

## Itération 3 : Analyse, tests, segmentation

### 1. Analyse des résultats VGG16 en fine Tunning

Sur les données d'évaluation de la même base d'images

Très bon score de 98.3%

Précision moindre (93%) sur les IG

Rappel sur toutes les classes plus faible : prédictions en IG

Reconnaissance des IG difficiles, qui s'explique par des caractéristiques pouvant se retrouver dans d'autres types de cellules.

Les prédictions sont faites avec des pourcentages supérieures à 85%

#### Prédictions sur une autre base d'image

Score beaucoup plus faible

Confusion importante monocyte/platelet

Des rappels très faibles sur certaines classes : basophils, eosinophils, ig, lymphocytes et étonnant sur les platelets également

Précision faible sur les IG : classification à tort en IG de beaucoup d'autres classes

### 2. Etude sans les cellules IG

Au vu de la confusion apportée par les cellules de type IG, l'idée est de réaliser un modèle avec les 7 autres types de cellules et d'analyser comment sont classés les IG dans ce cas.

#### Modèle VGG16 sur les 7 types de cellules

Très bon score de 98.8%

#### Prédictions sur des IG

Presque la moitié des images sont prédites en monocyte, 20% en neutrophil et 20% en basophil.

La prédiction en neutrophil et basophil était attendu : les IG étant des cellules immatures de types neutrophil, basophil et eosinophil.

Cependant le classement en monocyte pose question : caractéristiques communes identifiées sur ces 2 types de cellules.

#### GradCam

Retrouve des caractéristiques propres au type de cellule prédite, mais manque des caractéristiques propres au type : ex prise en compte uniquement noyau et non cytoplasme.

Identification de caractéristiques en dehors de la cellule

### 3. Regroupement des 2 bases d'images, 10 classes

Ajout des images issues de la base des patients leucémiques

Ajout de 2 classes à identifier : blast et smudge (cellules éclatées)

#### Modèle base MobileNet Transfer Learning

Freeze de toutes les couches, utilisation de Dropout

Score de 88% mais entraînement plus rapide

#### 4. Segmentation / Détourage de la cellule

L'idée est de supprimer les informations autour de la cellule pouvant être utilisé à tort comme caractéristiques.

##### Génération automatique des masques : OpenCV

Sélection d'images contenant des cellules bien isolées, pas de globules rouges agglomérés

Détection du contour et génération des masques

##### Modèle détection des cellules

Utilisation de la base de 1400 images avec masques générés automatiquement pour modèle de segmentation.

Modèle basée sur VGG16

Contour très bien identifié sur les cellules de la base d'entraînement/test

Test sur nouvelles cellules contenant des globules rouges agglomérés : bon résultat, suppression de la totalité des globules rouges et du fond, suppression rarement d'une partie de la cellule

##### Modèle MobileNet avec detourage des cellules

Test du détourage au fur et à mesure de l'entraînement, mais beaucoup trop long

Choix d'enregistrement toutes les images détournées