Laboratoire de Mathématiques et Informatique pour la Complexité et les Systèmes
MICS

Présente

L'AVIS DE SOUTENANCE

De **Monsieur Othmane Laousy**
Laboratoire MICS, CentraleSupélec, Université Paris Saclay, soutiendra publiquement ses travaux de thèse de doctorat intitulés :

"Deep learning methods for localization, segmentation and robustness in medical imaging."

Sous la Direction de Madame Marie-Pierre Revel et Madame Maria Vakalopoulou.

Le 9 juillet 2024 à 15h00

À l'école CentraleSupélec, en **amphi III** - Bâtiment Eiffel.

Membres du jury :

Jose DOLZ, Rapporteur & Examinateur, ETS Montréal
Prateek PRASANNA, Rapporteur & Examinateur, Stony Brook University
Isabelle BLOCH, Examinatrice, Sorbonne Université, CNRS
Laure FOURNIER, Examinatrice, Université Paris Cité
Marco LORENZI, Examinateur, INRIA Epione

Résumé :

Ces dernières années, l'apprentissage profond a considérablement progressé en médecine, notamment dans l'imagerie médicale, introduisant des défis significatifs demandant une attention particulière.
Un défi majeur est la nécessité de disposer de bases de données étendues et diversifiées pour l'entraînement de modèles d'Intelligence Artificielle (IA) de plus en plus sophistiqués. Une autre préoccupation est la vulnérabilité des modèles d'IA aux attaques adversariales, où de légères

modifications des données d'entrée peuvent perturber considérablement leur fiabilité, posant des risques dans les applications médicales.
Cette thèse aborde ces défis, offrant des perspectives pour atténuer la rareté des données et améliorer la fiabilité des modèles d'IA face aux attaques adversariales. Elle explore l'application de l'apprentissage par renforcement profond pour résoudre les problèmes de localisation et de segmentation en imagerie médicale. En améliorant la prise de décision grâce au 'Deep Q-learning' et en utilisant l'optimisation de politique proximale, nous avons mis au point une méthode de quantification de la sarcopénie sur des scans CT. Cette approche a été validée en milieu clinique pour l'évaluation de la sarcopénie chez des patients atteints de cancer en état critique.
Le deuxième axe de cette thèse porte sur l'amélioration de la sécurité des modèles d'IA en incorporant une robustesse certifiée contre les attaques adversariales. En tirant parti des modèles de diffusion et du lissage aléatoire, nous développons des méthodes qui garantissent une robustesse certifiée contre les manipulations adverses. Ces méthodes offrent non seulement une protection contre les violations de sécurité potentielles, mais instaurent également une plus grande confiance dans le déploiement des systèmes de santé pilotés par l'IA.

Abstract :

Recent advancements in deep learning and medicine, particularly in medical imaging, have introduced significant challenges requiring careful consideration. One major challenge is the need for extensive, diverse datasets to train increasingly sophisticated models, which are essential for accurately representing varied medical scenarios. Another concern is the vulnerability of AI models to adversarial attacks, where minor input changes can significantly disrupt model reliability, posing risks in real-world medical applications.
This thesis addresses these challenges, proposing solutions for data scarcity and enhancing model trustworthiness against adversarial attacks. It explores the application of deep reinforcement learning to medical image analysis, specifically to solve localization and segmentation problems. By optimizing decision-making through Deep Q-learning and using proximal policy optimization, we devise a method for sarcopenia quantification on CT scans. This approach was validated by assessing sarcopenia in critically ill solid cancer patients.
Additionally, the thesis focuses on enhancing AI model safety by incorporating certified adversarial robustness. To mitigate adversarial vulnerability, we explore and develop approaches that guarantee a level of robustness against adversarial manipulations by leveraging the power of denoising diffusion models and randomized smoothing. These methods not only provide a safeguard against potential security breaches but also instill greater confidence in the deployment of AI-driven healthcare systems.
Overall, this research contributes to the development of deep learning in medicine by tackling data scarcity and adversarial threats, aiming for a more resilient and trustworthy future in medical imaging.