

# Oil Identification on Beaches Using Deep Learning Techniques

Ramoni Reus Barros Negreiros  
Federal Institute of Paraíba  
ramoni.reus@academico.ifpb.edu.br

Rafael Araújo dos Santos  
Federal Institute of Paraíba  
rafael.santos@academico.ifpb.edu.br

André Luiz Firmino Alves  
Federal Institute of Paraíba  
andre.alves@ifpb.edu.br

Anderson Almeida Firmino  
Universidade Federal de Campina Grande  
andersonalmeida@copin.ufcg.edu.br

**Abstract**—The oil spill on the beaches of the Brazilian north-east coast has caused disastrous consequences, both for us human beings, as well as for the marine biodiversity. Bearing in mind that its identification is not easy, we propose to do it in this article, using a machine learning approach, a type of artificial intelligence widely used in the area of image classification, alongside with a Convolutional Neural Network. In this regard, we have as main objective the training of a model able to distinguish images, automatically, in three types of classes: (i) normal beaches; (ii) beaches with gulf-weed, a biological indicator of pollution; and (iii) beaches with oil. In the best scenario evaluated, we achieved an average accuracy of 91%.

## I. INTRODUCTION

The oil spill in the Brazilian Northeast has caused one of the biggest ecological disasters in the country's history. It happened at first in Jacumã and Tambaba beaches - in the State of Paraíba -, on August 30th, 2019<sup>1</sup>. The incident spread and hit beaches from the northeastern region and the State of Espírito Santo, including in total 111 towns, 107 of which are from the northeastern region. Since September 2nd of the same year, members of IBAMA (Brazilian Institute for the Environment and Natural Resources), ICMBio (Chico Mendes Institute for the Biodiversity Conservation), the Brazilian Navy, and other entities have been working on the oil removal.

Regarding the oil cleaning process in Brazil, one of the biggest challenges is the size of the country's coastline, which therefore requires a lot of time for the oil to be found and removed. Until they discover the first signs of leakage, damage to the environment may have already taken on large proportions.

Another challenging factor in this process is that personal protective equipment (PPE) is not always available for manual cleaning. Thus, if an individual is not equipped, he/she risks suffering some intoxication, since the contact of the oil with the skin irritates [1]. Another issue is regarding the storage of sediments. The raw oil, because it is a toxic and flammable substance, requires an appropriate container for the storage of oil (e.g., plastic drums). However, volunteers and environmentalists who work to remove oil stains often

dispose of sediment in ordinary garbage. In this way, the material can come into contact with the soil, contaminating groundwater, besides emanating toxicity and allowing leakage into the environment [1].

Thus, we show the problem that affects a part of the Brazilian coast, affecting marine biological diversity, and the human being himself in terms of health, economy, and tourism in these areas. Thus, the aim of this work is to apply the concepts of Deep Learning, a sub-area of machine learning, within image identification, to enable the automatic recognition of oil spills on beaches. The built model can identify and distinguish images in three types of classes: (i) normal beaches; (ii) beaches with seaweed, a biological indicator of pollution; and (iii) beaches with oil.

Therefore, this work aims to assist in the identification of oil on beaches, providing a model that receives aerial images of beaches and can distinguish if there is any leakage problem, thus allowing the competent bodies and authorities to take the measures and act to remove sediment as soon as possible.

## II. LITERATURE REVIEW

Papers related to Machine Learning and image classification through Deep Learning and Convolutional Neural Networks (CNNs), applied to different fields, are widely found in literature [2]. In Zhang et al [3] e.g a CNN was used to automatically identify chromosomes, reaching 97.5% of accuracy. In the medical field, Carvalho and Machado [4] conducted a study on the CNNs' potential for breast cancer diagnosis on BI-RAIDS scale, proposing a four group classification.

In the environmental context, which is the focus of this paper, there also are significant contributions. A low-cost system was developed by [5] to monitor and identify incidents on coastal zones, such as oil leakage, alerting coastguard officers when an incident is identified. Another research by [6] used CNNs and different models of image segmentation to analyze and classify oil spills on the ocean through photos obtained from synthetic aperture radars in satellites and drones. Still regarding marine environment, Jin et al [7] explored the ability of a refined CNN with Global Average Pooling layer, instead of fully connection layer, that improved training speed,

<sup>1</sup>Available at: <<https://oglobo.globo.com/sociedade/salles-diz-que-mais-de-100-de-borra-de-petroleo-foram-recolhidas-no-nordeste-1-24002287>>. Viewed on: May 8th, 2020.

for oil spill detection with fully plyometric Synthetic Aperture Radar data.

Therefore, our paper is the only one that deals with multi-class image classification, with four classes: clean beaches, beaches with oil and beaches with seaweed. It presents good results even between the classes beaches with seaweed and beaches with oil, which would require more training data to avoid errors in the classification step.

### III. METHODOLOGY

We based the proposed approach on the use of artificial intelligence - i.e., machine learning. The developed application can learn and adapt to unorthodox situations [8]. In our specific context, it learns visual patterns to discern certain images. For humans, recognizing or classifying objects is a trivial task; however, for machines, this task is complex, despite the recent advances in computer vision (CV) techniques [9].

Among the different methods of image recognition, we highlight the Deep Learning (DL) methods, a machine learning subarea that uses algorithms to teach a computer how to recognise objects. The teaching process can be done in two ways: in the first, the task is delivered to the algorithm and the pictures (dataset) are already labeled with the categories well defined. This is called supervised learning. The second type of training only requires the digital picture, delivering the dataset with no labels and leaving the task of determining the object category to the algorithm after the extraction of common features, which is known as unsupervised learning [9].

We can make an analogy associating supervised learning to a baby, who classifies his toys as cars, dolls and balls, with his mother's supervision. On the other hand, we relate a robot to unsupervised learning: it learns by itself how to categorize different plants, i.e., how to find the similarities between them and divide them into groups that were not provided by its creator. In our case, we use supervised learning.

Once we explain the initial concepts, we will proceed describing the development of our application. We use the online environment "Google Colaboratory" (*Google Colab*), a free *cloud* platform based on Jupyter Notebook, suitable for machine learning, also with GPU support (Graphics Processing Unit). The GPU is a specialised computer *hardware* originally created for high frame rate<sup>2</sup> images rendering. In the present case, GPUs are extremely useful for doing optimized floating point operations (FLOPS), required by matrix calculus of deep learning.

The methodology proposed in this paper is shown in Figure 1 and it is composed of three stages: images capture, training and model evaluation. The next subsections will detail the step-by-step of the program's development.

#### A. Images capture

Following the process illustrated by Figure 1, in the first step (images capture) we have decided to use 450 images

<sup>2</sup>Available at: <[https://course.fast.ai/gpu\\_tutorial.html](https://course.fast.ai/gpu_tutorial.html)>. Viewed on: May 10th, 2020.



Fig. 1. Proposed methodology flowchart.

for training and validation (*gold data*), downloaded through *google-images-download*<sup>3</sup> library. The following terms were used to collect the images: Seaweed Beach, Oil Beach and Beach. Once the images were collected, the researchers filtered them according to their categories, leaving the *gold data* free of dirt (pre-processing stage).

With the intention of training the model to recognise beaches with three different features, the dataset was equally divided into three (150 pictures for each category). The beach with seaweed category was included so the classifier learned how to distinguish the oil from seaweed. From the dataset, 80% and 20% were used to train and validate the classifier respectively.

#### B. Model training

After the images capture stage, we begin the stage of training the model that will categorize the beach images. The *fastai*<sup>4</sup> library was used in order to build the model, it simplifies the training of fast and accurate neural networks, and more specifically its *vision* module, which contains all necessary functions to define a dataset and train a model for computer vision tasks. By importing it, we have all the classes and objects we need. Among them, the *imagedatabunch* object, which is responsible for creating the training and validation datasets from the files containing the respective images, and standardizing them in the same size. Another fundamental class in a neural network definition is the *learner* module, responsible for defining the *cnn\_learner* method, that facilitates obtaining a suitable model for transfer learning.

According to Silva and Costa [10], transfer learning consists of taking advantage of previous knowledge to improve speed or performance of learning a task (target). In short, it would be to use a trained model in a large dataset (usually the ImageNet in computer vision [11]) and then adapt it to your own dataset<sup>5</sup>.

For easier understanding, see the following example: a person who knows how to play guitar wants to learn how to play electric guitar. Since he already knows some music theory and how to apply it in the guitar, this person can transfer part of this knowledge to another instrument, facilitating the learning.

After understanding what is transfer learning, we will proceed with the choice of which neural network we will use in our approach. Thus, we used the *ResNet-34* architecture, which belongs to Residual Networks, a set of pre-trained neural networks, as a variation of this set.

The Convolutional Neural Networks (CNNs) were inspired on human viewing system structure and are extremely successful in computer vision programs, such as facial recognition,

<sup>3</sup>Available at: <<https://google-images-download.readthedocs.io/>>.

<sup>4</sup>Available at: <<https://docs.fast.ai/>>.

<sup>5</sup>Available at: <<https://docs.fast.ai/vision.learner.html>>.

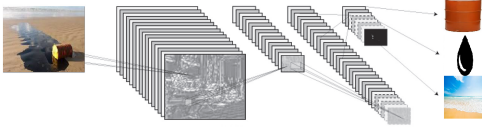


Fig. 2. Filter application. Adapted from: Voulodimos, A., et al. Deep Learning for Computer Vision: A Brief Review. Volume 2018.

objects detection, robotics, and driverless cars [12]. CNN’s are able to handle particularly pictures, since convolution allows filtering the images considering their two-dimensional structure (spatial). [13].

The main advantage of ResNet family compared to others CNNs is its expertise on computer vision tasks, possible due to the capability of creating shortcut connections which allow the simplification of information between the layers, optimizing the training [14]. Therefore, the ResNet-34 plays an important role as in Lau et al [15], that proposes a model instructed to identify cracks in sidewalks. That’s because the irregular standards and lighting conditions make the process difficult, which is mitigated by ResNet-34 functions.

Figure 2 demonstrates a CNN architecture detecting a beach with oil, pointing out necessary filters to break down a certain image. A CNN contains three main types of layers. These are: (i) convolutional layers, (ii) *pooling* layers and (iii) *fully connected* layers. Each type of layer plays a different role [13].

In a convolutional layer, each neuron is a filter applied to an input picture and each filter is a weight matrix. Kernels are applied to the picture, also known as convolution matrices - small matrices used to apply effects such as blurring, sharpness and borders detection. They are also used in machine learning for “resources extraction”, or convolution, a technique to determine the most important parts of an image [12].

The *pooling* layer, on the other hand, is responsible for reducing spatial dimensions (width x height) of the input volume. It should be noted that the *pooling* layer does not affect the depth dimension of the volume. This operation can also be named *subsampling* or *downsampling*, since the size reduction implies a simultaneous loss of data. However, this loss is beneficial to the network because the size reduction leads to less computational overhead to the next neural network layers [13].

Finally, in the *fully connected* layer (FC), each neuron has a weight associated to each element of the input vector. In CNNs, FC layers are used positioned after multiple convolutional layers [12]. The transition between a convolutional layer (that produces a tensor) and a FC layer, requires that the tensor be vectorized. For example, if the convolutional layer before a FC layer creates a tensor  $4 \times 4 \times 40$ , we resize this data so that it has the size of  $1 \times (4 \times 4 \times 40) = 1 \times 640$ . Therefore, each neuron in the FC layer must have 640 weights with the purpose of creating a linear combination of the vector [12].

		Predicted			FN
		Beach	Oil Beach	Sargassum Beach	
Actual Class	Beach	31	0	1	1
	Oil Beach	0	22	1	1
	Sargassum Beach	2	4	29	6
FP		2	4	2	8

#Oil Beach	
(TP) 22	(FN) 1
(FP) 4	(TN) 60

#Sargassum Beach	
(TP) 29	(FN) 6
(FP) 2	(TN) 53

#Beach	
(TP) 31	(FN) 1
(FP) 1	(TN) 51

Fig. 3. Confusion matrix.

### C. Model evaluation

Once the training stage is over, we enter the model evaluation stage. To make a visual analysis of the model results, we used a tool called confusion matrix, in which errors and hits can be visualized. On the main diagonal are located the correct predictions, while the others positions represent the wrong predictions. From the confusion matrix, True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) can be extracted, which are used in calculations of the main metrics of evaluation in systems of Data Recovery, especially classification problems.

To evaluate the model we use the accuracy metric (Acc), precision (*Precision*), *recall* and *F1 score* (or *f-measure*). The accuracy works to calculate the percentage of the model’s general hits. On the other hand, the *recall* indicates the frequency with which the classifier finds examples of the same class. The precision refers to the amount of hits classified. Finally, the *F1 score* represents the harmonic average of precision and *recall*.

In order to train the model, 10 epochs were used, because we observed that it presented better performance to get the networks parameters, avoiding *overfitting* (when the model exaggeratedly adapts itself to the training data) and *underfitting* (when the trained model was not trained enough).

## IV. RESULTS AND DISCUSSION

Figure 3 shows the confusion matrix to evaluate the trained model based on the validation data (20% of the 450 images). We note that the model does not show many errors, and that the biggest error amount (6) was counted when the model tried to predict the “seaweed beach” category. This indicates an efficacy a little smaller than the other categories, which counted one prediction error for each.

The evaluation metrics shown in Table I supplies a more detailed analysis. It is clear that the beach’s image classifier shows very satisfactory results, obtaining a 91% accuracy. It is important to note that the model indicates better results on images of clean beaches, reaching a 94% precision and 97% recall. This is due to the fact that there are more pictures of clean beaches in the validation set. Despite this, the results of beaches with oil are similar: 85% of precision; 96% of recall, and 90% of F1 score, which can be improved by adding more pictures.

TABLE I  
EVALUATION METRICS

Metrics	Categories			
	Clean Beach	Seaweed Beach	Oil Beach	Average
Precision	0.94	0.94	0.85	0.91
Recall	0.97	0.83	0.96	0.92
F1 Score	0.95	0.88	0.90	0.91
Accuracy	0.91			

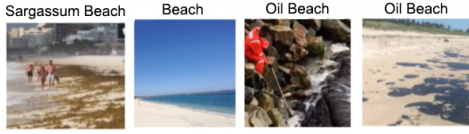


Fig. 4. Classification example made by our model.

Figure 4 presents an example of how classification is done by our model, with a label on each image. Therefore, it is possible to automatically classify any new beach image with a 91% accuracy, considering the three classes that the model was trained in.

## V. CONCLUSION

This paper presented an image classifier model based on a pre-trained CNN (resnet34) in a huge database (ImageNet), which contains images from different domains. Therefore, the transfer learning was carried out to adapt the knowledge acquired to our own database. It contains 450 images in total, with images of beaches with seaweed, beaches with oil, and clean beaches (without seaweed or oil). To each image class there is a set of 150 images. From this set, 80% was used for training the model and 20% for validation.

As a main difficulty we can mention the lack of images previously labeled, very valuable for the training optimization. With the shortage of this raw material, we chose to search for images on internet browsers, demanding more time. Moreover, about half of this process was made by the researchers themselves, without the use of a library to automatically collect the images (since over the course of the study the google\_images\_download was disabled), leading to a longer retention at this stage than expected.

Subsequently, the trained model was evaluated using the main classification metrics from Information Retrieval, namely: Accuracy, Recall, Precision, and F1 Score. The developed model presented excellent classification results, with 91% overall F1 Score, and 90% F1 Score for Oil Beach class (which is the focus of this paper). Thereby, we note that our model present an excellent result regarding oil spill identification on beaches.

We conclude that our model, based on machine learning, is able to present a solution to a contemporary problem, such as the oil recognition due the disaster initiated on August 2019, whose consequences are still experienced today.

Additionally, we also approached some of the typical concepts from machine learning and image classification literature, such as supervised and unsupervised learning and trans-

fer learning. Thus, aiming at demystifying such knowledge through analogies, which at first glance look too complex and intimidating to those unfamiliar with it. We highlighted the CNNs, regarding their importance to machine learning and their functioning, evidencing that this technology is among the most promising in this area.

As future work, we highlight the increasing of the dataset used, aiming to optimize the accuracy, precision, and other evaluation metrics. Moreover, it is intended to further specialise the model to detect and classify more types of images, such as fish death toll, muddy waters, caused by organic waste excess, and solid waste, like plastic. Therefore, the classifier would become specialized in beach anomalies identification, facilitating the work of automatic monitoring of a country's coast, such as Brazil.

**Source code.** We provide the project source code publicly in the following address: <https://github.com/RamoniReus/PraiaPetroleo>

## REFERENCES

- [1] "Guia de cuidados para moradores e voluntários que estão em áreas afetadas pelo derramamento de petróleo," Sociedade Brasileira de Dermatologia, Tech. Rep., 10 2019.
- [2] C. A. R. Pacheco and N. S. Pereira, "Deep learning conceitos e utilização nas diversas Áreas do conhecimento," *Revista Ada Lovelace*, vol. 2, pp. 34–49, dez. 2018.
- [3] W. Zhang, S. Song, T. Bai, Y. Zhao, F. Ma, J. Su, and L. Yu, "Chromosome classification with convolutional neural network based deep learning," in *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2018, pp. 1–5.
- [4] M. F. Carvalho and A. Machado, "Classificação da densidade mamária em mamografias utilizando redes neurais convolucionais," in *Anais Estendidos da XXXII Conference on Graphics, Patterns and Images*. Porto Alegre, RS, Brasil: SBC, 2019, pp. 184–187.
- [5] A. Correia, M. Simões-Marques, and R. Graça, *Automatic Classification of Incidents in Coastal Zones*, 07 2020, pp. 123–129.
- [6] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sensing*, vol. 11, no. 15, p. 1762, Jul 2019. [Online]. Available: <http://dx.doi.org/10.3390/rs11151762>
- [7] Z. Jin, L. Qingli, L. Yu, F. Hao, and W. Jujie, "Oil spill detection using refined convolutional neural network based on quad-polarimetric sar images," in *2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*. IEEE, 2019, pp. 528–536.
- [8] E. Alpaydin, *Introduction to Machine Learning*, ser. Adaptive Computation and Machine Learning series. MIT Press, 2020.
- [9] F. Arias Del Campo, O. O. Vergara, V. G. Cruz, L. A. García, and M. Nandayapa, "Influence of image pre-processing to improve the accuracy in a convolutional neural network," *International Journal of Combinatorial Optimization Problems and Informatics*, vol. 11, no. 1, pp. 88–96, Jan. 2020.
- [10] A. H. R. C. F. L. Silva, "A survey on transfer learning for multiagent reinforcement learning systems," *Journal of Artificial Intelligence Research*, vol. 64, pp. 645–703, 2019.
- [11] T. Akiba, S. Suzuki, and K. Fukuda, "Extremely large minibatch SGD: training resnet-50 on imagenet in 15 minutes," *CoRR*, vol. abs/1711.04325, 2017.
- [12] A. Vouloudimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational Intelligence and Neuroscience*, vol. 2018, p. 7068349, Feb 2018.
- [13] M. Antonelli Ponti and G. B. Paranhos da Costa, "Como funciona o Deep Learning," *arXiv e-prints*, p. arXiv:1806.07908, Jun. 2018.
- [14] S. Targ, D. Almeida, and K. Lyman, "Resnet in resnet: Generalizing residual architectures," *arXiv preprint arXiv:1603.08029*, 2016.
- [15] S. L. Lau, E. K. Chong, X. Yang, and X. Wang, "Automated pavement crack segmentation using u-net-based convolutional neural network," *IEEE Access*, vol. 8, pp. 114 892–114 899, 2020.