

Modèle de Deep Learning pour la Détection du Covid 19 au Sein de Radiographies Pulmonaires

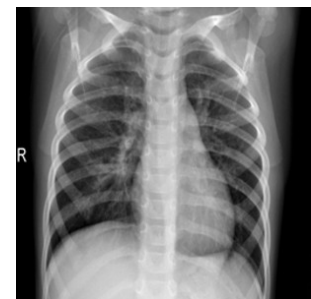
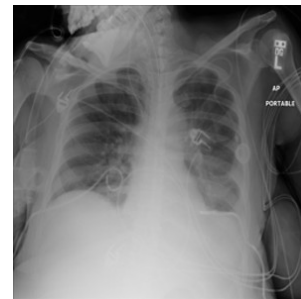
Jérôme Larose
Thomas Pagesy

Sommaire

- Présentation du projet et des données
- Problématiques et solutions apportées
 - Déséquilibre des données
 - Images et masques de tailles différentes
 - Biais de luminosité
 - Sources de biais au sein des images
- Sélection des modèles
 - Modèle « from scratch »
 - Modèle « finetuned »
- Résultats et interprétation
- Pistes d'amélioration

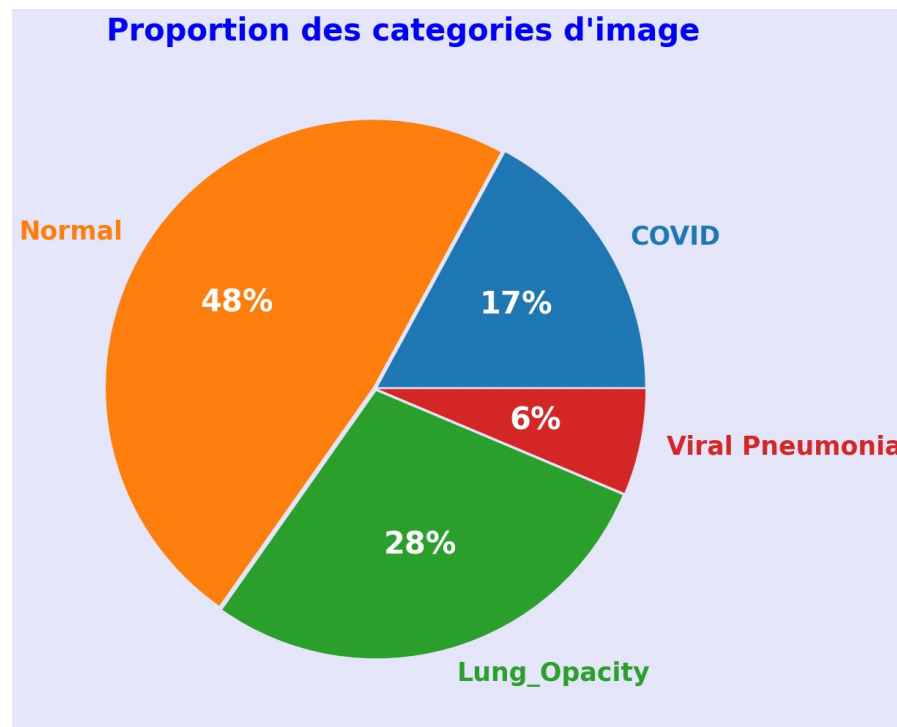
Présentation du projet et analyse des données

- Le dataset se compose:
 - D'images de radiographie de poumons de taille **299x299**
 - De leur masque binaire de taille **256x256** ne laissant visibles que les poumons
- Chaque couple image / masque se trouve dans l'un des 4 sous-dossiers suivant: ***COVID, Normal, Lung Opacity, Viral Pneumonia***
- Le but du projet est de détecter la classe Covid des trois autres
=> Classification **COVID Vs Non-COVID**



Problématiques et solutions apportées: Déséquilibre des données

Proportions des classes déséquilibrées au sein du dataset:
3616 Covid, **10192** Normales, **6012** Lung Opacity et **1345** Pneumonies.



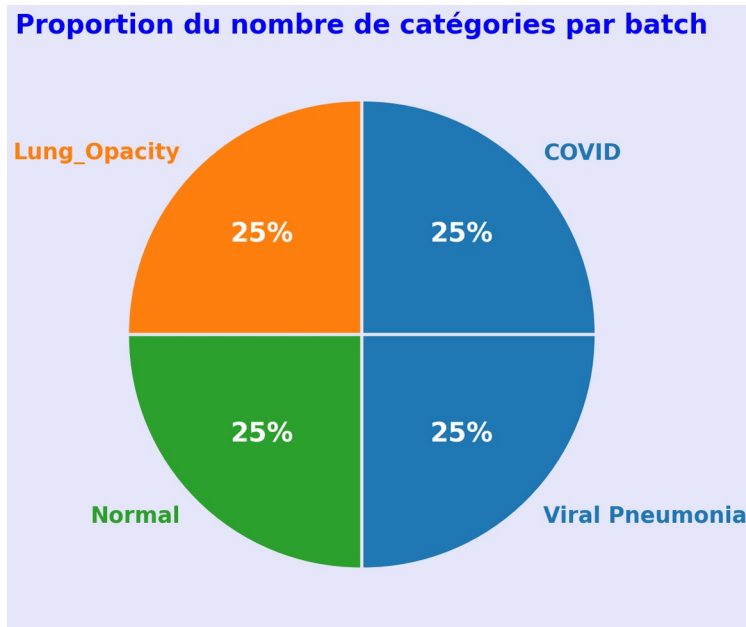
Covid Vs Non-Covid: 17% contre 83%

Problématiques et solutions apportées:

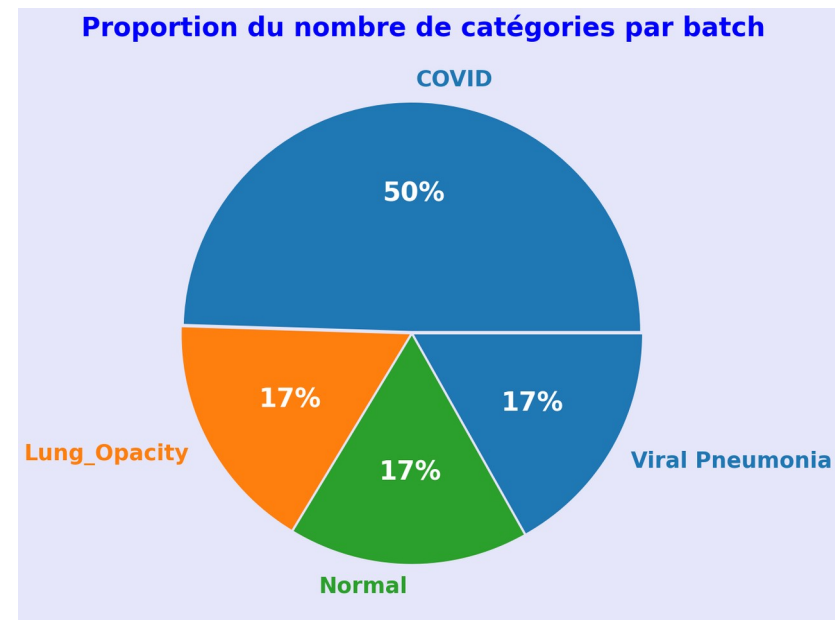
Déséquilibre des données

Solutions apportées:

- Augmentation des données (RandomFlip H-V + RandomRotation de 10%)
- Création d'un générateur de rééquilibrage des batchs



Répartition 4 Classes



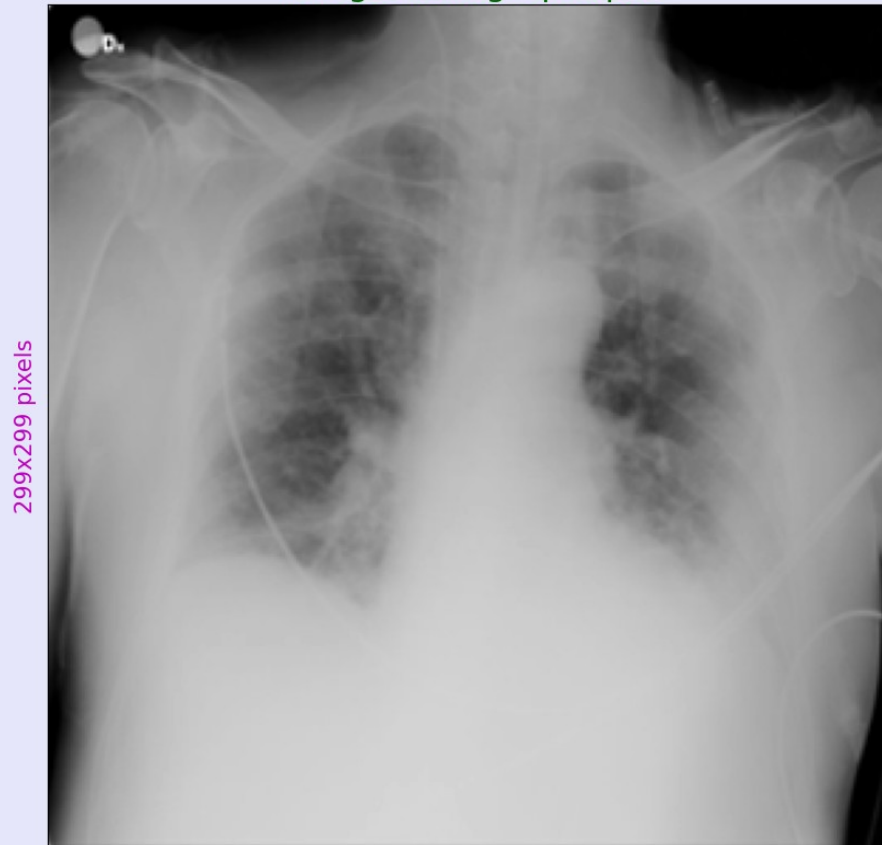
Répartition 2 classes **COVID** Vs **non-COVID**

Problématiques et solutions apportées: Images et masques de tailles différentes

Les images mesurent **299x299** et les masques **256x256**.

Mise en évidence de la différence de proportion entre une image et son masque

Image radiographique



299x299 pixels

299x299 pixels

Masque de l'image radiographique



256x256 pixels

256x256 pixels

Problématiques et solutions apportées:

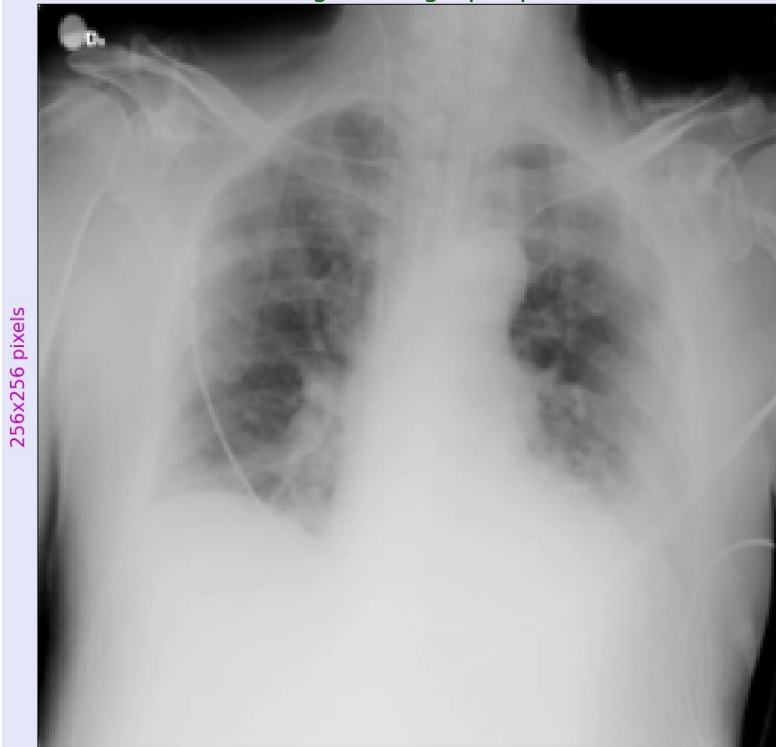
Images et masques de tailles différentes

Solutions apportées:

- les images ont été redimensionnées pour correspondre aux masques **256X256**.

Mise en évidence de la proportion similaire entre une image et son masque

Image radiographique



256x256 pixels

256x256 pixels

Masque de l'image radiographique



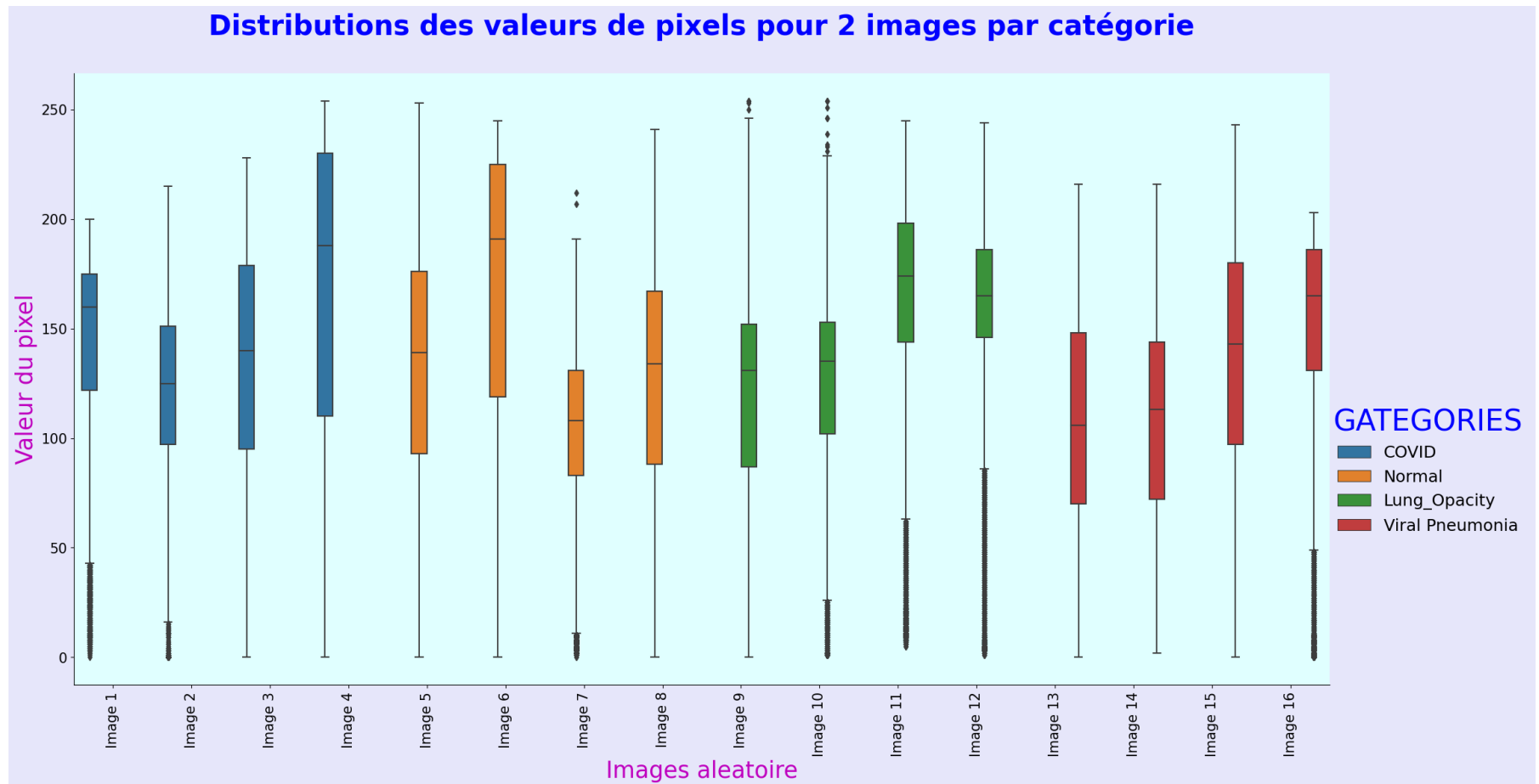
256x256 pixels

256x256 pixels

Problématiques et solutions apportées:

Biais de luminosité

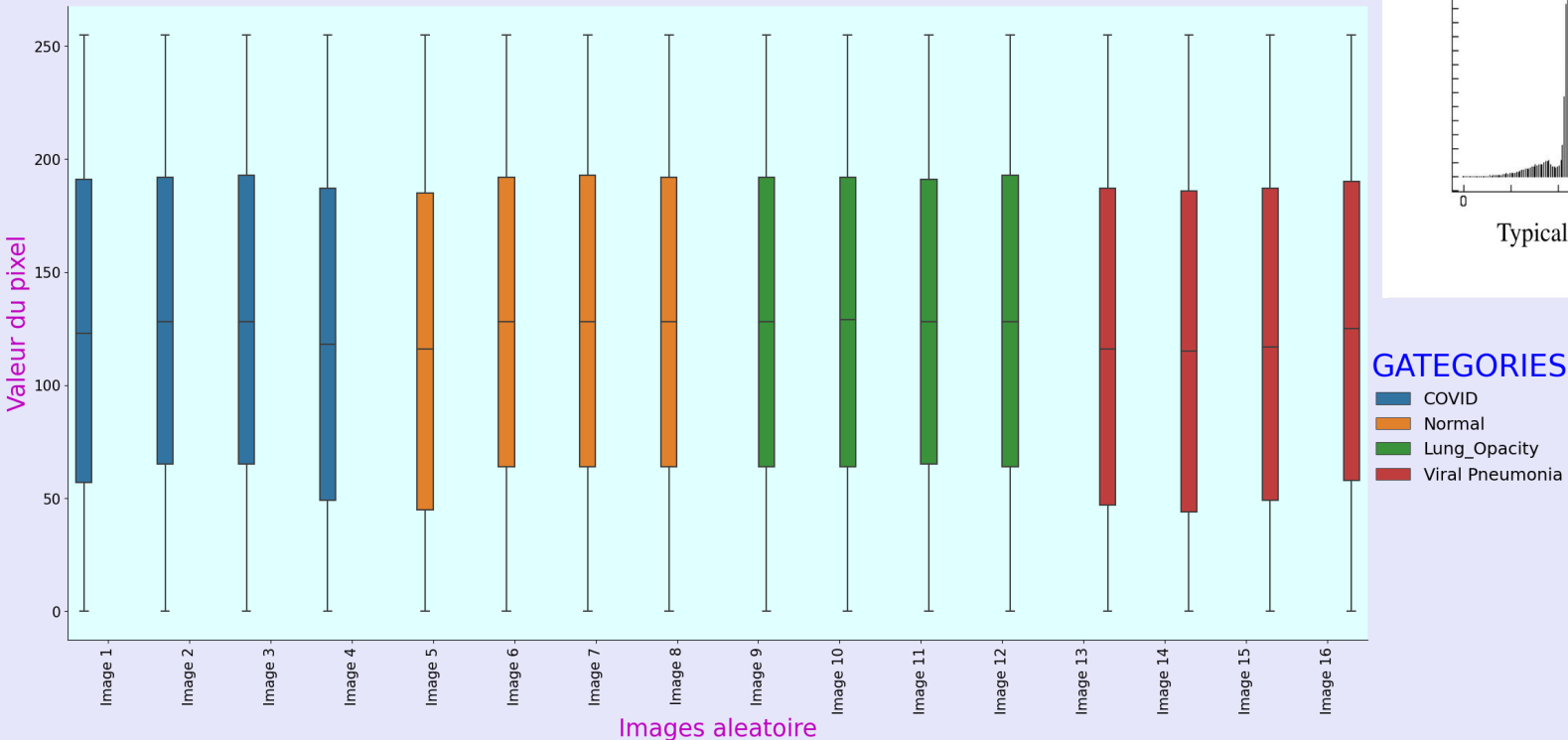
Répartition de la luminosité par pixels entre les images : les **pixels** des images ne sont pas répartis sur la **même échelle**.



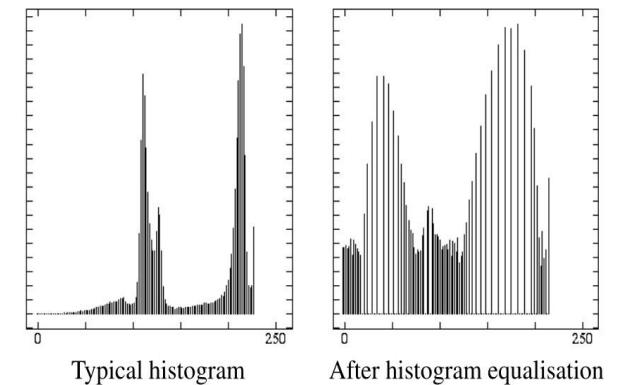
Problématiques et solutions apportées: Biais de luminosité

Égalisation par histogramme => meilleure répartition des
pixels (0,255) sur l'ensemble des niveaux de gris disponibles

Distributions des valeurs de pixels pour 2 images normalisées par catégorie



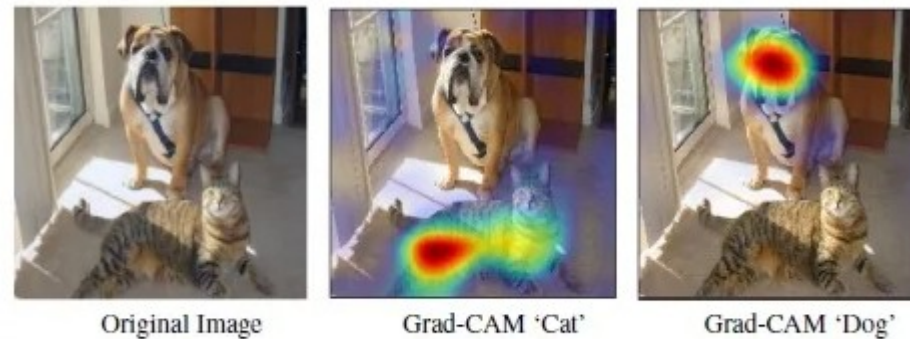
Histogram equalisation examples



Problématiques et solutions apportées:

Sources de Biais au sein de l'image

- Création d'un algorithme de **Grad-CAM** pour visualiser la heatmap des zones d'intérêt du modèle
- Vérifier que le modèle se focalise sur les poumons
- Permet une meilleure **interprétabilité** des résultats pour l'utilisateur



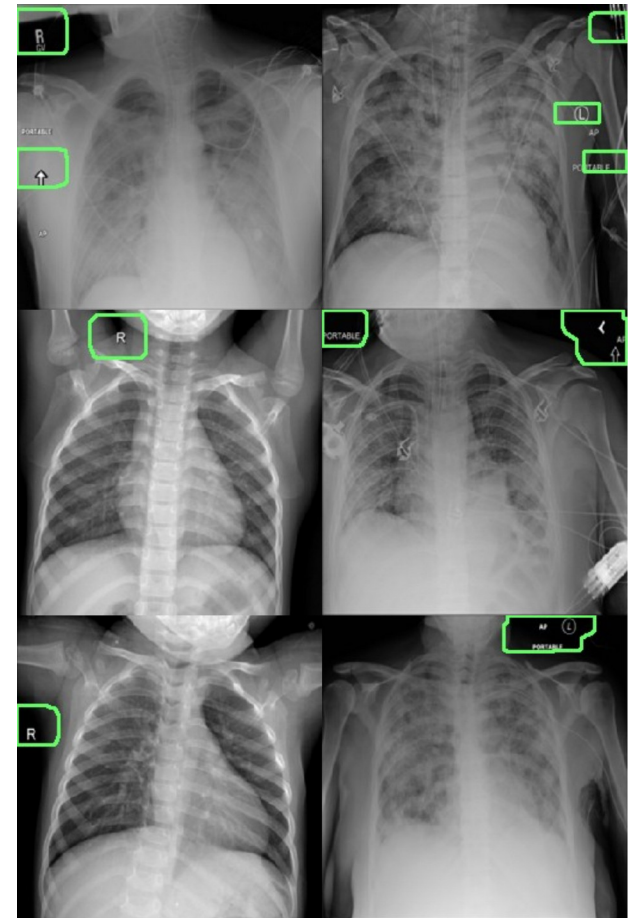
Exemple sur un modèle pour la détection de chiens ou de chats

Problématiques et solutions apportées:

Sources de Biais au sein de l'image

Visualisation avec **Grad-CAM** des zones d'intérêt d'un **CNN naïf** entraîné sur les images non masquées

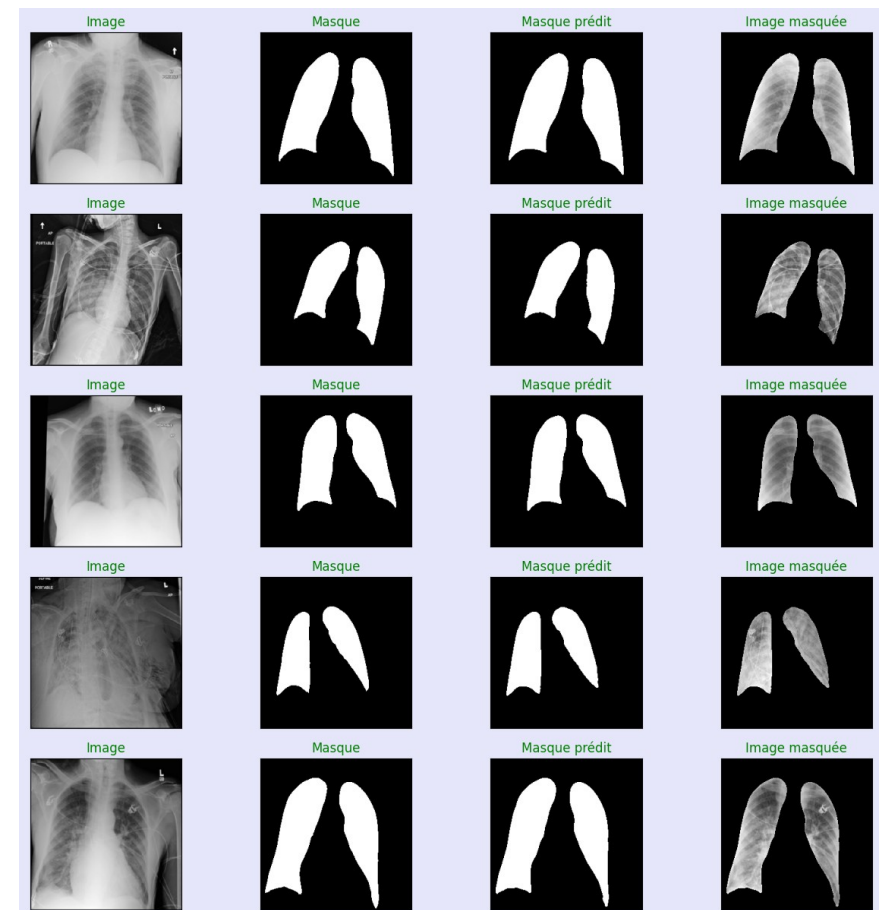
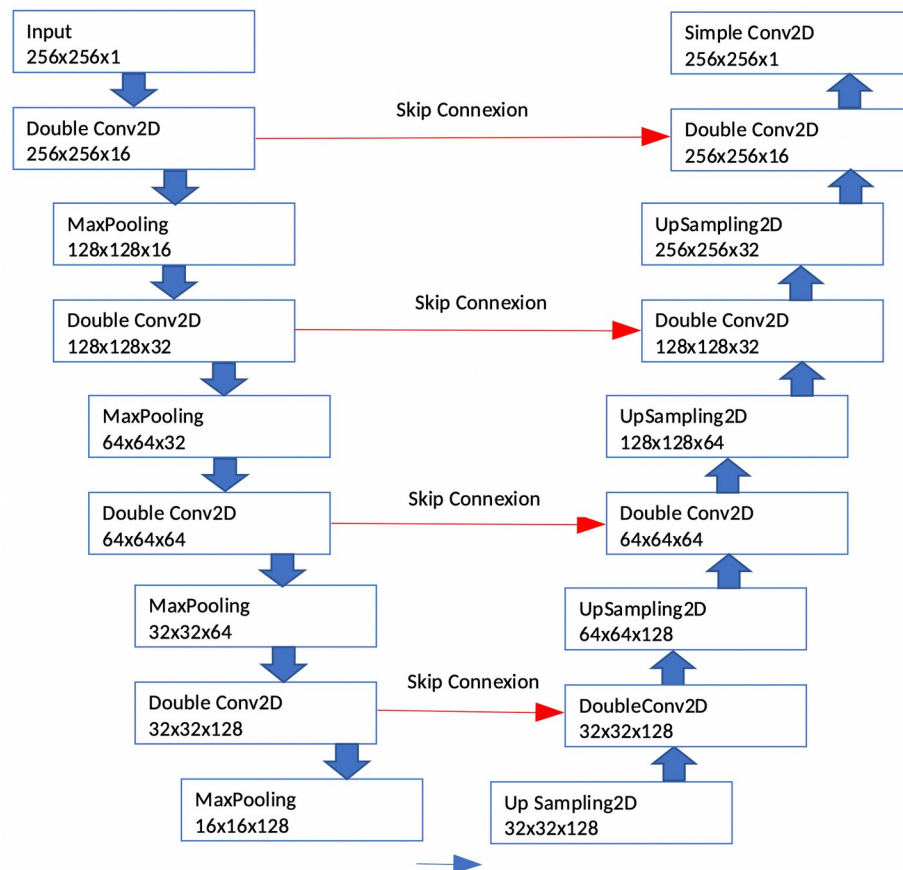
- Le modèle se focalise sur des inscriptions relatives aux sources des images
- Risque de biais lors de la catégorisation des images
- Nécessité de segmenter les images en calculant les masques des poumons



Problématiques et solutions apportées:

Sources de Biais au sein de l'image

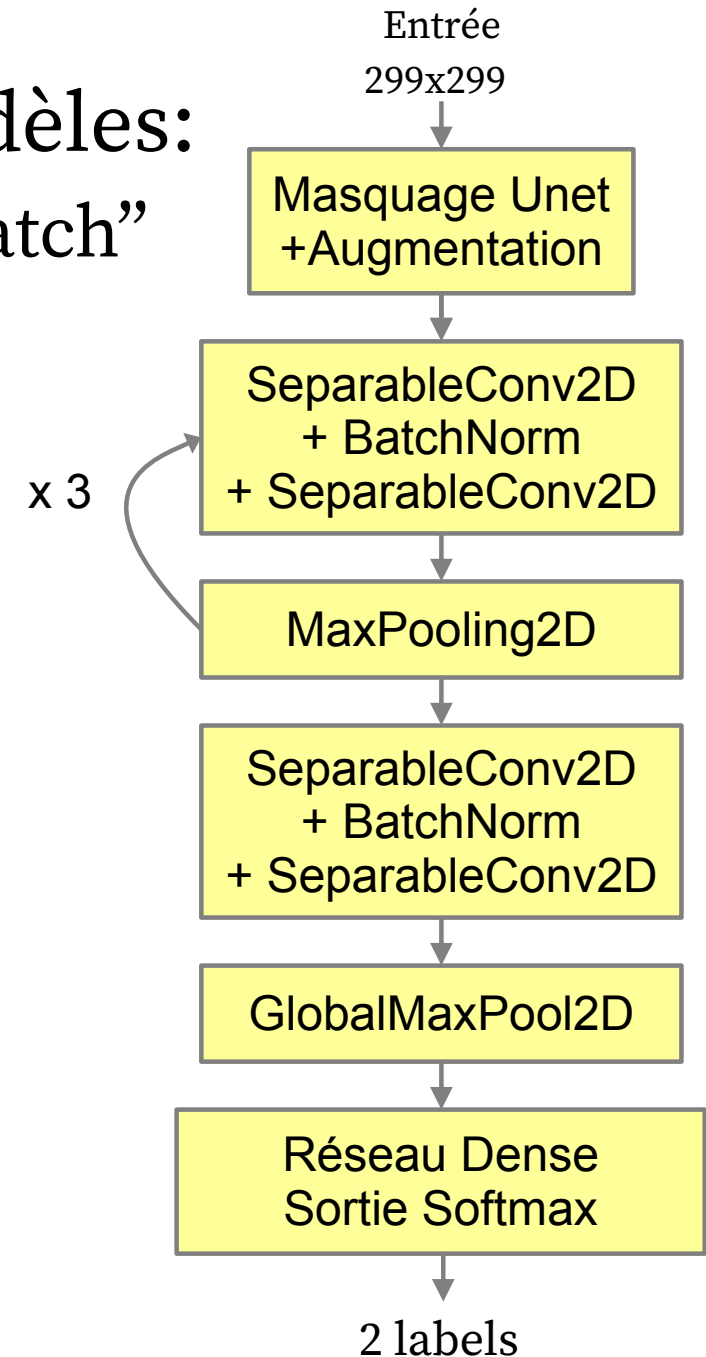
Création d'un modèle **Unet** entraîné sur 90% du dataset pour prédire les masques associés à chaque image => accuracy finale de **99,2%**



Sélection des modèles:

Modèle “from scratch”

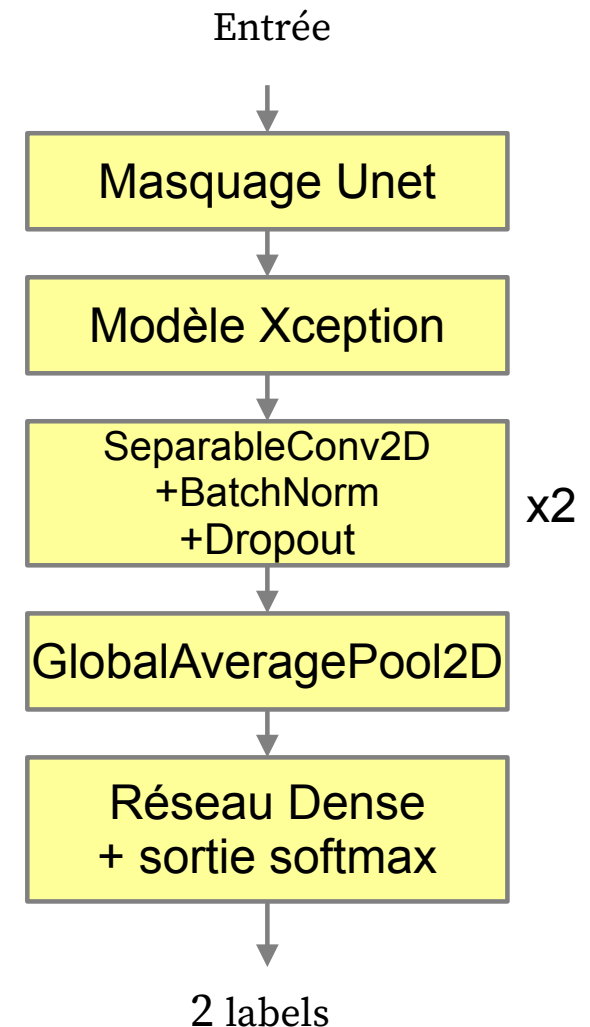
- Utilisation du **Unet** précédemment développé pour calculer et masquer l'image passée en entrée + augmentation des données
- Modèle **CNN** équipé de couches SeparableConv2D (technologie MobileNet/Xception) doublées
- Dimensions de filtres allant de 16 à 128 filtres par couches par puissances de 2
- 678483 paramètres au total (environ 70000 entraînables) => modèle très léger



Sélection des modèles:

Modèle “finetuned”

- Utilisation du **Unet** précédemment développé pour calculer et masquer l'image passée en entrée + augmentation des données
- Modèle **Xception** tronqué à sa **45ème couche**
- Utilisation de couches SeparableConv2D, d'un GlobalAveragePooling2D et d'un réseau dense pour classer la sortie de la partie **CNN** du modèle **Xception**

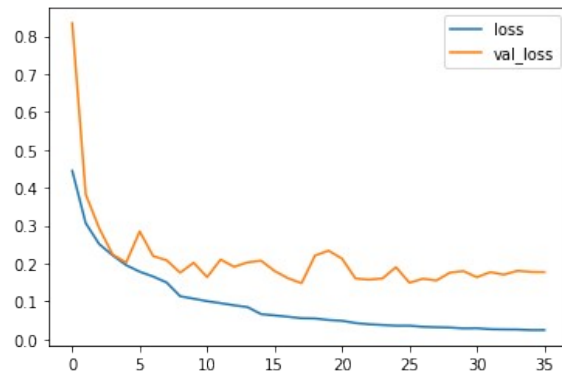


Résultats et interprétation

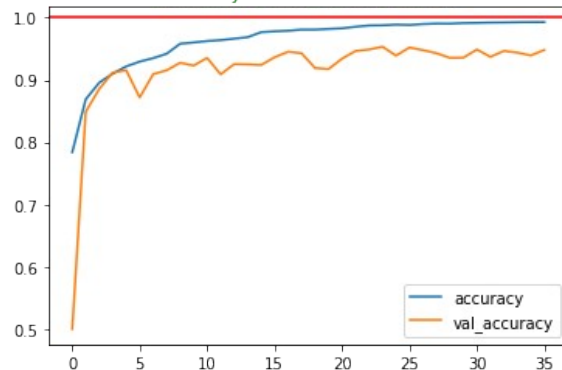
Courbes d'évolution de la fonction de perte et de l'accuracy pendant l'entraînement

Score d'accuracy final sur le set de test de **95,3%** pour le **CNN** from scratch, **97,78%** pour le modèle **Xception finetuned**

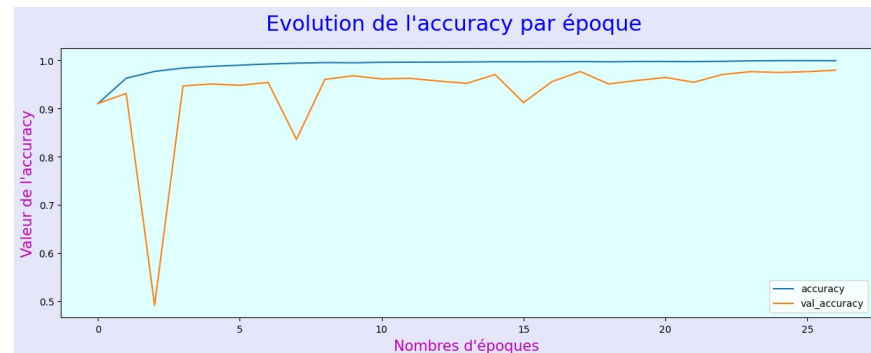
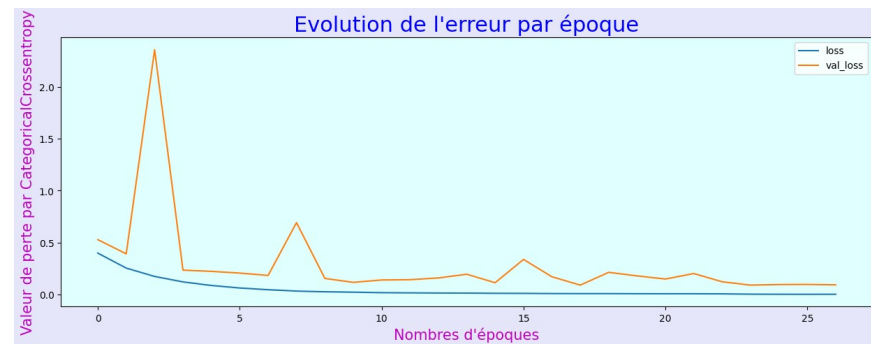
Modèle CNN from scratch



Accuracy: 0.9533442258834839



Modèle Xception finetuned



Résultats et interprétation

Matrices de confusion et scores sur le set de test

Modèle CNN from scratch:

Confusion matrix:

```
[[169  11]
 [ 37 839]]
```

	precision	recall	f1-score	support
0	0.82	0.94	0.88	180
1	0.99	0.96	0.97	876
accuracy			0.95	1056
macro avg	0.90	0.95	0.92	1056
weighted avg	0.96	0.95	0.96	1056

Modèle Xception finetuned:

Confusion matrix:

```
[[ 337   24]
 [   23 1731]]
```

classification report:

	precision	recall	f1-score	support
0	0.94	0.93	0.93	361
1	0.99	0.99	0.99	1754
accuracy			0.98	2115
macro avg	0.96	0.96	0.96	2115
weighted avg	0.98	0.98	0.98	2115

Résultats et interprétation

Commentaires

- Compte tenu du fait que nous avons appliqué une **normalisation par histogramme** et **segmenté les images**, les modèles semblent effectuer un **diagnostic réel et non biaisé**
- Les résultats obtenus sont **très satisfaisants** et sont confirmés à la fois sur les données de validation et de test (**95,3 % et 97,78 %**)
- Le **Grad-CAM** est meilleur pour le modèle **Xception finetuned**: image plus nette et mieux définie. Le fait que la heatmap soit absente de la plupart des images non-COVID, semble indiquer que le modèle fait de la **détection** et non de la classification.
- Les deux modèles sont **complémentaires** en terme de ratio **performances/stockage** (**5,5 mo et 90,4 mo**)
- Une petite démonstration?

Pistes d'amélioration

- Utilisation des deux modèles de manière complémentaire :
 - Avec un **voting classifier**
 - **Extraction de features** des dernières convolution + classifier quelconque
- Entraînement de **3 autres modèles** de type Normal Vs All, Lung Opacity Vs All et Pneumonia Vs All, puis utilisation d'un **Voting Classifier** à l'aide des 4 modèles
- **Généraliser la problématique** du projet à **4 classes** (et non 2) voire encore plus en incluant **d'autres maladies** (extension du dataset)

Merci !

Avez vous des questions ?

Convolutions Séparables

- **Depthwise convolution:** produit de convolution sur chaque channel séparé
- **Pointwise convolution:** produit de convolution par un kernel de taille $(1,1,N)$ où N représente le nombre de filtres de la couche
- **Réduit le nombre de multiplications** à effectuer en effectuant les opérations sur 2 dimensions plutôt que 3 dimensions simultanément

