



Rapport n°1

Projet Apprentissage profond

Classification d'images selon le nombre de pokémons, grâce un à réseau de neurones

Groupe A

AFKER Samy
BAURIAUD Laura
BOCANDÉ Thomas
GUIHUR Lilian
MALINGE Romain

Table des matières

1	Description du sujet	3
1.1	Préambule	3
1.2	Énoncé	3
2	Acquisition des données	4
2.1	Images générées	4
2.2	Images du jeu	5
3	Partition des données	6
4	Notre pronostic	6

Table des figures

1	Échantillon d'images générées	4
2	Échantillon d'images du jeu	5

1 Description du sujet

1.1 Préambule

Un matin, Samy s'est réveillé et comme à son habitude il a commencé à compter ses pokémons pour être sûr qu'ils étaient tous là. Malheureusement, il a constaté qu'il avait perdu ses lunettes !

Pour aider le pauvre Samy, nous nous proposons de construire un réseau de neurones grâce au conseils du grand Professeur Carlier.

1.2 Énoncé

Le but de notre projet est de classifier à l'aide d'un réseau de neurones des images suivant le nombre de Pokémons présents dessus. Nous nous sommes donc limités aux quatres classes suivantes :

- 1 Pokémon
- 2 Pokémons
- 3 Pokémons
- 4 Pokémons

Nous ferons l'entraînement et la validation de notre réseau sur des images générées puis nous le testerons sur des images provenant du jeu d'origine (Pokémon version Noir). Vous pouvez retrouver notre base de données sur le [GitHub de notre projet](#).

2 Acquisition des données

2.1 Images générées

Pour les ensembles d'entraînement et de validation, nous avons récupéré une trentaine d'arrière-plan sur des sites divers ainsi que les sprites avant et arrière de 649 pokémons sur le site [Pokekalos](#).

Ensuite, nous avons implanté un programme [generateur.py](#) qui permet de générer un ensemble d'images contenant des Pokémons avec les paramètres suivants :

- Nous choisissons le nombre de Pokémons, de 1 à 4, par image
- Les Pokémons sont placés de façon aléatoire
- Les Pokémons peuvent déborder de $\frac{1}{3}$ de leur taille sur les bords de l'image
- Les Pokémons ont une proximité maximale de $\frac{1}{2}$ de leur taille (ce qui permet la superposition de Pokémons)
- Les Pokémons sont agrandis avec une échelle aléatoire entre 1 et 1,5
- Les Pokémons sont placés de face ou de dos de façon aléatoire
- L'images finale a une dimension de 256 x 192

Le programme est écrit de manière à enregistrer les images dans le dossier de la classe correspondante ce qui permet d'annoter les données automatiquement. Voici un échantillon d'images générées par notre programme :



FIGURE 1 – Échantillon d'images générées

2.2 Images du jeu

Pour l'ensemble de test, les images ont été obtenues directement dans le jeu *Pokémon version Noir* sur l'émulateur DeSmuME. Les captures d'écran ont été réalisées avec le script [auto-screenshot.py](#), puis elles ont été recadrées avec le script [rognage.py](#).

Pour les images du jeu, nous avons fait une annotation manuelle. Voici un échantillon des images obtenues du jeu.



FIGURE 2 – Échantillon d'images du jeu

3 Partition des données

Comme nous disposons d'un programme pour générer automatiquement des images, nous avons pu prendre un grand nombre d'images. Notre répartition pour le projet est la suivante :

Nom de l'ensemble	Nombre d'images	Précision
Entraînement	4000	Générer par le programme
Validation	1000	Générer par le programme
Test	100	Issues du jeu <i>Pokémon Version Noire</i>

TABLE 1 – Partition des données

Nous avons choisi de placer un grand nombre d'images dans l'ensemble d'apprentissage pour augmenter nos chances que le réseaux réponde correctement au problème. Il y a peu d'images dans l'ensemble de test car les possibilités de disposition sont limités.

4 Notre pronostic

Le problème du projet est relativement difficile car certains Pokémon peuvent se superposer ou sortir de l'image. Cela rend leur dénombrement plus difficile même pour un humain. De plus, certains Pokémon ont des couleurs très proches du fond, rendant leur identification encore plus difficiles.

Nous pensons que notre réseau de neurones parviendra à bien classifier les images que nous avons générées. Cependant, il est possible que les résultats sur l'ensemble de test soient moins bons car celui-ci est composé d'images avec des éléments supplémentaires comme des boîtes de texte ou des effets de particules.