



AVIS DE SOUTENANCE DE THESE

Le Doyen de la Faculté des Sciences Dhar El Mahraz –Fès – annonce que

Mr **TAGNAMAS Jaouad**
Soutiendra : le **Samedi 27/12/2025 à 10H00**
Lieu : **FSDM – Département de Géologie**

Une thèse intitulée :

Exploring Hybrid Deep Learning Models for Medical Image Segmentation: Lightweight and Attention-based Approaches

En vue d'obtenir le **Doctorat**

FD : Sciences et Techniques
Spécialité : Informatique

Devant le jury composé comme suit :

Nom et prénom	Etablissement	Grade	Qualité
TAIRI Hamid	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Président
DAROUICHI Aziz	Faculté des Sciences et Techniques, UCA, Marrakech	MCH	Rapporteur
FARHAOUI Youssef	Faculté des Sciences et Techniques, UMI, Errachidia	PES	Rapporteur
RIFFI Jamal	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Rapporteur
EL ALLAOUI Ahmad	Faculté des Sciences et Techniques, UMI, Errachidia	PES	Examineur
EL FAZAZY Khalid	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Examineur
RAMADAN Hiba	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Co-directrice de thèse
YAHYAOUY Ali	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Directeur de thèse



Résumé :

La segmentation d'images médicales est une tâche fondamentale en vision par ordinateur, visant à automatiser et faciliter la délimitation précise des structures anatomiques et des régions pathologiques dans les images médicales. Les architectures d'apprentissage profond se sont révélées être parmi les techniques de segmentation d'images médicales les plus efficaces, particulièrement lorsqu'elles sont appliquées à de grands jeux de données médicaux avec diverses modalités d'imagerie. L'objectif de cette thèse est de développer des architectures d'apprentissage efficaces, légères et robustes pour la segmentation d'images médicales en exploitant des approches hybrides qui combinent les réseaux de neurones convolutifs (CNN), les Vision Transformers (ViTs) et les mécanismes d'attention. La première approche proposée introduit un cadre multi-tâches qui effectue simultanément la segmentation et la classification des tumeurs mammaires dans les images échographiques. Cette méthode emploie une architecture à double encodeur comprenant EfficientNetV2 et des encodeurs Vision Transformer adaptés, intégrés par un module novateur de Fusion d'Attention de Canal (CAF) pour capturer efficacement les informations spatiales locales et les caractéristiques sémantiques de haut niveau. Le modèle incorpore uniquement MLP-Mixer pour la classification tumorale, marquant sa première application dans l'analyse de lésions échographiques mammaires. Les résultats obtenus ont démontré 83.42% de coefficient de Dice pour la segmentation et 86% de précision pour la classification, prouvant une performance compétitive comparée aux méthodes de l'état de l'art. Pour répondre aux contraintes computationnelles du déploiement clinique, la seconde approche présente SCA-InceptionUNeXt, une architecture réseau légère qui incorpore un mécanisme d'Attention Spatiale-Canal (SCA) avec des blocs InceptionNeXt modifiés. Ce concept unifie les processus d'attention spatiale et de canal en un module cohésif, permettant des interactions de caractéristiques cross-dimensionnelles tout en réduisant significativement la complexité du modèle grâce aux convolutions séparables en profondeur et ponctuelles. Le SCA-InceptionUNeXt a atteint des résultats remarquables avec 81.66% de coefficient de Dice sur le jeu de données BUSI tout en utilisant 26.11 millions de paramètres de moins que l'U-Net conventionnel, établissant un équilibre optimal entre précision de segmentation et efficacité computationnelle. La troisième approche explore l'intégration des architectures MLP-Mixer avec des mécanismes d'attention de canal au sein du modèle U-Net établi. Cette investigation démontre la versatilité et l'efficacité des approches basées sur MLP au-delà des tâches de classification traditionnelles, atteignant 81.45% de coefficient de Dice et établissant de nouveaux standards pour la segmentation d'images médicales basée sur MLP. Les évaluations complètes sur plusieurs jeux de données d'imagerie médicale, incluant les images échographiques mammaires (BUSI), les régions de polypes dans les séquences de coloscopie, les lésions cutanées dans les images dermoscopiques, et multiples organes dans les scanners abdominaux CT, démontrent constamment la supériorité des méthodes proposées comparées aux approches existantes de l'état de l'art. Les architectures développées exhibent d'excellentes capacités de généralisation à travers diverses modalités d'imagerie et structures anatomiques, fournissant des solutions pratiques qui répondent aux exigences cliniques réelles tout en maintenant une haute précision de segmentation et une efficacité computationnelle.

Mots clés :

Segmentation d'Images Médicales, Mécanismes d'Attention, CNN-Transformer, MLP-Mixer, Réseaux Légers, Apprentissage profond.



EXPLORING HYBRID DEEP LEARNING MODELS FOR MEDICAL IMAGE SEGMENTATION: LIGHTWEIGHT AND ATTENTION-BASED APPROACHES

Abstract :

Medical image segmentation is a fundamental task in computer vision, aiming to automate and facilitate the precise delineation of anatomical structures and pathological regions in medical images. Deep learning architectures have proven to be among the most effective medical image segmentation techniques, particularly when applied to large medical datasets with diverse imaging modalities. The objective of this thesis is to develop efficient, lightweight, and robust learning architectures for medical image segmentation by leveraging hybrid approaches that combine convolutional neural networks (CNNs), Vision Transformers (ViTs), and attention mechanisms. The first proposed approach introduces a multi-task framework that simultaneously performs segmentation and classification of breast tumors in ultrasound images. This method employs a dual-encoder architecture featuring EfficientNetV2 and adapted Vision Transformer encoders, integrated through a novel Channel Attention Fusion (CAF) module to effectively capture both local spatial information and high-level semantic features. The framework uniquely incorporates MLP-Mixer for tumor classification, marking its first application in breast ultrasound lesion analysis. The achieved results demonstrated 83.42% Dice coefficient for segmentation and 86% accuracy for classification, proving competitive performance compared to state-of-the-art methods. To address computational constraints in clinical deployment, the second approach presents SCA-InceptionUNeXt, a lightweight network architecture that incorporates a novel Spatial-aware Channel Attention (SCA) mechanism with modified InceptionNeXt blocks. This design unifies spatial and channel attention processes into a cohesive module, enabling cross-dimensional feature interactions while significantly reducing model complexity through depthwise and pointwise separable convolutions. The SCA-InceptionUNeXt achieved remarkable results with 81.66% Dice coefficient on the BUSI dataset while using 26.11 million fewer parameters than conventional U-Net, establishing an optimal balance between segmentation accuracy and computational efficiency. The third approach explores the integration of MLP-Mixer architectures with channel attention mechanisms within the established U-Net framework. This investigation demonstrates the versatility and effectiveness of MLP-based approaches beyond traditional classification tasks, achieving 81.45% Dice coefficient and establishing new benchmarks for MLP-based medical image segmentation. Comprehensive evaluations across multiple medical imaging datasets, including breast ultrasound images (BUSI), polyp regions in colonoscopy sequences, skin lesions in dermoscopy images, and multiple organs in abdominal CT scans, consistently demonstrate the superiority of the proposed methods compared to existing state-of-the-art approaches. The developed architectures exhibit excellent generalization capabilities across diverse imaging modalities and anatomical structures, providing practical solutions that address real-world clinical requirements while maintaining high segmentation accuracy and computational efficiency.

Key Words :

Medical Image Segmentation, Attention Mechanisms, CNN-Transformer, MLP-Mixer, Lightweight Networks, Deep Learning.