

EL HAJ EBOU

Junior ML / Computer Vision Engineer | Medical Imaging, Deep Learning, Self-Supervised Learning

📍 Paris, Île-de-France 📞 07 59 58 62 25 ✉️ hadjteyib2@gmail.com 🌐 El Haj Ebou 📱 hajteyib 📁 Portfolio

PROFIL

Junior ML / Computer Vision Engineer, avec une formation d'ingénieur en informatique et un M2 spécialisé en Vision et Machine Intelligente. Expérience en deep learning appliqué à l'imagerie médicale 3D, à l'apprentissage auto-supervisé et à la construction de pipelines IA complets, du traitement des données à l'évaluation des modèles. Profil hybride recherche appliquée / engineering, capable de travailler sur des données complexes, d'expérimenter des architectures modernes et de construire des solutions IA exploitables.

EDUCATION

- **Université Paris Cité, Paris, France** Sept. 2025 – Août. 2026
Master 2 (M2) - Vision et Machine Intelligente
- **Institut National des Sciences Appliquées de Lyon - INSA Lyon** 2024-2025
Élève-ingénieur au Département Informatique
- **École Supérieure Polytechnique (ESP), Nouakchott, Mauritanie** 2022-2024
Élève-ingénieur en Informatique, Réseaux et Télécommunications

EXPÉRIENCE PROFESSIONNELLE

Stage - Segmentation et classification de nodules pulmonaires sur CT thoraciques Mars 2026 – Août 2026
AP-HP, Hôpital Saint-Louis, Paris, France

- Développement d'un pipeline reproductible de segmentation 3D de nodules pulmonaires sur CT thoraciques, appliqué à une cohorte clinique rare de patients immunodéprimés.
- Structuration et exploitation d'un jeu de données privé d'environ **200 cas exploitables**, incluant annotations de nodules, métadonnées cliniques et harmonisation DICOM/NIfTI.
- Entraînement et évaluation d'une baseline robuste de segmentation guidée avec **nnU-Net**, atteignant **0,83 de Dice moyen** sur un held-out test set interne.
- Adaptation de **MedSAM** par fine-tuning paramètre-efficace avec **LoRA**, améliorant la segmentation guidée avec **0,87 de Dice moyen, HD95 1,58 mm et ASSD 0,58 mm** sur le test interne.
- Préparation du volet aval combinant masques segmentés, radiomics et variables cliniques pour la classification étiologique des nodules infectieux.

Compétences techniques : *Python, PyTorch, nnU-Net, MedSAM, LoRA, SimpleITK, DICOM/NIfTI, TCIA/LIDC-IDRI, pandas, scikit-image, Git*

Stage - Apprentissage auto-supervisé pour la reconnaissance d'écriture manuscrite Mai 2025 – Septembre 2025
LIRIS, Lyon, France

- Conception d'une approche SSL pour la reconnaissance d'écriture manuscrite, combinant apprentissage contrastif avec SimCLR et reconstruction masquée avec MAE/Vision Transformer sur données non annotées.
- Constitution d'un corpus d'entraînement à grande échelle par web scraping de documents historiques Gallica/BNF et génération de données synthétiques.
- Pré-entraînement auto-supervisé puis fine-tuning supervisé sur données annotées, avec évaluation comparative face à des baselines internes entraînées directement en supervisé.
- Réduction du Taux d'Erreur Caractère jusqu'à environ **5%** après fine-tuning, avec une amélioration nette par rapport aux modèles de référence internes.

Compétences techniques : *Python, PyTorch, SimCLR, MAE/ViT, OpenCV, NumPy, BeautifulSoup*

PROJETS

- **Brain Tumor Segmentation — U-Net, ResNet-34 & Knowledge Distillation** Mars 2025
Projet académique — Deep Learning appliqué à l'imagerie médicale
 - Conception d'un pipeline complet de segmentation de tumeurs cérébrales sur IRM, de la préparation des données à l'évaluation des masques prédits.
 - Implémentation d'un modèle U-Net avec encodeur ResNet-34, adapté aux contraintes de segmentation médicale où la précision spatiale est essentielle.
 - Mise en place d'une approche de knowledge distillation pour transférer les connaissances d'un modèle enseignant vers un modèle étudiant plus léger, dans une logique de compression et d'efficacité du modèle.
 - Évaluation sur un ensemble de test avec Dice Score 0.915, IoU 0.887 et 98.2% de validation accuracy.
 - **Stack :** Python, PyTorch, U-Net, ResNet-34, OpenCV, NumPy, medical image segmentation.
- **Chest X-ray Pathology Detection — EfficientNet, Uncertainty Handling** 2025 - 2026
Projet académique — Deep Learning appliqué à l'imagerie médicale
 - Développement d'un pipeline de classification multi-label de radiographies thoraciques sur le dataset CheXpert, visant la détection automatique de plusieurs pathologies pulmonaires et cardiaques.
 - Entraînement d'un modèle EfficientNet-B0/DenseNet-121 avec gestion du fort déséquilibre de classes et des labels incertains, via pondération de loss et stratégies U-Zero / U-Ignore.
 - Évaluation du modèle avec mean AUC 0.86, AUC 0.91 sur pleural effusion, 0.89 sur cardiomegaly et 0.88 sur edema, complétée par F1-score, sensibilité et spécificité.
 - **Stack :** Python, PyTorch, TorchVision, EfficientNet, scikit-learn, OpenCV, medical image classification.

PROFIL COMPLÉMENTAIRE

Langues: Français (DALF C2), Anglais (TOEIC C1)

Centres d'intérêt Football, tennis, cinéma, jogging

Expériences collaboratives : Projets en équipe multidisciplinaire, mentoring d'étudiants juniors