



Centre d'Etudes Doctorales : Sciences et Techniques et Sciences Médicales

Avis de Soutenance

THESE DE DOCTORAT

Présentée par

Monsieur ABDELWAHID OUBAALLA

Discipline : Informatique
Spécialité : Informatique

Sujet de la thèse

Approches intelligentes pour la classification automatique des données fondées sur le machine Learning et le Deep Learning

Formation Doctorale " Sciences de l'Ingénieur, Sciences Physiques, Mathématiques et informatique"

Thèse présentée et soutenue **le samedi 23 mai 2026 à 15h30** à l'Ecole Nationale des Sciences Appliquées de Fès, devant le jury composé de :

NOM ET PRÉNOM	TITRE	ÉTABLISSEMENT	
ADIL KENZI	PES	Ecole Nationale des Sciences Appliquées de Fès	Président
RAFIK TAOUIL	PES	Faculté des Sciences de Tétouan	Rapporteur
ABDELHAFID MESSAOUDI	MCH	Ecole Supérieure de Technologie d'Oujda	Rapporteur
ADIL JEGHAL	MCH	Ecole Nationale des Sciences Appliquées de Fès	Rapporteur
ABDESLAM EL AKKAD	PES	Centre Régional des métiers de l'Education et de la Formation de Fès	Examineur
HICHAM EL MOUBTAHIJ	MCH	Ecole Supérieure de Technologie d'Agadir	Examineur
ABDELLATIF EZZOUHAIRI	PES	Ecole Nationale des Sciences Appliquées de Fès	Examineur
ABDELHAK BOULAALAM	PES	Ecole Nationale des Sciences Appliquées de Fès	Examineur
KHALID HADDOUCH	PES	Ecole Nationale des Sciences Appliquées de Fès	Examineur
NABIL EL AKKAD	MCH	Ecole Nationale des Sciences Appliquées de Fès	Directeur de Thèse

Laboratoire de recherche : Sciences appliquées et technologies émergentes
Etablissement : Ecole Nationale des Sciences Appliquées de Fès



Centre d'Etudes Doctorales : Sciences et Techniques et Sciences Médicales

Résumé de la thèse

L'intégration de l'intelligence artificielle dans le domaine de l'oncologie clinique représente une évolution majeure pour la détection précoce du cancer. Les techniques de vision par ordinateur constituent désormais un outil d'assistance cognitive essentiel au diagnostic histopathologique de précision. Dans un contexte marqué par une incidence croissante des cancers à l'échelle mondiale, caractérisée par environ 20 millions de nouveaux cas et 9,7 millions de décès recensés en 2022 selon les estimations du Centre international de recherche sur le cancer, la pathologie numérique apparaît comme une réponse stratégique aux limitations inhérentes à l'examen microscopique conventionnel. Ce dernier, bien qu'il demeure la référence diagnostique, est affecté par une variabilité inter-observateur significative (estimée entre 20 et 30 % selon les études) ainsi que par des délais de rendu pouvant compromettre sérieusement le pronostic des patients. Au Maroc, où l'on recense plus de 63 000 nouveaux cas annuels, ces contraintes sont exacerbées par une pénurie marquée de pathologistes spécialisés, accentuant ainsi les disparités territoriales et socio-économiques d'accès aux soins de qualité.

C'est dans ce cadre que s'inscrit la présente thèse, qui vise à concevoir et à optimiser des approches d'apprentissage profond pour la classification automatisée des images histopathologiques des cancers du poumon et du côlon, en s'appuyant sur le jeu de données LC25000. Deux problématiques principales ont structuré les travaux : l'exploitation optimale de modèles pré-entraînés dans un contexte de données médicales annotées limitées, et la réduction de la variabilité chromatique caractéristique des colorations hématoxyline-éosine (H&E).

La première contribution consiste en une étude comparative systématique portant sur douze configurations issues de quatre familles d'architectures de réseaux convolutifs pré-entraînés (VGG-16, ResNet, DenseNet, NASNet). L'analyse s'est focalisée sur l'influence de la profondeur du fine-tuning, c'est-à-dire sur le nombre de couches réentraînées. Les résultats démontrent qu'une adaptation intermédiaire des couches permet d'obtenir le meilleur équilibre entre performance et généralisation, conduisant à une exactitude globale (accuracy) de 99,4 % sur le jeu de données LC25000. Ces observations confirment la pertinence de l'apprentissage par transfert (transfer learning) comme stratégie efficace et économe en ressources computationnelles dans le domaine de la pathologie numérique.



Centre d'Etudes Doctorales : Sciences et Techniques et Sciences Médicales

La seconde contribution propose une méthodologie hybride novatrice combinant un prétraitement non supervisé, basé sur le modèle k-means et la correction gamma, et une stratégie d'apprentissage par ensemble (ensemble learning) fusionnant les probabilités issues de VGG16 et DenseNet via une moyenne pondérée. Ce pipeline permet de renforcer les marqueurs morphologiques discriminants et d'atteindre une exactitude globale de 99,96 %, un F1-score moyen de 0,9996 et une confusion marginale entre les classes les plus proches sur le plan morphologique. Au-delà des performances numériques élevées, l'intérêt majeur réside dans la robustesse accrue face aux variations de coloration et dans la complexité computationnelle relativement modérée, compatible avec un déploiement dans des environnements à ressources limitées.

L'évaluation des approches proposées a été conduite de manière rigoureuse sur des partitions train/validation/test du jeu de données LC25000, en mobilisant un ensemble complet de métriques classiques et complémentaires telles que l'exactitude (accuracy), la précision, le rappel, le F1-score et la matrice de confusion, ainsi que des analyses qualitatives comme la visualisation des erreurs restantes. Les performances obtenues se situent parmi les meilleures rapportées sur ce jeu de données, tout en révélant des limites structurelles : dépendance à un corpus relativement homogène, absence de validation externe sur des lames réelles issues de laboratoires marocains, et nécessité d'améliorer l'explicabilité des décisions algorithmiques.

En conclusion, ces travaux démontrent qu'il est possible d'atteindre une assistance diagnostique de très haute fiabilité au moyen de méthodes pragmatiques, reproductibles et économes en ressources. L'objectif final demeure d'ordre clinique et sociétal : réduire les délais de rendu diagnostique, atténuer les inégalités d'accès au diagnostic précoce et soutenir les pathologistes dans leur mission fondamentale.

Mots clés : Intelligence artificielle ; Apprentissage profond ; Vision par ordinateur ; Pathologie numérique ; Classification d'images histopathologiques ; Apprentissage par transfert ; Apprentissage par ensemble ; Cancer du poumon et du côlon.