

Reconnaissance automatique des anomalies chromosomiques afin d'accélérer le diagnostic pronostic dans la prise en charge des cancers du sang (Part I)

Ludivine, Erwan, Céline, Patricia

Août 2021

Microsoft IA/Simplon

Sommaire :

Introduction	3
Enjeux de la reconnaissance des chromosomes	3
Données disponibles	3
Classification des images des chromosomes	4
Segmentation des images des chromosomes	6

Introduction

Enjeux de la reconnaissance des chromosomes

A l'heure actuelle, les cancers du sang, hémopathies malignes, représentent 30% des cancers de l'enfant et 5% des cancers chez l'adulte. Afin de pouvoir classer ces cancers, il est nécessaire de prendre en compte le tissu d'origine, la morphologie, les marqueurs protéiques, ainsi que les données chromosomiques de ces cancers. Les cancers du sang sont diagnostiqués grâce au caryotype, car ils sont dus à une translocation de gènes, dans laquelle les chromosomes 9 et 22 ont échangé réciproquement des fragments de chromosomes.

Pour pouvoir mettre en évidence ce type d'anomalie chromosomique, il est donc nécessaire de classer les chromosomes afin d'établir le caryotype, puis de l'analyser. Ce projet est composé de deux parties, la première se charge de la segmentation et de la classification (Part I), et la deuxième se consacre à la détection d'anomalies (Part II). Dans ce brief, on va traiter la première partie.

Un caryotype est réalisé à partir d'une photographie d'une cellule en vue microscopique lors de la métaphase de la mitose. A cette étape, la chromatine est condensée ce qui rend les chromosomes visibles. Dans un caryotype, les chromosomes sont classés par paire, par taille et en fonction de la position du centromère. Il y a 23 paires de chromosomes autosomes (de 1 à 22) et une paire de gonosomes (XX pour les femmes et XY pour les hommes).

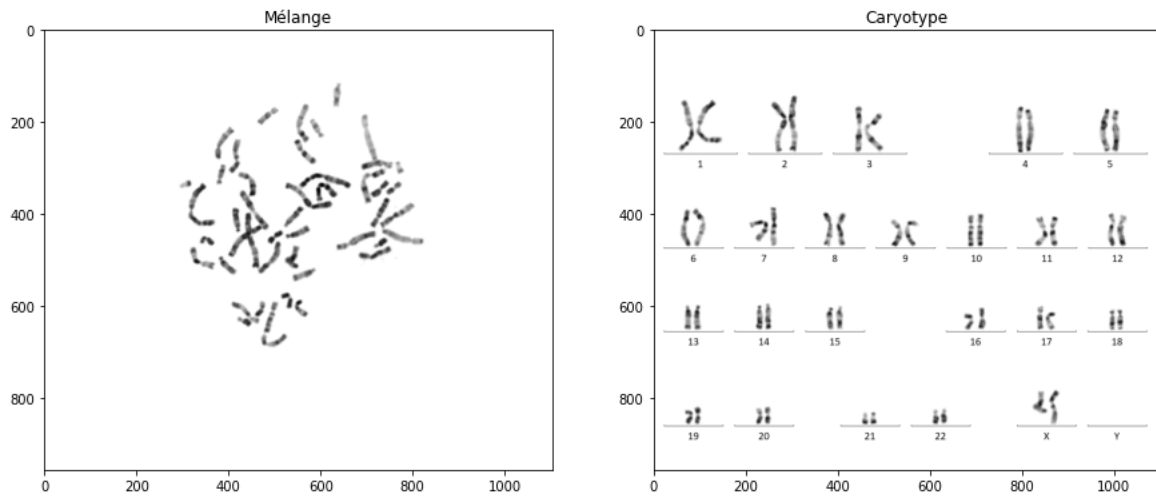
Les logiciels de réalisation des caryotypes actuels sont semi-automatiques et nécessitent encore l'intervention des techniciens de laboratoire qui savent reconnaître les chromosomes.

L'objectif de ce projet est d'automatiser entièrement ce processus. Pour ce faire, il est nécessaire de procéder en deux étapes. Il s'agit d'abord d'isoler chacun des chromosomes présents dans les vues microscopiques des cellules en phase de mitose (que nous appellerons "mélange"), puis de les classer, afin d'établir le caryotype final.

Données disponibles

203 couples d'images mélange/caryotype ont été mis à disposition par le CHRU pour la réalisation de ce projet

Visualisation d'un exemplaire
des 203 paires mélange/caryotype
disponibles pour l'analyse



Classification des images des chromosomes

Le travail de classification se fait en utilisant uniquement les caryotypes.

En premier lieu, un travail est réalisé pour extraire des images de chromosomes de chacun des 203 caryotypes. On obtient donc environ $23 \times 203 = 4669$ images de chromosomes (plus ou moins, en fonction des anomalies qui ont pu être rencontrées : trisomies, monosomies, ...).

Les images obtenues, extraites et enregistrées accompagnées de leur label (leur n° chromosomique), permettront d'entraîner un modèle de classification.

Ces images sont similaires à celles visibles ci-dessous :

visualisation des cromosomes séparés



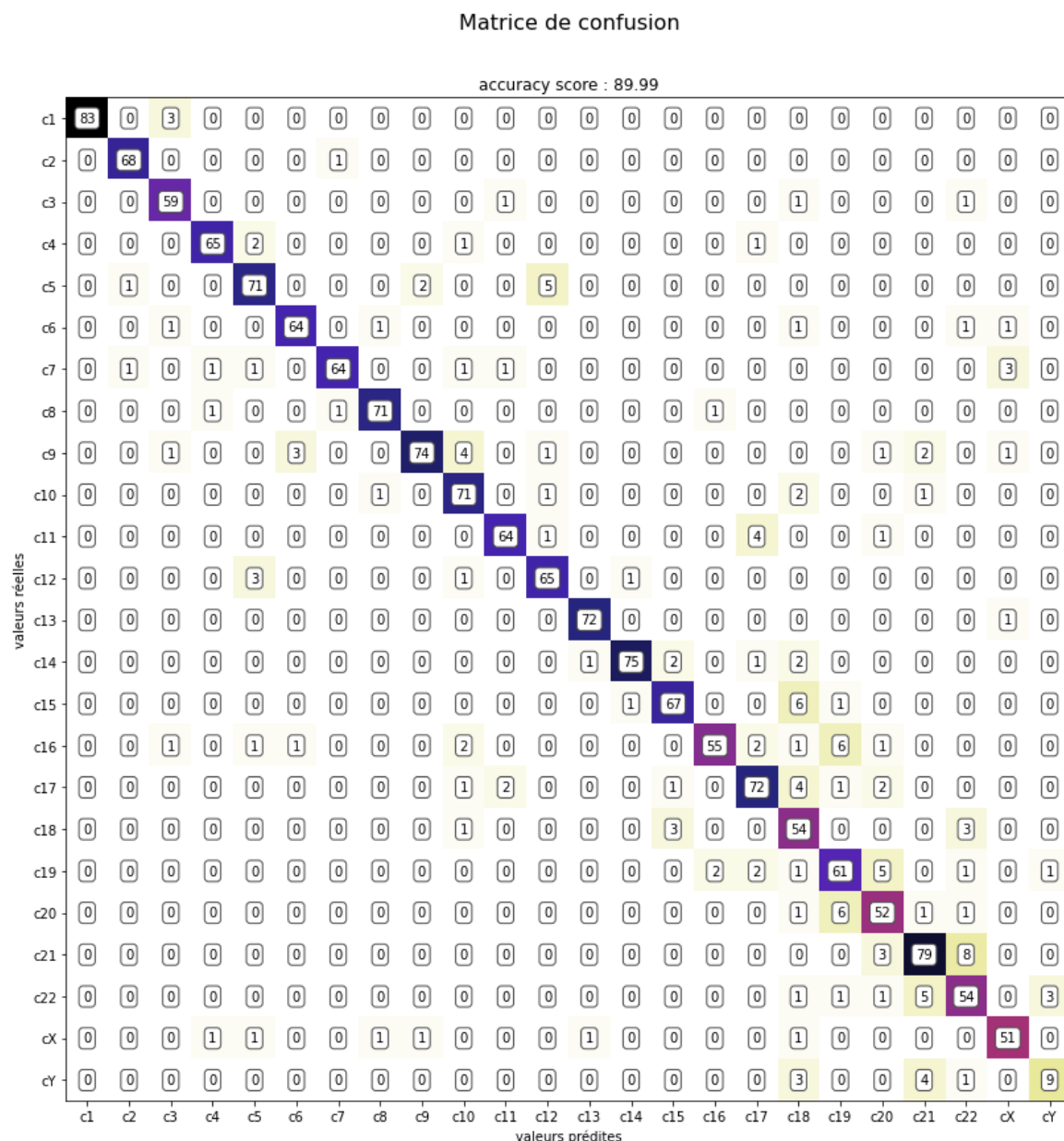
Le code pour obtenir les images et leur label est décrit dans le notebook [chromosomes_extraction.ipynb](#)

Grâce à ces images on peut ensuite entraîner un modèle de classification efficace.

Une data-augmentation a été réalisée pour donner une rotation aléatoire à chaque image, les chromosomes détectés dans les mélanges n'étant pas dirigés tous dans le même sens.

Un modèle séquentiel de CNN convolutif est ensuite entraîné sur les images.

Les résultats en sortie du modèle sont plutôt bons, avec une accuracy finale de 90% sur la base de test.



Segmentation des images des chromosomes