



Identifying floating plastic marine debris using a deep learning approach

Kyriaki Kylii¹ · Ioannis Kyriakides¹ · Alessandro Artusi² · Constantinos Hadjistassou¹

Received: 21 December 2018 / Accepted: 9 April 2019 / Published online: 18 April 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Estimating the volume of macro-plastics which dot the world's oceans is one of the most pressing environmental concerns of our time. Prevailing methods for determining the amount of floating plastic debris, usually conducted manually, are time demanding and rather limited in coverage. With the aid of deep learning, herein, we propose a fast, scalable, and potentially cost-effective method for automatically identifying floating marine plastics. When trained on three categories of plastic marine litter, that is, bottles, buckets, and straws, the classifier was able to successfully recognize the preceding floating objects at a success rate of $\approx 86\%$. Apparently, the high level of accuracy and efficiency of the developed machine learning tool constitutes a leap towards unraveling the true scale of floating plastics.

Keywords Image classification · Convolutional Neural Networks · Data processing · Deep learning · Marine debris · Plastics · Monitoring

Introduction

One of the biggest environmental challenges currently threatening our ecosystems is marine debris and especially pollution from plastics (Eriksen et al. 2014; Suaria and Aliani 2014). Plastics are entering the marine environment through various channels, such as being carried by wind, shipping, coastal activities, or through effluent water discharge (Woodall et al. 2014). Consequently, the world's oceans contain an enormous amount of plastic debris. More precisely, the five subtropical gyres harbor the largest accumulations of plastics in comparison to other ocean zones (Cózar et al. 2014). The Great Pacific Gyre alone, also known as the Great Pacific Garbage Patch (GPGP), is believed to contain at least 79 thousand tons of marine plastics that are floating in an area of 1.6 million km² (Lebreton et al. 2018)—roughly the size of Mongolia.

Boat surveys constitute the primary means for retrieving information about plastics floating on (or near) the water surface, such as, estimating their type and density (Barnes et al. 2009). Meanwhile, vessel surveys are time demanding, expensive, and limited in coverage because of the subsequent processing of the material and data that are collected, notwithstanding the risk of tainting observations with human errors. Attempting to identify floating debris, other researchers have proposed a sensing platform fitted with an onboard camera (Wang et al. 2015). However, devising an automated way for detecting and classifying marine debris, grounded on a deep learning approach, offers some added advantages. Such a method holds the potential to save valuable search time, lower expedition costs, and increase the level of accurately identifying large-size plastic debris. To automate the process of macro-plastic detection, we employ a machine learning method, like deep learning, which can yield a high level of accuracy.

Deep learning, which belongs to the family of machine learning methods, constitutes a tool that powers many aspects of today's society and finds applications ranging from content filtering in web searches to object identification and speech recognition. Domains of the scientific community, the business world, and governments are utilizing deep learning techniques mainly owing to their excellent performance and ability to conduct novel tasks.

Responsible editor: Philippe Garrigues

✉ Constantinos Hadjistassou
Hadjistassou.c@unic.ac.cy

¹ Marine & Carbon Lab, Department of Engineering, University of Nicosia, 46 Makedonitissas Avenue, CY-2417 Nicosia, Cyprus

² KIOS Research Center, University of Cyprus, CY-1678 Nicosia, Cyprus

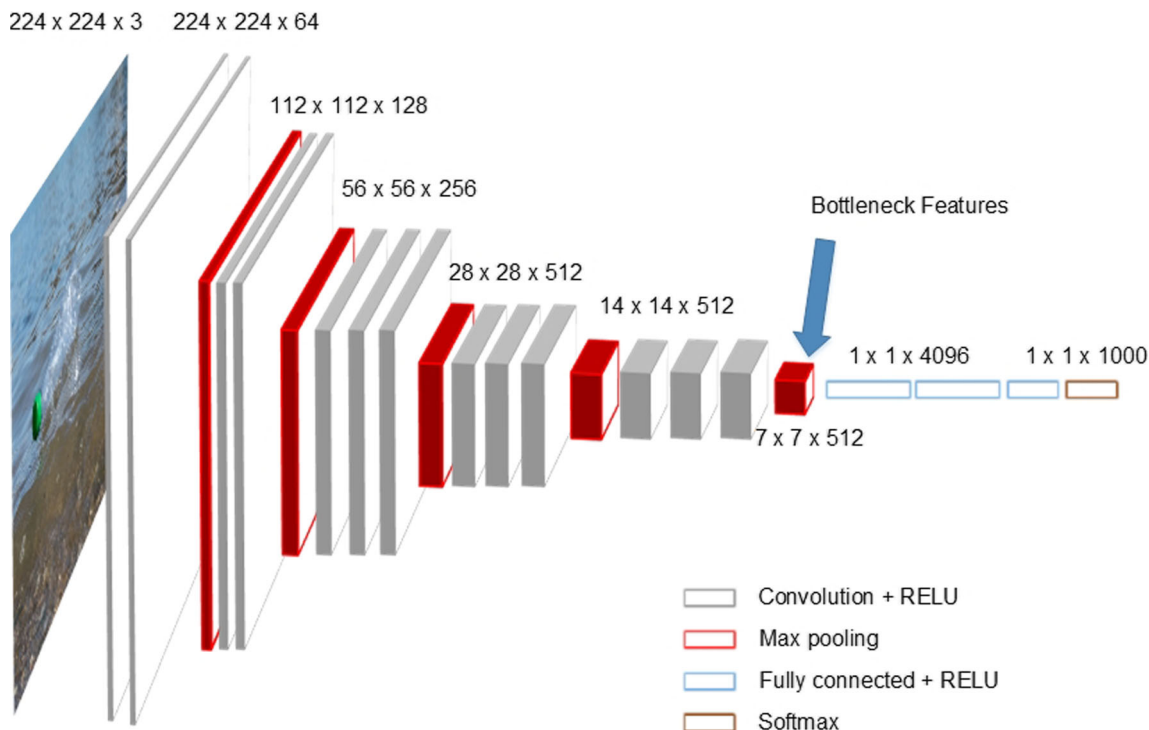


Fig. 1 The VGG16 model architecture as applied to the bottleneck method. The image on the left depicts a floating bottle (middle) in seawater. Algorithm VGG16 accepts as input RGB images of marine debris

of resolution 224×224 . Figures 4096 and 1000 denote the number of nodes while the neural network executes 3×3 convolutions

Applying deep learning methods in the context of marine debris research can help create faster and more accurate tools (LeCun et al. 2015) for detecting and classifying floating litter.

Several methods are prevalent in sampling macro-plastics, measuring > 20 mm in diameter (Barnes et al. 2009), from the marine environment. Among them are visual observations from boat expeditions which aim to pinpoint the areas with the largest accumulation of marine macroplastic (Chambault et al. 2018). Other surveys, such as aerial scans, gather airborne imagery from seashores including remote islands. The collected aerial visual material is pieced together to construct mosaics of images which help manually identify coastal debris bigger than 0.5 m^2 (Moy et al. 2018). More recently, the deployment of drones for conducting airborne surveys has gained considerable traction. Drones are suitable for surveying small coastal areas of remote locations as a way of visually sampling and monitoring the abundance of plastic litter (Deidun et al. 2018; Hengstmann et al. 2017).

The implementation of machine learning techniques in search of marine debris is a rather new research direction. According to the directive on monitoring marine litter in the European Seas (MSFD Technical Subgroup on Marine Litter 2013), the development of computerized camera surveys is considered of low maturity and is still at the

development stage. Other initiatives, such as the Joint Research Center (JRC) Floating Litter Monitoring

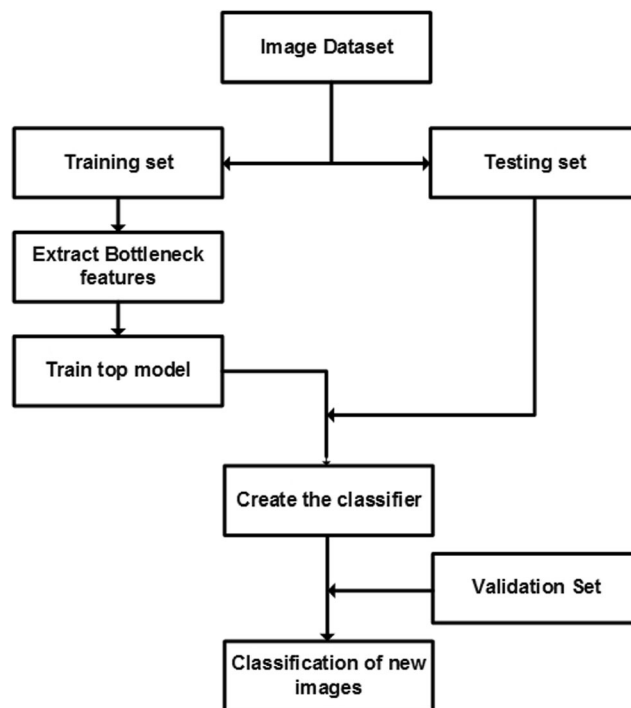


Fig. 2 Flowchart depicting the overall structure of the transfer learning approach proposed in this research

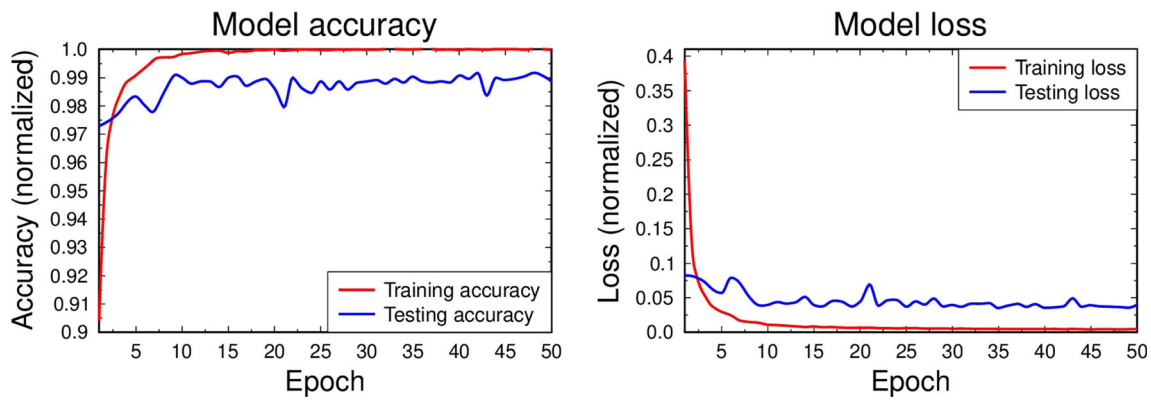


Fig. 3 Plots display the training and test accuracy (left) and respective loss (right) over 50 epochs

Application, are a software tool that collects data of debris in river estuaries. Users of this application can capture pictures of litter located at a river from an elevated position (e.g., a bridge) and tag the object class from a drop-down list of items and size categories (González-Fernández and Hanke 2017). To the best of our knowledge, our paper puts forward one of the first attempts to implement a machine learning method for detecting and classifying floating macro-plastic debris.

Herein, we propose a Convolutional Neural Network (CNN) approach that is able to train itself on images of plastic objects larger than a few centimeters, also known as, macro-plastics. Subsequently, the method can automatically predict the class of new images of macro-plastic objects floating at sea with very high accuracy. The underlying intention was to create a prototype device to be mounted onboard marine vessels. With the aid of a camera, the system would be able to scan the sea surface and automatically detect and recognize floating macro-plastic debris without calibrating the images geometrically. Considering the sheer size of our oceans and the large volume of macro-plastic debris, it is imperative that intelligent systems take over the gargantuan task of localizing floating plastic marine debris rather than relying on manual and time-consuming means.

Materials and methods

An unconventional machine learning tool, like deep learning, is a representation-learning method that feeds a computer with raw data and allows the automatic discovery of the representations needed for classification or detection purposes. This depiction mastering method exhibits multiple levels of abstraction, which consist of relatively simple but non-linear modules (Simonyan and Zisserman 2014). Each module transforms the representation at one level and feeds the raw data into a depiction at a higher and more abstract level. Interestingly, these manipulations allow the

method to undertake more complex functions and tasks (LeCun et al. 2015). The research study described in this paper utilizes a CNN architecture as suggested recently (Kylili et al. 2018). The CNN is a deep, feed-forward network that trains easily generalized networks with full connectivity between adjacent layers (LeCun et al. 1990; LeCun et al. 1998). The layers of the CNN architecture are divided into two types: (a) convolutional layers and (b) pooling layers. In the convolutional layer, the units are organized into feature maps, within which each unit is connected to local patches in the feature maps of the previous layer through a set of weights. All feature maps are sharing the same set of weights. As for the pooling layers, they compute the maximum of a local patch of units in one feature map (LeCun et al. 2015).

The CNN architecture used in this paper is the Visual Geometry Group-16 (VGG16) model pre-trained on the large-scale ImageNet dataset (Simonyan and Zisserman 2014). Due to this pre-training, the model has learned certain image features from this large dataset. Making use of transfer learning, these assimilated image features are applied to the debris image classification problem. This is achieved by removing the fully connected layer from the VGG16 model while training it by means of an image dataset related to our classification problem. The new image’ attributes, called bottleneck features, are then used to train the fully connected layer to tackle the specific debris classification problem described herein.

Figure 1 depicts the level of the VGG16 model at which the bottleneck features are stored from an image of a plastic bottle floating at sea. The gray color layer (Fig. 1) represents the convolutional layer while the Rectified Linear

Table 1 Summary of the levels of accuracy attained by the CNN

Results	
Training accuracy (%)	≈ 100
Test accuracy (%)	≈ 99
Validation accuracy (%)	≈ 86

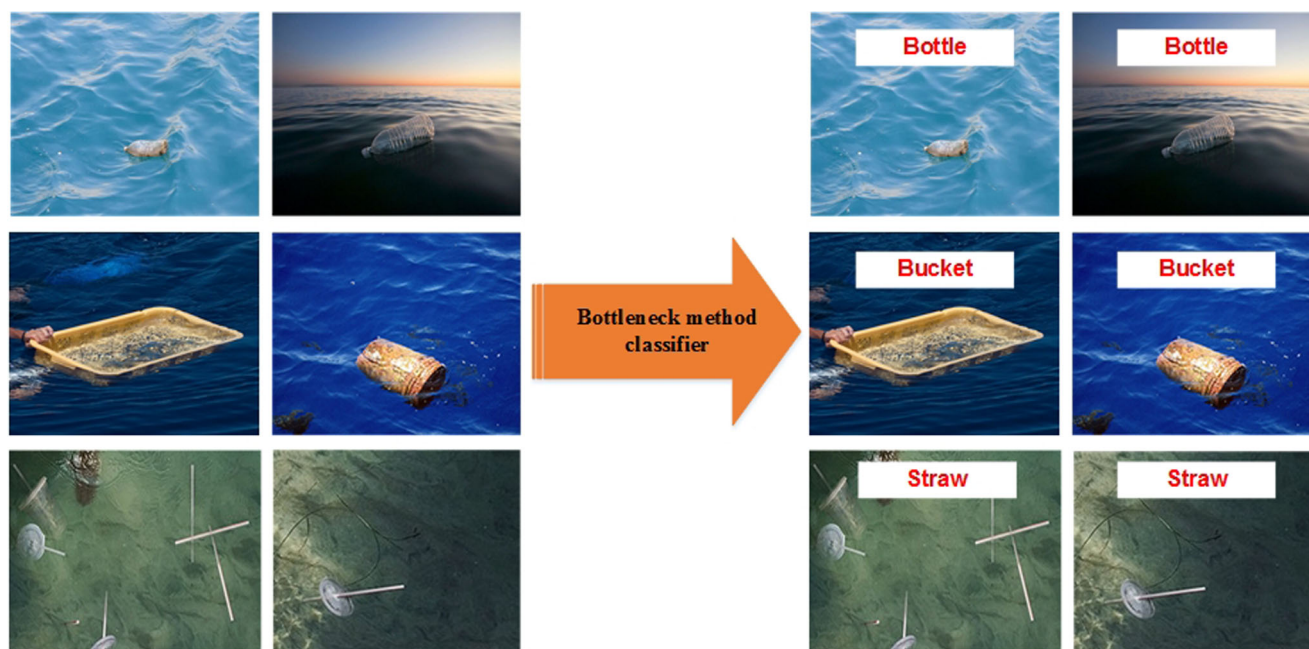


Fig. 4 Matrix of images demonstrates the ability of the bottleneck method classifier to correctly recognize plastic bottles, buckets, and straws

Unit (RELU), which constitutes its activation function, is given by:

$$f(x) = \max(0, x) \tag{1}$$

where x is the output weight. Activation functions are essential in neural networks as they endow non-linear properties to a neural system. The red color layer displays the maximum pooling type of layers where the sample-based discretization process occurs. The light blue color indicates the fully connected layers with the RELU being their activation function (Fig. 1). Finally, the brown color refers to the layer within which the Softmax function is implemented. Here, the Softmax function determines the probability of each target class over all of the possible target classes C , such that:

$$f(x)_i = \frac{e^{x_i}}{\sum_j^C e^{x_j}} \tag{2}$$

The marine debris image classification problem presented in this work distinguishes between three categories of litter, namely, plastic bottles, plastic buckets, and plastic straws. Samples of these images will be presented in the “Results and discussion” section. One of the major challenges of the real-world image classification problem is the need for a large enough dataset key to the training process. Typically, several thousand of images are used in the training process, which lend a high level of accuracy to the classification approach (Sun et al. 2017). However, in real-world applications such as the one presented in this paper, large datasets are not always available. For instance, each of the three categories used in this research consisted of 250 images, hence, making it difficult to recognize marine debris with an extremely high level of classification accuracy. To overcome this issue, we have used data augmentation. That is, the number of images is increased through the adoption of geometrical transformations, e.g., zooming, shifting, flipping, rotation, etc. In other words, using data augmentation, we can populate a larger

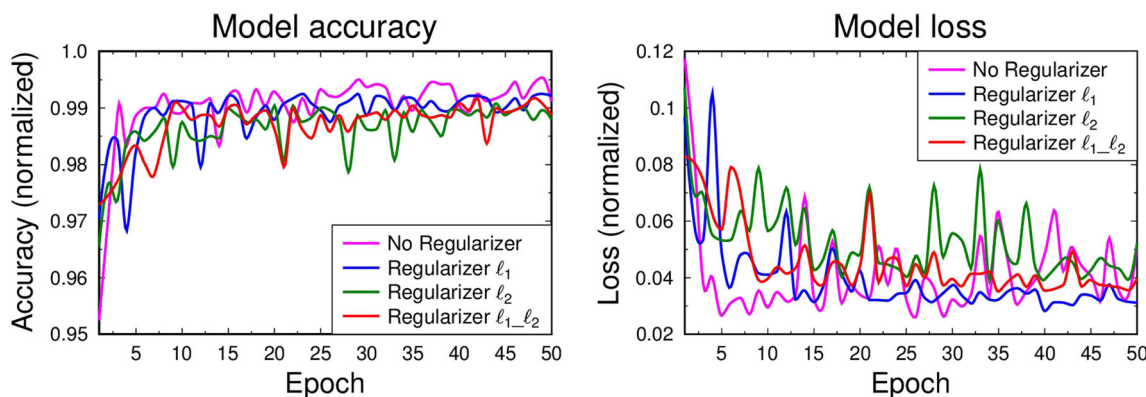


Fig. 5 The performance of the marine debris classifier in the absence of a regularizer and featuring different regularizers

database emanating from a relatively small image collection. In aggregate, the augmented dataset reached a total of 4000 images in each category.

Subsequently, this large dataset was randomly split in two subsets: (a) the training set and (b) the test set which consisted of 3200 and 800 images per category, respectively. As its name implies, the training set was used to train the model and the test set to test the model’s performance. During the model’s training stage, a set of weights was generated and employed in the validation process. For the record, the validation phase assesses the generalized performance capabilities of the classifier. It accepts as input new images of plastic marine debris drawn from the validation set with the ultimate goal of identifying all input images. Simply put, the classifier created from this method can recognize images that the trained model never encountered or trained on them before. The preceding process is summarized in Fig. 2.

Results and discussion

The classifier presented in the previous section was trained on 9600 images (training set) and tested on 2400 images (test set). The performance of the classifier was assessed in terms of the training accuracy, the test accuracy, and the validation accuracy. Here, training accuracy refers to the accuracy of training the classifier on the training set, whereas the test accuracy relates to the accuracy from testing the classifier on the test set. Lastly, the validation accuracy is defined to be the rate of successfully classifying a newly seen image embedded in the validation set. Moreover, the respective losses were recorded and taken into account when determining the performance of our method. Figure 3 displays the performance of the model with respect to accuracy (left plot) and loss (right plot). The accuracy and loss results were obtained from the training and test sets over 50 epochs. The number of epochs indicates the times that the entire dataset is passed forward and backward through the network and this process refines the model weights so as to yield the best possible classification outcome. Converging trends between training accuracy and loss reinforces the credibility of results.

Computationally, the model took about 1 h to conclude on an Intel Xeon (CPU 2.40 GHz) processor with NVIDIA (Quadro K4200) graphics card. Collectively, this is the time necessary to complete the data augmentation process, to store the bottleneck features extracted from the training set and the test set, and finally, validate the classifier. The training accuracy of the model reached a maximum of $\approx 100\%$ with a loss of about 1%, while the test accuracy topped at $\approx 99\%$ with a loss of $\approx 4\%$, as illustrated in Table 1. The test loss constitutes a metric on how good the predictions of the model are. That is, the smallest the test loss, the more trustworthy model results are.

Table 2 Results of *case 1* which uses different regularizers, 4000 images/category of which 80% in the training set and 20% in the test set

Without a regularizer	Training accuracy	$\approx 99.97\%$
	Test accuracy	$\approx 99.12\%$
	Validation accuracy	86%
Regularizer ℓ_1	Training accuracy	$\approx 100\%$
	Test accuracy	$\approx 99.21\%$
	Validation accuracy	86%
Regularizer ℓ_2	Training accuracy	$\approx 100\%$
	Test accuracy	$\approx 98.79\%$
	Validation accuracy	82.42%
Regularizer $\ell_1_{\ell_2}$	Training accuracy	$\approx 99.97\%$
	Test accuracy	$\approx 98.87\%$
	Validation accuracy	85.45%

It is worth mentioning that the validation dataset consisted of 55 images of marine debris per category or a total of 165 validation samples. Interestingly, when the classifier was tested on the validation image set, it correctly identified 53 plastic bottles, 55 plastic buckets, and 33 plastic straws. Overall, recognizing a total of 141 out of 165 newly provided images of plastic marine debris resulted in a validation accuracy of $\approx 86\%$. Being relatively high, the validation accuracy lends credibility to the effectiveness of the proposed classifier. A selection of correctly traced (labeled) plastic marine debris images is illustrated in Fig. 4.

As part of rigorous effort to scrutinize the trustworthiness of our classifier, it was deemed necessary to alter some structural parameters of the algorithmic model. At such, three scenarios were examined. The first one compared different types of regularizers as applied to the context of the training process. Secondly, a parametric investigation explored the performance of the program (code) by varying the number of images

Table 3 Findings of *case 2* which utilizes a different number of images from the test set. It employs the $\ell_1_{\ell_2}$ regularizer with 4000 images/category

Twenty percent (20%) in test set	Training accuracy	$\approx 99.97\%$
	Test accuracy	$\approx 98.87\%$
	Validation accuracy	85.45%
Thirty percent (30%) in test set	Training accuracy	$\approx 99.98\%$
	Test accuracy	$\approx 98.22\%$
	Validation accuracy	83.63%
Forty percent (40%) in test set	Training accuracy	$\approx 99.99\%$
	Test accuracy	$\approx 98.4\%$
	Validation accuracy	83%
Fifty percent (50%) in test set	Training accuracy	$\approx 100\%$
	Test accuracy	$\approx 98.12\%$
	Validation accuracy	$\approx 80.6\%$

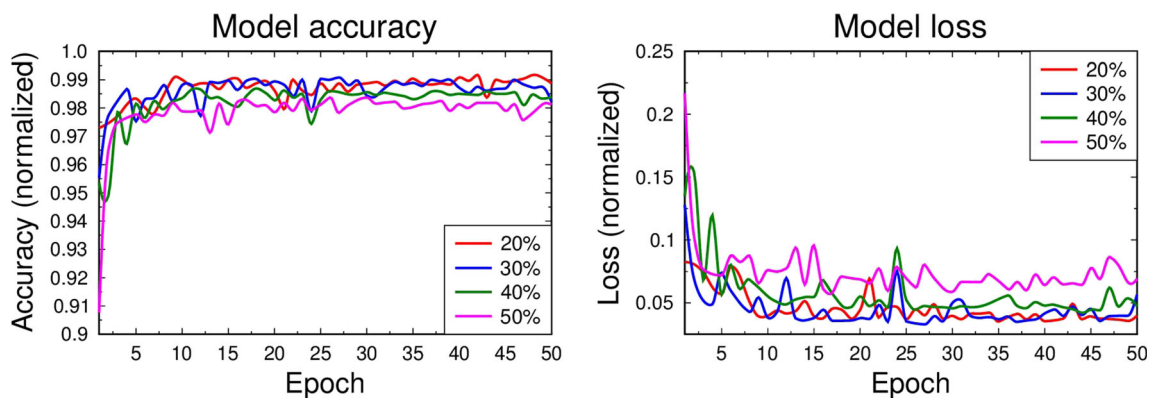


Fig. 6 Left graph displays the test accuracy, over 50 epochs, by increasing the proportion of the images used as test set from 20 to 50%, in increments of 10%. Right plot demonstrates the respective test loss, which increases for the same test set

in the test set. Finally, the performance of the classifier was tested as a function of the number of images generated from the data augmentation process.

The first case was designed to assess the capabilities of the classifier by utilizing different regularizers in the structure of our model which conducted the training process. Regularizers permit the assignment of “penalties” on the parameters of the model’s layers or on a layer’s activity during the optimization of the model. These penalties are added in the categorical cross-entropy loss function used to optimize the network parameters. The categorical cross-entropy loss function is described by:

$$\text{Cross-entropy loss} = \sum_{C=1}^M y_{O,C} \log(p_{O,C}) \tag{3}$$

where M is the number of image categories, y is a binary indicator (0 or 1) if class label C is the correct classification for observation O , and p is the probability observation O is of class C .

Accordingly, a particular regularizer can improve the performance of the classifier and lower the risk of encountering overfitting in the learning process. Overfitting is the flaw where the model embraces the details and noise in the training set, such that it adversely affects the performance of the model when processing new data. Consequently, the noise and random fluctuations in the training data are identified and conceptualized by the model which negatively impacts the model’s ability to generalize.

Overall, four cases were examined, namely, ℓ_1 , ℓ_2 , $\ell_1 \ell_2$, and the absence of a regularizer. Symbol ℓ_1 refers to Ridge regression (Hastie et al. 2013) that attaches a penalty on the weights similar to Laplace’s distribution. Regularizer ℓ_2 is the Lasso regression (Hastie et al. 2013) that forces the weights to adopt a Gaussian pattern. Denoted by $\ell_1 \ell_2$, this regularizer combined the two previous regression methods. Case four (4) was run without a regularizer. All cases utilized 4000 images, per category, as obtained from the data augmentation procedure. Of these 4000 images, 20% were dedicated to the test set. The performance of the proposed classifier using different regularizers is depicted in

Fig. 5. Although the case lacking a regularizer yields the highest test accuracy, yet its increasing loss over the epochs seems to point to overfitting. For that reason, this case was discarded from our investigation. The case which blended the Ridge and the Lasso regressions seems to perform better than using each regression method distinctively. Notably, regularizers $\ell_1 \ell_2$ and ℓ_1 produced the best performance results. When both the test accuracy and the loss curves were considered, regularizer $\ell_1 \ell_2$ was selected as it fared slightly better than regularizer ℓ_1 , as demonstrated in Table 2.

The second line of investigation varied the number of images of the test set. Specifically, as part of a parametric analysis, the percentage of the images fed into each test set comprised 20%, 30%, 40%, and 50% of each augmented image set. In all cases, the augmented data for each image category amounted to 4000 images whereas each case utilized regularizer $\ell_1 \ell_2$. For the 20% case, 3200 images were used as a training set and 800 images as a test set. In an analogous fashion, for the 30% case, the training set consisted of 2800 images while the training set included 1200 images. Correspondingly, for the 40% case, 2400 images made-up the training set and 1600 the test set. Finally, in the 50% case,

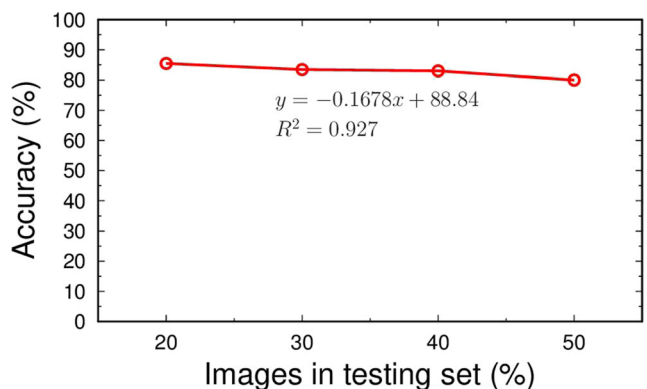


Fig. 7 Validation accuracy of the proposed classifier when recognizing new images of marine plastic debris. The equation of the straight line demonstrates the decreasing trend in the validation accuracy as the number of images in the training set is reduced

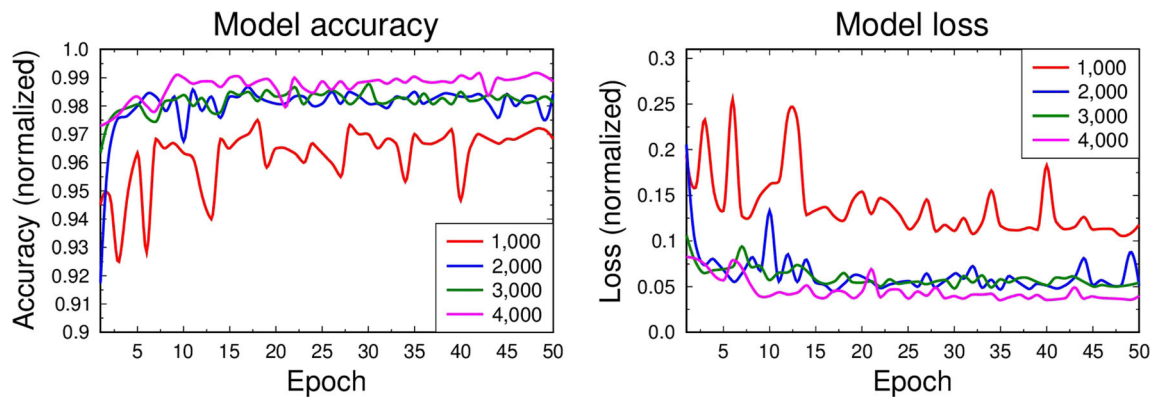


Fig. 8 The left graph indicates the test accuracy along progressing epochs for an increasing number of images ranging from 1000 to 4000, as generated by data augmentation. The right plot depicts the respective test losses

the number of images, as shown in Table 3, was equally split into 2000 in the training and 2000 in the test set.

The left graph in Fig. 6 shows the fluctuations in the testing accuracy over the number of epochs for all four cases which utilized 20%, 30%, 40%, and 50% of the augmented image set as a test set. Since the training accuracy for all cases reached a value of almost 100%, for comparison purposes, we focused on the test set accuracy. Likewise, because the training loss was almost negligible, it was decided to exclude these findings from Fig. 6.

As anticipated, the smallest the size of the image set, which made-up the test set, the better was the performance of the model. This result emerged because the training set had a larger pool of images. Further increase in the number of images in the test set led to a small drop in the test accuracy. This behavior was reasonable and expected as the method’s accuracy improves when the classifier processes more representations of images in the training set and trains on them.

The same observation is supported by Fig. 7 which displays the validation accuracy as a function of the number of images of the test set. Here, validation accuracy refers to the number of correctly identified images of plastic marine debris.

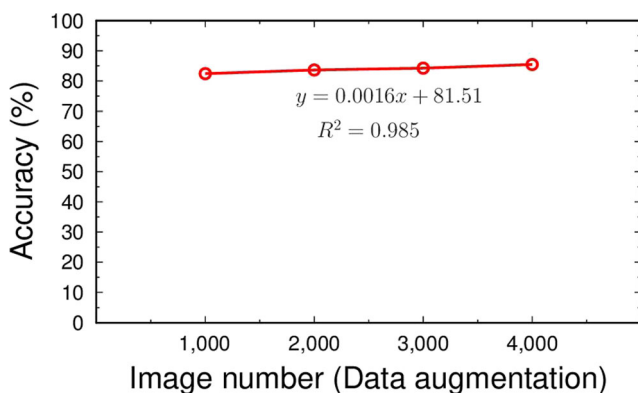


Fig. 9 Plot of classifier validation accuracy versus an increase in the number of images obtained from data augmentation. The case of 4000 images for each image category exhibits the highest validation accuracy of $\approx 86\%$

By fixing the fraction of images in the training set to 0.8 with the remaining 0.2 allocated to the test set, the validation accuracy reached its highest value at $\approx 86\%$. In other words, our method successfully identified 141 out of the 165 images that formed the validation set.

The third and final test was intended to define the capabilities of our method. To do so, the number of images was altered considering data augmentation manipulations, which expanded the image pool for each distinct image category: plastic straws, buckets, and bottles. As already mentioned in the “Materials and methods” section, for deep learning techniques to generate superior test and validation accuracy results compared to other methods, they must be trained on large datasets. To test this assertion, the number of images in each category was increased from 1000 to 4000 in increments of 1000. Each case was run with regularizer $\ell_1\text{-}\ell_2$ while the test scenario encompassed 20% of the augmented images, as illustrated in Fig. 8.

Referring to the left graph (Fig. 8), one can observe that the larger the number of images in the initial image

Table 4 Results obtained from case 3 which employs data augmentation to generate a variable number of images. It uses the $\ell_1\text{-}\ell_2$ regularizer and dedicates 80% of the images to the training set and 20% to the test set, respectively

4000 images	Training accuracy	$\approx 99.97\%$
	Test accuracy	$\approx 98.87\%$
	Validation accuracy	85.45%
3000 images	Training accuracy	$\approx 100\%$
	Test accuracy	$\approx 98.11\%$
	Validation accuracy	84.24%
2000 images	Training accuracy	$\approx 100\%$
	Test accuracy	$\approx 98.42\%$
	Validation accuracy	83.63%
1000 images	Training accuracy	$\approx 100\%$
	Test accuracy	$\approx 96.83\%$
	Validation accuracy	$\approx 82.42\%$

category, the better was the test accuracy of the method. Owing to the larger image set, the classifier becomes more efficient in discovering marine debris. Similarly, a comparable trend can be noticed in Fig. 9 where the validation accuracy improves as the number of images per category increases. The process of collecting unique marine images proved particularly tedious. Employing data augmentation helped partly alleviate the problem and improved the ability of the proposed classifier in identifying plastic debris. Still data processing limitations constrained the image pool to 4000 snapshots. Table 4 summarizes the results of the training accuracy and validation accuracy based on regularizer $\ell_1\text{-}\ell_2$.

Conclusions

Motivated by the need to tackle the problem of marine debris, we have applied the transfer learning method (Pan and Yang 2010) as a way of recognizing macro-plastic objects larger than a few centimeters, floating at the sea-water surface. Having parametrically tested the performance of the abovementioned method, an augmented dataset of 4000 marine debris images was generated as a preferable image population. Eighty percent of these images was used to construct the training set while the remaining 20% was reserved for the test set. Justified by its superior performance, regularizer $\ell_1\text{-}\ell_2$ was used to process new floating plastic images. Remarkably, the method yielded a training accuracy of 100% at a loss of 1%. Similarly, a test accuracy of $\approx 99\%$ was attained at a loss of $\approx 4\%$. Of major interest was the performance of our classifier, which when trained on the same three categories of plastic marine litter (plastic bottles, plastic buckets, and plastic straws) was able to successfully identify new plastic objects with a validation accuracy of $\approx 86\%$.

The deep learning method proposed herein can help automate the process of recognizing marine debris, which is currently mostly done manually. Not only is the method capable of enhancing the detection accuracy, it can also expedite the process of identifying plastic litter. In the nearby future, we plan to improve the competency of the classifier to realize an even higher classification accuracy as well as to distinguish between more classes of marine debris, such as, plastic bags and nets. An increase in the number of images is envisaged to boost the accuracy of the technique proposed herein to identify marine debris. Depending on the number of images, a trade-off is sought between the accuracy of the deep learning method and the computational time. Concluding, we are planning to validate the capabilities of the classifier under real sea-state conditions by building and installing a prototype device onboard a marine vessel.

Funding information K. K. acknowledges financial support from the Universitas Foundation, the A. G. Leventis Foundation, and the University of Nicosia. A. A.'s work is financially supported by the European Union's Horizon 2020 research and innovation program under grant agreement No 739551 (KIOS CoE) and from the Government of the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development.

References

- Barnes DKA, Galgani F, Thompson RC, Barlaz M (2009) Accumulation and fragmentation of plastic debris in global environments. *Philos Trans R Soc Lond B Biol Sci* 364:1985–1998
- Chambault P, Vandepierre F, Machete M, Lagoa JC, Pham CK (2018) Distribution and composition of floating macro litter off the Azores archipelago and Madeira (NE Atlantic) using opportunistic surveys. *Mar Environ Res* 141:225–232
- Cózar A, Echevarría F, González-Gordillo JJ, Irigoien X, Úbeda B, Hernández-León S, Palma ÁT, Navarro S, García-de-Lomas J, Ruiz A, Fernández-de-Puelles ML, Duarte CM (2014) Plastic debris in the open ocean. *Proc Natl Acad Sci* 111:10239–10244
- Deidun A, Gauci A, Lagorio S, Galgani F (2018) Optimising beached litter monitoring protocols through aerial imagery. *Mar Pollut Bull* 131:212–217
- Eriksen M, Lebreton LCM, Carson HS, Thiel M, Moore CJ, Borerro JC, Galgani F, Ryan PG, Reisser J (2014) Plastic pollution in the world's oceans: more than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea. *PLoS ONE* 9:1–15
- González-Fernández D, Hanke G (2017) Toward a harmonized approach for monitoring of riverine floating macro litter inputs to the marine environment. *Front Mar Sci* 4
- Hastie TJ, Tibshirani RJ, Friedman JH (2013) The elements of statistical learning: data mining, inference, and prediction. Springer, New York
- Hengstmann E, Gräwe D, Tamminga M, Fischer EK (2017) Marine litter abundance and distribution on beaches on the Isle of Rügen considering the influence of exposition, morphology and recreational activities. *Mar Pollut Bull* 115:297–306
- Kylili K, Artusi A, Kyriakides I, Hadjistassou C (2018) Tracking and identifying floating marine debris, 6th International Marine Debris Conference, San Diego, California, US
- Lebreton L, Slat B, Ferrari F, Sainte-Rose B, Aitken J, Marthouse R, Hajbane S, Cunsolo S, Schwarz A, Levivier A, Noble K, Debeljak P, Maral H, Schoeneich-Argent R, Brambini R, Reisser J (2018) Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. *Sci Rep* 8:4666
- LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Handwritten digit recognition with a back-propagation network, Proceedings of the 2nd International Conference on Neural Information Processing Systems. MIT Press, pp. 396–404
- LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86:2278–2324
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444
- Moy K, Neilson B, Chung A, Meadows A, Castrance M, Ambagis S, Davidson K (2018) Mapping coastal marine debris using aerial imagery and spatial analysis. *Mar Pollut Bull* 132:52–59
- MSFD Technical Subgroup on Marine Litter (2013) Guidance on monitoring of marine litter in European seas. JRC83985, Joint Research Centre Scientific and Policy Reports, Luxembourg
- Pan SJ, Yang Q (2010) A survey on transfer learning. *IEEE Trans Knowl Data Eng* 22:1345–1359

- Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. ICLR 2015
- Suaria G, Aliani S (2014) Floating debris in the Mediterranean Sea. *Mar Pollut Bull* 86:494–504
- Sun, C., Shrivastava, A., Singh, S., Gupta, A., 2017. Revisiting Unreasonable Effectiveness of Data in Deep Learning Era, 2017 IEEE International Conference on Computer Vision (ICCV), pp. 843–852
- Wang Y, Wang D, Lu Q, Luo D, Fang W (2015) Aquatic debris detection using embedded camera sensors. *Sensors* 15:3116–3137
- Woodall LC, Sanchez-Vidal A, Canals M, Paterson GLJ, Coppock R, Sleight V, Calafat A, Rogers AD, Narayanaswamy BE, Thompson RC (2014) The deep sea is a major sink for microplastic debris. *R Soc Open Sci* 1

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.