

ED STIC - Proposition de Sujets de Thèse  
pour la campagne d'Allocation de thèses 2011

**Titre du sujet :**

**Mention de thèse :**

**HDR Directeur de thèse inscrit à l'ED STIC :**

---

**Co-encadrant de thèse éventuel :**

**Nom :**

**Prénom :**

**Email :**

**Téléphone :**

---

**Email de contact pour ce sujet :**

**Laboratoire d'accueil :**

---

**Description du sujet :**

Les images de cellule sont acquises selon différentes modalités (microscopes confocaux, bi-photon, électronique) et pour des objectifs variés (présence ou absence d'une protéine, localisation d'un élément tel que l'iode dans le cadre de la recherche contre le cancer...). Une tâche souvent effectuée au cours des études biologiques est la classification des cellules en classes ou catégories pré-définies. L'exemple le plus simple est celui de deux classes : réponse positive d'une cellule (à une expérience) et réponse négative. La classification peut alors être utilisée pour calculer des mesures statistiques ou construire des modèles. Les classes de cellule peuvent correspondre à différents états d'évolution, conduisant à des modèles dynamiques.

En routine d'analyse biologique, la classification est réalisée "manuellement" (c'est-à-dire, visuellement) par les biologistes. Cette tâche est fastidieuse ce qui, en partie, limite le nombre d'expériences qui peuvent être menées. Pour obtenir des mesures statistiques fiables

ou construire des modèles précis, il faut traiter un grand nombre d'images de cellule, ce qui ne peut être raisonnablement fait qu'automatiquement. La classification automatique est un domaine de recherche très actif encore ouvert à de nombreuses améliorations. Un cadre de travail a particulièrement du succès. Il repose sur l'apprentissage supervisé fondé sur les Support Vector Machines (SVM) [1] : étant donné un ensemble d'images (appelé ensemble d'apprentissage) qui ont été classées par un expert, il s'agit de déterminer une fonction (définie par des SVM) qui donnera la classe d'une nouvelle image présentée en entrée. L'apprentissage se compose des étapes suivantes :

- extraction de descripteurs locaux dans les images de l'ensemble d'apprentissage,
- recherche parmi tous les descripteurs locaux d'un ensemble de descripteurs représentatifs dits "prototypes",
- pour chaque image de l'ensemble d'apprentissage, affectation de chaque descripteur local au prototype le plus proche et construction de l'histogramme d'occurrence des prototypes ainsi obtenus. Cet histogramme représente le descripteur global de l'image,
- pour deux classes données, recherche de la séparation optimale par Support Vector Machine (SVM) des descripteurs globaux des images de l'ensemble d'apprentissage dans ces classes.

Lorsque plus de deux classes sont prédéfinies, l'approche "one vs. all" est préférée. Il s'agit de calculer un SVM pour chaque classe en l'opposant à l'union des autres classes.

L'alternative est l'approche "one vs. one" qui nécessite beaucoup plus de SVM.

A noter que les trois premières étapes ci-dessus (technique appelées "bag-of-features" [3]) peut être remplacée par le calcul direct d'un descripteur d'image globale tel que le descripteur Gist [2].

Un fois l'apprentissage effectué, la classification de nouvelles images se fait par le calcul de leur descripteur global (en utilisant les prototypes appris) puis le test d'appartenance (ou les tests s'il y a plus de deux classes) correspondant au SVM (ou aux SVM).

En général, la classification d'images est appliquée à des images classiques telles que des photographies. En imagerie biologique, le traitement d'images de cellule a trois spécificités. Tout d'abord, elles sont souvent bruitées. Ensuite, elles ont tendance à avoir moins de contours francs que des images traditionnelles. Enfin, elles contiennent des éléments visuels moins diversifiés (par exemple en termes de couleur). Cela signifie que des descripteurs et des métriques entre descripteurs doivent être développés spécifiquement pour ce domaine. Cela constituera le premier volet du sujet. Trois options seront explorées : des descripteurs globaux tel que Gist [2], des descripteurs calculés à partir de descripteurs locaux tel que SIFTs (Scale-Invariant Feature Transform) [4] ou SMPs (Sparse Multiscale Patches) [6], et des descripteurs calculés à partir de descripteurs locaux et incluant des caractéristiques de géométrie relative. Cette dernière option sera certainement privilégiée.

Les deuxième et troisième volets du sujet concernent les SVM. Tout d'abord, des alternatives aux approches "one vs. all" (qui nécessite de travailler avec des classes virtuelles union de plusieurs classes) et "one vs. one" (qui nécessite le calcul de beaucoup de SVM) en présence de plus de deux classes vont être étudiées. Ensuite, un SVM étant un séparateur linéaire, il peut être employé pour de la séparation non linéaire uniquement

grâce à un noyau [5] : le produit scalaire entre descripteurs est remplacé par une fonction noyau. Cela a pour effet d'envoyer les descripteurs dans un espace à plus grande dimension où une séparation linéaire est possible. Bien qu'efficace, cette technique est difficile à appréhender. En particulier, à chaque noyau correspond une frontière courbe dans l'espace d'origine. Toutefois, la relation entre noyau et frontière n'est pas triviale. De plus, plusieurs noyaux peuvent produire une séparation correcte des descripteurs. L'étude de cette technique devrait fournir des indications sur la façon de choisir un noyau approprié à des descripteurs donnés ou pourrait aider à en trouver une alternative.

[1] Cortes, Vapnik, Support vector networks, *Machine Learning*, 20:273-297, 1995.

[2] Oliva, Torralba, Modeling the shape of the scene: a holistic representation of the spatial envelope, *International Journal of Computer Vision*, 42:145-175, 2001.

[3] Csurka, Dance, Fan, Willamowski, Bray, Visual categorization with bags of keypoints, *Workshop on Statistical Learning in Computer Vision, ECCV*, 2004.

[4] Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60:91-110, 2004.

[5] Shawe-Taylor, Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, 2004.

[6] Piro, Anthoine, Debreuve, Barlaud, Sparse Multiscale Patches (SMP) for image categorization, *International Conference on Multimedia Modeling*, 2009.

### **English version:**

TITLE: Classification of cell images using machine learning

LABORATORY: Laboratory I3S, Sophia Antipolis (<http://www.i3s.unice.fr/I3S/>)

SUPERVISION: Eric Debreuve

CONTEXT: Collaboration with a research unit in cell biology, Université de Nice-Sophia Antipolis.

BACKGROUND: Cells are imaged using different modalities (confocal, bi-photon, electronic microscopes) and for various purposes (presence or absence of some protein, localization of some element such as iodine in the scope of research on cancer...). One task that is often involved in biological studies is the classification of cells into pre-defined classes or categories. The simplest example corresponds to two classes: positive response of a cell (to some experiment) and negative response. The classification can then be used to compute statistical measures or to build models. The cell classes may correspond to different evolution states, leading to dynamic models.

PURPOSE: In biological routine analysis, classification is performed "manually" (i.e., visually) by biologists. This task is tedious, which, in a way, limits the number of experiments that can be conducted. To obtain more reliable statistical measures or to build more accurate models, it is necessary to process large populations of cells and, therefore, to do it automatically. Automatic classification is an active field of research with still many opportunities for improvement. One framework is particularly successful. It relies on supervised learning using Support Vector Machines (SVMs) [c95]: given a set of images

(called learning set) that have been classified by an expert, determine a function (relying on SVMs) that will output the class of a new image presented as input. The learning is composed of the following steps:

- extraction of local descriptors in the images of the learning set,
- look among all these descriptors for a set of representative descriptors called prototypes,
- for each image in the learning set, assign to each local descriptor the closest prototype and build the histogram of occurrence of these prototypes. This histogram represents the global descriptor of the image,
- For two given classes, determine the optimal separation using an SVM of the global descriptors of the learning set images belonging to these classes.

When there are more than two classes, the "one vs. all" approach is preferred. It corresponds to computing one SVM per class, opposing it to the union of the other classes.

The alternative is the "one vs. one" approach and it requires much more SVMs.

Note that the first three steps above (called the bag-of-features technique [c04]) can be replaced with the direct computation of a global image descriptor such as the so-called Gist descriptor [o01].

Once the learning stage has been done, the classification of new images is achieved by computing their global descriptor (using the learned prototypes) and then testing whether it belongs to one class or another using the learned SVM(s).

Traditional image classification is applied to full-size images such as photographs. In biological imaging, dealing with cell images has three specificities. First, they are usually noisy. Second, they tend to have less sharp edges than traditional images. Third, they exhibit less diversity (for example in terms of colors) than traditional images. This means that descriptors and descriptor metrics have to be tailored to this field. Therefore, this will be the first topic of research. Three options will be explored: global descriptors such as the above-mentioned Gist [o01], descriptors computed from local descriptors such as SIFTs (Scale-Invariant Feature Transform) [l04] or SMPs (Sparse Multiscale Patches) [p09], and descriptors computed from local descriptors and including relative, soft geometric information. More energy should probably be put on this last option.

The second and third topics of research will focus on the SVM framework. First, alternatives to the "one vs. all" (which requires to deal with virtual classes composed of the union of several classes) and the "one vs. one" (which requires to compute a lot of SVMs) approaches will be studied when dealing with more than two classes. Second, an SVM being a linear separator, it can deal with nonlinear separation only using the kernel trick [st04]: the inner product among descriptors is replaced with a kernel function. This has the effect of transferring the descriptors into a higher dimensional space where linear separation is possible. Although efficient, this technique is difficult to apprehend. In particular, to each kernel corresponds a curved frontier in the original space. However, the relation between both is far from straightforward. Moreover, several kernels might produce a correct separation of the descriptors. Studying this technique could provide hints on how to choose appropriate kernels for given descriptors or could help finding an alternative to it.

- [c95] Cortes, Vapnik, Support vector networks, Machine Learning, 20:273-297, 1995.
- [o01] Oliva, Torralba, Modeling the shape of the scene: a holistic representation of the spatial envelope, International Journal of Computer Vision, 42:145-175, 2001.
- [c04] Csurka, Dance, Fan, Willamowski, Bray, Visual categorization with bags of keypoints, Workshop on Statistical Learning in Computer Vision, ECCV, 2004.
- [l04] Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60:91-110, 2004.
- [st04] Shawe-Taylor, Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, 2004.
- [p09] Piro, Anthoine, Debreuve, Barlaud, Sparse Multiscale Patches (SMP) for image categorization, International Conference on Multimedia Modeling (MMM), 2009.