

### Présentation de la thématique :

Dans une image, l'information est transmise à différentes fréquences : les hautes fréquences vont plutôt représenter les détails de l'image alors que les basses fréquences vont représenter plutôt la structure globale. Partant de constat, l'idée est de séparer ces différentes composantes dans les couches de convolution d'un CNN et de les traiter différemment. Les informations basses fréquences qui évoluent plus « lentement » peuvent être traitées avec une résolution plus faible, réduisant ainsi la mémoire utilisée ainsi que le temps de calcul. Contrairement aux autres méthodes « multi-échelles », celle-ci, appelée Octave Convolution (OctConv), est assez transparente puisqu'elle peut remplacer de manière générique l'opérateur de convolution classique sans avoir à toucher à l'architecture du réseau.

Il se trouve qu'en plus de réduire la consommation de mémoire et le coût de calcul, cette méthode a expérimentalement obtenue de meilleures performances de classification par-rapport à la plupart des CNN les plus performants actuellement et sur le dataset d'ImageNet.

### Objectif :

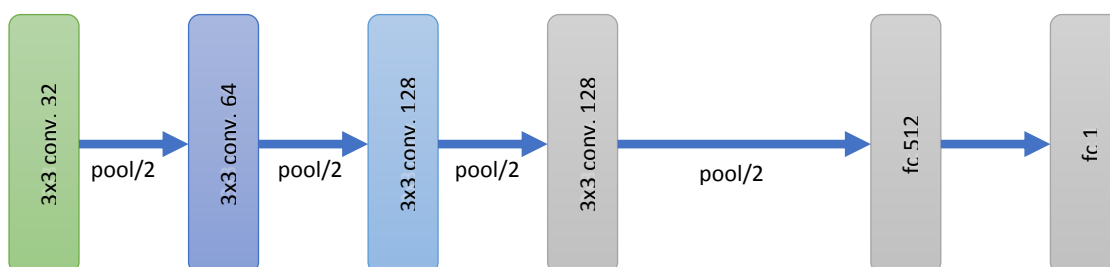
Mettre en place et tester cette couche de convolution OctConv sur le jeu de données du projet 7, et sur une architecture de CNN différente.

### Données :

Le dataset est celui du dernier projet, le [Stanford Dogs Dataset](#) qui contient 20 580 images de chiens pour 120 races différentes, et dont l'objectif est de réussir à classifier les différentes races correctement.

### Baseline :

Pour comparer les résultats de la méthode OctConv, on utilisera un modèle utilisé lors du projet précédent. A savoir, un CNN composé de 4 couches de convolutions (avec max pooling à chaque fois), suivies de 2 couches fully-connected.



### Sources :

- [Drop an Octave : Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution](#) (arXiv, avril 2019)
- [Multigrid Neural Architectures](#) (arXiv, mai 2017)