Rapport intermédiaire

# Préambule

Ce rapport intermédiaire contient l’état de l’art des technologies disponibles et utiles à ce travail de Bachelor. Il est divisé en plusieurs sections, chacune traitant un bloc composant l’application.

Pour rappel, nous parlerons ici des différents types de modèle deep learning se prêtant au mieux au contexte énoncé, puis nous réaliserons une synthèse des différents jeux de données trouvés en exposant leurs forces et limitations. Une fois ces deux éléments choisis, nous pourrons statuer sur une architecture pré-entrainée adaptée aux mobiles.

Ensuite, nous nous concentrerons sur la recherche de solutions existantes ou non permettant la réalisation de l’intégration d’un modèle deep learning entrainé à une application cross platform. Dans un premier temps, cette section sera indépendante du modèle deep learning choisi puisque nous sommes, dans l’idéal, à la recherche d’une solution faisant abstraction de la couche machine learning.

Finalement, une analyse architecturale mettra les éléments retenus jusqu’ici en commun et explicitera leurs intégrations dans l’application prototype finale.

# Modèles deep learning

Plusieurs fois au cours de ce rapport, nous avons fait mention de modèle « deep learning ». Notons ici que le terme est générique et ne définit pas un modèle à proprement parlé. Il s’agit plutôt d’une famille d’apprentissages automatiques fondée sur l’apprentissage de représentations de données[[1]](#footnote-1). Dit autrement, le terme définit une technique utilisée par un ensemble de modèles comme l’extraction de caractéristique dans une image ou dans un son. Toutefois, par souci de simplicité, nous utiliserons ce terme pour parler de modèles de machine learning traitant des images au travers d’un réseau neuronal convolutif.

Le réseau de neurones convolutifs (CNN[[2]](#footnote-2)) sont un élément fondamental pour la vision par ordinateur. Depuis son invention en 1988 par Kunihiko Fukushima, il a été largement amélioré et agrémenté de nouvelles possibilités au cours des 12 dernières années avec l’augmentation de puissance des processeurs graphiques (GPU) et la démocratisation de l’intelligence artificielle générative. Ils ont donc la possibilité de répondre à diverses tâches comme la classification d’images, la détection d’objet et la segmentation d’images[[3]](#footnote-3). Nous allons développer ces trois dernières tâches car elles représentent des solutions pertinentes à notre situation.

Une image contenant texte, capture d’écran, diagramme

Description générée automatiquement

Figure - Exemples de type de problèmes de vision par ordinateur[[4]](#footnote-4)

## État de l’art

### Classification d’images

Nous entendons par classification d’image, le processus permettant de catégoriser l’appartenance d’une image à une classe parmi celles d’un ensemble prédéfini. Plus particulièrement dans cette tâche, il s’agit de porter un regard sur l’ensemble de l’image sans spécifier au modèle les régions d’intérêts permettant de prédire la classe correctement. Par conséquent, et comme le démontre la Figure 1, cette technique.

#### Avantages

Expliquer ce qui rend ça bien et efficace pour le contexte

#### Limitations

Expliquer ce qui rend difficilement applicable au contexte

### Détection d’objets

idem

#### Avantages

idem

#### Limitations

idem

# Dataset

Complètement dépendant du point 1.

# Architectures de Réseau de neurones

Complètement dépendant du point 1.

# Intégration de modèle dans une application cross-platform

TODO

# Modélisation et architecture de l’application

TODO

# Solution choisie

1. Apprentissage profond (2024, 2 juillet). In *Wikipedia.* https://fr.wikipedia.org/wiki/Apprentissage\_profond [↑](#footnote-ref-1)
2. Convolutional neural networks [↑](#footnote-ref-2)
3. Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., Ghayvat, H., (2021). *CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope.* MDPI. https://www.mdpi.com/2079-9292/10/20/2470 [↑](#footnote-ref-3)
4. Mrinal Walia. (Sep 28, 2022). *Object Detection vs. Image Classification vs. Keypoint Detection*. Roboflow Blog: https://blog.roboflow.com/object-detection-vs-image-classification-vs-keypoint-detection/ [↑](#footnote-ref-4)