



HAL
open science

Méthodes de classification du cancer du sein pour les microscopes à réalité augmentée

Robin Heckenauer, Jonthan Weber, Cédric Wemmert, Michel Hassenforder, Pierre-Alain Muller, Germain Forestier

► **To cite this version:**

Robin Heckenauer, Jonthan Weber, Cédric Wemmert, Michel Hassenforder, Pierre-Alain Muller, et al.. Méthodes de classification du cancer du sein pour les microscopes à réalité augmentée. RFIAP 2022 (Congrès Reconnaissance des Formes, Image, Apprentissage et Perception), Jul 2022, Vannes, France. hal-03687423

HAL Id: hal-03687423

<https://hal.science/hal-03687423v1>

Submitted on 3 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Méthodes de classification du cancer du sein pour les microscopes à réalité augmentée

R. Heckenauer¹
M. Hassenforder¹

J. Weber¹
P-A. Muller¹

C. Wemmert²
G. Forestier¹

¹ IRIMAS, Université de Haute-Alsace, France

² ICube, Université de Strasbourg, France

robin.heckenauer@uha.fr

Résumé

Un microscope à réalité augmentée (ARM) permet d'afficher des informations supplémentaires sur les tissus en cours d'analyse par un pathologiste. Les méthodes d'analyse d'image utilisées par ces microscopes doivent être robustes aux changements de niveau de grossissement pour suivre les déplacements dans les lames de verres.

Nous résolvons ce problème en tirant parti des caractéristiques présentes dans certains niveaux de grossissement clefs qui permettent de mieux comprendre d'autres niveaux de grossissement.

Nous proposons une méthode temps réel robuste aux changements de niveaux de grossissement en entraînant des réseaux de neurones profonds sur certains niveaux clefs d'images de lames (WSI) et en les testant sur les niveaux d'un microscope.

Nous montrons que notre approche surpasse ou équivaut les méthodes naïves sur un jeu de données de cancer du sein pour l'architecture d'apprentissage profond Inception-ResNet-v2.

Mots Clef

Classification en temps réel du cancer du sein, pathologie numérique, microscope à réalité augmentée, réseaux convolutifs profonds.

Abstract

An Augmented Reality Microscope (ARM) displays additional information about the tissues being analyzed by a pathologist. The image analysis methods used by these microscopes must be robust to changes in magnification level in order to follow the displacements in the glass slides.

We solve this problem by taking advantage of features present in some key magnification levels that improve the results of other magnification levels.

We propose a real-time method robust to changes in magnification levels by training deep neural networks on some key Whole Slide Images (WSI) and testing them on common microscope levels.

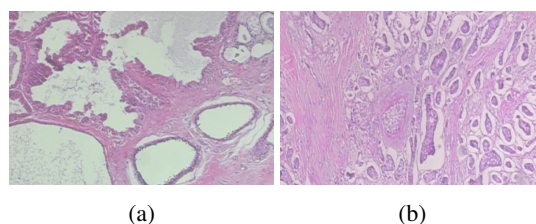


FIGURE 1 – Patches de tumeurs bénignes et malignes. Exemples de tumeur bénigne (a) et de tumeur maligne (b) telles qu'elles apparaissent dans une WSI d'un niveau de grossissement 40X, dans le cas d'un cancer du sein (coloration H&E).

We show that our approach outperforms or equals naive methods on a breast cancer dataset for the Inception-ResNet-v2 deep learning architecture.

Keywords

Real-time breast cancer classification, digital pathology, augmented reality microscope, deep convolutional networks.

1 Introduction

1.1 Pathologie numérique

L'introduction des scanners de lames vers la fin des années 1990 a permis l'avènement de la pathologie numérique [2]. En effet, les scanners ont rendu possible la numérisation des lames de verre en créant des images de tissus à haute résolution (WSI). Un exemple de telles images est montré en Figure 1.

La numérisation des lames a permis de résoudre de nombreux problèmes liés à l'étude microscopique. D'abord, contrairement aux lames de verres, il est possible de conserver indéfiniment et sans détérioration les tissus. Par conséquent, l'archivage des images médicales permet de créer de grandes bases de données autour de pathologies particulières, ce qui facilite les travaux de recherches

concernant ces maladies. Ensuite, il devient possible de partager les WSI à des fins d'examen par les pairs pour obtenir plusieurs diagnostics à distance d'une même lame de verre. D'une manière différente, la virtualisation d'un microscope et le partage de WSI peuvent être utilisés à des fins de formations. Effectivement, pour former des spécialistes en pathologie, il n'est plus nécessaire d'avoir des lames de verre, ni même un microscope à disposition. Un microscope virtuel avec des WSI suffit, permettant ainsi un accès facilité à l'éducation. Ces avancées technologiques s'accompagnent de nouveaux enjeux liés à la taille des WSI. À cause de leur grande résolution, les WSI pèsent plusieurs gigaoctets. C'est pourquoi, il faut des systèmes informatiques adaptés pour stocker, traiter et partager les WSI.

1.2 Pathologie numérique & IA

Suite aux résultats encourageants qu'ont montrés les intelligences artificielles (IA) dans le domaine de la vision par ordinateur, ces méthodes ont été rapidement utilisées sur des images médicales. Dans le domaine de la pathologie numérique, les IA ont le potentiel d'apporter des solutions dans l'éducation, le contrôle qualité, les diagnostics cliniques, l'analyse d'images [15, 16, 22]. Cependant, face à ces nouvelles opportunités, les défis sont nombreux : le manque de données annotées ; la consommation élevée en ressources de calcul ; le manque de transparence et d'explicabilité des prédictions des IA ; les questions éthiques et légales.

Parmi les méthodes d'IA, on trouve notamment les méthodes d'apprentissage profonds. Les réseaux de neurones profonds ont rapidement été appliqués à diverses pathologies, dont le cancer du sein, qui est le plus meurtrier chez les femmes [5, 6].

Ainsi, des études ont proposé des méthodes de classification de patches de tumeur dans le cas du cancer du sein [3, 10, 11, 14, 25]. Ces travaux ont montré la supériorité des approches d'apprentissage profond comparées aux approches d'apprentissage machine.

1.3 Microscope virtuel & réalité augmentée

Poussés par l'évolution de la pathologie numérique, les microscopes virtuels sont apparus, permettant ainsi d'émuler les fonctions d'un microscope optique sur un ordinateur [1, 9]. En plus de s'affranchir d'un microscope, les microscopes virtuels permettent de zoomer et dézoomer dans les tissus sans-délai, contrairement aux microscopes optiques, qui ne proposent que quelques niveaux de grossissement et nécessitent une mise au point du tissu observé à chaque changement de niveau.

Malgré ces avantages, dans un contexte de travail redondant tel que le diagnostic de lame de verre, l'utilisation du microscope virtuel n'est pas adaptée, car elle entraîne une perte de temps lors de la numérisation de la lame de verre, puis lors de la phase de diagnostic [23]. Une approche basée sur des outils de réalité augmentée intégrée aux microscopes (ARM), où l'on affiche des informations sup-

plémentaires lors de l'étude microscopique semble mieux adaptée pour aider les pathologistes dans leur routine.

Lors d'une étude microscopique réalisée avec un microscope optique, le pathologiste analyse une lame de verre à la recherche de structures d'intérêts pour établir un diagnostic. La lame de verre est notamment étudiée à différents niveaux de grossissement. Pour changer de niveau de grossissement, les microscopes optiques disposent d'un nombre limité d'objectifs. Nous pensons que les niveaux de grossissement intermédiaires inaccessibles avec un microscope optique sont riches en caractéristiques et peuvent permettre de mieux comprendre les tissus aux niveaux de grossissement accessibles par un microscope.

En suivant les précédents travaux [12] sur la détection d'objets en temps-réel dans des WSI, nous proposons de :

- Classer en temps-réel des patches de tumeurs bénignes et malignes dans le cas du cancer du sein à l'aide de l'architecture Inception-ResNet-v2 [21].
- Utiliser les niveaux de grossissement intermédiaires d'une WSI afin d'améliorer les performances d'une méthode ARM qui utilise des niveaux de grossissement standard d'un microscope optique (20X, 40X, 100X).
- Comparer trois approches de prédiction de tumeur sur différents niveaux de grossissement : Entraîner des réseaux sur plusieurs niveaux ; Entraîner des réseaux sur chaque niveau ; Entraîner des réseaux sur un niveau intermédiaire.

Dans un souci de reproductibilité, les expérimentations ont été effectuées sur le dataset public BreakHis [20], et le code source est disponible ici ¹.

L'article est organisé de la manière suivante. D'abord, nous introduisons trois méthodes afin de traiter des images de plusieurs niveaux de grossissement dans le cas d'une application ARM dans la section 2.1. Ensuite, nous présentons le jeu de données public BreakHis [20] et les jeux de données créés à partir de BreakHis dans la section 2.2. Nous comparons les méthodes présentées en 2.1 dans la section 3, avant d'apporter une analyse sur les expérimentations réalisées en section 4.

2 Matériels et méthodes

2.1 Méthodes

Ces dernières années, l'évolution rapide des capacités de calcul des cartes graphiques (GPU) a largement contribué à l'apparition de nouvelles approches améliorant les performances globales des méthodes d'apprentissage profond permettant de les envisager plus sérieusement pour des applications temps réel tel que l'analyse microscopique en réalité augmentée [4, 19]. L'un des défis de l'utilisation

1. <https://github.com/RobinHCK/Breast-cancer-classification-methods-for-augmented-reality-microscopes>

du microscope à réalité augmentée (ARM) est les changements fréquents du niveau de grossissement lors des séances de diagnostic. En effet, changer le grossissement modifie les structures, les formes et les bordures des tissus observés. Pour pallier ce problème, nous envisageons deux approches naïves pour développer une méthode temps réel robuste donnant de bons résultats à différents grossissements.

La première approche consiste à entraîner un seul réseau sur plusieurs grossissements, afin qu’il généralise sa compréhension des structures à plusieurs niveaux. Cette méthode ne nécessite qu’un seul modèle à entraîner, cependant, les performances sont largement liées à la capacité de l’architecture à généraliser sur plusieurs grossissements.

La deuxième approche, consiste à entraîner un réseau par niveau. Dans ce cas, plusieurs modèles sont nécessaires, selon la plage de niveaux de grossissement étudiée.

Nous proposons ensuite une troisième approche où nous entraînons un réseau sur un niveau proche du niveau utilisé pour la phase de test.

Parmi les architectures qui ont montré les meilleurs résultats sur la tâche de classification de patches de tumeur bénigne et maligne dans le cas du cancer du sein, on trouve : Inception-v3, ResNet, GoogleNet, VGG. Inception-v3 semble montrer des résultats légèrement supérieurs aux autres méthodes bien que d’un jeu de données à l’autre cette différence peut s’avérer minime. L’architecture Inception-ResNet-v2, qui est une évolution de l’architecture Inception-v3, propose d’ajouter des connexions résiduelles aux modules Inception, afin d’accélérer l’entraînement et de surpasser les architectures Inception sans connexions résiduelles.

Nous comparons les trois approches en étudiant les résultats de différents modèles Inception-ResNet-v2, sur une large plage de niveaux de grossissement pour pouvoir sélectionner les meilleurs réseaux sur une plage donnée.

2.2 Jeux de données

BreakHis. Le jeu de données public BreakHis [20] a été utilisé car il contient de nombreuses images à différents niveaux de grossissement. Il est composé de 7 909 patches colorés en H&E (Hématoxyline et Éosine), de taille 700x460 pixels, à des niveaux de grossissement 40X, 100X, 200X et 400X. Ce jeu de données est destiné à la tâche de classification du cancer du sein. 2 480 patches de tumeurs bénignes et 5 429 patches de tumeurs malignes ont été prélevés sur 82 patients.

Création des niveaux de grossissement intermédiaires. Afin de couvrir une plus large plage de niveaux de grossissement, nous interpolons des patches de 20 niveaux de grossissement intermédiaires à partir du jeu de données d’origine.

Pour cela, chaque patch original est recadré selon l’Équation 1 où *taille recadrage* est la taille du patch à conserver lors du recadrage, *taille patch* est la taille du patch qui sera utilisé par le réseau de neurones profonds et *plus proche ni-*

veau est le niveau de grossissement inférieur le plus proche qui existe dans les patches du jeu de données d’origine.

$$\text{taille recadrage} = \frac{\text{taille patch} * \text{plus proche niveau}}{\text{niveau souhaité}} \quad (1)$$

Ensuite, nous choisissons l’interpolation bicubique [13], qui est un bon compromis entre qualité de l’interpolation et temps de traitement [17], pour dimensionner les patches à la dimension *taille patch*. Ainsi, on obtient un jeu de données contenant des patches de taille 350x230 pixels² pour chacun des 24 niveaux de grossissement de notre plage (de 20X à 400X). Nous créons un dernier dataset nommé allX à partir des patches 40X, 100X, 200X et 400X.

Enfin, nous appliquons des méthodes d’augmentation de données aux jeux de données pour améliorer la diversité des données et lutter contre le manque de données qui est un problème récurrent lors de l’utilisation de données médicales. Les patches subissent aléatoirement des rotations, des transformations horizontales et verticales.

3 Expérimentations

Dans la section 2.1, nous avons présenté trois approches robustes aux changements de niveaux de grossissement. Afin d’étudier et comparer ces approches, nous entraînons des réseaux Inception-ResNet-v2 sur des jeux de données construits à partir de BreakHis sur une large plage de niveaux de grossissement.

Avant de commencer l’entraînement, nous transférons les poids d’un modèle pré-entraîné sur ImageNet [7] (ImageNet est la plus grande base de données d’images, couramment utilisée pour les tâches de classification, contenant 14 197 122 images organisées en 21 841 classes) à notre modèle. Nous gelons les poids du modèle, avant de réentraîner les 5 dernières couches du réseau sur nos données. Cette méthode est une procédure standard pour affiner l’entraînement des modèles avec peu de données. Cela permet d’économiser du temps d’entraînement et de réduire le surapprentissage.

Nous entraînons les réseaux avec les mêmes paramètres et hyperparamètres pour comparer plus facilement les résultats : taux d’apprentissage = 0,01 ; *momentum* = 0,9 ; époques = 30 ; taille de *batch* = 64 ; optimiseur = *SGD* ; fonction de coût = *softmax* ; régularisation = *dropout*. De plus, nous utilisons *EarlyStopping* pour arrêter l’entraînement lorsqu’aucune amélioration n’a lieu sur la *loss* de validation au cours des 15 dernières époques pour éviter le surapprentissage. Afin de tirer pleinement parti des données dont nous disposons et de lutter contre le surapprentissage, nous effectuons une validation croisée 5 blocs ($K = 5$) pour chaque niveau de grossissement. Les patches sont répartis dans chaque bloc en respectant l’organisation

2. La taille des patches du dataset BreakHis est de 700x460 pixels. Afin d’obtenir des patches de niveau de grossissement plus faible, il faut découper et redimensionner les patches d’origine. De ce fait on perd une partie du contexte du patch d’origine.

donnée par BreakHis. Chaque bloc est utilisé pour tester une fois les résultats obtenus sur un réseau entraîné sur tous les autres blocs. Ainsi, chaque patch du jeu de données est classé une fois.

Nous effectuons tous les traitements informatiques sur des GPU fournies par le centre de calcul de l’Université de Strasbourg. Au total, à raison de 10 minutes d’entraînement par réseau, il a fallu environ 1 jour de calcul cumulé pour entraîner les 125 réseaux (validation croisée $K = 5 * 25$ jeux de données) que compte cette expérimentation.

Enfin, nous testons chaque modèle sur tous les niveaux de grossissement. Pour cette étape, nous choisissons d’utiliser la précision définie dans l’Équation 2 comme métrique d’évaluation puisque nos classes sont équilibrées.

$$\text{précision} = \frac{\text{patches bien classés}}{\text{nombre total de patches}} \quad (2)$$

Nous construisons une carte de chaleur illustrée en Figure 2 à partir des précisions obtenues par les réseaux.

En parallèle, nous mesurons le nombre de patches traités par seconde lors de l’étape de test sur un sous-échantillon de 10 réseaux. Ensuite, en posant l’hypothèse que ce calcul est linéaire, nous convertissons ce nombre de patches par seconde en nombre d’image d’une définition de 720p et 1080p testés par seconde (IPS) pour obtenir le Tableau 1. Ces mesures ont été réalisées en utilisant un processeur (CPU)³ puis deux GPU⁴.

TABLE 1 – Nombre d’IPS traitées par Inception-ResNet-v2. Le nombre moyen d’images par seconde (IPS) testées par un réseau selon le type de ressource utilisé (GPU ou CPU) et selon la définition d’image (720p ou 1080p).

Définition d’image	720p	1080p
CPU	1.3	0.6
GPU 1080	3.5	1.7
GPU 3080	7.3	3.2

4 Résultats et discussion

4.1 Temps réel

L’affichage d’informations en réalité augmentée sur un microscope nécessite une méthode en temps réel capable de suivre sans entrave les mouvements d’un pathologiste. Nous définissons la notion de temps réel en nous appuyant sur [8]. Ainsi, pour la tâche de classification de patch de tumeur, nous parlons de temps réel pour dire que notre méthode est capable de prédire le champ de vue observé sous un microscope plusieurs fois par seconde comme montré dans le Tableau 1. En effet, l’architecture Inception-ResNet-v2 traite en moyenne 7.3 IPS à une résolution de 720p avec un GPU, ce qui peut permettre de classer les images assez rapidement pour un résultat fluide.

Avec 55 873 736 paramètres et une profondeur de 572, Inception-ResNet-v2 fait partie des réseaux gourmands en ressource de calcul. Malgré sa taille, il est possible de réaliser des tâches en temps réel à l’aide d’une seule GPU de la dernière génération. Grâce aux récentes avancées en matière de capacité de calcul, il est à présent possible d’utiliser en temps réel des réseaux d’apprentissage profond en embarquant une GPU du commerce sur un microscope.

4.2 Carte de chaleur

Tout d’abord, il est évident que les réseaux entraînés sur un niveau de grossissement ont les meilleures précisions sur les niveaux de grossissement proches, comme le montre la Figure 2. Il apparaît clairement qu’un réseau sera généralement meilleur sur un niveau de grossissement proche de celui utilisé lors de l’étape d’apprentissage. Cependant, en regardant la diagonale centrale, on peut voir que la meilleure précision à un niveau de grossissement n’est pas toujours trouvée par le réseau entraîné sur ce niveau de grossissement. En d’autres termes, un léger zoom avant ou arrière pendant l’entraînement améliore les résultats du réseau. Il semble que le contexte proche contient des caractéristiques riches qui améliorent les résultats d’un réseau. Les résultats montrent également qu’à un niveau de grossissement donné, les réseaux entraînés à ce niveau et les niveaux proches obtiennent les meilleures performances. De plus, ces réseaux ne se trompent pas forcément sur les mêmes cas. Cela signifie que chaque niveau de grossissement ne fournit pas les mêmes caractéristiques.

La deuxième observation concerne le gradient de part et d’autre de la diagonale de cases les plus foncées. Plus le niveau de grossissement d’un jeu de données testé est éloigné du niveau de grossissement de l’ensemble de données utilisé pour l’entraînement, plus la précision est faible. Ceci semble confirmer l’intuition que le contexte proche contient les caractéristiques les plus utiles pour un réseau. Troisièmement, nous remarquons que plus le niveau de grossissement est élevé, moins bonnes seront les précisions des réseaux entraînés sur un seul niveau de grossissement. Avec un niveau de grossissement élevé, on perd le contexte spatial qui permettait de comprendre la structure d’une tumeur. De ce fait, observer des tissus à un niveau de grossissement élevé n’est pas pertinent dans la classification de tumeur dans le cas du cancer sein. On remarque que les niveaux de grossissement entre 20X et 80X sont plus riches en caractéristiques pour nos réseaux dans notre cas. En pratique, cela coïncide avec le comportement du pathologiste, qui passe plus de temps dans les niveaux de grossissement faibles [18].

Quatrièmement, notons que les meilleurs réseaux donnent des précisions proches de 80%. Ces performances sont plus faibles que l’état de l’art en matière de classification de patch de tumeur dans le cas du cancer du sein sur BreakHis [24]. Cela s’explique par la manière dont nous construisons nos jeux de données à partir de BreakHis. En effet, afin d’étudier une grande plage de niveaux de gros-

3. i5-11600KF 3.90GHz 3.91 GHz

4. NVIDIA GeForce GTX 1080 & RTX 3080

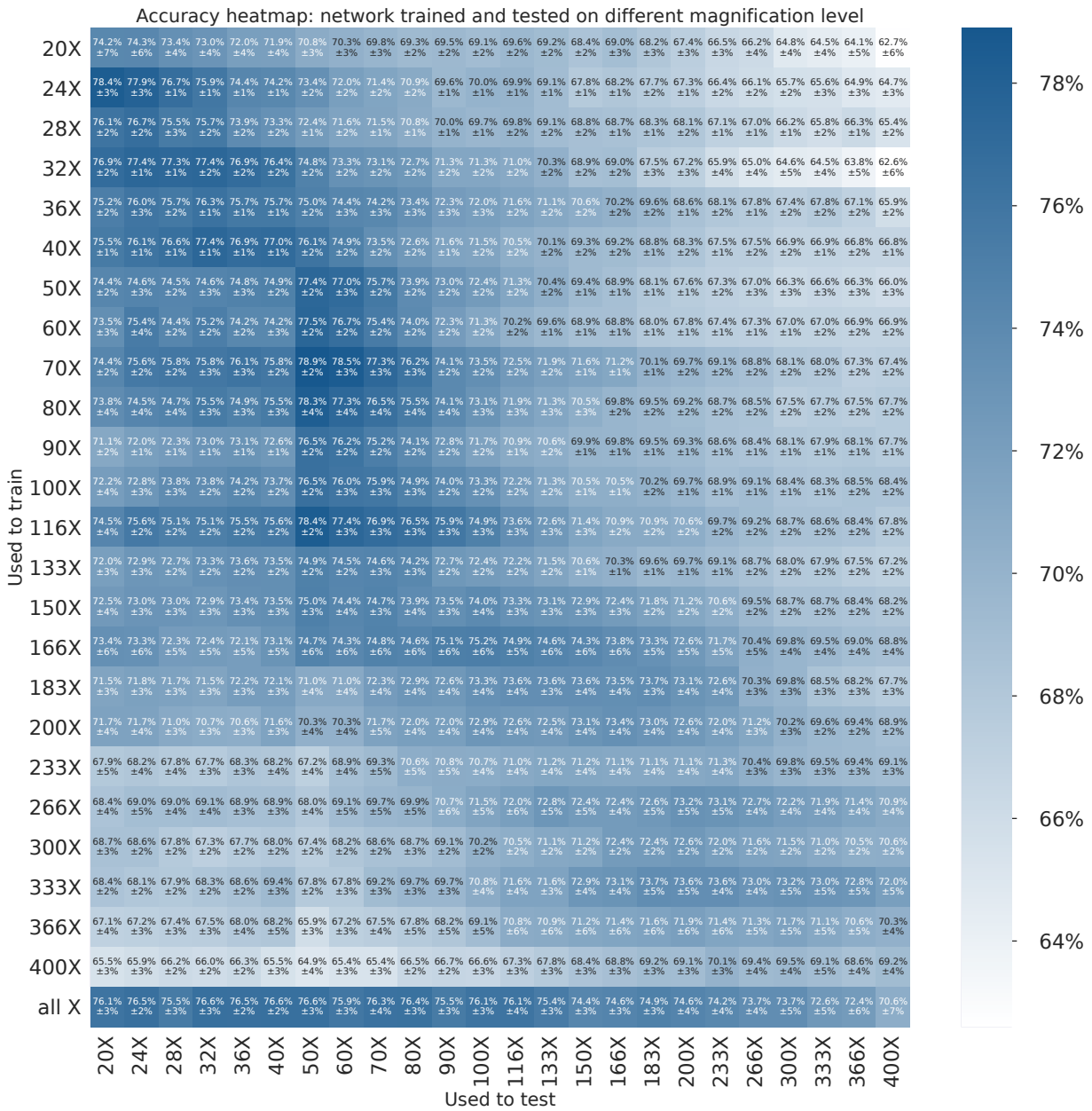


FIGURE 2 – Carte de chaleur des précisions des réseaux Inception-ResNet-v2. Indique les précisions (%) obtenues par des réseaux entraînés testés individuellement sur tous les jeux de données dans une plage de niveaux de grossissement de 20X à 400X. Plus un réseau obtient une précision élevée sur un jeu de données, plus la case est sombre, et inversement.

issement, nous construisons des patches de nouveaux niveaux, de ce fait une partie du contexte a été supprimé, ce qui impacte les performances des réseaux. On remarque également que dans le pire des cas, la plus faible précision obtenue est de 62,6%. Cela signifie que dans le cas du cancer du sein, les caractéristiques permettant de distinguer les tumeurs bénignes et malignes sont présentes à tous les niveaux de grossissement de la plage d'étude.

4.3 Comparaison des approches

À partir de la carte de chaleur de la Figure 2, nous pouvons comparer notre approche avec deux méthodes naïves comme rapporté dans le Tableau 2.

Premièrement, nous constatons que la première approche, où l'on entraîne les réseaux sur le niveau de grossissement utilisé pour le test, parvient à obtenir les meilleures précisions sur certains niveaux de la plage 20X à 70X.

Deuxièmement, les réseaux entraînés sur tous les niveaux de grossissement obtiennent des meilleures précisions sur la plage de 100X à 300X. Cette deuxième approche réussit à généraliser les caractéristiques apprises sur plusieurs niveaux, obtenant ainsi des meilleures précisions sur des niveaux de grossissement élevés. Cependant, cette méthode échoue à surpasser les réseaux spécialisés sur des niveaux de grossissement plus faibles (20X à 100X).

Troisièmement, nous constatons que l'approche proposée, où nous sélectionnons les meilleurs réseaux entraînés sur un niveau de grossissement quel qu'il soit, obtient les meilleures précisions pour la plage 20X à 90X. Un niveau de grossissement contient des caractéristiques qui permettent parfois d'obtenir de meilleurs résultats sur des niveaux de grossissement proches.

En moyenne, l'approche proposée (75,4%) surpasse ou équivaut les approches naïves (74% & 75,1%) sur la plage de niveaux de grossissement étudié avec l'architecture Inception-ResNet-v2 sur le jeu de données BreakHis. Notre approche montre de meilleurs résultats sur des niveaux de grossissement faibles (20X à 90X). À partir de 100X, l'approche allX devient meilleure. Au-delà d'un certain palier de niveau de grossissement, les caractéristiques des tumeurs présentent dans l'image perdent de leur pertinence. Cependant, les caractéristiques propres à un niveau de grossissement faible aident significativement la compréhension des caractéristiques sur des niveaux de grossissement élevé.

En observant les meilleurs réseaux par niveaux de grossissement montrés dans le Tableau 2, on observe que les réseaux 24X, 32X, 40X, 70X, 116X, 333X et allX permettent de couvrir toute la plage de niveau de grossissement étudié, tout en obtenant les meilleurs résultats.

Dans le cadre d'une méthode ARM, on souhaite afficher des informations supplémentaires sur le tissu observé sur les quelques niveaux de grossissement que propose le microscope optique. Prenons par exemple un microscope avec des niveaux de grossissement 20X, 40X et 100X. Afin d'obtenir les meilleurs résultats dans notre cas, il faut utili-

ser : le modèle 24X sur le niveau 20X ; le modèle 40X sur le niveau 40X ; le modèle allX sur le niveau 100X.

Il paraît intéressant d'utiliser des réseaux entraînés sur un seul niveau de grossissement pour les niveaux de grossissement les plus faibles jusqu'à 90X. Ensuite, à partir de 100X, il convient d'utiliser des réseaux entraînés sur plusieurs niveaux de grossissement.

Enfin, il ne faut pas oublier que ces résultats ont été obtenus avec l'architecture Inception-ResNet-v2 sur un jeu de données de cancer du sein. Selon la méthode et les données utilisées, les conclusions peuvent changer.

5 Conclusion

Ces dernières années, l'évolution rapide des capacités de calcul des cartes graphiques a largement contribué à l'apparition de nouvelles approches améliorant les performances globales des méthodes d'apprentissage profond. Ces avancées nous permettent d'envisager plus sérieusement les méthodes d'apprentissage profond pour les outils embarqués de microscope en temps réel. De telles applications peuvent être utilisées pour assister les pathologistes dans leur routine quotidienne.

Nous avons proposé une approche qui tire profit des niveaux de grossissement intermédiaires des WSI afin d'améliorer les performances d'une méthode ARM qui utilise des niveaux de grossissement standard d'un microscope optique.

Nous pensons que notre approche peut être applicable à d'autres organes, d'autres pathologies et tout autre application nécessitant un microscope optique à réalité augmentée. Les travaux futurs devraient se concentrer sur l'utilisation de méthodes d'apprentissage ensembliste pour tirer pleinement parti des caractéristiques de chaque niveau de grossissement. Enfin, il serait intéressant de voir des approches temps réel être mises en oeuvre sur différentes tâches dans diverses applications médicales et autres. Cela permettra de prendre conscience des défis liés à l'utilisation de méthodes temps réel d'apprentissage profond en condition réelle.

6 Remerciements

Les auteurs tiennent à remercier le Centre de Calcul Haute Performance de l'Université de Strasbourg pour avoir soutenu ce travail en fournissant un soutien scientifique et un accès aux ressources informatiques. Une partie des ressources informatiques ont été financées par le projet Equip@Meso (Programme Investissements d'Avenir) et le CPER Alsacalcul / Big Data.

Références

- [1] A. Afework, M. Beynon, F. Bustamante, S. Cho, A. Demarzo, R. Ferreira, R. Miller, M. Silberman, J. Saltz, A. Sussman, Digital dynamic telepathology—the Virtual Microscope, *Proceedings of the AMIA Symposium*, pp. 912, 1998.

TABLE 2 – Comparaison des approches. Comparaison des précisions (%) de trois approches de prédiction de tumeur sur différents niveaux de grossissement : Entraîner des réseaux sur plusieurs niveaux ; Entraîner des réseaux sur chaque niveau ; Entraîner des réseaux sur un niveau intermédiaire.

Niveaux de grossissement testés	Réseaux entraînés & testés sur un niveau	Réseaux entraînés sur 40X, 100X, 200X, 400X (allX)	Réseaux entraînés sur un niveau intermédiaire (Méthode proposée)	Le meilleur réseau
20X	0.742	0.761	0.784	24X
24X	0.779	0.765	0.779	24X
28X	0.755	0.755	0.773	32X
32X	0.774	0.766	0.774	32-40X
36X	0.757	0.765	0.769	32-40X
40X	0.770	0.766	0.770	40X
50X	0.774	0.766	0.789	70X
60X	0.767	0.759	0.785	70X
70X	0.773	0.763	0.773	70X
80X	0.755	0.764	0.765	116X
90X	0.728	0.755	0.759	116X
100X	0.733	0.761	0.752	allX
116X	0.736	0.761	0.749	allX
133X	0.715	0.754	0.746	allX
150X	0.729	0.744	0.743	allX
166X	0.738	0.746	0.738	allX
183X	0.737	0.749	0.737	allX
200X	0.726	0.746	0.736	allX
233X	0.713	0.742	0.736	allX
266X	0.727	0.737	0.730	allX
300X	0.715	0.737	0.732	allX
333X	0.730	0.726	0.730	333X
366X	0.706	0.724	0.728	333X
400X	0.692	0.706	0.720	333X
Moyenne	0.740	0.751	0.754	

- [2] S. Al-Janabi, A. Huisman, P. Van Diest, Digital pathology : current status and future perspectives, *Histopathology*, Vol. 61, pp. 1-9, 2012.
- [3] H. Chen, C. Li, X. Li, M. Rahaman, W. Hu, Y. Li, W. Liu, C. Sun, H. Sun, X. Huang, IL-MCAM : An interactive learning and multi-channel attention mechanism-based weakly supervised colorectal histopathology image classification approach, *Computers in Biology and Medicine*, pp. 105265, 2022.
- [4] P-H. Chen, K. Gadepalli, R. MacDonald, Y. Liu, K. Nagpal, T. Kohlberger, J. Dean, G. Corrado, J. Hipp, M. Stumpe, Microscope 2.0 : an augmented reality microscope with real-time artificial intelligence integration, *arXiv preprint arXiv :1812.00825*, 2018.
- [5] G. Chugh, S. Kumar, N. Singh, Survey on machine learning and deep learning applications in breast cancer diagnosis, *Cognitive Computation*, Vol. 13, pp. 1451-1470, 2021.
- [6] T. Debelee, F. Schwenker, A. Ibenthal, D. Yohannes, Survey of deep learning in breast cancer image analysis, *Evolving Systems*, Vol. 11, pp. 143-163, 2020.
- [7] J. Deng, W. Dong, R. Socher, L-J. Li, K. Li, L. Fei-Fei, Imagenet : A large-scale hierarchical image database, *2009 IEEE conference on computer vision and pattern recognition*, pp. 248-255, 2009.
- [8] R. Dodhiawala, N. Sridharan, P. Raulefs, C. Pickering, Real-Time AI Systems : A Definition and An Architecture, *IJCAI*, pp. 256-264, 1989.
- [9] R. Ferreira, B. Moon, J. Humphries, A. Sussman, J. Saltz, R. Miller, A. Demarzo, The virtual microscope, *Proceedings of the AMIA Annual Fall Symposium*, pp. 449, 1997.
- [10] A. Golatkar, D. Anand, A. Sethi, Classification of breast cancer histology using deep learning, *International conference image analysis and recognition*, pp. 837-844, 2018.
- [11] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, S. Li, Breast cancer multi-classification from histopathological images with structured deep learning model, *Scientific reports*, Vol. 7, pp. 1-10, 2017.
- [12] R. Heckenauer, J. Weber, C. Wemmert, F. Feuerhake, M. Hassenforder, P-A. Muller, G. Forestier, Real-time detection of glomeruli in renal pathology, *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 350-355, 2020.
- [13] R. Keys, Cubic convolution interpolation for digital image processing, *IEEE transactions on acoustics, speech, and signal processing*, Vol. 29, pp. 1153-1160, 1981.
- [14] S. Khan, N. Islam, Z. Jan, I. Din, J. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, *Pattern Recognition Letters*, Vol. 125, pp. 1-6, 2019.
- [15] A. Madabhushi, G. Lee, Image analysis and machine learning in digital pathology : Challenges and opportunities, , Vol. 33, pp. 170-175, 2016.
- [16] M. Niazi, A. Parwani, M. Gurcan, Digital pathology and artificial intelligence, *The lancet oncology*, Vol. 20, pp. 253-261, 2019.
- [17] P. Parsania, P. Virparia, A comparative analysis of image interpolation algorithms, *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 5, pp. 29-34, 2016.
- [18] R. Randell, R. Ruddle, C. Mello-Thoms, R. Thomas, P. Quirke, D. Treanor, Virtual reality microscope versus conventional microscope regarding time to diagnosis : an experimental study, *Histopathology*, Vol. 62, pp. 351-358, 2013.
- [19] N. Razavian, Augmented reality microscopes for cancer histopathology, *Nature medicine*, Vol. 25, pp. 1334-1336, 2019.
- [20] F. Spanhol, L. Oliveira, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification, *IEEE transactions on biomedical engineering*, Vol. 63, pp. 1455-1462, 2015.
- [21] C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [22] H. Tizhoosh, L. Pantanowitz, Artificial intelligence and digital pathology : challenges and opportunities, *Journal of pathology informatics*, Vol. 9, 2018.
- [23] D. Treanor, N. Jordan-Owers, J. Hodrien, J. Wood, P. Quirke, R. Ruddle, Virtual reality Powerwall versus conventional microscope for viewing pathology slides : an experimental comparison, *Histopathology*, Vol. 55, pp. 294-300, 2009.
- [24] A. Tufail, Y-K. Ma, M. Kaabar, F. Martínez, AR Junejo, I. Ullah, R. Khan, Deep learning in cancer diagnosis and prognosis prediction : a minireview on challenges, recent trends, and future directions, *Computational and Mathematical Methods in Medicine*, 2021.
- [25] D. Wang, A. Khosla, R. Gargeya, H. Irshad, A. Beck, Deep learning for identifying metastatic breast cancer, *arXiv preprint arXiv :1606.05718*, 2016.