

CNN in compressed domain

Thierry FLÜCK

Travail de Bachelor 2023

Filière Informatique et Systèmes de Communication - Orientation Ingénierie des Données

Professeur: Javier BRACAMONTE

Expert: Jean-Christophe Mathez

Description

La classification d'image consiste à donner à un ordinateur une image et lui demander ce qu'il voit dessus. Par exemple on pourrait vouloir savoir quel animal apparaît sur une photo.

La classification d'images est une tâche fréquente qui peut nécessiter des flux de données importants ainsi qu'une grande puissance de calcul.

Il est donc intéressant de trouver une solution pour réduire les besoins inhérents à cette tâche en effectuant le traitement de ces images sur des appareils n'ayant pas une grande puissance de calcul sans perdre en qualité de prédiction.

Ce projet vise à étudier s'il est bénéfique d'utiliser un modèle de classification ne traitant que des images dans le domaine compressé pour réduire les besoins en puissance de calcul.

Déroulement

Description

L'objectif de ce projet est de trouver des modèles de classification qui fonctionnent avec moins de ressources. C'est-à-dire moins de mémoire, moins de puissance de calcul et moins de temps.

Déroulement

Il a tout d'abord fallu trouver une méthodologie permettant de déterminer quel programme (on parle plutôt de modèle en général) est mieux qu'un autre. Ensuite, il a fallu développer ces modèles, puis pour finir, les tester et les comparer.

Résultats

Résultats

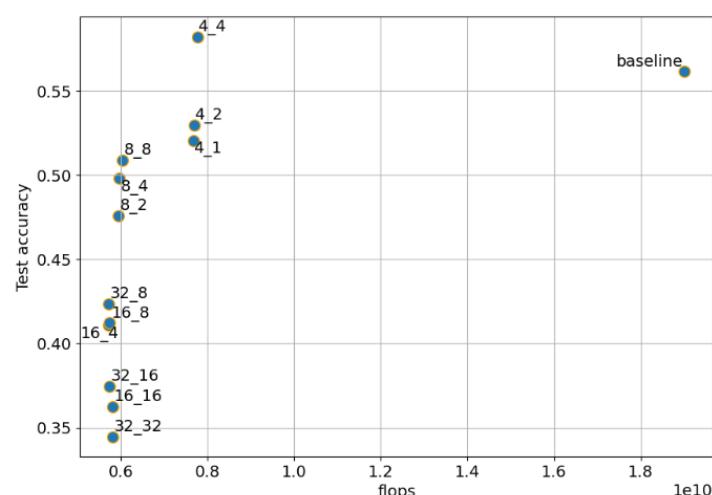
Nous avons réussi à trouver des modèles qui demandent 2,4 fois moins de puissance de calcul que ceux qui fonctionnent avec des images normales et qui classifient correctement 4% plus souvent.

Voici quelques explications pour mieux comprendre le graphique : tout d'abord, chaque point représente un modèle. Le point labellisé « baseline », c'est le modèle qui utilise des images normales. Les autres sont des modèles avec lesquels on a utilisé des données dans le domaine compressé. Chaque modèle a utilisé des données qui ont été transformées différemment.

Sur l'axe vertical, se trouve le score de classification, c'est-à-dire le rapport entre le nombre total de classifications et les classifications correctes. Plus le modèle est haut dans le graphique, mieux c'est.

Sur l'axe horizontal, on retrouve le nombre de calculs nécessaires. Plus un modèle est à gauche sur le graphique, mieux c'est.

Test accuracy vs flops with model ResNet18 on cifar10



Discussion : Conclusions et perspectives

Il est possible que nous ayons trouvé des modèles de classification d'images qui, avec un peu plus de recherches, pourraient grandement réduire les ressources nécessaires aux tâches de classification d'images.

De plus amples recherches seraient nécessaires pour trouver des algorithmes réellement efficaces. Et nous espérons que d'autres personnes décideront d'investiguer le sujet.