

Fusion de segmentations d'images microscopiques par SVM dans différents espaces couleur selon la théorie de l'évidence

Christophe CHARRIER, Gilles LEBRUN, Olivier LÉZORAY

Université de Caen-Basse Normandie, LUSAC EA-2607, Groupe Vision et Analyse d'Images
120, route de l'exode, 50000 Saint-Lô, France

{c.charrier,g.lebrun,o.lezoray}@chbg.unicaen.fr

Résumé – Dans cet article, un outil d'aide au diagnostic par segmentation automatique d'images microscopiques cellulaires est proposé. Cet outil repose sur la mise en place d'une méthode de segmentation automatique combinant (selon la théorie de l'évidence) les résultats issus des différentes segmentations par Machines à Support de Vecteurs—*Support Vector Machines (SVM)*—en intégrant des incertitudes et des redondances liés aux espaces couleur. Ces incertitudes sont obtenues par calcul d'une probabilité *a posteriori* qui est fonction des résultats générés par les SVMs. Une amélioration sensible de la qualité de la segmentation est finalement réalisée par la prise en compte de l'incohérence de plusieurs classifications des pixels.

Abstract – In this paper, a tool for diagnosis assistance by automatic segmentation of microscopic cellular images is introduced. This method is based on an automatic segmentation technique combining (with the Dempster-Shafer rule) the results obtained by Support Vector Machines (SVM) applied within different color spaces. This combination is performed by integrating uncertainties and redundancies for each one of the color spaces. These uncertainties are computed as *a posteriori* probabilities according to the SVM obtained results. An improvement of the final segmentation quality is performed by taking into account the inconsistencies of several pixel classification.

1 Introduction

L'analyse d'images microscopiques dans les laboratoires d'anatomie et de cytologie pathologiques tient une place très importante pour la détection des cellules anormales afin d'établir un diagnostic. Les cellules sont visuellement évaluées par un technicien lors d'une phase de screening. Cette analyse est d'un intérêt capital pour le pathologiste car son diagnostic dépend de la bonne reconnaissance des cellules anormales ou suspectes. Or le screening visuel est difficile et le faible nombre de cellules anormales comparativement aux nombre de cellules normales implique une grande concentration du technicien. A cause de ce facteur subjectif des faux-négatifs peuvent apparaître. Pour pallier ceci, une approche consiste à aider le technicien par un système de screening semi-automatique afin de confirmer l'analyse visuelle. Dans cette optique, nous désirons concevoir une méthode de segmentation des cellules d'images de microscopie, ces cellules étant constituées d'un cytoplasme et d'un noyau. Cependant leur configuration spatiale et colorimétrique est extrêmement variable, ce qui rend difficile leur extraction précise et une méthode de segmentation peut bien fonctionner sur certaines images et moins bien sur d'autres.

En général, un seul schéma de segmentation est appliqué afin d'obtenir des segmentations les plus fidèles possibles par rapport aux vérités terrains disponibles. Or améliorer les taux de classification, et donc la qualité de la segmentation induit inéluctablement la définition et le développement de protocoles de classification de plus en plus sophistiqués, et par la même de plus en plus sensibles à la moindre variation de la structure de l'image.

Une autre voie peut cependant être envisagée : celle de la combinaison des résultats obtenus par différents schémas. Ainsi

nous proposons d'utiliser différents schémas de segmentation par classification de pixels dans différents espaces couleur afin d'obtenir plusieurs segmentations possibles d'une même image en trois classes : fond, cytoplasme, noyau. La classification est ainsi opérée séparément dans cinq espaces couleur à savoir RVB , $L^*a^*b^*$, YUV et deux systèmes de représentation colorimétrique antagonistes YC_bC_r et HSL . Ces différentes segmentations sont ensuite fusionnées afin d'obtenir une unique carte de segmentation.

2 Création et fusion des cartes de segmentation

La figure 1 illustre le synopsis du schéma de segmentation proposé. Une première étape de segmentation par classification des pixels est réalisée dans les cinq espaces couleur retenus. A l'issue de cette phase, et pour chacune des segmentations, une probabilité *a posteriori* p_i est calculée, ce qui permet d'obtenir un indice de confiance de la classification pour chaque pixel de l'image par rapport aux classes finales. Puis, dès lors qu'il y a divergence de classification pour un pixel, une étape de fusion est opérée en prenant en compte la probabilité *a posteriori*. Une carte de segmentation finale est ainsi générée.

2.1 Segmentation par classification des pixels

Parmi toutes les techniques de segmentation qui sont offertes, les segmentations par les SVM ont été retenues car une précédente étude a permis de mettre en exergue des niveaux de classification très satisfaisants [1].

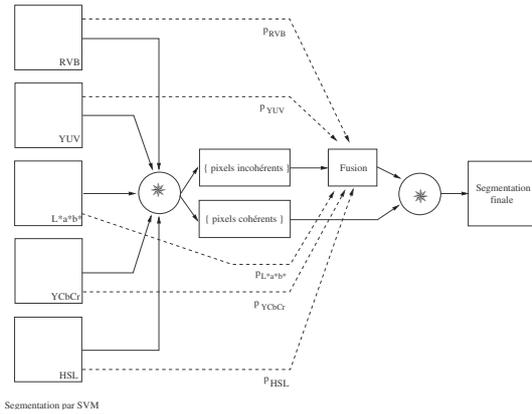


FIG. 1: Synopsis du schéma proposé où p_i représente la probabilité a posteriori calculée dans l'espace couleur i .

2.1.1 Principe des SVM

La théorie de l'Apprentissage Statistique de Vapnik et de Chervonenkis [2] a conduit au développement d'une classe d'algorithme connu sous le nom de SVM. Ils permettent de réaliser des estimations en classification. Une des originalités de la méthode est de produire une fonction de décision qui n'utilise qu'un sous-ensemble de la base d'apprentissage. Les éléments de ce sous-ensemble sont nommés Vecteurs de Support (SV).

Soit un ensemble d'apprentissage $A = \{(x_1, y_1), \dots, (x_k, y_k)\}$ composé de k couples avec $x_i \in \mathcal{R}^n$ et $y_i \in \{-1, +1\}$. L'algorithme des SVMs projette les vecteurs x_i dans un espace de travail \mathbf{H} à partir d'une fonction non linéaire $\phi : \mathcal{R}^n \rightarrow \mathbf{H}$. L'hyperplan optimal de séparation des deux classes dans l'espace \mathbf{H} est ensuite recherché. Cet hyperplan (\mathbf{w}, b) matérialise la frontière de séparation entre les deux classes. La classe y d'un nouvel exemple x est définie par $y = \text{sign}(\mathbf{w} \cdot \phi(x) + b)$. L'hyperplan est optimal s'il maximise la distance qui le sépare des exemples dont il est le plus proche. Cette distance est usuellement appelée marge du classifieur. Il a été démontré [2] que maximiser cette marge correspond à maximiser le "pouvoir" généralisateur du classifieur. En choisissant une fonction noyau $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$, il a également été montré que la résolution du problème dual correspondant produit une fonction de décision de la forme $f(x) = \sum_{x_i \in \text{SV}} \alpha_i y_i K(x_i, x) + b$ où $\{\alpha_i\}$ et b correspondent à la solution optimale du problème dual.

2.1.2 Calcul et utilisation des probabilités a posteriori

Les SVMs étant des classifieurs binaires, la résolution d'un problème multi-classes est effectuée en le transformant en une combinaison de problèmes binaires [3]. Ainsi, dans notre cadre applicatif, trois fonctions de décision seront utilisées pour réaliser la classification en trois classes {noyau, fond, cytoplasme} correspondant à leurs discriminations en un contre tous. Cependant les SVMs ne permettent pas d'aboutir directement à des probabilités a posteriori de classification. Au lieu d'estimer les densités conditionnelles des classes $p(f|y)$, un modèle paramétrique est utilisé afin d'obtenir les probabilités a posteriori $p(y = 1|f)$ directement où f représente la valeur obtenue à la sortie des SVMs. PLATT [4] propose une méthode de calcul de ces probabilités à partir des paramètres obtenus par les SVMs. Ainsi les probabilités a posteriori sont déterminées à

l'aide d'une fonction paramétrique sigmoïde par :

$$p(y = 1|f) = \frac{1}{1 + \exp(Bf + C)}, \quad (1)$$

dans laquelle les paramètres B et C sont estimés en utilisant la procédure du maximum de vraisemblance.

Un ensemble d'images segmentées constituant une vérité terrain des classes de pixels sert de base d'apprentissage. Chacune des trois fonctions de décision réalise un apprentissage par espace couleur, ce qui permet d'obtenir cinq cartes de segmentation distinctes avec pour chaque pixel de chaque carte une probabilité d'appartenance a posteriori p_i pour chaque classe. Cette probabilité correspond à un indice de confiance pour chaque classe et pour chaque espace couleur. L'objectif est ensuite de fusionner ces résultats afin d'aboutir à une classification finale, et donc une segmentation en trois classes. Parmi tous les outils qui s'offrent à nous dans ce domaine, nous décidons d'utiliser la théorie de l'évidence (aussi appelée théorie de Dempster-Shafer ou théorie des fonctions de croyance) [5], qui permet d'une part de traiter des informations incertaines, et d'autre part de combiner des informations provenant de plusieurs sources. Dans le cadre de cette théorie, plusieurs règles de décision sont définies pour nous permettre de sélectionner la classe à laquelle le pixel sera finalement affecté.

Afin de procéder à la fusion, une nouvelle carte de segmentation est générée par intersection des cartes obtenues dans les cinq espaces couleur. Seuls les pixels présentant des incohérences de classification seront traités par la théorie de l'évidence. Un "pixel incohérent" correspond à un pixel classé dans plus d'une classe lorsque l'union de toutes les cartes est réalisée.

2.2 Traitement des pixels incohérents

2.2.1 Elements de la théorie de l'évidence

Soit $\Omega = \{\omega_1, \dots, \omega_N\}$, l'ensemble des N classes finales possibles pour un pixel incohérent, appelé cadre de discernement. Dans notre cadre applicatif $N = 3$ et Ω correspond aux trois classes $\{\omega_F, \omega_C, \omega_N\}$, respectivement pour le fond de l'image, le cytoplasme et le noyau. Au lieu de restreindre ses mesures à cet ensemble (comme le ferait la théorie des probabilités contrainte par son axiome d'additivité), la théorie de l'évidence s'étend sur l'ensemble puissance de Ω , noté 2^Ω , l'ensemble des 2^N sous-ensembles de Ω . On définit alors une fonction de masse m qui représente la croyance que l'on accorde aux différents états du système, à un instant donné. Cette fonction est également connue en tant que masse de croyance élémentaire. Elle est définie de 2^Ω dans $[0, 1]$ et vérifie

$$\sum_{A \subseteq \Omega} m(A) = 1 \text{ et } m(\emptyset) = 0 \quad (2)$$

où $m(A)$ représente la croyance que l'on place dans la proposition A . Il est à noter que A représente soit un singleton ω_n , soit une disjonction d'hypothèses. Ici, $m(A)$ quantifie la croyance que l'on accorde au fait que la classe cherchée appartienne au sous-ensemble A de Ω (et à aucun autre sous-ensemble de A). Les propositions A telles que $m(A) > 0$ sont appelées éléments focaux. De plus, dans le cas où l'ensemble des hypothèses est exhaustif et exclusif, la masse de l'ensemble vide \emptyset est nulle.

Deux fonctions de masse initiales m_1 et m_2 représentant les informations respectives de deux sources différentes, peuvent être combinées selon la règle de Dempster [6] :

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K}, \quad \forall A \in \Omega \quad (3)$$

$$A \neq \emptyset.$$

K est appelé le facteur de conflit et représente le désaccord entre les deux sources. Il correspond à la masse de l'ensemble vide $K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$. Notons que la combinaison de Dempster, appelée aussi somme orthogonale et notée $m = m_1 \oplus m_2$ est commutative et associative.

Après combinaison, il reste à prendre une décision quant à l'élément le plus "probable" de Ω . Plusieurs règles de décision sont alors possibles, mais une des plus utilisées reste celle du maximum de probabilité "pignistique". Cette règle de décision présentée par SMETS [7] utilise la transformation pignistique, qui permet de transformer une masse m_i en une distribution de probabilité permettant la prise de décision telle que :

$$\omega^* = \text{Arg} \left\{ \max_{\omega \in \Omega} \left[\sum_{A \subseteq \Omega} \frac{m(A)}{|A|} \right] \right\}. \quad (4)$$

$|A|$ étant le cardinal de A .

2.2.2 Modélisation des fonctions de masse

L'une des difficultés réside dans le choix de la modélisation des fonctions de masse. Parmi les modélisations existantes, celle proposée par DENÈUX [8] a été utilisée car elle est basée sur la notion de distance aux voisins et permet d'intégrer divers critères de voisinage. Ainsi la masse $m(\{\omega_n\})$ est définie comme une fonction décroissante de la distance d entre le pixel à classifier et le barycentre de la classe :

$$\begin{cases} m(\omega_n) = \alpha \exp(-\gamma_n d^2) \\ m(\Omega) = 1 - m(\omega_n) \end{cases} \quad (5)$$

où $0 < \alpha < 1$ est une constante. Dans cette étude, α est déterminé à partir des probabilités *a posteriori* obtenues par les SVM pour la classe ω_n dans les différents espaces couleur : $\alpha = p_i$. $\gamma_n > 0$ dépend de la classe ω_n et est calculé par minimisation d'un critère d'erreur selon l'algorithme EM utilisé dans [9].

Les 5 fonctions de masses initiales (m_i) $i \in [1, \dots, 5]$ sont donc créées après calcul des k régions candidates à la fusion et avant décision du vote majoritaire. Ainsi pour une même carte de segmentation, un pixel qui a été initialement affecté à une seule des trois classes $\{\omega_F, \omega_C, \omega_N\}$ peut, en relâchant la contrainte d'affectation, être affecté à une classe parmi au plus les trois classes. La contrainte correspond simplement à la création d'un intervalle calculé à partir de la valeur maximale de la probabilité *a posteriori* p_i tel que $[p_i, p_i - \delta_i]$, avec δ_i fonction de l'espace couleur associé. Ainsi toutes les classes dont les probabilités *a posteriori* sont incluses dans cet intervalle sont considérées comme des classes candidates à l'affectation du pixel.

Les fonctions de masses permettent de prendre en compte une incertitude associée à chacune des cartes de segmentation. Ainsi, on regroupe dans le même élément focal les classes qui se trouvent très proches les unes des autres, et la prise de décision ne s'effectue qu'après avoir combiné les résultats des autres projections.

Dans cette étude deux formulations de la distance d ont été étudiées. Il s'agit 1) de la distance euclidienne, 2) de la distance MAHALANOBIS permettant de comparer la dispersion spatio-colorimétrique de deux nuages couleurs, définie par

$$d = \sqrt{(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)} \quad (6)$$

où μ_i représente le vecteur moyenne et Σ_i^{-1} la matrice de covariance inverse associée à l'hypothèse ω_n .

Lors de l'utilisation de la distance euclidienne seule, la formulation de la distance euclidienne pondérée préconisée en 1994 par la Commission Internationale de l'Éclairage (CIE) lors de l'utilisation de l'espace couleur $L^*a^*b^*$ a été utilisée. Même si l'expression du calcul de la distance entre couleurs doit s'exprimer dans un espace couleur autre que l'espace RVB , ces formules ne sont valables que pour des petites distances et leurs utilisations en dehors de ce contexte n'est pas validée par la CIE.

3 Résultats expérimentaux

3.1 Évaluation de la qualité de la segmentation

L'évaluation de la qualité d'un traitement est un domaine de recherche mobilisant de nombreuses équipes. Que les champs d'investigations concernent la compression ou la segmentation des images couleur, deux classes se distinguent en fonction de la présence d'une image de référence ou non. Dans le cas de la segmentation, le terme "vérité terrain" est utilisé en lieu et place d'image de référence.

Lorsque l'on dispose d'une vérité terrain, les mesures de qualité sont toutes basées sur la prise en compte de la taille des régions et du taux de recouvrement entre les régions R_i de l'image segmentée I et les régions V_j de l'image de référence V . Parmi toutes les métriques disponibles, la mesure définie par MARTIN a été retenue puisqu'elle permet de mesurer la dissimilarité entre la segmentation résultat et la segmentation de référence.

Cette mesure est basée sur l'erreur $E(s)$ calculée en chaque pixel s par :

$$E(s) = \frac{\text{card}(V_j \setminus R_i)}{\text{card}(V_j)}. \quad (7)$$

ou

$$E'(s) = \frac{\text{card}(R_i \setminus V_j)}{\text{card}(R_i)}. \quad (8)$$

La dissimilarité est ensuite fournie par l'erreur locale de cohérence fournissant une mesure de qualité :

$$\text{ELC}(I, V) = \frac{1}{h \times l} \sum_s \min\{E(s), E'(s)\} \quad (9)$$

où h et l représentent respectivement la hauteur et la largeur en pixels de l'image V .

Il existe une erreur globale de cohérence qui n'a pas été retenue ici car cette dernière a le défaut de favoriser une sur-segmentation ou une sous-segmentation [10].

3.2 Résultats

La technique proposée est appliquée sur une base d'images médicales composée d'images microscopiques cellulaires. La

figure 2 présente les pixels incohérents obtenus après intersection des cartes de segmentation dans les cinq espaces couleur (figure 2(b)). Les désaccords interviennent principalement au niveau du cytoplasme et des noyaux lorsque l'on compare les résultats avec l'image initiale (figure 2(a)).

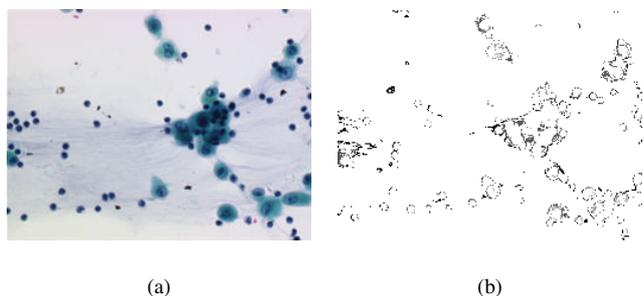


FIG. 2: *Pixels incohérents obtenus par intersections des diverses cartes de segmentation en fonction d'une image originale.*

Le tableau 1 présente les moyennes en pourcentages des pixels incorrectement et correctement classifiés ainsi que la moyenne en pourcentage des pixels présentant une incohérence de classification. On note qu'il subsiste un pourcentage de pixels (9.4%) pour lesquels il n'est plus possible de revenir sur la classification puisque ces derniers ont été associés à une classe erronée de manière unanime (selon la comparaison avec les images de vérité terrain). Ainsi, le taux de classification maximal que l'on puisse atteindre est fonction de ces affectations. Seuls les pixels pour lesquels subsistent un désaccord entre les classificateurs pourront subir une révision de classification, soit de 3.47% au maximum. Ainsi pour améliorer le schéma de classification, il est également nécessaire d'améliorer les pré-traitements, à savoir, les résultats obtenus par les SVMs.

Moyenne des classification par SVMs dans les différents espaces couleur		
correcte	incorrecte	incohérente
87.13 ± 6.12	9.4 ± 5.32	3.47 ± 1.13

TAB. 1: *Pourcentage des pixels classifiés correctement, incorrectement selon les vérités terrain.*

Le tableau 2 présente la moyenne des taux de classification correctes obtenus sur la base pour les SVMs et la méthode proposée ainsi que la mesure de qualité de la segmentation finale.

Méthode	moyenne TR	Moyenne qualité
SVM	87.13 ± 6.12	0.45 ± 0.04
SVM combinés (d_1)	87.67 ± 5.62	0.39 ± 0.03
SVM combinés (d_2)	88.12 ± 5.43	0.38 ± 0.04

TAB. 2: *Moyennes du taux de reconnaissance et de qualité selon le critère de Martin pour la segmentation par les SVMs, la combinaison des SVMs intégrant la distance euclidienne (d_1) et la distance de Mahalanobis (d_2).*

Les résultats obtenus démontrent que la combinaison des probabilités des pixels incohérents permet d'améliorer sensiblement la qualité de la segmentation basée sur un processus de classification des pixels. On note effectivement une augmentation du pourcentage de reconnaissance pour chacune des trois

classes. L'incohérence se situe principalement sur les pixels appartenant au cytoplasme et aux noyaux. Le gain de qualité se portant essentiellement sur la classification de ces pixels, la qualité de la segmentation d'un point de vue cytologique est grandement améliorée.

4 Conclusion

Dans cet article, un outil d'aide au diagnostic d'images microscopiques cellulaires est présenté. L'algorithme sous-jacent repose sur une combinaison de segmentations réalisées dans divers espaces colorimétriques. A partir des cartes de segmentation ainsi obtenues, une probabilité *a posteriori* est calculée et permet ainsi de quantifier le degré de confiance que l'on accorde à la segmentation. Une combinaison, selon la théorie de Dempster-Shafer, est ensuite réalisée de manière à prendre en compte les incertitudes liées à chacune des classifications dans les divers espaces couleur. Les résultats permettent de montrer une amélioration des taux de classification.

Références

- [1] G. Lebrun, C. Charrier, O. Lezoray, C. Meurie, and H. Cardot, "Fast pixel classification by SVM using vector quantization, tabu search and hybrid color space," in *the 11th International Conference on CAIP*, (Rocquencourt, France), 2005. to appear.
- [2] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
- [3] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 3, pp. 415–425, 2002.
- [4] J. Platt, "Probabilities for SV machines," in *Advances in Large-Margin Classifiers* (D. S. Peter J. Bartlett, Bernhard Schölkopf and A. J. Smola, eds.), Neural Information Processing Systems, MIT Press, 2000.
- [5] G. Shafer, *A mathematical theory of evidence*. Princeton University Press, 1976.
- [6] A. Dempster, "Upper and Lower Probabilities Induced by Multivalued Mapping," *Ann. Math. Statist.*, vol. 38, pp. 325–339, 1967.
- [7] P. Smets, "Constructing the pignistic probability function in a context of uncertainty," *Uncertainty in Artificial Intelligence*, vol. 5, pp. 29–39, 1990. Elsevier Science Publishers.
- [8] T. Denoeux, "A k-nearest neighbor classification rule based on dempster-shafer theory," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 5, pp. 804–813, 1995.
- [9] L. M. Zouhal and T. Denoeux, "An evidence-theoretic k-NN rule with parameter optimization," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 28, pp. 263–271, 1998.
- [10] S. Philipp-Foliguet, "Evaluation de méthodes de segmentation d'images couleur," in *Actes de l'école de printemps du GdR-ISIS, Opération Imagerie Couleur*, (Pau, France), 2001.