

Annotation et classification d'images : présentation et exemple d'application en cancérologie

G. Gauthier¹ et D. Elie¹

¹ ADCIS S.A., 10 avenue de Garbsen, 14200 Hérouville Saint-Clair, France

Email: gervais.gauthier@adcis.net, damien.elie@adcis.net

Résumé

L'image numérique est de plus en plus présente dans de nombreux domaines comme la caractérisation, le contrôle de qualité, le diagnostic médical, la gestion de documents, etc. Le nombre d'images à exploiter croît de façon significative rendant l'analyse automatique des images nécessaire voire indispensable. Les outils de classification sont une des réponses possibles pour automatiser la prise de décision à l'issue de l'analyse des images. Un exemple d'application en cancérologie de la classification supervisée basée sur une architecture de type réseaux de neurones illustre l'intérêt de la classification appliquée aux images. L'annotation d'images est également présentée puisque c'est un outil nécessaire pour la phase d'apprentissage de toute classification supervisée.

Mots-clés : image, couleur, classification, annotation, réseaux de neurones, cancérologie

1 Introduction

Les systèmes d'acquisition numérique d'images ont en quelques années envahi les laboratoires et services de recherche et développement, les sites de production ainsi que notre quotidien. Le moyen d'observation qu'est l'image numérique est devenu l'un des plus utilisés en analyse grâce à la richesse des informations contenues et à son coût abordable suite à l'essor commercial des appareils photographiques numériques. En parallèle, l'augmentation des capacités des moyens de stockage a permis la création de base de données images. La taille croissante des informations stockées rend l'exploitation des bases de données impossible manuellement dans un délai raisonnable. Aussi, il est nécessaire d'exploiter automatiquement ces données. Nous présenterons donc des outils permettant d'atteindre cet objectif et basés sur une méthode de classification supervisée de type réseaux de neurones.

2 Généralités sur la classification appliquée à l'image numérique

2.1 Quelles données peuvent être classées dans une image ?

La classification de données peut être appliquée aux images à plusieurs étapes de leur traitement. Elle peut être appliquée pour le tri des pixels dans le but d'extraire des zones de l'image, i.e. l'étape de

segmentation, ou pour le tri des zones extraites de l'image que nous nommerons *objets* dans la suite de cet article.

Dans le premier cas, nous utiliserons les attributs associés à chaque pixel (intensité ou couleur, position, fonctions dépendant des pixels voisins) comme données permettant la classification. L'outil de classification de pixels le plus connu est le seuillage qui peut être interactif (le seuil est fixé par un opérateur humain) ou automatique (le seuil est prédéfini ou calculé à partir des attributs associés aux pixels).

Dans le second cas, nous utiliserons les attributs de chaque objet (statistiques sur les intensités ou les couleurs, position, taille, mesures de forme, mesures de texture, etc.). Dans l'exemple présenté dans la suite de cet article, la classification est appliquée aux objets (cellules) extraits des images par segmentation automatique.

2.2 Types de classification

Il existe deux types de classification : la classification non supervisée et la classification supervisée.

La classification non supervisée est une classification sans a priori, i.e. aucune classe n'est définie avant le calcul des classes. Ce type de classification a pour but de découvrir des groupes cohérents d'individus ou d'objets.

La classification supervisée, quant à elle, trie les données en catégories grâce à la modélisation de la relation entre les observations et les catégories établies à partir d'exemples connus (ensemble de

couples observations-catégories). Ce tri permettra d'associer une catégorie à chaque nouvelle observation (pixel ou objet pour une image).

3 Apprentissage

Le tri par classification supervisée n'est possible qu'après avoir modélisé la relation entre les observations (valeurs en entrée du classifieur) et les catégories (valeurs en sortie du classifieur). Il est nécessaire, pour cela, de définir les catégories à associer à chaque élément d'un ensemble d'observations que nous appellerons base d'apprentissage.

3.1 Annotation

L'association d'une catégorie ou étiquetage à chaque observation sera obtenue lors de la phase d'apprentissage par annotation de l'image par un ou plusieurs experts capables de définir la catégorie de chaque observation.

3.1.1 Outil d'annotation

Afin de simplifier la tâche d'étiquetage des observations, la société ADCIS a développé des applications d'annotation d'images adaptées à différents domaines, dont un exemple est présenté ci-dessous.

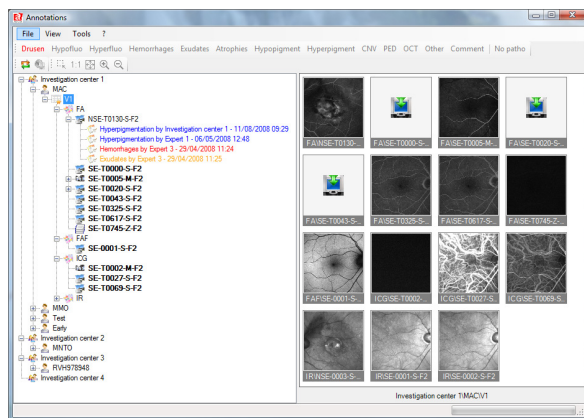


Figure 1 : interface graphique du logiciel d'annotation pour l'étiquetage d'images de la rétine.

L'application d'annotation doit gérer les données de façon homogène depuis le traitement d'image et l'extraction de zones caractéristiques de l'image (objets) jusqu'à la classification des objets. Il est souhaitable que la manipulation des données (images, objets, attributs, catégories) soit la plus transparente possible pour les experts chargés de l'annotation. Les objets sont représentés graphiquement sur les images dont ils sont issus. L'expert peut sélectionner à l'aide de la souris un ou plusieurs objets appartenant à la même catégorie et associer à ce ou ces objets leur catégorie en la cochant parmi la liste prédéfinie ou permet de définir une nouvelle catégorie. Cette

dernière est alors associée aux objets sélectionnés. Dans certains cas, l'application peut également être utilisée pour créer une base d'objets destinée à tester et valider une méthode de segmentation. Les objets ne sont alors plus générés par segmentation et importés dans l'application d'annotation, mais générés par l'expert à l'aide de la souris grâce à des outils de contourage d'objet. L'évaluation de différentes méthodes de segmentation sera alors obtenue par comparaison des objets issus de la segmentation automatique et des objets dessinés par le ou les experts. L'application présentée en Figure 1 propose des outils de contourage pouvant à la fois générer les objets et leur associer une catégorie et des commentaires.

3.1.2 Mise à disposition des données

Tout outil d'annotation doit être adapté au contexte de travail afin que la phase d'enregistrement des données d'apprentissage soit la moins fastidieuse possible, surtout si celle-ci requiert l'intervention d'experts. Il est également parfois nécessaire de sécuriser l'accès à certaines données à caractère confidentiel (par exemple les données patient dans le cas d'informations associées à un dossier médical). Enfin, si les données de la base d'apprentissage doivent être enregistrées par des experts géographiquement éloignés, il est nécessaire que l'application d'annotation permette la gestion de données présentes sur un serveur distant. L'application présentée en Figure 1 intègre les outils pour interroger une base de données distante et annoter les images présentes dans cette base de données.

3.2 Calcul des paramètres du classifieur

La modélisation de la relation entre l'observation et la catégorie est obtenue par un classifieur dont les paramètres sont réglés grâce à l'ensemble d'observations étiquetées manuellement par le ou les experts (base d'apprentissage). La base d'apprentissage est divisée en deux groupes d'observations étiquetées : l'un sert au calcul des paramètres du classifieur, l'autre est utilisé pour évaluer les performances du classifieur paramétré.

4 Choix du classifieur

Le choix du classifieur dépend de plusieurs critères : type de données d'observation, type de classification attendue, taille de l'échantillon à classer, performances (efficacité et vitesse de classification).

Les méthodes de classification les plus répandues sont : analyse en composantes principales (ACP), analyse factorielle des correspondances (AFC), classification basée sur le théorème de Bayes, machines à vecteurs de support (SVM), K plus

proches voisins (KNN), classification vectorielle K-means, réseaux de neurones.

L'exemple suivant présente un classifieur basé sur des réseaux de neurones et les résultats obtenus avec celui-ci.

5 Exemple d'application en cancérologie : CLOCYT

5.1 Présentation de CLOCYT

ADCIS a développé pour une application de tri cellulaire un ensemble de composants orientés microscopie référencés sous le nom de CLOCYT, évolution de l'application ARCTIC^[1], pour le dépistage du cancer.

Cette application gère l'ensemble de la chaîne d'analyse d'image depuis l'acquisition d'images jusqu'aux résultats d'analyse. La préparation des échantillons constitue également une étape importante pour l'analyse. En effet, il est nécessaire que les échantillons traités par CLOCYT soient comparables. Aussi, pour obtenir une coloration reproductible et homogène, les échantillons biologiques sont préparés par un automate appliquant une coloration de Papanicolaou.

Le microscope et l'acquisition d'images sont automatisés pour capturer les images de l'ensemble des lames biologiques fixées sur la platine (jusqu'à 8 lames par session).

5.2 Segmentation des cellules et mesures

Une correction colorimétrique est appliquée aux

images capturées pour que les images aient des couleurs comparables quels que soient la caméra utilisée et son état électronique (voir Figure 2)

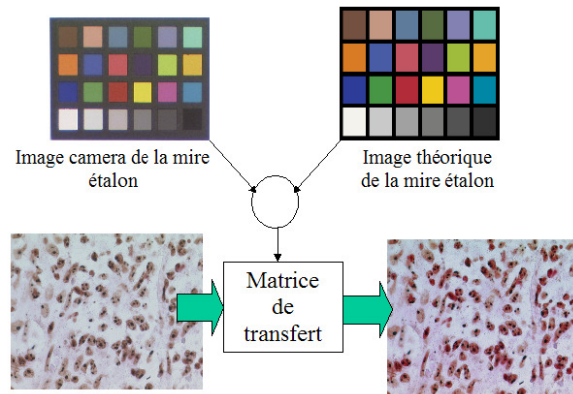


Figure 2 : calibration couleur des images.

Les noyaux et les cytoplasmes sont extraits des images corrigées par des opérateurs de morphologie mathématique et par des opérations de traitement d'images parmi lesquelles : le seuillage par la méthode d'Otsu et la ligne de partage des eaux à partir des marqueurs des noyaux. Le schéma global de la segmentation des images est présenté en Figure 3.

Des mesures colorimétriques et des mesures de forme sont calculées sur le noyau et le cytoplasme de toutes les cellules. Certaines de ces mesures sont des rapports entre des mesures du cytoplasme et celles du noyau.

5.3 Annotation et apprentissage

Une fois que les cellules (noyaux et cytoplasmes) ont été segmentées, les cellules doivent être classées pour qu'un diagnostic puisse être donné à partir des images

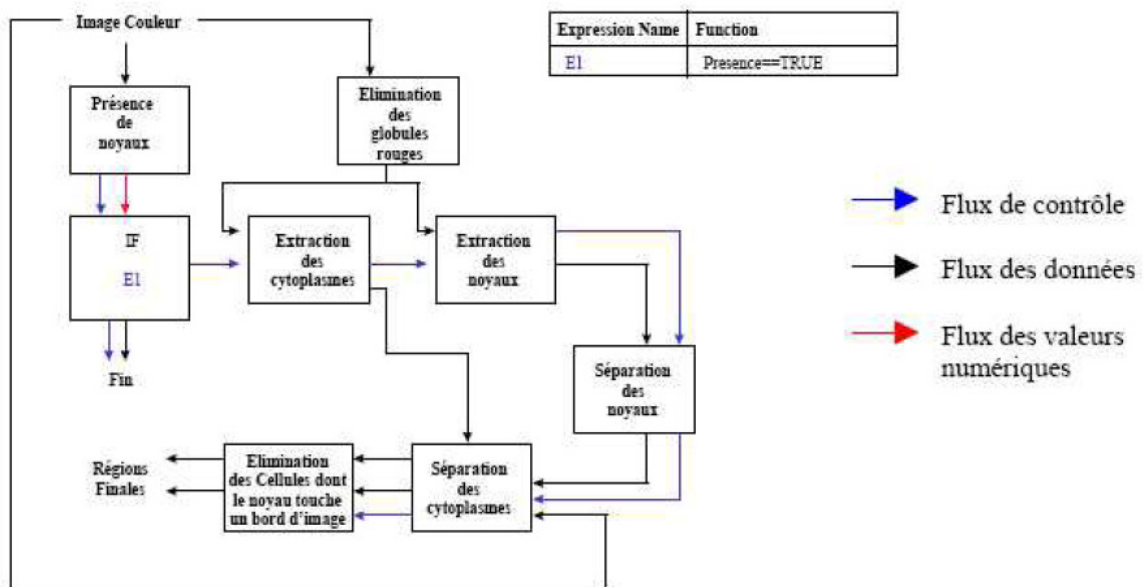


Figure 3 : schéma du traitement appliqué pour la segmentation

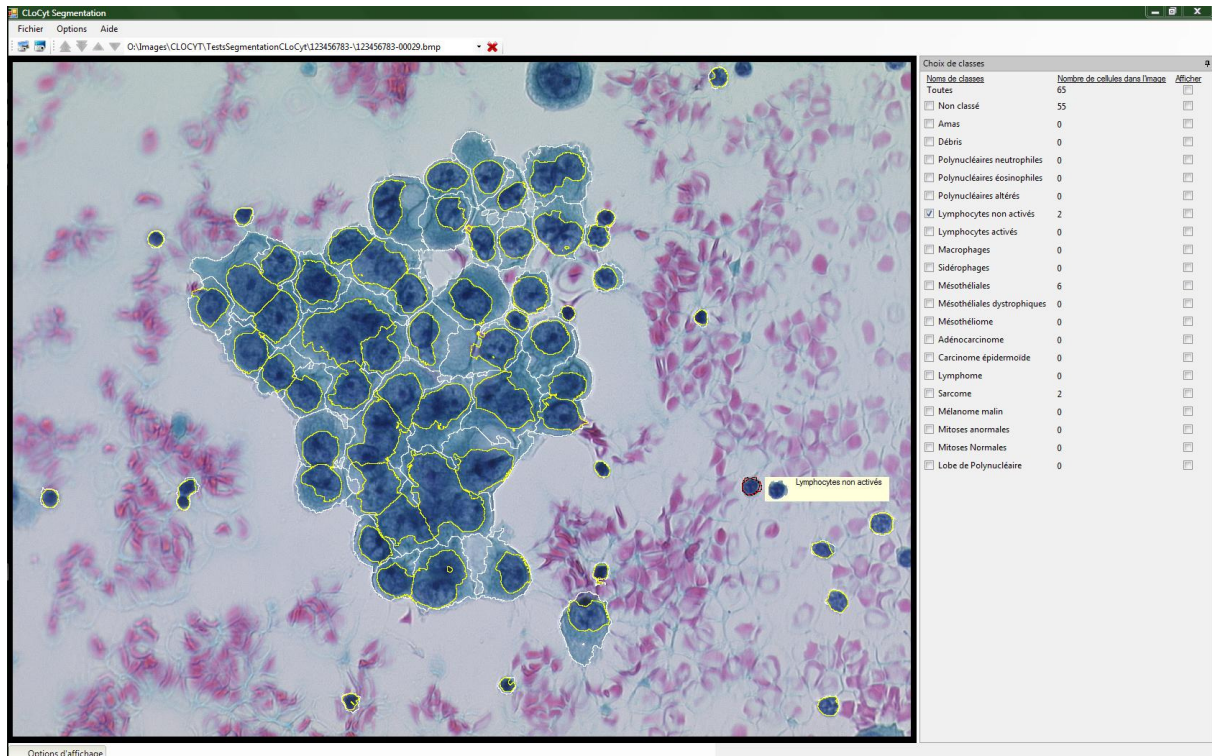


Figure 4 : interface d'étiquetage manuel des cellules

de l'échantillon analysé. Le paramétrage de l'outil de classification est obtenu par une phase d'apprentissage. Deux lots de cellules sont définis pour cette étape. Le premier lot est utilisé pour le calcul des paramètres du classifieur alors que le second permettra d'évaluer la qualité du tri effectué par le classifieur ainsi paramétré.

CLOCYT propose une interface d'annotation des cellules pour permettre aux anatomopathologistes d'associer une catégorie à chaque cellule de la base d'apprentissage en plaçant manuellement les cellules dans différentes catégories (Figure 4). Le premier lot de cellules de la base d'apprentissage est ensuite exploité par le classifieur pour le calcul de ses paramètres. Le second lot de cellules est trié par le classifieur. L'étiquetage des cellules par ce tri est revu par les anatomopathologistes pour la validation de la classification.

5.4 Méthode de classification de CLOCYT

L'outil de classification MONNA (Multiple Ordinate Neural Network Architecture)^[2] (Figure 5) développé par le Groupe de Recherche En Informatique Image Instrumentation de Caen (GREYC) a été intégré à l'environnement Aphelion, logiciel de traitement et d'analyse d'images développé par la société ADCIS. Le composant logiciel issu de cette intégration permet de créer, configurer et utiliser le classifieur MONNA. Le tri des cellules dans l'application CLOCYT est effectué par ce composant. MONNA est un classifieur permettant de combiner les résultats de

classifieurs binaires de type réseau de neurones simple couche. Cette architecture simplifie le problème de chaque réseau de neurone et améliore la performance de tri de chaque sous-réseau.

Une généralisation de ce classifieur à tout type de classifieur à réponse binaire est également disponible : MOCA (Multiple Ordinate Classifier Architecture).

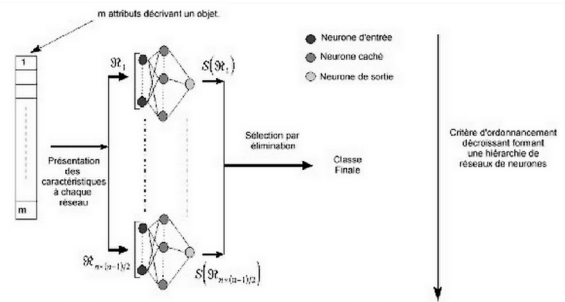


Figure 5 : architecture du classifieur MONNA

La création d'un classifieur MONNA recherche l'architecture la mieux adaptée (nombre de neurones) pour chaque sous-réseau, évalue l'ensemble de mesures le plus pertinent pour le tri par l'algorithme SFFS (Sequential Forward Floating Selection) et calcule les poids de chaque réseau de neurones. La réponse du classifieur MONNA pour chaque objet (cellule) est la combinaison des résultats des classifieurs binaires (sous-réseaux) selon l'une des trois méthodes suivantes : vote majoritaire, élimination, méthode de Price.

5.5 Résultats

Un ensemble de 926 cellules a été trié à l'aide de MONNA. Lors de l'apprentissage, les cellules ont été classées en douze catégories : polynucléaires neutrophiles, polynucléaires éosinophiles, polynucléaires altérés, lymphocytes non activés, lymphocytes activés, macrophages, mésothéliales, mésothéliales dystrophiques, mésothéliomes, adénocarcinomes, mitoses anormales, lobes de polynucléaire. D'autres catégories de cellules existent mais le nombre d'échantillons présents étaient insuffisants pour permettre l'évaluation et le test de la classification de celles-ci. A terme, le nombre de cellules du lot d'apprentissage sera augmenté afin d'avoir une représentativité plus importante de toutes les catégories de cellules et la phase d'apprentissage sera relancée pour affiner les paramètres du classifieur et améliorer ainsi les performances de tri.

Le premier tri évalué a été obtenu en classant les cellules par la méthode de combinaison des résultats des classifieurs binaires du vote majoritaire (*Tableau 1*). L'analyse des résultats a montré que les erreurs de classement sont dues à des similitudes entre cellules (la majorité des résultats de classification erronés se porte sur une ou deux classes de cellules dont l'aspect est très proche).

Tableau 1 : classement en 12 catégories

Type de cellules	Nombre total	Nombre de cellules mal classées	Nombre de cellules bien classées	% cellules mal classées	% Cellules bien classées
Polynucléaires neutrophiles	86	55	31	63.95%	36.05%
Polynucléaires éosinophiles	69	19	50	27.54%	72.46%
Polynucléaires altérés	52	24	28	46.15%	53.85%
Lymphocytes non activés	175	42	133	24.00%	76.00%
Lymphocytes activés	86	22	64	25.58%	74.42%
Macrophages	154	70	84	45.45%	54.55%
Mésothéliales	141	28	113	19.86%	80.14%
Mésothéliales dystrophiques	10	7	3	70.00%	30.00%
Mésothéliome	7	3	4	42.86%	57.14%
Adénocarcinome	96	15	81	15.63%	84.38%
Mitoses anormales	8	4	4	50.00%	50.00%
Lobe de Polynucléaire	42	19	23	45.24%	54.76%
Total	926	308	618	33.26%	66.74%

L'affectation de certaines cellules à une catégorie par un expert est parfois difficile à cause de la similitude d'aspect. L'apprentissage a été effectué en permettant à l'expert d'annoter les cellules dans plusieurs

catégories. Les paramètres du classifieur ont été calculés en minimisant l'erreur par rapport à la classe la plus proche parmi les classes associées à chaque cellule. Le *Tableau 1* présente le résultat de la classification des données de test en considérant uniquement les cellules associées à une seule catégorie. Les tests en considérant les cellules associées indifféremment à une ou plusieurs catégories donnent 29,10% d'erreur contre 33,26% en considérant uniquement les cellules associées à une seule catégorie car les cellules à catégorie multiple sont mieux classées par les experts.

Certaines catégories de cellules peuvent être regroupées car le diagnostic final de l'anatomopathologiste ne changera pas si une cellule est associée à l'une ou l'autre des catégories du regroupement. Les catégories inventoriées ont donc été regroupées en quatre catégories : polynucléaires (polynucléaires neutrophiles, polynucléaires éosinophiles, polynucléaires altérés et lobes de polynucléaires), lymphocytes (lymphocytes activés et non activés), macrophages mésothéliales (macrophages et mésothéliales) et anormales (mésothéliales dystrophiques, mésothéliomes, adénocarcinomes et mitoses anormales). L'évaluation du tri obtenu en classant les cellules dans l'une des quatre catégories est présentée dans le *Tableau 2*. Les résultats montrent que les erreurs observées dans le tri en douze catégories étaient principalement dues à un classement des cellules dans une catégorie d'un même groupe.

Tableau 2 : regroupement en 4 catégories

Type de cellules	Nombre total	Nombre de cellules mal classées	Nombre de cellules bien classées	% cellules mal classées	% Cellules bien classées
Polynucléaires	249	39	210	15.66%	84.34%
Lymphocytes	261	38	223	14.56%	85.44%
Macrophages Mésothéliales	295	57	238	19.32%	80.68%
Anormales	121	7	114	5.79%	94.21%
Total	926	141	785	15.23%	84.77%

Le tableau présente le résultat de la classification avec les données de test en considérant uniquement les cellules associées à une seule catégorie. Les tests en considérant les cellules associées indifféremment à une ou plusieurs catégories donnent 13,35% d'erreur car les cellules à catégorie multiple sont mieux classées par les experts.

Un nouveau regroupement de catégories a été opéré pour valider la classification. Les cellules ont été considérées comme bien classées cette fois-ci si leur catégorie était dans le même groupe que la catégorie de la réponse en utilisant deux catégories : cellules normales (polynucléaires, lymphocytes, macrophages,

mésotéliales et mitoses normales) et anormales. Les erreurs de tri commises par MONNA avec le découpage en douze catégories associent les cellules à une autre catégorie appartenant au même regroupement de catégories ; l'application permet donc de bien distinguer les cellules anormales des cellules normales. Le résultat du regroupement est présenté dans le *Tableau 3*.

Tableau 3 : regroupement en 2 catégories

Type de cellules	Nombre total	Nombre de cellules mal classées	Nombre de cellules bien classées	% cellules mal classées	% Cellules bien classées
Normales	805	15	790	1.86%	98.14%
Anormales	121	7	114	5.79%	94.21%
Total	926	22	904	2.38%	97.62%

Le tableau montre le résultat de la classification avec les données de test en considérant uniquement les cellules associées à une seule catégorie. Les tests en considérant les cellules associées indifféremment à une ou plusieurs catégories donnent 2.40% d'erreur car les cellules à catégorie multiple sont moins bien classées dans ces deux groupes. Toutefois, l'erreur dans les deux cas est faible.

6 Conclusion

La mise en œuvre d'outils d'annotation et de classification devient nécessaire pour la prise de décision par l'analyse automatique d'images. Les outils que nous avons intégrés dans l'application CLOCYT permettent d'obtenir un diagnostic pour le dépistage du cancer. Ces mêmes outils ont été également appliqués à la reconnaissance de caractères et à l'indexation d'images et ont donné des résultats probants.

7 Remerciements

Le développement des outils de classification MONNA et réseaux de neurones dans l'environnement Aphelion[®] a été rendu possible grâce à la collaboration avec le Groupe de Recherche En Informatique Image Instrumentation de Caen (GREYC).

L'application CLOCYT a été développée dans le cadre de la collaboration avec le GREYC et le Centre Hospitalier Public du Cotentin (qui a fourni les images, validé les techniques de segmentation et classification, et qui testera dans les prochains mois l'application complète).

Références

- [1] O. Lezoray, A. Elmoataz, H. Cardot, M. Revenu, *A.R.C.T.I.C : Un système de tri cellulaire par analyse d'images*, Proceedings of Vision Interface 99, Trois-Rivières, pp. 312-319, 18-21 Mai, Canada
- [2] O. Lezoray, H. Cardot, D. Fournier, M. Revenu, *MONNA : a Multiple Ordinate Neural Network Architecture*, Proceedings of EIS'2000, pp. 47-53, Paisley, Ecosse, Juin 2000