

Régression logistique pour la classification d'images à grande échelle

Thanh-Nghi Do*, François Poulet**

*Université de Can Tho, Rue 3/2, Ninh Kieu,
Can Tho, Vietnam, dtngchi@cit.ctu.edu.vn

**Université de Rennes I - IRISA
Campus de Beaulieu, 35042 Rennes Cedex, France
francois.poulet@irisa.fr

Résumé. Nous présentons un nouvel algorithme parallèle de régression logistique (PAR-MC-LR) pour la classification d'images à grande échelle. Nous proposons plusieurs extensions de l'algorithme original de régression logistique à deux classes pour en développer une version efficace pour les grands ensembles de données d'images avec plusieurs centaines de classes. Nous présentons un nouvel algorithme LR-BBATCH-SGD de descente de gradient stochastique de régression logistique en batch équilibré avec un apprentissage parallèle (approche un contre le reste) multi-classes sur de multiples cœurs. Les résultats expérimentaux sur des ensembles de données d'ImageNet montrent que notre algorithme est efficace comparés aux algorithmes de classification linéaires de l'état de l'art.

1 Introduction

La classification d'images vise à assigner automatiquement une catégorie prédéfinie à une image. Parmi ses nombreuses applications, on peut citer la reconnaissance de caractères manuscrits, la reconnaissance d'empreintes digitales ou la reconnaissance de visages. Le nombre d'images stockées dans les différentes bases de données ne cesse de croître, par exemple, Facebook contient aujourd'hui plus de 90 milliards d'images, il a été estimé que les utilisateurs moyens d'appareils photos numériques prendront environ 100 000 photos au long de leur vie. Sur internet 65% du volume des données chargées correspond à des images.

La catégorisation d'images par le contenu est un véritable challenge d'importance aujourd'hui. Les approches les plus performantes de ces dernières années utilisent un modèle de "sac de mots visuels" (Bag of Words (Bow)) construits sur des descripteurs locaux des images. Le "sac de mots visuels" est une adaptation du sac de mots utilisé en catégorisation de textes où l'on a un ensemble de documents, chaque document contenant un ensemble de mots. Le modèle de sac de mots correspond aux nombres d'occurrence des mots dans les documents. Dans le cas du sac de mots visuels, on commence par calculer des descripteurs de bas niveau en des points particuliers de l'image. Les plus populaires sont les SIFTs (Scale Invariant Feature Transform (Lowe, 2004)), les SURFs (Speeded Up Robust Features (Bay et al., 2008)) and DSIFTs (Dense SIFTs) (Bosch et al., 2007). Ces méthodes d'extraction de descripteurs