

Problématique d’instanciation de réseaux de neurones profonds sur système embarqué pour la vision multi-vues.

Direction : Encadrant : François BERRY – PR, **Co-encadrant :** Christophe BLANC - MCF
Équipes associées : Axe ISPR, équipe DREAM

Résumé : Les Réseaux de Neurones Convolutifs profonds (CNNs) ont connu un large succès au cours de la dernière décennie, devenant un standard de la vision par ordinateur. Ce succès s’est fait au détriment d’un large coût de calcul, rendant le déploiement des CNNs une tâche ardue, surtout sous des contraintes de temps réel. En effet, les techniques d’apprentissage profond sont très intensives en terme de calculs, impliquant jusqu’à 30 milliards d’opérations pour classer une seule image. Ces charges de calculs élevées empêchent l’implémentation de l’apprentissage profond dans des systèmes alliant performances temps réels et faible consommation énergétique

Afin de rendre ce déploiement possible, la littérature exploite le parallélisme important de ces algorithmes, ce qui nécessite l’utilisation de plateformes matérielles dédiées. Dans les environnements soumis à des contraintes de consommation énergétique, tels que les nœuds des caméras intelligentes, les cœurs de traitement à base de FPGAs sont reconnus comme des solutions de choix pour accélérer les applications de vision par ordinateur. Ceci est d’autant plus vrai pour les CNNs, où les traitements se font naturellement sur un flot de données, rendant les architectures matérielles à base de FPGA encore plus pertinentes.

Une première thèse a abordé les problématiques liées à l’implémentation des CNNs sur FPGAs. En particulier, ces travaux ont visé à améliorer l’efficacité des implantations grâce à deux principales stratégies d’optimisation : la première explore le modèle et les paramètres des CNNs, tandis que la seconde se concentre sur les architectures matérielles adaptées au FPGA.

Dans la présente thèse, nous proposons d’aborder le problème sous une forme distribuée en termes de calcul et de perception. L’idée est donc d’utiliser plusieurs caméras avec un calculateur léger embarqué pour percevoir une scène sous différents points de vues. De cette manière, nous essaierons de motiver que le gain en informations visuelles peut permettre une dégradation des réseaux embarqués au sein de chaque caméra.

Ceci constitue un premier axe de cette thèse où il pourrait être intéressant d’intégrer la géométrie des différentes caméras comme informations supplémentaires.

Toutefois, il n’est actuellement pas envisageable d’intégrer un réseau complet dans une série de caméras distribuées et il est nécessaire de devoir fusionner ces différents flux en un calculateur central. Pour cela, nous ambitionnons d’utiliser un système SOC/FPGA utilisant la reconfiguration dynamique partielle de la zone FPGA. Cette étude sur les architectures reconfigurables dynamiquement pour les CNN constitue donc le second axe de cette thèse.