

Classification d'images par rayons X pour la détection de la pneumonie

Hamou Mohamed Nadir

Mansouri Maël

Résumé : Ce rapport fournit un aperçu détaillé d'un projet de classifications d'images afin de détecter si une personne est saine ou atteinte de pneumonie. Nous y verrons l'approche abordée, les méthodes utilisées et les problèmes rencontrés.

Résumé du sujet

La pneumonie est une infection pulmonaire sérieuse qui affecte les poumons et peut être diagnostiquée par radiographie thoracique. Ce projet vise à développer un système automatisé de détection de la pneumonie à partir d'images de radiographie pulmonaire, en utilisant des approches combinant Machine Learning et Deep Learning.

Nous utilisons le jeu de données **Chest X-Ray Pneumonia**, disponible sur Kaggle ([lien ici](#)), qui contient des radiographies de patients sains et atteints de pneumonie.

Méthodologie adoptée :

- Extraction de caractéristiques :** Utilisation de descripteurs d'image tels que HOG, LBP, GLCM, histogrammes, statistiques d'ordre 1 et filtres de Gabor pour capturer les textures et structures importantes.
- Modèles de Machine Learning :** Entraînement de KNN, Naïve Bayes et SVM, avec validation croisée et optimisation des hyperparamètres pour améliorer les performances.
- Techniques de rééquilibrage des données :** Lutter contre le déséquilibre des classes (3 fois plus de cas de pneumonie que de cas normaux) à l'aide d'oversampling et d'undersampling.
- Deep Learning & Fine-Tuning :** Expérimentation avec des modèles pré-entraînés comme **ViT (Vision Transformer)** et **VGG16** pour améliorer la détection automatique des cas pathologiques. Un ajustement de la fonction de coût a été appliqué pour mieux détecter les cas normaux.

Les ajustements sur les données et les architectures de modèle ont permis d'améliorer la précision globale du système, avec un **meilleur équilibre entre les classes et une robustesse accrue en généralisation**.

Table des matières

Table des matières

Classification d’images par rayons X pour la détection de la pneumonie 1

Résumé du sujet 1

1. Introduction 4

2. Présentation de la classification d’images radiographiques 4

3. Utilisation des descripteurs d’image 5

 3.1 Types de descripteurs utilisés 5

4. Classification des images 6

 4.1 Modèles classiques de classification 6

 4.1.1 Analyse des résultats des classifications 6

 8

 4.1.2 Optimisation de l’HOG 9

 4.2 Classification par deep learning 12

 Comparaison des performances avant et après rééquilibrage 13

 Problèmes rencontrés et solutions 14

5. Analyse comparative des méthodes 16

6. Conclusion et perspectives 16

 Perspectives 16

 Analyse des images mal classées 17

 Roc curve 19

1. Introduction

La pneumonie est une infection pulmonaire pouvant avoir des origines virales, bactériennes ou fongiques. Son diagnostic repose principalement sur l'analyse des radiographies thoraciques (X-ray). Avec les avancées en vision par ordinateur et intelligence artificielle, la classification automatisée des images radiographiques permet d'améliorer la détection précoce et la prise en charge des patients.

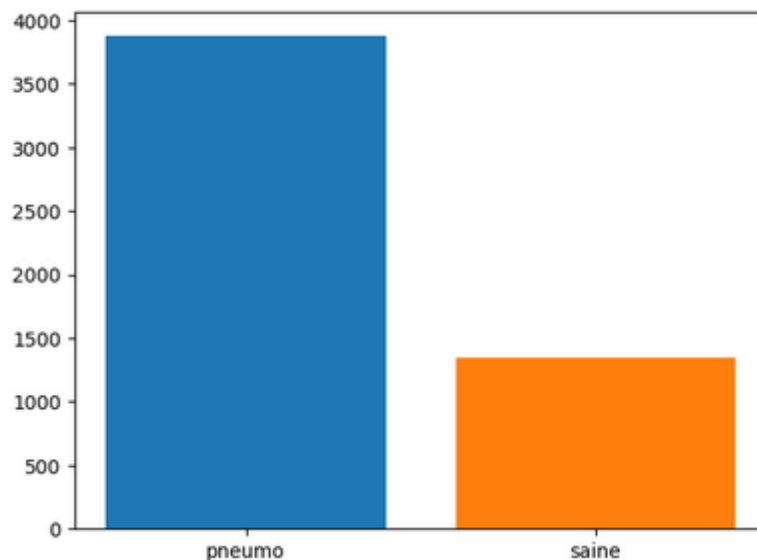
Ce rapport présente une méthode de classification des images par rayons X en utilisant des descripteurs d'image pour l'extraction des caractéristiques et leur classification.

2. Présentation de la classification d'images radiographiques

La classification d'images consiste à attribuer une étiquette (état sain ou pneumonie) à une image en se basant sur des caractéristiques visuelles. Cette approche repose sur deux étapes majeures :

1. **Extraction des caractéristiques** : Représentation des images par des descripteurs pertinents.
2. **Classification** : Utilisation de modèles d'apprentissage pour distinguer les classes d'images.

Les données utilisées dans cette étude sont des images de tailles variables de radiographies thoraciques, comprenant des patients sains et d'autres atteints de pneumonie. Cependant, notre jeu de données présente un **déséquilibre important** : nous avons **trois fois plus d'images de patients atteints de pneumonie que d'images de patients sains**. Cette imbalance peut impacter les performances des modèles et nécessitera l'utilisation de techniques d'ajustement telles que le **rééchantillonnage** ou la **pondération des classes**.



Voici des exemples de radiographies thoraciques :

- Image d'une personne saine :



- Image d'une personne atteinte de pneumonie :



3. Utilisation des descripteurs d'image

Les descripteurs d'image sont essentiels pour représenter efficacement les structures visuelles des radiographies. Ils permettent d'extraire des informations discriminantes sur la texture, la forme et les motifs de l'image.

3.1 Types de descripteurs utilisés

- **Descripteurs de texture :**
 - *Local Binary Pattern (LBP)* : Analyse de la texture en capturant les variations locales de pixels.
 - *Histogram of Oriented Gradients (HOG)* : Capture la distribution des orientations des gradients.
 - *Grey-Level Co-occurrence Matrix (GLCM)* : Analyse de la distribution spatiale des niveaux de gris.
 - *Gabor Filters* : Détection des structures à différentes échelles et orientations.
- **Descripteurs basés sur l'intensité :**
 - *Histogramme des niveaux de gris* : Analyse de la répartition des intensités des pixels.
 - *First Order Statistics* : Mesures statistiques comme la moyenne, l'écart-type, etc...

4. Classification des images

Une fois les descripteurs extraits, des modèles de classification sont appliqués pour distinguer les images saines de celles atteintes de pneumonie.

4.1 Modèles classiques de classification

- **k-Nearest Neighbors (k-NN)** : Classifie en fonction des échantillons les plus proches.

- **Naive Bayes** : Utilise un modèle probabiliste basé sur le théorème de Bayes pour effectuer la classification.
- **Support Vector Machine (SVM)** : Utilise des hyperplans pour séparer les classes dans un espace de caractéristiques. SVM est particulièrement efficace pour les jeux de données à grande dimension, comme les images médicales, et peut être utilisé avec des noyaux pour gérer des séparations non linéaires.

Approche utilisée :

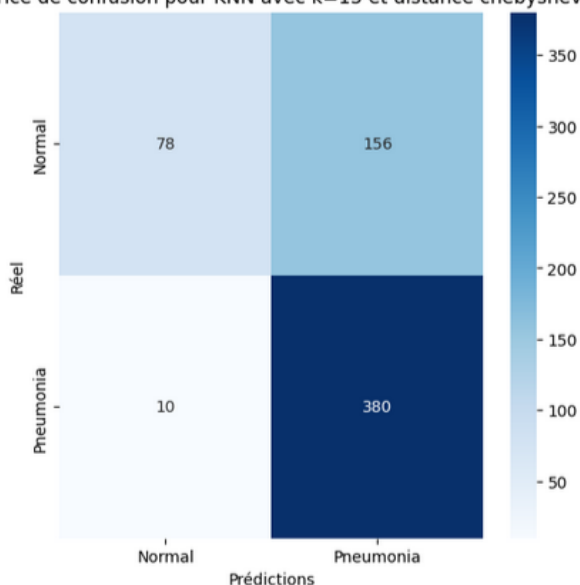
Pour chaque descripteur, nous avons extrait des vecteurs de caractéristiques afin de les utiliser comme entrée pour les modèles de classification. Cependant, pour le descripteur **HOG**, la taille du vecteur généré était particulièrement élevée (**8100 dimensions**), ce qui ralentissait considérablement l'entraînement des modèles et augmentait la complexité computationnelle. Pour remédier à ce problème, nous avons appliqué l'**Analyse en Composantes Principales (ACP)** afin de réduire la dimension à **1500**, permettant ainsi d'accélérer l'entraînement tout en conservant un maximum d'informations discriminantes. Nous avons ensuite testé plusieurs modèles de **machine learning**, notamment **K-Nearest Neighbors (KNN)**, **Naïve Bayes (NB)** et **Support Vector Machine (SVM)**. Afin d'optimiser les performances de **KNN**, nous avons utilisé la validation croisée **k-folds**, ce qui nous a permis d'ajuster les hyperparamètres, notamment le choix de la **métrique de distance** pour améliorer la classification.

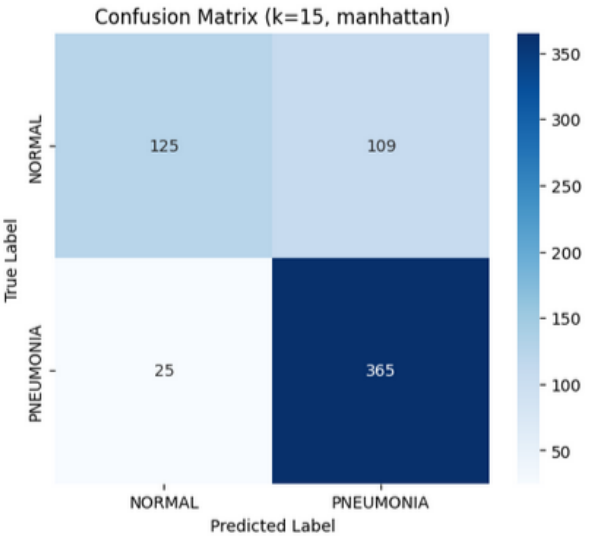
Toutefois, un problème majeur du dataset était le **déséquilibre des classes**, avec environ **trois fois plus d'images de pneumonie que d'images normales**. Cette disparité entraînait un biais dans les modèles, qui avaient tendance à privilégier la classe majoritaire et à ignorer la classe minoritaire, réduisant ainsi le **recall** des images normales. Pour y remédier, nous avons appliqué différentes **techniques d'équilibrage des données**, en combinant **oversampling** et **undersampling**. Ces méthodes ont permis d'améliorer significativement la performance des modèles, en rendant la classification plus équitable entre les deux classes et en augmentant la robustesse globale du système.

4.1.1 Analyse des résultats des classifications

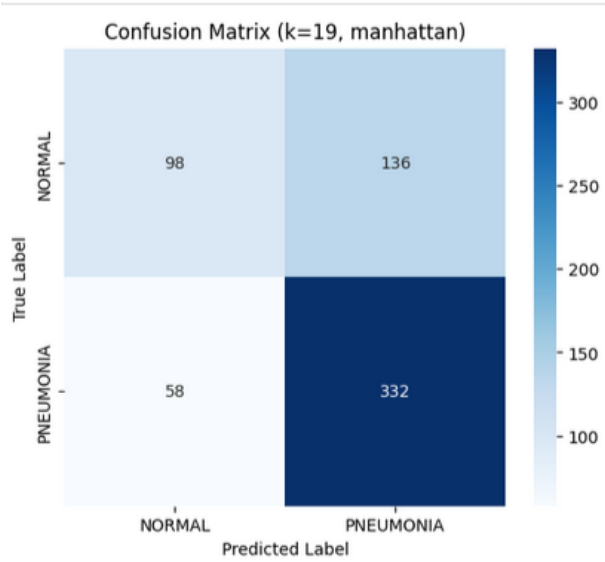
Nous allons comparer l'efficacité des approches basées sur descripteurs classiques en termes de précision, rappel et F1-score.

Matrice de confusion pour KNN avec k=15 et distance chebyshev

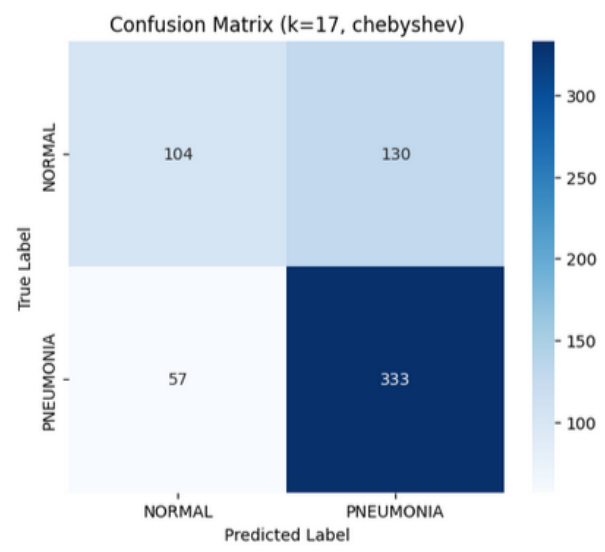




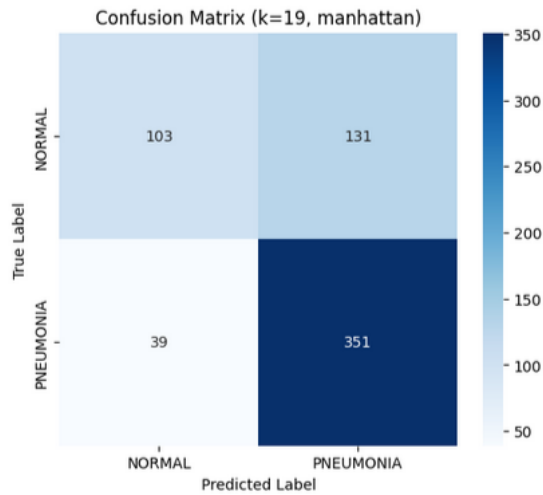
Histogram feature 78.53%



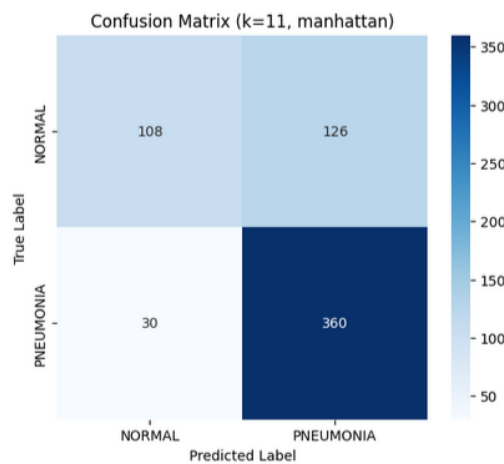
First order feature 68.91 %



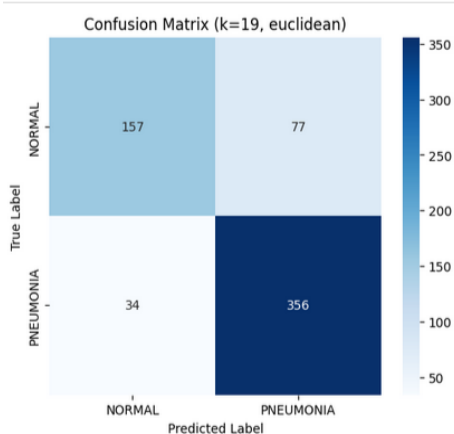
GLCM 70.03 %



Filtre de gabor 72.76 %



LBP feature 75 %



Valeur combiné 82.21 %

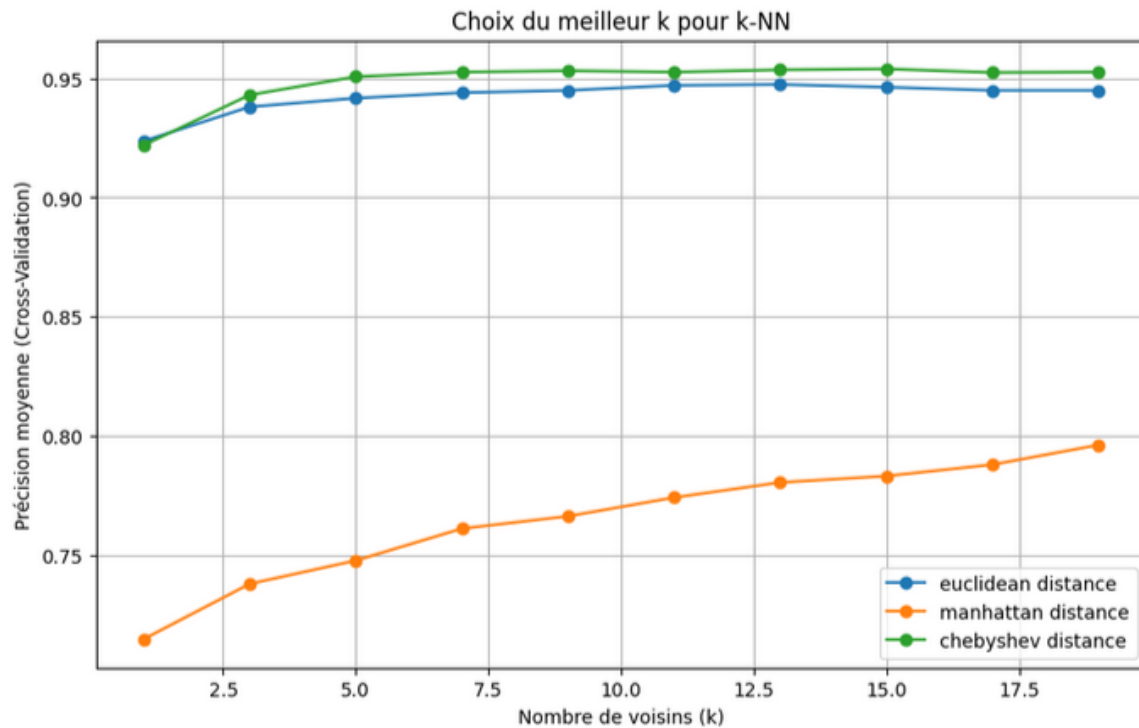
On a donc 4 classificateurs qui ont l'air utile,

l'HOG, l'histogramme et le LBP et le combiné.

Pour le choix d'un des 4 comme « vrai » descripteurs, le meilleur serait d'utiliser HOG, car l'orientation des contours est très utile dans ce cas-là. Nous allons donc rentrer en détail pour ce descripteur.

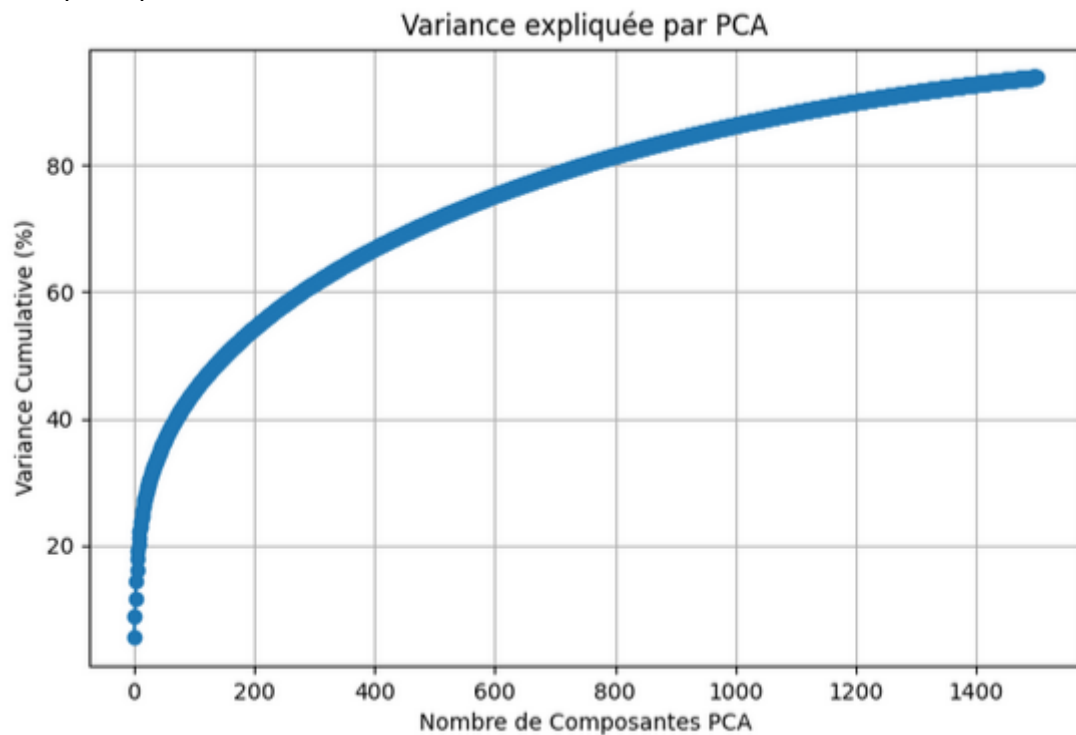
4.1.2 Optimisation de l'HOG

Tout d'abord, on veut optimiser le choix du nombre de voisins et le modèle des distances pour avoir une précision optimale, c'est pour cela qu'en utilisant la technique des k-folds cela nous donne :



Ce qui permet de choisir $k = 15$ et une distance de chebyshev. Ce qui va grandement améliorer la précision

Les vecteurs renvoyés par l'HOG étant grand, on peut aussi faire une analyse en composantes principales pour réduire les vecteurs

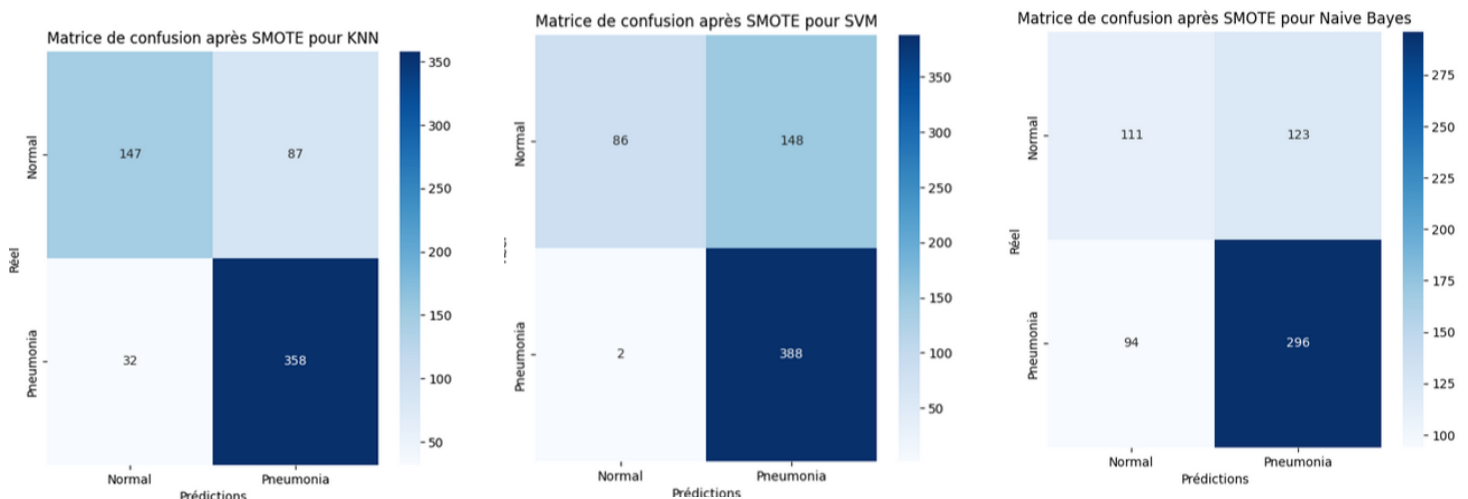


Et enfin
il y a
aussi
un

problème de data imbalance. Que l'on peut corriger ici soit par de l'undersampling(on divise le nombre d'image de pneumonie par 3) ou de l'oversampling(on génère des images artificiellement ressemblant au images de personnes saines)

Le tout combiné, cela nous donne une précision de 80.93 %

Pour ce qui est du modèle, **KNN** reste le meilleur avec une précision de 80.93 %, suivi du SVM avec 76 % et enfin naïve bayes avec 65 %.



4.2 Classification par Deep Learning

Les réseaux de neurones convolutifs (CNN) permettent d'apprendre automatiquement des caractéristiques discriminantes à partir des images.

1. Contexte et objectif

Contexte

- Le projet vise à détecter automatiquement la pneumonie à partir de radiographies thoraciques.
- Les données sont fortement déséquilibrées entre les classes "NORMAL" et "PNEUMONIA", ce qui pose un défi majeur pour le modèle.

Objectif

- Construire un modèle capable de classer correctement les images en deux catégories : "NORMAL" (pas de pneumonie) et "PNEUMONIA" (présence de pneumonie).
- Évaluer les performances du modèle sur l'ensemble de test après avoir appliqué des techniques pour résoudre le déséquilibre des classes.

2. Description du modèle

A. Choix du modèle

Nous avons utilisé un modèle basé sur VGG16, un réseau convolutionnel pré-entraîné sur ImageNet :

1. Pourquoi VGG16 ?

- Avantages :

- Architecture simple et robuste, bien adaptée aux problèmes de classification d'images.
- Pré-entraîné sur un grand ensemble de données (ImageNet), ce qui permet de tirer parti des caractéristiques apprises sur des millions d'images.

- Limites :

- Relativement lourd en termes de nombre de paramètres, ce qui peut entraîner des temps d'entraînement plus longs.
- Peut sur-apprendre si les données ne sont pas suffisamment variées.

1. Architecture personnalisée :

- Nous avons ajouté des couches supplémentaires au-dessus de VGG16 pour adapter le modèle à notre problème de classification binaire :
 - GlobalAveragePooling2D : Réduit la dimensionnalité des caractéristiques extraites par VGG16.

- Couches denses (Dense) avec activation ReLU : Apprennent des représentations spécifiques à notre jeu de données.
- Couche de sortie (Dense) avec activation sigmoïde : Produit une probabilité pour la classe "PNEUMONIA".

B. Fonction de perte et optimiseur

- Fonction de perte : `binary_crossentropy`
 - Adaptée à la classification binaire.
 - Pénalise les erreurs de classification en fonction de la différence entre les probabilités prédites et les vraies étiquettes.
- Optimiseur : Adam
 - Algorithme d'optimisation efficace qui ajuste dynamiquement le taux d'apprentissage.
 - Taux d'apprentissage initial : 0.0001 (faible pour éviter le surapprentissage).

C. Métriques utilisées

- Accuracy : Mesure globale de la proportion de prédictions correctes.
- Precision, Recall, F1-Score : Métriques détaillées pour évaluer les performances sur chaque classe.
- Matrice de confusion : Visualisation des erreurs spécifiques (faux positifs, faux négatifs).

2. Étapes du fine-tuning

A. Préparation des données

1- Prétraitement des images :

- Redimensionnement à (128, 128) pixels pour standardiser la taille des images.
- Conversion en RGB (3 canaux) pour correspondre à l'entrée attendue par VGG16.
- Suppression des canaux supplémentaires (par exemple, 9 canaux) pour ne conserver que les 3 premiers.

2- Rééquilibrage des classes :

- Oversampling : Duplication des images de la classe minoritaire ("NORMAL") pour égaliser le nombre d'échantillons dans chaque classe.
- Résultat : Après oversampling, nous avons obtenu 1248 images pour chaque classe dans l'ensemble d'entraînement.

3- Augmentation des données :

- Application de transformations (rotation, zoom, retournement horizontal) pour enrichir l'ensemble d'entraînement et réduire le risque de surapprentissage.

B. Entraînement du modèle

1- Freezing des couches pré-entraînées :

- Les couches de VGG16 ont été figées pendant l'entraînement initial pour éviter de perturber les caractéristiques apprises sur ImageNet.

2- Fine-tuning :

- Certaines couches supérieures de VGG16 ont été dégelées pour adapter davantage le modèle à nos données spécifiques.

3- Validation :

- Utilisation d'un ensemble de validation pour surveiller les performances du modèle pendant l'entraînement et ajuster les hyperparamètres (nombre d'époques, taux d'apprentissage).

C. Évaluation

- Le modèle a été évalué sur l'ensemble de test pour mesurer ses performances globales et analyser les erreurs spécifiques.

4. Analyse des performances

A. Avant rééquilibrage

- **Résultats :**
 - Accuracy globale : 73.38% .
 - Classe "NORMAL" : Precision et recall nuls (aucune prédiction correcte).
 - Classe "PNEUMONIA" : Bonne performance (precision = 0.73, recall = 1.00).
- **Problème :**
 - Le modèle était biaisé vers la classe majoritaire ("PNEUMONIA"), car il prédisait systématiquement cette classe pour maximiser l'accuracy globale.

B. Après rééquilibrage

- **Résultats :**
 - Accuracy globale : 91.47% .
 - Classe "NORMAL" : Amélioration significative (precision = 0.79, recall = 0.93).
 - Classe "PNEUMONIA" : Performance toujours élevée (precision = 0.97, recall = 0.91).
- **Explications :**
 - L'oversampling a corrigé le déséquilibre des classes, permettant au modèle de mieux détecter la classe minoritaire ("NORMAL").

- L'augmentation des données a réduit le surapprentissage et amélioré la généralisation.

5. Pourquoi ces performances ?

Points forts

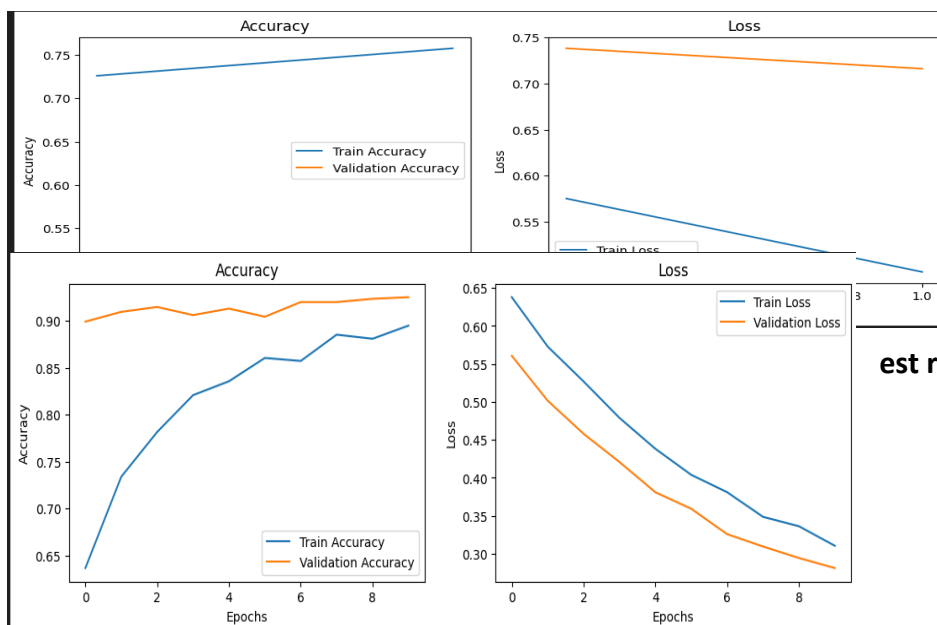
1. **Rééquilibrage des classes** : L'oversampling a permis de corriger le déséquilibre initial, améliorant ainsi la détection de la classe minoritaire.
2. **Modèle pré-entraîné** : VGG16 a tiré parti des caractéristiques apprises sur ImageNet, réduisant le besoin de grandes quantités de données propres.
3. **Augmentation des données** : Les transformations enrichissent l'ensemble d'entraînement, réduisant le risque de surapprentissage.

Limites

1. **Qualité des données** : Certaines images mal annotées ou bruitées peuvent affecter les performances.
2. **Complexité du modèle** : VGG16 est relativement lourd, ce qui peut poser des problèmes pour le déploiement.
3. **Faibles données de validation** : Seulement 8 images par classe dans l'ensemble de validation, rendant difficile une évaluation fiable.

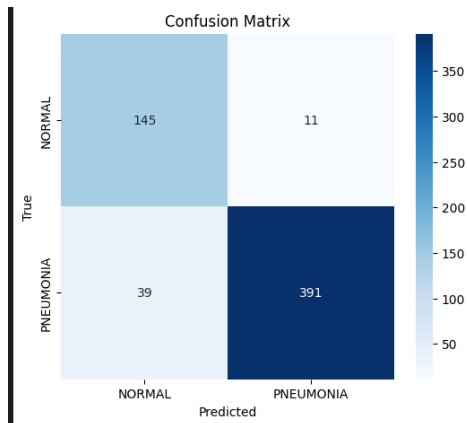
Conclusion

- Le modèle a atteint une accuracy globale de 91.47% après le rééquilibrage des données, avec des performances satisfaisantes pour les deux classes.
- Les métriques détaillées (precision, recall, F1-score) montrent que le modèle est maintenant capable de détecter correctement la classe "NORMAL", contrairement à l'approche initiale.



L'évolution du loss et accuracy du modèle avant l'équilibrage des données

Avec les données équilibrées l'évolution est remarquable



Matrice de confusion du model

Model ViT (Vision Transformers) :

❓ Pourquoi choisir ViT ?

- Le modèle **ViT** est particulièrement puissant pour les tâches de classification d'images grâce à sa capacité à capturer des relations complexes dans les images à travers les transformers. Il est capable de traiter de grandes quantités d'informations spatiales, ce qui le rend bien adapté aux images de radiographies thoraciques.
- Le pré-entraînement du modèle sur **ImageNet** a permis de tirer parti des caractéristiques visuelles générales déjà apprises, ce qui accélère le processus d'apprentissage et améliore la performance, même avec un nombre relativement limité de données spécifiques à la tâche.

❓ Fine-tuning :

- **Congélation des couches inférieures** : Au début, les couches inférieures du modèle pré-entraîné ont été gelées pour conserver les caractéristiques apprises sur ImageNet, et seules les couches supérieures ont été fine-tunées sur les images de radiographies.
- **Ajustement des poids de classe** : Afin de gérer l'imbalance des classes, un poids plus élevé a été attribué à la classe minoritaire (les images normales), ce qui a amélioré le rappel pour cette classe.

Fine-Tuning et Entraînement

Le modèle ViT a été fine-tuné avec les étapes suivantes :

- **Entraînement** : Le modèle a été entraîné avec une fonction de perte **CrossEntropyLoss** modifiée pour inclure des poids de classe. Cela a permis de réduire le biais du modèle envers la classe majoritaire (Pneumonia).
- **Optimisation** : L'optimiseur Adam a été utilisé avec un taux d'apprentissage de 0.001.
- **Validation** : La validation a été effectuée après chaque époque pour évaluer la performance sur l'ensemble de test.

Résultats du Fine-Tuning

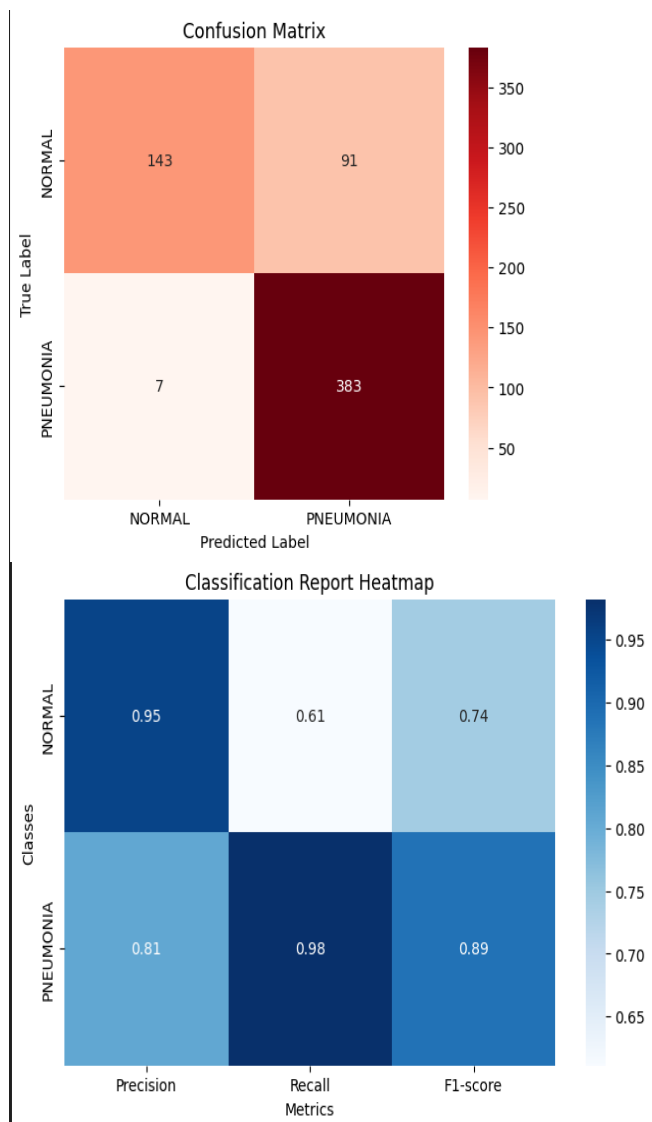
Les résultats de l'entraînement et de la validation après fine-tuning sont les suivants :

Entraînement

- **Exactitude d'entraînement** : Le modèle a atteint une **exactitude de 84.29%** sur l'ensemble de validation.

- **Perte d'entraînement** : La perte a progressivement diminué au fur et à mesure de l'entraînement.

Évaluation sur l'ensemble de test



Analyse des Performances

- **Précision et Rappel** :
 - **Précision** pour la classe normale : 0.95, mais un faible **rappel** (0.61), ce qui signifie que le modèle a du mal à identifier toutes les images normales.
 - **Précision** pour la classe pneumonie : 0.81 et un très bon **rappel** (0.98), ce qui montre que le modèle détecte bien les images de pneumonie.
- **Impact des Poids de Classe** : En attribuant des poids plus élevés à la classe minoritaire (normale), la performance pour cette classe a été améliorée. Toutefois, il reste des marges d'amélioration, notamment en ajustant davantage les techniques de rééchantillonnage et d'augmentation des données.

Conclusion

Le modèle **Vision Transformer (ViT)** a montré une bonne capacité de généralisation pour la détection de pneumonie sur les radiographies thoraciques, atteignant une **exactitude de 84.29%**

avec un bon rappel pour la classe pneumonie. Cependant, des ajustements supplémentaires sont nécessaires pour améliorer la détection de la classe normale, notamment en affinant l'approche de rééchantillonnage et d'augmentation des données.

Améliorations possibles :

- **Augmenter la diversité des données** par des techniques de **data augmentation** (rotation, zoom, etc.) pour aider le modèle à mieux généraliser.
- **Approches supplémentaires de rééchantillonnage**, comme le sur-échantillonnage spécifique aux classes ou l'utilisation de techniques de génération de données.

5. Analyse comparative des méthodes

Approche	Précision
k-NN (HOG)	80.93 %
SVM	76 %
Naive Bayes	65 %
CNN (VGG16)	91.2 %

L'approche **CNN surpasse largement les modèles classiques** grâce à sa capacité à apprendre des caractéristiques complexes. Cependant, elle nécessite plus de puissance de calcul et un grand volume de données annotées.

6. Conclusion et perspectives

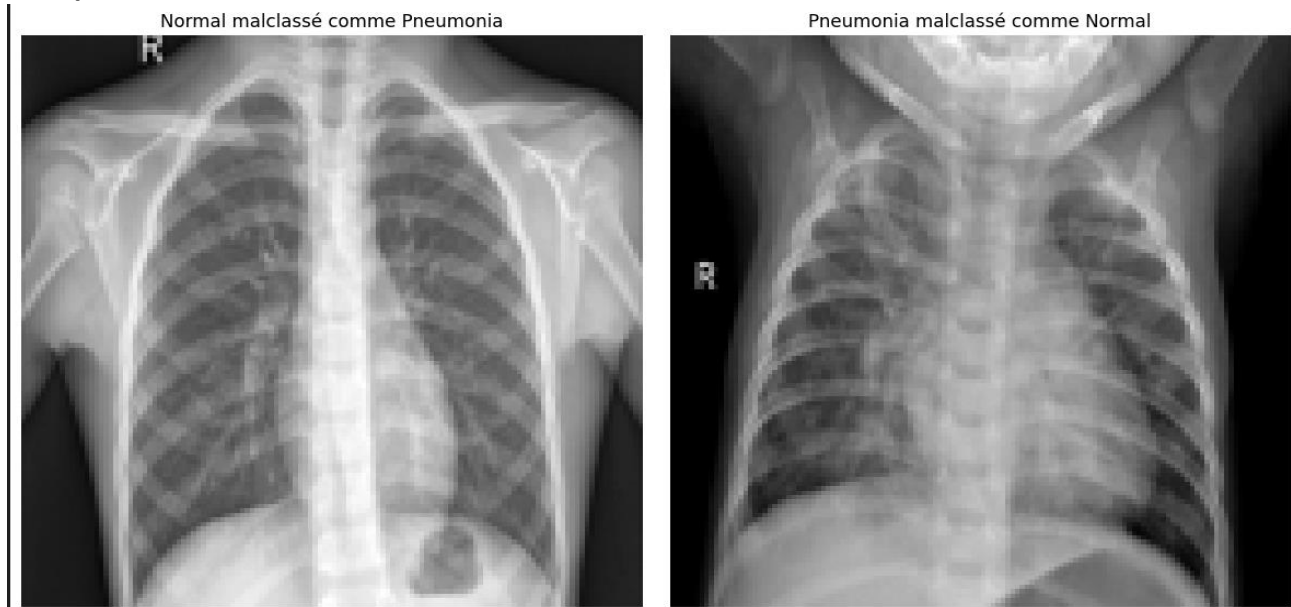
Ce projet a exploré différentes approches de classification des images radiographiques pour la détection de la pneumonie. Nous avons montré que :

- **Les descripteurs classiques (HOG, LBP, histogramme)** permettent une classification efficace avec k-NN, atteignant **80.93 % de précision**.
- **Les CNN surpassent les modèles classiques**, atteignant **91.2 % de précision** après rééquilibrage des données.
- **Le déséquilibre des classes impacte les performances**, mais peut être compensé par des techniques d'augmentation des données et de pondération des erreurs.

Perspectives

- ⑩ Test d'autres architectures CNN (ResNet, EfficientNet).
- ⑩ Amélioration de la robustesse du modèle sur d'autres jeux de données.
- Augmentation de la quantité des données pour bien entraîner le modèle.

Analyse pour les images mal classées :



L'analyse de ces deux radiographies révèle des difficultés rencontrées par le modèle dans la classification des cas normaux et pneumoniques.

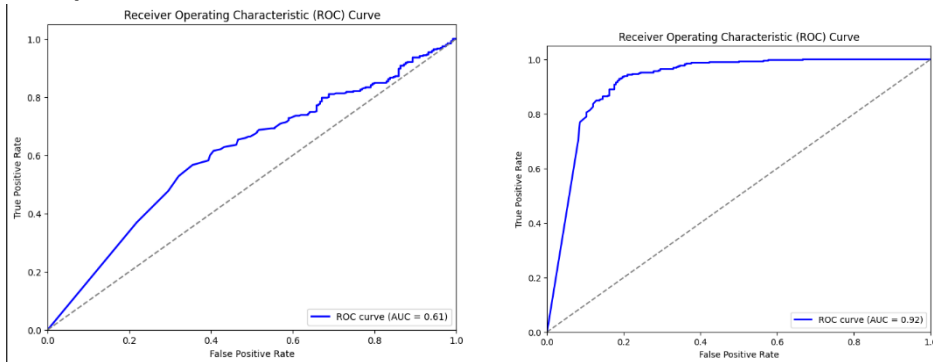
1. Image normale classée comme pneumonie

- L'image semble présenter un léger flou ou un contraste élevé qui pourrait donner l'impression d'opacités pulmonaires, caractéristiques de la pneumonie.
- Certains artefacts ou superpositions osseuses peuvent également tromper le modèle en mimant une atteinte pulmonaire.
- Il est aussi possible que le modèle se base sur des caractéristiques non pertinentes comme l'éclairage ou des variations anatomiques naturelles.

2. Image de pneumonie classée comme normale

- L'opacité pulmonaire semble moins marquée que dans certains cas évidents de pneumonie, ce qui pourrait conduire à une confusion avec une radiographie normale.
- L'image présente une luminosité relativement homogène, ce qui peut masquer certains motifs pathologiques et réduire la sensibilité du modèle à détecter les zones affectées.
- Certaines pneumonies atypiques ou à un stade précoce peuvent ne pas présenter d'opacités évidentes, rendant leur détection plus difficile.

Analyse des Roc curve avant et après l'équilibrage des données :



- **Avant équilibrage** : L'AUC faible (0.61) et la courbe proche de $y = x$ montrent que le modèle n'est pas efficace pour séparer les deux classes, probablement à cause du déséquilibre des données.
- **Après équilibrage** : L'AUC élevée (0.92) et la courbe ROC éloignée de $y = x$ montrent que l'équilibrage des données a permis au modèle de mieux distinguer les deux classes, particulièrement la classe minoritaire, en réduisant le biais vers la classe majoritaire.

Conclusion :

Le **rééquilibrage des données** a permis d'améliorer considérablement la capacité du modèle à identifier les instances de la classe minoritaire (PNEUMONIA), ce qui a conduit à une **AUC de 0.92** et une courbe ROC qui s'éloigne de la droite $y = x$, ce qui est un signe clair d'amélioration de la performance du modèle.