

Exploration visuelle d'espaces paramétriques pour l'analyse d'images

Sujet de Stage de Master 2016-2017

Encadrants :

Romain Vuillemot <http://romain.vuillemot.net/>

Stéphane Derrode <http://perso.ec-lyon.fr/derrode.stephane/>

Contact :

prenom.nom@ec-lyon.fr

Contexte :

Les modèles de traitement automatique (classification, prédiction, ..) deviennent de plus en plus complexes. Ainsi, pour les choisir et configurer, dans le but d'optimiser leur sortie (précision, taux d'erreur, etc.), l'humain est souvent impliqué dans le choix des paramètres d'entrées (comme les valeurs de seuillage, choix de jeux d'essais, etc.). Il utilise ainsi des valeurs classiques ou admises par la communauté, mais doit cependant toujours les valider au moyen d'essais et d'erreurs, ce qui consomme du temps. Des méthodes automatiques existent (grid search, random search, Bayesian optimization, etc.) afin de rechercher les valeurs dans l'espace paramétrique qui est immense, mais l'humain doit encore pour définir des intervalles de valeurs ou une manière de discrétiser les valeurs. Et il est toujours nécessaire d'effectuer des tests.

Ces approches ne sont pas toujours efficaces, pour plusieurs raisons : elles ne permettent pas de comprendre et parcourir complètement l'espace de recherche des paramètres; ne permettent que l'optimisation de paramètres quantitatifs déjà connus; même si une solution optimale est trouvée, il n'est pas possible d'expliquer pourquoi telle décision est prise (dans les cas complexes).

Objectif du stage :

L'objectif de ce stage de Master est de proposer de nouvelles techniques visuelles interactives permettant à l'humain de comprendre rapidement et naviguer de manière efficace dans ces grands espaces de recherche paramétrique. Des exemples de techniques similaires sont Google Tensorflow Playground <http://playground.tensorflow.org/>

Le travail se décompose en 3 parties :

- **La première phase consistera à se focaliser sur des exemples de traitements simple** en analyse d'image (K-means, ..), mais qui présentent déjà une combinatoire importante. Nous mettrons en place un benchmark de méthodes (de

classification par exemple) en décrivant leurs espaces d'entrée et mesures possibles, à la fois liées à l'exécution intermédiaire ou "runtime" (variabilité, ..) et aux sorties "outputs" (temps de calcul, précision, ..).

- **La seconde phase consistera à identifier les tâches pour lesquelles l'humain est performant** (détection de patterns, recherche d'optimums locaux/globaux, validation d'un résultat qualitatif type texte ou image) face à la machine. Cela permettra la conception de visualisations interactives et le test avec le benchmark précédemment constitué.
- **La troisième phase consistera à répéter les phases 1 et 2 afin de généraliser à d'autres modèles et de monter en complexité.** Idéalement, nous souhaitons aller au delà du domaine de l'analyse d'image (et ainsi valider les nouvelles propositions en fouille de données, analyse de graphes, etc.), dans des domaines qui soit partagent les mêmes modèles, soit en utilisant de nouveaux. Les visualisations et interactions seront également à optimiser en performance afin de représenter de grands volumes de données tout en gardant une réactivité du système très rapide (< 100ms).

Le principal rendu attendu est une publication scientifique décrivant les démarches proposées, en les contextualisant dans les domaines de l'analyse d'images, d'interaction homme-machine et de *visual analytics*. Un prototype des interfaces proposées est également attendu pour la phase de validation, avec des jeux de données réels et représentatifs de plusieurs scénarios différents.

En fonction de l'excellence du candidat, le stage de Master pourrait éventuellement se poursuivre dans le cadre d'une thèse.

De bonnes connaissances en programmation sont requises, de préférence en Python et/ou en technologies Web (mais il est possible d'utiliser tout autre langage si le candidat en a une très bonne maîtrise). Un grand sens de la curiosité est souhaité, ainsi que la capacité à se familiariser rapidement avec différents domaines scientifiques et ceci de manière autonome.

Lieu du stage :

Laboratoire LIRIS UMR CNRS 5205 - Ecole Centrale de Lyon (Ecully)

Période de stage :

Février / Juin 2017

Profil recherché :

Master 2 Informatique

Rémunéré :

Oui (~550 euros / mois)