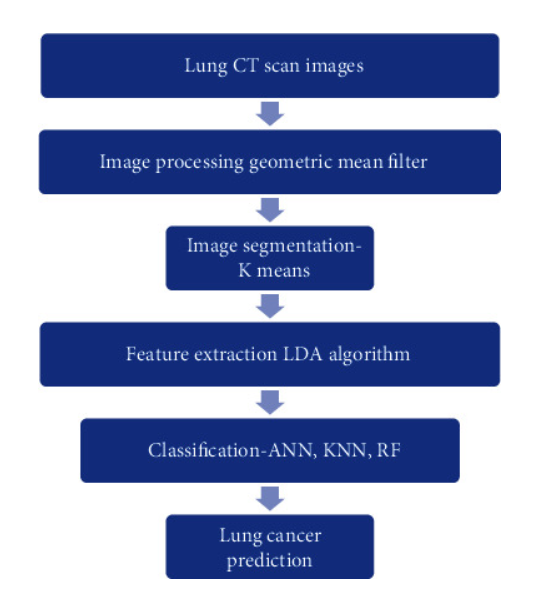
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9424001/>

Etant une des maladies les plus mortelles dans le monde, le cancer du poumon cause des millions de morts dans le monde.

Repérables à partir de scans, l’objectif est de pouvoir identifier les nodules cancéreuses des poumons le plus rapidement possible.

Pour cela, des techniques de détection et prédiction sur des images de scans ont été développées. Ci-dessous un schéma du processus utilisé pour détecter les cancers.



1. Preprocessing :

Des filtres sont appliqués sur les images en entrée pour les redimensionner mais aussi réduire les imperfections et les bruits des images. Pour cela, des méthodes de filtrage sont utilisées comme le [geometric mean filter](https://en.wikipedia.org/wiki/Geometric_mean_filter)

1. Segmentation :

La segmentation permet de regrouper les éléments de l’image qui se ressemblent, qui se rapprochent au niveau de leur aspect. Pour cela, la technique la plus utilisée, surtout dans le domaine médical, est le [K-means clustering](https://en.wikipedia.org/wiki/K-means_clustering)

1. Extraction des caractéristiques

A partir de ces groupes, l’algorithme [LDA](https://en.wikipedia.org/wiki/Linear_discriminant_analysis) permet d’extraire et regrouper les différentes caractéristiques dans d’autres caractéristiques plus générales.

1. Classification

ANN : Aritificial Neural Network ⇒ Réseau de neurones artificiels

Le réseau de neurones parvient à identifier la catégorie de l’image en analysant les différentes caractéristiques. Avec les données qu’il a acquises pendant la phase d’entraînement, le réseau peut catégoriser l’image donnée

KNN : K-nearest neighbors network - apprentissage supervisé

Le réseau de neurones compare l’image donnée avec un dataset d’images dont on connaît les labels.

Il détermine la distance avec chaque image du dataset (distance en fonction de la ressemblance ?) et choisit les plus proches.

Avec les images les plus ressemblantes, le réseau de neurones attribue la catégorie correspondante.

D’après les résultats, la méthode de classification ANN est la plus précise.

