FRUGAL-WSI : comprendre la compression des lames numériques pour une IA frugale et robuste : application à la prédiction de la réponse à l'immunothérapie

Porteur : **Hamid Ladjal** (LIRIS, UMR 5205 CNRS, Université Claude Bernard Lyon 1)
Partenaires : Nazim Benzerdjeb, pathologiste, Hospices Civils de Lyon et CICLY – Centre pour l'Innovation en Cancérologie de Lyon (UR 3738), Université Claude Bernard Lyon 1
Laboratoire : LIRIS-Université Claude Bernard Lyon 1
Composante : Département INFO

Nature du financement demandé : Stage de M2 (6 mois)
(Pour les bourses d'excellence prendre le template dédié en anglais)
Dates : février – juillet 2027 (présentation des résultats aux journées GI-EIF des 1er-2 juillet 2027)

Résumé :
Les lames entières numérisées (*whole slide images*, WSI) atteignent 1 à 3 Go par lame, soit plusieurs téraoctets pour une cohorte de 2 000 lames. Leur stockage, leur transfert et leur traitement sur GPU entraînent un coût énergétique et carbone significatif. Chaque scanner emploie un format propriétaire (SVS, MRXS, NDPI, iSyntax, DICOM) et un schéma de compression avec perte dont les paramètres réels sont rarement vérifiés. Sur nos acquisitions Leica GT450, le taux de compression effectif (26:1) excède ainsi largement le taux nominal (10:1).
Cette étude poursuit deux objectifs. Le premier est de caractériser la compression des WSI et d'en optimiser les réglages, afin de réduire la volumétrie et l'empreinte carbone des pipelines d'intelligence artificielle en pathologie. Le second est de quantifier l'effet de cette compression sur les prédictions des modèles et sur leur interprétation. Le protocole porte sur une cohorte de 2 000 lames de patients traités par immunothérapie. Trois familles de méthodes seront comparées : des codecs classiques (JPEG, JPEG 2000, JPEG XL), une compression neuronale apprise et une compression multi-échelle guidée par graphe. Les critères d'évaluation incluent le volume de données, l'énergie consommée mesurée par CodeCarbon, la fidélité des représentations extraites par les modèles de fondation les plus récents (UNI2, Virchow2, H-optimus-1, CONCH/TITAN) et la performance de prédiction de la réponse thérapeutique. Le résultat attendu est une recommandation quantifiée du compromis **compression – performance – carbone**, applicable à l'archivage et à une IA frugale en pathologie numérique.

Sujet développé :
Contexte. La pathologie numérique génère des volumes considérables : un fichier DICOM GT450 de notre laboratoire (114 484 × 68 070 px, 119 168 tuiles) pèse 846 Mo malgré une compression JPEG déjà agressive (3,8 % des données conservées). Les analyses par IA reposent aujourd'hui sur des modèles de fondation (MF) qui extraient des descripteurs de dizaines de milliers de tuiles par lame, puis sur un apprentissage multi-instances (MIL). Or ces modèles sont entraînés sur des lames hétérogènes en format et en compression : des travaux récents montrent que la compression avec perte biaise les modèles et que leurs performances chutent lorsque les schémas de compression

diffèrent entre entraînement et inférence [1,2]. L'effet de la compression propriétaire sur une prédiction clinique complexe, comme la réponse à l'immunothérapie, reste inconnu.

Objectif 1 – Comprendre la compression pour une IA frugale. (i) Caractériser finement la compression réelle des formats propriétaires (codec, qualité, sous-échantillonnage chromatique, taux mesuré vs déclaré). (ii) Comparer, à iso-débit (bits/pixel), codecs classiques (JPEG, JPEG 2000/HTJ2K, JPEG XL), reproduction de la représentation neuronale implicite CINR [3], compression apprise affinée sur nos tuiles (CLERIC-like [4]) et suppression du fond [5]. (iii) Proposer une compression multi-échelle guidée par graphe : chaque lame est représentée par un graphe dont les nœuds sont les tuiles à 5×, 10× et 20× reliées par des arêtes spatiales et inter-échelles, à la manière du MS-RGCN [6] ; l'importance des nœuds (attention/relevance) sert à allouer le débit, en conservant la haute résolution uniquement là où elle est informative, voire à n'archiver que le graphe d'embeddings. (iv) Mesurer pour chaque stratégie le gain de stockage, le temps de calcul et l'énergie/CO_2eq (CodeCarbon) de toute la chaîne stockage – extraction de descripteurs – entraînement.

Objectif 2 – Impact sur la prédiction et la compréhension des modèles. La cohorte de 2 000 lames de patients traités par inhibiteurs de points de contrôle immunitaires sera recodée selon les différents formats/codecs et niveaux de compression. Pour chaque condition : (a) fidélité des descripteurs des MF récents (UNI2, Virchow2, H-optimus-1, CONCH v1.5 / TITAN) par similarité cosinus et CKA par rapport à la lame source, métrique qui prédit la performance aval [2] ; (b) performance de prédiction de la réponse (ABMIL, TransMIL et MS-RGCN ; AUC, calibration, validation croisée) avec scénarios croisés « entraînement compression A / test compression B » ; (c) stabilité de l'interprétation : cartes d'attention, régions et types cellulaires (lymphocytes infiltrants, stroma, tumeur) retenus par le modèle. On identifiera ainsi le « point de rupture » au-delà duquel la compression altère la décision et les MF les plus robustes.

Rôle du stagiaire et calendrier. M1-2 : prise en main du pipeline (OpenSlide, TRIDENT/CLAM, P2CHPD), audit des formats. M2-3 : recodage de la cohorte, extraction des descripteurs MF et suivi carbone. M3-5 : compression multi-échelle par graphe et modèles MIL, analyses de robustesse et d'interprétabilité. M6 : synthèse, rapport et article.

Encadrement : laboratoire en informatique et apprentissage automatique LIRIS (porteur) et Co encadrement/collaboration avec un médecin pathologiste ; données pseudonymisées, hébergées sur les infrastructures HCL/Lyon 1.

Retombées. Recommandations chiffrées compression/performance/carbone pour l'archivage hospitalier et l'IA en pathologie ; un outil open source de compression multi-échelle ; une publication (MICCAI/MIDL ou J Pathol Inform) ; un socle pour un projet ANR/Europe sur l'IA médicale frugale, au cœur des enjeux énergie de la GI-EIF.

Références.

[1] Fischer M. et al., MOVI-MICCAI 2022, LNCS 13578:73-83.
[2] Fischer M. et al., J Pathol Inform 2025;17:100421.
[3] Lee et al., CINR, MICCAI 2024.
[4] Lee et al., CLERIC, arXiv:2503.23862, 2025.
[5] Codipilly et al., Mayo Clin Proc Digit Health 2023.
[6] Bazargani R. et al., Multi-scale relational graph convolutional network for MIL in histopathology images, Med Image Anal 2024.