

Réseaux de neurones convolutionnels multi-échelle pour la classification cellulaire

Pierre Buyssens, Abderrahim Elmoataz

Université de Normandie, UNICAEN, ENSICAEN, GREYC UMR CNRS 6072

6 Bd. du maréchal Juin, 14050, CAEN

{prénom.nom}@unicaen.fr

Résumé

Dans cet article, nous présentons une approche basée sur des réseaux de neurones convolutionnels multi-échelle (MCNN) pour la classification de cellules. Basée sur plusieurs réseaux profonds convolutionnels agissant à différentes résolutions, l'architecture proposée évite l'étape classique d'extraction manuelle de caractéristiques, en procédant à l'extraction de caractéristiques et à la classification en une fois au sein d'un même réseau de neurones. L'approche proposée fournit de meilleurs résultats de classification que les méthodes usuelles permettant une aide au diagnostic plus sûre du cancer de la plèvre.

Mots Clef

Classification, réseaux profonds, multi-échelle, cytologie

Abstract

We present a Multiscale Convolutional Neural Network (MCNN) approach for vision-based classification of cells. Based on several deep Convolutional Neural Networks (CNN) acting at different resolutions, the proposed architecture avoid the classical handcrafted features extraction step, by processing features extraction and classification as a whole. The proposed approach gives better classification rates than classical state-of-the-art methods allowing a safer Computer-Aided Diagnosis of pleural cancer.

Keywords

Classification, deep networks, multiscale, cytology

1 Introduction et contexte

L'aide informatique au diagnostic (*Computer-Aided Diagnosis*, CAD) à l'aide du traitement d'images [5, 8] suscite de plus en plus d'intérêt. L'extraction automatique de caractéristiques pertinentes par un ordinateur peut en effet faire gagner un temps précieux aux pathologistes, et, étant donné l'accroissement constant de données médicales virtuelles, cette aide informatique au diagnostic est devenue un élément essentiel pour l'obtention de diagnostics fiables, rapides et répétables.

Nos travaux s'inscrivent dans le cadre de l'analyse cytologique de cellules cancéreuses pour le diagnostic précoce du cancer de la plèvre. Les fluides corporels obtenus par

ponction peuvent être examinés au sein d'une préparation à l'aide d'un microscope. Le but de notre étude est de fournir aux pathologistes une classification automatique des cellules. Cette classification aide ensuite le cytopathologiste à distinguer les cellules mésothéliales anormales et dystrophiques qui se révèlent être très importantes lors d'un diagnostic précoce des cancers liés à l'amiante [1].

Les images traitées dans cette étude sont obtenues par coloration de Feulgen pour laquelle la chromatine est colorée en rose et le cytoplasme n'est pas coloré (ou très peu). En conjonction avec d'autres préparations, la coloration de Feulgen fournit des informations très utiles au cytopathologiste pour l'établissement du diagnostic. La classification de cellules sous coloration de Feulgen a déjà été étudiée dans [14]. Nos travaux proposent de traiter des lames virtuelles scannées à l'aide d'un scanner Leica SCN400 à la résolution 40×, rendant ainsi des images à pleine résolution pouvant atteindre la taille de 80000 × 90000 pixels.

Le fluide corporel contient plusieurs types de cellules pouvant présenter différents arrangements de leur chromatine [1]. Le but d'un outil de traitement automatique dans ce cas est de classer les centaines de milliers de cellules présentes sur la lame virtuelle en différentes classes prédéfinies : cellules malignes, cellules dystrophiques, cellules mésothéliales normales, macrophages, et toutes les cellules inflammatoires telles que les polynucléaires ou les lymphocytes. Pour réaliser un tel classement, la répartition de la chromatine au sein de la cellule est en fait un point critique pour distinguer les cellules malignes, dystrophiques, et mésothéliales normales.

Un traitement complètement automatique de lames virtuelles est classiquement décomposé en 3 étapes :

- Segmentation et extraction des objets cellulaires contenus dans la lame virtuelle,
- extraction de caractéristiques pour chaque cellule détectée,
- classification des cellules selon les caractéristiques extraites précédemment.

Dans cet article, nous proposons une approche basée sur la vision pour la classification cellulaire. Les deux dernières étapes sont fusionnées au sein de réseaux de neurones profonds qui réalisent à la fois l'extraction automatique des caractéristiques et leur classification. De plus, une approche

multi-échelle est proposée utilisant plusieurs réseaux profonds à différentes échelles (taille de l'entrée), et dont les sorties sont fusionnées.

La suite de l'article est organisée comme suit : La section 2 décrit l'approche profonde multi-échelle (MCNN pour *Multiscale Convolutional Neural Network*) proposée, tandis que les expérimentations et résultats sont présentés à la section 4. La section 5 conclut cet article et dresse quelques pistes de futurs travaux.

2 Réseaux de neurones convolutionnels multi-échelle

Les réseaux de neurones convolutionnels (CNN pour *Convolutional Neural Networks*) sont des réseaux de neurones multi-couches qui sont spécialisés dans des tâches de reconnaissance de forme [7], [4]. Ils sont réputés pour leur robustesse aux faibles variations d'entrée, le faible taux de prétraitement nécessaires à leur fonctionnement, et ne requièrent aucun choix d'extracteur de caractéristiques spécifiques. L'architecture proposée repose sur plusieurs réseaux de neurones profonds alternant entre couches de convolution et couches d'agrégation (*pooling*). Ces réseaux de neurones profonds appartiennent à une classe de modèles s'inspirant des travaux de Hubel et Wiesel sur le cortex visuel primaire du chat [2]. Le début de l'architecture consistant en une succession de couches de convolution et d'agrégation est dédiée à l'extraction automatique de caractéristiques, tandis que la seconde partie, composée de couches de neurones complètement connectés, est dédiée à la classification.

2.1 Couche de convolution

Une couche de convolution C^i (couche i du réseau) est paramétrée par son nombre N de cartes de convolution M_j^i ($j \in \{1, \dots, N\}$), la taille des noyaux de convolution $K_x \times K_y$ (souvent carrée), et le schéma de connexion à la couche précédente L^{i-1} . Chaque carte de convolution M_j^i est le résultat d'une somme de convolution des cartes de la couche précédente M_j^{i-1} par son noyau de convolution respectif. Un biais b_j^i est ensuite ajouté et le résultat est passé à une fonction de transfert non-linéaire $\phi(x) = 1.7159 \tanh(\frac{2}{3}x)$ [4]. Dans le cas d'une carte complètement connectée aux cartes de la couche précédente, le résultat est alors calculé par :

$$M_j^i = \phi \left(b_j^i + \sum_{n=1}^N M_n^{i-1} * K_n^i \right)$$

où $*$ est l'opérateur de convolution.

2.2 Couche de max-aggrégation

Dans les architectures classiques de réseaux de neurones convolutionnels, les couches de convolution sont suivies par des couches de sous-échantillonnage. Une couche de sous-échantillonnage réduit la taille des cartes, et introduit

de l'invariance aux (faibles) rotations et translations pouvant apparaître en entrée. Une couche de max-aggrégation (*max-pooling*) est une variante d'une telle couche qui a montré certains atouts dans le passé [11]. La sortie d'une couche de max-aggrégation est donnée par la valeur maximale d'activation au sein de la couche d'entrée pour différentes régions de taille $K_x \times K_y$ ne se chevauchant pas. De façon similaire à une couche de convolution, un biais est ajouté et le résultat est passé à la fonction de transfert ϕ définie plus haut.

2.3 Couche de neurones classique

Les paramètres des couches de convolution et de *pooling* sont choisis de sorte que les cartes d'activation de la dernière couche soient de taille 1, ce qui résulte en un vecteur 1D d'attributs. Des couches classiques complètement connectées composées de neurones sont alors ajoutées au réseau pour réaliser la classification. La dernière couche, dans le cas d'un apprentissage supervisé, contient autant de neurones que de classes désirées. Dans nos travaux, cette dernière couche contient 6 neurones, et une fonction d'activation de type softmax est utilisée afin d'obtenir des probabilités d'appartenance à chaque classe.

2.4 CNN multi-échelle (MCNN)

Il est généralement admis que la vision humaine est un processus multi-échelle. Dans ces travaux, nous créons N réseaux de neurones convolutionnels avec différentes tailles de rétine (i.e., la taille de la rétine correspond à la taille des images d'entrée). Une entrée donnée est redimensionnée N fois pour correspondre à la rétine de chaque réseau. La question de l'optimisation des sorties des classifieurs est une question récurrente dans le domaine de la reconnaissance de formes. Pour la reconnaissance de caractères manuscrits, il a été montré [6] qu'un simple moyennage des sorties de chaque classifieur permet d'obtenir des taux de classification au moins aussi bons qu'avec une combinaison linéaire dont les poids auraient été appris via une validation croisée [12]. Cependant, dans [6], il n'y a aucune raison de pondérer les classifieurs étant donné qu'ils agissent à la même résolution et diffèrent entre eux seulement en fonction des distorsions appliquées aux entrées. Dans notre cas, une hypothèse raisonnable est que le classifieur à la résolution la plus faible peut être moins pertinent que celui à la résolution la plus haute. Les résultats finaux sont donc calculés, dans notre cas, comme une combinaison linéaire des sorties des N CNN. L'architecture globale du MCNN proposée est montrée à la figure 1.

3 Expérimentations

Dans cette section, nous détaillons les différentes expérimentations que nous avons mené. Afin de comparer les performances de notre approche à d'autres schémas de classification (où les caractéristiques sont extraites manuellement), une étape préalable de segmentation manuelle des cellules a été réalisée.

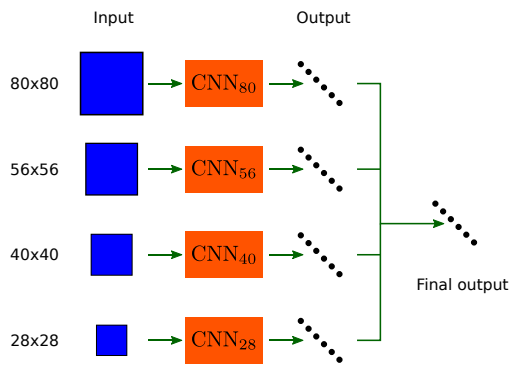


FIGURE 1 – Architecture globale du réseau de neurones convolutifs multi-échelle.

3.1 Acquisition des données et segmentation

Une base de données a été acquise par un cytopathologiste (quelques échantillons sont montrés à la figure 2) et est composée de :

- 215 *mésothéliales anormales* (notées C_1 dans la suite),
- 209 *mésothéliales dystrophiques* (C_2),
- 201 *mésothéliales normales* (C_3),
- 195 *macrophages* (C_4),
- 198 *polynucléaires* (C_5),
- 196 *lymphocytes* (C_6).

Cette base de données (et la segmentation de chaque cellule) est rendue disponible à des fins de recherche¹. Elle contient de nombreuses difficultés et potentiellement des erreurs d'annotation (en faible nombre toutefois). La figure 3 montre quelques cas difficiles où des cellules en apparence similaire ont été annotées comme appartenant à des classes différentes.

Une segmentation manuelle des cellules de la base est effectuée afin de séparer les pixels appartenant à la cellule de ceux appartenant au fond. La taille des imageries varie et ne sont pas nécessairement carrées. Après une inspection visuelle de la base de données, la taille de 80×80 est retenue pour la taille des images d'entrée étant donné qu'aucune cellule n'excède cette taille. Un padding est ensuite effectué sur toutes les images de la base de sorte qu'elles soient de cette taille.

3.2 Architecture des réseaux

Nous contruisons 4 CNN agissant à différentes résolutions. La taille des images à pleine résolution est de 80×80 . Cette taille est successivement divisée par un facteur $\sqrt{2}$, $\sqrt{2}^2$ et $\sqrt{2}^3$ selon chaque axe pour donner respectivement des images de taille 56×56 , 40×40 and 28×28 . Les réseaux, appelés dans la suite CNN_{80} , CNN_{56} , CNN_{40} et CNN_{28} , sont construits de manière similaire au réseau LeNet [4]. Ils diffèrent essentiellement par la taille de leur rétine et la taille de leurs noyaux de convolution. L'architecture dé-

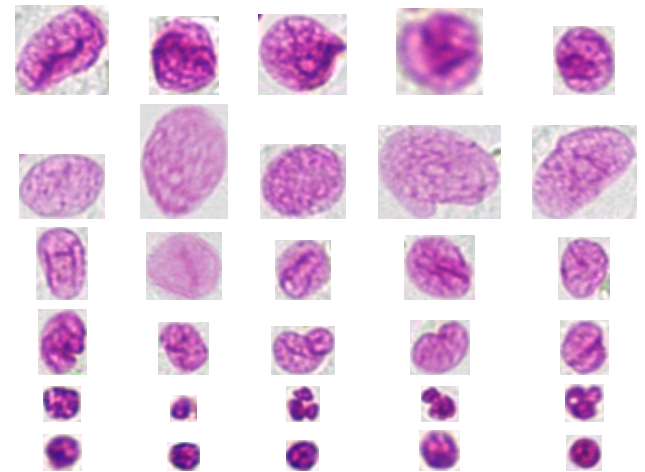


FIGURE 2 – Échantillons de la base de données. De haut en bas : mésothéliales anormales (C_1), mésothéliales dystrophiques (C_2), mésothéliales normales (C_3), macrophages (C_4), polynucléaires (C_5), lymphocytes (C_6).

TABLE 1 – Architectures des réseaux. Dans chaque case du tableau : Haut : taille du noyau, et Bas : taille de la carte en sortie.

	CNN_{80}	CNN_{56}	CNN_{40}	CNN_{28}
Taille rétine	80×80	56×56	40×40	28×28
C1	7×7 74×74	7×7 50×50	5×5 36×36	5×5 24×24
P2	2×2 37×37	2×2 25×25	2×2 18×18	2×2 12×12
C3	6×6 32×32	6×6 20×20	5×5 14×14	5×5 8×8
P4	4×4 8×8	4×4 5×5	2×2 7×7	2×2 4×4
C5	8×8 1×1	5×5 1×1	7×7 1×1	4×4 1×1

taillée du réseau à pleine résolution est schématisé à la figure 4. Les architectures des autres réseaux sont résumées au tableau 1.

3.3 Approches de l'état de l'art

Les approches classiques de la littérature pour la classification cellulaire reposent sur l'extraction de caractéristiques de forme, et de texture, ou de caractéristiques photométriques des cellules. Les caractéristiques de forme incluent la *surface*, le *périmètre*, la *compacité*, ou encore l'*elongation* [10]. Les caractéristiques photométriques incluent les *moyenne* et *écart-type* des couleurs selon le spectre choisi. La *densité optique intégrée* (DOI) est également un critère de choix pour les pathologistes et est conçue pour refléter une spécificité visuelle de marquage

1. <https://sites.google.com/site/pierrebuysens/cells-database>

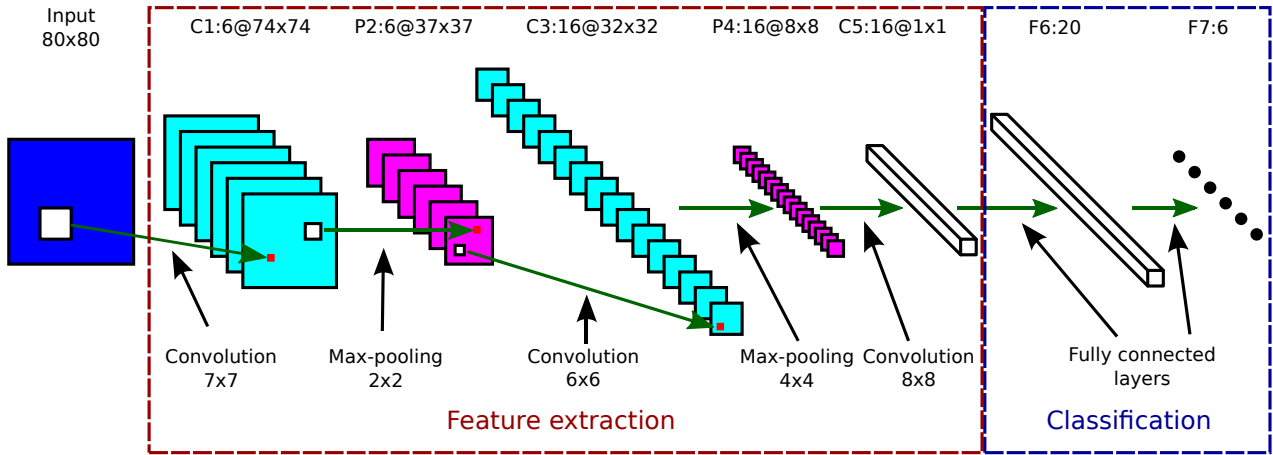


FIGURE 4 – Architecture du réseau CNN₈₀.

des cellules [10]. Elle est calculée par :

$$DOI = \sum_{x,y} DO(x,y) \quad \forall x,y \in \text{Cellule}$$

où DO est la *densité optique* définie par

$$DO(x,y) = -\log\left(\frac{I(x,y)}{I_0}\right)$$

avec I_0 la couleur moyenne du fond, et (x,y) le couple de coordonnées appartenant à la cellule. Finalement, l'analyse de texture a pour but d'extraire des caractéristiques importantes de la cellule, et en tout premier lieu des informations liées à la distribution de chromatine contenue dans le noyau, ce qui indique une malignité potentielle. Une telle analyse inclue des caractéristiques *morphologiques* [14], des caractéristiques d'*ondelettes* [5], des caractéristiques de forme et de couleur calculées sur des sous-régions [10], et des modèles hiérarchiques de texture [13] basés sur les textons [3].

Afin de comparer notre approche à l'état de l'art, nous avons implémenté la plupart de ces caractéristiques, et les avons extraites de chaque cellule (à pleine résolution) de la base. Certaines de ces caractéristiques sont extraites sur la cellule entière, tandis que d'autres ne sont calculées que sur des parties du noyau. Certaines images auxiliaires calculées pour l'extraction de ces caractéristiques sont montrées à la figure 5. En particulier, nous avons extrait certaines caractéristiques couleur sur des images de partitions des cellules (figures 5(c) et 5(f)) afin de refléter la distribution de chromatine à l'intérieur de la cellule : pour une cellule anormale, la chromatine tend à se concentrer près des bords de la cellule. Les cellules ont alors été divisées en rond concentriques (figure 5(c)) ou en régions géodésiques (basées sur des graphes, figure 5(f)) à l'aide d'un graphe d'adjacence de régions (RAG, figure 5(e)) calculé via un watershed sur la cellule (figure 5(d)).

Une validation croisée (à 10 ensembles) a été réalisée sur ces caractéristiques à l'aide d'un SVM et d'un noyau Gaus-

sien. Le taux d'erreur de classification final obtenu est présenté au tableau 2.

Une validation croisée similaire a également été réalisée sur ces caractéristiques à l'aide d'un réseau de neurones (NN), contenant 103 neurones en entrées (la taille d'un vecteur de caractéristiques extrait d'une cellule), 80 neurones sur la couche cachée, et 6 en sortie (le nombre de classes). Le taux d'erreur de classification final obtenu est présenté au tableau 2.

Finalement, ces caractéristiques ont également été classifiées à l'aide d'un SVM multi-noyaux (MKL) [9]. Ce schéma de classification utilise un noyau par caractéristiques (ou groupe de caractéristiques) et apprend la pondération adéquate. Le taux d'erreur de classification final obtenu est présenté au tableau 2.

Notons ici que toutes ces expérimentations ont été réalisées plusieurs fois afin de nous assurer des résultats fiables. Les taux d'erreur présentés dans au tableau 2 sont un moyennage des taux d'erreur calculés à chaque réalisation.

3.4 Résultats de classification

Avant l'apprentissage des réseaux, les images sont prétraitées (première ligne de la figure 6). Étant donné que la couleur complémentaire des cellules roses et le vert, seul ce canal est retenu dans la suite. Les valeurs des pixels sont normalisées de sorte que elles soient comprises entre -1 et 1 (les valeurs des pixels appartenant au fond sont alors proches de 0). Les images sont alors redimensionnées pour s'adapter aux différentes tailles de rétines.

Une validation croisée (à 10 ensembles) est effectuée afin de tester notre approche. Étant donné que la classification doit être invariante aux rotations, l'ensemble d'apprentissage est augmenté artificiellement en procédant à des rotations des images de l'ensemble d'apprentissage avec un pas de 10 degrés (deuxième ligne de la figure 6). Les ensembles d'apprentissage sont alors composés d'approximativement 39000 images. Les poids initiaux des réseaux sont initialisés au hasard avec de faibles valeurs, et un apprentissage

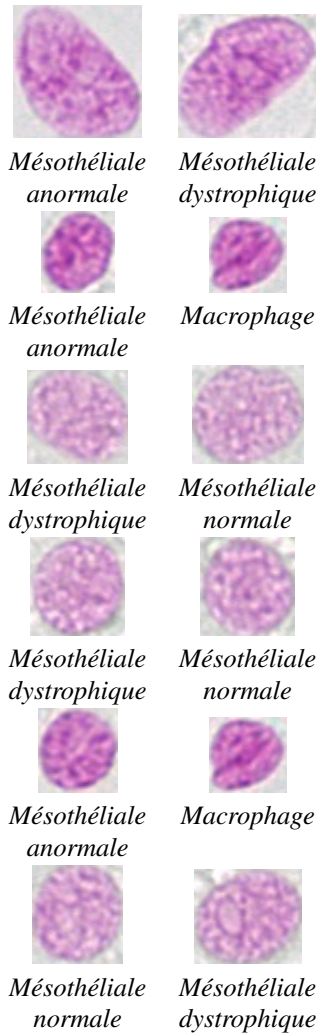


FIGURE 3 – En colonne, cellules similaires appartenant à des classes différentes. Le nom de la classe est inscrit sous chaque cellule.

stochastique est réalisé à l’aide de techniques classiques de momentum et de méthodes du second ordre pour l’accélération de l’apprentissage. Celui-ci se termine lorsque le taux d’erreur sur un sous-ensemble (ensemble de validation représentant 10% de la base) ne décroît plus (typiquement après 50 époques). Les taux d’erreur de classification finaux obtenus pour chaque réseau sont présentés au tableau 2. Comme attendu, ces taux diminuent lorsque la résolution des images d’entrée croît, étant donné que les hautes résolutions sont capables de capturer plus d’informations. Nous avons ensuite testé deux schémas de fusion des sorties des 4 réseaux :

- Le premier, et le plus simple, (MCNN dans le tableau 2) moyenne les sorties de chaque CNN ;
- Le second (w MCNN dans le tableau 2) effectue une pondération des sorties des réseaux.

Concernant ce dernier schéma, les poids sont calculés pour

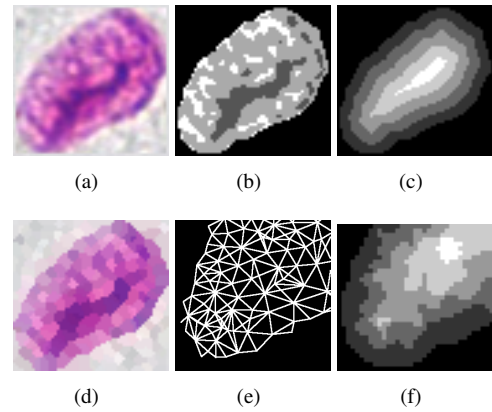


FIGURE 5 – (a) Cellule extraite, (b) image de textons ronds concentriques, (c) image partitionnée, (d) image partitionnée, (e) extrait du graphe d’adjacence de régions, et (f) extrait du partitionnement en régions basées graphe.

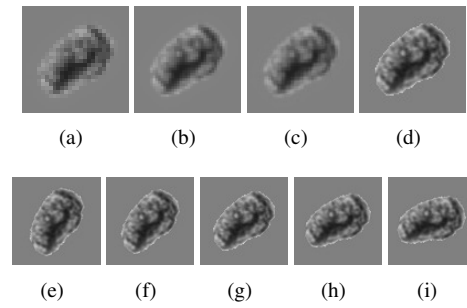


FIGURE 6 – Première ligne : une cellule anormale à différentes résolutions. Deuxième ligne : rotations artificielles de la cellule.

chaque CNN par :

$$w(CNN_x) = \frac{\text{moyenne}(cr_x)}{\text{std}(cr_x)}$$

où cr_x est le taux de classification du réseau CNN_x . Cette pondération donne un poids plus faible aux réseaux ayant de faibles taux de classification, et inversement, donne un poids plus élevé aux réseaux ayant de bons scores. Les poids sont ainsi croissants selon la résolution étant donné que les plus hautes résolutions sont plus pertinentes. Un schéma de vote majoritaire (MV) a également été testé et le taux d’erreur de classification est présenté au tableau 2. La meilleure combinaison se trouve être celle ne prenant pas en compte le réseau à la résolution la plus faible CNN_{28} . Il se trouve que le taux d’erreur le plus faible est obtenu à l’aide du deuxième schéma de fusion de données sans prise en compte du réseau CNN_{28} . Dans tout les cas, la fusion des sorties des réseaux (avec ou sans CNN_{28}) donne de meilleurs résultats (taux d’erreur inférieurs) que les approches de l’état de l’art considérées.

Le tableau 3 montre la matrice de confusion pour la

TABLE 2 – Évaluation des différentes méthodes (Écart-type entre parenthèses).

Méthode	Taux d'erreur
CNN ₂₈	9.04% (1.07)
CNN ₄₀	8.02% (1.02)
CNN ₅₆	8.01% (0.98)
NN	7.85% (0.98)
SVM	7.75% (0.95)
CNN ₈₀	7.55% (0.89)
MV /w CNN ₂₈	6.13% (0.84)
MKL	6.12% (0.83)
MCNN	6.02% (0.82)
^w MCNN	5.90% (0.80)
^w MCNN /w CNN ₂₈	5.74% (0.79)

meilleur approche (pondération des réseaux sans prise en compte de celui à la résolution la plus faible : ^wMCNN /w CNN₂₈). On peut ainsi observer que les cellules mal classées sont principalement confondues avec celles des classes voisines (valeurs différentes de 0 proches de la diagonale dans la matrice de confusion). Le bon taux de classification des cellules mésothéliales *anormales* (classe C_1), qui sont les plus importantes lors d'un diagnostic précoce du cancer de la plèvre, est encourageant. Les échantillons mal classés sont principalement confondus avec des cellules mésothéliales dystrophiques (classe C_2) ou des cellules inflammatoires (macrophages, classe C_4). À noter que les différences entre ces classes peuvent parfois être minces (voir la figure 3 pour quelques exemples de cas difficiles).

4 Discussion

L'approche proposée dans cet article évite la conception et/ou l'extraction *manuelle* de caractéristiques. Les approches classiques sont typiquement confrontées à des problèmes difficiles tels que :

- La conception de caractéristiques robustes et faciles à calculer,
- l'évaluation de ces caractéristiques et leur pertinence pour la séparation des classes,
- et le cas plus général de la sélection des caractéristiques pertinentes.

Notre approche basée sur les réseaux de neurones convolutionnels contourne ces difficultés et évite l'étape délicate liée aux caractéristiques.

L'approche consistant à utiliser plusieurs réseaux de neurones convolutionnels pour la même tâche a déjà été utilisée dans [6]. L'idée sous-jacente est que chaque réseau va *apprendre* différentes caractéristiques d'une même entrée. Afin de tendre vers ce but, les auteurs de [6] appliquent différentes distorsions (couleur, forme) à l'ensemble d'apprentissage. Cette approche ne peut être utilisée ici étant donné que l'application de distorsions à des cellules peut changer leur taille ou leur distribution de

TABLE 3 – Matrice de confusion de la classification (en pourcentages) pour la meilleure méthode. En lignes : vérité terrain, en colonnes : classe trouvée.

	C_1	C_2	C_3	C_4	C_5	C_6
C_1	97.41	0.90	0.77	0.90	0	0
C_2	2.39	93.75	3.85	0	0	0
C_3	1.24	12.58	80.22	5.94	0	0
C_4	0.14	0	2.56	97.15	0.14	0
C_5	0	0	0	0.42	99.14	0.42
C_6	0	0	0	0.28	1.82	97.89

chromatine, ce qui n'est pas viable pour la classification. Le moyen simple que nous proposons pour obtenir plus d'informations est donc l'utilisation d'une approche multi-échelle et la construction de réseaux adéquats.

Un autre problème difficile (que nous n'avons pas traité ici) et la nécessité d'avoir un outil de segmentation robuste étant donné que les caractéristiques d'une cellule sont principalement calculées sur une version masquée (selon la segmentation) de celle-ci. Par exemple, si une cellule est mal segmentée et que son masque contient des pixels du fond, sa surface sera plus grande, et une simple mésothéliale normale peut alors être confondue avec une mésothéliale dystrophique. La segmentation doit également être rapide étant donné que les lames virtuelles à haute résolution peuvent contenir des centaines de milliers d'objets cellulaires. Dans le cadre des travaux présentés dans cet article, un (relativement) faible nombre de cellules a été segmenté manuellement par un cytopathologiste. De futurs travaux impliquent la mise au point et l'extension de notre approche pour la détection automatique des cellules (et ainsi éviter l'étape de segmentation, remplacé par un module de détection), ainsi que des tests de robustesse de notre approche lorsque les cellules ne sont plus segmentées (fond non mis à zéro).

5 Conclusion

Nous avons présenté dans cet article une approche de classification basée sur un ensemble de réseaux de neurones convolutionnels (MCNN). Reposant sur plusieurs réseaux agissant à des résolutions différentes, notre approche évite l'étape fastidieuse d'extraction manuelle de caractéristiques. L'approche proposée atteint les meilleurs taux de classification sur une base de données de cellules par rapport aux méthodes classiques de l'état de l'art.

Les travaux futurs incluent la mise au point d'un module de détection de cellules afin d'éviter l'étape fastidieuse de segmentation. De plus, notre approche est largement parallélisable, et une implémentation sur GPU devrait permettre d'accélérer grandement le processus.

Références

- [1] A. Churg, P. T. Cagle, and V. L. Roggli. *Tumors of the Serosal Membrane*. AFIP Atlas of Tumor Pathology

– Series 4, American Registry of Pathology, 2006.

- [2] D. H. Hubel and T. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Physiol*, 160 :106–154, 1962.
- [3] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290 :91–97, March 1981.
- [4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of IEEE*, 86(11) :2278–2324, November 1998.
- [5] Jihene Malek, Abderrahim Sebri, Souhir Mabrouk, Kholdoun Torki, and Rached Tourki. Automated breast cancer diagnosis based on gvf-snake segmentation, wavelet features extraction and fuzzy classification. *J. Signal Process. Syst.*, 55(1-3) :49–66, April 2009.
- [6] Ueli Meier, Dan Claudiu Cireșan, Luca Maria Gambardella, and Jürgen Schmidhuber. Better digit recognition with a committee of simple neural nets. In *ICDAR*, pages 1250–1254. IEEE, 2011.
- [7] Claus Nebauer. Evaluation of convolutional neural networks for visual recognition. *IEEE Transactions on Neural Networks*, 9(4) :685–696, 1998.
- [8] J. H. Puls, H. E. Dytch, and M. R. Roman. Automated screening for cervical cancer using image processing techniques. In *Proceedings of 1980 FPS Users Group Meeting*, April 1980.
- [9] Alain Rakotomamonjy, Francis Bach, Stéphane Canu, and Yves Grandvalet. SimpleMKL, January 27 2008.
- [10] Karsten Rodenacker. A feature set for cytometry on digitized microscopic images. *Cell. Pathol*, 25 :1–36, 2001.
- [11] Dominik Scherer, Andreas Müller, and Sven Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *ICANN*, pages 92–101, 2010.
- [12] N Ueda. Optimal linear combination of neural networks for improving classification performance. *PAMI*, 22(2) :207–215, 2000.
- [13] G. Wolf, M. Beil, and H. Guski. Chromatin structure analysis based on a hierarchic texture model. *Analytical and quantitative cytology and histology*, 17(1) :25–34, 1995.
- [14] Martin Zarzycki, Timna E. Schneider, Dietrich Meyer-Ebrecht, and Alfred Böcking. Classification of cell types in feulgen stained cytologic specimens using morphologic features. In *Bildverarbeitung für die Medizin*, pages 410–414, 2005.