

Résumé

Dans un contexte d'économie budgétaire mais aussi de pénurie de médecins et de paramédicaux, les professionnels de santé ont besoin d'outils efficaces, leur permettant d'offrir un service de qualité accessible pour tous. Les diagnostics assistés par ordinateur émergent comme des leviers technologiques essentiels. À l'aube de cette quatrième révolution industrielle, nous assistons à la naissance de nouveaux algorithmes basés principalement sur l'apprentissage profond offrant des solutions prometteuses pour renforcer l'efficacité des professionnels de la santé et répondre aux besoins croissants de la population.

Dans ce travail de recherche, nous proposons des architectures neuronales destinées à deux défis majeurs, chacun orienté vers des finalités distinctes : le cancer du sein, pour lequel nous utilisons des images 2D spatiales, et la maladie coronarienne, pour laquelle nous nous appuyons sur des images spatiales 2D également, mais temporelles. Cette démarche a conduit à quatre contributions significatives. La première, sur la qualité et le traitement des données médicales privées (mammographies et artères coronaires) partant de l'anonymisation des données (confidentielles et très sensibles) avec les bonnes pratiques de chiffrement jusqu'à la labellisation au travers d'une application réalisée sur mesure afin d'optimiser cette entreprise chronophage. La deuxième, concerne la suspicion d'anomalies appliquées aux lésions précoces du cancer du sein et aux lésions coronaires. Respectivement, elle vise à classifier les tumeurs de manière binaire (négatives et positives), à localiser les tumeurs sur la mammographie, ainsi qu'à détecter les sténoses présumées et à classifier le type d'artère coronaire. Ces tâches sont réalisées à l'aide de diverses architectures neuronales consolidées par l'intégration de techniques d'explicabilité et d'interprétabilité. La troisième, l'apprentissage fédéré, évalué dans le cadre de la détection précoce du cancer du sein avec une architecture distribuée permettant de traiter des données confidentielles en collaboration avec plusieurs partenaires, sans avoir à transférer les ensembles de données à la coalition. La quatrième, la cybersécurité pour l'IA, dans le contexte de l'apprentissage fédéré, plus particulièrement pour la protection des données et des algorithmes, respectivement avec la confidentialité différentielle et le chiffrement homomorphique. Les diverses contributions ont permis d'arriver à des résultats très prometteurs. Pour la suspicion des anomalies précoces du cancer du sein (en se concentrant sur les lésions), avec une courbe ROC dépassant les 90% pour la classification binaire des tumeurs via des approches basées sur l'apprentissage traditionnel et fédéré. De même pour la suspicion des lésions coronaires et le type d'artère coronaire, nous avons respectivement atteint une courbe ROC de 82 % et 93 %. Quant à l'architecture sécurisée fédérée proposée, elle permet de répondre de manière adéquate aux multiples vulnérabilités inhérentes à l'architecture distribuée. Cette proposition vise à atténuer les risques d'attaques par empoisonnement et inférence, tout en préservant rigoureusement la confidentialité des données. En effet, cette approche garantit que les données demeurent sous la supervision de leurs propriétaires respectifs, ce qui prévient toute perte de contrôle sur ces éléments informationnels sensibles. Nos architectures présentent l'avantage d'être très génériques, ce qui les rend applicables à d'autres cas d'utilisation, aussi bien dans le secteur hospitalier que dans d'autres secteurs.

Il est important d'insister sur le fait que ces technologies ne prétendent pas remplacer le médecin, mais plutôt d'augmenter ses performances en combinant les deux intelligences : la première, artificielle, apportant une nouvelle lecture avec une première analyse, et la seconde, humaine, pour interpréter les résultats, poser le bon diagnostic et à fortiori d'entreprendre un traitement.