

Structure du Mémoire

Methods for large-scale image classification and application to biomedical data

BUI QUANG PHUONG Quang Linh - ULB ID : 000427796
MA2 Computer Sciences

Novembre 2019

Ce qui a été réalisé en **Master 1** :

1. **Introduction** : présentation des concepts, historique, des exemples d'application.
2. **General concepts of AI** : présentation du machine learning et deep learning, supervised/un-supervised learning, définition du problème de classification d'images.
3. **Material and Methods** : présentation des neural network, CNN, SVM ainsi que des méthodes d'évaluation et de validation.
4. **Breast cancer detection** : présentation d'un cas connu à partir d'un article, présentation de ses méthodes et résultats

Ce qui va être réalisé en **Master 2** :

5. **Model prototype** : création d'un premier modèle et application sur des données non-médicales (CIFAR10). Présentation du dataset CIFAR10, présentation du modèle, des différents paramètres:
 - Tester différents optimiseurs : adam, SGD, ...
 - Tester avec et sans dropout + test de différentes valeurs de dropout
 - Tester avec weight decay (kernel_regularizer)
 - Tester avec data augmentation
 - Tester différentes valeurs de hyperparameters (epochs, batch_size, ...)

→ Manipuler et comparer/présenter les résultats obtenus en fonction de ces différentes valeurs et voir quelle configuration obtient les meilleurs résultats. + justification en expliquant les différents paramètres / variables.
6. **A specific case: Identification of Leukemia Subtypes** (ou autre article/problème): présentation du problème en question, essayer notre modèle sur le dataset de l'article, voir si les paramètres optimisés sont les mêmes qu'avec un dataset classique (CIFAR10), appliquer les différentes méthodes/filtres utilisés (data augmentation) dans l'article afin de reproduire (voir améliorer) les résultats si possible.
7. **Exclusive data analysis** : création d'un modèle pour des données exclusives, présentation des données et des résultats obtenus. (encore un peu flou, à compléter)

8. **Interactive image labelling** : progresser sur l'annotation automatique d'images (active learning) (encore un peu flou, à compléter)
9. **Combining different types of input data** : combinaison de différents types d'input lors de la création du modèle → Mixed Data (encore un peu flou, à compléter)

N.B: L'ordre des sections peut être modifié au cours du travail.

A première vue, les points 5 ainsi qu'une partie (voir totalement) du point 6 vont être réalisés pendant le premier quadri tandis que les 3 derniers points seront réservés au second quadri.