

Apprentissage, Intelligence Artificielle et Optimisation (AIAO)

RSNA Screening Mammography Breast Cancer Detection

BETHUEL Elouan
GHRAIZI Marwa
SALAUN Nicolas

*M2 Biologie Informatique
Année 2023/2024*

Introduction

Le cancer du sein en chiffre

- Cancer le plus observé chez la femme
- 2,3 millions de nouveaux cas en 2020 à l'échelle mondiale et 685 000 décès
- 5 millions de dépistage en France pour l'année 2028

Projet de deep learning

- Importance de déceler le stade précoce
- Appuie au diagnostic clinique
- Identification de structures complexes
- Organisée par la RSNA sur la plateforme Kaggle

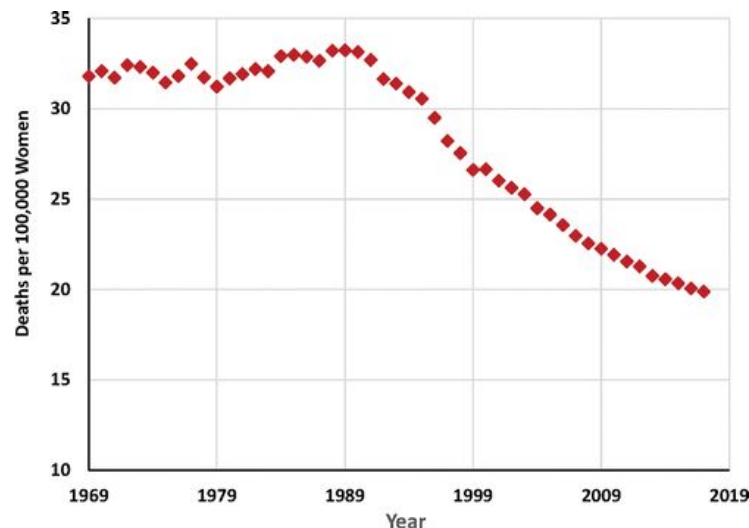


Diagramme illustrant la baisse de la mortalité du cancer du sein suite à une campagne de dépistage massive par mammographies aux USA

Matériels

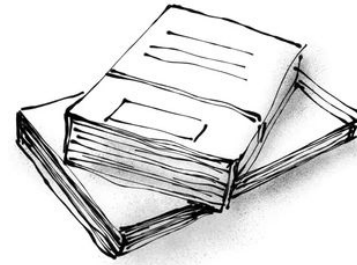
Données kaggle

- Mammographies au format .dcm
- Métadonnées fichier .csv
- 53548 images saines et 1158 cancéreuses
- Volume total de **314.72 Go**



Données Croppées

- Plusieurs résolutions au format .png
(256, 512, 768 et 1024)
- Métadonnées fichier .csv
- 53548 images saines et 1158 cancéreuses
- Volume total **entre 1 Go et 14 Go**



Méthodologie: Pré-traitement

Pre-processing: Cropping des images

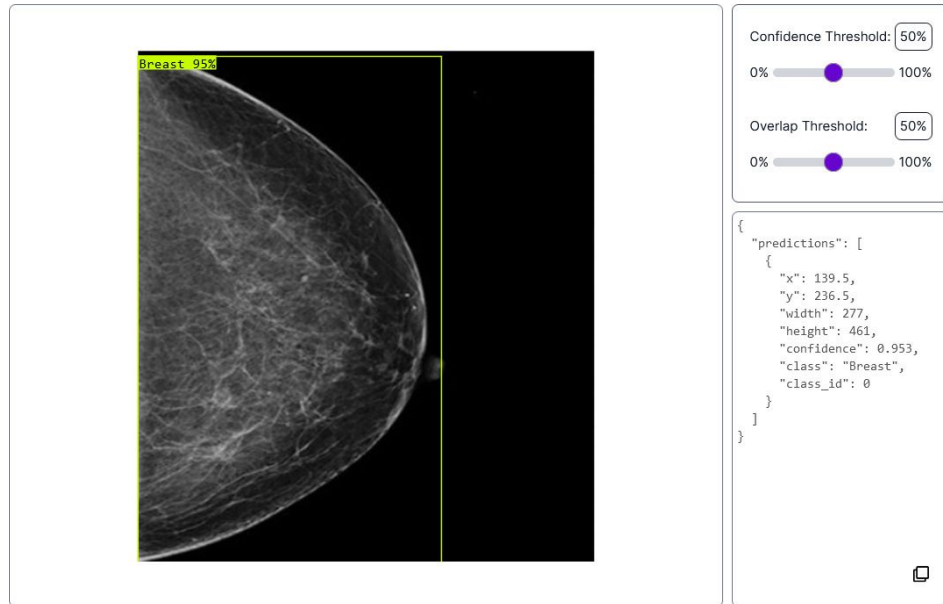
Utilisation de YOLOX :

- Annotations manuelles de 320 mammographies
- Divisons en 250 images d'apprentissage et 70 images de test
- Entraînement du modèle

Résultats

- Précision de 99,7%,
- Obtention des ROI (Region Of Interest)

breast-detection-vdydv/2 (latest)



Résultat du modèle YOLOX sur la prédiction de la zone d'intérêt

Pre-processing: T-distributed Stochastic Neighbor Embedding (T-SNE)

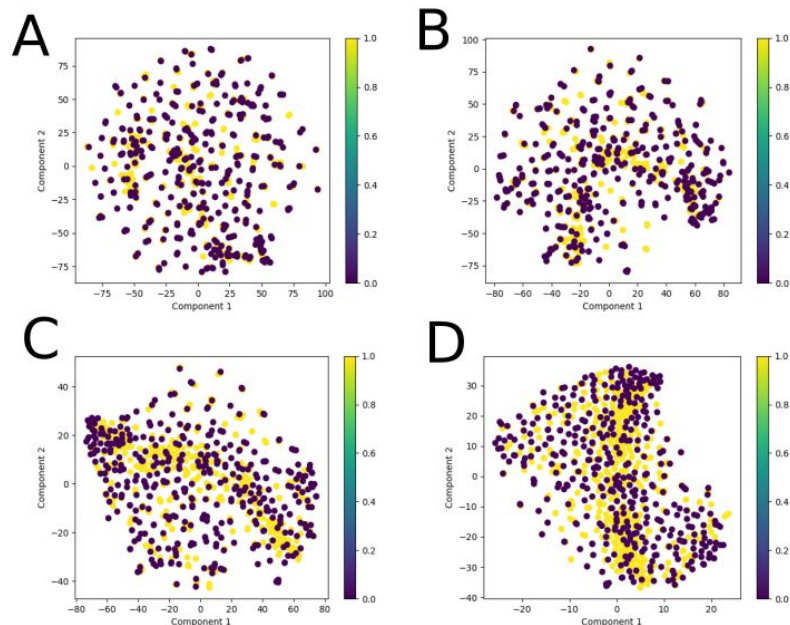
Méthode de visualisation non supervisée

Intérêts

- Réduction de dimensions non linéaire
- Préserve les structures locales

Réalisé avec :

- 400 images saines et 400 cancéreuses
- Valeurs de perplexités croissantes (2, 5, 10, 50)



*T-SNE sur les deux premières dimensions. En jaune les images cancéreuses et en violets les saines
A) perplexité 2 B) perplexité 5 C) perplexité 10 D) perplexité 50*

Pre-processing: Normalisation et contraste

Différences de répartition des niveaux de gris

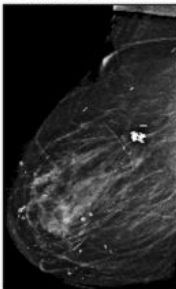
Normalisation globale :

- Z-score
- Par quantile

Accentuation du contraste :

- Égalisation d'histogramme
- CLAHE

image cancéreuse, contours marqués



Histogramme de répartition des pixels

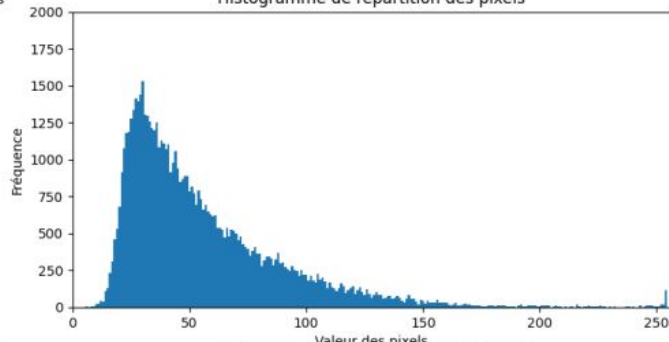
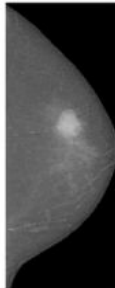
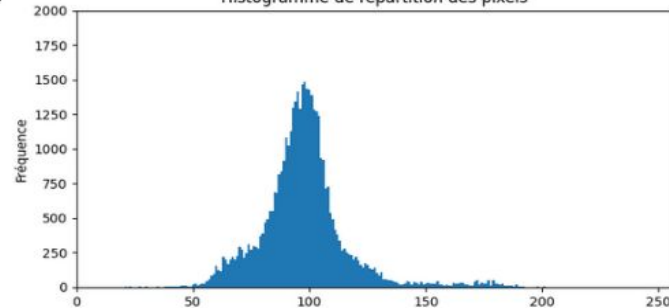


image cancéreuse, contours diffus

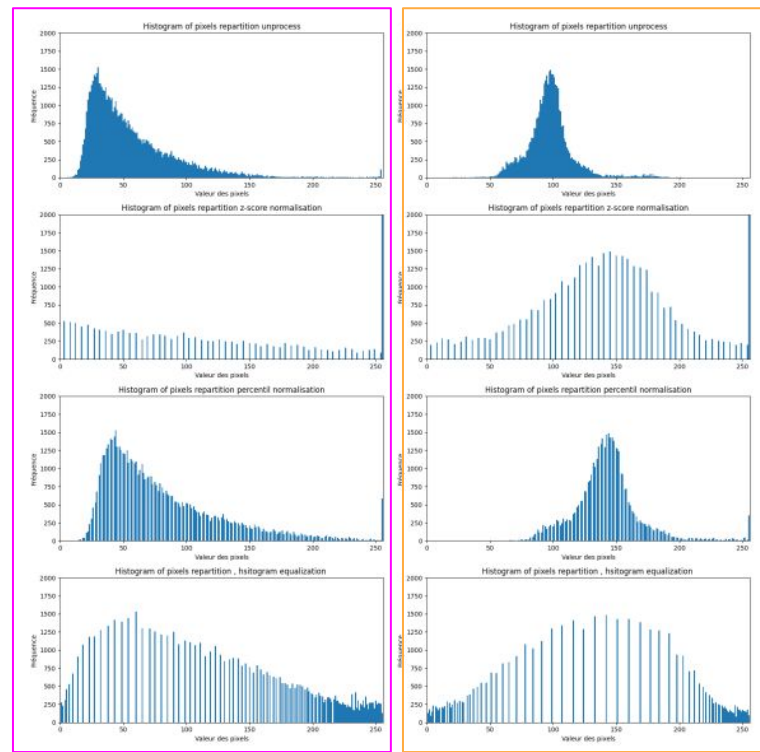
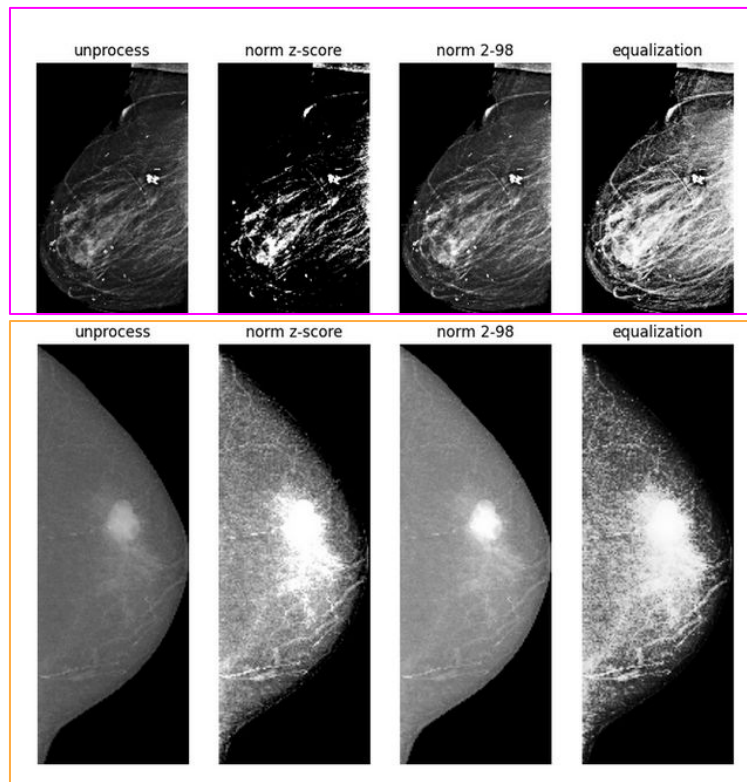


Histogramme de répartition des pixels



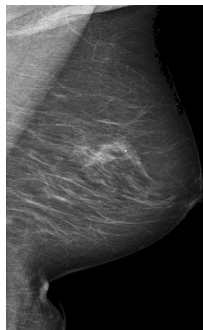
Figures illustrant la différence de luminosité entre les images, et les histogrammes de répartition des pixels associées

Pre-processing: Normalisation et contraste

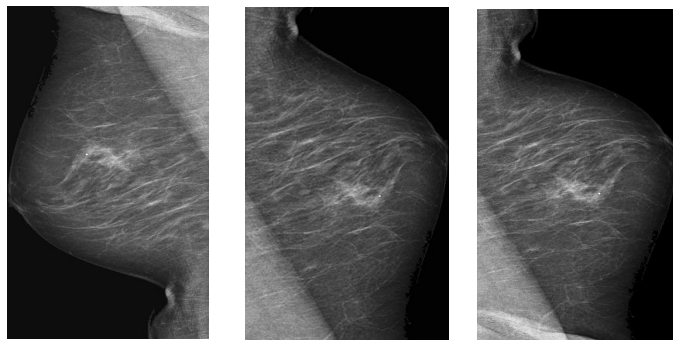
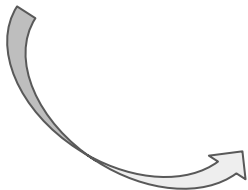
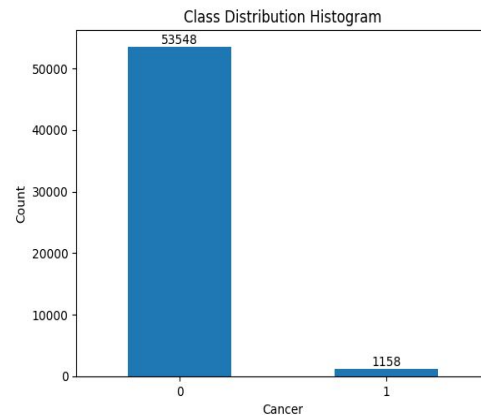


Figures illustrants l'impact de chaque traitement sur deux images à la luminosité différente et leurs histogrammes associées

Pre-processing: Data Augmentation

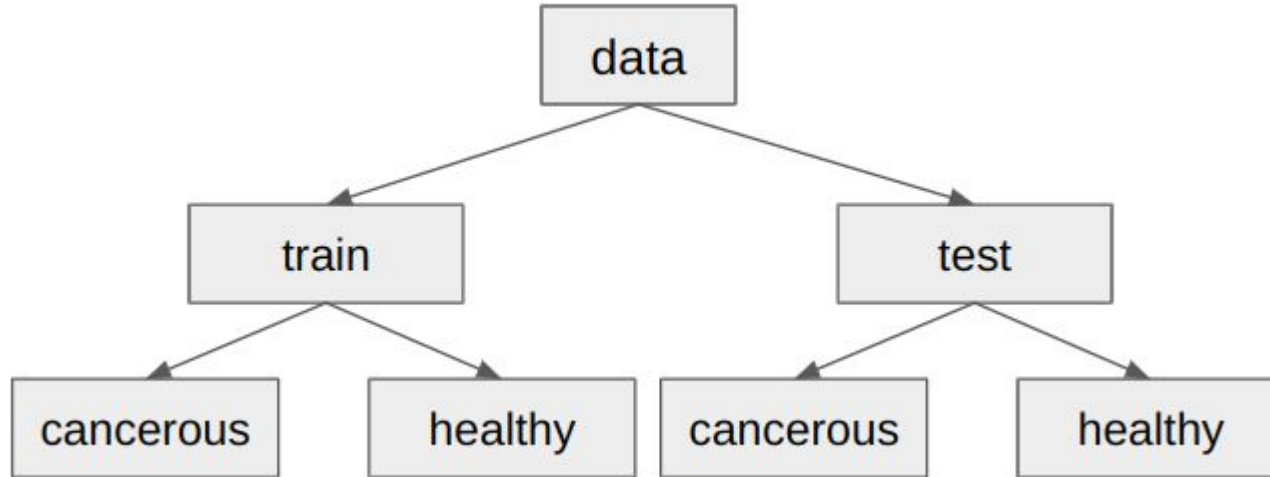


- Solution au déséquilibre des classes
- S'appuie sur un état de l'art de la compétition
- Meilleures expositions des tumeurs



- Modifications aléatoires apportées :
 - rotations horizontales / verticales aléatoires
 - contraste, magnitudes 0.8
 - luminosité, magnitudes 1.2
 - distorsions
 - déviations, magnitudes 0.6
- **965** images par classe → **2895** images par classe

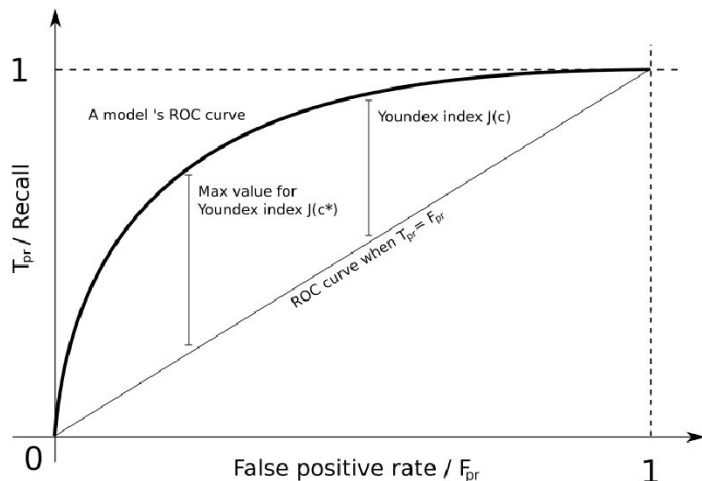
Organisation des données d'entrées



Evaluation des Performances

Cancer: 1
Healthy: 0

		Reality	
Confusion matrix		Negative : 0	Positive : 1
Prediction	Negative : 0	True Negative : TN	False Negative : FN
	Positive : 1	False Positive : FP	True Positive : TP



Receiver Operating Characteristic (ROC) Curve

Performance
Metric for IDS

$$\text{Recall} = \frac{TP}{TP + FN} (\times 100\%)$$

$$\text{Precision} = \frac{TP}{TP + FP} (\times 100\%)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} (\times 100\%)$$

$$F - \text{Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Specificity} = \frac{TN}{TN + FP} (\times 100\%)$$

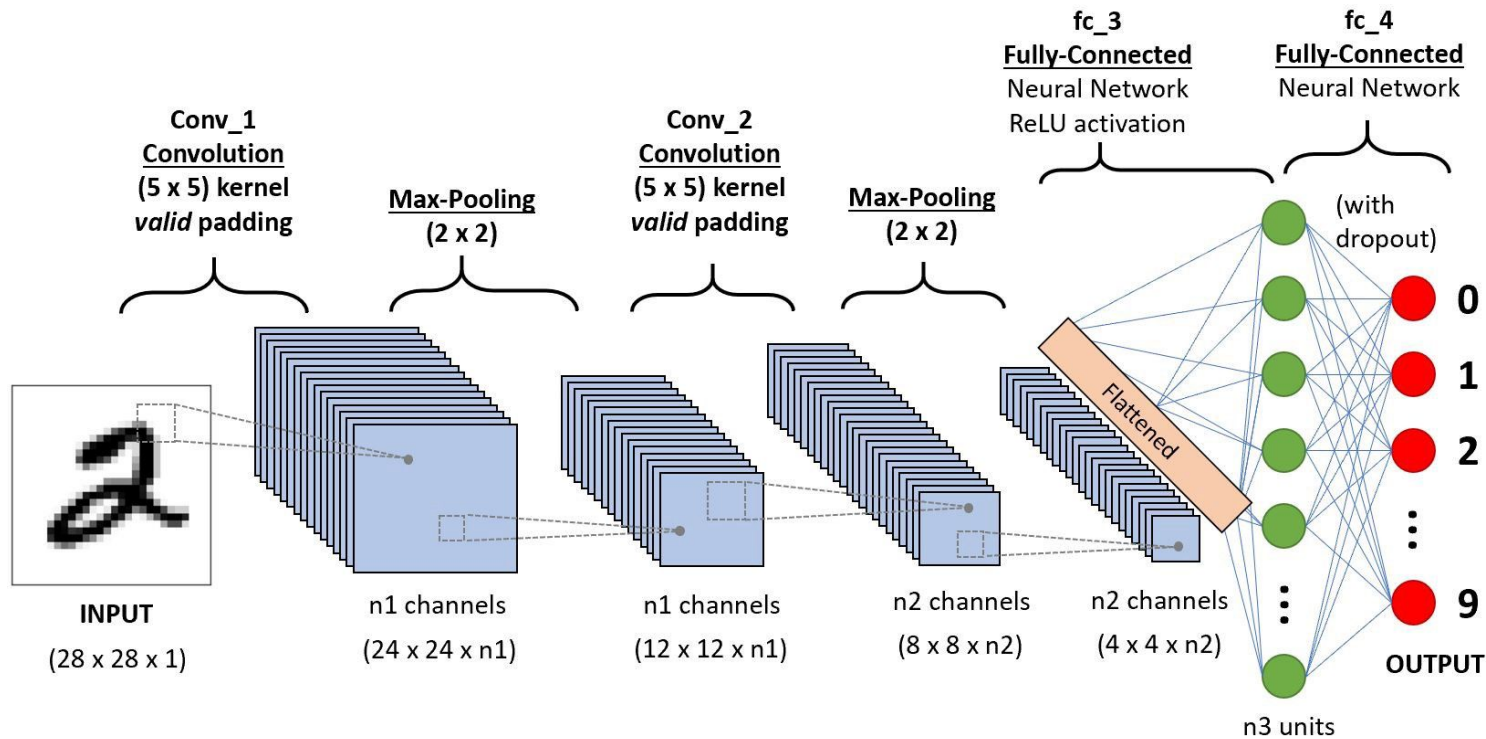
$$\text{True Negative rate} = \frac{TN}{TN + FP}$$

$$\text{True Positive rate} = \frac{TP}{TP + FN}$$

$$\text{False Positive rate} = \frac{FP}{TN + FP}$$

Méthodologie: Modèles

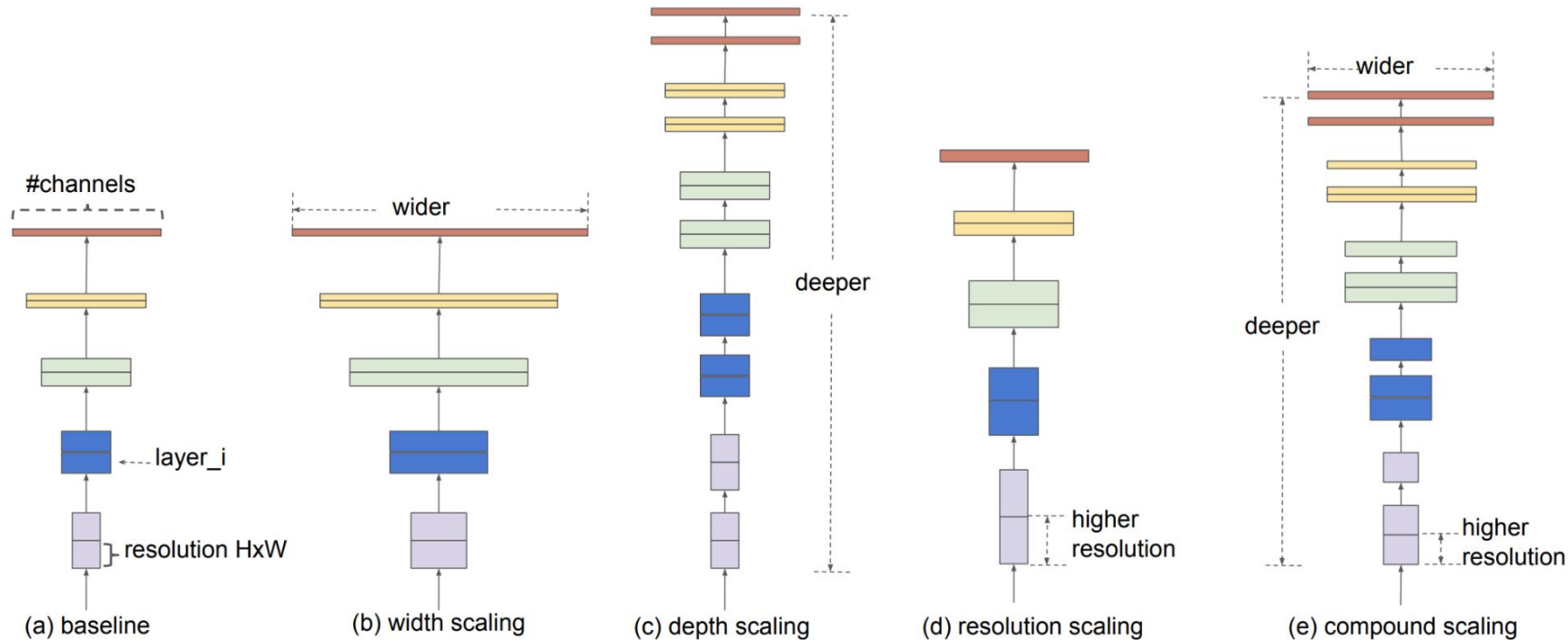
CNN classique



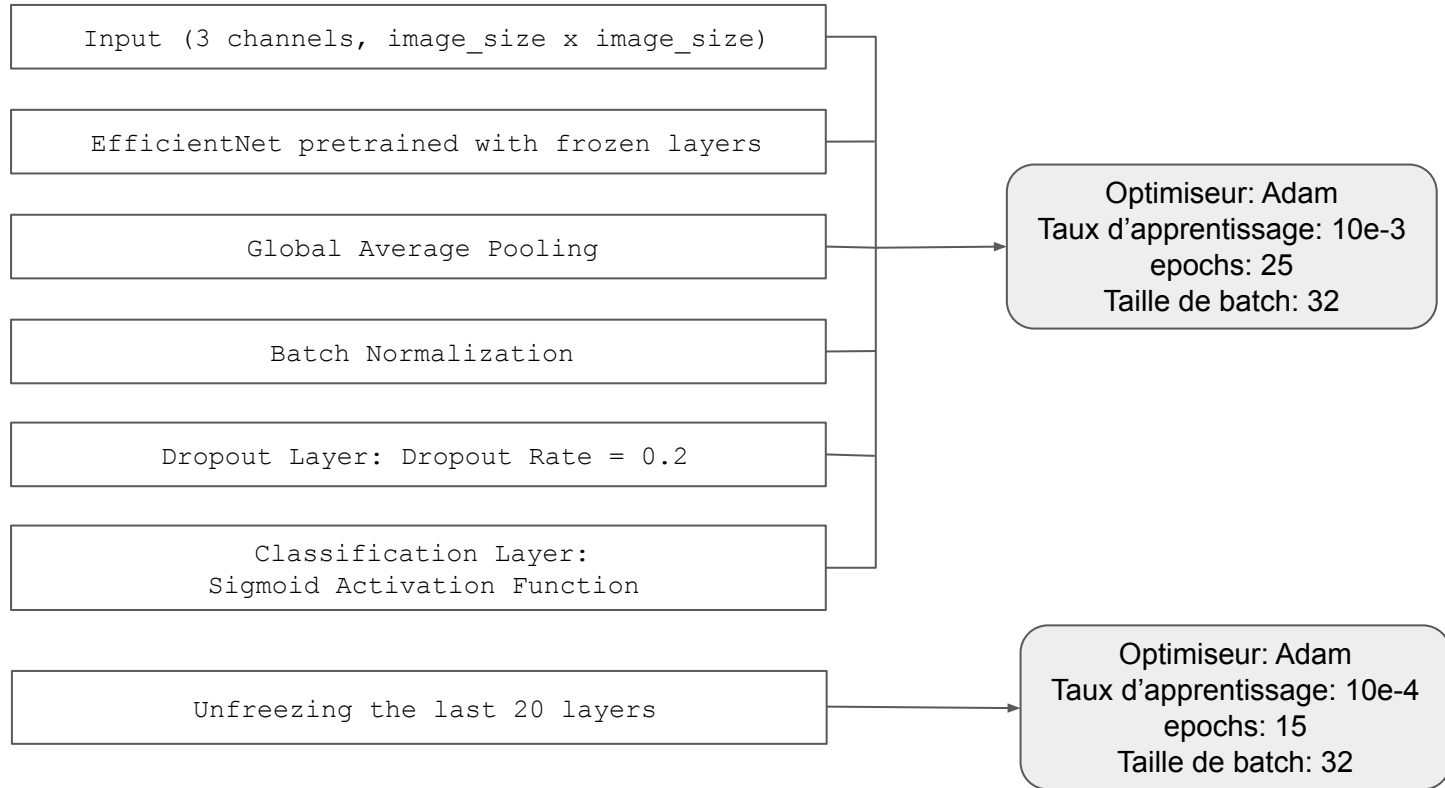
CNN classique

- Test de 4 architectures de CNN pour 32, 64 et 128 neurones par couche :
 - Type I: conv / maxpool / conv / maxpool / dense
 - Type II: conv / maxpool / conv / maxpool / dense / dense
 - Type III: conv / maxpool / conv / maxpool / conv / maxpool / dense / dense
 - Type IV: conv / maxpool / conv / maxpool / conv / maxpool / conv / maxpool / dense / dense
- Augmentation croissante de la complexité
- Réduction de dimensions avec le pooling

Modèle: EfficientNet



Modèle: EfficientNet



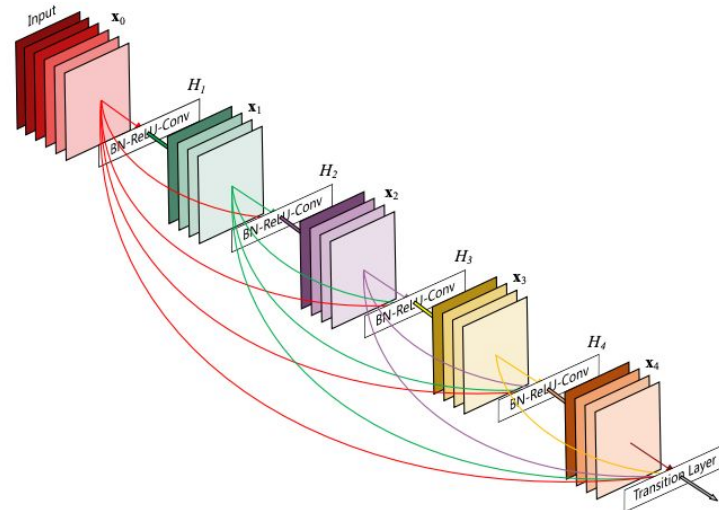
Modèle: DenseNet

Architecture 'dense', chaque couche a en entrée les sorties des couches précédentes

Avantages :

- Meilleures performances sur des tâches complexes
- Diminue le sur-apprentissage
- Utilisé en analyse d'images médicales

Approche similaire à Resnet (gèle des couches)
Augmentation des taux de dropout (0.5 - 0.6)



Résultats

Résultats: CNN

Model	Type I	Type II	Type III	Type IV
32 neurons / layers	0,5	0,47	0,51	0.5
64 neurons / layers	0,51	0,49	0,53	0.52
128 neurons / layers	0,53	0,53	0,55	0.53

F1 scores des différents modèles en fonction du nombre de neurones par couche

- Résultats globalement moyens
- Meilleur modèle de CNN → Type III
- Complexité \neq meilleur résultats
- Beaucoup de paramètres

Résultats: EfficientNet

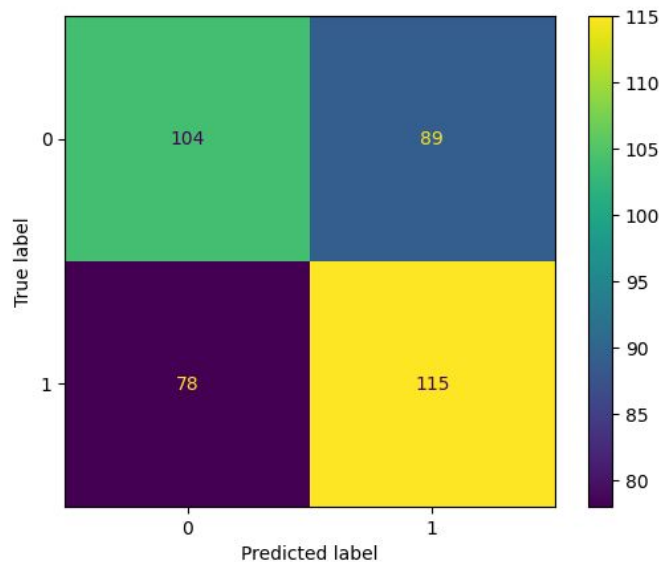
Model	EfficientNet B2 Resolution 256		EfficientNet B4 Resolution 380		EfficientNet B6 Resolution 512			EfficientNet V2S Resolution 512	
Data	un-processed	enhanced	un-processed	enhanced	un-processed	enhanced	augmented	un-processed	enhanced
F1 score	0.56	0.48	0.56	0.34	0.58	0.57	0.57	0.63	0.52
Accuracy	0.58	0.56	0.60	0.54	0.57	0.57	0.56	0.59	0.57
Precision	0.56	0.56	0.60	0.54	0.57	0.57	0.56	0.59	0.56
Recall	0.54	0.41	0.51	0.23	0.59	0.58	0.59	0.70	0.48

Resultats des modeles Efficient Net B2, B4, B6 et V2S sur des images de résolution 256, 380 et 512 prétraitées ou non

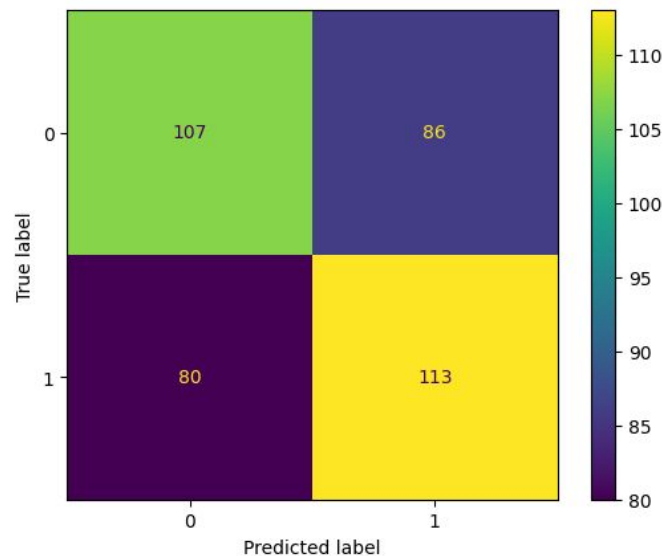
- Performance s'améliore à la mesure de l'augmentation de la résolution
- Les images enhanced créent des biais vers les images saines (faible recall)
- Efficient Net B6 avec les images 512 non traitées produit les meilleures résultats

Résultats: EfficientNet

Matrices de confusion du meilleur modèle:
Efficient Net B6 avec les images de résolution 512



images non traitées

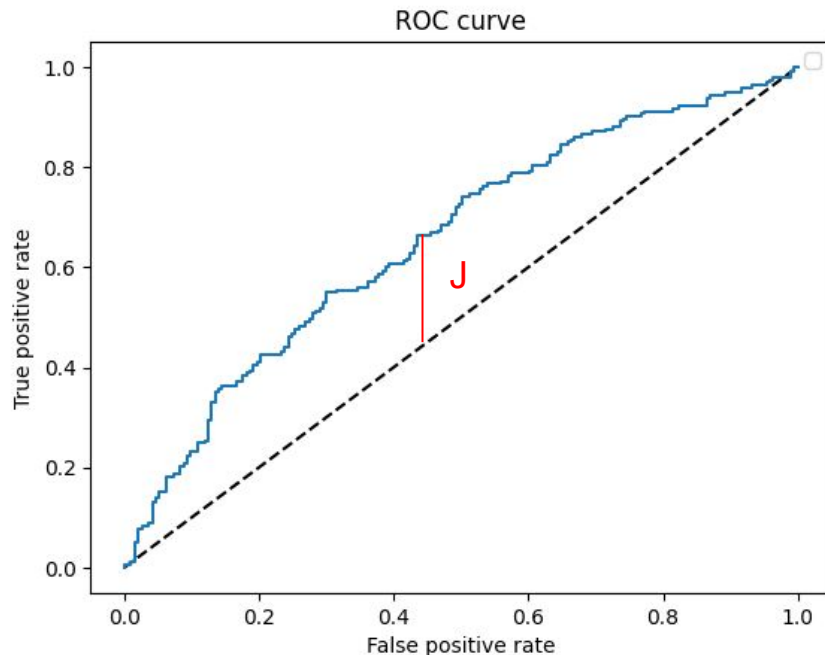


images enhanced

Cancer: 1 & Healthy: 0

Résultats: EfficientNet

ROC Curve et Indice de Youden



Seuil optimal de classification selon l'indice de Youden:
0.46

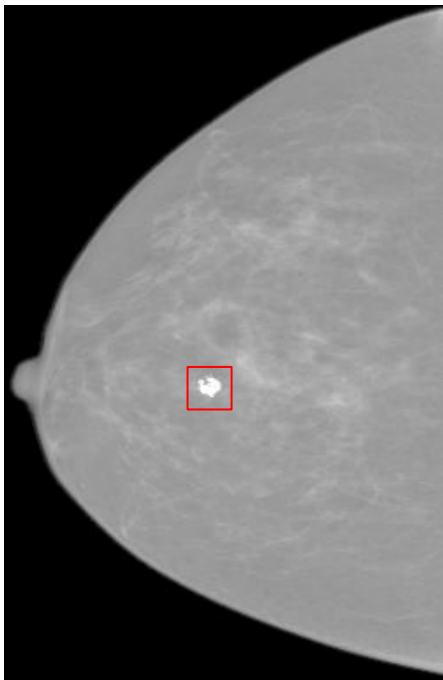
	seuil = 0.5	seuil = 0.46
F1 score	0.51	0.59
Accuracy	0.57	0.61
Precision	0.56	0.60
Recall	0.48 0.5376	0.55

performances avant et après la détermination du seuil optimal

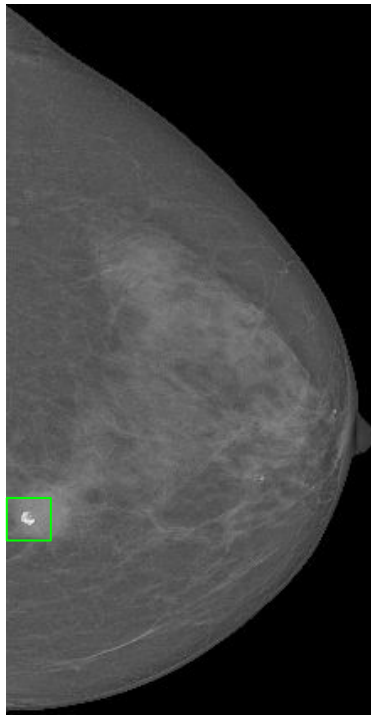
*Courbe ROC des résultats de prédiction du modèle
EfficientNet V2S sur les images non traitées de résolution 512*

Résultats: EfficientNet

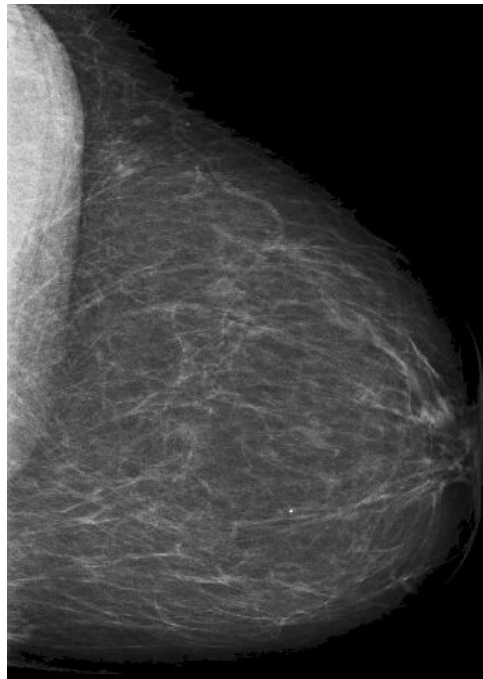
Exemples



False Positive



True Positive



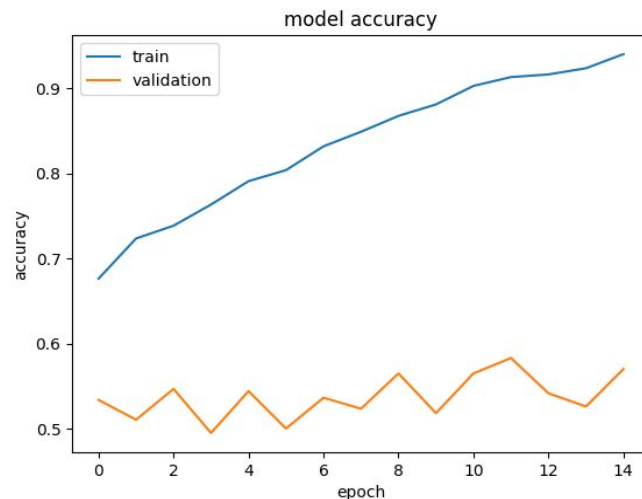
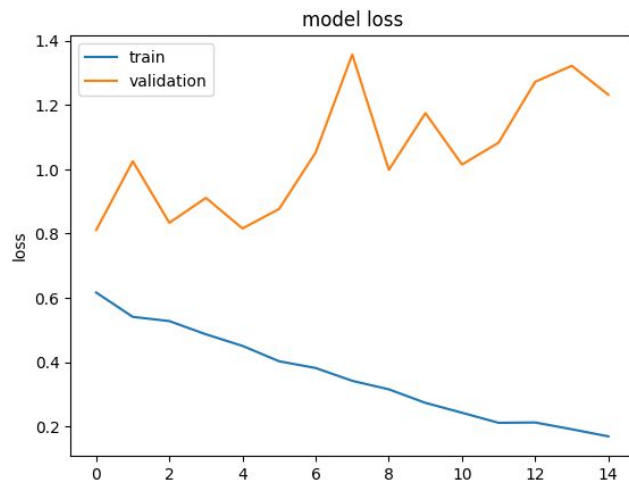
True Negative



False Negative

Résultats: EfficientNet

Meilleur Modèle: Efficient Net B6 avec les images de résolution 512



- Valeurs de perte de validation augmente à travers les epochs
- Différence significative entre l'accuracy train supérieure à l'accuracy validation

➡ sur-apprentissage

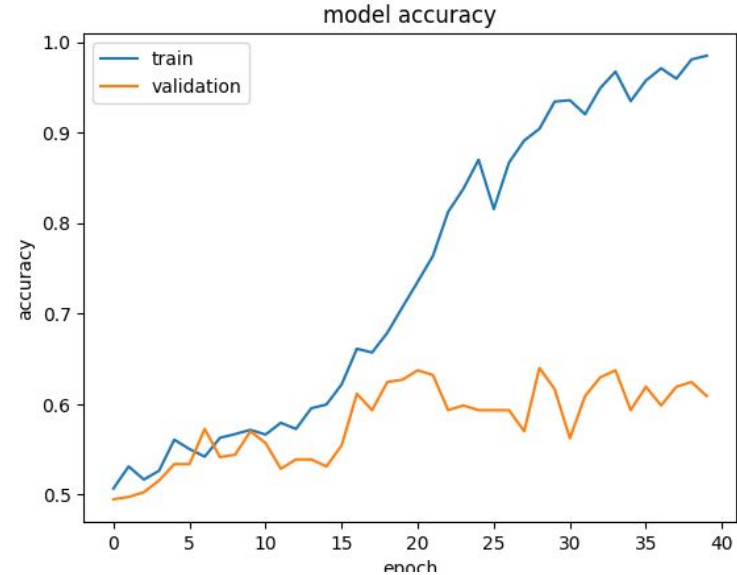
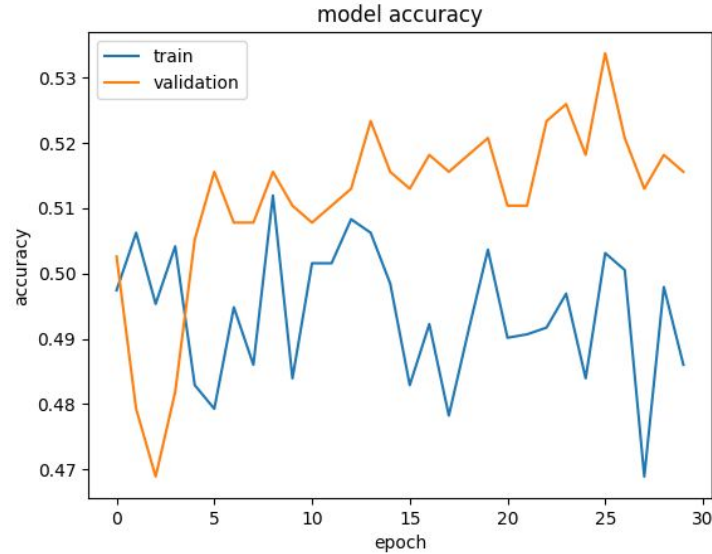
Résultats: DenseNet

Modèle	DenseNet121			
Data	non traitées	norm z-score	norm 2-98	égalisation
F1 score	0.61	0.50	0.53	0.52
Accuracy	0.64	0.56	0.61	0.58
Precision	0.68	0.57	0.66	0.59
Recall	0.56	0.45	0.44	0.46

Résultats du modèle Densenet121 sur des images de résolution 256 avec différentes normalisations

- Les meilleures performances sont observées sur les données non traitées
- On constate une chute importante du Recall avec la normalisation par le 2ème et 98ème percentile

Résultats: DenseNet



Evolution de l'accuracy au cours des epochs pour les données non-traitées en résolution 256, avec Densnet121 : à gauche gelé, à droite dégelé.

- Le sur-apprentissage apparaît au delà de 20 epochs

Conclusion

- Trois méthodes de pré-processing testées
 - Normalisation / Enhancement / Data Augmentation
- Trois modèles de Deep Learning :
 - CNN, EfficientNet, DenseNet
 - Meilleur modèle : Efficient Net B6 (F1 score = 0.58)
- Perspectives:
 - SHAP
 - Nombres d'images cancéreuses plus important (dataset extérieur)
 - Modèle pré entraîné sur des images médicales
 - Méthodes de segmentations
 - Utilisation d'images de plus haute résolution

Merci de votre attention

EfficientNet V2S unprocessed

