MONUMENTS RECOGNITION

2019-2020

Table of Contents

1 Introduction

First of all, i would like to thank EMINES School of Industrial Management for this priceless opportunity that leads me to conclude that Morocco needs just to be conducted by right persons. FSKM UiTM for the hospitality and the great tutoring especially my supervisors : Assoc. Prof. Zaidah Ibrahim, Dr Azliza Mohd Ali and Prof. Dr. Yap Bee Wah the head of Advanced Analytics Engineering Centre (AAEC).
The report will contains a brief description of my personal journey in Malaysia, and the technical details of my project.

2 Malaysia wasn't my first choice but definitely the best choice

2.1 Weather Conditions

The first remarkable thing in Malaysia was the weather, it's categorised as equatorial, being hot and humid throughout the year. It was very difficult for me to get familiar with this situation, since in Morocco we have the four seasons and the humid weather is located in just few cities.

FIGURE 1 – Raining day in Kuala Lumpur the capital of Malaysia

2.2 The touristic Malaysia

Malaysia is known as one of the most attractive places over the world, it contains plenty of islands, jungles and waterfalls.

2.2.1 Langkawi island

I had the chance to visit Lankgawi a duty free island.
Oriental Village was the first place to see, is an open air all-in-one complex. It's full of stalls retail stores selling all kinds of items like clothes, souvenirs, bags, sunglasses, art craft items and lot more. Other offerings include several food beverage outlets restaurants, host of activities, rides tours, animal exhibits, spa and even a boutique hotel.

FIGURE 2 – Oriental Village

SkyBridge is one of the famous places in Langkawi, completed in 2004, this suspended bridge is built on top of the Machinchang mountain. Accessible from the Top Station, the bridge is suspended from a 82m high single pylon, hangs at about 100m above ground and it can accommodate up to 250 people at the same time.

FIGURE 3 – SkyBridge, Oriental Village

I also tried some maritime activities, such as parasailing. it costs about 400 DH for 15 min in Saidia (touristic place in the east of Morocco) whereas 250 DH for the same duration in Langkawi.
The north of Morocco if full of amazing beaches, however there is no direct roads to the most of them, so it is almost impossible to visit them with family.

2.2.2 Malacca

One the historical cities in Malaysia (700 years of existence), it contains several monuments, most of them from the portugal colonisation. (The famous one is A'Famosa).

FIGURE 4 – A'Famosa Monument

The city is also well known by the Jonker Walk which is the Chinatown street of Melaka. The road starts from across Melaka River near the Stadthuys. The road is filled with historical houses along its left and right sides dating back to 17th century. It also has shops selling antiques, textiles, foods, handicrafts and souvenirs such as keychains and shirts.

FIGURE 5 – Jonker Walk

3 FSKM UiTM

My internship took place in Faculty of Computer and Mathematical Science in Univeristi Teknologi of Mara, located in Shah Alam, the university gives the students with good background in Mathematics and Computer Science a 4 years teaching in IT and Data Science (where i was affected), an advanced tutoring in their final year projects that i had the chance to visit in their exhibition.

FIGURE 6 – The exhibition of the final year projects

During my training, Prof. Dr. Yap Bee Wah the head of Advanced Analytics Engineering Centre (AAEC) invite us to the summer workshops that are related to the data science.
The first was about Deep Learning Using Matlab, 2 days of intensive courses about image classification and transfer learning.
The second one covered the data analytics using IBM SPSS, it had included the exploratory data analysis, the different tests (normality, ARIMA...).

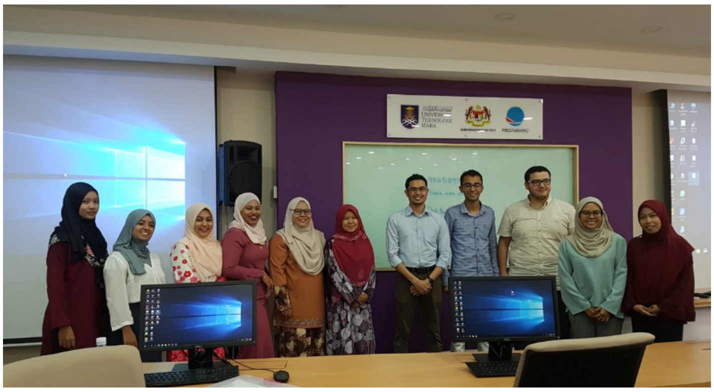

FIGURE 7 – SPSS Worshop

Beside the tutoring provided by our supervisors Assoc. Prof. Zaidah Ibrahim and Dr Azliza Mohd Ali, they invite us several time to try the local food and fruits, and to discover the malaysian weeding.

FIGURE 8 – Outdoor activities

3.0.1 Personal Experience

My first time in an asian country after 17 hours traveling from Morocco, it was a great experience to discover a place that is far away from my own one. The purpose was (beside working on a technical project) going outside of my comfort zone.
The first thing that comes in mind in a new country is the crime level :
Source : https ://www.numbeo.com/crime/rankings_by_country.jsp

— Malaysia is ranked 20 over 123 countries with an index of crime of 60.66, and safety index of 39.34.
— Morocco is ranked 39 over 123 countries with an index of crime of 49.53, and safety index of 50.47.

Kuala lumpur, the capital of Malaysia is very famous by the big malls, its ground floor is dedicated generally to small sellers in order to prevent the spread of peddlers.

FIGURE 9 – The ground floor of PAVILION mall in KL

This section will be discussed in details during my presentation.

4 Monuments Recognition Using Deep Learning

4.1 Introduction

Image Classification is one of the famous application of deep learning, and it refers to a process in computer vision that can classify an image according to its visual content. For example, an image classification algorithm may be designed to tell if an image contains a human figure or not.

4.2 Context of the project

The idea of the project cames from a need of an application that can guide the tourist to well know the famous monuments of a country before visiting it. In fact, anyone of us can find an image of a monument in social media and want to know its name and ideally its location.

4.3 State of art

Google Cloud offers two computer vision products that use machine learning (AutoML Vision and API Vision), it can gives you the different labels in your image with a certain probability.

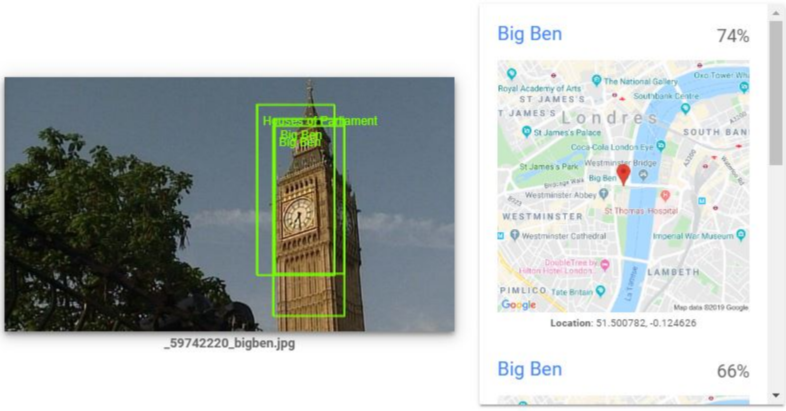

FIGURE 10 – Example of execution in Google Cloud

The image above shows that the Google Cloud gives the name of the monument (Big Ben 78%) with its location.

4.4 Convolutional Neural Network (CNN)

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.
The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.I will use CNN in this project.
The CNN architecture consists most of time of 3 layers (Convolutional Layer,Pooling Layer and the output Layer)

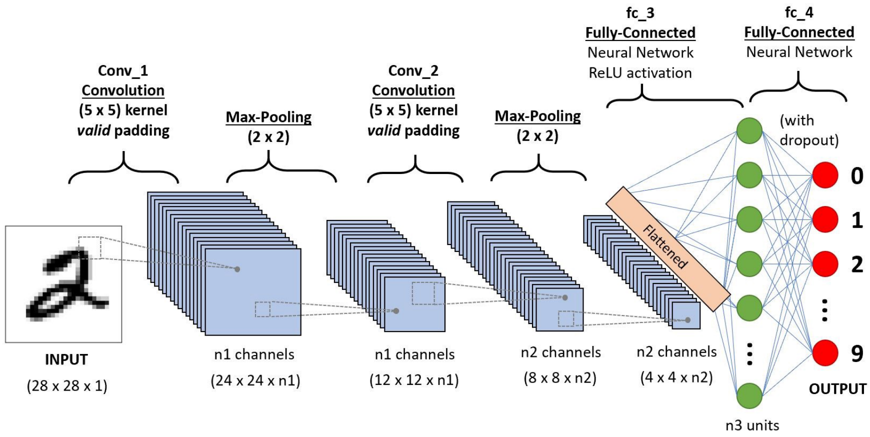

FIGURE 11 – A CNN sequence to classify handwritten digits

4.5 Dataset

4.5.1 Data gathering

The Dataset of the project is a combination of images that I photographed and downloaded from google images using a plugin of google chrome that help to scrap images from a given web page zipped in one folder, the steps are as follows :

— A query (name of the monument) in Google Images should be written in the search box ;
— A zipped folder is uploaded to the local machine.

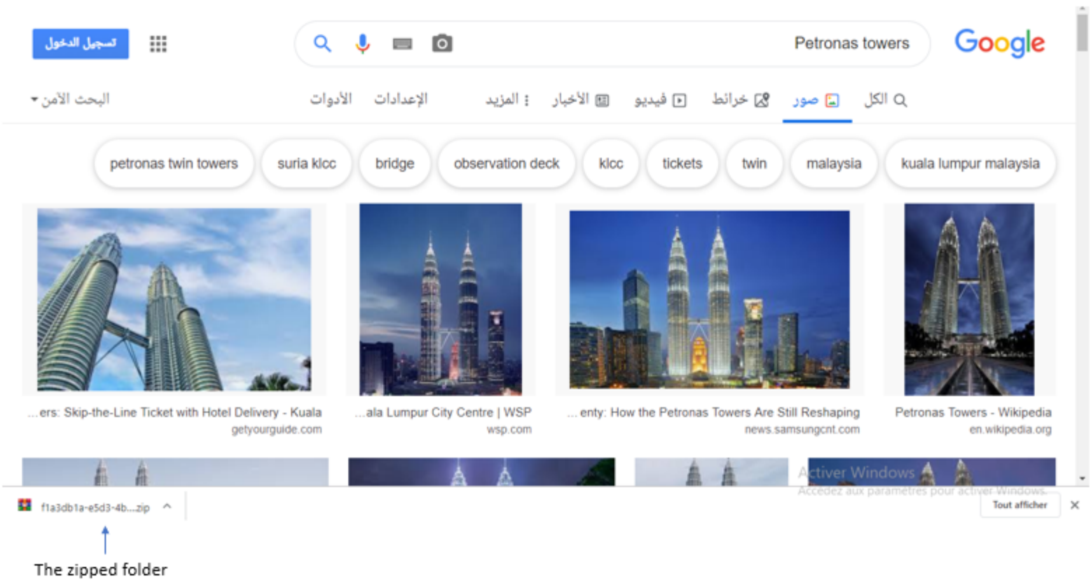

FIGURE 12 – Example of the process

The Dataset consists of monument classes :

— Eiffel Tower, France : 451 pictures ;
— Big Ben, England : 359 pictures ;
— Petronas Towers, Malaysia : 537 pictures ;
— Statue of Liberty, USA : 635 pictures ;
— Taj Mahal, India : 420 pictures.

The final dataset is a folder that contains sub-folders as follow :

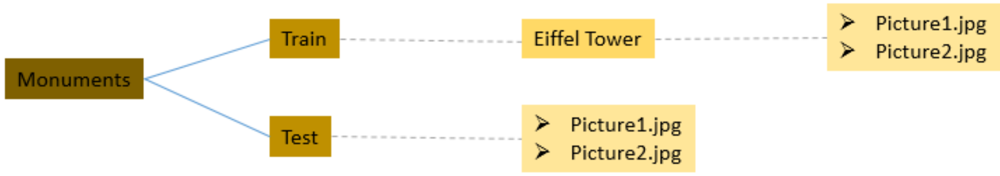

FIGURE 13 – Architecture of the dataset

Let's plot some images from the dataset.

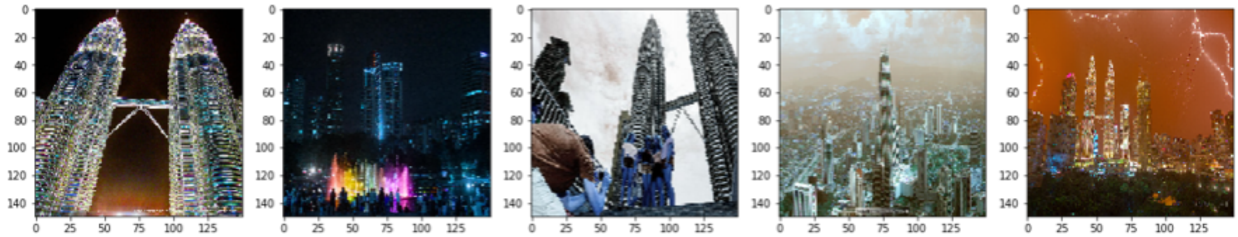

FIGURE 14 – Example of images from the dataset

4.5.2 Data processing

The next step is to resize the image in order to prepare it as an input of the model, it depends of the default shape of different models, let's plot the original and the resized image from the dataset.

FIGURE 15 – Example of resizing an image

Let's see the number of images per class (monument) in the next graph.

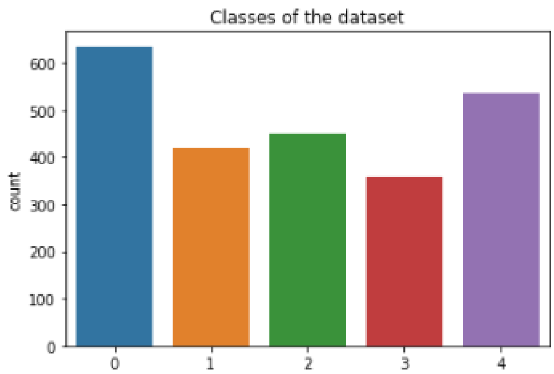

FIGURE 16 – Classes of the dataset

We have 5 classes for 5 monuments distributed as follows :
- Class 0 : Status of Liberty ;
- Class 1 : Taj Mahal ;
- Class 2 : Eiffel Tower ;
- Class 3 : Big Ben ;
- Class 4 : Petronas Tower.

4.6 Experiments

4.6.1 Parameters of the study

The aim of this study is to determine the best hyper-parameters for the CNN Model. Mainly, we have 3 hyper parameters :
- Epochs : one epoch is one forward pass and one backward pass of all the training examples ;
- Dropout : a technique where randomly selected neurons are ignored during training. They are "dropped-out" randomly ;
- Batch Size : is the number of samples that will be passed through to the network at one time.

4.6.2 General procedure

The experiments will begin by running the program 10 times, each time 2 parameters will be fixed to a random value while changing the target. After that we choose the best validation accuracy, finally we plot the graph of different validation accuracies in order to pick the the highest one.

4.6.3 Epochs

4.6.3.1 Hypothesis

The higher the epochs number, the higher the accuracy we get. For this experiment, we fix some parameters (just a choice) : 3 layers, batch_size=64, dropout=0.2, kernel_size=2 and pooling_size=2.

4.6.3.2 Results

The results of this experiment are shown in the next table :

Epochs	Validation accuracy
5	0.7344
10	0.8215
20	0.8796
30	0.8921
50	0.9087
60	0.9211
80	0.9211
100	0.917
150	0.917
200	0.917
250	0.9211
300	0.9128
350	0.9253
400	0.9211
500	0.8921
600	0.9253
800	0.8879

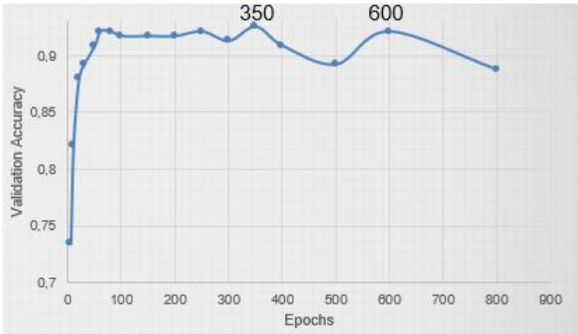

FIGURE 17 – Evolution of the validation accuracy in function of the epochs

The graph shows that the validation accuracy reaches its maximum (92.53%) for two values of epochs 350 and 600, i will work with the first value since it consumes less time in the training phase.

4.6.3.3 Conclusion

The previous hypothesis isn't true in this case, there is a fluctuation of the values of the validation accuracy.

4.6.4 Dropout

4.6.4.1 Hypothesis

The main purpose of this experiment is to fix the problem of overfitting which means that the training accuracy is far away from the validation accuracy, the hypothesis is formulated as follows :
The higher dropout, the lowest the difference between the training and validation accuracies. We fix some parameters (just a choice) : 3 layers, epochs = 100, batch_size=64, kernel_size=2 and pooling_size=2.

4.6.4.2 Results

Since the dropout is a probability between 0 and 1, I will choose a step of 0.1 (9 points).

Dropout	Difference between accuracies (%)
0.1	11.68
0.2	12.16
0.3	11.3
0.4	11.42
0.5	9.45
0.6	9.03
0.7	5.78
0.8	3.1
0.9	1.81

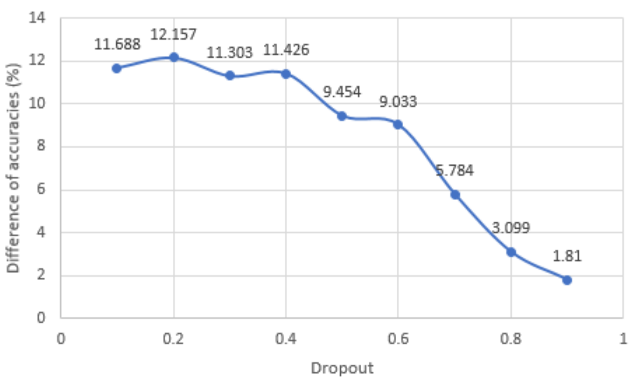

FIGURE 18 – Evolution of Difference between accuracies in function of the Dropout

The general trend of the graph is a decrease of the difference between training and validation accuracies, however we should pay attention to the validation accuracy since it measures the performance of the model.
Let's have a look on the maximum validation accuracy reached for every dropout :

Dropout	Maximum validation accuracy (%)
0.4	0.9087
0.5	0.9128
0.6	0.917
0.7	0.9336
0.8	0.8755
0.9	0.8215

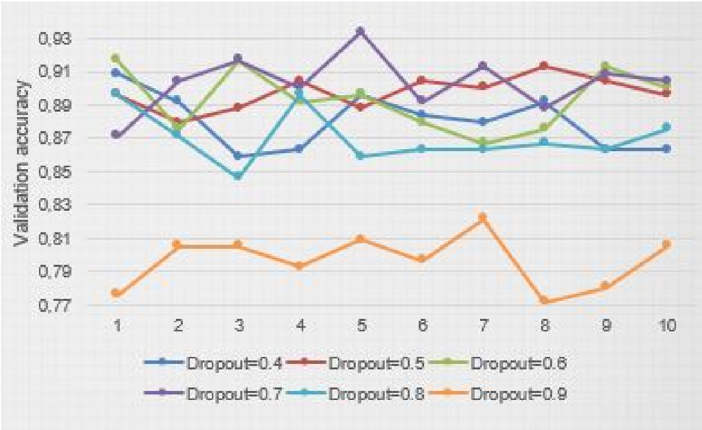

FIGURE 19 – Evolution of maximum validation accuracy

We observe from the last graph that the validation accuracy is maximum for a dropout of 70%.

4.6.4.3 Conclusion

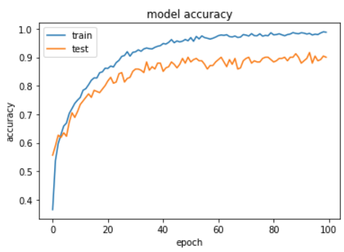

FIGURE 20 – Dropout = 0.6

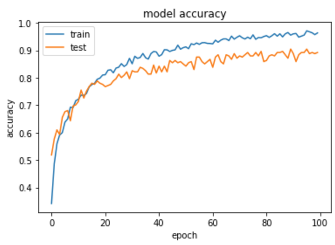

FIGURE 21 – Dropout = 0.7

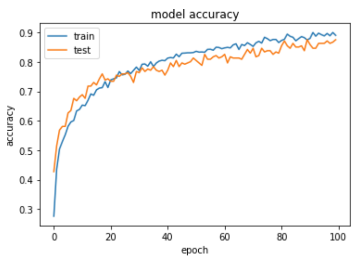

FIGURE 22 – Dropout = 0.8

As a result, I will choose a value of 0.7 because the validation accuracy stays high while the gap between the training and validation accuracies is small (6%).

4.6.5 Batch size

4.6.5.1 Hypothesis

The higher batch size, the higher validation accuracy becomes. We fix some parameters (just a choice) : 3 layers, epochs = 10, Dropout=64, kernel_size=2 and pooling_size=2.

4.6.5.2 Results

Batch Size	Validation accuracy
8	0,8921
10	0.9004
12	0,8713
15	0,8838
20	0,9045
30	0,8755
32	0,9087
40	0,7634
45	0,8506
50	0,8547
55	0,8547
60	0,8298
64	0,7427
70	0,7759
75	0,8340
80	0,7883
85	0,7883
90	0,8049

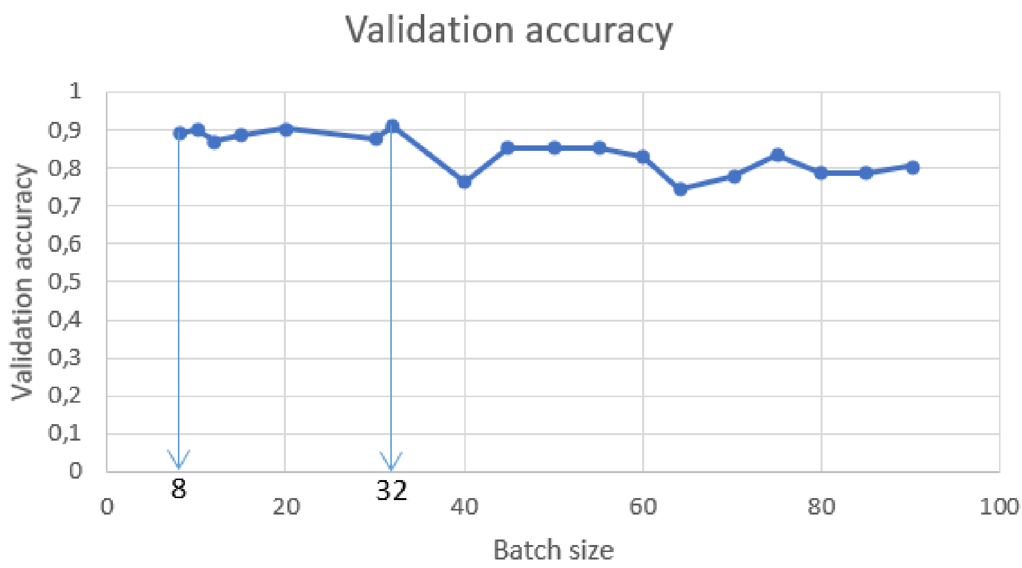

FIGURE 23 – Evolution of validation accuracy in function of the batch size

The validation accuracy increases between two values 8 and 32, and it fluctuates after that. The maximum value is reached at 32.

4.6.5.3 Conclusion

I will choose as a value of the batch size 32 since it gives the higher validation accuracy.

4.6.6 Comments

The experiments established previously allow us to choose wisely the best values of the tree hyper-parameters :
- Epochs : 350 ;
- Dropout : 70% ;
- Batch size : 32.

Our first model will be called model from scratch, because we define by ourselves the layers and assigned values to each one of them.

4.7 Transfer Learning

4.7.1 Object

The object of this section is to compare the validation accuracy of different pretrained models, model from scratch and the augmented one.

4.7.2 Motivation

Humans search always to benefit from their previous knowledge on doing new tasks, Conventional machine learning and deep learning algorithms, so far, have been traditionally designed to work in isolation. These algorithms are trained to solve specific tasks. The models have to be rebuilt from scratch once the feature-space distribution changes. Transfer learning is the idea of overcoming the isolated learning paradigm and utilizing knowledge acquired for one task to solve related ones.
Transfer learning can be implemented on a small data that contains items that are not equally distributed to the sub-folders.

4.7.3 Pretrained Models

We can take a pretrained image classification network that has already learned to extract powerful and informative features from natural images and use it as a starting point to learn a new task. The majority of the pretrained networks are trained on a subset of the ImageNet database, which is used in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) . These networks have been trained on more than a million images and can classify images into 1000 object categories, such as keyboard, coffee mug, pencil, and many animals. Using a pretrained network with transfer learning is typically much faster and easier than training a network from scratch. We will use six pretrained models (Xception, VGG19, VGG16, Inception V3, MobileNetV2).

4.7.3.1 VGG16 and VGG19

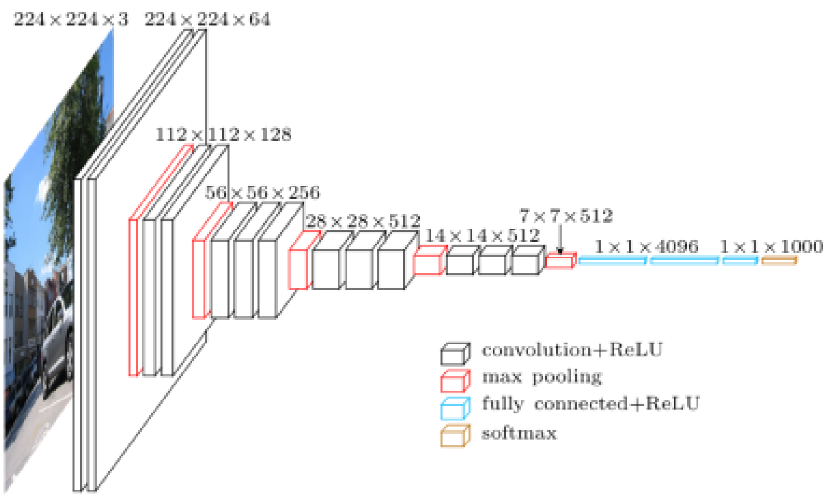

Figure 24 – A visualization of the VGG architecture

This network is characterized by its simplicity, using only 3×3 convolutional layers stacked on top of each other in increasing depth. Reducing volume size is handled by max pooling. Two fully-connected layers, each with 4,096 nodes are then followed by a softmax classifier.
The “16” and “19” stand for the number of weight layers in the network.

4.7.3.2 Inception V3

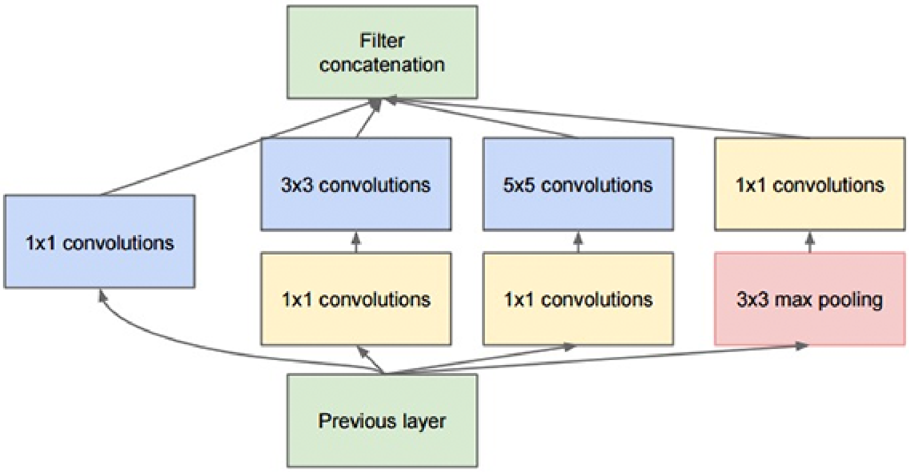

FIGURE 25 – The original Inception module used in GoogLeNet

The goal of the inception module is to act as a "multi-level feature extractor" by computing 1×1, 3×3, and 5×5 convolutions within the same module of the network — the output of these filters are then stacked along the channel dimension and before being fed into the next layer in the network.

4.7.3.3 Xception

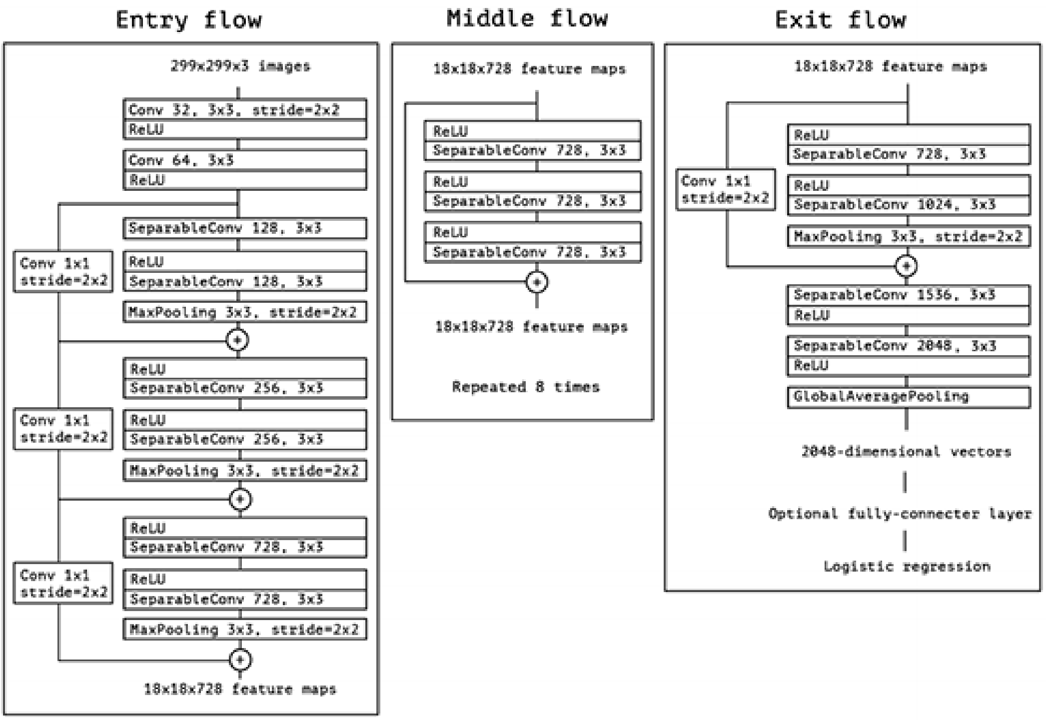

FIGURE 26 – The Xception architecture

Xception was proposed by none other than François Chollet himself, the creator and chief maintainer of the Keras library.
Xception is an extension of the Inception architecture which replaces the standard Inception modules with depthwise separable convolutions.

4.7.4 Implementation

We will try now to implement each of the pretrained models described before, the split ration will be 0.3 between the train and validation set.

Model	Epochs	Maximum validation accuracy
Xception	63	0.9274
VGG19	196	0.9622
VGG16	192	0.9535
Inception V3	78	0.9259
MobileNetV2	53	0.8955
From scratch	112	0.8682
From scratch (augmented)	122	0.8447

We note that the Model 'From scratch (augmented)' has the same architecture as 'From scratch' but different is in the dataset in terms of number of samples, we called that data augmentation which is a technique that can be used to artificially expand the size of a training dataset by creating modified versions of images in the dataset.
The modification includes a range of operations from the field of image manipulation, such as shifts, flips, zooms, and much more.

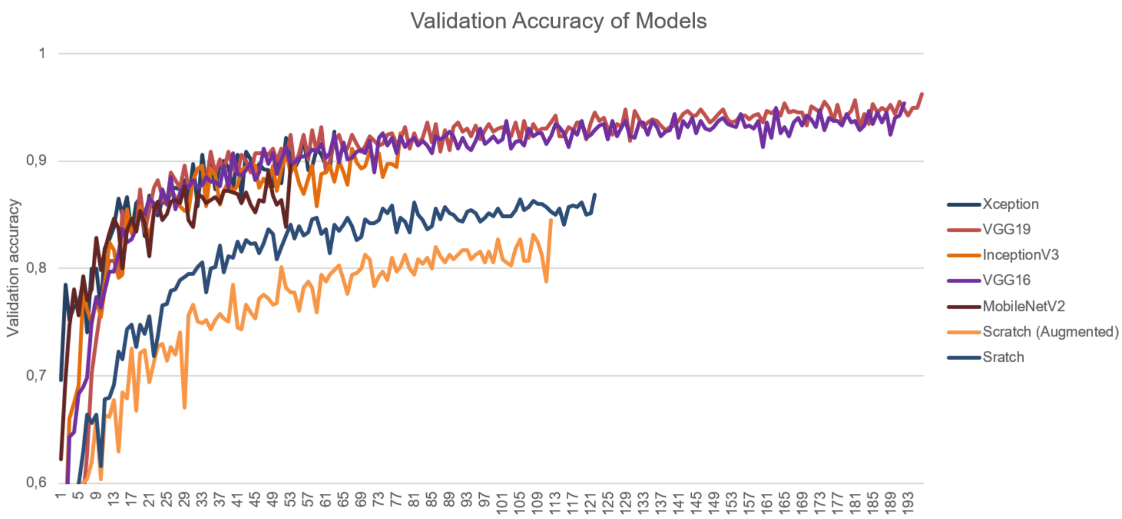

Figure 27 – The Xception architecture

The previous graph shows the validation accuracy evolution for all models, obviously the best ones in terms of accuracies are VGG19 and VGG16.

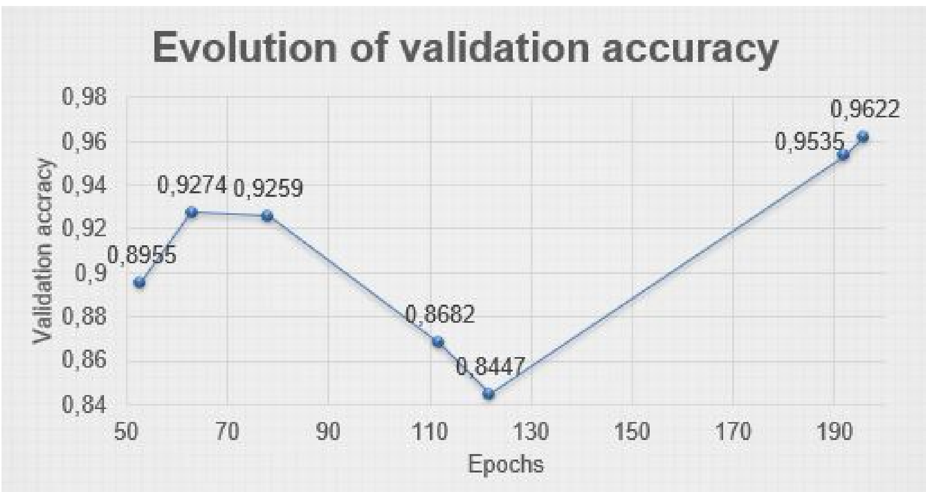

Figure 28 – Evolution of maximum validation accuracy for all models

In the last graph we have two options :

- Choose the model (VGG19) that gives the higher validation accuracy but more time consuming for the training (196 epochs).
- Choose the model (Xception) that gives a validation accuracy higher than 90% and less time consuming (78 epochs).

4.8 Testing the model

4.8.1 The process

We will choose the pretrained model VGG19 for predicting the new images.

- The first step is to resize the input image to the default shape of the model (224x224), and convert it into an array of pixels.

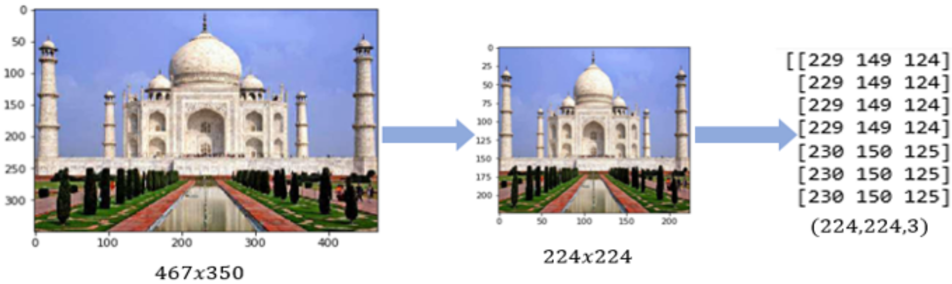

Figure 29 – The resizing process

- The model attributes for each class a certain probability.

Figure 30 – Set of the probabilities for each class

- The model chooses the higher probability and gives it's name as an output.

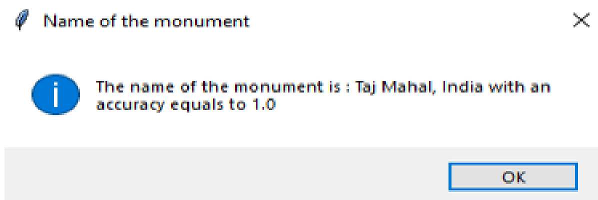

Figure 31 – Name of the monument

- The final feature is getting the location of the monument on a map.

Figure 32 – Name of the monument

4.8.2 Graphical User Interface

4.8.2.1 Description

All the features described previously are gathered in one Graphical User Interface as shown in the image below.

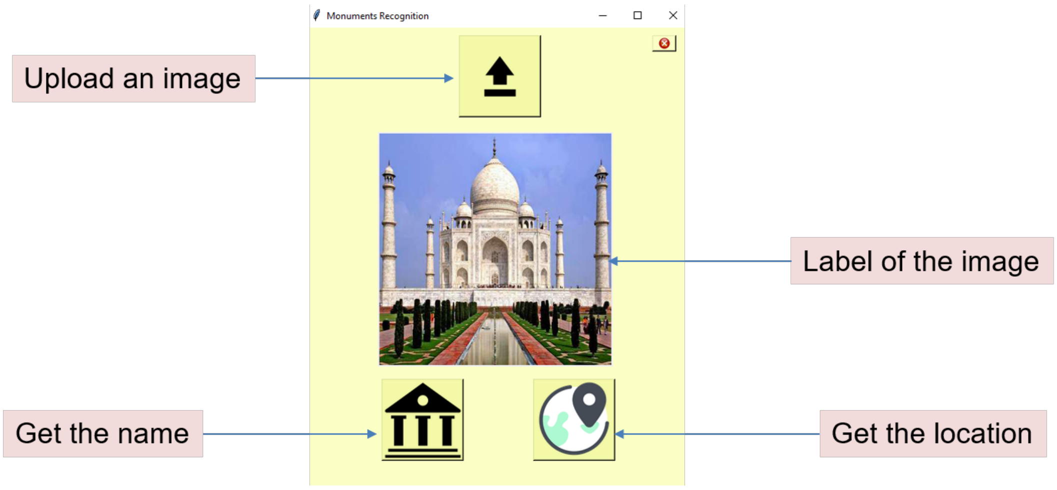

FIGURE 33 – Graphical User Interface

4.8.2.2 Extreme cases

We can have a different input images that don't represent the monuments (such as : flowers, animal, humans...). To deal with that, we define a threshold for the probability of belonging to a class.

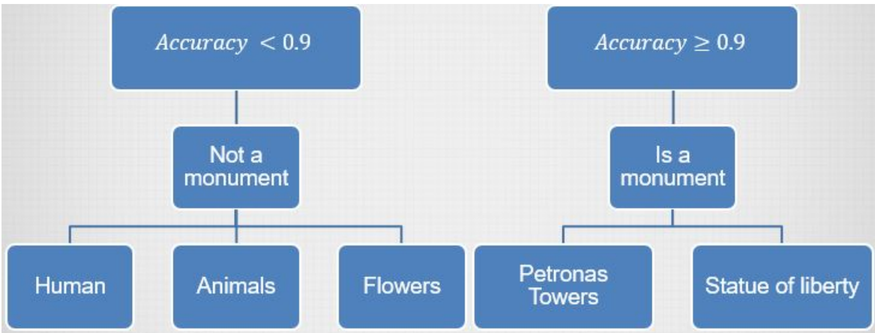

FIGURE 34 – Process of selecting the right class

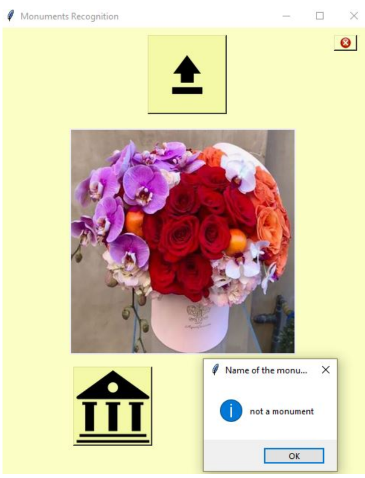

FIGURE 35 – Flowers

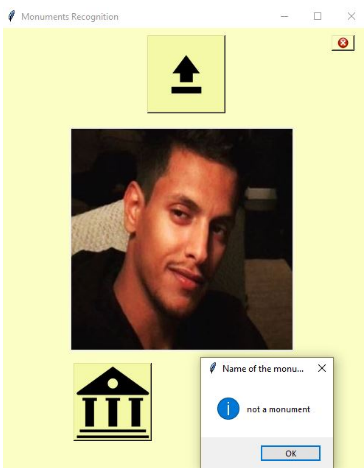

FIGURE 36 – Humans

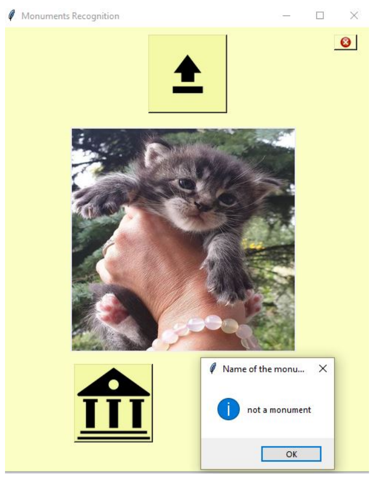

FIGURE 37 – Animals

4.8.2.3 Prospects

— Add more monuments in the dataset ;
— Add a button for training the new images ;
— Design mobile application for the interface.

4.9 Presentation of the project

By the end of the internship, i presented my work in presence of my supervisors and the dean of the university. i shared with them the Moroccan culture and highlighted the difficulties that i suffered from during my stay in Malaysia.

Figure 38 – Certificate of completion of internship