TP Réseau de neurones convolutifs: CNN

1 Comptes rendus à rendre

1.1 Objectif

Dans ce projet nous allons développer un réseau de neurones capable de classer des photos de chats et de chiens. Bien que le problème semble simple, il n'a été résolu efficacement que ces dernières années grâce à l'utilisation de réseaux neuronaux convolutifs à apprentissage profond. Nous allons apprendre et pratiquer comment développer, évaluer et utiliser des réseaux neuronaux convolutifs d'apprentissage profond pour la classification d'images à partir de zéro. Pour se faire, nous étudierons les couches classiques de ce type de réseaux et nous mettrons en place un premier réseau appris sur le jeu de données standard images de chats et de chiens extraites à partir du challenge Kaggle www.kaggle.com/c/dogs-vs-cats/data. Une fois le premier réseau testé nous aborderons différentes techniques permettant d'améliorer le processus d'apprentissage du réseau : normalisation, data augmentation, dropout et batch et transfert learning.

L'objectif de ce TP est de vous familiariser avec :

- La préparation des photos de chiens et de chats pour l'apprentissage.
- Le développement un réseau neuronal convolutif pour la classification de photos à partir de zéro et l'amélioration des performances du modèle par la data augmentation.
- L'amélioration du processus d'apprentissage par les techniques de transfert learning.

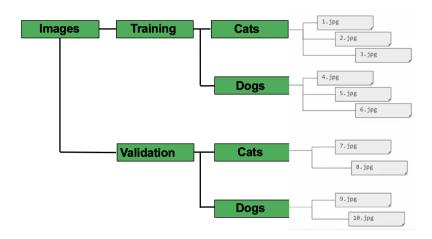
2 Préparation des données

L'ensemble de données "chiens/chats" fait référence à un ensemble de données utilisé pour un concours d'apprentissage par machine Kaggle organisé en 2013. L'ensemble de données peut être téléchargé gratuitement sur le site web de Kaggle https://www.kaggle.com/c/dogs-vs-cats/data. Si vous n'avez pas de compte Kaggle, inscrivez-vous d'abord. Téléchargez l'ensemble de données en vous rendant sur la



FIGURE 1 – Quelques images de chats et de chiens extraits de www.kaggle.com/c/dogs-vs-cats/data

page données sur les chiens et les chats et cliquez sur le bouton "Télécharger tout". Cela permettra de télécharger le fichier de 850 mégaoctets "chiens-vs-cats.zip" sur votre poste de travail. Dézippez le fichier et vous verrez train.zip, test1.zip et un fichier .csv. Dézippez le fichier train.zip, car nous nous concentrerons uniquement sur cet ensemble de données. Vous aurez alors un dossier appelé "train/" qui contient 25 000 fichiers .jpg de chiens et de chats. Les photos sont étiquetées par leur nom de fichier, avec le mot "chien" ou "chat". La figure 2 propose une convention de dénomination des fichiers.



 ${\tt Figure}\ 2-{\tt Organisation}\ des\ données\ d'apprentissage\ et\ de\ validation.$