

Détection automatique de glomérules en pathologie digitale

Raphaël MARÉE^{1,2}, Jean-Christophe OLIVO-MARIN¹, Vannary MEAS-YEDID¹

¹Institut Pasteur, Unité d'Analyse d'Images Biologiques
25 rue du docteur Roux, 75015, Paris, France

²Université de Liège, Systèmes et Modélisation,
Institut Montefiore & GIGA-Research, Grande Traverse 10, 4000 Liège Sart-Tilman, Belgique
raphael.maree@pasteur.fr, jcolivo@pasteur.fr, vmeasyed@pasteur.fr,

Résumé – Dans cet article, nous proposons une méthodologie de détection de glomérules au sein d'images de biopsies rénales. Elle combine des techniques de traitement d'images et d'apprentissage supervisé. Nous évaluons l'approche sur des images présentant plusieurs sources de variations et montrons que les comptages automatiques sont très bien corrélés avec les comptages des experts.

Abstract – In this paper, we propose a methodology to automatically detect glomeruli in renal biopsy images. It combines image processing and supervised learning algorithms. We evaluate the approach on images with various sources of variability and show that automatic counting is highly correlated with human expert counting.

1 Introduction

1.1 Contexte et motivations

La biopsie rénale est un outil de diagnostic clé dans l'analyse de pathologies rénales aussi bien pour les reins natifs que pour les reins transplantés. En effet, l'examen de la biopsie permet de détecter d'éventuelles lésions avant l'apparition de symptômes cliniques. Ce diagnostic est actuellement réalisé par plusieurs anatomopathologistes à partir d'un examen visuel de lames de biopsies colorées par différents marquages, au travers d'un microscope. Cet examen visuel est très fastidieux, sujet à de nombreuses variations et se révèle peu reproductible.

Dans ce travail, nous nous intéressons à la détection automatique des glomérules dans le cortex rénal qui est un élément essentiel du rein. En effet, le glomérule est la composante principale de l'unité fonctionnelle rénale qui sert à filtrer le sang. Pour l'analyse de la biopsie, son énumération est importante. Comme une biopsie rénale ne représente qu'une faible fraction de l'organe entier (l'échantillon de tissu est inférieur à 20 mm^3 , ce qui représente environ $1/10000$ du volume total du rein), une première vérification avant toute analyse consiste donc à s'assurer que l'échantillon prélevé est représentatif de l'organe étudié. Le critère de représentativité retenu par les experts est le nombre de glomérules présents sur l'échantillon (au minimum 7 selon la classification de Banff [14]).

De manière générale, une difficulté majeure dans l'analyse d'images histologiques est la variation importante des images à traiter. Ces variations peuvent être générées par plusieurs facteurs : la préparation des échantillons (la coupe, qualité du marquage), l'échantillon lui-même, ainsi que l'outil d'acquisition d'images. La figure 1 illustre des variations possibles sur des

images de glomérules colorés par le trichrome de Masson.

1.2 État de l'art

Le domaine de la pathologie digitale [12, 8] est très actif depuis quelques années. Dans ce travail, nous considérons le problème de la détection de glomérules au sein d'images de lames scannées de biopsies. À notre connaissance, peu de travaux visent à automatiser la détection de glomérules. [9] a utilisé un détecteur de contours et des algorithmes génétiques mais a seulement évalué l'approche sur 13 glomérules. Plus récemment, [7] a proposé un détecteur de glomérules utilisant une SVM-linéaire sur des vecteurs de caractéristiques de HOG [1]. Sur des lames immunohistochimiques de reins de souris acquises avec un seul scanner de lames, ils obtiennent de l'ordre de 85% de taux de reconnaissance. Leurs images présentent moins de sources de variations que les nôtres issues de tissus humains, provenant de plusieurs laboratoires, et acquises à l'aide de trois scanners de lames. De plus, la construction des caractéristiques de HOG et le classificateur sont appliqués exhaustivement sur les fenêtres 200×200 pixels des images entières de lames (potentiellement plusieurs millions de positions), ce qui est coûteux.

2 Méthode

Dans cette section, nous décrivons notre procédure de détection des glomérules. Celle-ci repose sur les propriétés géométriques des glomérules qui sont exploitées par une première étape de traitements d'image qui propose un nombre réduit de candidats glomérules par rapport au nombre total de pixels dans

une image. Ensuite, une étape de classification est appliquée pour réduire le nombre de faux positifs.

2.1 Détection de glomérules candidats

Le glomérule est un peloton de capillaires qui est généralement entouré d'un espace de Bowman (zones de lumière) et peut être approximé par une ellipse. Nous proposons une première approche pour la détection des candidats glomérules avec zones de lumière, et une seconde approche pour les glomérules sans.

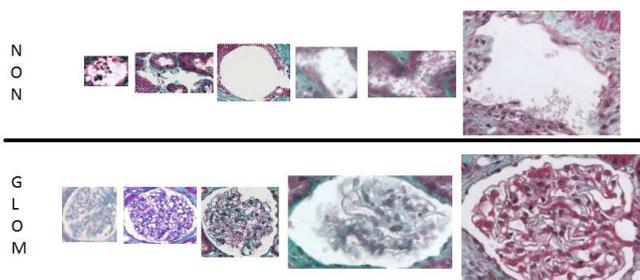


FIGURE 1 – Exemples des variations d'apparence (taille, couleur, texture) des glomérules (en bas) et de non-glomérules (en haut) issus de notre base de données. Tous ces objets sont détectés comme glomérules candidats par la première étape de notre méthodologie.

2.1.1 Avec lumière

Nous utilisons la composante de luminance de l'image couleur seuillée par l'algorithme de Otsu pour détecter les zones de lumière [13]. Les régions détectées retenues sont rendues convexes afin d'améliorer la détection des ellipses. Pour ce faire nous avons implémenté la méthode d'estimation d'ellipse de Fitzgibbon [5]. Elle repose sur l'estimation par moindres carrés sous contrainte du discriminant $4ac - b^2 = 1$ du polynôme des coniques. L'avantage de cette méthode est qu'elle est robuste, efficace et simple à implémenter.

2.1.2 Sans lumière

La seconde approche pour détecter les glomérules sans lumière se base sur la partition de l'image en différentes régions par le regroupement de pixels similaires appelés *superpixels*. Nous utilisons l'algorithme de coupure de graphe de Felzenszwalb [4]. La construction de graphe relie chaque pixel avec ses plus proches voisins dans l'espace à 5 dimensions (x, y, r, g, b) . 3 paramètres sont à régler : σ du filtre Gaussien, k : paramètre d'échelle d'observation et min , la taille minimale des régions retenues. Pour notre étude, nous avons choisi les valeurs suivantes $\sigma = 0.6$; $k = 600$, $min = 500$.

2.2 Classification automatique des candidats

La première étape de notre approche produisant une liste de régions incluant des faux positifs (par exemple des vaisseaux

et tubes détectés comme des glomérules, voir Figure 1), nous proposons une seconde étape qui filtre cette liste. Pour ce faire, une méthode de classification supervisée est proposée afin de distinguer les glomérules des non-glomérules parmi les zones (imassettes) extraites par la première étape. Nous utilisons pour cela une des variantes de [10] qui combine l'extraction de fenêtres aléatoires, la construction d'un ensemble d'arbres aléatoires [6], et la classification finale par un SVM-linéaire. Cette méthode est générique dans le sens où elle ne fait pas d'hypothèse sur les objets à reconnaître et ne nécessite pas la définition de caractéristiques. En réalité, elle tente d'apprendre d'elle-même les caractéristiques qui permettent de discriminer les catégories d'objets sur base d'un ensemble d'apprentissage. Ces caractéristiques correspondent aux fréquences d'apparition des fenêtres en chaque noeud terminal de l'ensemble d'arbres. Cette nouvelle représentation à haute dimension (plusieurs milliers de variables) est injectée dans une SVM-linéaire qui réalise la classification finale.

2.2.1 Paramètres

L'influence des paramètres des variantes de cette méthode a été largement étudiée dans [10]. Dans ce travail, nous avons également évalué l'influence de plusieurs combinaisons de paramètres mais au final nos observations corroborent celles de cette précédente étude. Les paramètres retenus sont les suivants. Un millier de fenêtres aléatoires sont extraites par imasette. Ces fenêtres sont carrées, et dans notre cas elles ont une taille tirée aléatoirement entre 5% et 50% de la taille des imassettes (afin de capturer la texture spécifique des glomérules) et une position tirée aléatoirement de manière à être toujours complètement incluse dans les imassettes. Ces fenêtres sont redimensionnées en 16×16 pixels (interpolation par plus proche voisin) et ceux-ci sont encodés dans un espace de couleur RGB normalisé qui s'avère plus robuste aux changements d'illumination comme déjà observé par [10]. Au total, ce sont donc 768 variables qui décrivent chaque fenêtre. Cet ensemble d'apprentissage est injectée dans l'algorithme d'arbres extrêmement aléatoires [6] dont les paramètres sont les suivants : $T = 10$ arbres, un nombre de tests évalués par noeud interne égal à la racine carrée du nombre de variables décrivant les fenêtres ($K = 28$), un nombre minimal de fenêtres par noeud pour poursuivre son développement $n_{min} = 100$. Une fois l'ensemble d'arbres construit, on représente une imasette par un vecteur dont la dimension est le nombre total de feuilles terminales de l'ensemble d'arbres, et dont la valeur de chaque élément est la proportion des fenêtres extraites de l'imasette qui aboutissent dans ce noeud terminal. Cette représentation avec les paramètres utilisés contient sur nos jeux de données environ 50000 éléments dont au plus 10000 (le nombre de fenêtres extraites multipliées par le nombre d'arbres) sont non-nuls par image. Ce nouvel ensemble d'apprentissage (les imassettes décrites par cette représentation à haute-dimension) est utilisé en entrée d'un SVM-linéaire ($C = 1.0$) qui produit le classificateur binaire final. Lors de l'étape de prédiction, pour chaque imasette candidate

on extrait similairement de manière aléatoire 1000 fenêtres, celles-ci sont propagées dans l'ensemble d'arbres et la représentation ainsi créée est classée par le SVM-linéaire en glomérule ou non-glomérule.

2.2.2 Implémentation

Pour l'étape 1, nous avons développé des plug-ins au sein de Icy [2] (<http://icy.bioimageanalysis.org/>). Pour l'étape 2, nous avons utilisé la même implémentation Java que [10] de l'algorithme d'extraction de fenêtres aléatoires et de construction de caractéristiques par ensembles d'arbres aléatoires. L'étape finale de classification fait appel à la librairie Liblinear [3].

3 Résultats

3.1 Données

Un ensemble de 9 lames ont été utilisées pour notre évaluation. Bien que ce nombre soit réduit, notre base de données exhibe une variabilité relativement importante. En effet, ces lames de biopsies ont été colorées au trichrome Masson au sein de 3 laboratoires hospitaliers et elles ont été numérisées à l'aide de 3 scanners de lames : le Nanozoomer de la société Hamamatsu (objectif 20X, NA=0.75, résolution de $0.46\mu m$ par pixel), le MiraxScan de la compagnie 3DHitech (objectif 20X, NA=0.8, $0.33\mu m$ par pixel), et l'AxioScan.Z1 de la compagnie Zeiss (objectif 20X, NA=0.80, $0.22\mu m$ par pixel). Pour accélérer et faciliter les traitements, les images originales ont été réduites, passant de plusieurs dizaines de milliers de pixels de large à maximum 11000 pixels de large, dépendant de la taille de l'échantillon et de la résolution d'acquisition.

3.2 Protocole d'évaluation

L'objectif de notre approche est de détecter et d'estimer au mieux le nombre de glomérules présents dans chaque section. Afin d'évaluer l'approche dans une configuration réaliste où une nouvelle section est appelée à être traitée sans avoir été utilisée pour mettre au point la méthode, nous proposons un protocole de type "leave-one-slide-out" et plusieurs métriques d'évaluation.

Ce protocole consiste à appliquer la procédure d'apprentissage sur toutes les lames sauf une ($n - 1$), ensuite on teste la procédure complète sur la section restante, et on répète cette évaluation pour les $n = 9$ lames. Plus concrètement, pour l'apprentissage, on commence par appliquer l'étape 1 sur les $n - 1$ lames. Les imagerie de glomérules et non-glomérules de ces $n - 1$ lames (qu'on considère labellisées par les experts) sont utilisées comme ensemble d'apprentissage du modèle de classification. Ensuite, on valide l'approche sur la n ème section en lui appliquant l'étape 1 suivie de la classification automatique des imagerie par le modèle construit sur les $n - 1$ autres lames. Pour chaque section on calcule pour l'étape 1 et la combinaison de l'étape 1 et 2 le nombre de glomérules détectés et

on le compare avec le nombre réel de glomérules au sein de la section, ce qui permet de calculer le pourcentage de vrais positifs et faux positifs.

Au total, l'étape 1 est appliquée sur les images des 9 lames et les 9 modèles de l'étape 2 sont appliqués sur les 9 partitions d'imagerie parmi le total de 183 imagerie candidates (114 glomérules et 69 non-glomérules).

3.3 Taux de reconnaissance

L'étape 1 donne un taux de faux positifs de 37.71% (69 non-glomérules détectés comme glomérules). La table 1 présente les résultats de classification de l'étape 2 sous la forme d'une matrice de confusion qui cumule les prédictions pour l'ensemble des imagerie des 9 lames selon le protocole de "leave-one-slide-out". Après l'étape 2, le taux de faux positifs est de 7.65% (14 non-glomérules détectés comme glomérules). Le taux de reconnaissance du classificateur est de 89.62%, en particulier de 79.71% (55/69) pour les non-glomérules et de 95.61% pour les glomérules (109/114). L'utilisation d'un classificateur permet donc de réduire sensiblement le pourcentage de faux positifs détectés par l'étape 1, comme souhaité, au détriment dans certains cas d'une légère augmentation du taux de faux négatifs puisque l'étape 2 classe 5 glomérules parmi 114 (4.39%) comme non-glomérules. La table 2 détaille les résultats lame par lame et montre que l'étape 1 surestime pour toutes les lames (sauf une) le nombre de glomérules par rapport à la vérité terrain. La combinaison avec l'étape de classification permet de se rapprocher de la vérité terrain fournie par l'expert. La figure 2 illustre une section, le résultat de la détection de l'étape 1, et le résultat final avec l'étape de classification.

TABLE 1 – Matrice de confusion cumulée pour l'ensemble des imagerie des 9 lames selon le protocole "leave-one-slide-out".

	NONGLOM	GLOM
NONGLOM	55	14
GLOM	5	109

4 Conclusions et perspectives

Dans ce travail nous avons proposé et évalué une approche combinant le traitement d'images et la classification supervisée pour la détection de glomérules au sein de biopsies rénales. Les résultats obtenus sont encourageants.

Nous souhaitons comme travail futur mener une évaluation de l'approche sur des centaines de lames provenant de plus de centres hospitaliers et présentant davantage de sources de variations. Une telle évaluation à grande échelle est nécessaire afin d'étudier la robustesse de l'approche (notamment pour l'évaluation des faux négatifs).

De plus, nous envisageons la classification de plusieurs catégories de glomérules telles que les glomérulites. Pour travailler à cette échelle, nous combinerons les logiciels Icy [2] et Cyto-

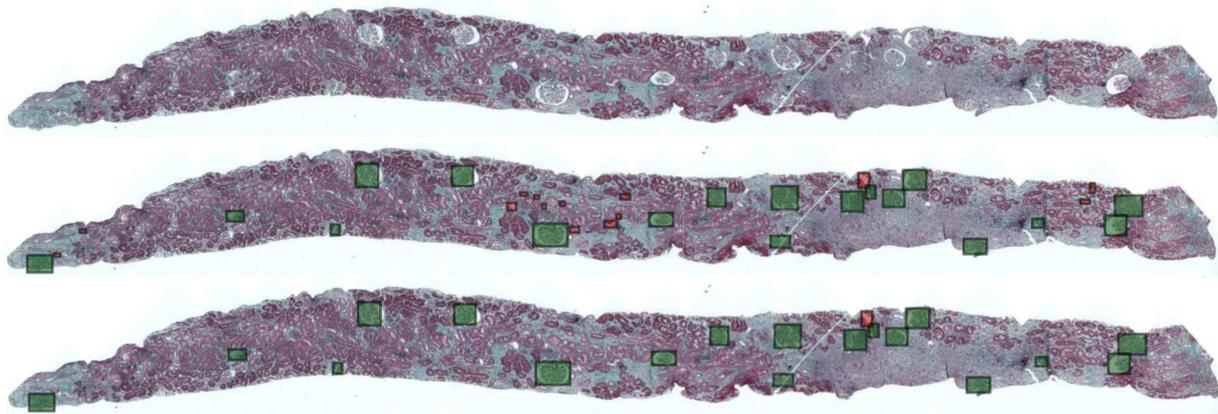


FIGURE 2 – Résultats de détection sur la section 4. Image originale (haut), détections de l'étape 1 avec 14 faux positifs (milieu), détections après l'étape 2 avec 1 faux positif (bas). En vert les glomérules correctement détectés, en rouge les non-glomérules détectés comme glomérules.

TABLE 2 – Nombre (NB) de glomérules détectées pour chaque section par l'expert, l'étapes 1 seule, les étapes 1 et 2 combinées, vrais positifs (TP), faux positifs (FP), vrais négatifs (TN), et faux négatifs (FN).

Section	# Expert	# Etape 1	# Etape 1+2
	NB	NB (TP/FP/TN/FN)	NB (TP/FP/TN/FN)
1	6	8 (6/2/0/0)	6 (6/0/2/0)
2	7	21 (7/14/0/0)	8 (7/1/13/0)
3	10	17 (10/7/0/0)	10 (10/0/7/0)
4	18	32 (18/14/0/0)	19 (18/1/13/0)
5	12	16 (12/4/0/0)	12 (11/1/3/1)
6	17	26 (17/9/0/0)	22 (16/6/3/1)
7	9	9 (9/0/0/0)	9 (9/0/0/0)
8	13	16 (13/3/0/0)	14 (13/1/2/0)
9	22	38 (22/16/0/0)	23 (19/4/12/3)

mine [11] afin de faciliter le travail des pathologistes pour corriger les erreurs de nos algorithmes et itérativement améliorer notre approche.

5 Remerciements

Nous tenons à remercier l'Institut Pasteur Carnot Maladies Infectieuses pour la bourse postdoctorale de R. Marée ainsi que nos collègues médecins des hôpitaux Necker enfants malades, européen Georges Pompidou et Edouard Herriot de Lyon pour nous avoir fourni les lames de biopsies.

Références

[1] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. *Proc. CVPR*, pages 886–893, 2005.

[2] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Herve, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J. C. Olivo-Marin. Icy : an open bioimage informatics platform for extended reproducible research. *Nat Methods*, 9(7) :690–6, 2012.

[3] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. Liblinear : A library for large linear classification. *Journal of Machine Learning Research*, 9 :1871–1874, 2008.

[4] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *IJCV*, 59(2) :167–181, 2004.

[5] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least-squares fitting of ellipses. *IEEE Transactions on PAMI*, 21(5) :476–480, May 1999.

[6] P. Geurts, D. Ernst, and L. Wehenkel. Extremely randomized trees. *Machine Learning*, 36(1) :3–42, 2006.

[7] Y. Hirohashi, R. Relator, T. Kakimoto, R. Saito, Y. Horai, A. Fukunari, H. Utsumi, K. Okada, and T. Kato. Automated quantitative image analysis of glomerular desmin immunostaining as a sensitive injury marker in spontaneously diabetic torii rats. *Journal of Biomedical Image Processing*, 1(1) :20–28, May 2014.

[8] S. Kothari, J.H. Phan, T.H. Stokes, and M.D. Wang. Pathology imaging informatics for quantitative analysis of whole-slide images. *Journal of the American Medical Informatics Association*, 20(6) :1099–1108, November-December 2013.

[9] J. Ma and Zhang J. an Hu J. Glomerulus extraction by using genetic algorithm for edge patching. In *Proc. of IEEE Congress on Evolutionary Computation*, pages 2474–2479, 2009.

[10] R. Marée, P. Geurts, and L. Wehenkel. Towards generic image classification : a extensive empirical study. Technical report, University of Liège, December 2014. Available online at <http://orbi.ulg.ac.be/handle/2268/175525>.

[11] R. Marée, B. Stevens, L. Rollus, N. Rocks, X. Moles-Lopez, I. Salmon, D. Cataldo, and L. Wehenkel. A rich internet application for remote visualization and collaborative annotation of digital slide images in histology and cytology. *BMC Diagnostic Pathology*, 8 S1, 2013. <http://www.cytomine.be/>.

[12] MT. McCann, C. Castro, JA. Ozolek, B. Parvin, and J. Kovacevic. Automated histology analysis : opportunities for signal processing. *IEEE Signal Processing*, 32(1) :78–87, 2014.

[13] V. Meas-Yedid, A. Servais, L. H. Noel, C. Panterne, P. Landais, N. Herve, N. Brousse, H. Kreis, C. Legendre, E. Thervet, J. C. Olivo-Marin, and E. Morelon. New computerized color image analysis for the quantification of interstitial fibrosis in renal transplantation. *Transplantation*, 92(8) :890–9, 2011.

[14] B. Sis et al. Banff '09 meeting report : antibody mediated graft deterioration and implementation of banff working groups. *Am J Transplant*, 10(3) :464–71, 2010.