

Auteurs : Pohyer Valentin (Valentin.Pohyer@hotmail.fr), stagiaire en Master 2 bio-informatique ;
David Pau (david.pau@roche.com), Chef de Projet Statistiques et Data Science

Titre : Développements d'algorithmes ciblés d'analyse d'images numérisées de prélèvements tumoraux dans le cancer du rein afin de déterminer le statut du biomarqueur PD-L1.

Objectifs : A partir d'un échantillon de 250 patients de la cohorte nationale UroCCR, identifier grâce à des algorithmes les différentes zones importantes qui composent les lames numérisées des échantillons de tumeur réutilisés afin de déterminer le statut PD-L1 avec le test Ventana SP142.

Méthodes : Trois types de lames histologiques Hematoxylin & Eosin staining (H&E), Negative Control (NC) et PD-L1 ont été numérisées au format WSI (microscope virtuel) par le scanner 3DHistech panoramic 250 de plateforme Imag'IN de l'Institut Universitaire du Cancer de Toulouse.

La segmentation des images par application de filtres colorimétriques a été utilisée pour: A faible grossissement, 1/ identifier le tissu présent sur les lames, 2/ Identifier et supprimer certains artefacts trouvés de façon récurrente sur les lames via leurs caractéristiques colorimétrique (Figure 1), 3/ Extraire des données (en nombre de pixels ou en pourcentage de tissu) sous la forme d'indicateurs permettant de quantifier ces différents segments présents sur les lames, 4/ Supprimer le bruit autour des images (zones non exploitables telles qu'un stylo sur les lames). A fort grossissement, identifier les cellules immunitaires marquées permettant de déterminer la surexpression du biomarqueur PD-L1.

Toutes les analyses ont été réalisées avec le langage Python 3.7.4 et la librairie Openslide.

Discussion : L'approche de segmentation des images par niveau de couleur RGB a d'abord été testé sur un nombre très limité d'image avant d'être étendu à un sous-ensemble de 60 lames (training dataset), puis d'être déployé sur le reste de la cohorte (validation dataset).

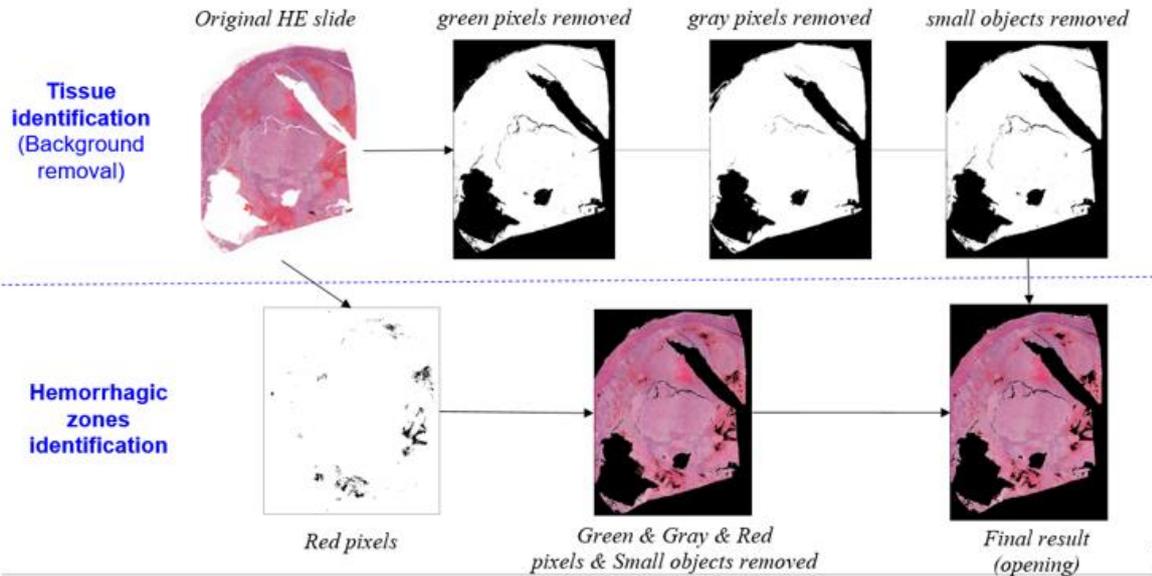
Le processing décrit dans la partie méthode est une étape importante pour l'obtention final d'un statut PD-L1 grâce à des algorithmes. L'utilisation de la segmentation des images par RGB a déjà permis d'identifier différentes zones importantes dans le cadre du test PDL1 sur les lames de cancer métastatique du rein.

La première étape permet d'identifier pour chaque patient la minorité de surface des lames correspondant à du tissu. La seconde étape permet d'éliminer les surfaces inexploitables sur les lames afin de pouvoir réaliser le test PD-L1.

L'avantage principal de l'utilisation d'algorithmes est d'éviter les biais liés à l'analyse mentale humaine actuellement en vigueur sur ce test dans cette pathologie, ainsi que l'obtention de résultats inchangés d'une exécution à l'autre.

Figure 1 : application de filtres colorimétriques sur les différents types d'images

1 - Processing initial des images Hematoxylin & Eosin (H&E)



2 - Processing initial des images PDL1 et Negative Control

