

État de l'art : deep learning au service de la santé

Introduction

L'intelligence artificielle révolutionne le monde du XXI^e siècle dans tous les domaines. Le domaine médical n'y échappe pas, ainsi les évolutions techniques du deep learning permettent une amélioration de la détection de maladie, mais aussi une assistance pour le personnel médical. Nous avons étudié trois récents articles traitant du sujet, et au cours de cette synthèse, nous explorons les dernières méthodes de deep learning et leurs applications médicales.

Méthodes Générales

Les réseaux de neurones artificiels (ANN) sont constitués de plusieurs couches de neurones artificiels. Une couche reçoit des informations et les traite puis les transmet à la couche suivante. Chaque neurone traite les informations en les multipliant par des valeurs de poids et en ajoutant une valeur de biais, puis en appliquant une fonction d'activation pour introduire de la non-linéarité dans le traitement. Cela permet aux réseaux de neurones d'apprendre des relations complexes entre les données d'entrée et de sortie. Pendant l'entraînement, le réseau ajuste les valeurs de poids pour minimiser l'erreur entre les prédictions et les résultats réels. En pratique, les réseaux de neurones peuvent avoir de nombreuses couches cachées et peuvent être utilisés pour résoudre des problèmes de classification, de régression et de prédiction.

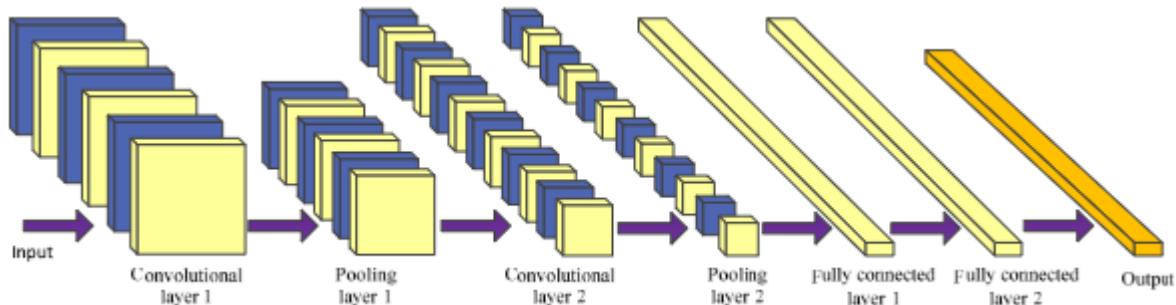
FM-deep learning

C'est dans ce cadre de réseau de neurones artificiel que le FM-deep learning s'inscrit. En 2016 apparaît le modèle FNN qui utilise un réseau de neurone profond, permettant d'obtenir le vecteur dense de chaque caractéristique en pré-entraînant le modèle FM d'obtenir la densité. Les Factorisations Machines (FM) modélisent les interactions entre les variables en utilisant des produits scalaires entre les facteurs correspondants. L'utilisation de ces vecteurs cachés permet de résoudre le problème de la rareté des données. Tous les échantillons contenant des combinaisons de caractéristiques non nulles peuvent être utilisés pour apprendre les vecteurs cachés correspondants.

CNN

Le Convolutional Neural Network utilisent des filtres (noyaux de convolution) qui balayent l'image pour extraire des caractéristiques locales telles que les bords, les coins et les textures, en utilisant un processus appelé convolution.

Les CNN comprennent également des couches de normalisation par lots, des couches d'activation et des couches de pooling.



La couche de normalisation par lots normalise les données en lots pour les faire correspondre à une distribution normale standard, ce qui évite efficacement la disparition du gradient et accélère la convergence.

La couche d'activation traite de manière non linéaire l'entrée à l'aide d'une fonction d'activation, ce qui permet au réseau neuronal de s'adapter à n'importe quelle fonction. La formule est la suivante : $y = ax + b$, où a est la fonction d'activation, x est l'entrée, et b est un paramètre de poids.

La couche de pooling réduit la taille de l'image en échantillonnant uniquement les valeurs maximales ou moyennes dans une région donnée, ce qui permet de réduire encore le nombre de paramètres et de prévenir le surapprentissage.

L'un des premiers modèles développés par LeCun en 1989 est LeNet-5. Cette méthode de deep-learning est populaire et comporte actuellement un large nombre de variantes et de modèles.

RNN

Un Recurrent neural network (RNN) est une classe de réseaux de neurones récursifs qui en prend des données séquentielles comme entrée, effectue une récursion sur la direction de progression de cette séquence et relie les unités cycliques dans une chaîne. Ce type de modèle est adapté pour traiter des problèmes liés à la synchronisation tels que la vidéo, la voix et le texte. De plus, afin d'obtenir une meilleure accuracy, RNNs sont souvent combinés avec d'autres réseaux de neurones.

(image)

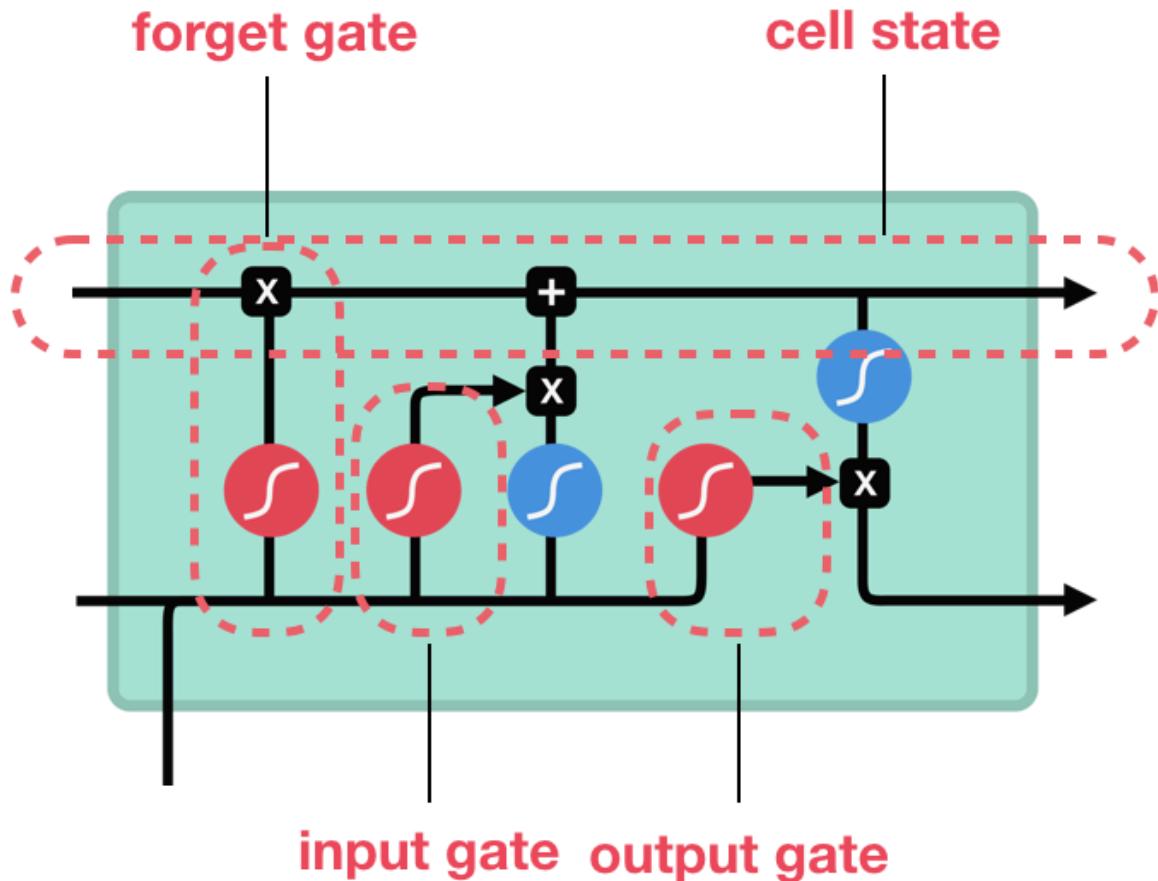
Dans les RNNs, la sortie actuelle d'une séquence est également liée à la sortie précédente. Le réseau se souviendra des informations précédentes et les appliquera pour calculer la sortie actuelle. C'est-à-dire que les noeuds entre les couches masquées sont connectés et que l'entrée de la couche masquée inclut la sortie de la couche d'entrée et la sortie de la couche masquée précédentes.

En théorie, les RNNs peuvent traiter des données séquentielles de toutes longueurs. Cependant, on suppose généralement que l'état actuel est uniquement lié aux états précédents pour réduire la complexité dans la pratique.

LSTM

Long and Short-term memory (LSTM) est un type de RNN introduit en 1997, et qui vise à résoudre le problème du “vanishing gradient” (vanishing gradient problem) présent dans les RNN. Son insensibilité à la longueur de l'écart (gap lenght) constitue son avantage par rapport aux autres RNNs et aux autres méthodes d'apprentissage de séquence.

Ce modèle vise à fournir une mémoire à court terme aux RNN qui peut durer des milliers de pas de temps.



Les LSTM introduisent une nouvelle unité appelée cellule mémoire, permettant au réseau de stocker et d'accéder à des informations sur une période étendue. Cette cellule mémoire est composée de plusieurs portes : une porte d'entrée (input gate), une porte de sortie (output gate) et une porte d'oubli (forget gate). Ces portes régulent le flux d'informations et permettent ainsi de contrôler les informations à retenir et celles à oublier.

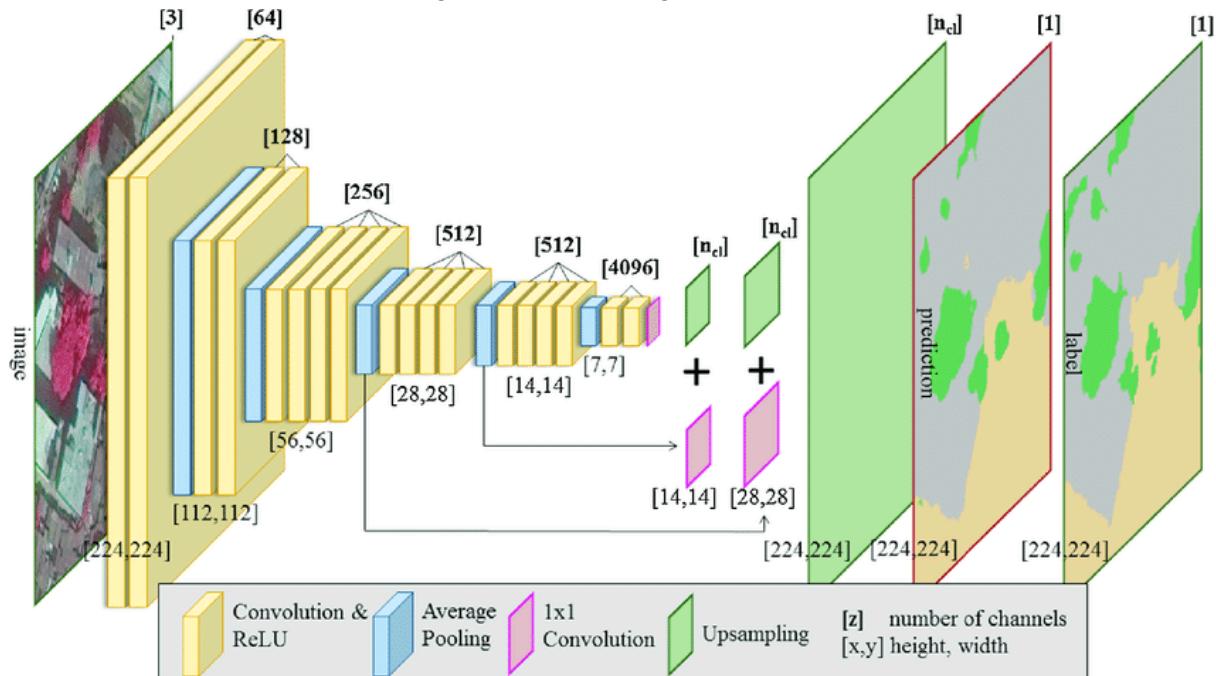
Dans un LSTM, la première étape est la porte d'oubli qui détermine quelles informations doit être oubliées. La prochaine étape est la porte d'entrée qui détermine qui nouvelles informations vont être sauvegardées. Enfin, l'étape finale est la porte de sortie qui détermine les valeurs de sortie.

CNN-LSTM

L'architecture CNN-LSTM implique l'utilisation de couches CNN pour l'extraction de caractéristique sur des données d'entrée combinées avec des LSTM pour faire une prédiction de séquence.

FCN

Un fully-convolutionnal network (FCN) est un framework pour la segmentation sémantique d'image proposé en 2015. Il remplace la couche entièrement connectée derrière le CNN traditionnel par une couche convective. C'est donc un réseau convolutif sans couche complètement connectée qui réalise une classification au niveau des pixels des images, résolvant ainsi le problème de segmentation d'image au niveau sémantique.



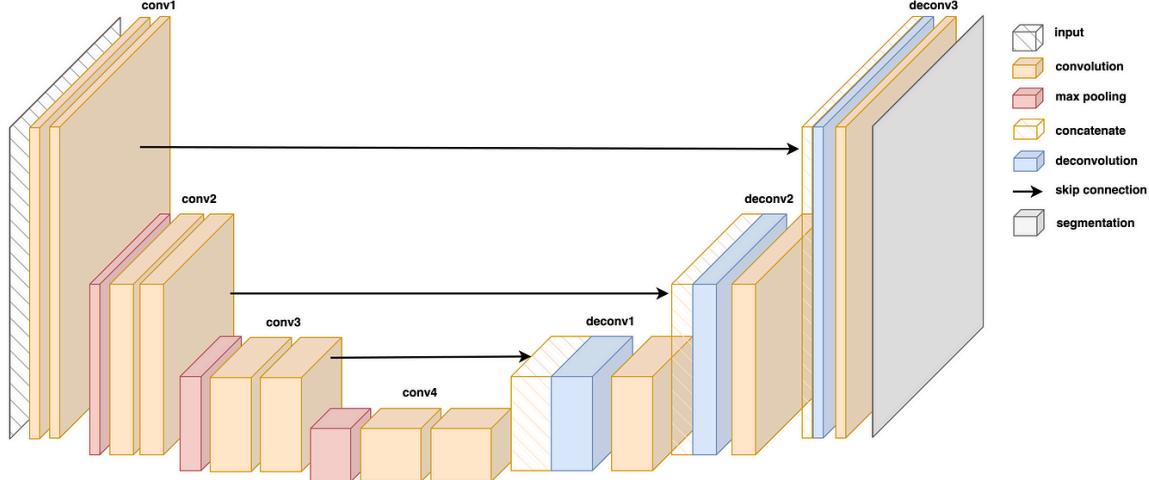
La structure d'un FCN est divisé en deux parties : une partie de convolution et une partie de déconvolution. La partie de convolution comprend des réseaux CNN utilisée pour extraire les caractéristiques. La déconvolution est introduite dans le processus de suréchantillonnage pour échantillonner les caractéristiques extraites de la dernière couche de convolution pour la restaurer à la même taille que l'image d'entrée afin que la prédiction puisse être générée pour chaque pixel tout en conservant l'information spatiale dans l'image d'entrée d'origine. Enfin, la classification des pixels est réalisée sur les caractéristiques échantillonnées et la sortie est une image qui sera labellisée. Un FCN utilise la correspondance entre la sortie et l'entrée, pour donner la classification d'une région sur l'image d'entrée.

Cependant, les FCNs ont encore des limitations. Par exemple, il ne prend pas en compte certaines informations globales, ne peut pas résoudre le problème de segmentation d'instance et n'est pas sensible aux détails de l'image.

U-net

L'architecture U-net, proposée en 2015, est réseau neuronal convolutif développé pour la segmentation d'image médical. Il s'appuie et étend l'architecture du FCN et permet de

classifier chaque pixel d'une image et renvoie en sortie une image segmentée selon la classification des pixels.



Cette architecture consiste en deux chemins, un “contracting path” permettant de capturer le contexte et un “expanding path” (ou chemin d’expansion) symétrique permettant une propagation des informations de contexte vers les couches de résolution supérieure. Le “contracting path”, consiste en plusieurs convolutions, regroupement et échantillonnage de l’image, similaire aux réseaux convolutifs réguliers et permet de fournir les informations de classification.

L’ “expanding path” (ou chemin d’expansion) combine les caractéristiques haute définition du “contracting path” avec la sortie suréchantillonnée pour restaurer la forme de l’image originale, donnant une prédiction par pixels.

L’étape de suréchantillonnage dans l’architecture U-net ajoute un grand nombre de fonctionnalités, permettant de propager plus d’informations concernant la texture de l’image originale dans les couches de haute résolution.

U-net ne possède pas de couche entièrement connectée et peut être utilisée pour la convolution. C'est-à-dire, que la segmentation ne contiendra que des pixels afin que le résultat ne puisse pas être basé sur des informations de contexte manquantes, de sorte que la taille des images d’entrée et de sortie ne soient pas trop grandes.

Application Médicales

L’application médicale de ces différentes méthodes est diverse et variée. Ici nous allons détailler des cas d’applications à commencer par un modèle de FM profond utilisé pour diagnostiquer la maladie d’Alzheimer en combinant simplement un réseau de neurone et une machine de factorisation. Les RNN avec unités LSTM sont également utilisées dans le traitement du signal électrocardiogramme (ECG) et dans l’analyse de données de dossiers médicaux avec des caractéristiques temporelles par exemple c’est utilisé une méthode RNN-LSTM pour prédire la probabilité de réadmission des patients atteints de lupus dans les 30 jours en exploitant la relation temporelle.

Le CNN est applicable dans plein de cas médicaux, car les évolutions technologiques ont amené à utiliser l'imagerie à être omniprésente dans le médical. Ainsi, CNN a déjà été utilisé dans des études pour de l'analyse d'électroencéphalogramme (EEG) ou encore pour de créer un modèle utilisant deux ResNets de 50 couches classant des images de lésions cutanées.

Cancers

Un diagnostic de cancer est un problème de classification dont la nature exige une grande précision de classification. CNN est le modèle ayant le plus de succès parmi les méthodes de deep-learning. Le CNN a été utilisé pour identifier automatiquement les cancers de la peau et les lésions bénignes. De plus, l'utilisation d'un réseau ResNet101 pré-entraîné pour le diagnostic de plusieurs lésions cancéreuses a permis de classer les photos de cancer du côlon et du poumon d'un ensemble de 25 000 images histopathologiques de lésions cancéreuses en cinq catégories.

Le modèle de segmentation d'images FCN appliqué à des images obtenues via des IRM a permis de détecter des cancers de la prostate avec une segmentation marquant les différents états de ce cancer et son évolution. De plus, un réseau neuronal entièrement convolutif hybride (HFCNN) dérivé du FCN a permis de réaliser une segmentation du foie et de détecter les métastases hépatiques sur celui-ci.

L'architecture U-net étant une amélioration et une extension du FCN a permis d'obtenir une meilleure précision de classification avec moins de données d'entraînement. Cette architecture a été utilisée pour segmenter le foie et les tumeurs à partir de CT abdominale. Pour cela, le nombre de filtres et les couches de l'architecture U-net original ont été modifiés pour réduire la complexité et améliorer les performances de segmentation.

Cependant, ces méthodes ont toujours du mal à détecter les petites tumeurs et les tumeurs irrégulières.

Génétique

La classification de séquence d'ADN dans le domaine biomédical est d'une importance cruciale. Cela permet en effet de catégoriser de potentielles menaces comme celle de la COVID par exemple.

Le dataset utilisé dans l'article de recherche est constitué de plus de 80 000 séquences. Cependant certaines séquences sont plus représentées (39 000 pour la Covid-19 contre 9000 pour la Dengue) c'est pourquoi il faut d'abord passer par une étape d'Adaptive Synthetic Sampling Approach afin de réduire les disparités dans le dataset. Trois modèles sont utilisés CNN CNN-LSTM et CNN-LSTM bidirectionnel.

Ils sont testés avec trois encodages :

- label-encoding : les séquences sont représentés par un index données par la fonction LabelBinarizer()
- one-hot vector: une seule bit à 1 les autres à 0
- k-mer encoding : transforme les séquences d'ADN en phrases en langue anglaise on utilise ensuite ensuite un modèle de NLP dessu

Le CNN se révèle être le plus efficace que les deux autres quelle que soit la méthode d'encodage utilisée.

Limitation

Même si les applications médicales du deep-learning sont nombreuses, on dénote différents problèmes.

Premièrement, la donnée médicale, qui permet de nourrir les ANN, est en faible quantité et rarement en bonne qualité. Une très bonne qualité de la donnée va permettre au modèle d' excellentes performances. Les aberrations et erreurs de la donnée ou données manquantes sont fréquentes dans le domaine et nécessitent un expert médical pour labelliser la donnée. De part leur sensibilité, les données médicales sont rarement publiques. Une grande partie des dataset ne sont pas utilisables en pratique, car non représentatif de la situation. De plus, la donnée médicale est déséquilibrée car il y a moins de personnes malades que de personnes en bonne santé. Cela engendre des modèles classent toujours les échantillons dans la classe majoritaire. Par exemple, si un modèle est entraîné à prédire si un patient a une tumeur, lorsque le nombre d'échantillons négatifs (patients sans tumeur) dans l'ensemble d'entraînement est beaucoup plus élevé que le nombre d'échantillons positifs, le modèle diagnostiquera toujours le patient comme n'ayant pas de tumeur. De plus, concernant les cancers, la rareté de certains types de cancers va mener a l'obtention datasets ayant des disparités de classes et les modèles entraînés sur ce type de datasets vont être biaisés vers la plus grande majorité.

Deuxièmement, l'interprétation des résultats est faible. Effectivement en comparaisons à des méthodes statistiques simples, offrant une interprétabilité fiable aux médecins, les algorithmes de deep learning sont souvent considérés comme une boîte noire, car le processus d'extraction et de sélection de caractéristiques n'est pas visible. Cela peut entraîner des résultats peu fiables et des risques pour le patient.

Conclusion

Dans cette synthèse, nous avons vu qu'il existe différentes méthodes de deep-learning. FM-deep learning, CNN, RNN, LSTM, CNN-LSTM, FCN et U-net qui ne sont qu'un échantillon des dernières méthodes reconnu par le monde scientifique. Nous avons aussi exploré les applications médicales qui sont nombreuses, notamment dans la détection de cancers, la génétique, et de façon générale dans t. Mais nous constatons aussi des limitations par le manque d'interpretation médicales des résultats, mais aussi par la donnée d'entrée qui reste en quantité limitée et de faible qualité. Cependant il reste des espoirs pour améliorer la santé de tous. Le monde médical utilise de plus en plus le deep-learning, et la donnée devrait logiquement devenir plus intéressante.

References

- Jiang, Xiaoyan, Zuojin Hu, Shuihua Wang, et Yudong Zhang. « Deep Learning for Medical Image-Based Cancer Diagnosis ». *Cancers* 15, n° 14 (janvier 2023): 3608. <https://doi.org/10.3390/cancers15143608>.
- Miah, Jonayet, Eftekhar Hossain Ayon Ayon, Bishnu Padh Ghosh Ghosh, Md Tuhin Mia, Badruddowza, Md Shohail Uddin Sarker Sarker, et MD Tanvir Islam. « Enhancing Viral DNA Sequence Classification Using Hybrid Deep Learning Models and Genetic Algorithm Optimization ». SSRN Scholarly Paper. Rochester, NY, 11 janvier 2024. <https://papers.ssrn.com/abstract=4692259>.
- Yu, Zengchen, Ke Wang, Zhibo Wan, Shuxuan Xie, et Zhihan Lv. « Popular Deep Learning Algorithms for Disease Prediction: A Review ». *Cluster Computing* 26, n° 2 (1 avril 2023): 1231-51. <https://doi.org/10.1007/s10586-022-03707-y>.