

Hamza El Massaoudy

IA médicale fiable | Apprentissage multimodal | Validation externe

Nanjing, China | elmassaoudy.hamza@gmail.com | hamzaelmassaoudy.site
github.com/hamzaelmassaoudy | linkedin.com/in/hamza-elmassaoudy

PROFIL

Étudiant en master d'ingénierie en informatique à NUIST, spécialisé en IA médicale fiable. Mes travaux associent apprentissage multimodal, audit des données et validation externe, avec des logiciels reproductibles et une analyse rigoureuse des défaillances. Mon mémoire porte sur la détection de la fibrillation atriale avec des modèles de fondation ECG sous décalage de distribution.

AXES DE RECHERCHE

Validation externe et décalage de distribution ; apprentissage de raccourcis et audit des données ; calibration et prédiction sélective ; apprentissage multimodal ; explicabilité avec des limites clairement définies.

FORMATION

Master d'ingénierie en informatique | NUIST 2025 - 2027
Nanjing University of Information Science & Technology. Diplôme prévu en juin 2027. GPA : 3.69/4.0.
Bachelor d'ingénierie en informatique | Dalian Polytechnic University
2021 - 2025
Dalian, Chine. Formation dispensée en anglais.

MANUSCRIT

Wen X., El Massaoudy H. (auteur correspondant). Multimodal Classification of Primary Liver Tumors on Multiphase CT: Internal Performance, Acquisition Artifacts, and Limited External Transportability. *Journal of Intelligent Medicine and Healthcare*.
Under revision.

EXPÉRIENCE DE RECHERCHE

Pipeline CT hépatique et validation externe [Code public] 2025 - 2026

- Développé et publié un pipeline ML Python reproductible (v2.4.2) avec pytest, GitHub Actions CI, smoke tests et contrôles de publication.
- Audit 1,112 examens et identifié des artéfacts d'acquisition confondus avec la classe. Des tests contrefactuels ont vérifié que le CNN exploitait ces artéfacts.
- Testé le modèle sur 203 patients externes et analysé le décalage d'acquisition comme source de défaillance.

Détection fiable de la fibrillation atriale avec des modèles de fondation ECG

Sept. 2026 - présent

- Étudie la sélection de dérivations selon la confiance et le renvoi des cas incertains à une revue humaine sous décalage de distribution. Encadrant : Wen Xuezh.
- Compare les encodeurs gelés ECGFounder et ECG-JEPA à une référence ResNet1D avec prédiction sélective.
- Développement sur PTB-XL et évaluation externe prévue sur CPSC2018 et Georgia avec tests de stress au bruit de mouvement des électrodes. Résultats non encore publics.

Trustworthy Lead Selection and Deferral for Atrial Fibrillation Detection with ECG Foundation Models under Distribution Shift

Hamza El Massaoudy

CV académique suite

EXPÉRIENCE PROFESSIONNELLE

gethsk | Cofondateur

Janv. 2026 - présent

- Cofondé une application d'apprentissage du chinois pour le HSK, en développement. Conçois les parcours et les cours, et développe le code associé.

AsiaScholar | Développeur full-stack et coordinateur technique (stage)

Nov. 2024 - juin 2025

- Piloté le développement et la maintenance du site de l'agence et créé un système interne de gestion des opérations, en coordonnant le travail technique entre équipes.

Indépendant | Développeur web freelance

2023 - 2025

- Créé un Product Management System PHP/MySQL pour factures, commandes, paiements et analyses, vendu plus de 100 fois au Maroc, en Arabie saoudite et aux Émirats arabes unis.
- Livré une boutique en ligne avec paiement et un système de gestion des élèves en Java/Spring Boot/SQL pour École Al Faraj, Taza.

PROJET LOGICIEL SÉLECTIONNÉ

Prédiction explicable de la survie dans le HCC [Code public]

2026

- Développé un package Python et une interface en ligne de commande avec démonstration Streamlit locale, tests pytest sur données synthétiques, pre-commit et ruff.
- Mis en place une validation croisée imbriquée sans fuite, avec calibration et choix du seuil limités à l'entraînement. Documenté méthodes et limites dans des fiches modèle et données.

COMPÉTENCES

Imagerie médicale: CT multiphasique, CNN 3D MONAI, PyRadiomics, LightGBM, DICOM/NIfTI, SimpleITK.

ECG et IA fiable: Modèles de fondation ECG, ResNet1D, transfert avec encodeurs gelés, robustesse au bruit, prédiction sélective, calibration, Grad-CAM.

Évaluation: Validation croisée imbriquée sans fuite, jeux de test réservés, tests de DeLong, correction des comparaisons multiples, importance par permutation, fiches modèle et données.

Python et ML: Python, PyTorch, entraînement GPU, scikit-learn, NumPy, pandas, SciPy, Matplotlib, TensorFlow/Keras.

Ingénierie: pytest, GitHub Actions CI, pre-commit, Git, Linux, packaging Python ; Docker (travaux pratiques uniquement).

Logiciel: Django, Spring Boot, Flask, Java, JavaScript, PHP, MySQL, PostgreSQL, Streamlit.

DISTINCTIONS

Bourse Longshan, première classe, NUIST (2026 - 2027).

CERTIFICATIONS SÉLECTIONNÉES

AI for Medical Diagnosis, DeepLearning.AI / Coursera (mars 2026).

Supervised Machine Learning: Regression and Classification, DeepLearning.AI et Stanford (février 2026).

HSK niveau 4 et HSKK intermédiaire (janvier 2026).

LANGUES

Arabe et amazigh (langues maternelles), anglais (C1), chinois (HSK 4), français (B1-B2).

FORMATION ANTÉRIEURE

BTS Multimédia et Conception Web, Lycée Technique de Taza, Maroc (2019 - 2021).