

Applications de l'IA en cytologie : où en sommes-nous?

Pr. Thomas Boyer

Journées du Collège d'Hématologie des Hôpitaux

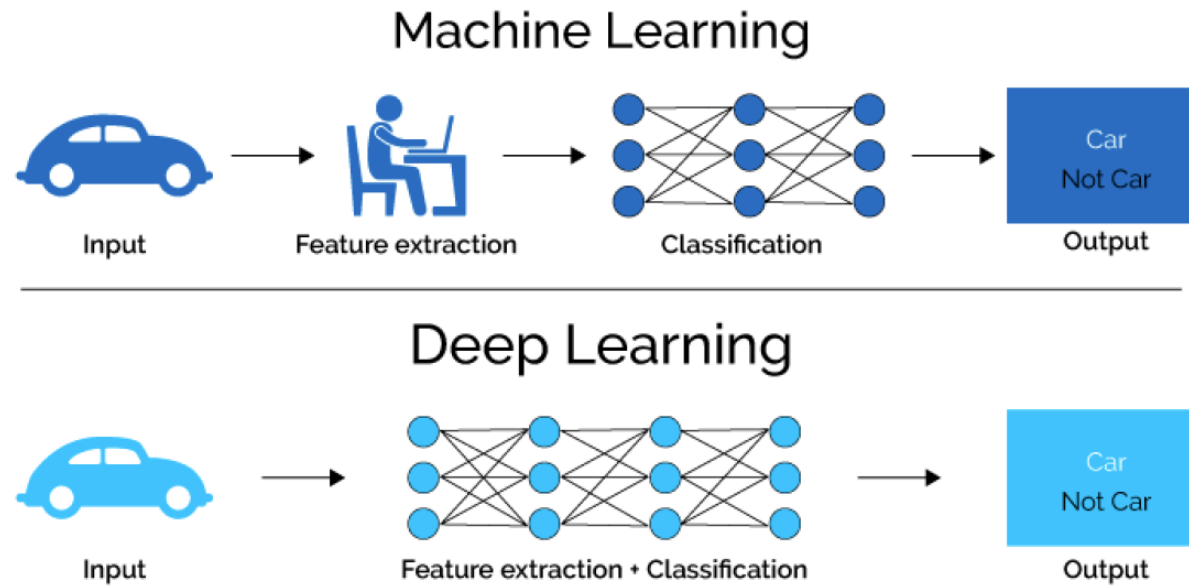
Faculté de Pharmacie - Paris

Mardi 9 décembre 2025

IA : définition

Ensemble des théories et des techniques mises en œuvre pour réaliser des machines dont le fonctionnement s'apparente à celui du cerveau humain

IA : généralités



<https://lawtomated.com/tag/unsupervised-learning/>

IA en hématologie biologique

Revue générale et analyse prospective

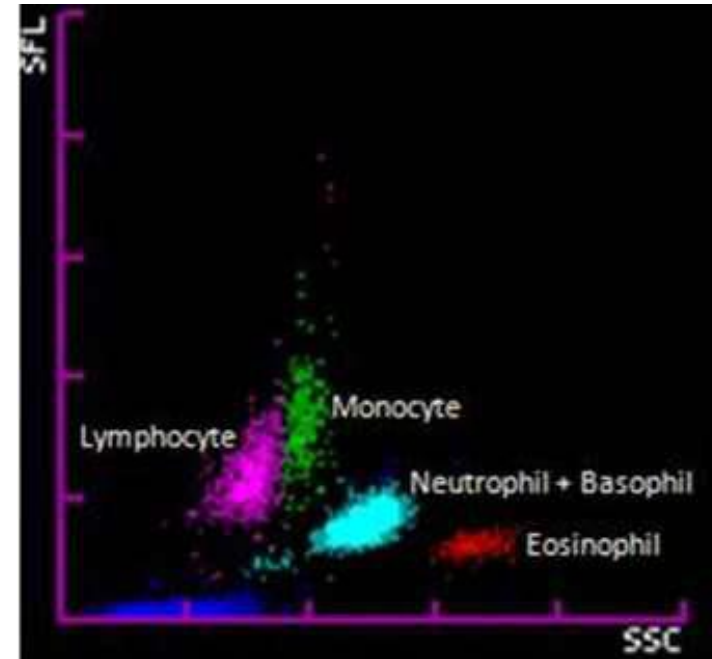
Validation informatique en hématologie

Computer validation in hematology

J.X. Corberand

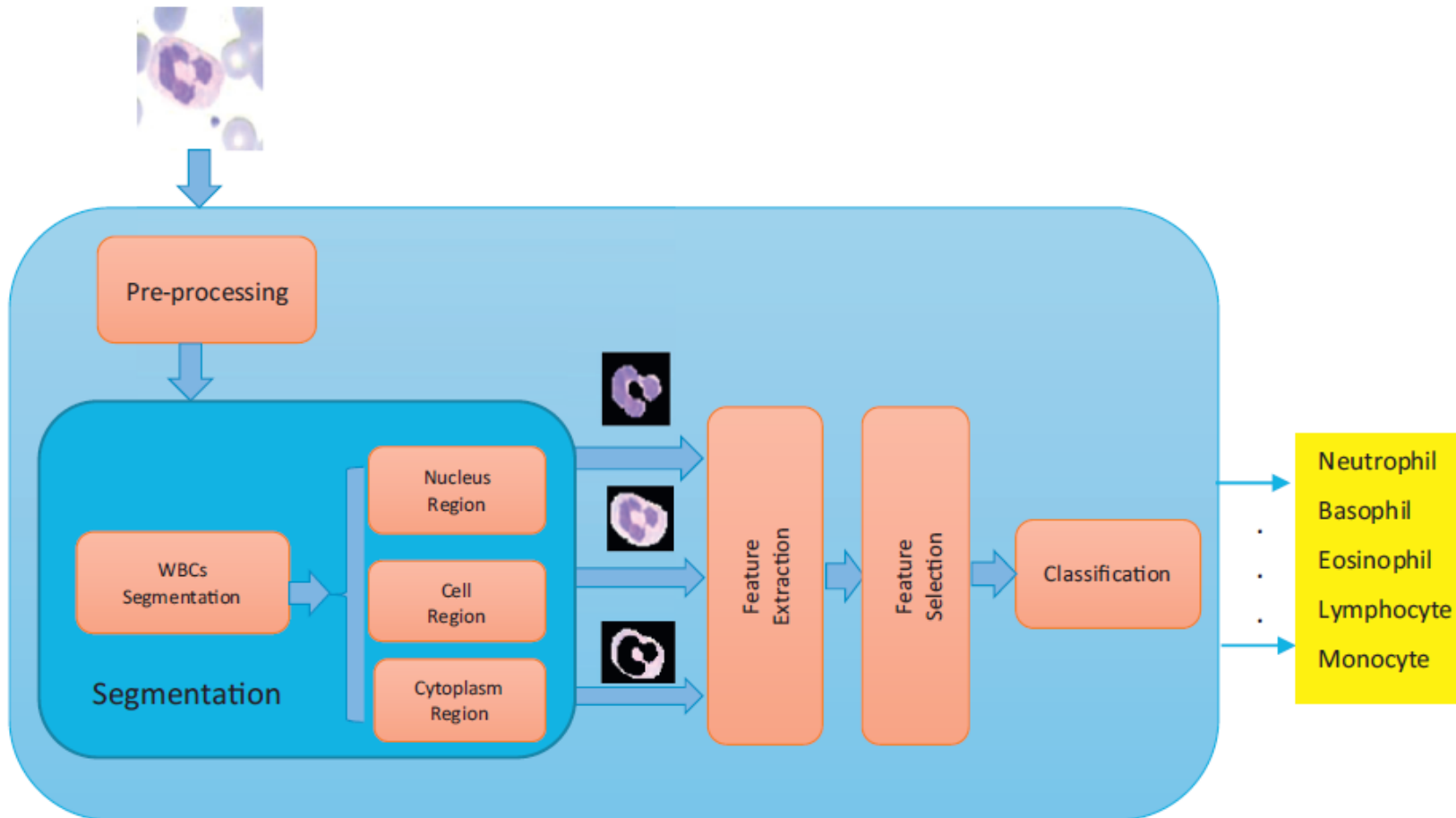
Laboratoire d'hématologie, hôpital Rangueil, 1, avenue Jean-Poulhès, 31405 Toulouse cedex 04, France

Reçu le 21 janvier 2003 ; accepté le 15 février 2003



« Par ma foi! Il y a plus de 40 ans que je dis de la prose sans que je n'en susse rien, et je vous suis le plus obligé du monde de m'avoir appris cela » (Molière, Le Bourgeois Gentilhomme)

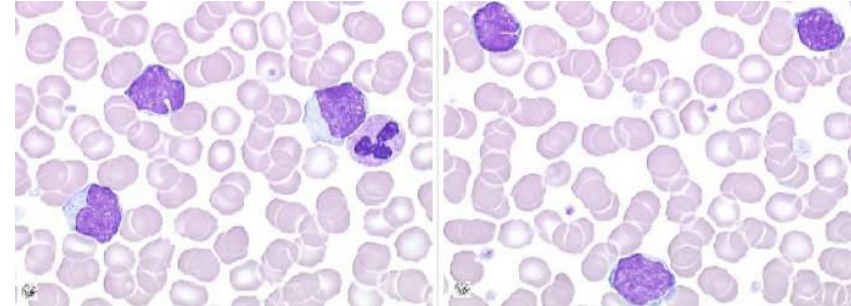
IA et hématologie cellulaire



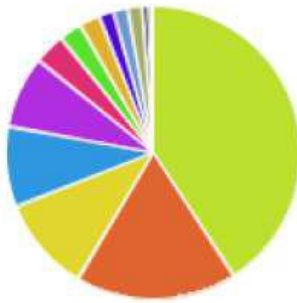
Etapes traditionnelles de machine learning pour la reconnaissance des leucocytes

IA et cytologie

Association de Biologie Praticienne, Campagne
2017 (CYT-17-3-2), Pr M.Zandecki



Nombre de réponses analysées: 852



- Leucémie lymphoïde chronique avec atypies morphologiques (40.7%)
- Dissémination sanguine d'un lymphome à cellules du manteau (17.6%)
- Dissémination sanguine d'un lymphome folliculaire (10.4%)
- Dissémination sanguine d'un lymphome (à petites cellules matures) (8.8%)
- Leucémie lymphoïde chronique (LLC) aspect typique (7.9%)
- Réponse en texte libre (3.4%)
- Hémopathie lymphoïde chronique non classable (2.3%)
- Dissémination sanguine d'un lymphome de la zone marginale à lymphocytes non villeux (2.1%)
- Suspicion de lymphocytose B monoclonale (1.5%)
- LLC forme prolymphocytoïde (1.5%)
- Je ne sais pas (préciser à quel niveau se situe votre problème) (1.5%)
- Lymphocytose réactionnelle (0.5%)
- Dissémination sanguine d'un lymphome à grandes cellules (0.5%)
- Leucémie prolymphocytaire (0.4%)
- Dissémination sanguine d'un lymphome splénique à lymphocytes villeux (0.2%)
- Aspect évoquant une leucémie aiguë monoblastique (LAM5 - FAB) (0.1%)
- Aspect évoquant une leucémie aiguë lymphoblastique (0.1%)
- Suspicion de macroglobulinémie de Waldenström (0.1%)
- Syndrome de Sézary (0.1%)
- Leucémie à tricholeucocytes (0.1%)

IA et cytologie

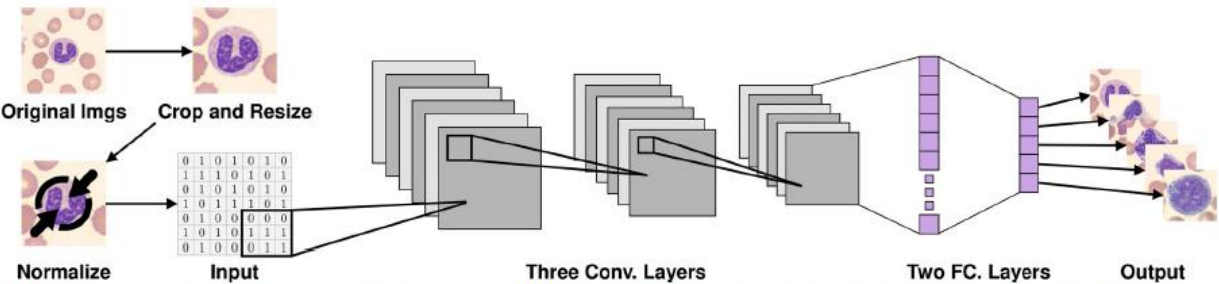


Fig. 2 An overview of the pre-processing and the proposed CNN-based architecture for WBC image classification. The pre-processing consists of cropping, re-sizing and normalizing. Three convolutional layers (including three pooling layers) are in charge of extracting and learning features, and two fully connected layers are in charge of classification

	Original Images					Synthesized Images				
Neutrophil										
Eosinophil										
Basophil										
Lymphocyte										
Monocyte										

	NE. (%)	EO. (%)	BA. (%)	LY. (%)	MO. (%)
Fold-0	100	95	92	99	96
Fold-1	98	99	94	100	100
Fold-2	96	93	100	95	98
Fold-3	97	99	100	95	96
Fold-4	100	100	97	98	97
Fold-5	100	98	94	97	98
Fold-6	100	98	94	97	91
Fold-7	95	98	94	98	96
Fold-8	100	93	94	97	99
Fold-9	98	100	91	95	97
Avr. Acc.	98	97	95	97	97

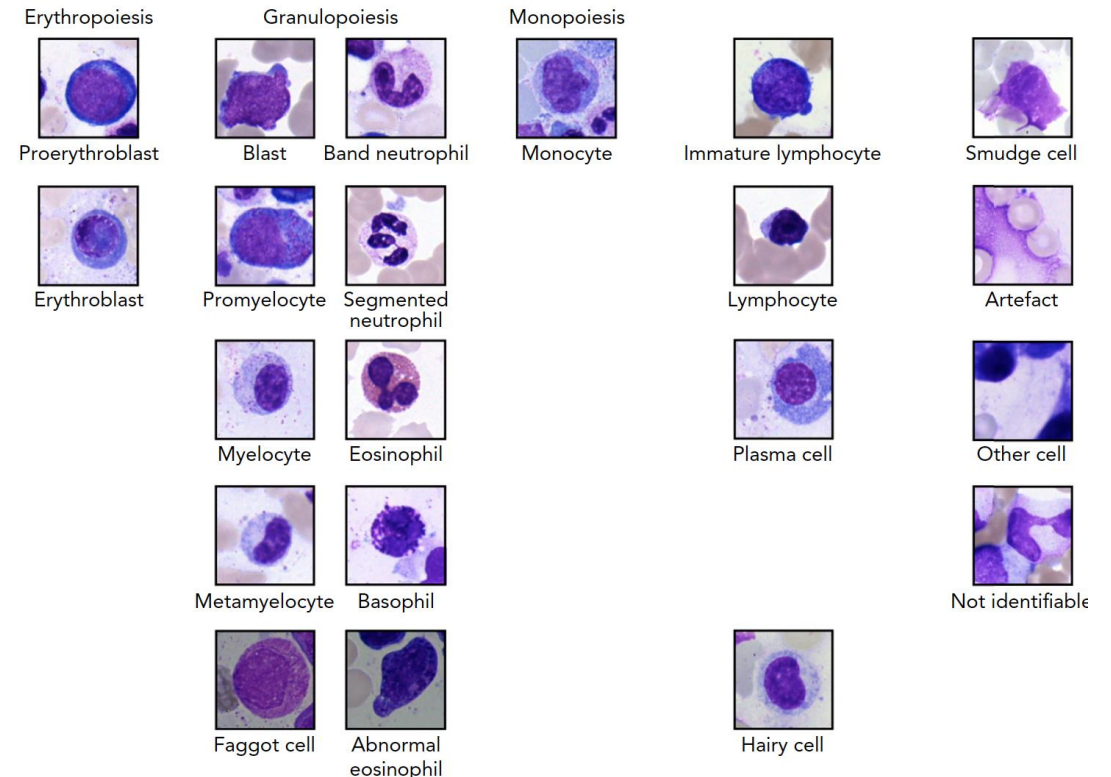
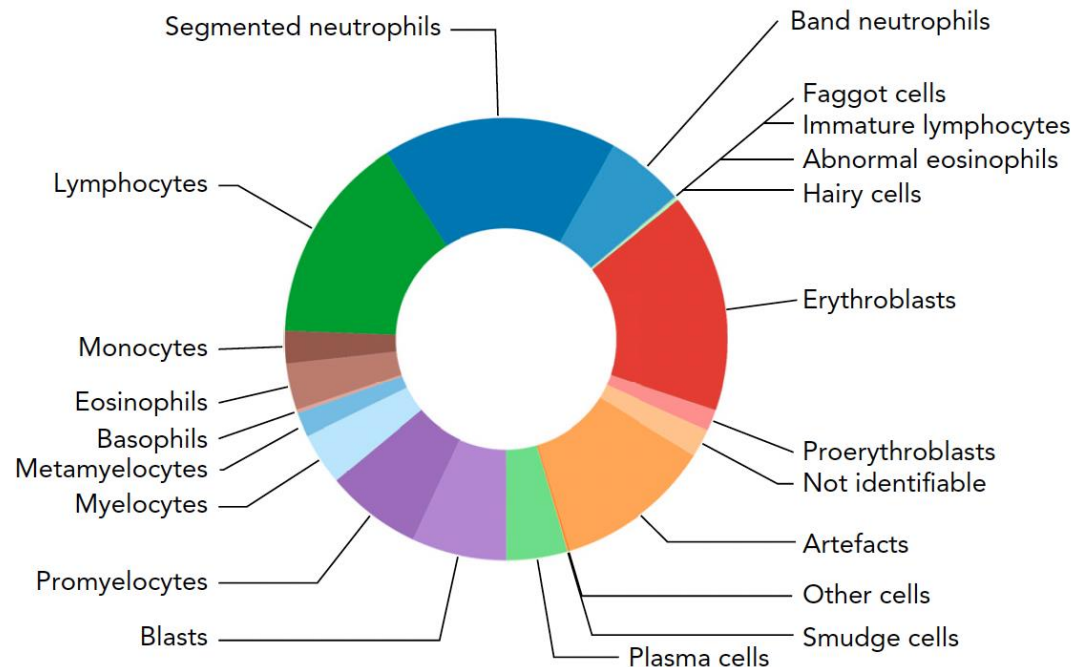
The average accuracy for five classes is 97%

Utilisation CNN pour génération d’images pour leur
algorithme d’IA

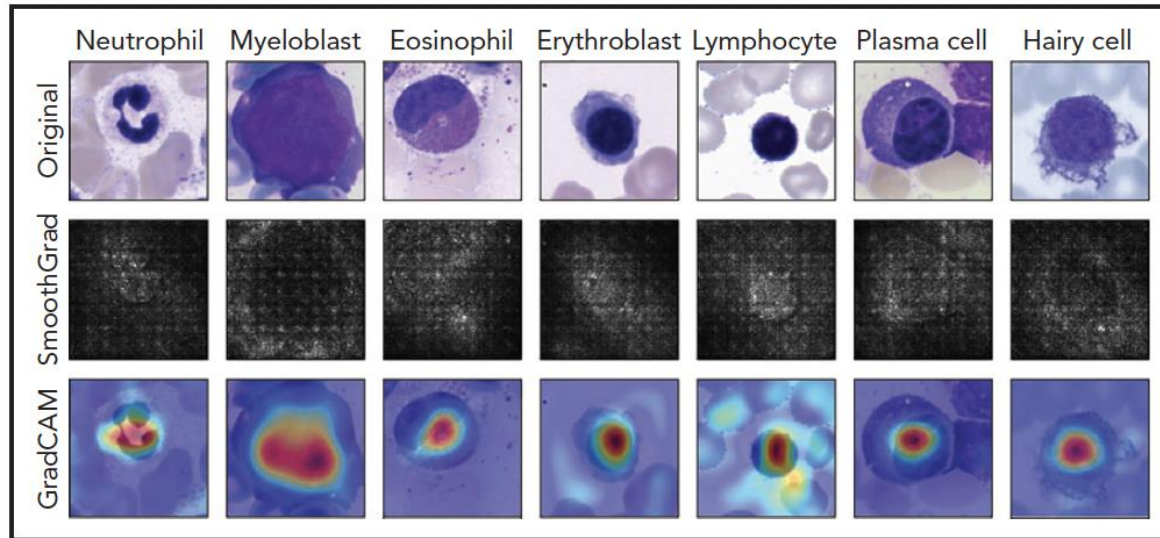
IA et cytologie médullaire

Highly accurate differentiation of bone marrow cell morphologies using deep neural networks on a large image data set

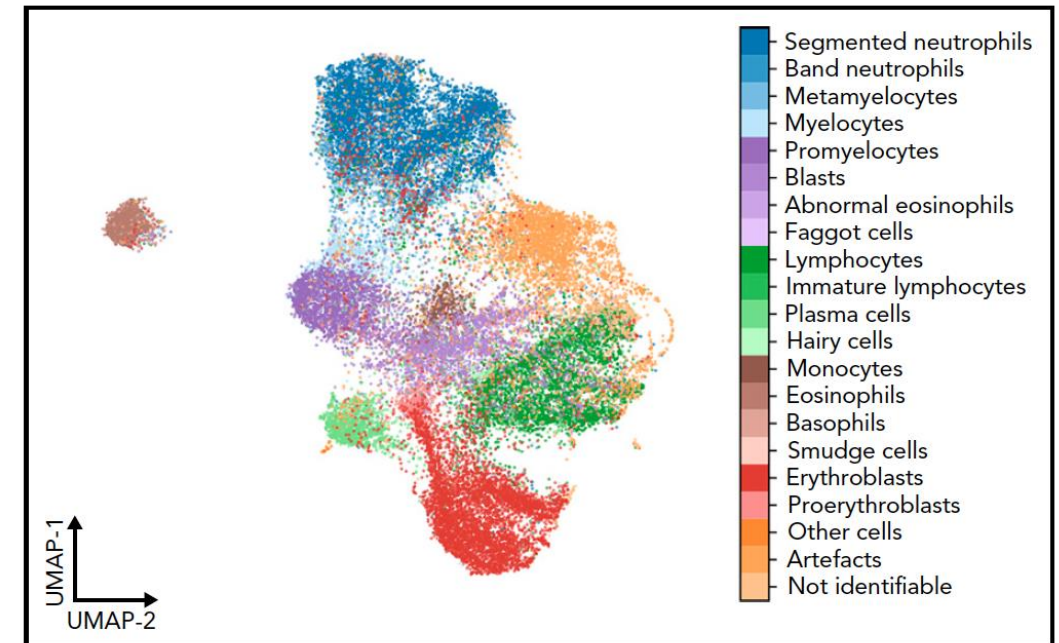
21 classes morphologiques, > 170000 images



IA et cytologie médullaire

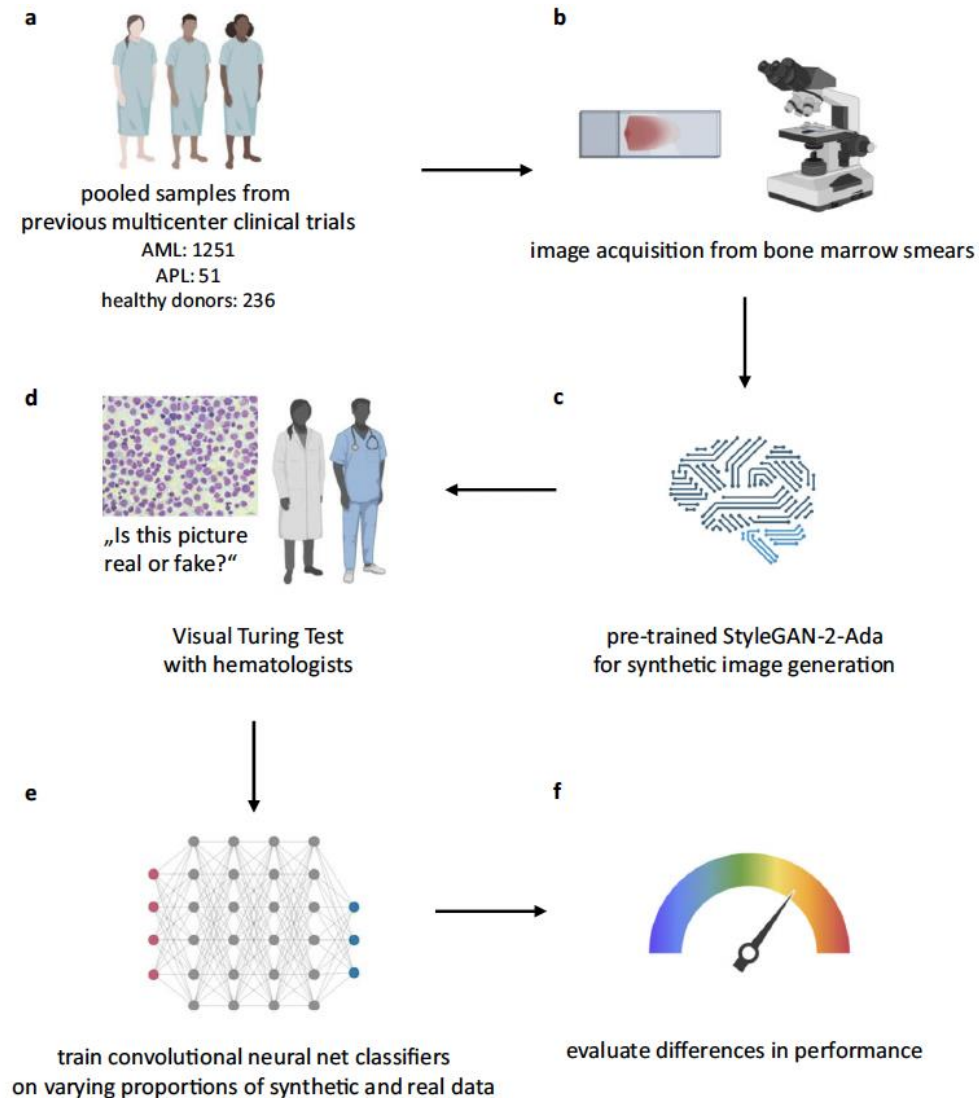


Heatmap des régions d'intérêt ayant permis
l'identification des cellules



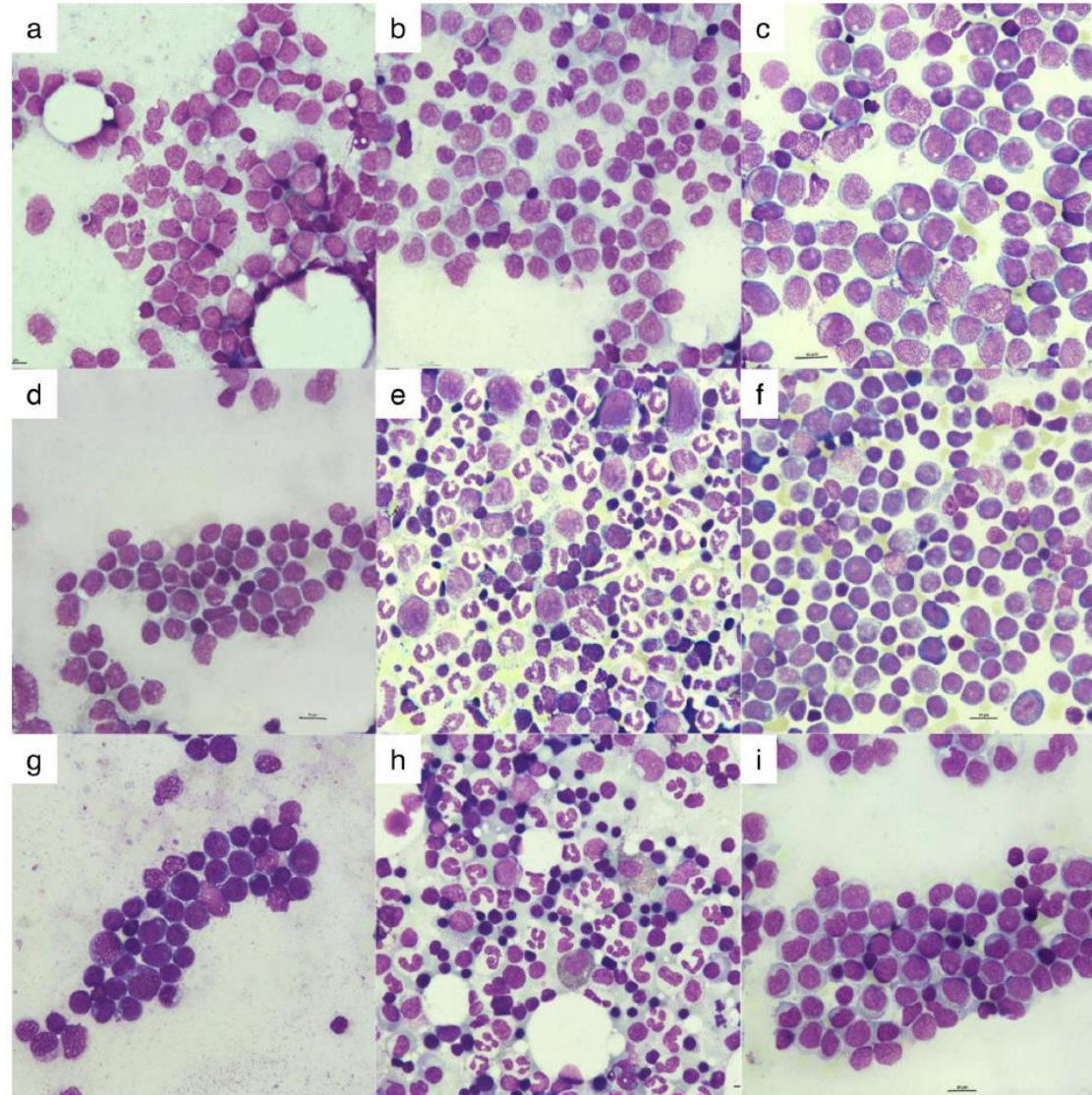
Réduction de dimensionnalité de type UMAP

IA et cytologie médullaire

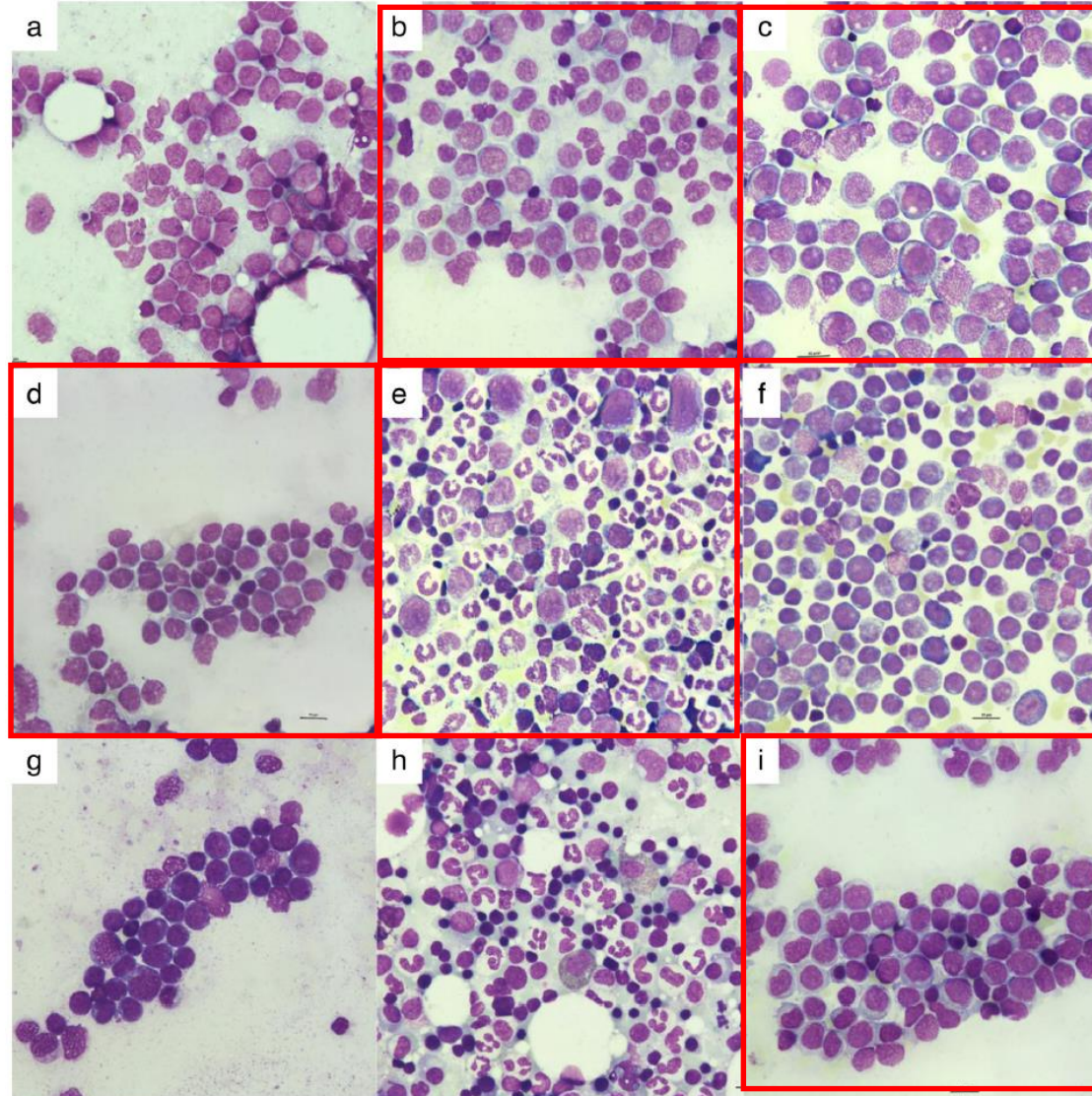


Synthetic bone marrow images augment real samples in developing acute myeloid leukemia microscopy classification models

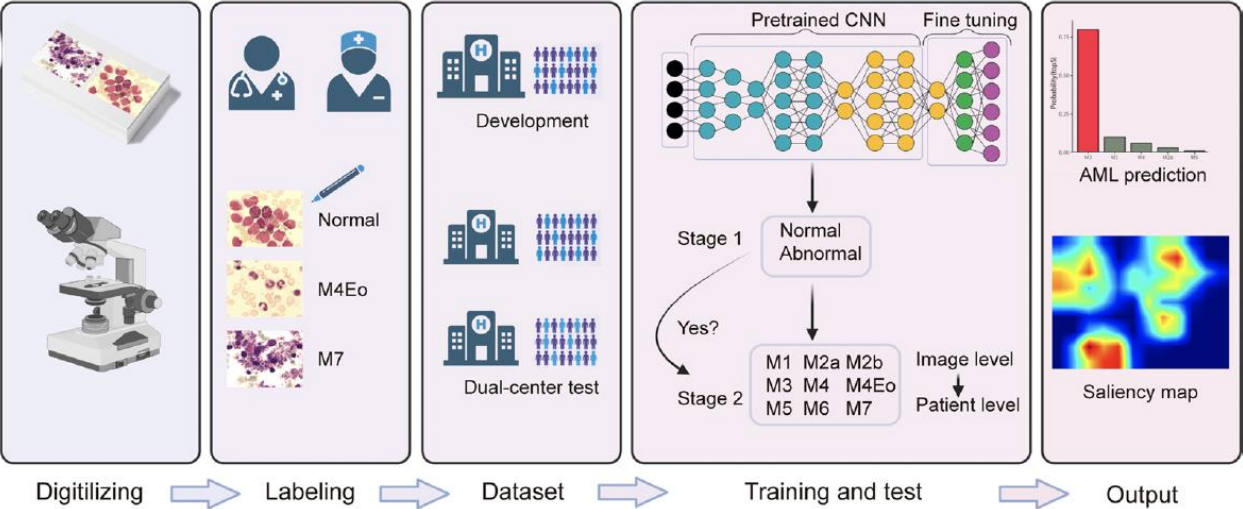
IA et cytologie médullaire



IA et cytologie médullaire

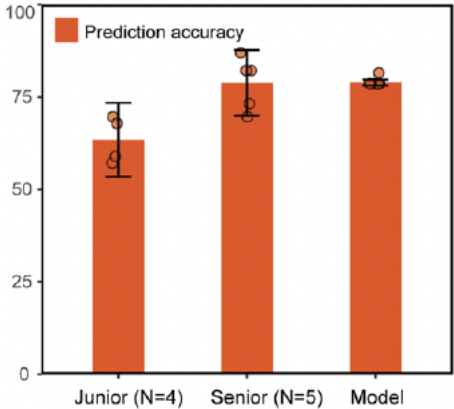


IA et LAM

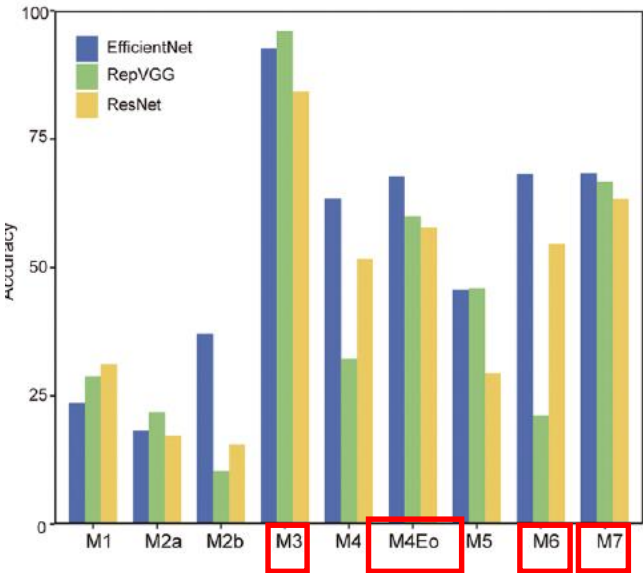


AMLnet (2 modules)

Classification + heatmap régions d'intérêt

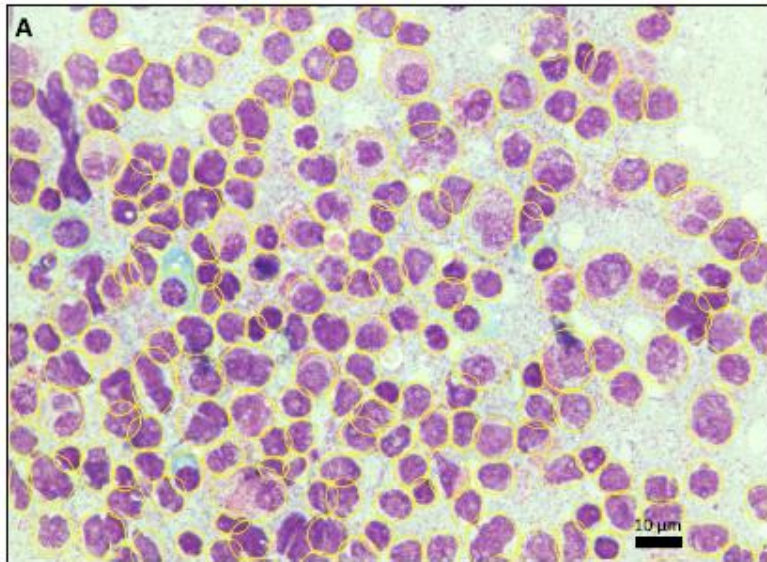


Performances comparables à celles des pathologistes séniors (mais mieux que les juniors!)



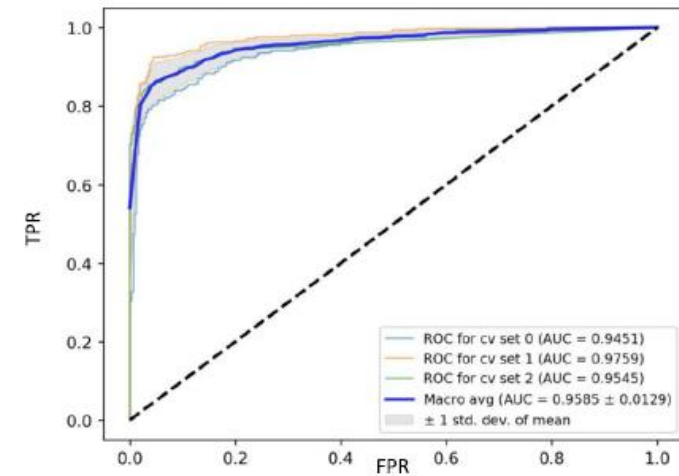
IA et LAM3

Deep Learning

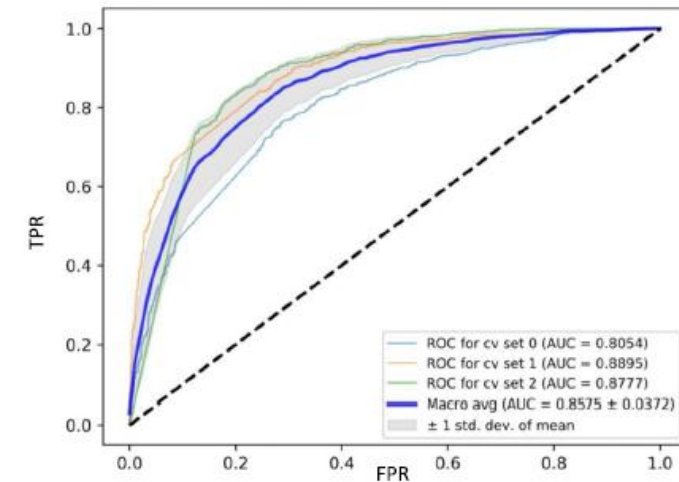


51 moelles LAP
1048 moelles LAM non LAP
236 moelles donneurs sains

LAP vs MO normales

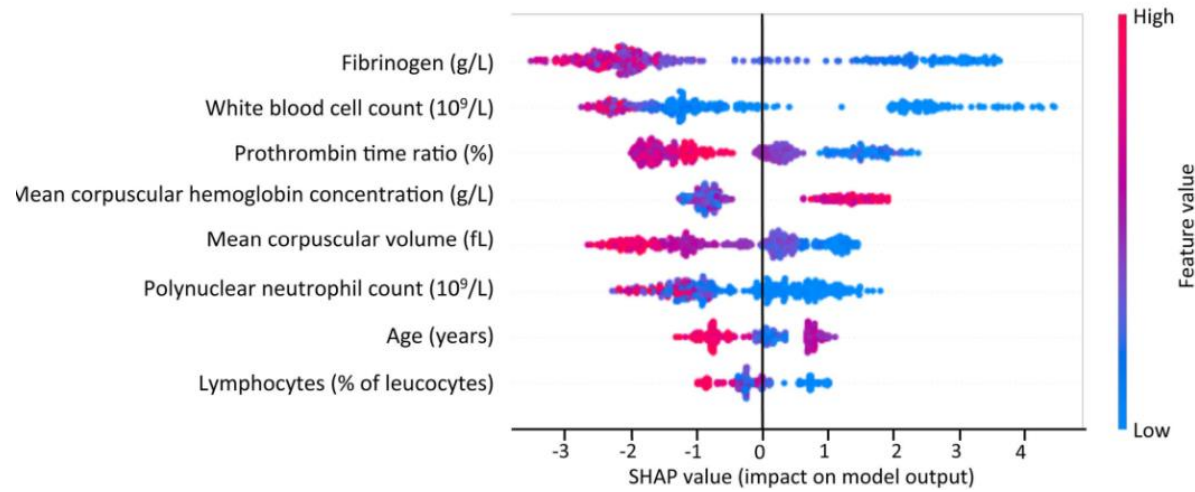
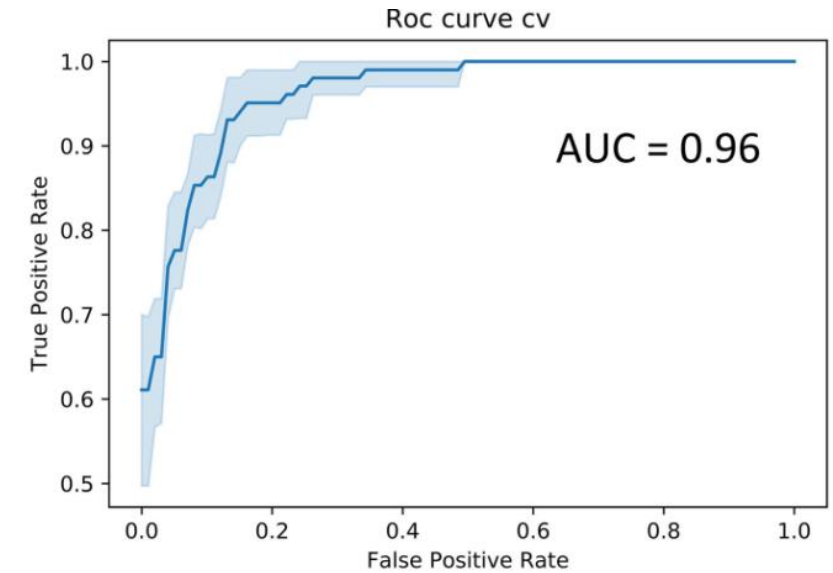
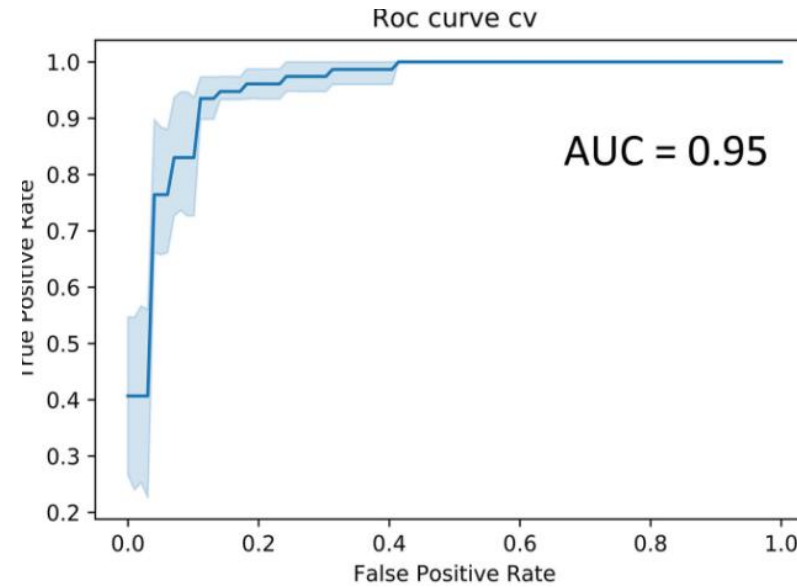


LAP vs LAM non LAP



IA et LAM3

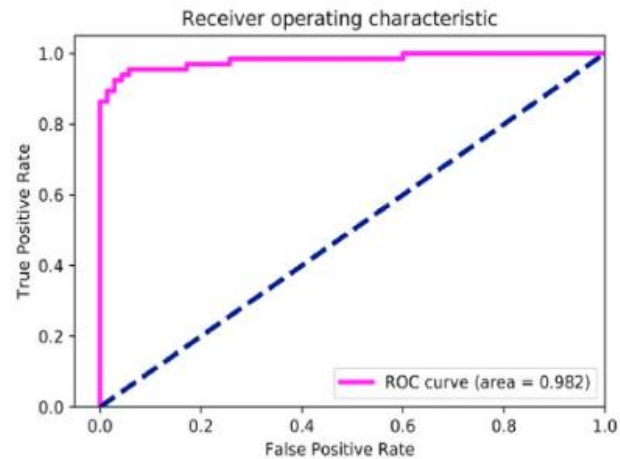
Algorithme XGBoost (ML)



https://share.streamlit.io/nico-facto/leukemia-apl-classification/main/Leucemie_app.py

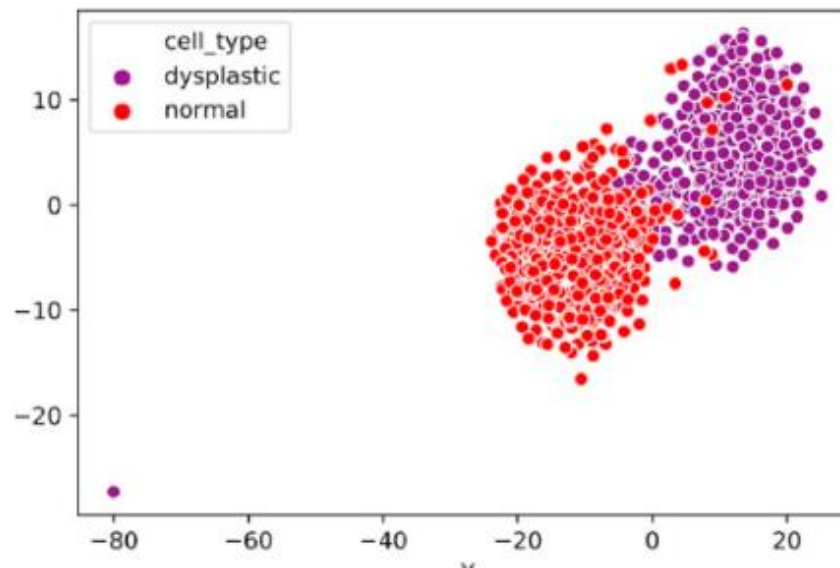
IA et syndromes myélodysplasiques

Algorithme DysplasiaNet (CNN)



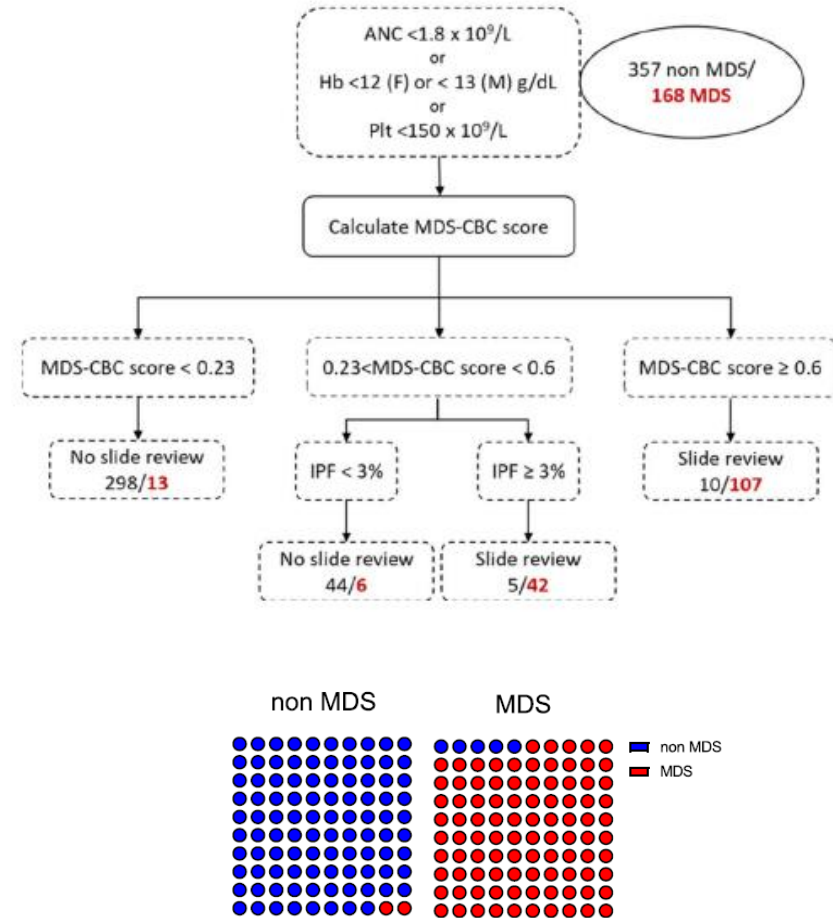
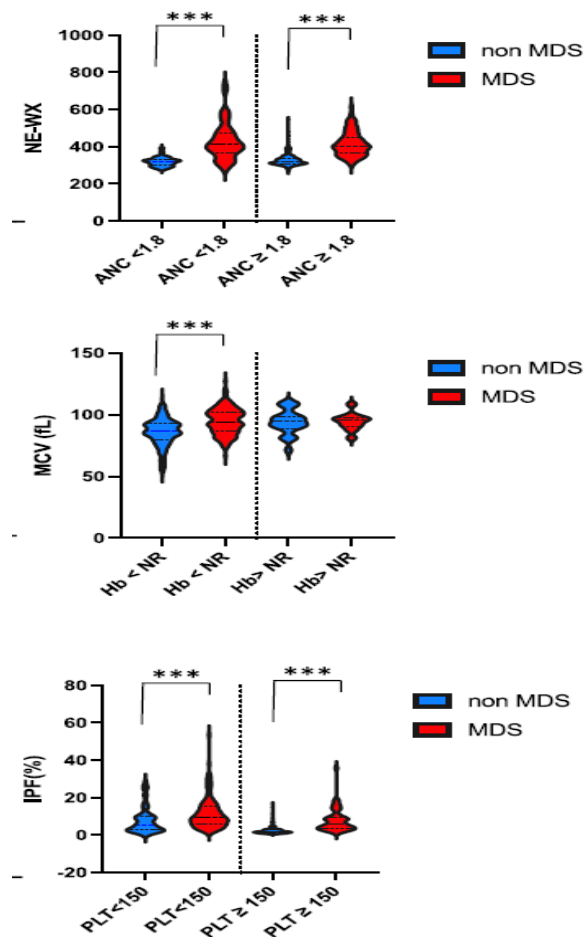
		Predicted	
		Dysplastic	Normal
True	Dysplastic	63 (95.5 %)	3 (4.5 %)
	Normal	4 (5.7 %)	66 (94.3 %)

t-SNE avec réduction de dimensionnalité



IA et syndromes myélodysplasiques

Random Forest + CART (ML)



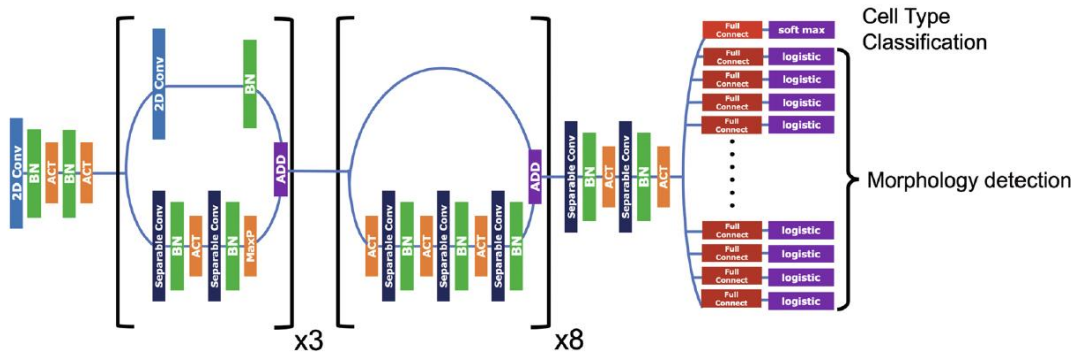
98% des Non MDS et 95% des MDS
correctement classés

IA et syndromes myélodysplasiques

Deep Learning + XG Boost

17 types cellulaires

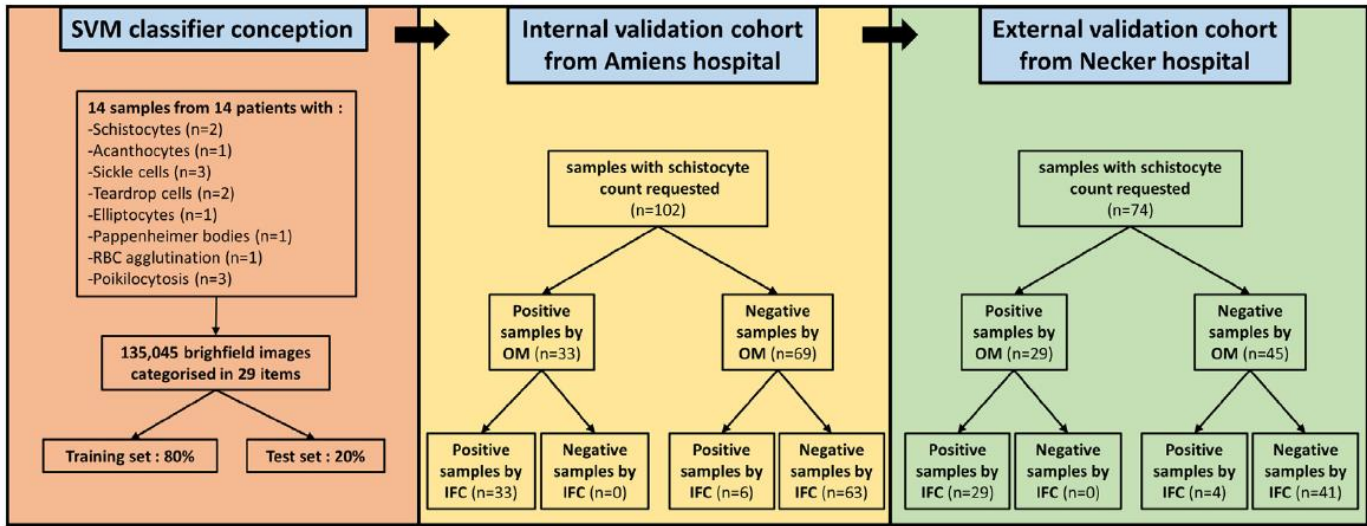
97 anomalies morphologiques



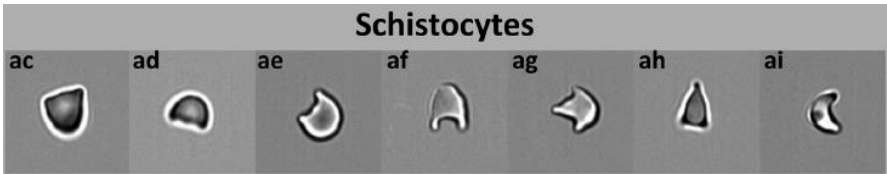
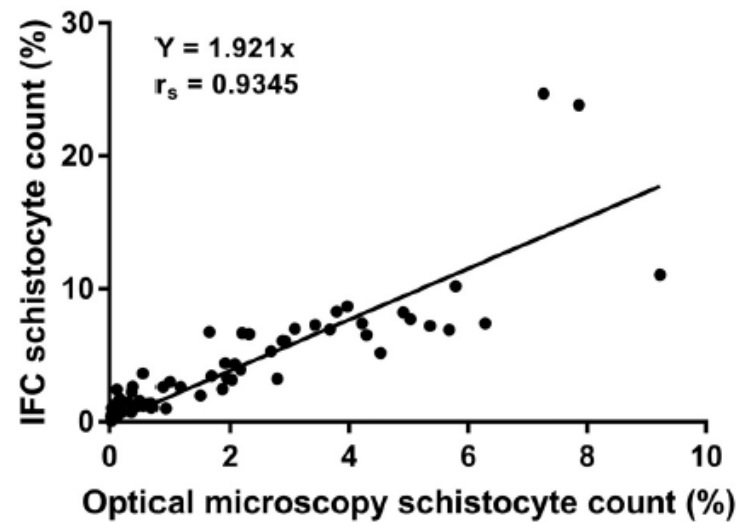
Elaboration algorithme pour différencier MDS et AA
(Se=96% et Sp=100%)



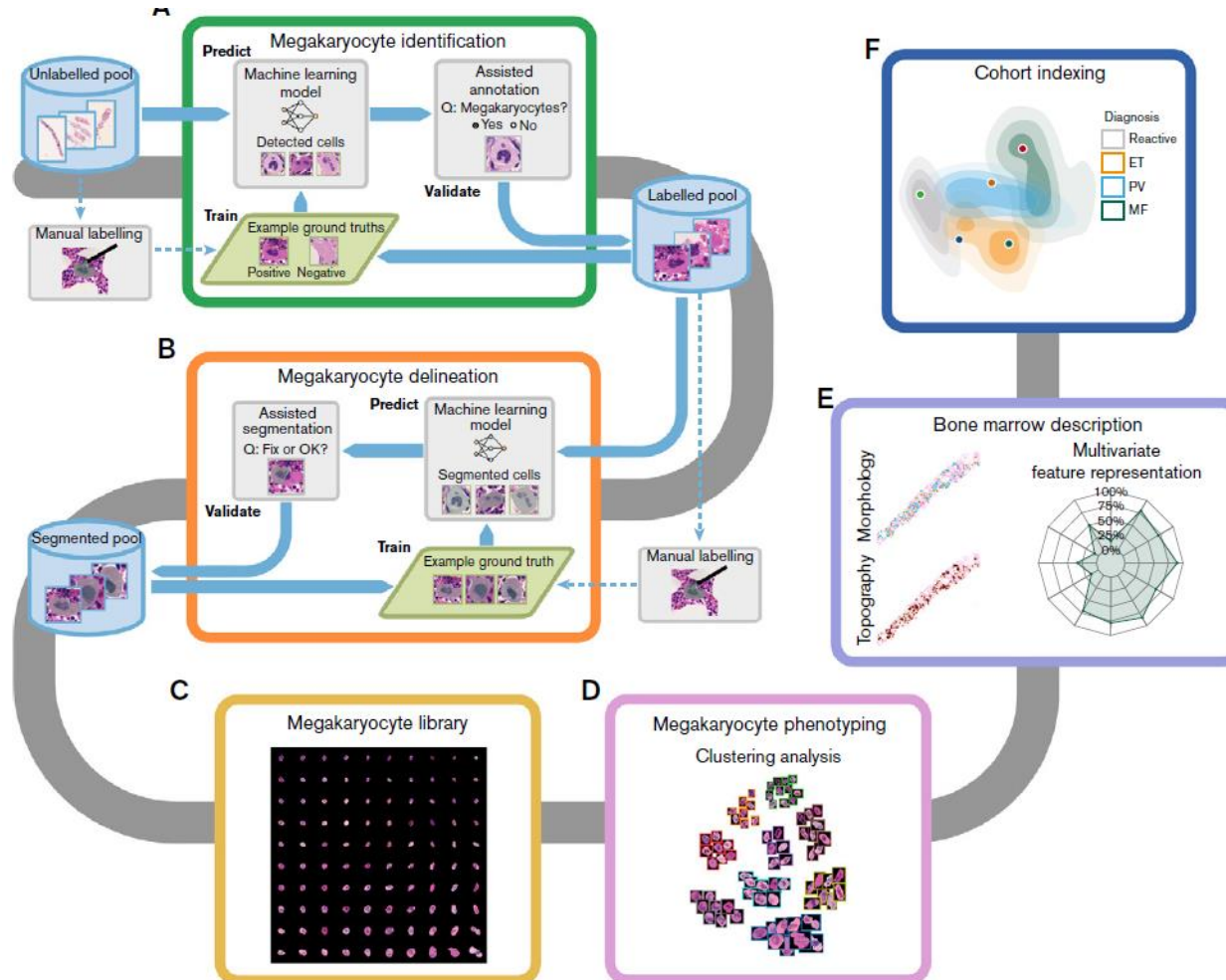
L'IA ne voit pas rouge?



SVM



IA et syndromes myéloprolifératifs



A partir d'images de BOM

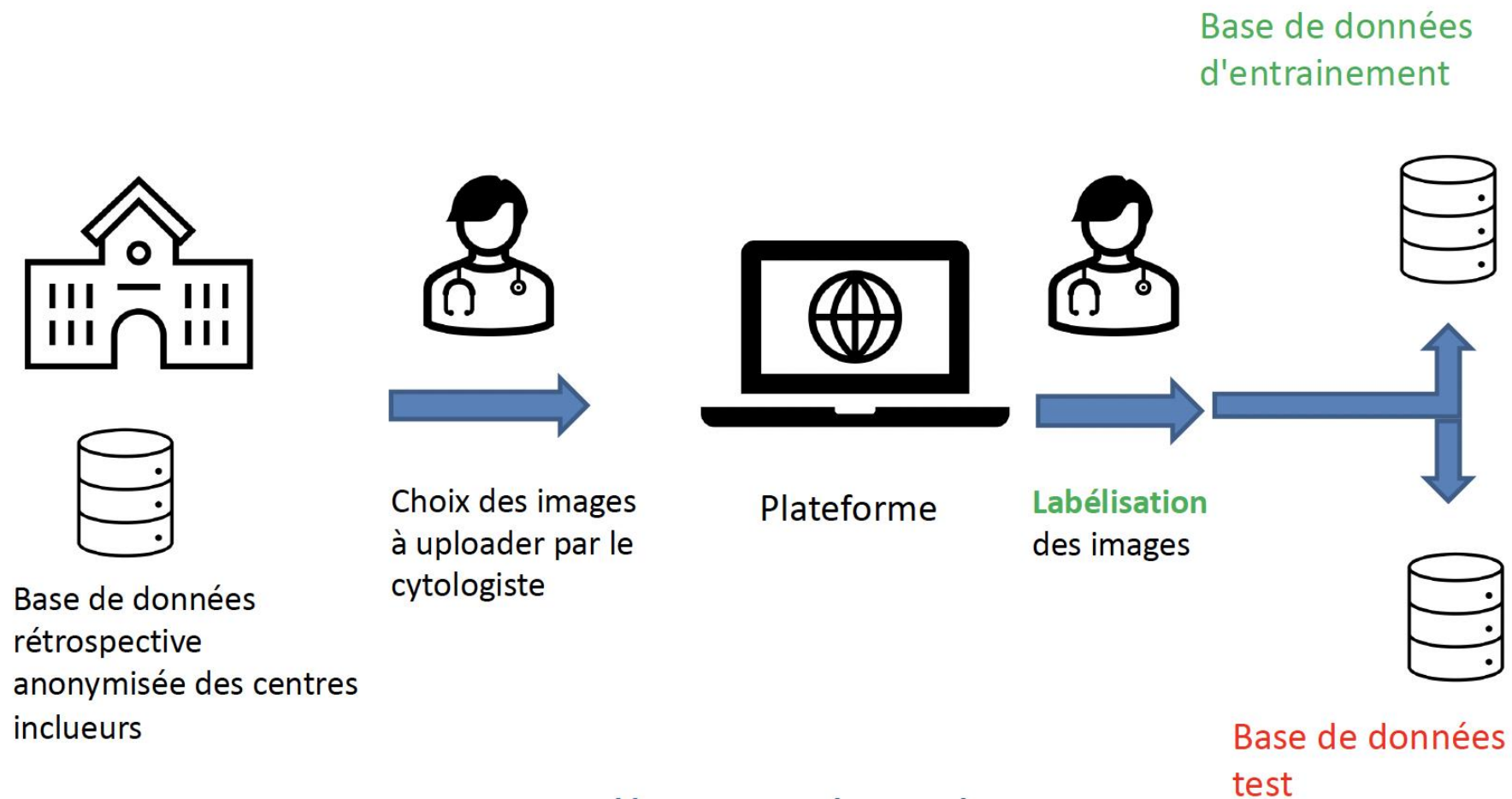
Random Forest classifier

AUC : 0,95

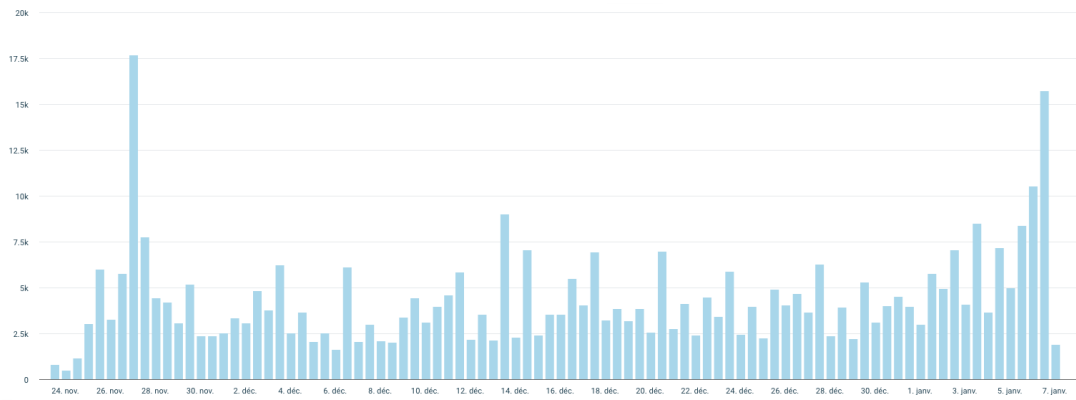
Datachallenge Cytologia



Datachallenge Cytologia

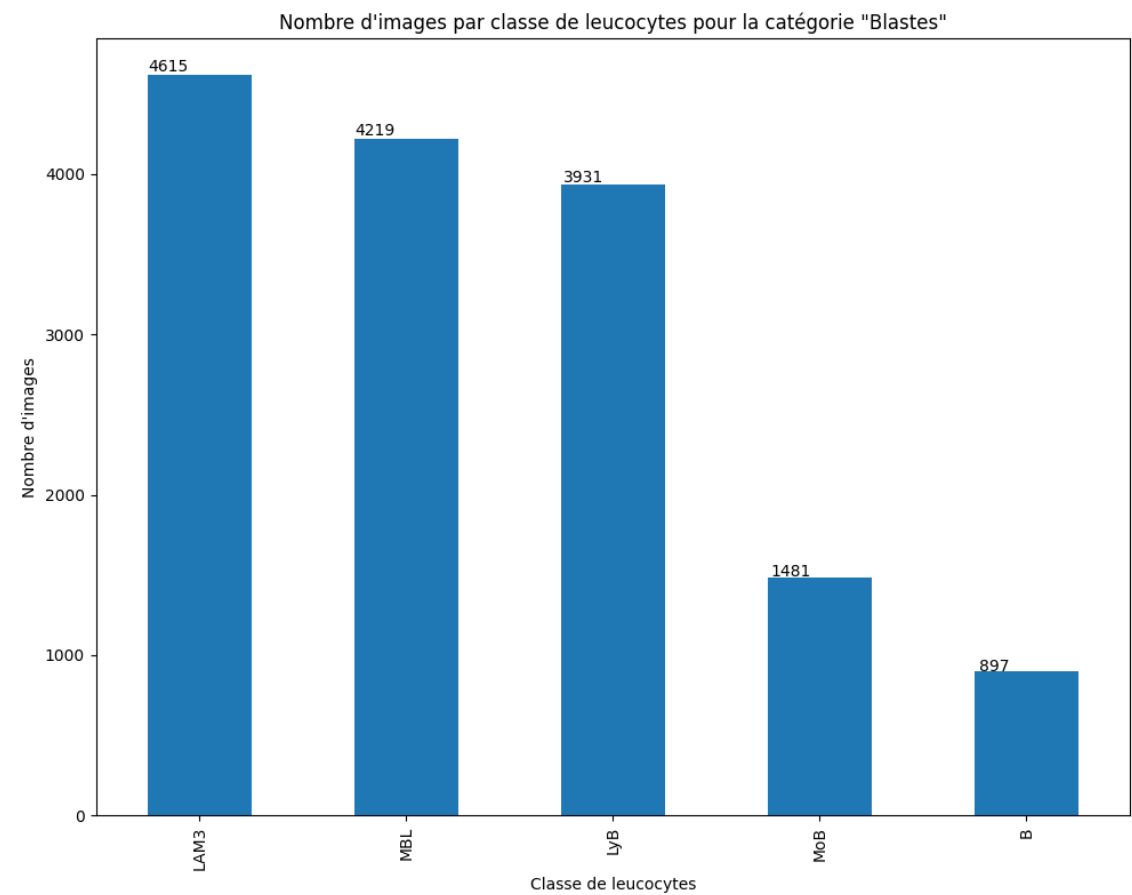
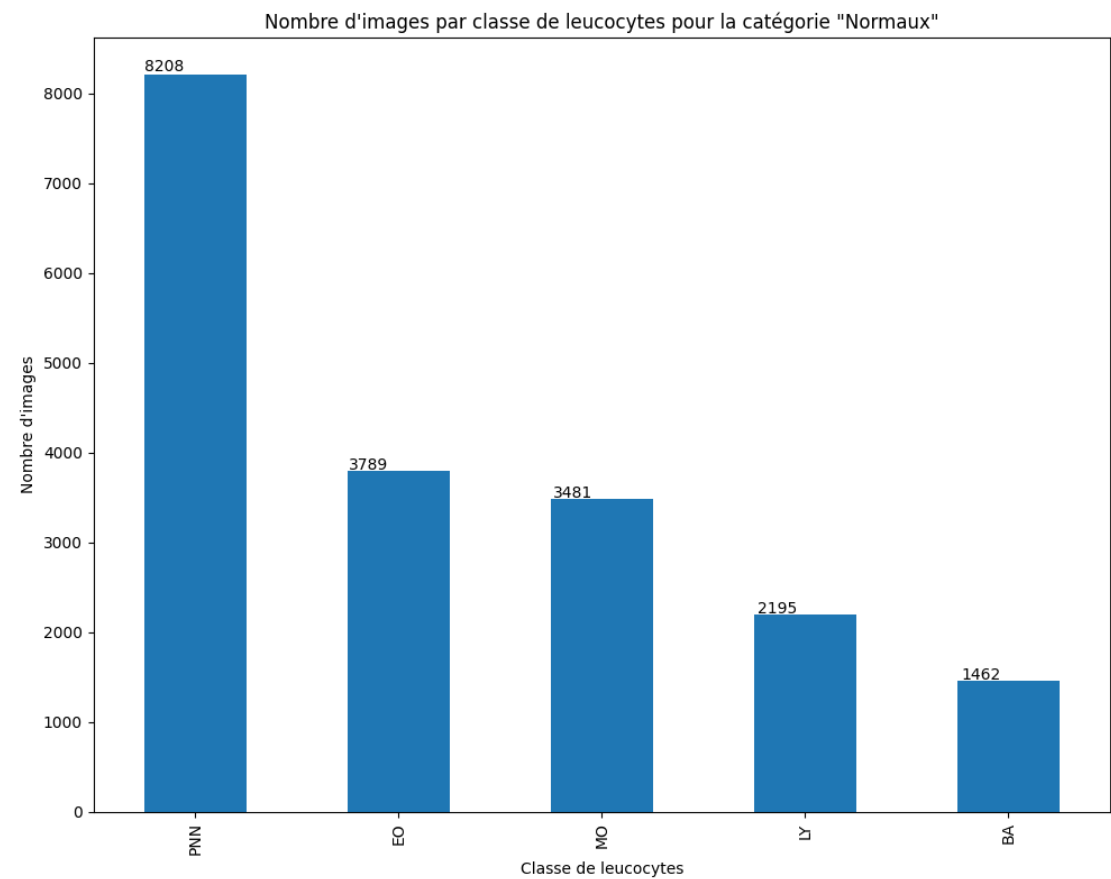


Datachallenge Cytologia

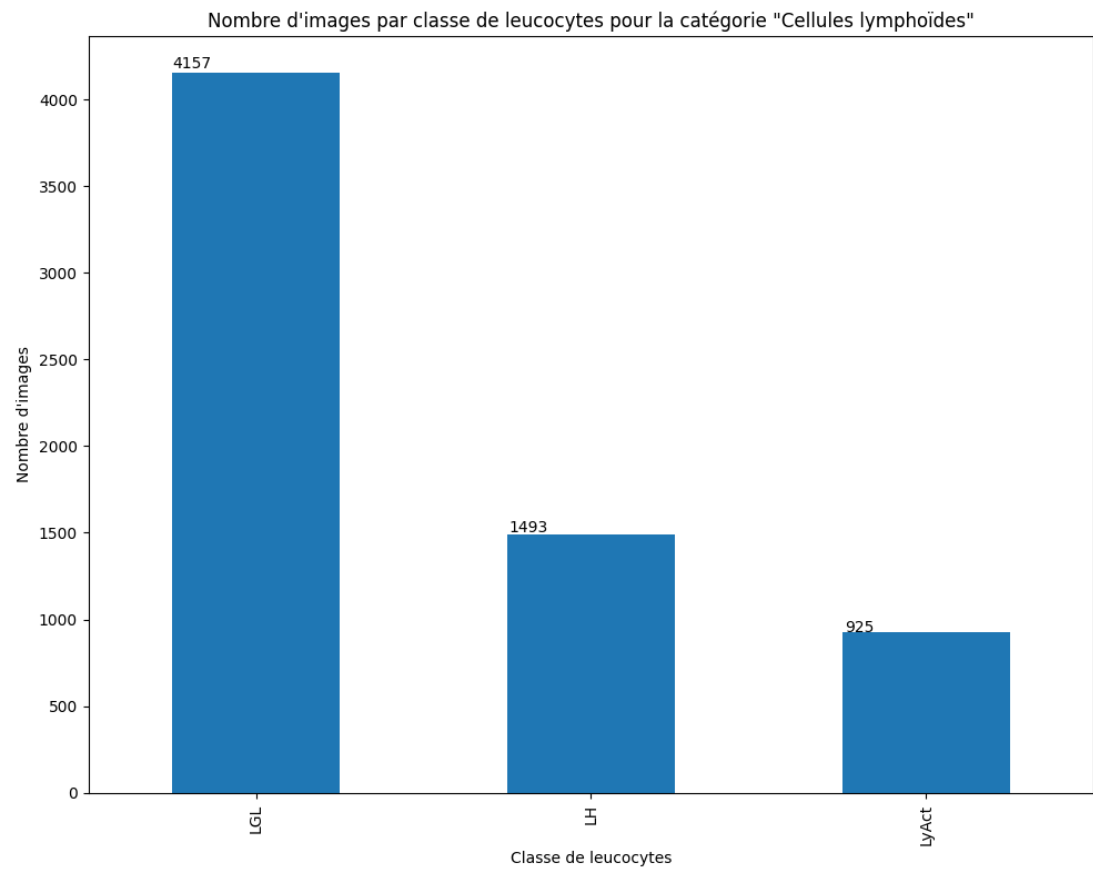
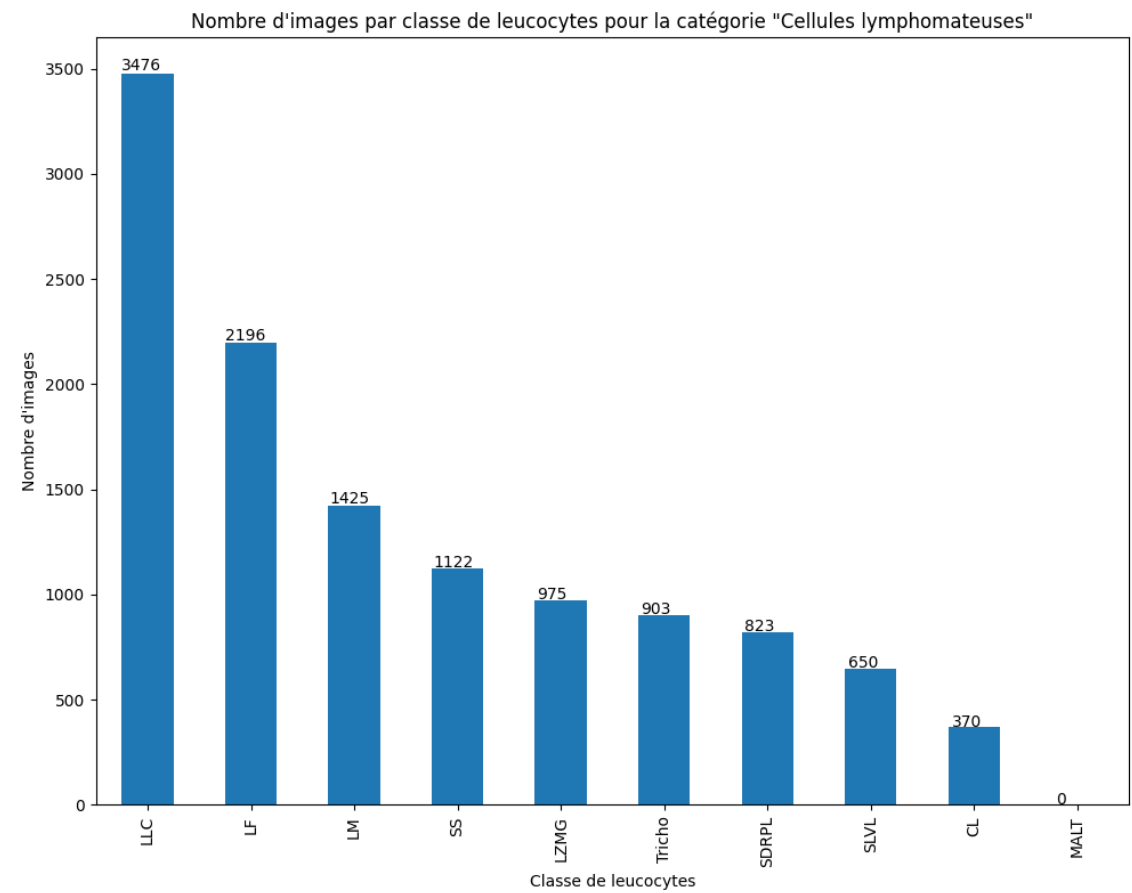


- 6 semaines : du 29/11/2024 au 06/01/2025
- **245** participants
- **69168 images** réparties en **23 classes**
- Base de données en **Open Data**
- Algorithmes d'IA en **Open Source**
- **25000€** à partager entre les trois premiers

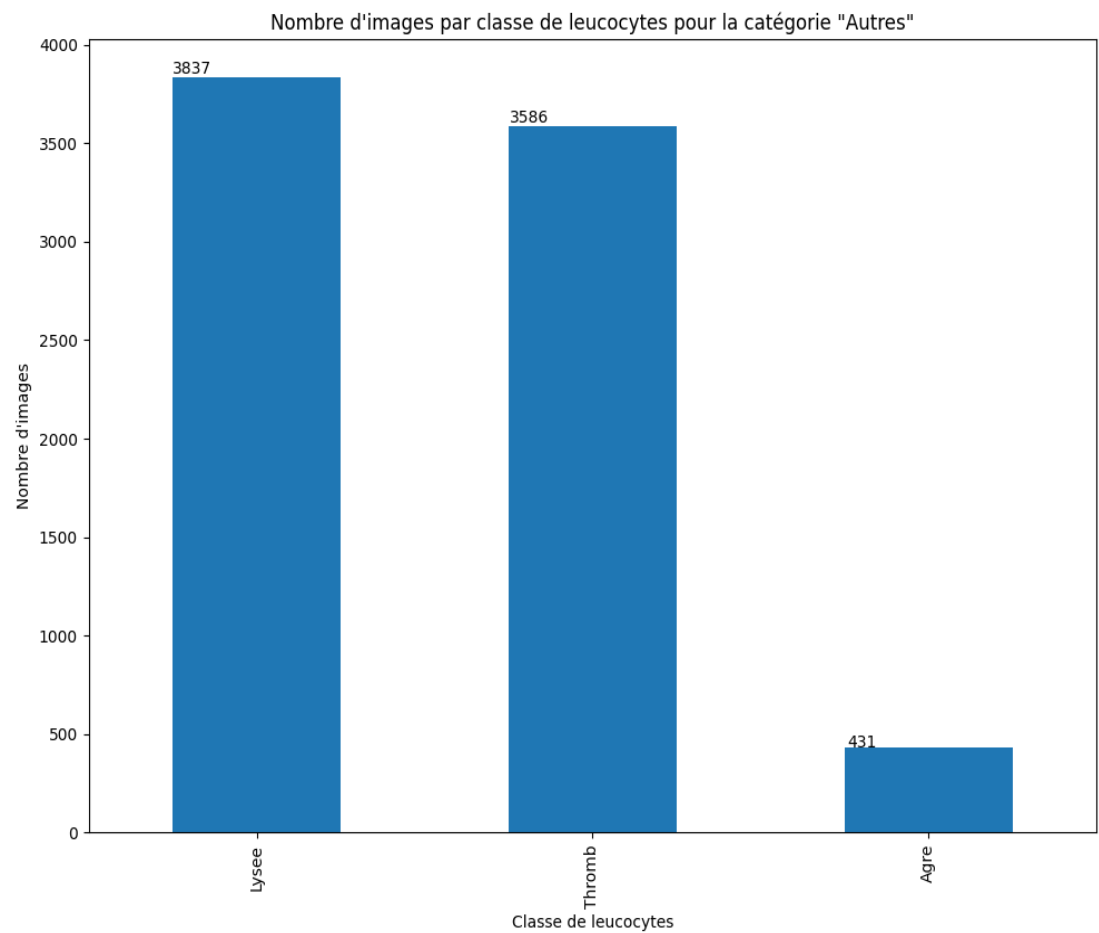
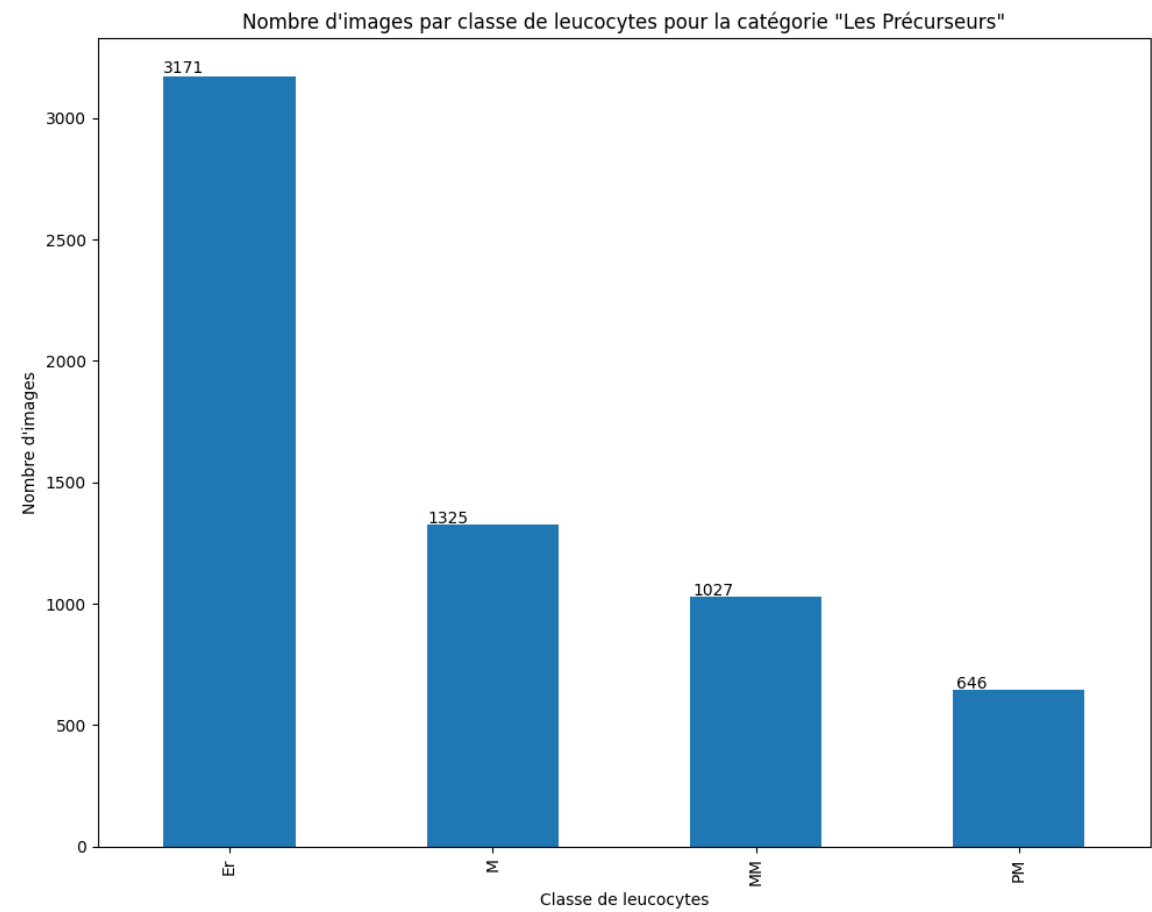
Datachallenge Cytologia : classes de cellules



Datachallenge Cytologia : classes de cellules

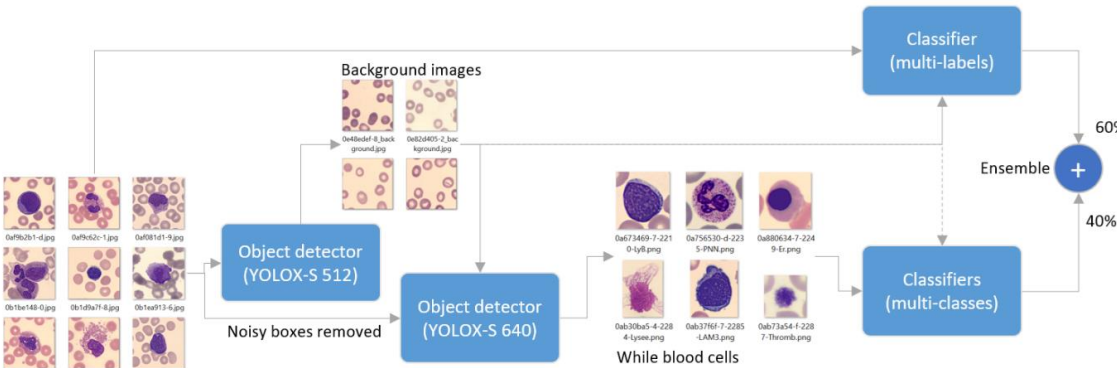
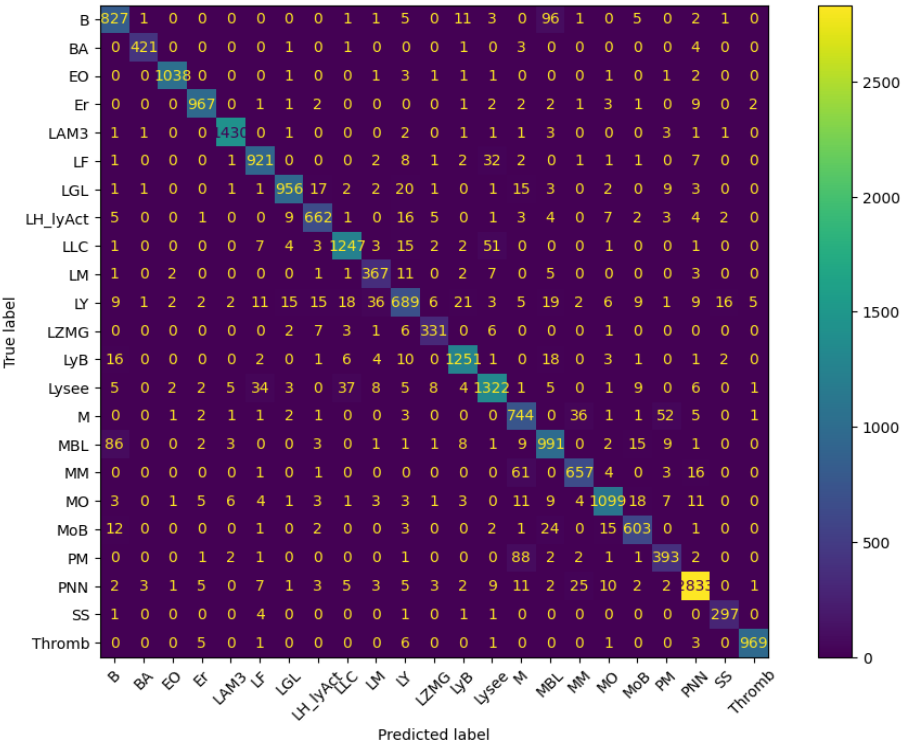


Datachallenge Cytologia : classes de cellules



Les résultats!

	precision	recall	f1-score	support
B	0.85	0.87	0.86	954
BA	0.98	0.98	0.98	431
EO	0.99	0.99	0.99	1050
Er	0.97	0.97	0.97	994
LAM3	0.99	0.99	0.99	1446
LF	0.92	0.94	0.93	980
LGL	0.96	0.92	0.94	1035
LH_lyAct	0.92	0.91	0.92	725
LLC	0.94	0.93	0.94	1337
LM	0.85	0.92	0.88	400
LY	0.85	0.76	0.80	902
LZMG	0.92	0.93	0.92	357
LyB	0.95	0.95	0.95	1316
Lysee	0.91	0.91	0.91	1458
M	0.78	0.87	0.82	851
MBL	0.84	0.87	0.86	1133
MM	0.90	0.88	0.89	743
MO	0.95	0.92	0.93	1193
MoB	0.90	0.91	0.91	664
PM	0.81	0.80	0.80	494
PNN	0.97	0.97	0.97	2935
SS	0.93	0.97	0.95	305
Thromb	0.99	0.98	0.99	986
accuracy			0.93	22689
macro avg	0.92	0.92	0.92	22689
weighted avg	0.93	0.93	0.93	22689

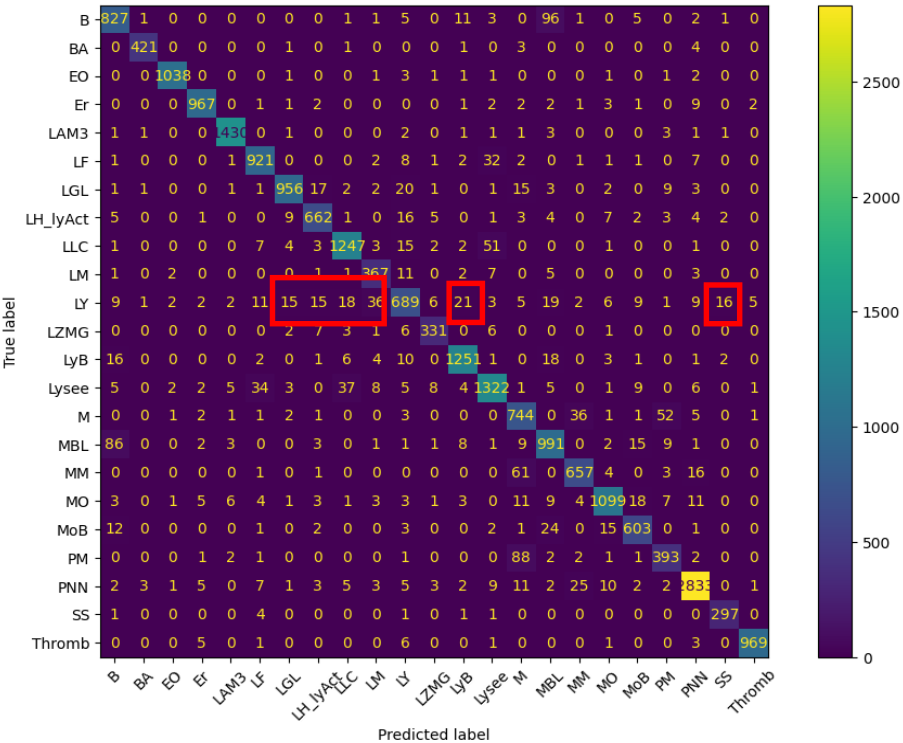


1^{er} prix : MPWARE
f1 score : 0,9371

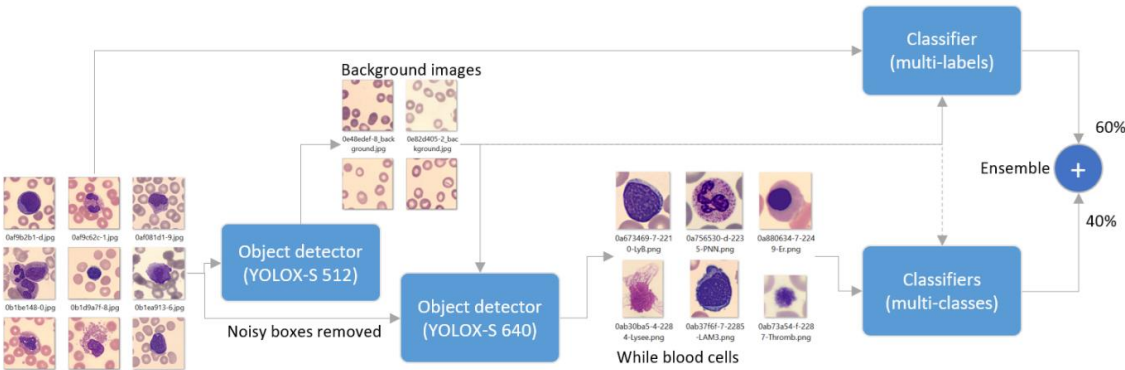
<https://github.com/MPWARE-TEAM/Cytologia/tree/main>

Les résultats!

	precision	recall	f1-score	support
B	0.85	0.87	0.86	954
BA	0.98	0.98	0.98	431
EO	0.99	0.99	0.99	1050
Er	0.97	0.97	0.97	994
LAM3	0.99	0.99	0.99	1446
LF	0.92	0.94	0.93	980
LGL	0.96	0.92	0.94	1035
LH_lyAct	0.92	0.91	0.92	725
LLC	0.94	0.93	0.94	1337
LM	0.85	0.92	0.88	400
LY	0.85	0.76	0.80	902
LZMG	0.92	0.93	0.92	357
LyB	0.95	0.95	0.95	1316
Lysee	0.91	0.91	0.91	1458
M	0.78	0.87	0.82	851
MBL	0.84	0.87	0.86	1133
MM	0.90	0.88	0.89	743
MO	0.95	0.92	0.93	1193
MoB	0.90	0.91	0.91	664
PM	0.81	0.80	0.80	494
PNN	0.97	0.97	0.97	2935
SS	0.93	0.97	0.95	305
Thromb	0.99	0.98	0.99	986
accuracy			0.93	22689
macro avg	0.92	0.92	0.92	22689
weighted avg	0.93	0.93	0.93	22689

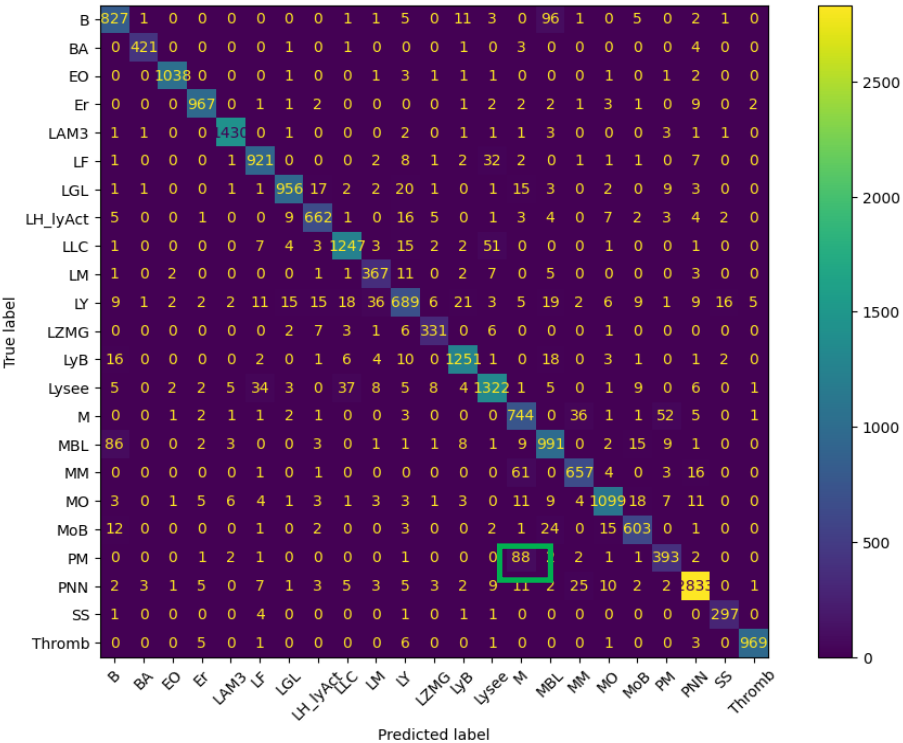


<https://github.com/MPWARE-TEAM/Cytologia/tree/main>

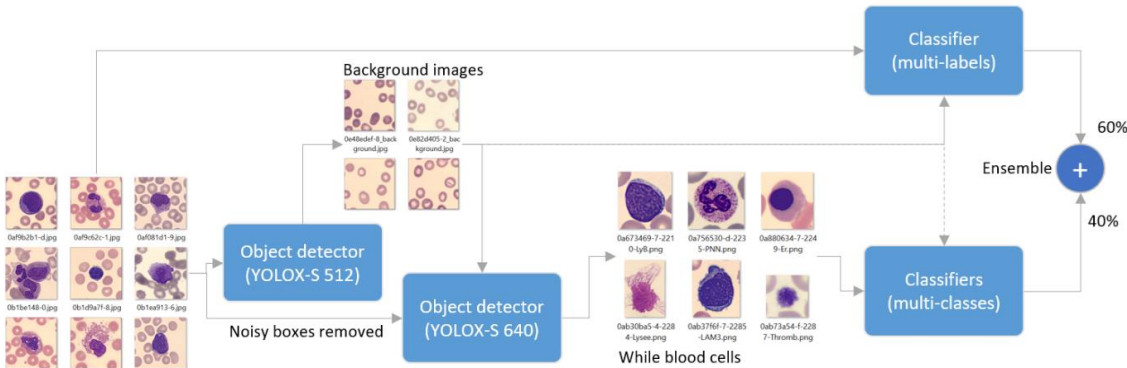


Les résultats!

	precision	recall	f1-score	support
B	0.85	0.87	0.86	954
BA	0.98	0.98	0.98	431
EO	0.99	0.99	0.99	1050
Er	0.97	0.97	0.97	994
LAM3	0.99	0.99	0.99	1446
LF	0.92	0.94	0.93	980
LGL	0.96	0.92	0.94	1035
LH_lyAct	0.92	0.91	0.92	725
LLC	0.94	0.93	0.94	1337
LM	0.85	0.92	0.88	400
LY	0.85	0.76	0.80	902
LZMG	0.92	0.93	0.92	357
LyB	0.95	0.95	0.95	1316
Lysee	0.91	0.91	0.91	1458
M	0.78	0.87	0.82	851
MBL	0.84	0.87	0.86	1133
MM	0.90	0.88	0.89	743
MO	0.95	0.92	0.93	1193
MoB	0.90	0.91	0.91	664
PM	0.81	0.80	0.80	494
PNN	0.97	0.97	0.97	2935
SS	0.93	0.97	0.95	305
Thromb	0.99	0.98	0.99	986
accuracy			0.93	22689
macro avg	0.92	0.92	0.92	22689
weighted avg	0.93	0.93	0.93	22689

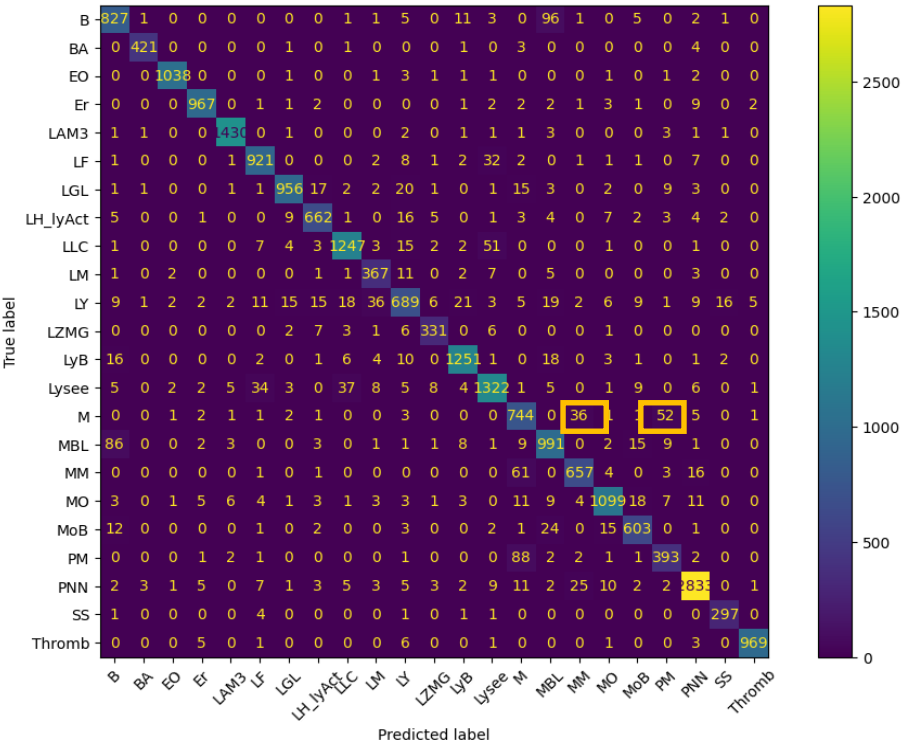


<https://github.com/MPWARE-TEAM/Cytologia/tree/main>

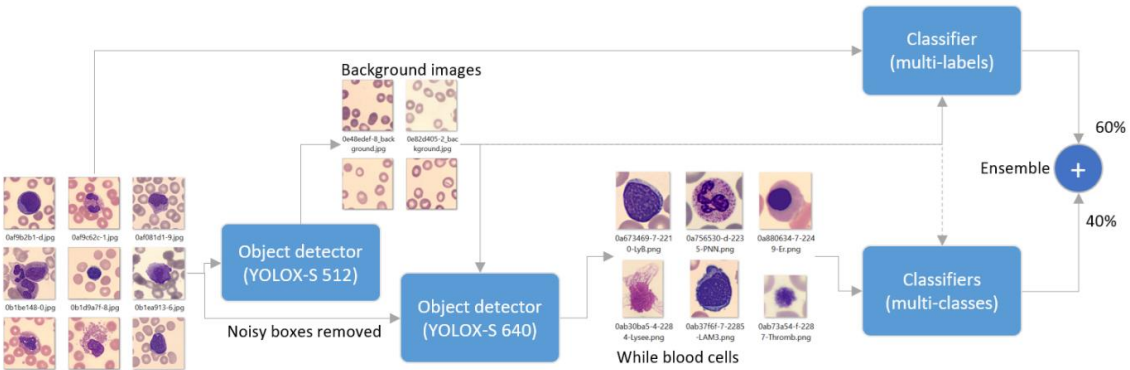


Les résultats!

	precision	recall	f1-score	support
B	0.85	0.87	0.86	954
BA	0.98	0.98	0.98	431
EO	0.99	0.99	0.99	1050
Er	0.97	0.97	0.97	994
LAM3	0.99	0.99	0.99	1446
LF	0.92	0.94	0.93	980
LGL	0.96	0.92	0.94	1035
LH_lyAct	0.92	0.91	0.92	725
LLC	0.94	0.93	0.94	1337
LM	0.85	0.92	0.88	400
LY	0.85	0.76	0.80	902
LZMG	0.92	0.93	0.92	357
LyB	0.95	0.95	0.95	1316
Lysee	0.91	0.91	0.91	1458
M	0.78	0.87	0.82	851
MBL	0.84	0.87	0.86	1133
MM	0.90	0.88	0.89	743
MO	0.95	0.92	0.93	1193
MoB	0.90	0.91	0.91	664
PM	0.81	0.80	0.80	494
PNN	0.97	0.97	0.97	2935
SS	0.93	0.97	0.95	305
Thromb	0.99	0.98	0.99	986
accuracy			0.93	22689
macro avg	0.92	0.92	0.92	22689
weighted avg	0.93	0.93	0.93	22689



<https://github.com/MPWARE-TEAM/Cytologia/tree/main>

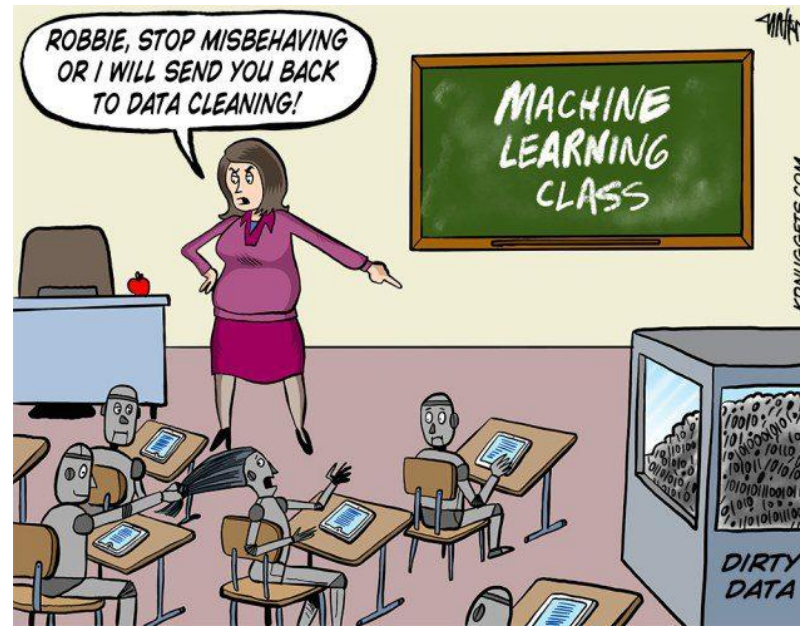


Après le Datachallenge

- Base de données Cytologia : <https://www.data.gouv.fr/datasets/cytologia/>
- Article under review (HemaSphere)
- Participation à plusieurs Webinaires organisés par le HdH, remise des prix le 04/06/25
- Collaboration avec industriels?

Conclusion

- **Très nombreuses applications de l'AI en hématologie :**
 - Diagnostic/diagnostic différentiel
 - Morphologie
 - Onco-hématologie
 - Hématologie bénigne
- **Importance de l'anonymisation des données**
- Etudes multicentriques +++ (généralisation limitée pour la plupart des modèles)
- **Diagnostic le plus fin possible pour le patient**



Merci de votre attention