

Abapathi

Fish vs plastic classifier for ocean cleanup

Author 1 **AUTHOR1@POLYTECHNIQUE.CM**

*Computer Science Department
National Advanced School of Engineering of Yaounde
Yaounde, Cameroon
P.O. Box 8390 Yaounde*

Author 2 **AUTHOR2@POLYTECHNIQUE.CM**

*Computer Science Department
National Advanced School of Engineering of Yaounde
Yaounde, Cameroon
P.O. Box 8390 Yaounde*

1 Introduction

Plastic pollution is pollution caused by the accumulation of plastic waste in the environment. The global system of production, use and disposal of plastics is a broken system. Plastic pollution is correlated with the low cost of plastic, which leads to its massive and disposable use. It is also due to the low degradability of plastics. This pollution can have detrimental effects on land, seas and oceans, and waterways by affecting wildlife, habitat and secondarily or through feedback from humans. It is estimated that 24 to 35 million tons of plastic waste enter the aquatic environment each year. Living organisms, particularly marine animals, can be impacted, either through mechanical effects such as entanglement in plastic objects and problems associated with ingestion of plastic waste, or through exposure to chemicals in plastics that interfere with their physiology. Effects on humans include disruption of various hormonal mechanisms. The pollution of the oceans by plastics is therefore a serious problem to which a solution must be found.

In this paper we will build a fish vs plastic classifier that will be able to distinguish between fish images and plastics images, these images being taken in ocean waters. A robot equipped with such a model would be able to scour the depths of ocean waters and retrieve plastic waste without endangering marine animals.

2 Related work

Regarding the problem of cleaning the oceans, at this time the automatic solutions proposed are robots that scour the surface of the water to recover floating plastic waste. The best of them are equipped with an Artificial Intelligence allowing them to detect

objects, which allows a more optimal movement on the surface of the water, like the one proposed by Razer [1].

The limit of this solution is that it can only recover the plastics present on the surface, but some of this waste remains at depth, and represents a real danger for marine animals, and consequently for humans. For a robot to be able to carry out this cleaning in deep waters, it would have to be equipped with a program allowing it to recognize and distinguish plastic waste from marine animals.

Concerning what has already been done, on the one hand a fish detection algorithm has been published by Yu Zhou and Alt [2]. On the other hand, Gautam Tata proposed an algorithm for the detection of marine plastics, realized with YOLO, and this one having an accuracy of 85% [3].

The model we propose will be able to distinguish between fish and plastics with a very good accuracy so that the robot in charge of cleaning will make as few mistakes as possible.

3 Methodology and implementation

To successfully achieve this project, we followed a methodology based on four (06) main stages:

- The definition and understanding of the problem
- Data collection
- Data preprocessing
- Building Deep Learning/Machine Learning algorithm using convolutional neural network
- Training and testing
- Result

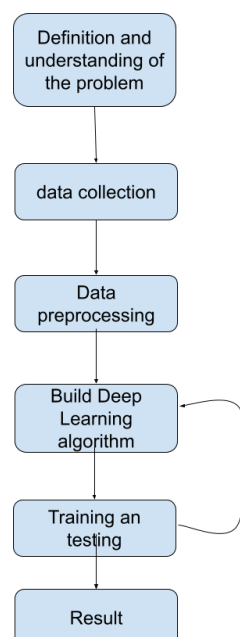


Figure 1: Methodology process

3.1 Definition an understanding of the problem

Most African coastal countries open to the ocean such as Senegal, Ghana, South Africa, Kenya and Cameroon are facing the problem of ocean water pollution. According to UNEP[4], every year humans produce 300 million tons of plastic garbage, and 11 million of them end up in the oceans. end up in the oceans. Without rapid action on a large scale, these volumes triple by 2040. The plastic never really disappears. It turns into fine particles, which are swallowed by fish or farm animals and eventually consumed by humans through their food and tap water. This leads to an increased risk of cancer and hormonal problems.

As some consequences of pollution of oceans by plastics in the african context we can list :

- The ocean provides the primary source of protein for more than one billion people. The threat of marine wildlife poses a danger to the food supply of these populations.
- Pollution threatens the job security of workers in the artisanal fishing sector.
- A country like Senegal has as its main source of income the fishing industry.
- Strangulation of marine animals.
- Waste that has not sunk or been washed ashore, drifts and drift and wash up on our beaches. In addition to constituting a major ecological challenge, they are a visual pollution with real impacts on the local economy. It is thus all the tourist activity that can suffer the consequences: a decrease in the number of tourists, negative image of the town, etc.

Following these observations, it is trivial that the pollution of ocean waters is an urgent problem that must be addressed.

3.2 Data collection

Data collection is one of the first and most important tasks of a Machine Learning project, since the algorithm will process this data to provide results. The quality of the data used in the construction of the model therefore has a great impact on the quality of the results obtained.

To build our fish vs plastic classifier, we needed as input fish an plastic images taken inside ocean waters. We built our dataset by collecting images from several sources:

- Almost 10% of both type of images (fish and plastic) has been selected manually on google, pixabay, unsplash.
- 55% of fish has been taken into the dataset “Fish Species Image Data” on Kaggle[5] and 40% from the dataset “Blue Bot Dataset: Parrot fish”[6].
- 95% of plastic images is from the drive “DeepPlastic”[7].

This allowed us to have a balanced dataset of 3550 fish images, and 3600 plastic images.

3.3 Data preprocessing

Tensorflow allows to label data automatically. For our project, we have two classes: a plastic class and a fish class. To label our images, we simply created two folders with the exact names of our classes ("fish" and "plastic") and in each of these folders we respectively put the corresponding images. When loading the images, tensorflow automatically assigns to each image as a class the name of the folder that contains it, and it is also possible to indicate the proportion that will be used for training and for testing.

Below we have some examples of the dataset images.



Figure 2: examples of plastic images



Figure 3: examples of fish images

3.4 Building Deep Learning/Machine Learning algorithms

The problem we are trying to solve falls into the category of image processing. We tried several approaches to solve it, including some Machine Learning algorithms like k-Nearest Neighbor Classifier, SVM and a deep learning algorithm, precisely a Convolutional Neural Networks (CNN). We evaluated the performance of each of these models and finally chose the best one.

The main challenges faced when solving the image classification problem are: view-point orientation, scale variation, deformation, occlusion, illumination conditions, background clutter and intra-class variation. Thus, a good model has to overcome all, if not most of these challenges.

- K-Nearest Neighbor Classifier

Clearly the simplest approach, it works as follow: the image to classify will be compared to every single image of the training images, and the predicted label is voted by the label of the k closest images of the training images.

In our case, to compare two images, we tried two distances: the L1(Manhattan) distance and the L2(Euclidean) distance. Given two images I_1 and I_2 of size $n \times m$:

$$d_1(I_1, I_2) = \sum_{i=1, \dots, n; j=1, \dots, m} |I_{1(i,j)} - I_{2(i,j)}|$$

$$d_2(I_1, I_2) = \sqrt{\sum_{i=1, \dots, n; j=1, \dots, m} (I_{1(i,j)} - I_{2(i,j)})^2}$$

To choose the best value of k , we varied k in the interval 10-100, observing the accuracy of the model each time.

- SVM

This is one of the Linear Classifiers. With this approach, we try to model a linear function f that take an image x as input, and return L numbers giving class scores, L being the number of label classes.

$$f(x, W) = Wx + b$$

W and b are the parameters of the model to be learned.

In the case of SVM this linear function f is the equation of the decision boundary used to classify a new input.

For non-linearly separable data, it is often used to define an affine application called a kernel to reduce the data to a higher dimensional space where they are linearly separable, which makes classification easier.

- Convolutional Neural Networks(CNN)

The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. In mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other. This kind of neural network uses convolution in place of general matrix multiplication in at least one of their layers. Their architecture involves several types of layers, each playing a specific role.

- Convolutional layer

It is the core building block of the CNN, and is where the majority of computation occurs. To work it requires an input data, a filter or kernel and a feature map. The filter will move across the receptive fields of the image, checking if the feature is present. This process help to detect features present on the image such as edges, color, gradient orientation.

- Pooling layer

This layer conducts dimensionality reduction, reducing the number of parameters in the input. The pooling operation sweeps a filter across the entire input, but the difference is that this filter does not have any weights. Instead, the kernel applies an aggregation function to the values within the receptive field, populating the output array. There are mainly Max

Pooling and Average Pooling. Even if a lot of information is lost in the pooling layer, the benefit is that it helps to reduce the complexity, improve efficiency, and limit the risk of overfitting.

– Dense layer or fully-connected layer

For a such layer, each node in the output layer connects directly to a node in the previous layer. This layer performs the task of classification based on the features extracted through the previous layers and their different filters. This layer usually leverage a softmax activation function to classify inputs appropriately, producing a probability from 0 to 1, whereas convolutional and pooling layers tend to use ReLu functions.

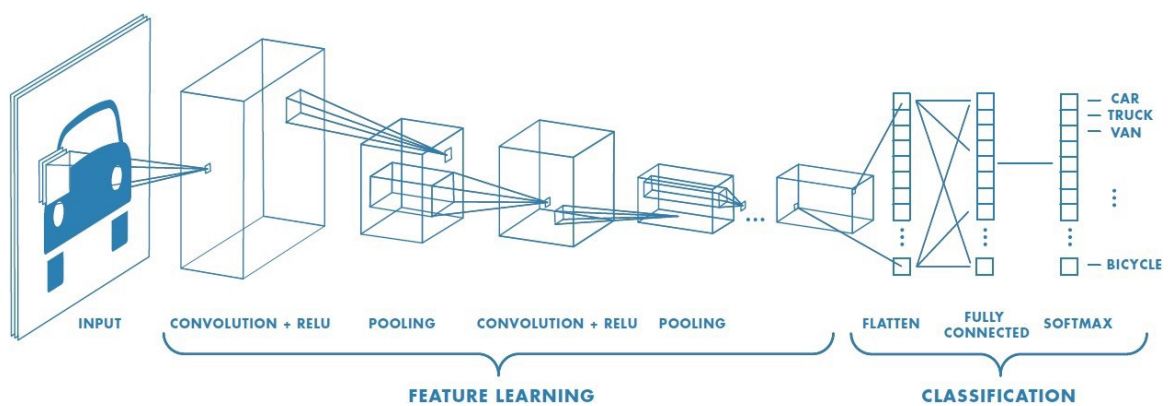


Figure 4: CNN architecture [8]

The architecture of the neural network we used is as follows:

- A rescaling layer to bring the pixel values in the interval $[0, 1]$.
- 04 series of convolution layer followed by a max pooling layer.
- A flatten layer.
- Then follow a dense layer for the classification
- And finally another dense layer for the output.

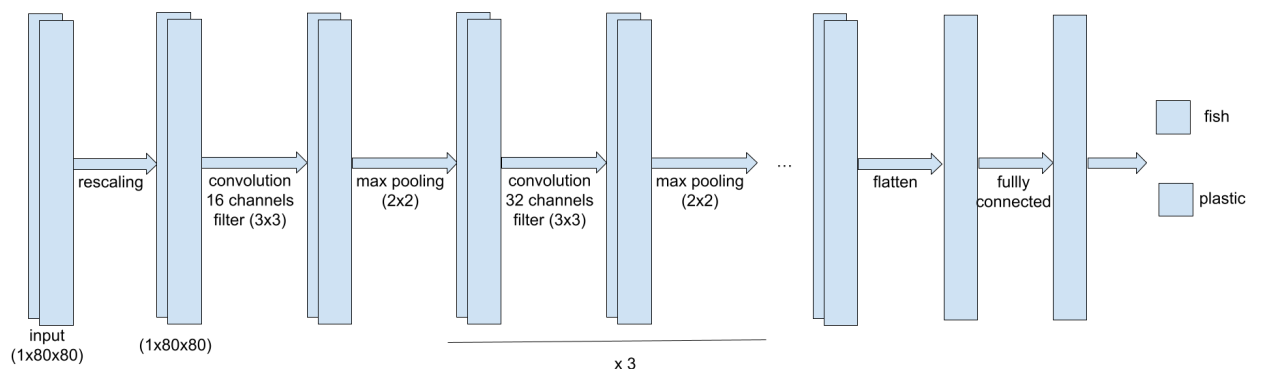


Figure 5: Our Neural Network architecture

3.5 Training and testing

After modeling our neural network, we used the tensorflow framework to implement it. And now it is the step of training and testing.

This is where performance of the algorithm, quality of data, and required output all appears out. From the huge dataset, collected 80% of the data was used for training and 20% for the test. Training is the process of making the machine to learn and giving it the capability to make further classifications based on the training it took. Whereas testing means already having a predefined data set with output also previously labelled and the model is tested whether it is working properly or not and is giving the right class or not. If maximum number of classificatios are right then model will have a good accuracy percentage and is reliable to continue with otherwise better to change the parameters of the model. Also, further new set of inputs and the classifications made by the model will be keep on adding to the dataset, which makes dataset more powerful and accurate.

We have trained our model on 30 epochs with 8 steps per epochs.

3.6 Result

Since we used a balanced dataset, we considered it relevant to use accuracy as a performance measure for our model.

Of all the approaches we used, the best result was obtained with the CNNs, with 98% of accuracy on testing.

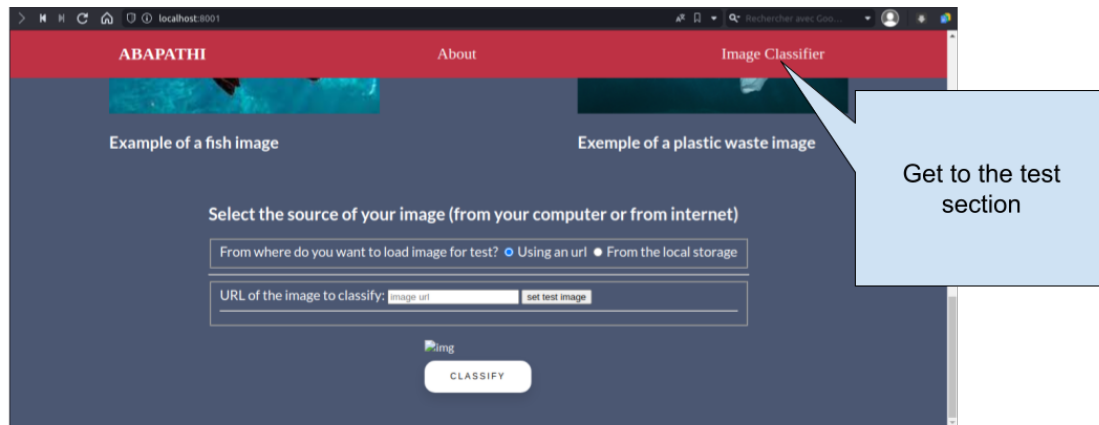
Model	accuracy	Model description
K-Nearest Neighbor Classifier	38.4%	k=10; L1 distance
K-Nearest Neighbor Classifier	40.2%	k=12; L2 distance
SVM	59.8%	kernel: RBF function
CNNs	98%	

3.6.1 Prototype

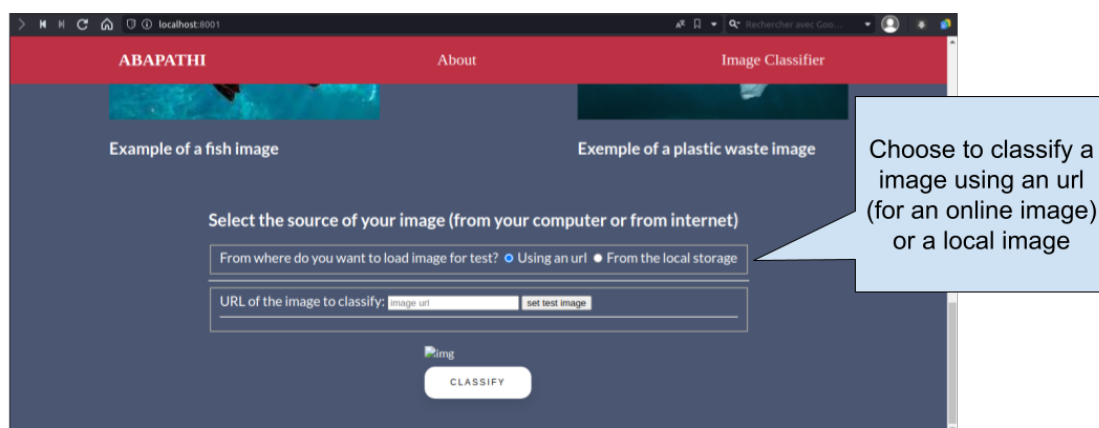
We have integrated our final model in a web application using Flask. we have included in the submission folder the source code of this application (in the app folder). You will find in this app folder a file "deployment_guide.pdf" which explains how to deploy the application locally in order to test it.

After following this guide and deploying the application, below is the user guide of the application.

1. Get to the test section



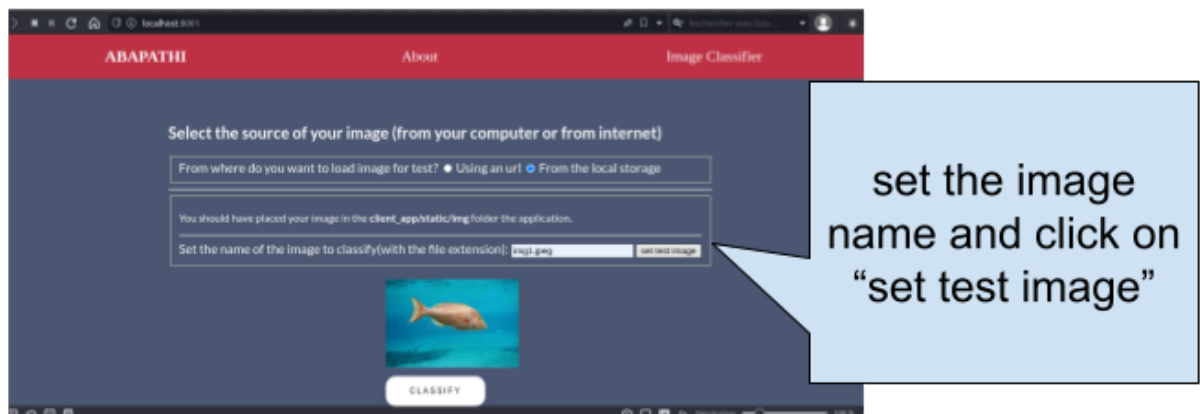
2. Select the image source



3. If you choosed to classify an image from the internet, go in your browser and copy the image's address and paste it in the form.



Otherwise, if you choosed to classify a local image, you have to put your image in the client_app/static/img folder, and set the image name with it's extension in the form.



4. Then hit the classify button to classify the image. You will then see the result appear below the image.



4 Conclusion and future work

In conclusion, this paper presents fish vs plastic classifier, for images taken in ocean waters. To achieve this result, we walked through a clear methodology that includes many steps (definition of the problem, data collection, data preprocessing, Deep Learning algorithm, training, testing and results).

Currently, our model works well only for fish with a specific shape like those presented on the dataset image examples. For the model to be usable it must be able to distinguish between plastics and marine animals in general, not just fish with a particular shape. To achieve this goal, it is necessary to have a large number of images of different marine species. This would greatly help to improve the performance of the model and make it usable in the real world. This is therefore a future perspective for the improvement and finalization of the model.

References

- [1] <https://gamerant.com/razer-robot-cleans-trash-ocean/>
- [2] <https://www.hindawi.com/journals/acisc/2020/3738108/#results-and-discussion>
- [3] <https://towardsdatascience.com/identifying-and-removing-marine-plastic-a-deep-learning-approach-6ad5e41bec41>
- [4] <https://www.unep.org/news-and-stories/story/tourism-pandemic-world-tackling-plastic-pollution>
- [5] <https://www.kaggle.com/datasets/sripaadsrinivasan/fish-species-image-data>
- [6] <https://www.kaggle.com/datasets/hiyaro/parrotfish>
- [7] https://drive.google.com/drive/folders/1fsS_u2QpbRGynYkP6-D6cfvq8r0hpjXI
- [8] <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>