ASTRO PI

MISSION SPACE LAB

Mission Space Lab Phase 4 Report

Team Name: Bergson

Chosen theme: Life on Earth

Organization name: Lycée Henri Bergson

Country: France

Introduction

Like everyone on the planet we felt concern about COVID-19 pandemic. The starting point of our experiment was the impact of COVID-19 on NO_2 worldwide emission, which is mapped by ESA Copernicus Sentinel-5P satellite.

We took this opportunity to evaluate whether the NO_2 concentration in the atmosphere could be correlated to the Normalized Difference Vegetation Index (NDVI) that can be processed by Astro Pi.

To do so, the Bergson team built an Artificial Intelligence (**AI**) model predicting NO_2 pollution level from NDVI pictures taken by Astro Pi Izzy computer.

We tackled two challenges at the same time throughout our experiment. First, we addressed a Data Science problem where we had to figure out how to correlate NDVI pictures data to NO_2 data. Secondly, we wanted to evaluate the technical feasibility of running our experiment in real time with the limited resource of Astro Pi. As a matter of fact, running inference from a Deep Learning AI is not easy when it comes to Raspberry Pi especially with Flight OS constraints.

Method

The basics of any Data Science project is to leverage training data that can be used for a model to learn to make prediction. Data we used are images from Earth taken by previous Astro Pi mission, on which we applied the NDVI pre-processing. We then manually associated a NO_2 category within ”low”, ”medium”, ”high” to each NDVI pictures to create our training dataset, a list of (NDVI, NO_2) couples.

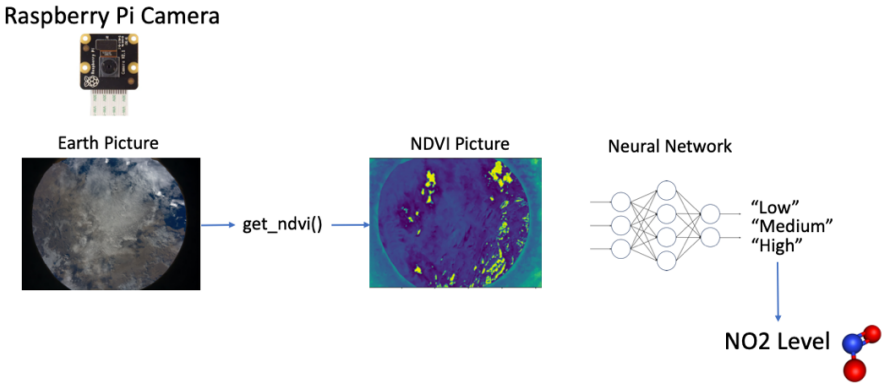


Figure 1 – Bergson Experiment Design

Because the training dataset is based on images, we decided to apply Computer Vision techniques to build our Artificial Intelligence model. We therefore leveraged a Convolutional Neural Network (CNN) architecture to build our 3-class classifier.

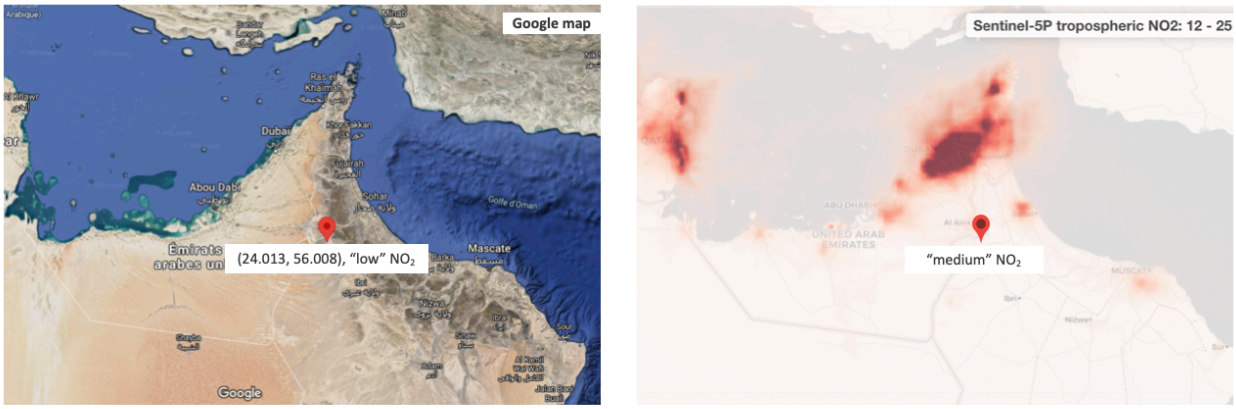
We got started with a simple CNN architecture but realized that our small dataset would probably lead to poor result in real condition. To overcome this limit, we introduced a transfer learning approach with Mobilenetv2 architecture, hopefully providing better prediction performance.

After our model was trained, the Bergson experiment method was straightforward. We take an image from the Astro Pi camera, this image is then pre-processed as an NDVI picture. The NDVI picture is fed into the trained Neural Network which finally output a NO_2 prediction category.

Results

Over the 3 hours run of our experiment, we took 490 x raw images and made 490 x NO_2 predictions saved on a CSV file, but kept only 245 x good samples that were not dark nor to much sunny.

To analyze our results, we post-processed our raw data to build an appropriate data representation as a data frame containing the AI predictions, the geolocation of the picture as well as the country name in a single clean file.



NO_2 AI prediction NO_2 actual

Figure 2 - NO_2 AI "low" vs NO_2 Actual "medium"

We then used the picture GPS coordinates to first locate our NO_2 AI predictions on Google Maps, then compared it with the timestamped "ground" truth value from Sentinel-5P tropospheric satellite. This process allowed us to measure the quality of our AI model, using typical classifier metrics like Confusion Matrix, Precision, Recall and Accuracy.

To build those metrics, we use extra code from scikit-learn library, which comes handy with functions to calculate them.

The first outstanding result is that our AI model predicted well "low" level of NO_2 pollution for 236 x pictures. When comparing to the Sentinel-5P tropospheric NO_2 value, we should have detected 236 x "low" and 9 x "medium".

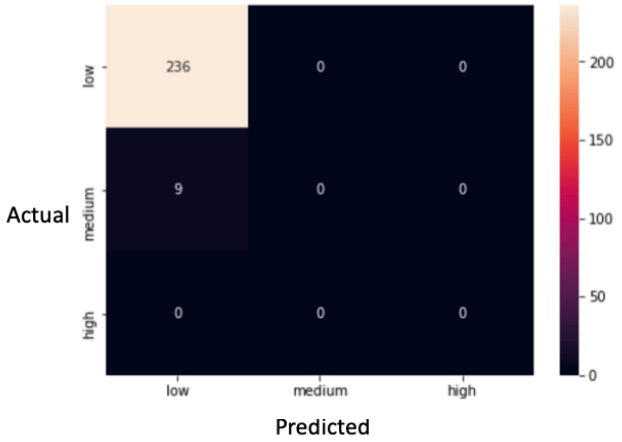


Figure 3 - Confusion Matrix

When looking at more advanced metrics, it shows however that our classifier did not perform well on predicting "medium" classes that should have been detected, because we got precision and recall value of 0 for this class.

	precision	recall	f1-score
low	0.96	1.00	0.98
medium	0.00	0.00	0.00
high	0.00	0.00	0.00
accuracy			0.96

Figure 4 - Classifier metrics

Overall, we got an Accuracy value of 96% on the prediction made by our AI model. This is a very promising result for a first iteration of the Bergson experiment. This result makes lot of sense provided that during the experiment run, the ISS flew over important ocean area where we naturally expect the NO_2 pollution to be low as no heavy industrial activity can happen there.

Conclusion

NDVI to NO_2 correlation
First, the Bergson team experiment has demonstrated that a clear correlation exists between the NDVI measure and the related NO_2 level of a pictured area. As a matter of fact, 96% of predictions made by the Bergson AI model were correct.

Deep Learning AI on Astro Pi
Secondly, we are very proud for running successfully a Deep Learning model on Astro Pi using transfer learning technique. We leveraged the pre-trained Mobilenetv2 neural network that comes with 2.2 Million parameters and took approximatively 30s to execute each pass forward on Astro Pi Izzy.

AI model Improvement
Finally, we want to propose two improvements to our AI model:

1. Get more data from previous & future mission to build a more important dataset with more diverse occurrence of "low", "medium" and "high". Typically, this would require picture from industrial areas like the US, Europe, China, Brazil.
2. Create both a training and validation set, then retrain the model to discuss further the classifier metrics to avoid overfitting.

For that, we think that a reasonable time window for our experiment would be 30 hours, from which we could expect around 2500 x pictures over main industrial areas.