



Projet TC5 - Traitement d'Image

Débruitage d'images par l'auto-encodeur convolutif

Trung VUTHANH
Nour NOUREDINE
Malik KAZI AOUAL

Master 2 AIC
Université Paris Saclays
Paris, France

4 janvier 2019

1 Introduction

L'implémentation de notre projet s'inspire des algorithmes et jeux de données de l'article "Medical image denoising using convolutional denoising autoencoders".

2 Préliminaires

2.1 Auto-encodeur

Les autoencodeurs sont un type spécial d'architectures de réseau neuronal dans lequel la sortie est la même que l'entrée. Les autoencodeurs sont formés de manière non supervisée afin d'apprendre les représentations de bas niveau extrêmement des données d'entrée. Ces caractéristiques de bas niveau sont ensuite déformées de nouveau pour projeter les données réelles. Un autoencodeur est une tâche de régression où le réseau est invité à prédire son entrée (en d'autres termes, modéliser la fonction d'identité). Ces réseaux ont un goulet d'étranglement (bottleneck) serré de quelques

neurones au milieu, les forçant à créer des représentations efficaces qui compressent l'entrée dans un code de basses dimensions qui peut être utilisé par le décodeur pour reproduire l'entrée d'origine. Une architecture d'autocodeur typique comprend trois composants principaux :

- *Architecture de codage* : l'architecture de codeur comprend des séries de couches avec un nombre décroissant de nœuds et se réduit finalement à une représentation de vue latente.
- *Représentation de la vue latente* : la vue latente représente l'espace de niveau le plus bas dans lequel les entrées sont réduites et les informations sont conservées.
- *Décodage architecture* : l'architecture de décodage est l'image miroir de l'architecture de codage, mais dans lequel le nombre de nœuds dans chaque couche augmente et, finalement, les sorties de l'entrée similaire.

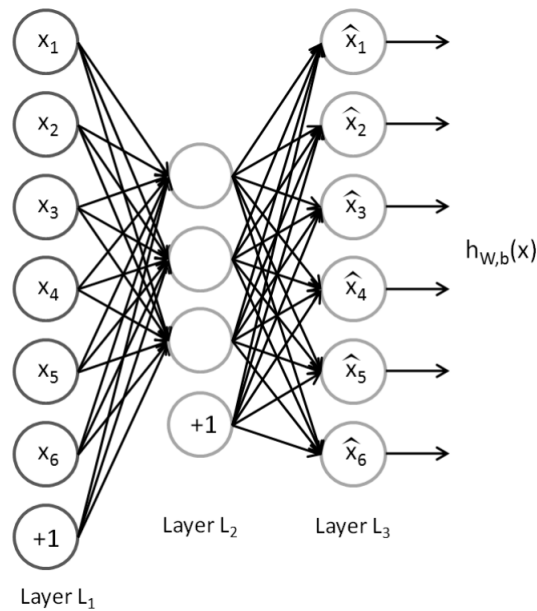


FIGURE 1 – Un auto-encodeur de base

2.2 Auto-encodeur débruiteur

Auto-encodeur débruiteur est une extension à autoencodeur classique, qui est nous forcer le modèle à apprendre la reconstruction de l'entrée compte tenu de sa version bruyante. Un processus de corruption stochastique définit aléatoirement certaines des entrées à zéro, forçant l'auto-encodeur débruiteur à prédire les valeurs manquantes (endommagées) pour les sous-ensembles sélectionnés aléatoirement de modèles manquants. Les autoencodeurs peuvent être empilés pour créer une réseau profond.

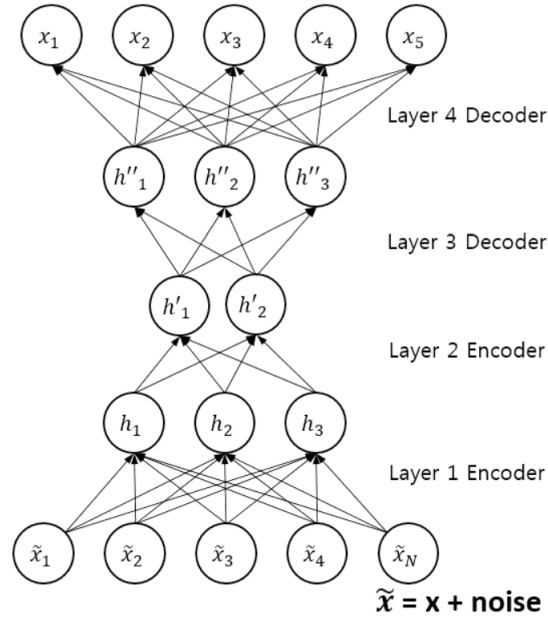


FIGURE 2 – Un Auto-encodeur débruiteur emplilé

2.3 Réseau neuronal convolutif

La principale limitation des réseaux neuronaux standard avec des couches entièrement connectées est le fait qu'ils ne sont pas à l'échelle bien pour les images. Ajout de couches ultérieures augmente le nombre énorme de poids encore plus. Evidemment, il peut rapidement conduire à surapprentissage. La solution à ce problème est de diminuer la taille d'une entrée de manière sage. Les réseaux neuronaux convolutionnels (CNN) sont utilisés explicitement dans les cas où l'entrée est une image. Ils utilisent des couches supplémentaires au début du réseau neuronal pour réduire la taille et prétraiter une image.

Comme l'architecture ci-dessous, les couches convolutionnelles et de regroupement (pooling) représentent des caractéristiques de haut niveau de l'image d'entrée. La couche de regroupement réduit la taille d'une image pour contrôler le surapprentissage. En outre, les couches convolutionnelles et de regroupement sont toujours valables à utiliser pendant l'algorithme de rétropropagation afin que le réseau neuronal peut être encore formé en utilisant des approches de descente de gradient.

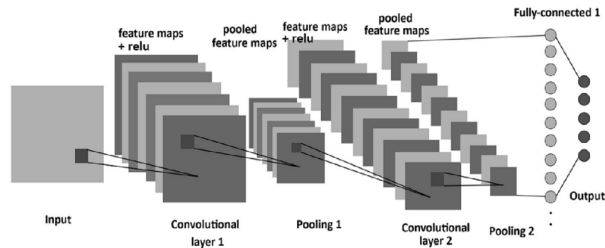


FIGURE 3 – Architecture de CNN

3 Prétraitement des données

3.1 Jeux des données

- Pour l'entraînement du réseau, nous utilisons deux datasets : une base de données de mammogrammes (MMM) et une base de données de radiographie dentaire (DX). MMM a 322 images de résolution 1024×1024 et DX a 520 des images de rayons X céphalométriques recueillies auprès de 400 patients avec une résolution de 1935×2400 . Totalemment, il y a 842 images pour le training. Les images aléatoires des deux jeux de données sont affichées ci-dessous :

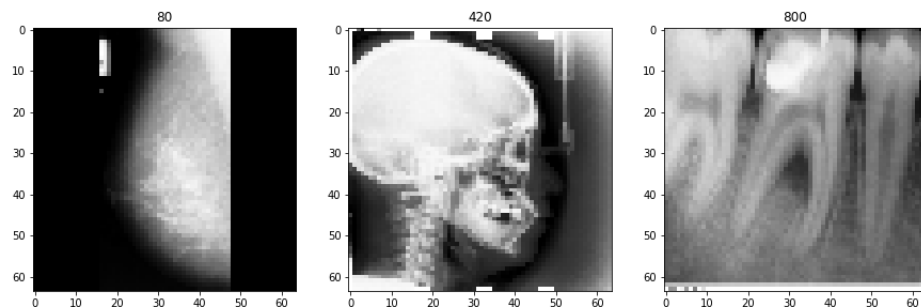


FIGURE 4 – Échantillon aléatoire d'images médicales à partir de jeux de données

3.2 Paramètres de bruitage

Toutes les images ont été traitées avant la modélisation. Prétraitement consistait à redimensionner toutes les images à 64×64 pour raisons de ressources informatiques. Nous rajoutons 2 types de bruits : Gaussien et Poisson.

type de bruit	paramètres de corruption
Gaussien	$\rho = 0.5, \mu = 0, \sigma = 9$
Poisson	$\rho = 0.5, \lambda = 20$

TABLE 1 – ρ est la proportion de bruit introduite, σ et μ sont l'écart type et la moyenne de la distribution normale et λ est la moyenne de la distribution de poisson

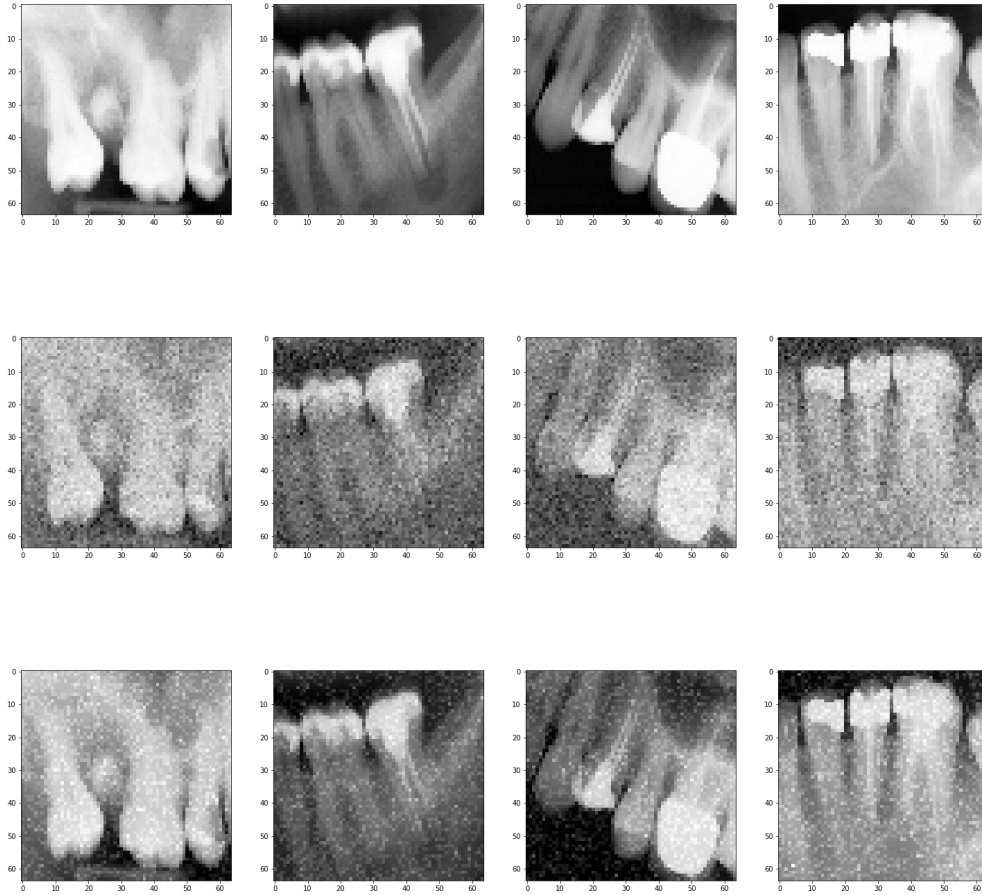


FIGURE 5 – La rangée supérieure affiche des images réelles, la deuxième rangée montrant la version bruyante gaussienne et la troisième rangée montre les images avec bruits de poisson

4 Auto-encodeur débruiteur convolutionnel (CNN DAE)

Après avoir simultanément perturbé toutes les images avec du bruit, des jeux de données corrompus ont ensuite été utilisés pour la modélisation. L'architecture ci-dessous a été utilisée :

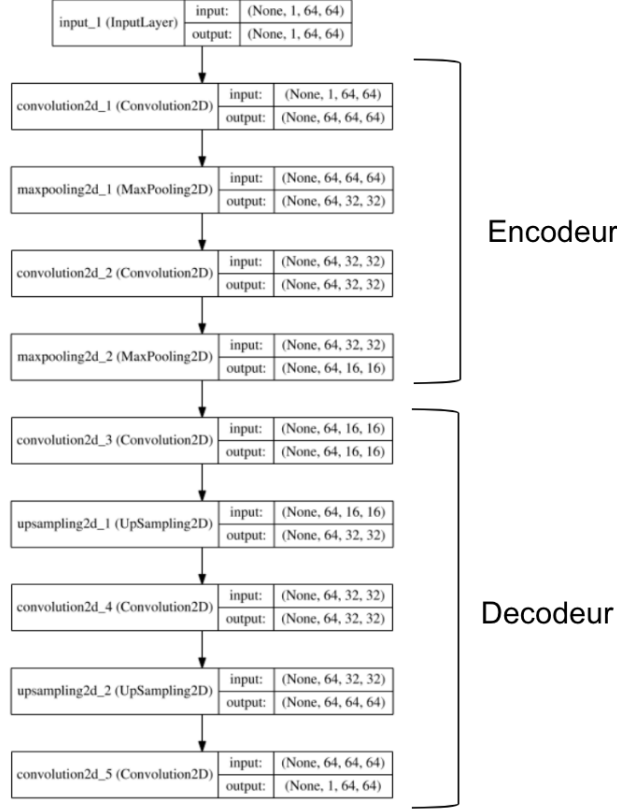


FIGURE 6 – L’architecture de CNN DAE utilisé

Les couches de suréchantillonnage dans le décodeur reconstruisent la taille d’entrée. Enfin, la dernière couche convolutionnelle avec fonction d’activation sigmoïde génère une image décodée.

Comme un roman, le ”leaky ReLU” est utilisé pour fixer les limitations du ReLU standard. Le ReLU standard $\max(0, x)$ provoque parfois un neurone pour rester à zéro et ne jamais être activé dans les prochaines itérations de descente de gradient. Le ”leaky ReLU” résout ce problème par la définition suivante de la fonction d’activation : $1_{x<0}(\alpha * x) + 1_{x \geq 0}(x)$.

Le dataset est divisé en 577 images pour training, 145 images pour validation et 120 images pour le test. Le réseau s’en sert de l’entropie croisée binaire pour la perte au niveau de pixel par pixel entre une image réelle et sa version reconstruite.

Pour l’évaluation de la performance du modèle, nous utilisons 2 métriques spéciales : PSNR et SSIM. Le terme PSNR (Peak signal-to-Noise Ratio) est une expression pour le rapport entre la valeur maximale possible (puissance) d’un signal et la puissance de distorsion du bruit qui affecte la qualité de sa représentation. Le PSNR est généralement exprimé en termes de l’échelle logarithmique :

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$

$$MSE = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

FIGURE 7 – Equation de Peak Signal to Noise Ratio

Le SSIM (structural similarity index measure) mesure la différence perceptuelle entre deux images similaires. En tant qu'indice composite de trois mesures, le SSIM estime les effets visuels des la luminance de l'image, le contraste et d'autres erreurs restantes, collectivement appelées changements structurels. Pour l'original et codé signaux x et y , le SSIM est donné :

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma$$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$s(x, y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

FIGURE 8 – α , β et $\gamma > 0$ contrôlent l'importance relative des chacun des trois termes dans SSIM et l , c et s sont luminance, composants de contraste et de structure. μ_x et μ_y représentent la moyenne des l'image codée, σ_x et σ_y sont des écarts types et σ_{xy} est la covariance de deux images.

5 Résultat empirique

Nous avons entraîné le modèle avec la taille de batch de 10 et 100 époques, Les fonctions de perte (loss) de bruit gaussienne et poisson sont ci-dessous :

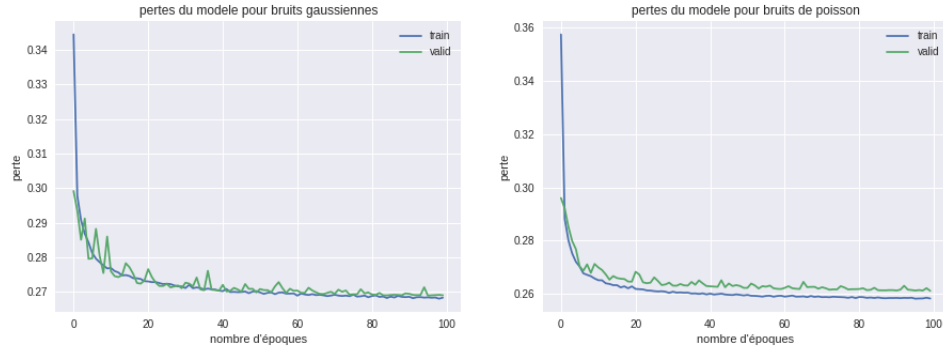


FIGURE 9 – Perte de training et de validation à partir de 100 époques en utilisant une taille de batch de 10.

Modèle a bien convergé pour les niveaux de bruit donnés et la taille de l'échantillon. La comparaison des performances de l'auto-encodeur débruiteur convolutif par rapport des autres méthodes est montrée via les scores de SSIM et PSNR :

Méthode	PSNR	SSIM
Image bruitée	20.17	0.42
Filtre médian	25.48	0.59
Débruitage par patchs (Non-local means)	20.18	0.42
CNN DAE	29.41	0.87

TABLE 2 – Scores moyen de SSIM pour les images de test pour le bruit gaussien

Méthode	PSNR	SSIM
Image bruitée	23.36	0.46
Filtre médian	26.46	0.57
Débruitage par patchs (Non-local means)	8.81	0.35
CNN DAE	31.75	0.91

TABLE 3 – Scores moyen de SSIM pour les images de test pour le bruit poisson

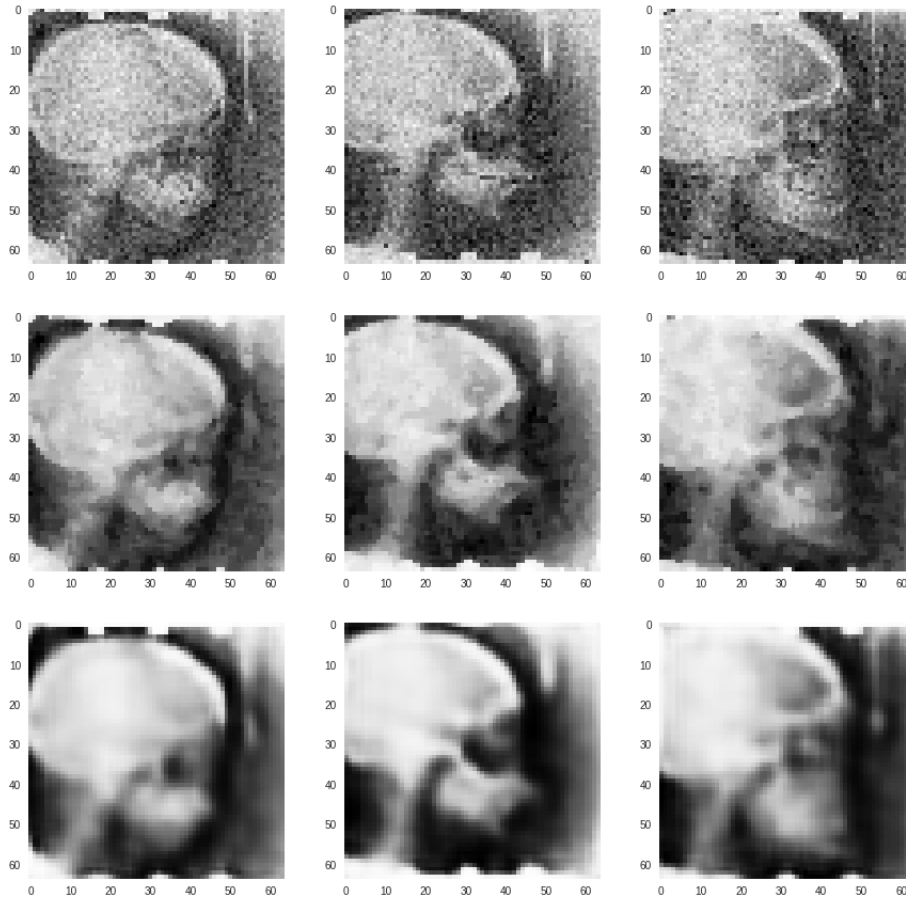


FIGURE 10 – Les performances de différents modèles sur bruit gaussien. La rangée supérieure montre les images bruitées, la deuxième rangée montre les résultats du filtre médian et la troisième rangée montre des images débruitées utilisant le CNN DAE

Les résultats montrent une performance exceptionnelle de l’auto-encodeur débruiteur convolutionnel sur les autres méthodes de débruitage d’images.

Selon une proposition dans l’article, nous avons testé une variante du modèle pour regarder si il est capable de booster la performance. Il s’agit de prétraiter les images d’entraînement par le filtrage médian avant l’utilisation de CNN DAE. Malheureusement, les résultats est pire que modèle original :

Méthode	PSNR gauss	SSIM gauss	PSNR poisson	SSIM poisson
CNN.DAE avec médian	26.05	0.84	31.19	0.89

TABLE 4 – Scores CNN DAE combiné avec filtre médian pour 2 types de bruits

6 Conclusion

Nous avons appris à construire et mis en place modèle de l’auto-encodeur débruiteur convolutionnel. Il est prouvé d’être efficace sur les données de images médicales. Pour l’étape suivante, pour améliorer le modèle, nous pouvons essayer de rendre le réseau plus profond ou utiliser d’auto-encodeur convolutionnel avec des connexions de saut symétrique (Symmetric Skip Connections).

Références

- [1] Lovedeep Gondara, ”Medical image denoising using convolutional denoising autoencoders” Department of Computer Science Simon Fraser University.
- [2] Building Autoencoders in Keras
- [3] Structural similarity index - SSIM