

Prévention des cancers et détection de tumeurs grâce à l'intelligence artificielle

Les progrès de l'informatique, notamment ceux ayant lieu en intelligence artificielle, devraient permettre des avancées dans de nombreux domaines, comme le domaine médical.

L'intelligence artificielle permet désormais de détecter des tumeurs malignes avec un meilleur taux de réussite qu'un radiologue. Une détection plus rapide et efficace de ces tumeurs permet de mieux les traiter, de prévenir l'avancement du cancer et d'améliorer les chances de guérison du malade.

Positionnement thématique (ETAPE 1)

INFORMATIQUE (Informatique pratique).

Mots-clés (ETAPE 1)

Mots-Clés (en français)	Mots-Clés (en anglais)
<i>Intelligence artificielle</i>	<i>Artificial intelligence</i>
<i>Base de données</i>	<i>Data set</i>
<i>Traitement d'images</i>	<i>Image processing</i>
<i>Propagation des cellules cancéreuses</i>	<i>Spread of cancer cells</i>
<i>Méthode des plus proches voisins</i>	<i>Nearest neighbour classifier</i>

Bibliographie commentée

La majorité des tumeurs, dans le cas d'un cancer du sein, sont désormais détectées de manière non palpées, par mammographie [2]. Dans ce cadre, de nombreuses lésions malignes passent inaperçues et d'autres lésions suspectes s'avèrent souvent bénignes [3].

Or, depuis 2019, grâce notamment au travail d'entreprises comme Google, l'intelligence artificielle est plus performante qu'un radiologue pour dépister un cancer (sein, poumon...) et réduit aussi le nombre de faux positifs [2][3]. La méthode égalait même les performances des diagnostics établis par deux médecins, et sa précision ne peut qu'augmenter.

L'outil s'avère alors extrêmement utile pour les radiologues (qu'il ne vise pas à remplacer, mais à aider) et permet la détection précoce de cancers, il est déjà autorisé par d'importantes instances officielles, notamment aux États-Unis [3].

Le type d'intelligence artificielle privilégié dans ce genre de problème de classification, où le nombre de données est important, est la méthode des plus proches voisins [4]. On pourra traiter les données grâce à des techniques de traitement d'images [1].

Pour l'utiliser, on sépare notre base de données entre bases d'entraînement, de test et de validation.

La méthode des plus proches voisins consiste alors à placer des points correspondant aux éléments de la base d'entraînement, dont on connaît le groupe (malade ou sain), sur un graphique en fonction de différents paramètres [1].

On peut attribuer un groupe aux éléments de la base de test selon les k plus proches voisins des points correspondant à ces éléments [1][4].

Enfin on vérifie sur la base de validation que le modèle n'est pas "trop spécifique" et "trop adapté" à la base de test [1].

On mesure aisément le taux de réussite et le taux de faux positifs du modèle, et on peut alors optimiser celui-ci en faisant varier ses "hyperparamètres" (type de calcul de distance, nombre de voisins utilisé...) [4]. On conserve le taux de réussite maximum pour obtenir l'IA la plus efficace possible.

Problématique retenue

Il s'agit de modéliser une base de données de radiographies et comprendre le fonctionnement d'une intelligence artificielle qui permet de détecter des tumeurs malignes.

Objectifs du TIPE

Mon objectif est de reproduire une intelligence artificielle permettant de détecter des tumeurs que j'aurais modélisées. Je mettrai en place un modèle simple représentant les radiographies d'individus sains et malades, grâce à des cellules infectées pouvant se propager (deux dimensions). Ensuite j'entraînerai et optimiserai une intelligence artificielle supervisée utilisant la méthode des plus proches voisins afin qu'elle détecte les tumeurs. Enfin je complexifierai le modèle étudié (trois dimensions).

Références bibliographiques (ETAPE 1)

[1] MACHINE LEARNIA : PYTHON SKLEARN: KNN, LinearRegression et SUPERVISED LEARNING (20/30) :

https://www.youtube.com/watch?v=P6kSc3qVph0&list=PLO_fdPEVlfKqMDNmCFzQISI2H_nJcEDJq&index=20

[2] CORALIE LEMKE : Une IA de Google surpasse les radiologues pour détecter le cancer du sein :

https://www.sciencesetavenir.fr/sante/une-ia-de-google-surpasse-les-radiologues-pour-detecter-le-cancer-du-sein_140225

[3] BRUNO BENQUE AVEC RSNA : IA et dépistage du cancer du sein : sensibilité accrue et temps de diagnostic raccourci : <https://www.thema-radiologie.fr/actualites/2804/ia-et-depistage-du-cancer-du-sein-sensibilite-accrue-et-temps-de-diagnostic-raccourci.html>

[4] P. BHUVANESWARI ET A. BRINTHA THERESE : Detection of Cancer in Lung with K-NN Classification Using Genetic Algorithm :

<https://www.sciencedirect.com/science/article/pii/S2211812815003156>

Références bibliographiques (ETAPE 2)

[1] HUGUES MEUNIER : Jeu de la vie épidémique en Python :

<https://www.hmeunier.com/2020/05/jeu-de-la-vie-epidémique-en-python.html>

DOT

[1] *juillet 2021 : ébauche du MCOT (sans bibliographie commentée et modifiée par la suite)*

[2] *juillet-octobre 2021 : réflexion sur les modélisations possibles et implémentation de ces modèles en Python (ils ne seront pas retenus)*

[3] *novembre-février 2021 : adaptation des modèles, en prenant compte des remarques et conseils du professeur encadrant, programmation en Python*

[4] *mai 2022 : réalisation du beamer, retouche des algorithmes, rédaction finale du MCOT et du DOT*