RER – CNN Réseaux de Neurones Convolutifs

# Contexte :

Apprendre à utiliser les réseaux de neurones convolutifs pour classifier les images (deep learning).

# Problématique(s) :

Quelles technologies utiliser pour construire un CNN ?

Comment améliorer les modèles avec les CNN ?

Comment fonctionne un CNN ?

Pourquoi et en quoi les CNN sont-ils plus adaptés à l’imagerie médicale ?

Comment utiliser le transfert-learning ?

# Mots clés :

* CNN (convolution) : les CNN (Convolutional Neural Network) sont des réseaux de neurones spécialisés dans l'analyse d'images, qui utilisent des filtres de convolution et des couches de pooling pour extraire des caractéristiques importantes de l'image en entrée.
* Convolution 2D : La convolution 2D est une opération mathématique couramment utilisée dans les réseaux de neurones convolutionnels (CNN) pour traiter des données en deux dimensions, comme des images. Elle consiste à appliquer un petit filtre (également appelé noyau) sur chaque pixel de l'image pour calculer une nouvelle valeur de pixel. Le filtre est déplacé sur toute l'image en effectuant une multiplication élémentaire et une sommation pour chaque zone de recouvrement, ce qui permet de créer une nouvelle image qui met en évidence certaines caractéristiques de l'image d'origine. La convolution 2D est une étape essentielle dans l'extraction de caractéristiques dans les réseaux de neurones convolutionnels et permet de détecter des motifs tels que des bords, des coins, des textures, etc.
* Pooling : Le pooling (ou sous-échantillonnage en français) est une opération couramment utilisée dans les réseaux de neurones, notamment dans les CNN (Convolutional Neural Networks), pour réduire la taille des données et de limiter la complexité du modèle.
* Padding : Le padding (ou remplissage en français) est une technique utilisée en traitement d'image et dans les réseaux de neurones pour ajouter des bordures de zéros autour d'une image ou d'une carte de caractéristiques.
* Fully Connected : En informatique, "fully connected" fait généralement référence à une architecture de réseau de neurones artificiels, également connue sous le nom de "réseau de neurones multicouches". Dans ce type d'architecture, chaque neurone de chaque couche est connecté à tous les neurones de la couche précédente et de la couche suivante, d'où le terme "fully connected" (entièrement connecté en français). Cela signifie que chaque entrée est connectée à chaque neurone de la couche cachée, et chaque neurone de la couche cachée est connecté à chaque sortie. Les réseaux de neurones fully connected sont couramment utilisés pour résoudre des problèmes de classification, de régression et de reconnaissance de modèles.
* Dropout : Dropout est une technique de régularisation couramment utilisée dans les réseaux de neurones artificiels pour éviter le surapprentissage (overfitting en anglais). Cette technique consiste à désactiver temporairement un certain nombre de neurones d'une couche pendant l'entraînement, de manière aléatoire, de sorte que le réseau ne puisse pas trop se spécialiser sur des caractéristiques spécifiques des données d'entraînement. Cela permet au réseau de mieux généraliser sur de nouvelles données et d'améliorer les performances de la prédiction. Pendant la phase de test, tous les neurones sont activés pour obtenir la prédiction finale.
* Transfert-learning : Le transfert-learning, ou "apprentissage par transfert" en français, est une technique d'apprentissage automatique qui consiste à utiliser un modèle pré-entraîné sur une tâche pour en résoudre une autre. Plutôt que de créer un nouveau modèle à partir de zéro pour chaque tâche, le transfert-learning permet de transférer la connaissance d'un modèle existant à une tâche similaire ou connexe. Cela permet de gagner du temps et des ressources, car le modèle pré-entraîné peut déjà avoir appris des caractéristiques utiles et des modèles de données générales qui peuvent être réutilisées pour résoudre une tâche différente. Le transfert-learning est largement utilisé dans la vision par ordinateur et le traitement du langage naturel.
* Fine tuning : Le fine tuning, ou "affinage" en français, est une technique d'apprentissage automatique qui consiste à prendre un modèle pré-entraîné et à continuer à l'entraîner sur un ensemble de données spécifique à une tâche donnée. Cette technique est souvent utilisée dans le contexte du transfert-learning, où un modèle pré-entraîné est ajusté pour mieux s'adapter à la tâche spécifique pour laquelle il est utilisé. Pendant l'entraînement de fine tuning, les poids du modèle sont ajustés pour minimiser l'erreur sur l'ensemble de données spécifique à la tâche, tout en conservant les caractéristiques de haut niveau apprises lors de l'entraînement préalable. Le fine tuning est souvent utilisé pour améliorer les performances des modèles pré-entraînés dans des domaines tels que la reconnaissance d'images, la reconnaissance vocale, et le traitement du langage naturel.
* Data hub : Un data hub est une plateforme de gestion de données qui permet de collecter, stocker, organiser et partager des données provenant de différentes sources. Il permet aux entreprises de centraliser l'ensemble de leurs données dans un emplacement unique et facilement accessible pour les applications et les utilisateurs. Les data hubs peuvent également inclure des outils de traitement des données, de visualisation, d'analyse et de gouvernance des données pour aider les utilisateurs à comprendre et à exploiter les données. Les avantages d'un data hub incluent une meilleure qualité des données, une amélioration de la productivité et de l'efficacité, et une réduction des coûts liés à la gestion des données.
* MedNet : MedNet est une plateforme open source pour l'entraînement de modèles de réseaux de neurones convolutionnels (CNN) appliqués à la classification d'images médicales, tels que les radiographies, les scanners et les IRM. Cette plateforme est spécifiquement conçue pour les problèmes médicaux, en tenant compte des particularités des données médicales telles que la faible quantité de données d'entraînement, la variabilité inter-patient, et la nécessité de modèles interprétables. MedNet est conçu pour permettre un entraînement de modèles rapide et efficace en utilisant des techniques telles que le transfert-learning et le fine tuning. Il est utilisé dans des applications telles que la détection de pathologies médicales, la segmentation d'images, et la classification de l'état de santé des patients.
* MedNIST : MedNIST est un ensemble de données d'images médicales destiné à être utilisé pour l'entraînement et l'évaluation de modèles d'apprentissage automatique tels que les réseaux de neurones convolutionnels (CNN). Cet ensemble de données contient 10 classes de données médicales différentes, notamment des radiographies de thorax, des scanners cérébraux, des échographies cardiaques, et des IRM. Chaque classe contient 1000 images de haute qualité, avec des annotations pour chaque image indiquant la catégorie à laquelle elle appartient. MedNIST a été créé pour faciliter la recherche et le développement de modèles d'apprentissage automatique pour la classification d'images médicales, en fournissant un ensemble de données standardisé et facilement accessible.
* Yolo : YOLO (You Only Look Once) est un algorithme d'apprentissage profond pour la détection d'objets en temps réel dans les images et les vidéos. Il utilise un réseau de neurones convolutionnels pour détecter des objets dans une image en une seule passe, sans nécessiter de régions d'intérêt prédéfinies ou de segmentation de l'image. L'algorithme fonctionne en divisant l'image en une grille et en appliquant des prédictions de boîtes englobantes pour chaque cellule de la grille. YOLO a été conçu pour être rapide et précis, permettant de détecter plusieurs objets en temps réel avec une haute précision. Il est souvent utilisé dans des applications telles que la surveillance vidéo, la reconnaissance de plaques d'immatriculation et la reconnaissance d'objets dans les véhicules autonomes.
* Data augmentation : La data augmentation est une technique utilisée dans l'apprentissage automatique pour augmenter la quantité de données d'entraînement en créant des versions modifiées des données existantes. Cette technique vise à améliorer la robustesse et la capacité de généralisation des modèles d'apprentissage automatique en introduisant des variations dans les données d'entraînement. Les modifications peuvent inclure des transformations telles que la rotation, la translation, le zoom, le changement de luminosité, le bruit, la distorsion, ou le changement de contraste, entre autres. En utilisant la data augmentation, les modèles d'apprentissage automatique sont capables d'apprendre à reconnaître les objets dans des situations différentes de celles rencontrées dans les données d'entraînement originales, ce qui améliore les performances du modèle sur des données de test ou de validation.
* GAN (réseaux antagonistes génératifs) :Les réseaux antagonistes génératifs (GAN) sont une classe d'algorithmes d'apprentissage automatique qui utilisent deux réseaux de neurones (générateur et discriminateur) pour générer de nouvelles données réalistes à partir de données d'entrée existantes. Le générateur crée des exemples de données synthétiques, tandis que le discriminateur essaie de distinguer les données générées des données réelles. Les deux réseaux sont entraînés ensemble, de sorte que le générateur apprend à produire des données plus réalistes, tandis que le discriminateur apprend à mieux distinguer les données générées des données réelles. Cette adversité permet de produire des données synthétiques de haute qualité, qui peuvent être utilisées dans une variété d'applications telles que la génération de musique, de texte, de photos ou de vidéos. Les GAN sont également utilisés dans la retouche d'images, la conversion de style d'image, et la synthèse de données pour les modèles d'apprentissage automatique.

# **Hypothèses :**

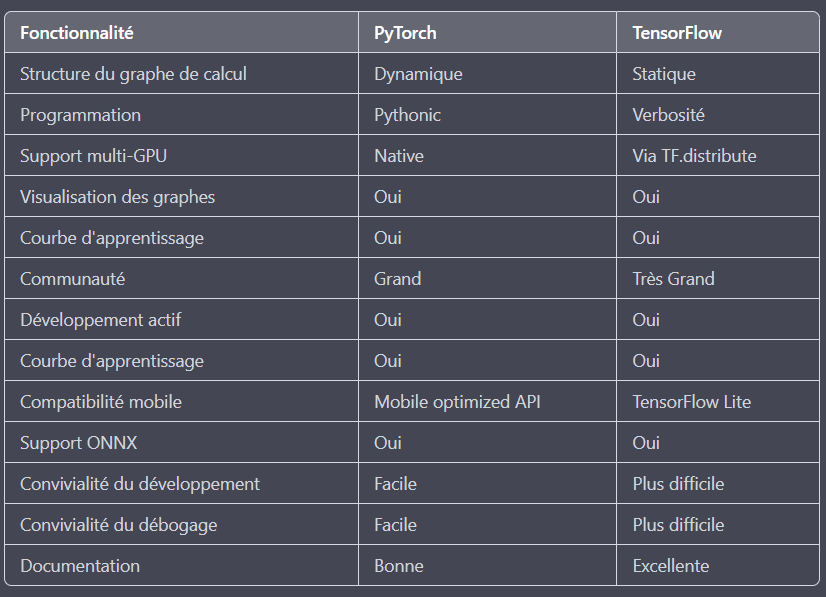
* Adrien : On peut utiliser les CNN avec d’autres entrées que des images
* Axel : On peut utiliser les CNN pour faire de la reconnaissance faciale

Vrai

* Etienne : On peut utiliser les CNN pour faire du réseau antagoniste génératif
* Loic : Le CNN nous permet d'éviter la totalité des actions que nous avons vues avec OpenCV
* Briand : Yolo a les mêmes fonctionnalités qu’OpenCV
* Adeline : Les CNN sont largement utilisés en viticulture de précision.
* Seydou : Les CNN sans transfert- learning ne sont pas efficaces dans la classification des images.
* Nicolas : CNN utilise la bibliothèque OpenCV pour le traitement de images / vidéos
* Loic 2 : Tensorflow est principalement utilisé pour le CNN et se spécialise uniquement dans la reconnaissance d'image.
* Aude : Pytorch est plus spécialisé que Tensorflow pour les CNN
* Osman : Les CNN sont différents des réseaux de neurones de Deep Learning.
* Jean-Paul : Les CNN utilisent les fonctions d’activation.
* Tetyana : On peut modifier ou court-circuiter les étapes du CNN
* Briand 2 : Améliorer un modèle MedNet revient à modifier des données d’entrée.

# Plan d'action :

* Explorer les ressources
* Traiter les mots clés
* Comparer Pytorch et TensorFlow



* Répondre aux problématiques
* Faire les workshops :
* Workshop 1 : la célèbre librairie de deepLearning de google, tensorFlow.
* Workshop 2 : Chat et Chien.
* Workshop 3 : MedNIST.
* Rendre les livrables : RER et workshops