table (LUT) generation, reflectance retrieval by atmospheric correction, spectral optimization, terrain and bidirectional reflectance distribution function (BRDF) correction, and verification. For a multi-spectral case over a flat terrain, no terrain correction and spectral optimization is required, so BRDF correction was not applicable for the present case. Atmospheric correction was applied to all multi-spectral images obtained in our campaigns, utilizing the DROACOR procedure. The image acquisition dates were aligned with the drone surveys conducted on 11 January, 29 March, 20 July, and 7–8 September 2023. All images captured between 11:00 and 16:30 UTC+9 were carefully distributed across diverse time zones. A singular dataset of Micasense reflectance panels (with an approximately 50% reflectance) obtained 1 m above ground was directly employed on the shore for panel reflection information essential for the radiometric sensor calibration. The sensor altitude of all multi-spectral images ranged from 10 to 170 m. Consequently, a dedicated atmospheric

correction procedure was implemented for each set of multi-spectral images, aiming to minimize atmospheric correction discrepancies to the utmost extent possible.

2.3.3. Image Processing

In this study, we initially collected images captured at altitudes ranging from 10 to 170 m. However, during the development of the deep learning model, we encountered significant challenges in recognizing objects in images. Consequently, we reorganized the dataset, including only images taken at altitudes between 10 and 40 m based on the GSD (2.77–6.94 cm/pixel). This adjustment enhanced the object recognition capabilities and overall accuracy of the model. We meticulously examined the FMML dataset and categorized a subset of the images into the four main groups: film, fiber, fragment, and foam. However, the initial number of images in each category (Table 3) was insufficient for effectively training the deep learning model. To address this issue, we employed image-augmentation techniques to expand the dataset [39]. Image-augmentation artificially increases the number of images by introducing variations, which helps to reduce the risk of overfitting and potentially improve the deep learning model performance [32,39].

Table 3. Label information on the number of multi-spectral images obtained from 10 to 40 m among the FMML dataset.

| Label Name | Contents | Number | Percentage (%) |
|------------|--|--------|----------------|
| Film | (White, Black) Plastic bags, (Orange) Film buoy | 354 | 40 |
| Fiber | (White) Fishing rope | 203 | 23 |
| Fragment | (White) Basket, (Transparent) Plastic bottles | 124 | 14 |
| Foam | (White) Styrofoam buoy | 206 | 23 |

In our image augmentation process, we implemented cropping, horizontal and vertical flips, and rotations of 70°, 140°, and 210° on the available images. Consequently, after the image augmentation, the total number of images in the FMML dataset increased to 4908 for the film, 2880 for the fiber, 1908 for the fragment, and 2844 for the foam. Subsequently, to expedite and enhance the classification and detection capabilities of the deep learning model, we resized the entire dataset to 128 × 128 × 5 pixels. Table 4 illustrates the outcomes of manually identifying and categorizing the dataset into four categories. Furthermore, we created three different training sets for the three CNN models to determine the impact of the number of images in each category of the FMML on the CNN model components: (1) CNN-1 (1000 images per category), (2) CNN-2 (1500 images per category), and (3) CNN-3 (4908 film images, 2880 fiber images, 1908 fragment images, and 2844 foam images) (Table 4).

Table 4. CNN model information with three different training image sets.

| Model | Training Image Number | | | |
|-------|-----------------------|-------|----------|------|
| | Film | Fiber | Fragment | Foam |
| CNN-1 | 1000 | 1000 | 1000 | 1000 |
| CNN-2 | 1500 | 1500 | 1500 | 1500 |
| CNN-3 | 4908 | 2880 | 1908 | 2844 |

2.4. Deep Learning Algorithm

In this study, we developed an algorithm for automatically detecting FMML in drone images using a deep learning approach based on CNN architecture. Unlike other machine learning methods that require a predefined set of features, deep learning models, specifically CNNs, learn and recognize the spatial patterns of a target using features embedded within its architecture. The architecture of the CNN model used in this study is shown below

(Figure 4). The CNN model was developed using MathWorks. The figure demonstrates the creation and training of a simple classification-based deep learning network. The CNN architectures used in this study are listed in Table 5.

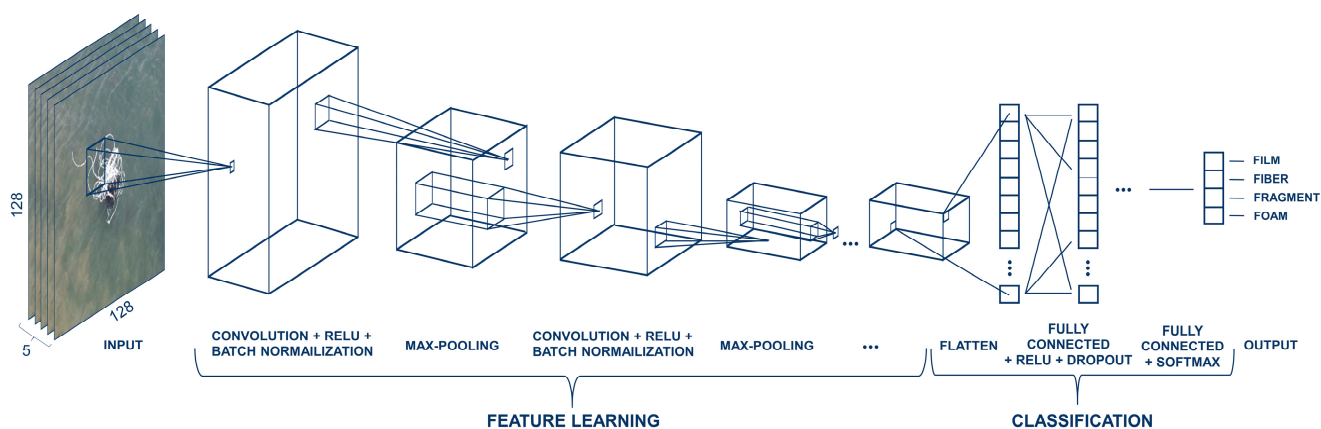


Figure 4. CNN architecture for the classification of FMML. The training, validation, and test sets comprised FMML datasets as input. The input image size was $128 \times 128 \times 5$. The output was labeled as film, fiber, fragment, and foam for the FMML. This network consisted of input, feature learning, classification, and output.

Table 5. CNN structures used in this study.

| Order | Layers | Options |
|-------|---------------------------|---|
| 1 | Image Input Layer | Input size: $128 \times 128 \times 5$ |
| 2 | Convolution 2D Layer | Filter size: 3×3 ; number of filters: 32; stride: 1×1 |
| 3 | Batch Normalization Layer | Mean decay: 0.1; variance decay: 0.1; epsilon: 0.00001 |
| 4 | ReLU Layer | Number of inputs: 1; number of outputs: 1 |
| 5 | Max-Pooling 2D Layer | Poll size: 2×2 ; stride: 2×2 |
| 6 | Convolution 2D Layer | Filter size: 3×3 ; number of filters: 32; stride: 1×1 |
| 7 | Batch Normalization Layer | Mean decay: 0.1; variance decay: 0.1; epsilon: 0.00001 |
| 8 | ReLU Layer | Number of inputs: 1; number of outputs: 1 |
| 9 | Max-Pooling 2D Layer | Poll size: 2×2 ; stride: 2×2 |
| 10 | Convolution 2D Layer | Filter size: 3×3 ; number of filters: 32; stride: 1×1 |
| 11 | Batch Normalization Layer | Mean decay: 0.1; variance decay: 0.1; epsilon: 0.00001 |
| 12 | ReLU Layer | Number of inputs: 1; number of outputs: 1 |
| 13 | Max-Pooling 2D Layer | Poll size: 2×2 ; stride: 2×2 |
| 14 | Fully Connected Layer | Output size: 512 |
| 15 | ReLU Layer | Number of inputs: 1; number of outputs: 1 |
| 16 | Dropout Layer | Number of layers: 19; number of connections: 18 |
| 17 | Fully Connected Layer | Output size: 4 |
| 18 | Softmax Layer | Number of inputs: 1; number of outputs: 1 |
| 19 | Classification | Classes: 4 |

Our network initially consisted of a sequence of layers, including three 2D convolutional layers, three batch normalization layers, four rectified linear unit (ReLU) layers, and

three max-pooling 2D layers for feature extraction. The convolution process is mathematically expressed using Equation (1) [40].

$$y_n = f \left(\sum_n x_n \times \omega_n + b_n \right) \quad (1)$$

The training and test sets consisted of drone images with inputs from five spectral bands. The input images contained preprocessed FMML datasets with the dimensions $128 \times 128 \times 5$. The output images were labeled as film, fiber, fragment, and foam. The network architecture consisted of input, feature extraction, classification, and output components, as listed in Table 5. In Equation (1), y , x , w , and b represent the pixel values of the output and input images, filter pixels, and bias term, respectively. Following the convolution, the ReLU activation function (f) was applied after each convolution layer to introduce nonlinear operations. The pooling layers were used to down-sample data from the previous layers. In this study, we employed max-pooling, which involves selecting the maximum value from the corresponding pixels in an input image to generate a new pixel value. Fully connected and SoftMax layers were added. Fully connected layers linked the output neurons from the previous layers to each neuron in the subsequent layers, creating a network of interconnected neurons. The results from this layer were normalized into a probability distribution for classification using the SoftMax layer. The final output was assigned to the classification layer.

After defining the network architecture, we specified the options for training, validation, and testing. We randomly divided the model split ratios for each CNN model into 60% for training, 20% for validation, and 20% for testing. These images were not duplicated and were repeatedly used in the process. We used the adaptive moment estimation (Adam) optimizer based on adaptive estimates of the first and second moments [41]. The initial learning rate for the Adam was set to 0.01, and the maximum number of epochs was set to 50, with each epoch representing a complete training cycle over the entire training dataset. A shuffle option was applied to each epoch. Throughout the training process, the software consistently evaluated the accuracy of the validation data. The network was trained using a high-speed GPU to maximize processing efficiency.

2.5. Grad-CAM

In various deep learning applications related to imaging analysis, it is very important to make deep learning models interpretative and enhanced. Grad-CAM, a generalized method derived from the class activation map (CAM) among object recognition techniques, provides an explainable perspective on deep learning models [42,43]. Visual explanations using Grad-CAM have been conducted as previously described [43]. Therefore, in this study, we applied Grad-CAM to check the possibility of enhancing the proposed deep learning model and to analyze its interpretability. Grad-CAM uses multi-spectral images containing preprocessed FMML datasets as an input. Grad-CAM techniques were subsequently applied using the CNN model and then applied to any of the convolutional (Conv) layers. Consistently with previous studies [43,44], Grad-CAM was applied to the last Conv layer. This method indicates that Grad-CAM is a valuable tool for visually explaining the results obtained from a CNN model, which may enhance the accuracy of deep learning models.

3. Results

3.1. Reflectance Estimation of Different Flight Altitudes Through Atmospheric Correction

Figure 5a shows the drone image and spectral data according to the altitude obtained on 29 March 2023. The spectral data of the FMML were obtained continuously at the same location from at flight altitudes of 23 to 170 m. During the 10 min of data acquisition, the solar zenith angle varied slightly between 32.9° and 35.7° . We conducted atmospheric correction and analyzed the following results with the orange film buoy, the largest object in the FMML dataset of this study. Before atmospheric correction, images acquired at different

altitudes showed differences in the image value (Figure 5b). However, the images showed consistent reflectance for all flight altitudes between 23 and 170 m above the ground after atmospheric correction using the DROACOR method (Figure 5c). Reflectance retrieval by atmospheric correction for the other FMML datasets, except for the orange film buoy, was also conducted, and the results showed consistent reflectance regardless of the altitude. The average deviation quantified the effect of the atmosphere on the observed data [35]. The average deviation for each band covering all seven altitudes ranged from 0.71 to 1.32% (Table 6).

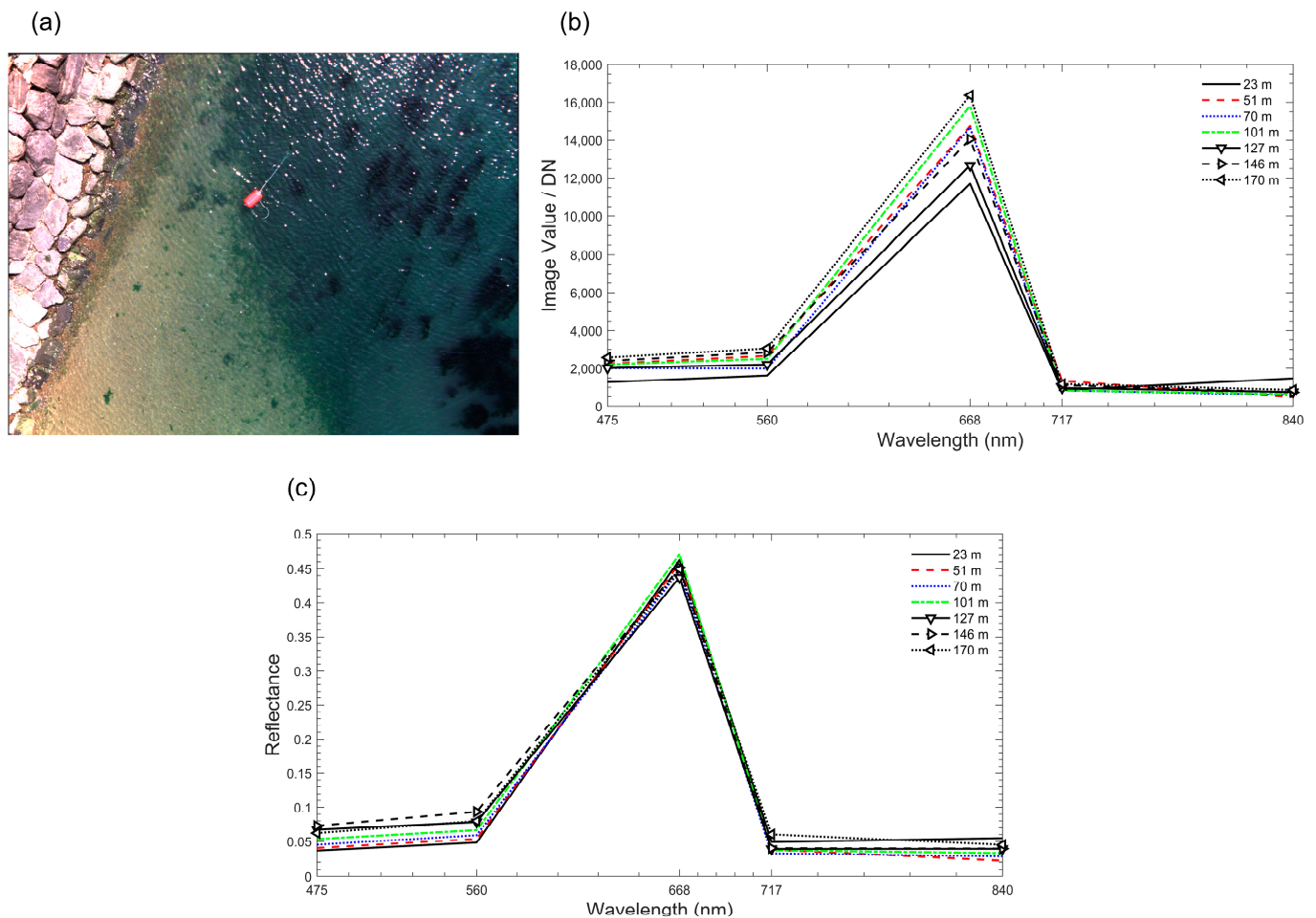


Figure 5. Reflectance analysis of flight altitude through atmospheric correction. (a) A multi-spectral image was obtained on 29 March 2023 (true color RGB; R: 668 nm; G: 560 nm; B: 475 nm; a 51 m flight altitude). Images for atmospheric correction were acquired at altitudes of 23, 51, 70, 101, 127, 146, and 170 m. (b) The image values for each altitude of the orange film buoy image before atmospheric correction were compared. (c) The reflectance for each altitude of the orange film buoy image using a DROACOR atmospheric correction processor were compared.

Table 6. Average deviations of reflectance outputs (units: reflectance) at various flight altitudes after atmospheric correction.

| Band wavelength (nm) | 475 | 560 | 668 | 717 | 840 | Mean |
|-----------------------|------|------|------|------|------|------|
| Average deviation (%) | 1.15 | 1.32 | 0.86 | 0.71 | 0.82 | 0.97 |

3.2. Analysis of FMML Reflectance

Since we obtained consistent reflectance after atmospheric correction, as demonstrated in Section 3.1, atmospheric correction was applied to subsequent datasets. Figure 6 shows

the reflectance spectra of each marine litter within the FMML dataset used in this study. We used two fragments consisting of hard plastic materials that tend to float in seawater: transparent fragment plastic bottles and the white fragment basket. The former exhibited a reflective shape similar to that of the seawater spectral signals because of their transparency. The latter also showed a spectrum that partially resembled seawater, with a reflectance ranging from 0.18 to 0.24 in the blue, green, and red edge bands but dropping sharply to approximately 0.02 in the red and 0.04 in the NIR bands. For the film label, seawater did not significantly affect the black and white plastic bags. However, the orange film buoy, which naturally floats well in seawater, exhibited a reflectivity value of 0.4, particularly in the red edge band. The white fiber fishing rope was used for the fiber label, and the reflectance of the fishing rope fell between 0.17 and 0.19 in the blue and green bands but dropped to less than 0.03 in the red and NIR bands. The reflectance of the white foam buoy, a common floater in seawater, was notably high, exceeding 0.15 in the blue, green, and red edge bands, but declining to approximately 0.02 in the red and NIR bands.

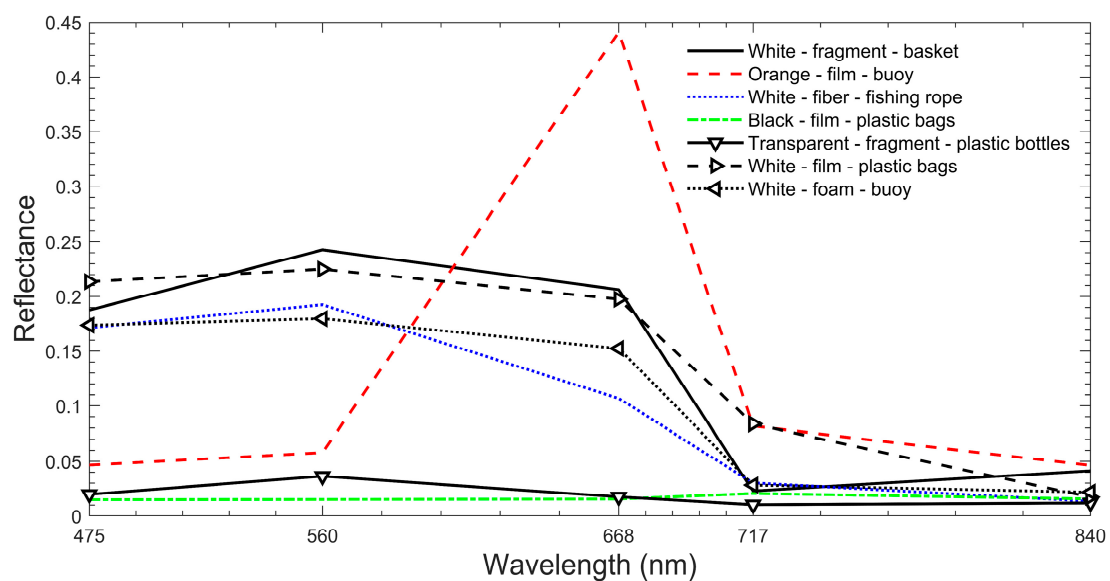


Figure 6. Spectra of all FMML lists in the dataset from the DROACOR-calculated reflectance.

3.3. Performance of CNN Models

To assess the classification accuracy of the FMML images, we developed and evaluated three different CNN models, each trained with a different number of images. The total number of images for the three CNN models (representing the number of images in the FMML dataset) was used for training, validation, and testing during the classification. During this process, 60% of the images were allocated for training, 20% for validation, and 20% for testing during cross-validation. A confusion matrix was used for analysis [45]. In the confusion matrix, the representation of “false” and “true” corresponded to the presence or absence of FMML in the actual dataset. The overall performance of the model results was evaluated for both processes using four parameters: accuracy, precision, recall, and the F1 score. Accuracy (Equation (2)) represents the fraction of all processed images that were correctly classified as containing or not containing FMML. In contrast, precision (Equation (3)) represents the fraction of images classified as containing FMML belonging to that class. In contrast, recall (Equation (4)) represents the fraction of correctly labeled images within each class. The accuracy, precision, and recall values ranged from 0 to 1. The F1 score (Equation (5)) represents the balance between precision and recall [46], and its value increases according to the performance of the model [47]. The repeatability of the method was tested by performing ten runs on randomly selected image sets.

$$\text{Accuracy} = \frac{(TP + TN)}{N} \quad (2)$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (3)$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (4)$$

$$\text{F1 score} = \frac{(2 * TP)}{(2 * TP + FP + FN)} \quad (5)$$

where TP is true positive, where images with FMML are well classified; TN is true negative, where images without FMML are well classified; FP is false positive, where images without FMML are misclassified; FN is false negative, where images with FMML are misclassified; and N is total number of images of FMML.

Table 7 lists the performances of the three CNN models for the FMML. We created 12,540 images through drone surveys and image pre-processing, including manual selection and image augmentation. The datasets were separated into three trained CNN models. CNN-1 and -2 were trained using randomly selected images from the entire dataset, whereas CNN-3 was trained using all the images. Each CNN model produced different results based on the training data. The CNN-1 model, trained with 1000 images from each label, yielded a precision of 0.57 and an F1 score of 0.54. The CNN-2 model, trained with 1500 images from each label, yielded a precision of 0.73 and an F1 score of 0.73. Meanwhile, the CNN-3 model, trained with the highest number of images, exhibited a precision of 0.90 and an F1 score of 0.89.

Table 7. Performance of evaluation of three CNN models using test set for FMML.

| Model | Training Image Number | | | | Precision | Recall | Accuracy | F1 Score |
|-------|-----------------------|-------|----------|------|-----------|--------|----------|----------|
| | Film | Fiber | Fragment | Foam | | | | |
| CNN-1 | 1000 | 1000 | 1000 | 1000 | 0.57 | 0.51 | 0.51 | 0.54 |
| CNN-2 | 1500 | 1500 | 1500 | 1500 | 0.73 | 0.72 | 0.73 | 0.74 |
| CNN-3 | 4908 | 2880 | 1908 | 2844 | 0.90 | 0.88 | 0.90 | 0.89 |

Figure 7 shows the confusion matrix results for the CNN-3 model. Among the four classification labels, foam had the highest precision value at 95.4%, whereas fragment had the lowest recall value at 73.0%. The three performance evaluation indicators are presented in Table 8.

3.4. CNN Model Evaluation Using Grad-CAM

We visualized images from the FMML dataset, which consisted of four labels with complex combinations, including colors and types, using Grad-CAM on the CNN-3 model (Figure 8) to evaluate how the CNN model performed. In Figure 8, the left image represents the class activation map, and the right image represents the Micasense multi-spectral image of band five. The class activation map represents the high-intensity visuals (red and yellow) that reflected the area of interest to our model at the time of prediction. Each image was obtained from confident (Figure 8a–d) and unconfident (Figure 8e–h) detections of each label, depending on the heatmap outcomes. The heatmap generated by the Grad-CAM emphasizes crucial areas in the input image that have a significant impact on the final prediction of the network. The heatmap outcomes in Figure 8a–d show the clear visibility of the images for each label. The confident detection demonstrated that our model visualized the correct heatmap corresponding to the characteristics of the dataset images. For instance, Figure 8a,c represent the confident detection achieved despite the turbid seawater conditions and high flight altitude, respectively. In contrast, sun glint and the movement of seawater led to unconfident detections and the visualization of incorrect heatmap outcomes. In contrast, the heatmap outcomes in Figure 8e–h show the obscured visibility of the labeled images. The sun glint and movement of seawater led to unconfident detections and the visualization of incorrect heatmap outcomes. Figure 8e represents

unconfident detections owing to the influence of sun glint, which led to misinterpretation. Figure 8h shows that the movement of seawater around the white foam buoy generated an incorrect heatmap pattern. Consistent with our findings, previous studies have shown that potential adverse factors, such as fog, sun glint, and shadow, lead to unconfident detections [48–50].

Confusion Matrix

| Predicted Class | Actual Class | | | | |
|-----------------|---------------|---------------|---------------|----------------|----------------|
| | Fiber | Film | Foam | Fragment | |
| Fiber | 533 21.2% | 12 0.5% | 26 1.0% | 42 1.7% | 86.9% 13.1% |
| Film | 20 0.8% | 948 37.8% | 20 0.8% | 51 2.0% | 91.2% 8.8% |
| Foam | 3 0.1% | 12 0.5% | 517 20.6% | 10 0.4% | 95.4% 4.6% |
| Fragment | 20 0.8% | 10 0.4% | 6 0.2% | 279 11.1% | 88.6% 11.4% |
| | 92.5% 7.5% | 96.5% 3.5% | 90.9% 9.1% | 73.0% 27.0% | 90.8% 9.2% |

Figure 7. A confusion matrix of the CNN-3 model (x -axis: recall; y -axis: precision). The green box indicates correct classification by the model, and the red box indicates incorrect classification.

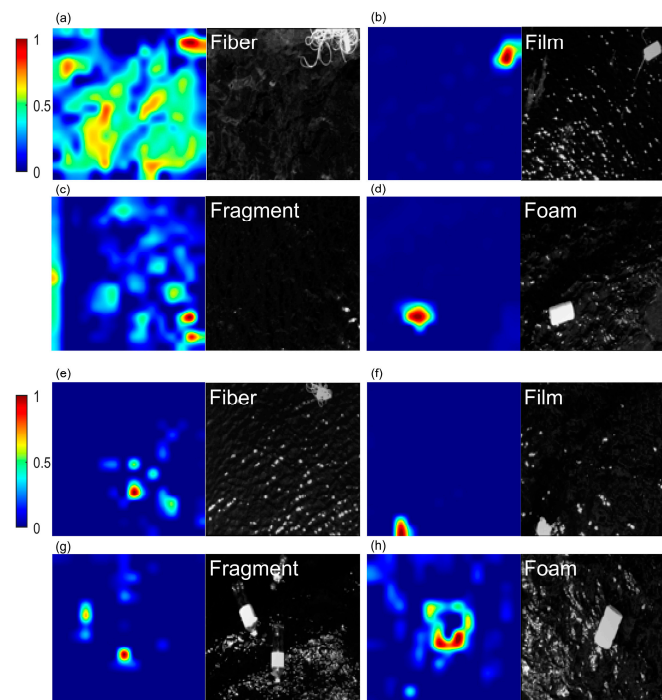


Figure 8. Visualization of FMML using Grad-CAM on CNN-3 model. (a–d) Confident detections of FMML dataset labels. (e–h) Unconfident detections of FMML dataset labels.

Table 8. Precision, recall, and F1 score of four labels in CNN-3 model.

| CNN-3 | Precision | Recall | F1 Score |
|----------|-----------|--------|----------|
| Fiber | 0.86 | 0.92 | 0.89 |
| Film | 0.91 | 0.96 | 0.93 |
| Foam | 0.95 | 0.90 | 0.93 |
| Fragment | 0.88 | 0.73 | 0.80 |

4. Discussion

Despite the high spatial resolution of the drone's multi-spectral images, it was very difficult to directly classify the FMML as having various colors and types. Thus, we developed a classification model for this study with the CNN models to apply it to the multi-spectral imagery. We found that our model was successful in distinguishing FMML of different types with the same color, and our model also demonstrated a high accuracy in the current dataset. Nevertheless, the model's performance may be limited by the diversity of the FMML training dataset. While the results are promising, we cannot confidently assert that the model would maintain an equally high accuracy when applied to datasets with a broader range of categories or different domains. This limitation highlights the need for further testing and validation using more diverse datasets to ensure the model's generalizability and robustness across various contexts. To advance a deep learning model for FMML classification, future work may increase the number of FMML categories and investigate geographical and environmental conditions (e.g., turbidity and water movement) across different regions of the world. This may lead to different outcomes. Moreover, long-term monitoring could provide insights into seasonal variations in marine litter distribution and help trace the destination of FMML, thereby improving our understanding of its movement. In this study, we only considered the four different types of FMML (film, fiber, fragment, and foam) classification and used five multi-spectral band spectra (blue, green, red, red edge, and NIR). This Section describes FMML classification using reflectance, two different CNN model comparisons based on RGB and multi-spectral imageries with the same parameter, and a qualitative evaluation of our CNN model.

4.1. FMML Classification Using Reflectance

Our results show the reflectance spectra of different types and colors within the FMML dataset, including fragment, foam, fiber, and film types. Dataset labels were established based on the National Coastal Litter Monitoring Program of the MOF. We used an automatic processor, called DROACOR, for atmospheric correction to calculate the reflectance of the FMML, with each item of FMML exhibiting a unique reflectance shape. Using panel data directly for reflectance retrieval is a valid approach as long as the panel data are acquired in similar atmospheric conditions, e.g., before or after data acquisition. Our results show a good agreement between the two reflectance retrievals (Figure 5b,c). Using the panel for sensor calibration, as performed in DROACOR, instead of as a reflectance reference, offers several advantages. An absolute sensor calibration can be applied to a long series of data acquisitions, eliminating the need for repeated panel measurements. Additionally, temporal changes in solar illumination and atmospheric conditions can be better accounted for with absolute calibrated data and appropriate physical modeling, as within DROACOR. This is especially useful for analyzing dark objects such as shaded areas, suspended sediments, or water bodies [51,52].

It is important to note that the reflectance results' interpretation process can be influenced by seawater [53], potentially resulting in variations in the reflectance of the actual objects. The reflectance of the FMML with similar colors showed remarkable similarities. However, our findings highlight the potential of material-based classification and that the types of plastics had different reflectance spectra [54]. Moreover, the outstanding accuracy in classifying white objects, including fiber rope, film plastic bags, and a fragment basket, as demonstrated in the confusion matrix of the CNN-3 model, offers compelling evidence that the model's performance is not solely dependent on color. Notably, the transparent

fragment plastic bottles and black film plastic bags exhibited the lowest reflectance within the FMML dataset. This outcome is likely attributable to their direct interaction with seawater. The transparent fragment plastic bottles showed a higher reflectance in the blue and green bands, consistent with the reflective properties of seawater. Conversely, the black film plastic bags resulted in a reduced reflectance, which was attributed to light absorption according to the color. These reflectance results within the FMML dataset emphasize the significance of color and type properties when utilizing deep learning models alongside multi-spectral images. Therefore, each FMML reflectance value obtained from a multi-spectral sensor is crucial to effectively leverage deep learning.

4.2. Model Comparison Between RGB and Multi-Spectral Imagery Bases

We initially hypothesized that the multi-spectral images would outperform the RGB images in marine litter detection due to their broader spectral information. While RGB images excel in the shape-based identification of specific objects, the diverse nature of marine litter necessitates a more comprehensive approach. In this study, we aimed to test this hypothesis and evaluate the effectiveness of multi-spectral images in distinguishing various colors and types of marine litter.

To identify the contribution of red edge and NIR bands to the classification of FMML, we compared the performance of the CNN model trained from RGB images. Three models were generated using the same training options and image numbers by exclusively extracting the RGB bands (R: 668 nm; G: 560 nm; B: 475 nm) from the dataset used in this study (Table 9). Compared to the RGB-3 model, the precision, recall, and F1 score of the CNN-3 increased by 11.11%, 8.64%, and 11.25%, respectively (Table 7). These results indicate that the CNN-3 model can classify each label of the FMML better than the RGB-3 model. Therefore, our findings suggest that the multi-spectral image model yields superior results. The deviation in the accuracy assessment values among the three models using only the RGB images was minimal. None of the RGB models showed any difference in the overall evaluation index, even when the number of trained images was changed. Therefore, our findings emphasize the superiority of multi-spectral images, which can capture a broad range of colors and types of litter. Additionally, they contribute to the accurate identification of FMML, thereby enhancing the performance of deep learning models. The comparison of deep learning models using RGB and multi-spectral images provides significant insights and important results, demonstrating the need for red edge and NIR bands. The RGB imagery demonstrates proficiency in tasks where color information is sufficient for feature extraction, and its limitations become apparent in tasks that require a deeper understanding of spectral information.

Table 9. Performance evaluation of three CNN models using RGB image of test set for FMML.

| Model | Training Image Number | | | | Precision | Recall | Accuracy | F1 Score |
|-------|-----------------------|-------|----------|------|-----------|--------|----------|----------|
| | Film | Fiber | Fragment | Foam | | | | |
| RGB-1 | 1000 | 1000 | 1000 | 1000 | 0.74 | 0.74 | 0.74 | 0.74 |
| RGB-2 | 1500 | 1500 | 1500 | 1500 | 0.77 | 0.76 | 0.76 | 0.76 |
| RGB-3 | 4908 | 2880 | 1908 | 2844 | 0.81 | 0.81 | 0.82 | 0.80 |

4.3. Qualitative Evaluation of CNN Model

CNN models provide visual interpretability of their results, which is a significant advantage of techniques like Grad-CAM [43]. In medical imaging, Grad-CAM has been effectively applied to various tasks, demonstrating its utility in visualizing important regions for model decisions [55]. Figure 8a–d visually represent the model’s focus areas quite well, particularly for confident detections where the heatmaps correctly highlighted the characteristics of each label. This visual feedback is crucial for validating the model’s decision-making process and ensuring that its focus aligns with expert knowledge. However, Grad-CAM also revealed limitations in our model’s performance under challenging environmental conditions. Unconfident

detections, as shown in Figure 8e–h, demonstrated that factors such as sun glint, water movement, and turbidity could lead to incorrect heatmap patterns and potential misinterpretations. These findings align with previous studies highlighting the impact of adverse environmental factors on detection accuracy [48–50]. While Grad-CAM proves valuable for marine debris analysis, offering both high accuracy and interpretability for confident detections. However, it also reveals areas for improvement, particularly in handling complex environmental conditions. This dual insight into the model's strengths and weaknesses is crucial for enhancing the reliability of AI-assisted environmental monitoring systems and guiding future improvements in model performance.

Figure 9 shows a multi-spectral image of band five of the well-classified and misclassified cases in the confusion matrix results of the CNN-3 model. These results represent a qualitative evaluation based on the confusion matrix. We analyzed the classifications included in the FMML dataset using the following resulting images: The results of the correct classifications are shown in Figure 9a,f,k,p, with the FMML displayed in a green circle. The green circles represent the well-classified FMML while the red circles represent the misclassified FMML of each category. The model of this study, which was trained differently at various flight altitudes and with various seawater influences, with various sizes, colors, and shapes of the FMML in the images, showed results as a quantitatively good classification model. Nevertheless, there are five reasons for this misclassification. First, there were cases where the shapes were similarly or incorrectly classified. Second, misclassification was caused by the model not showing many different colors. Third, the size of the FMML in the images caused misclassification of the image if it was too small. Fourth, external factors, such as sun glint and foam caused by the background of the image and not the influence of the FMML, resulted in misclassification. Fifth, one image of FMML in this study had an incorrect labeling mistake upon image preprocessing. As shown in Figure 9b, it is a fiber, but it is detected as a film, and many sun glints in the resulting image are thought to have hindered correct classification. In Figure 9c, a white fiber fishing rope exists in the image, but it was included in the dataset without confirming whether a white foam buoy was included in the image processing stage. Therefore, it was misclassified as a white foam buoy. In the case of Figure 9d, the size of the white foam buoy in the image was small; therefore, it was misclassified. In the case of Figure 9e, the image of the black film plastic bags did not clearly show the characteristic points owing to the color; therefore, it was confirmed that the wrong classification was made. In the case of Figure 9g,h, the images are similar in shape to the white foam buoy and transparent fragment plastic bottles, confirming that they were misclassified as foam and fragment. As shown in Figure 9i,j,l, the size of the object in the image is too small to be misclassified. In addition, the objects in Figure 9i,j appear to be small and more affected by the land. In the case of Figure 9m, the shape of the transparent fragment plastic bottles in the image appears to be similar to that of the rope; therefore, it appears that the misclassification occurred because of the morphological characteristics. In the case of Figure 9n, it was confirmed that the shape of the transparent fragment plastic bottles, along with the effect of sun glint, caused them to be misclassified, as they showed a form similar to that of the white film plastic bags. In Figure 9o, it was confirmed that the characteristics of the seawater shown in the image were noticeably more significant than the characteristics of the FMML. Accordingly, misclassification was confirmed. Therefore, our findings showed that carefully controlled drones and manually removed images need to be considered during the image acquisition process that in which external factors (i.e., sun glint and seawater quality) have a high impact through the qualitative evaluation of well-classified and misclassified cases.

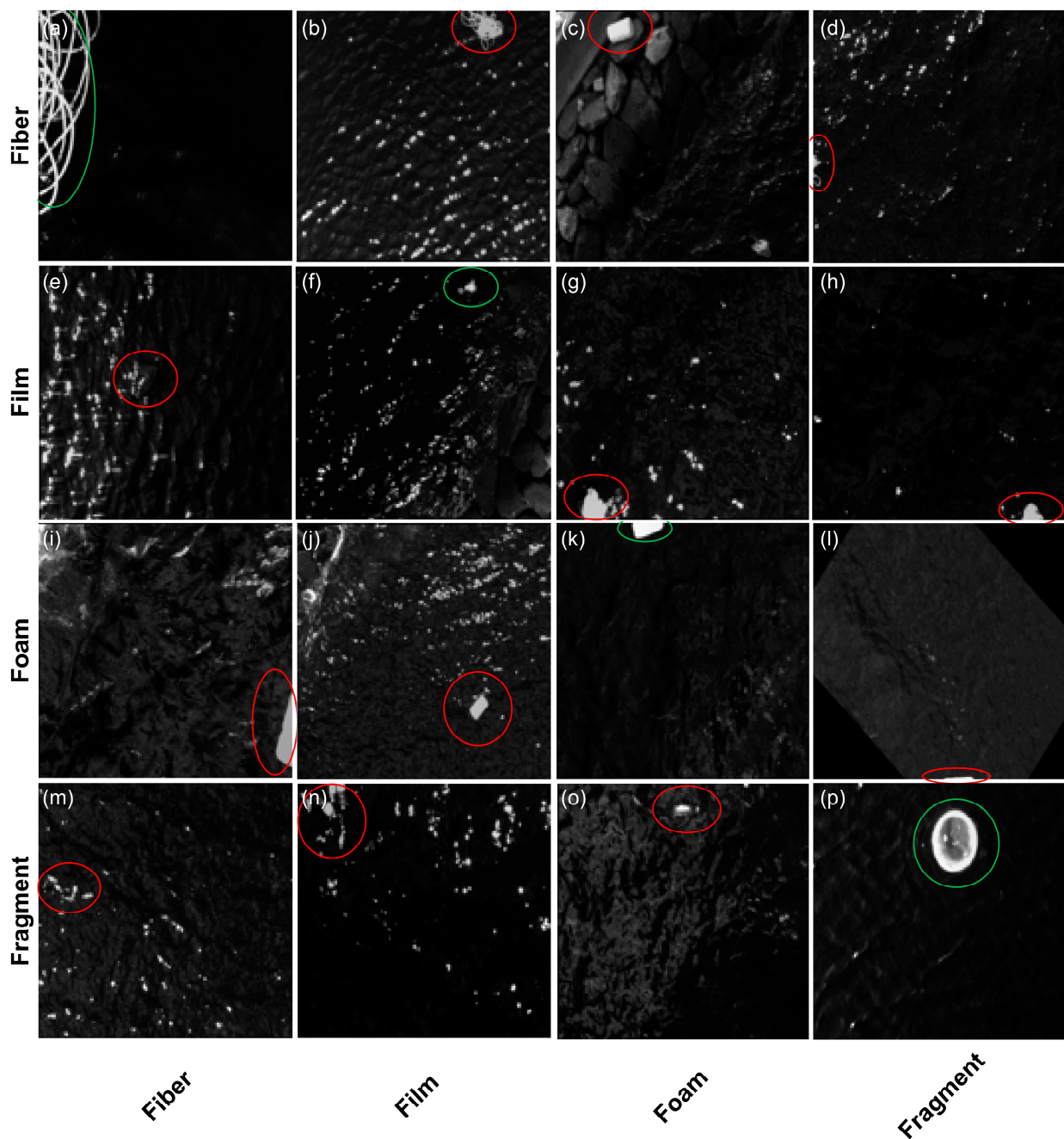


Figure 9. The well-classified and misclassified results of each category in the CNN-3 Model. All the images are Micasense multi-spectral images of band five. (a–d) Classified as fiber. (e–h) Classified as film. (i–l) Classified as foam. (m–p) Classified as fragment. Green and red circles indicate well-classified and misclassified results, respectively.

5. Conclusions

In this study, our CNN model was developed for different sizes, types, and colors of FMML classification from drone imagery with a multi-spectral sensor. The main findings of this study are as follows:

- (1) The atmospheric correction of multi-spectral images captured at seven flight altitudes revealed a significant reduction in atmospheric signals, showing only a reflectance deviation of 0.71 to 1.31%.
- (2) The FMML reflectance revealed distinct spectra variations in the different colors and types for each FMML. We trained and tested three CNN models to classify

the reflectance of each FMML. As a result, by leveraging these characteristics, we successfully developed the CNN-3 model with an accuracy and F1 score of 0.90 and 0.89, which showed the highest performance. We confirmed that the CNN model trained by reflectance classified different colors and types of FMML well.

- (3) We qualitatively analyzed the results of the CNN-3 model using Grad-CAM and visually represented the results obtained from the CNN-3 model. Most of the results were well classified, despite the conditions of turbid water and high-altitude images.

The novel findings presented in this study represent a significant advancement in marine litter detection, with three key contributions: (1) the atmospheric correction of multi-spectral images improved the reflectance accuracy across different flight altitudes, (2) the optimized CNN-3 model effectively classified FMML by leveraging distinct spectral characteristics, and (3) the application of Grad-CAM enhanced visualization, enabling the precise identification of FMML colors and types, even under challenging environmental conditions. These results highlight the potential of combining drone-based multi-spectral sensors with deep learning techniques to achieve accurate and efficient marine litter monitoring. The classified FMML data generated in this study offer critical insights that can support strategies to reduce marine litter and mitigate its impacts along coastal regions.

Author Contributions: Conceptualization, Y.J., J.S. and Y.-H.J.; methodology, Y.J., J.S. and Y.-H.J.; software, D.S.; validation, Y.J., J.S., D.S. and Y.-H.J.; formal analysis, Y.J.; investigation, Y.J., J.-S.L., J.-Y.B. and S.-Y.K.; resources, Y.-H.J.; data curation, Y.J., J.S. and Y.-H.J.; writing—original draft preparation, Y.J.; writing—review and editing, Y.J., J.S., J.-S.L., J.-Y.B., D.S. and Y.-H.J.; visualization, Y.J.; supervision, J.S., D.S. and Y.-H.J.; project administration, Y.J. and Y.-H.J.; funding acquisition, J.-Y.J. and Y.-H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project titled “Establishment of the Ocean Research Stations in the Jurisdiction Zone and Convergence Research” from the Ministry of Oceans and Fisheries, South Korea (grant number RS-2021-KS211502) and “Analysis for Subsurface Physical Oceanographic Meso-scale Phenomena based on Three Dimensional Data derived from Satellite Observations”, by the National Research Foundation of Korea (NRF), funded by the Korea Government (grant number RS-2023-00280650).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank ReSe Applications LLC for providing DROA-COR 2.0.1 software.

Conflicts of Interest: Author Daniel Schläpfer is affiliated with the company ReSe Applications LLC. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. UNEP. *Marine Plastic Debris and Microplastics: Global Lessons and Research to Inspire Action and Guide Policy Change*; United Nations Environment Programme: Nairobi, Kenya, 2016.
2. Galgani, F.; Hanke, G.; Werner, S.; De Vrees, L. Marine litter within the European Marine Strategy Framework Directive. *ICES J. Mar. Sci.* **2013**, *70*, 1055–1064. [[CrossRef](#)]
3. GESAMP. *Guidelines for the Monitoring and Assessment of Plastic Litter in the Ocean*; Rep. Stud. GESAMP No. 99; GESAMP: London, UK, 2019.
4. Commission, E.; Centre, J.R.; Vighi, M.; Ruiz-Orejón, L.; Hanke, G. *Monitoring of Floating Marine Macro Litter—State of the Art and Literature Overview*; Rep. EUR 31111 EN; Publications Office of the European Union: Luxembourg, 2022.
5. Tomás, J.; Guitart, R.; Mateo, R.; Raga, J.A. Marine debris ingestion in loggerhead sea turtles, *Caretta caretta*, from the Western Mediterranean. *Mar. Pollut. Bull.* **2002**, *44*, 211–216. [[CrossRef](#)] [[PubMed](#)]
6. De Stephanis, R.; Giménez, J.; Carpinelli, E.; Gutierrez-Exposito, C.; Cañadas, A. As main meal for sperm whales: Plastics debris. *Mar. Pollut. Bull.* **2013**, *69*, 206–214. [[CrossRef](#)] [[PubMed](#)]
7. Schuyler, Q.; Hardesty, B.D.; Wilcox, C.; Townsend, K. Global analysis of anthropogenic debris ingestion by sea turtles. *Conserv. Biol.* **2013**, *28*, 129–139. [[CrossRef](#)] [[PubMed](#)]
8. Domènech, F.; Aznar, F.J.; Raga, J.A.; Tomás, J. Two decades of monitoring in marine debris ingestion in loggerhead sea turtle, *Caretta caretta*, from the western Mediterranean. *Environ. Pollut.* **2019**, *244*, 367–378. [[CrossRef](#)]

9. Garcia-Garin, O.; Vighi, M.; Aguilar, A.; Tsangaris, C.; Digka, N.; Kaberi, H.; Borrell, A. Boops boops as a bioindicator of microplastic pollution along the Spanish Catalan coast. *Mar. Pollut. Bull.* **2019**, *149*, 110648. [CrossRef]
10. Garcia-Garin, O.; Vighi, M.; Sala, B.; Aguilar, A.; Tsangaris, C.; Digka, N.; Kaberi, H.; Eljarrat, E.; Borrell, A. Assessment of organophosphate flame retardants in Mediterranean Boops boops and their relationship to anthropization levels and microplastic ingestion. *Chemosphere* **2020**, *252*, 126569. [CrossRef]
11. Garcia-Garin, O.; Aguilar, A.; Vighi, M.; Víkingsson, G.A.; Chosson, V.; Borrell, A. Ingestion of synthetic particles by fin whales feeding off western Iceland in summer. *Chemosphere* **2021**, *279*, 130564. [CrossRef]
12. Heo, N.W.; Hong, S.H.; Han, G.M.; Hong, S.; Lee, J.; Song, Y.K.; Jang, M.; Shim, W.J. Distribution of small plastic debris in cross-section and high strandline on Heungnam beach, South Korea. *Ocean Sci. J.* **2013**, *48*, 225–233. [CrossRef]
13. Jang, Y.C.; Lee, J.; Hong, S.; Lee, J.S.; Shim, W.J.; Song, W.J. Sources of plastic marine debris on beaches of Korea: More from the ocean than the land. *Ocean Sci. J.* **2014**, *49*, 151–162. [CrossRef]
14. Lee, J.; Lee, J.S.; Jang, Y.C.; Hong, S.Y.; Shim, W.J.; Song, Y.K.; Hong, S.H.; Jang, M.; Han, G.M.; Kang, D.; et al. Distribution and Size Relationships of Plastic Marine Debris on Beaches in South Korea. *Arch. Environ. Contam. Toxicol.* **2015**, *69*, 288–298. [CrossRef] [PubMed]
15. Pichel, W.G.; Veenstra, T.S.; Churnside, J.H.; Arabini, E.; Friedman, K.S.; Foley, D.G.; Brainard, R.E.; Kiefer, D.; Ogle, S.; Clemente-Colón, P.; et al. GhostNet marine debris survey in the Gulf of Alaska—Satellite guidance and aircraft observations. *Mar. Pollut. Bull.* **2012**, *65*, 28–41. [CrossRef] [PubMed]
16. Guirado, E.; Tabik, S.; Rivas, M.L.; Alcaraz-Segura, D.; Herrera, F. Whale counting in satellite and aerial images with deep learning. *Sci. Rep.* **2019**, *9*, 14259. [CrossRef] [PubMed]
17. Topouzelis, K.; Papakonstantinou, A.; Garaba, S.P. Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018). *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *79*, 175–183. [CrossRef]
18. Biermann, L.; Clewley, D.; Martínez-Vicente, V.; Topouzelis, K. Finding Plastic Patches in Coastal Waters using Optical Satellite Data. *Sci. Rep.* **2020**, *10*, 5364. [CrossRef]
19. Moy, K.; Neilson, B.; Chung, A.; Meadows, A.; Castrence, M.; Ambagis, S.; Davidson, K. Mapping coastal marine debris using aerial imagery and spatial analysis. *Mar. Pollut. Bull.* **2018**, *132*, 52–59. [CrossRef]
20. Garcia-Garin, O.; Monleón-Getino, T.; López-Brosa, P.; Borrell, A.; Aguilar, A.; Borja-Robalino, R.; Cardona, L.; Vighi, M. Automatic detection and quantification of floating marine macro-litter in aerial images: Introducing a novel deep learning approach connected to a web application in R. *Environ. Pollut.* **2021**, *273*, 116490. [CrossRef]
21. Brooke, S.; Graham, D.; Jacobs, T.; Littnan, C.; Manuel, M.; O’Conner, R. Testing marine conservation applications of unmanned aerial systems (UAS) in a remote marine protected area. *J. Unmanned Veh. Syst.* **2015**, *3*, 237–251. [CrossRef]
22. Garcia-Garin, O.; Borrell, A.; Aguilar, A.; Cardona, L.; Vighi, M. Floating marine macro-litter in the North Western Mediterranean Sea: Results from a combined monitoring approach. *Mar. Pollut. Bull.* **2020**, *159*, 111467. [CrossRef]
23. Gonçalves, G.; Andriolo, U.; Pinto, L.; Duarte, D. Mapping marine litter with Unmanned Aerial Systems: A showcase comparison among manual image screening and machine learning techniques. *Mar. Pollut. Bull.* **2020**, *155*, 111158. [CrossRef]
24. Gonçalves, G.; Andriolo, U. Operational use of multispectral images for macro-litter mapping and categorization by Unmanned Aerial Vehicle. *Mar. Pollut. Bull.* **2022**, *176*, 113431. [CrossRef]
25. Pinto, J.O.; O’Sullivan, D.; Taylor, S.; Elston, J.; Baker, C.B.; Hotz, D.; Marshall, C.; Jacob, J.; Barfuss, K.; Pigué, B.; et al. The Status and Future of Small Uncrewed Aircraft Systems (UAS) in Operational Meteorology. *Bull. Am. Meteor. Soc.* **2021**, *102*, E2121–E2136. [CrossRef]
26. Wolf, M.; van den Berg, K.; Garaba, S.P.; Gnan, N.; Sattler, K.; Stahl, F.; Zielinski, O. Machine learning for aquatic plastic litter detection, classification and quantification (APLastic-Q). *Environ. Res. Lett.* **2020**, *15*, 114042. [CrossRef]
27. Cortesi, I.; Masiero, A.; Tucci, G.; Topouzelis, K. UAV-based river plastic detection with a multispectral camera. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2022**, *XLIII-B3-2022*, 855–862. [CrossRef]
28. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
29. Bak, S.H.; Hwang, D.H.; Kim, H.M.; Yoon, H.J. Detection and monitoring of beach litter using uav image and deep neural network. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2019**, *42*, 55–58. [CrossRef]
30. Takaya, K.; Shibata, A.; Mizuno, Y.; Ise, T. Unmanned aerial vehicles and deep learning for assessment of anthropogenic marine debris on beaches on an island in a semienclosed sea in Japan. *Environ. Res. Commun.* **2022**, *4*, 015003. [CrossRef]
31. Yamashita, R.; Nishio, M.; Do, R.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [CrossRef]
32. Kylili, K.; Kyriakides, L.; Artusi, A.; Hadjistassou, C. Identifying floating plastic marine debris using a deep learning approach. *Environ. Sci. Pollut. Res.* **2019**, *26*, 17091–17099. [CrossRef]
33. Freitas, S.; Silva, H.; Silva, E. Remote Hyperspectral Imaging Acquisition and Characterization for Marine Litter Detection. *Remote Sens.* **2021**, *13*, 2536. [CrossRef]
34. MOF (Ministry of Ocean Fisheries). National Coastal Litter Monitoring Program. 2021. Available online: <https://www.koem.or.kr/common/board/Download.do?bcIdx=31247&cbIdx=370&fileNo=1> (accessed on 31 December 2021).
35. Schläpfer, D.; Popp, C.; Richter, R. Drone Data Atmospheric Correction Concept for Multi- and Hyperspectral Imagery—The DROACOR Model. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2020**, *43*, 473–478. [CrossRef]

36. Karunaratne, S.; Thomson, A.; Morse-McNabb, E.; Wijesingha, J.; Stayches, D.; Copland, A.; Jacobs, J. The fusion of spectral and structural datasets derived from an airborne multispectral sensor for estimation of pasture dry matter yield at paddock scale with time. *Remote Sens.* **2020**, *12*, 2017. [[CrossRef](#)]
37. Andriolo, U.; Topouzelis, K.; van Emmerik, T.H.; Papakonstantinou, A.; Monteiro, J.G.; Isobe, A.; Hidaka, M.; Kako, S.I.; Kataoka, T.; Gonçalves, G. Drones for litter monitoring on coasts and rivers: Suitable flight altitude and image resolution. *Mar. Pollut. Bull.* **2023**, *195*, 115521. [[CrossRef](#)] [[PubMed](#)]
38. Joint Research Centre (European Commission); Fleet, D.; Vlachogianni, T.; Hanke, G. *Joint List of Litter Categories for Marine Macro-Litter Monitoring—Manual for the Application of the Classification System*; Rep. EUR 30348 EN; Publications Office of the European Union: Luxembourg, 2021.
39. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
40. Tekerek, A.; Yapici, M.M. A novel malware classification and augmentation model based on convolutional neural network. *Comput. Secur.* **2022**, *112*, 102515. [[CrossRef](#)]
41. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
42. Zhou, B.; Khosla, A.; Lapedriza, À.; Oliva, A.; Torralba, A. Learning Deep Features for Discriminative Localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016. [[CrossRef](#)]
43. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359. [[CrossRef](#)]
44. Panwar, H.; Gupta, P.K.; Siddiqui, M.K.; Morales-Menendez, R.; Singh, V. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos Solitons Fractals* **2020**, *138*, 109944. [[CrossRef](#)]
45. Kohavi, R. Glossary of terms. *Mach. Learn.* **1998**, *30*, 271–274.
46. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
47. Bekkar, M.; Djemaa, H.K.; Alitouche, T.A. Evaluation measures for models assessment over imbalanced data sets. *J. Inf. Eng. Appl.* **2013**, *3*.
48. Cai, C.; Gu, S. Research on Marine Garbage Detection Based on Improved Yolov5 Model. *J. Phys. Conf. Ser.* **2022**, *2405*, 012008. [[CrossRef](#)]
49. Zhang, Y.; Yin, Y.; Shao, Z. An Enhanced Target Detection Algorithm for Maritime Search and Rescue Based on Aerial Images. *Remote Sens.* **2023**, *15*, 4818. [[CrossRef](#)]
50. Zhu, Q.; Ma, K.; Wang, Z.; Shi, P. YOLOv7-CSAW for maritime target detection. *Front. Neurobot.* **2023**, *17*, 1210470. [[CrossRef](#)]
51. De Keukelaere, L.; Moelans, R.; Knaeps, E.; Sterckx, S.; Reusen, I.; De Munck, D.; Simis, S.G.; Constantinescu, A.M.; Scricciu, A.; Katsouras, G.; et al. Airborne drones for water quality mapping in inland, transitional and coastal waters—MapEO water data processing and validation. *Remote Sens.* **2023**, *15*, 1345. [[CrossRef](#)]
52. Lee, J.S.; Baek, J.Y.; Shin, J.; Kim, J.S.; Jo, Y.H. Suspended Sediment Concentration Estimation along Turbid Water Outflow Using a Multispectral Camera on an Unmanned Aerial Vehicle. *Remote Sens.* **2023**, *15*, 5540. [[CrossRef](#)]
53. Escobar-Sánchez, G.; Markfort, G.; Berghald, M.; Ritzenhofen, L.; Schernewski, G. Aerial and underwater drones for marine litter monitoring in shallow coastal waters: Factors influencing item detection and cost-efficiency. *Environ. Monit. Assess.* **2022**, *194*, 863. [[CrossRef](#)]
54. Balsi, M.; Moroni, M.; Chiarabini, V.; Tanda, G. High-resolution aerial detection of marine plastic litter by hyperspectral sensing. *Remote Sens.* **2021**, *13*, 1557. [[CrossRef](#)]
55. Musthafa M, M.; Mahesh, T.R.; Kumar V, V.; Guluwadi, S. Enhancing brain tumor detection in MRI images through explainable AI using Grad-CAM with Resnet 50. *BMC Med. Imaging* **2024**, *24*, 39. [[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.