Projet d'apprentissage faiblement supervisé IA et Multimedia

A. Carlier

2021

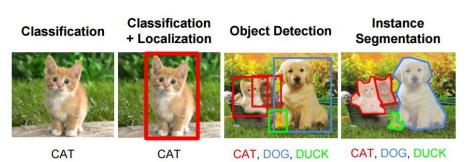
Données : des images...

- ullet 50 000 images de taille 64 imes 64.
- Représentant des animaux de 10 classes différentes (à vous de les identifier...)
- Pour corser le problème, les classes représentent parfois des animaux qui se ressemblent beaucoup!



Problème à résoudre : à la carte

- Classification d'image (/14)
- Détection d'objet par boîte englobante (/17)
- Segmentation d'image (/20)



Articulation du travail

Deux temps principaux :

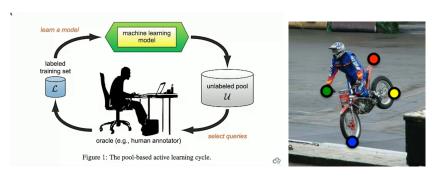
- Annotation des données
 - Un peu... (< 1000)
 - Beaucoup... (< 5000)</p>
 - 3 Passionnément! (> 5000)
- Entraînement du modèle
 - Approches auto-supervisées pour pré-entraîner un réseau, puis transfer learning sur une petite base supervisée.
 - Approches semi-supervisées (avec consistency loss, type Π-Modèle ou FixMatch) si vous avez assez peu de données supervisées.
 - Approches semi-supervisées (avec pseudo-labellisation, type Noisy-Student) si vous avez beaucoup de données supervisées.

Choix du degré d'annotation et de la méthode d'entraînement à la carte!

Pour l'annotation :

cf. Cours du 19/10/2021 :

- N'hésitez pas à considérer des approches d'apprentissage actif.
- Pensez aux interfaces d'annotation open-source.
- Pensez également aux interfaces "intelligentes".



Informations en vrac

- Vous travaillerez par **groupes de 3** (à former pendant la séance du 20/10).
- Vous disposerez de ressources (articles de recherche dont vous pouvez vous inspirer) sous Moodle.
- A mi-chemin du projet, je vous fournirai une base de test (labellisée par mes soins...)
- Critères d'évaluation :
 - Performance de votre réseau
 - Qualité du travail
 - ► Rapport : celui-ci doit me permettre, en plus de bien comprendre ce que vous avez fait, d'éventuellement pouvoir reproduire votre travail!