

# Apport de l'apprentissage profond pour le diagnostic en hématologie

Cédric De Almeida Braga



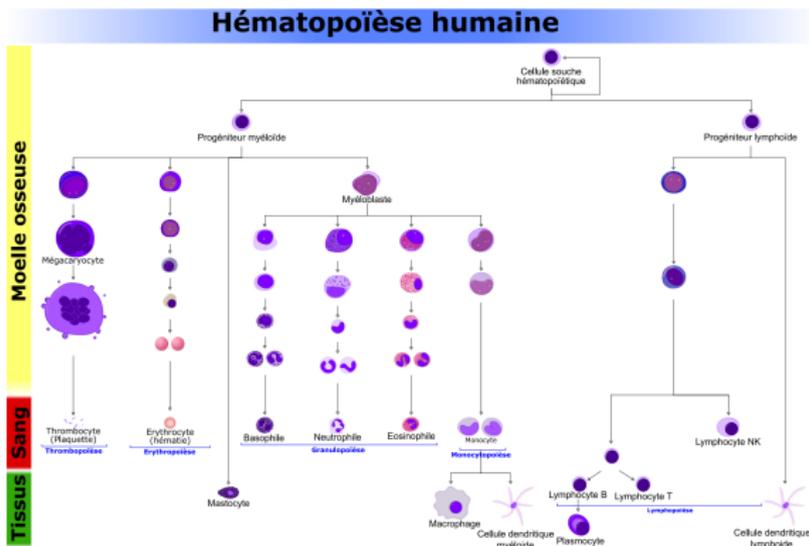
**BioInfoDiag**



# Hématologie et diagnostic

## Définitions

- Hématologie : étude du sang et de l'hématopoïèse, nombreuses pathologies



# Hématologie et diagnostic

## Définitions

- Hématologie : étude du sang et de l'hématopoïèse, nombreuses pathologies
- Diagnostic complexe, étude de l'héogramme et du myélogramme

### HEMOGRAMME

#### NUMERATION SANGUINE

*Système CHAINE MINROD (HG) - XED / XS (HGRU)*

Leucocytes	5,59	Giga/l			(4,0-10,0)
Hématites	3,03	Tera/l	*		(4,50-5,90)
Hémoglobine	10,7	g/dl	*		(13,5-17,5)
Hématocrite	32,5	%	*		(40,0-52,0)
VGM	107,3	Femto.l	*		(80,0-98,0)
CCMH	32,9	g/dl			(32,0-38,0)
TCMH	35,3	pg	*		(26,0-32,0)
Indice de distribution des GR	15,0	%	*		(12,0-14,5)
Plaquettes	213	Giga/l			(150-400)
				Numération par fluorocytométrie	
VPM	10,2	Femto.l			(7,4-11,7)
Réticulocytes	16	/1000 GR	Soit	48	Giga/l (25-100)

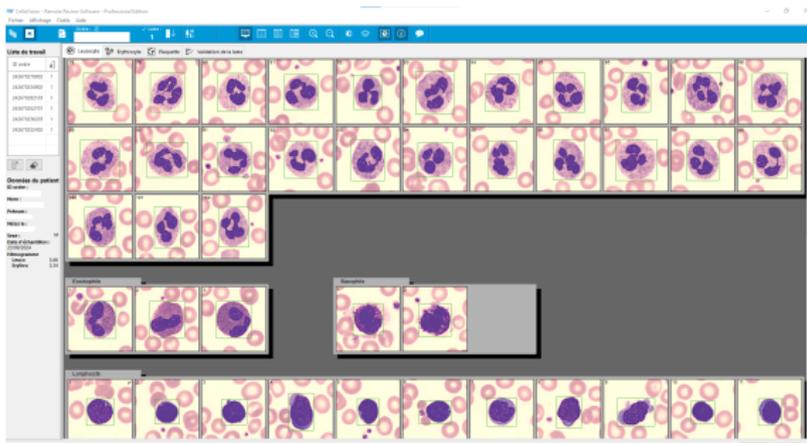
#### FORMULE LEUCOCYTAIRE

Polynucléaires neutrophiles	64,1	%	Soit	3,58	Giga/l (1,70-7,50)
Lymphocytes	27,9	%	Soit	1,56	Giga/l (1,50-4,00)
Monocytes	6,6	%	Soit	0,37	Giga/l (0,15-0,90)
Polynucléaires éosinophiles	0,9	%	Soit	0,05	Giga/l (0,04-0,50)
Polynucléaires basophiles	0,5	%	Soit	0,03	Giga/l (0,00-0,06)
Erythroblastes	0,0	/100 GB			(0,0-0,0)

# Hématologie et diagnostic

## Définitions

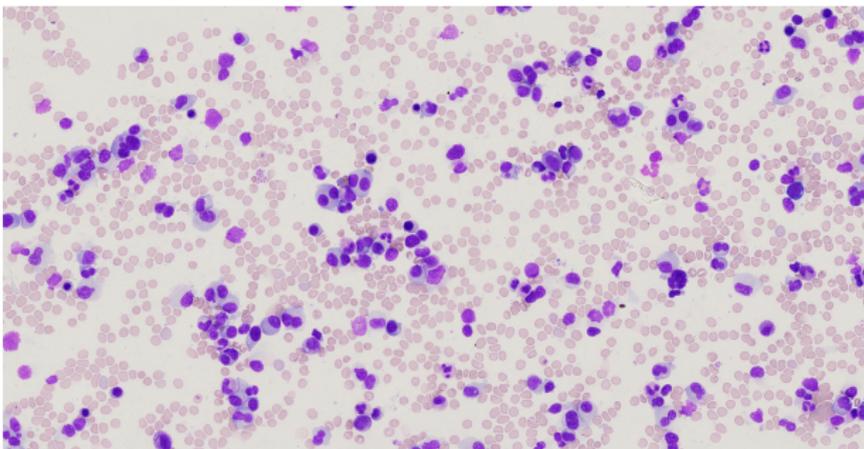
- Hématologie : étude du sang et de l'hématopoïèse, nombreuses pathologies
- Diagnostic complexe, étude de l'hémogramme et du myélogramme



# Hématologie et diagnostic

## Définitions

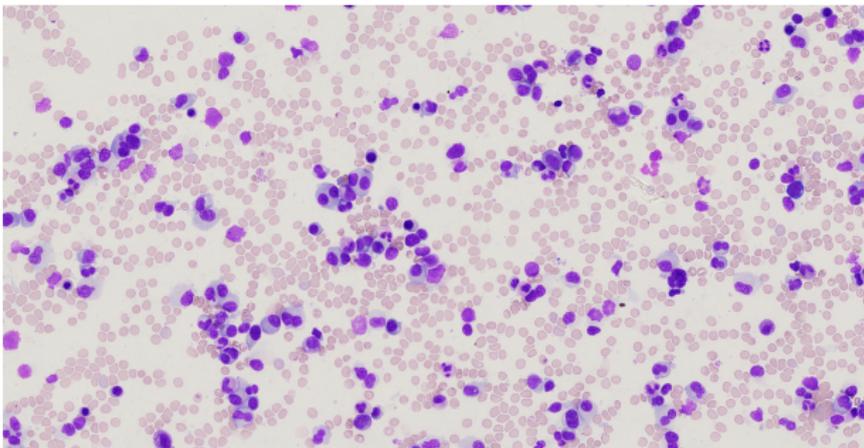
- Hématologie : étude du sang et de l'hématopoïèse, nombreuses pathologies
- Diagnostic complexe, étude de l'hémogramme et du myélogramme



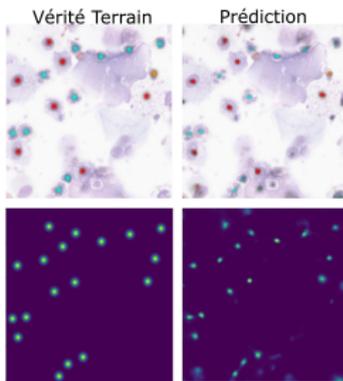
# Hématologie et diagnostic

## Définitions

- Hématologie : étude du sang et de l'hématopoïèse, nombreuses pathologies
- Diagnostic complexe, étude de l'hémogramme et du myélogramme
- Cytologie : comptage, reconnaissance de morphologies anormales



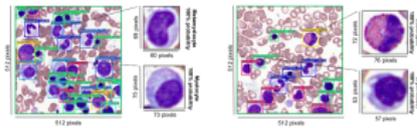
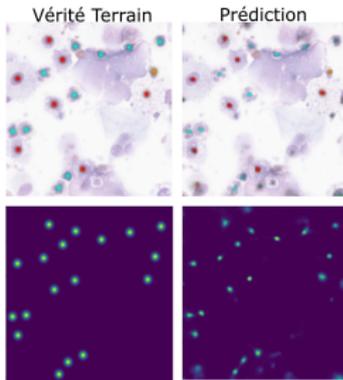
# Apprentissage profond en hématologie



## Applications en hématologies

- Comptage cellulaire, détection d'objet
- Segmentation d'instance ou sémantique, reconnaissance de type cellulaire
- Diagnostic de bout en bout

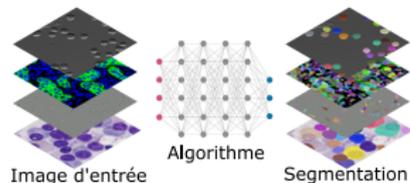
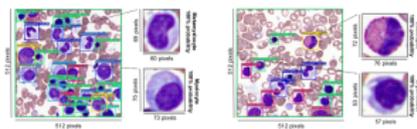
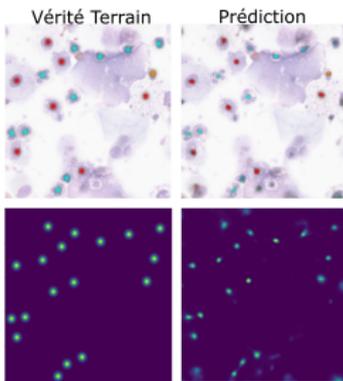
# Apprentissage profond en hématologie



## Applications en hématologies

- Comptage cellulaire, détection d'objet
- Segmentation d'instance ou sémantique, reconnaissance de type cellulaire
- Diagnostic de bout en bout

# Apprentissage profond en hématologie



## Applications en hématologies

- Comptage cellulaire, détection d'objet
- Segmentation d'instance ou sémantique, reconnaissance de type cellulaire
- Diagnostic de bout en bout

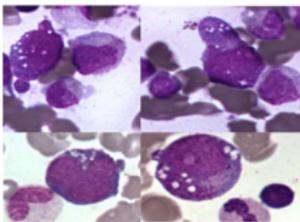
# Limitations et problématiques

## Limitations

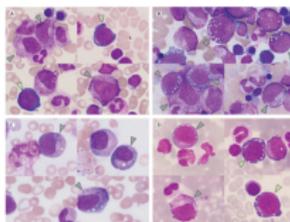
- Fort besoin en données et savoir expert
- Nécessité de garder l'humain dans la boucle diagnostique
- Contraintes matérielles pour le déploiement des modèles

# Syndrome de VEXAS

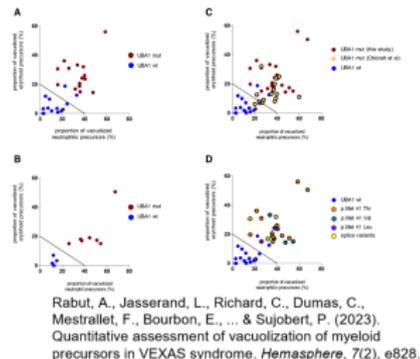
# Définition du syndrome



Templé, M.; Kosmider, O. VEXAS Syndrome: A Novelty in MDS Landscape. *Diagnostics* 2022, 12, 1590



Lacombe, V., Hadjadj, J., Geogin-Lavialle, S., Lavigne, C., Geneviève, F., & Kosmider, O. (2024). Vacuoles in bone marrow progenitors: VEXAS syndrome and beyond. *The Lancet Haematology*, 11(2), e160-e167.



## Caractéristiques et diagnostic

- Décrit en 2020<sup>1</sup>, affecte les hommes > **50 ans**
- Mutation du gène **UBA1**, impliqué dans la réponse inflammatoire
- Manque de critères sur les cellules circulantes

1. "Somatic Mutations in UBA1 and Severe Adult-Onset Autoinflammatory Disease". In : *New England Journal of Medicine* 383.27 (2020), p. 2628-2638.

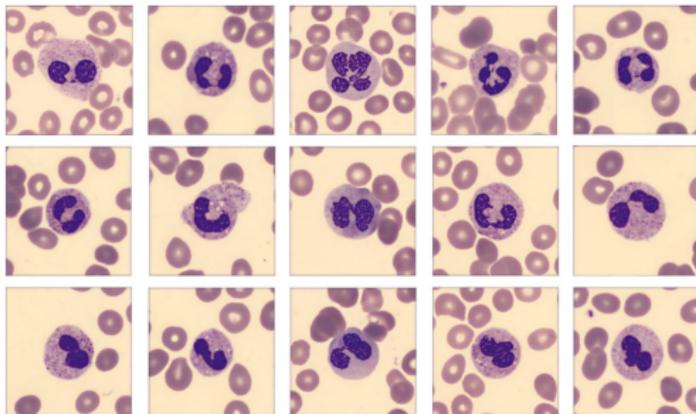
# Objectifs

## Analyse du sang périphérique

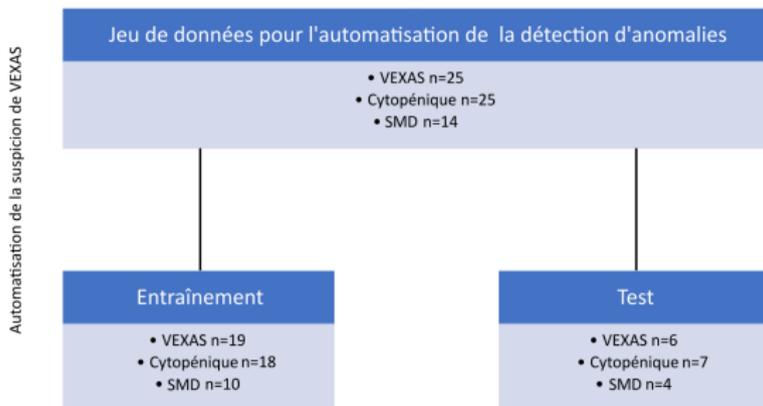
- Identifier des critères cytomorphologiques pour la suspicion de VEXAS dans le sang
- Automatiser la quantification de ces critères à l'aide des pipelines d'analyse pré-existants



# Patients et données

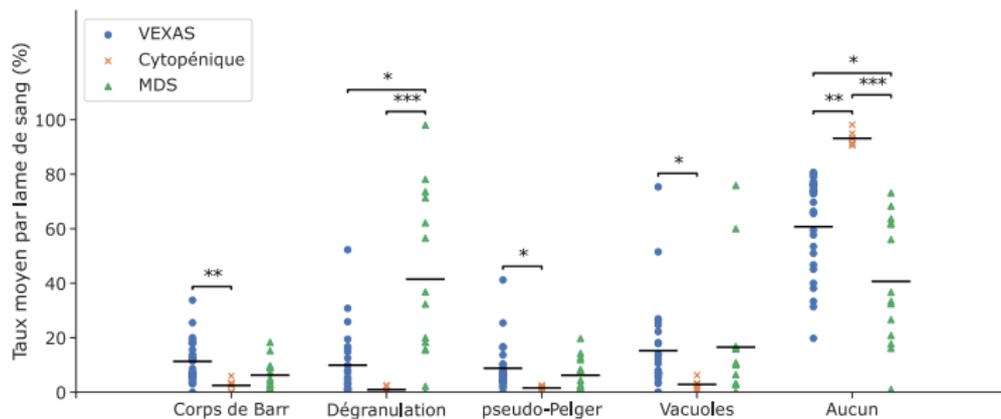
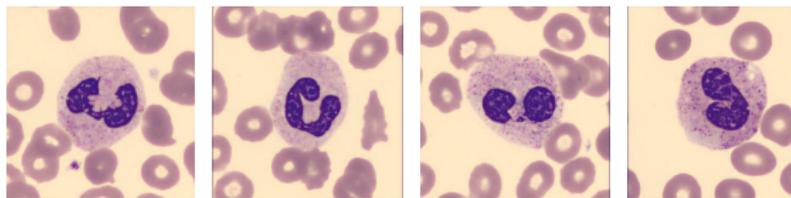


# Patients et données

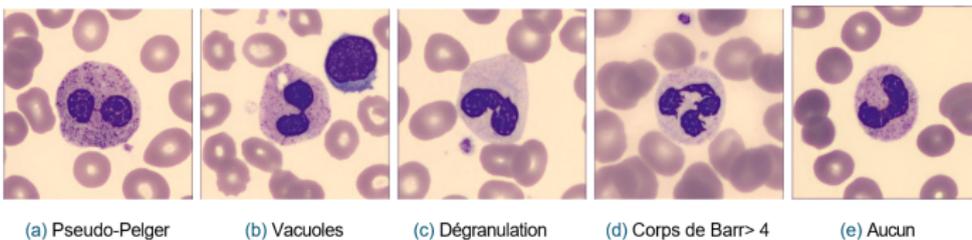
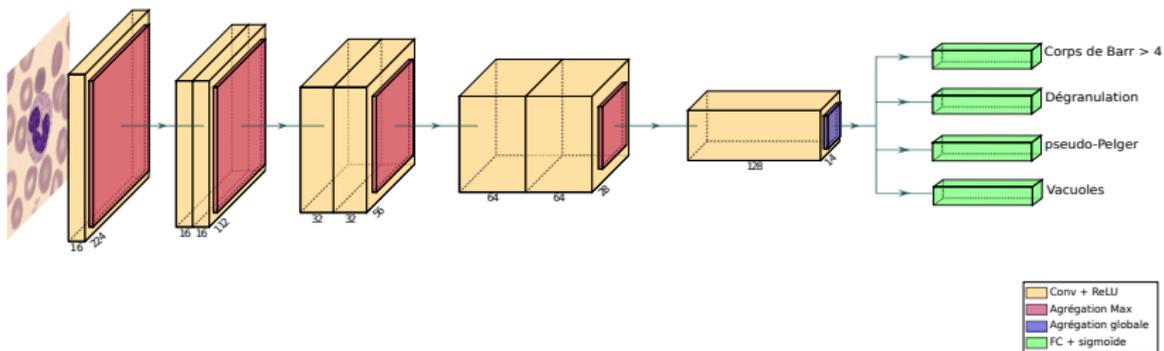


	<b>Entraînement</b>	<b>Test</b>	<b>Total</b>
patients VEXAS	19	6	25
patients Cytopéniques	18	7	25
patients SMD	10	4	14
Lames de sang	85	28	113
Images	7,244	2,270	9,514

# Anomalies morphologiques - Résultats

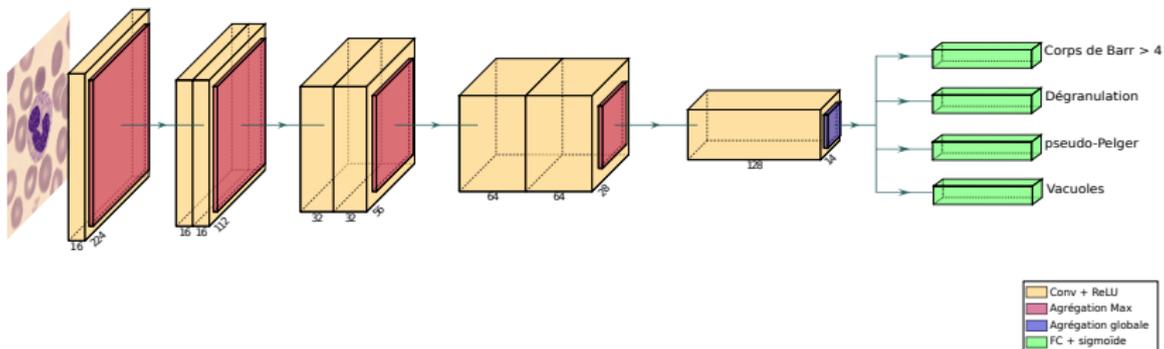


# Reconnaissance d'anomalies - Méthodes



Classification multi-label

# Reconnaissance d'anomalies - Méthodes



## Paradigme expérimental

- Pré-entraînement WBCAtt<sup>2</sup>
- 3 architectures testées : ResNet50, ResNet18, CNN-VEXAS
- Stratégie d'échantillonnage, fonction de coût asymétrique<sup>3</sup>

2. Satoshi TSUTSUI, Winnie PANG et Bihan WEN. "WBCAtt : a white blood cell dataset annotated with detailed morphological attributes". In : **Advances in Neural Information Processing Systems** 36 (2024).

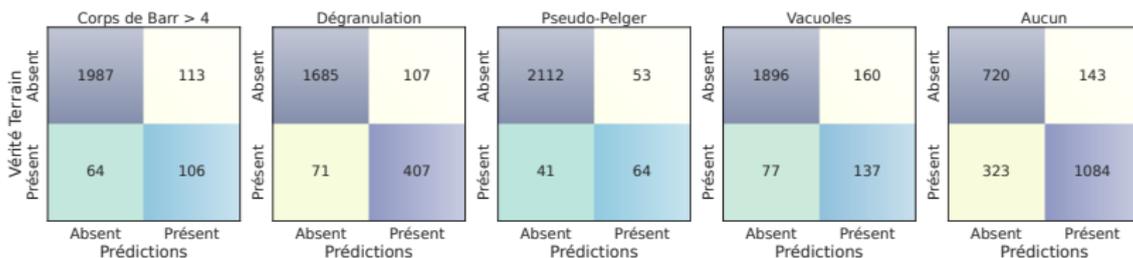
3. Tal RIDNIK et al. "Asymmetric loss for multi-label classification". In : **Proceedings of the IEEE/CVF international conference on computer vision**. 2021, p. 82-91.

# Reconnaissance d'anomalies - Résultats

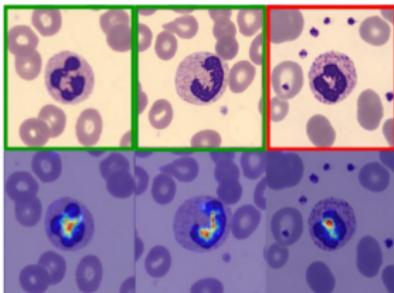
Model	# Params	GFLOPS	Macro-P	Macro-R	Macro-F1
ResNet18	11,18M	3,64	63,14	67,60	64,55
ResNet50	23,52M	8,21	<b>65,55</b>	<b>70,77</b>	<b>67,70</b>
CNN-VEXAS	<b>0,15M</b>	<b>0,37</b>	63,35	69,90	66,03

## Reconnaissance d'anomalies - Résultats

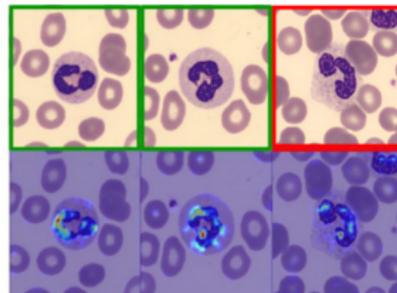
Model	# Params	GFLOPS	Macro-P	Macro-R	Macro-F1
ResNet18	11,18M	3,64	63,14	67,60	64,55
ResNet50	23,52M	8,21	<b>65,55</b>	<b>70,77</b>	<b>67,70</b>
CNN-VEXAS	<b>0,15M</b>	<b>0,37</b>	63,35	69,90	66,03



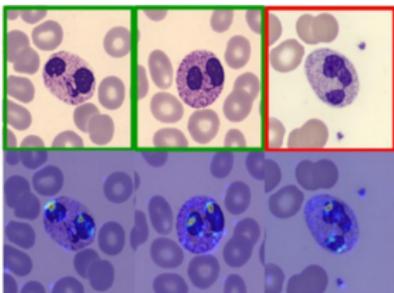
# Reconnaissance d'anomalies - Résultats



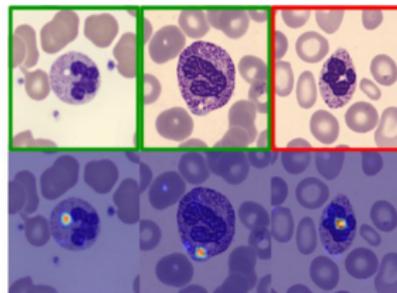
Corps de Barr



Dégranulation



Pseudo-Pelger

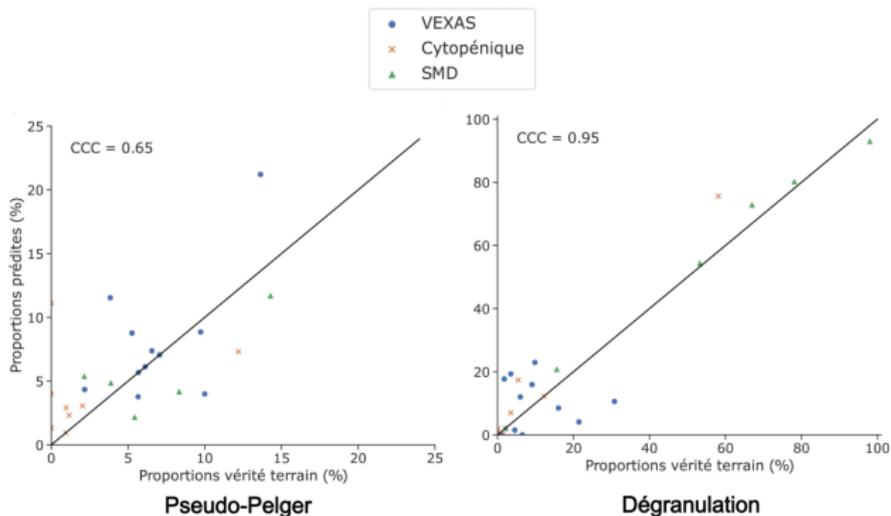


Vacuoles

# Pratique clinique

## Analyse exploratoire

- CC de Lin par classe dans l'intervalle [0.65;0.95]
- Accord inter-annotateur Dégranulation = 0.40, Pseudo-Pelger = 0.37<sup>4</sup>
- Dégranulation → SMD | Pseudo-Pelger → VEXAS

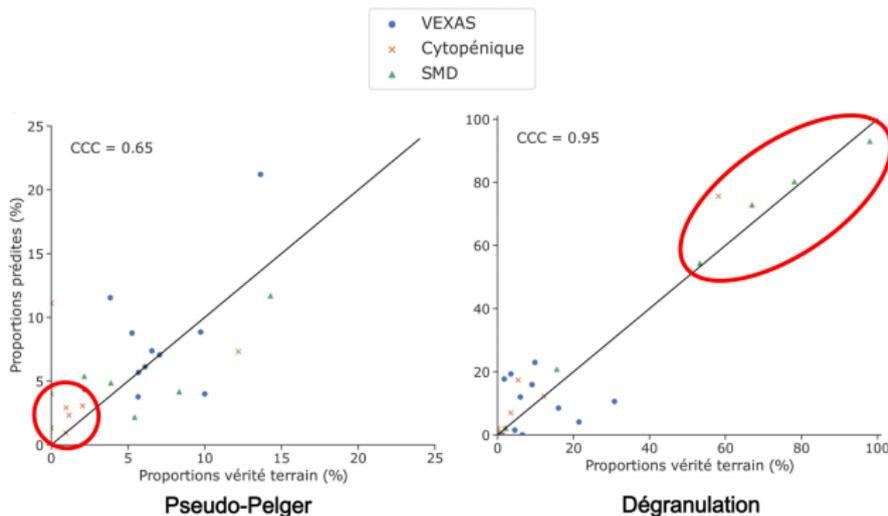


4. Keiko SASADA et al. "Inter-observer variance and the need for standardization in the morphological classification of myelodysplastic syndrome". In : *Leukemia research* 69 (2018), p. 54-59.

# Pratique clinique

## Analyse exploratoire

- CC de Lin par classe dans l'intervalle [0.65;0.95]
- Accord inter-annotateur Dégranulation = 0.40, Pseudo-Pelger = 0.37<sup>4</sup>
- Dégranulation → SMD | Pseudo-Pelger → VEXAS



4. Keiko SASADA et al. "Inter-observer variance and the need for standardization in the morphological classification of myelodysplastic syndrome". In : *Leukemia research* 69 (2018), p. 54-59.

## En résumé

### Contributions

- Collaboration inter-centres : Lyon, Tours, La Roche, St Nazaire, Nantes
- Critères cytomorphologiques pour la suspicion de VEXAS dans le sang
- Modèle pour la caractérisation des anomalies des neutrophiles
- Jeu de données pour la classification mutli-label de morphologies cellulaires dans un contexte clinique complexe

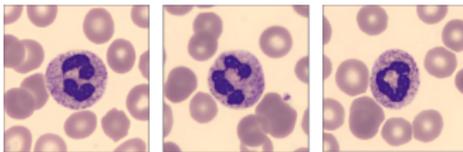
Perspective : Inclusion de paramètres supplémentaires pour la prédiction du diagnostic

# Syndrome Myélodysplasique

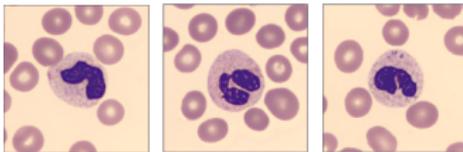
# Le Syndrome Myélodysplasique

## Définition

- Groupe hétérogène d'hémopathies clonales de la moelle osseuse, hématopoïèse inefficace
- Cytopénie + dysplasie >10% dans les lignées cellulaires
- Touche le patient âgé (70 ans), risque d'évolution en leucémie aiguë myéloïde

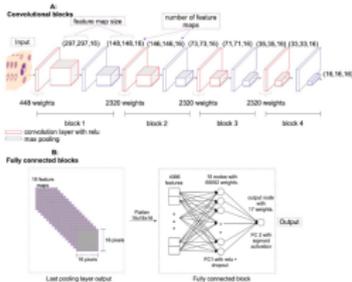


Neutrophiles normaux

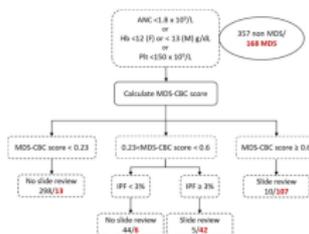


Neutrophiles dysplasiques

# État de l'art du dépistage sur le sang



Acevedo, Andrea, et al. "A new convolutional neural network predictive model for the automatic recognition of hypogranulated neutrophils in myelodysplastic syndromes." *Computers in Biology and Medicine* 134 (2021): 104479.



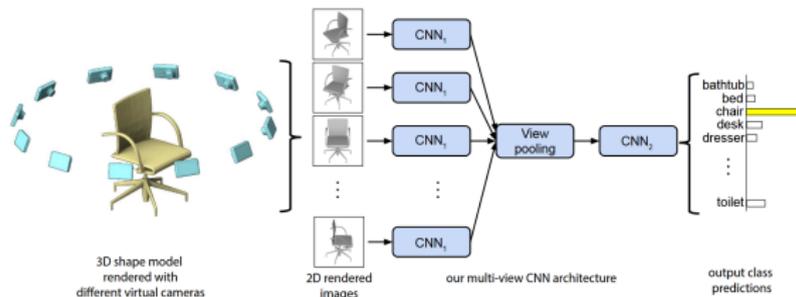
Zhu, Jaja, et al. "Machine learning-based improvement of MDS-CBC score brings platelets into the limelight to optimize smear review in the hematology laboratory." *BMC cancer* 22.1 (2022): 1-10

## Différentes modalités des données

- Diagnostic SMD automatisé soit par l'image, soit par la NFS
- Étude nantaise diagnostic SMD automatisé sur la NFS<sup>5</sup>
- Utilisation conjointe des modalités images et numériques

5. Robin BOUTAULT et al. "A novel complete blood count-based score to screen for myelodysplastic syndrome in cytopenic patients". In : *British journal of haematology* 183.5 (2018), p. 736-746.

# Quantification de la dysplasie dans l'image



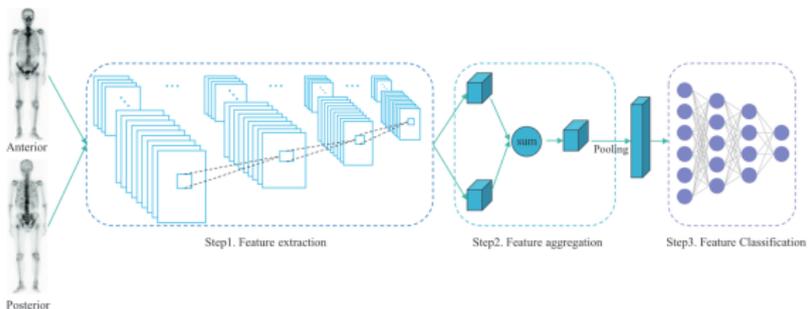
## Définition

- Seuil de dysplasie SMD défini par l'OMS = 10%
- Utilisation d'un CNN multi-vues<sup>6</sup>
- Appliqué avec succès aux images médicales pour le diagnostic<sup>7</sup>

6. Hang Su et al. "Multi-view convolutional neural networks for 3d shape recognition". In : **Proceedings of the IEEE international conference on computer vision**. 2015, p. 945-953.

7. Yong Pi et al. "Automated diagnosis of bone metastasis based on multi-view bone scans using attention-augmented deep neural networks". In : **Medical Image Analysis** 65 (2020), p. 101784.

# Quantification de la dysplasie dans l'image



## Définition

- Seuil de dysplasie SMD défini par l'OMS = 10%
- Utilisation d'un CNN multi-vues<sup>6</sup>
- Appliqué avec succès aux images médicales pour le diagnostic<sup>7</sup>

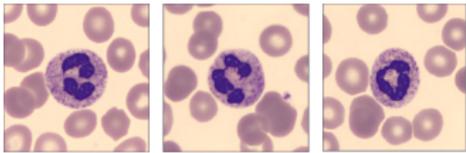
6. Hang Su et al. "Multi-view convolutional neural networks for 3d shape recognition". In : **Proceedings of the IEEE international conference on computer vision**. 2015, p. 945-953.

7. Yong Pi et al. "Automated diagnosis of bone metastasis based on multi-view bone scans using attention-augmented deep neural networks". In : **Medical Image Analysis** 65 (2020), p. 101784.

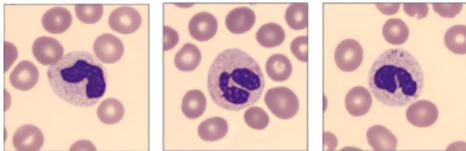
# Problématique

## Délimitation du sujet

- Dépistage du SMD à l'hémogramme
- Données multimodales = paramètres NFS + images neutrophiles
- Diagnostic de bout-en-bout, effort d'annotation minimal

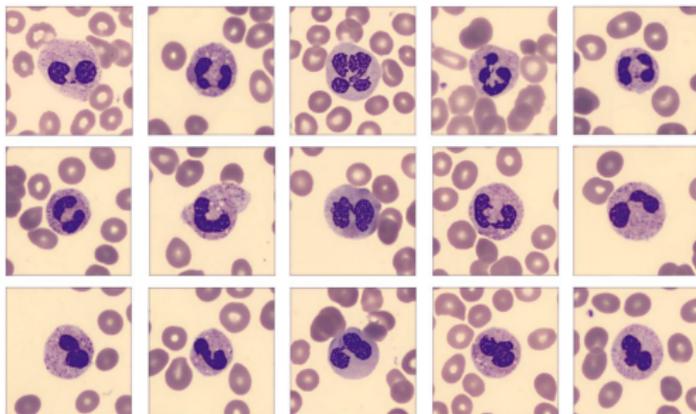


Neutrophiles normaux



Neutrophiles dysplasiques

# Données



## Paramètres NFS

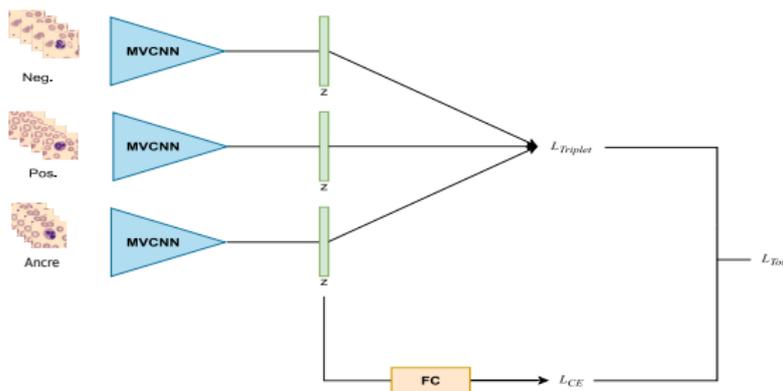
- Compte absolu de neutrophiles (ANC)
- Volume globulaire moyen (VGM)
- Neutro-WX

## Données

	jeu d'entraînement	jeu test
Patients SMD		
# patients	38	22
Âge médian (IQR)	76 (70-84)	77 (71-79)
# images	3 253	1 652
images/patients	85,61	75,09
Patients Contrôles		
# patients	35	22
Âge médian (IQR)	68 (61-80)	65 (58-72)
# images	3 369	2 123
images/patients	96,26	96,50

117 patients : 60 SMD et 57 contrôles, 10 397 images : 4 905 SMD et 5 492 contrôles

# Méthodes

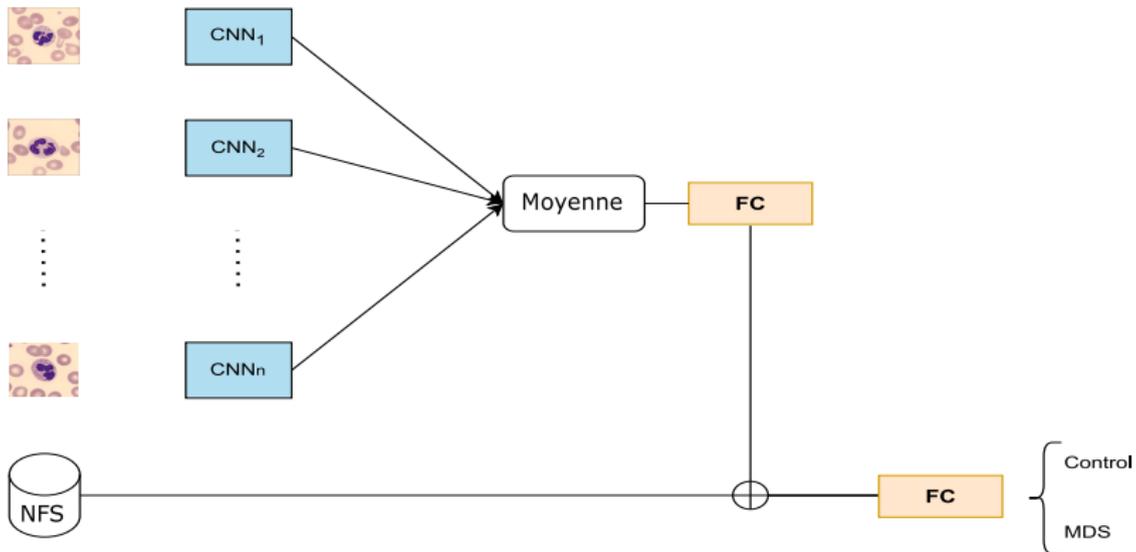


## Apprentissage multi-vues et multi-tâches

$$L_{tot} = L_{CE} + L_{Triplet}$$

- $L_{CE}$  : Classification binaire SMD/Contrôle à partir des  $n$  images d'un patient donné
- $L_{Triplet}(A, P, N)$  : Apprentissage réseau siamois, triplets → Contraintes dans l'espace latent du réseau + sur-échantillonnage

# Méthodes



# Résultats

Méthode	Accuracy	F1 score	Précision	Rappel
CBC-score <sup>8</sup>	75,00	74,88	72,00	81,82
MVCNN	72,73	72,21	67,86	86,36
MVCNN + Triplet	75,00	74,88	<b>78,95</b>	68,18
MVCNN + Triplet + NFS	<b>79,55</b>	<b>79,45</b>	76,00	<b>86,36</b>

8. Robin BOUTAULT et al. "A novel complete blood count-based score to screen for myelodysplastic syndrome in cytopenic patients". In : *British journal of haematology* 183.5 (2018), p. 736-746.

# En résumé

## Contributions

- Cohorte multi-modale pour le dépistage du SMD
- Paradigme d'entraînement multi-vues et multi-tâches
- Meilleures performances que le score en place, basé uniquement sur la NFS
- Collaboration inter-centres : Tours, Poitiers, Nantes

# En résumé

## Contributions

- Cohorte multi-modale pour le dépistage du SMD
- Paradigme d'entraînement multi-vues et multi-tâches
- Meilleures performances que le score en place, basé uniquement sur la NFS
- Collaboration inter-centres : Tours, Poitiers, Nantes

## Perspectives

- Inclusion de patients et centres partenaires (CHU de Toulouse)
- Étude de la lignée érythrocytaire
- Mécanisme de fusion des données (modules d'attention)
- Explicabilité de la décision pour le praticien

## Conclusion

# Conclusion générale

## Contributions

- Apprentissage profond pour le diagnostic de différentes pathologies
- Visualisation et interprétabilité de la décision dans un domaine critique
- Architectures légères et compétitives, plus facilement déployables

# Conclusion générale

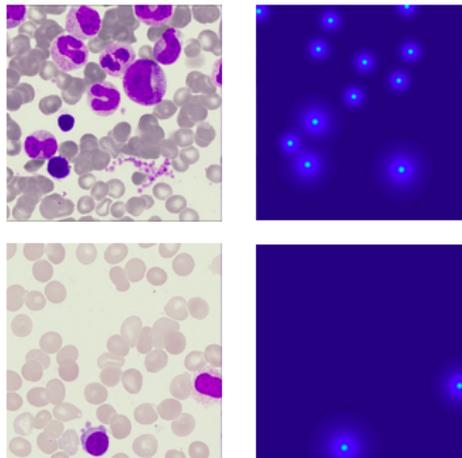
## Contributions

- Apprentissage profond pour le diagnostic de différentes pathologies
- Visualisation et interprétabilité de la décision dans un domaine critique
- Architectures légères et compétitives, plus facilement déployables

## Rayonnement

- Collaborations multi-centriques au niveau national
- Moteur de nombreux projets : thèse d'exercice, mémoire de DES, stage de master...
- Recherche translationnelle : Publications scientifiques et congrès

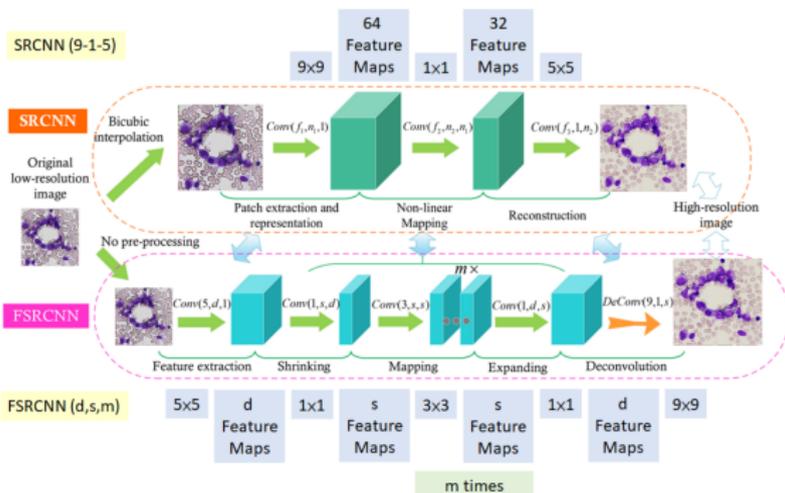
# Projet I - Comptage cellulaire différentiel pour le diagnostic



## Détection d'objet faiblement supervisée

- Benchmark architectures de l'état-de-l'art
- Étude comparative du niveau de supervision et de la résolution des images
- Impact diagnostic dans le cas du Myélome Multiple

# Projet II - Super-résolution des myélogrammes numérisés



## Modèles génératifs pour la génération d'images haute résolution

- Scanner à lame haut débit  $40 \times$  vs Microscope à huile  $60 \times$
- Modèles génératifs/de diffusion latente pour améliorer le  $40 \times$
- Ouverture possible sur la synthèse d'images de myélogrammes

# Remerciements

## Patients ayant participé aux études

### Centres Hospitaliers :

- CHU de Lyon
- CHU de Tours
- Centre hospitalier de Poitiers
- Centre hospitalier de La Roche sur Yon Hospital
- Centre hospitalier de St Nazaire

## Collaborateurs :

- Pr. M. Eveillard, CHU de Nantes
- Pr. Nicolas Normand, Polytech Nantes
- Dr. Perrine Paul-Gilloteaux, IR-SUN Nantes
- A. Lelard, PharmD, CHU de Nantes
- V Le Roc'h, PharmD, CHU de Nantes
- M. Bauvais, PharmD, CHU de Nantes



**BioInfoDiag**

