



INSTITUT UNIVERSITAIRE
DU CANCER DE TOULOUSE
Oncopole



Intitulé de l'offre de stage

Réalisation d'un pipeline de Machine Learning et Deep Learning pour l'aide au diagnostic de certaines lésions buccales potentiellement précancéreuses

Stagiaire en	Intelligence Artificielle (Deep Learning)
Affectation	IUCT Oncopole
Durée	5 à 6 mois
Rémunération	environ 670 € net mensuel
Date de la publication	15/11/2024
Date d'embauche prévue	mars ou avril 2025 suivant disponibilités
Lieu	IUCT Oncopole 1 avenue Irène Joliot-Curie, 31100 Toulouse

L'IUCT Oncopole

L'Oncopole regroupe sur un même site plus de 2000 professionnels engagés dans la lutte contre le cancer. Ils sont répartis entre l'Institut Universitaire du Cancer de Toulouse - Oncopole (IUCT-O) et le Centre de Recherche en cancérologie de Toulouse (CRCT). L'IUCT-Oncopole réunit l'expertise et l'excellence de l'Institut Claudius Regaud (Centre de Lutte Contre le Cancer) et de plusieurs services d'oncologie du CHU de Toulouse. Ses missions sont le soin, la recherche et l'enseignement.

L'Institut de Recherche en Informatique de Toulouse

L'Institut de Recherche en Informatique de Toulouse (IRIT), une des plus imposantes Unité Mixte de Recherche au niveau national, est l'un des piliers de la recherche en Occitanie avec ses 700 membres, permanents et non-permanents. De par son caractère multi-tutelle (CNRS, Universités toulousaines), son impact scientifique et ses interactions avec les autres domaines, le laboratoire constitue une des forces structurantes du paysage de l'informatique et de ses applications dans le monde du numérique, tant au niveau régional que national.

L'équipe d'accueil

L'informatique et l'intelligence artificielle sont appelés à prendre une place croissante dans le monde de la recherche médicale, et en particulier dans le monde de la recherche contre le cancer. Dans ce cadre, l'IRIT collabore activement avec le Centre de Recherche en Cancérologie de Toulouse, le CHU de Toulouse et l'Institut Claudius Régaud. Le but de ce travail commun est de dégager pour chaque projet un binôme d'encadrants appartenant d'un côté à l'IRIT et de l'autre au domaine médical. Cette collaboration a abouti à la création d'une équipe commune entre l'IRIT et le CRCT.

Le stagiaire bénéficiera d'un encadrement sur le plan clinique par l'équipe médicale composée du docteur Emmanuelle Vigarios et du docteur Delphine Comtesse-Maret.

Objet du stage

Les cancers de la cavité buccale font partie des cancers les plus fréquents dans le monde et sont de pronostic défavorable. En France, pour l'année 2018, le nombre estimé de nouveaux cas de cancer de la cavité buccale était de 4 677 dont 3 106 chez les hommes et de 1 571 chez les femmes. La plupart des patients sont diagnostiqués à un stade avancé et il n'y a pas d'augmentation du taux de survie à 5 ans depuis plus de 30 ans (Le Plan contre le Cancer, 2021). Les patients porteurs de lésions orales à potentiel malin (LOPM) ont une susceptibilité accrue de développer un cancer n'importe où dans leur bouche au cours de leur vie.

Les LOPM peuvent concerner n'importe quel site anatomique de la cavité buccale et peuvent être uni- ou multifocales. Les LOPM ont une évolution clinique imprévisible : elles restent statiques ou peuvent progresser ou régresser en quelques mois. Dans ce contexte, l'informatique et l'intelligence artificielle pourraient apporter une aide complémentaire au clinicien pour le dépistage précoce des cancers et pour le suivi des lésions orales à potentiel malin.

Il s'agit :

- > de réaliser une étude bibliographique afin d'avoir un état de l'art de la recherche en matière d'utilisation de techniques de Machine Learning et de Deep Learning dans le domaine des lésions orales à potentiel malin et des cancers de la muqueuse buccale
- > d'identifier et de récupérer les données/caractéristiques physiologiques d'intérêt à partir d'une base de photographies cliniques de lésions orales issues de consultations spécialisées du Dr Emmanuelle Vigarios (IUCT-Oncopole)
- > de les convertir dans un format exploitable par des paquetages standards de Deep Learning (TensorFlow, PyTorch)
- > de tester, à partir de la base annotée, l'étape de détournement avec une architecture type *Mask R-CNN* pour extraire des photos la zone d'intérêt de la langue
- > A partir des lésions annotées et de l'extraction de caractéristiques images ("features") des LOPM, de classer par un algorithme type *XGBoost* les zones saines des lésions avec une interprétation des variables via les valeurs de Shapley.
- > de tester, à partir de la base annotée, une architecture type *YoloV5* pour identifier et classer les zones saines des lésions de la langue et comparer cette approche à celle basée "features".

Formation

École d'ingénieur, de préférence avec spécialisation en informatique ou mathématiques (en année de césure ou stage long). Master 2 informatique ou mathématiques appliquées.

Compétences attendues

Des connaissances en science des données, traitement d'images, apprentissage et deep learning seront appréciées.

Modalité de candidature

- > CV à envoyer à: odonto@stages-medecine-numerique.fr
- > Date limite de candidature : 1/03/2025
- > Encadrants: Docteur Emmanuelle Vigarios (ICR-IUCT), Docteur Delphine Comtesse-Maret (CHU), Sandrine Mouysset (IRIT)

Une première sélection sera effectuée sur la base des CVs reçus. Les candidats seront immédiatement informés du résultat, et ceux qui seront retenus à l'issue de la première sélection seront invités à un ou deux entretiens individuels en téléconférence avant sélection définitive.

L'ensemble des offres de stage est disponible sur <http://www.stages-medecine-numerique.fr>.