

# **Apprentissage Profond**

## Rapport

### Constitution de la base de données

Paul ANACLET  
Nicolas CORMARY  
Sylvain CROUZET  
Théo DESPRATS

2A SN - L12

2022

# I - Présentation du projet

## 1. Informations générales

Le projet “Capitals Learning”, dont il est question dans ce rapport, sera porté par le groupe 1 des L1-L2, composé de Paul Anaclet, Nicolas Cormary, Sylvain Crouzet et Théo Desprats.

Nous allons, dans la suite de ce rapport, décrire le sujet de ce projet ainsi que les méthodes employées pour constituer la base de données (BDD) relative à celui-ci.

## 2. Description du sujet

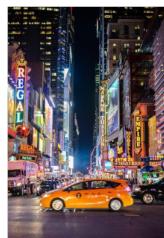
L'objectif de ce projet est d'entraîner un réseau de neurones (perceptron multicouches) pour la classification d'images de villes et notamment de capitales du monde. Comme montré dans la Figure 1 ci-dessous, l'intelligence artificielle devra, à partir d'une image d'une rue ou d'un monument, reconnaître et prédire dans quelle ville se situe cette rue ou ce monument.



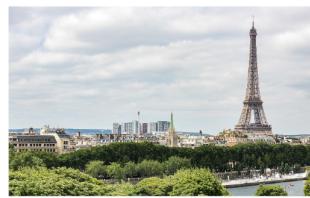
Figure 1: Représentation des résultats attendus

## II - Constitution de la base de données

**New York**



**Paris**



**Tokyo**



### 1. Acquisition et annotation des données

Nous avons pu acquérir une première partie des données en “scrappant” des images provenant de Google Images grâce au script suivant écrit par l’utilisateur “ohyicong” sur GitHub : <https://github.com/ohyicong/Google-Image-Scraper>.

Ce script nous a permis d’obtenir des images de villes à l’aide des mots clés de recherche suivants : “[nom de la ville]”, “[nom de la ville] ville” et “[nom de la ville] rue” que nous avons stockées dans notre BDD hébergée sur GitHub :

<https://github.com/TCD32/Projet-CapitalsLearning/tree/main/data>

L’annotation s’effectuera simplement en associant à l’image le nom de la ville à laquelle elle appartient par le biais du script de chargement fourni et légèrement remanié disponible sur le lien suivant :

[https://github.com/TCD32/Projet-CapitalsLearning/blob/main/src/features/load\\_data.py](https://github.com/TCD32/Projet-CapitalsLearning/blob/main/src/features/load_data.py)

Le nombre de classes de résultats différentes (villes) sera initialement peu élevé le temps de la calibration puis nous comptons l’augmenter au fur et à mesure et observer l’évolution des résultats.

## 2. Partition des images en ensembles d'entraînement / validation / test

Nous comptons répartir les images de la manière suivante pour chaque classe d'image différente : environ 50% dans ‘training’, environ 25% dans ‘test’ et environ 25% dans ‘validation’. Comme nous avons à peu près 500 images par ville, nous pensons prendre environ 250 images pour l’entraînement du réseau, en nous assurant d’avoir des images de nuit comme de jours, mais aussi de monuments, comme la tour Eiffel et de quartiers.

### III - Résultats attendus

Pour ce qui est de la prédiction de la ville lorsque certains bâtiments, monuments, emblématiques sont présents nous nous attendons à de bons résultats. La reconnaissance de Paris grâce à la tour Eiffel ou de l'arc de triomphe semble facile, comme celle de Tokyo grâce à la tokyo tower, ou New-York grâce à Times Square.

En revanche, nous sommes beaucoup moins confiant pour la reconnaissance des quartiers, les rues de Tokyo et de New-York notamment sont très récentes, compactes et bétonnées, donc les ressemblances sont moins évidentes qu'elles pourraient l'être avec Paris où il y a davantage de bâtiments à base de pierre.

Nous avons également noté certaines images qui si elles sont placées dans la base de donnée de vérification ou de test seront assurément fausses, comme par exemple celle-ci qui représente le Washington Square Arch, avec derrière le One World Center, et nous avons peu d'espoir que le réseau de neurone reconnaisse ce dernier derrière, et ignore l'apparent arc de triomphe au premier plan.



De plus, nous avons deux types d'environnement bien distincts pour chaque ville. En effet, il y a majoritairement des photos aériennes des villes ou bien des photos des rues qui les composent. Cela mènera sûrement à de moins bons résultats que si on ne s'était concentré sur qu'une seule des deux catégories.

## IV - Conclusion

Pour conclure, nous pensons que nous avons un problème plutôt complexe de par les éléments cités dans la partie III. Nous pensons donc que nous aurons des résultats vers les 75 % de bonnes réponses mais que ce pourcentage pourrait chuter énormément si on ajoutait d'autres villes notamment si elles sont proches des villes déjà choisies (Tokyo et Kyoto, Paris et Bordeaux, ...)