

RAPPORT DE PROJET H-419 SEGMENTATION DU CANCER DU SEIN AVEC ATTENTION

MEFENYA TAKOUMBO ROMEO JOEL

Table des matières

1	Résumé	2
2	Introduction	2
3	Jeu De Données Utilisé	2
4	Réseau	2
4.1	La convolution	3
4.2	Le max pooling	3
4.3	Couches entièrement connectées	3
4.4	Le modèle U-Net	4
5	Résultats et interprétation	6
5.1	Interprétation des courbes	6
5.2	Prédiction de quelques données	6
6	Conclusion et Perspectives	7
7	Références	8
8	Lien code, base de données originale	8

1 Résumé

Le projet de segmentation du cancer du sein se concentre sur l'utilisation des réseaux de neurones convolutionnels (CNN) pour détecter et segmenter les tumeurs dans des images de mammographies. En appliquant une méthode avancée de segmentation, le modèle utilisé est basé sur l'architecture U-Net qui est particulièrement efficace pour les tâches de segmentation d'images médicales. Ce modèle permet de prédire les zones suspectes d'une mammographie, contribuant ainsi à une détection précoce du cancer du sein. La performance du modèle est évaluée grâce à des métriques telles que la précision (accuracy), la perte (loss) et l'IoU (Intersection over Union) afin d'assurer la fiabilité du modèle dans un contexte clinique.

2 Introduction

Le cancer du sein reste l'une des principales causes de mortalité chez les femmes à l'échelle mondiale. La détection précoce de cette affection est cruciale pour améliorer les chances de survie et minimiser les traitements lourds. Dans ce contexte, les techniques de machine learning, surtout la segmentation d'images médicales, peuvent jouer un rôle clé. Ce projet se base sur des images mammographiques et exploite un réseau de neurones convolutifs (CNN) pour délimiter les zones suspectes et repérer les anomalies, ce qui pourrait permettre aux professionnels de la santé de détecter plus rapidement les cas à risque de cancer du sein. Notre but est donc de concevoir un modèle fiable et précis pour la segmentation automatique des tumeurs mammaires à partir des images radiologiques. La segmenta-

tion d'image constitue une étape essentielle en traitement d'image, visant à diviser une image en régions significatives pour en extraire des informations pertinentes. En médecine, la segmentation permet de définir des structures anatomiques ou des anomalies (comme les tumeurs), facilitant ainsi le diagnostic, la planification du traitement ou le suivi de l'évolution d'une maladie.

3 Jeu De Données Utilisé

Notre projet ainsi que nos données ont été pris sur Kaggle qui est une plateforme en ligne populaire pour les scientifiques des données et les praticiens de l'apprentissage automatique. Les données utilisées proviennent du jeu BUSI (Breast Ultrasound Images), un ensemble d'images échographiques du sein. Ce dataset est destiné à la segmentation des tumeurs mammaires. Il comprend :

- **Trois catégories d'images :**
 - **Benign** : images contenant des tumeurs bénignes.
 - **Malignant** : images contenant des tumeurs malignes.
 - **Normal** : images normales sans aucune anomalie détectée.

4 Réseau

Afin de mener à bien cette tâche, nous avons recours à un réseau de neurones convolutif (CNN), une architecture de deep learning particulièrement adaptée à l'analyse d'images. Différemment des réseaux entièrement connectés, les CNN tirent parti de la structure spatiale des images pour dénicher, de manière autonome, des caractéris-

tiques pertinentes.

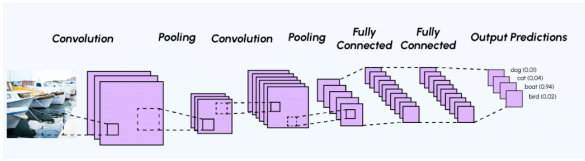


FIGURE 1 – Exemple d'architecture d'un CNN. [1]

La configuration traditionnelle d'un CNN comporte une série d'éléments, répartis en trois groupes principaux. Le premier groupe est constitué de couches de convolution, le second groupe intègre les couches de pooling, et le troisième groupe se focalise sur les couches d'activation.

4.1 La convolution

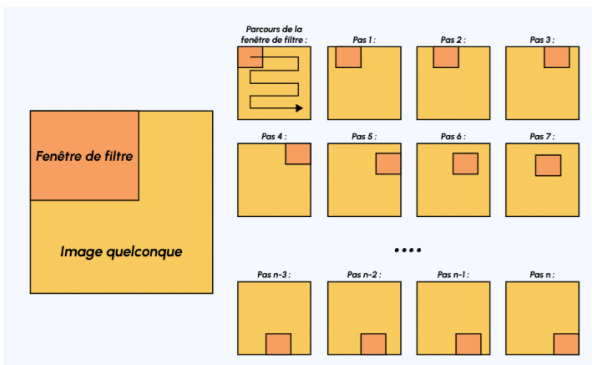


FIGURE 2 – Schéma du parcours de la fenêtre de filtre sur l'image [2]

La convolution est une opération mathématique simple généralement utilisée pour le traitement et la reconnaissance d'images. Sur une image, son effet s'assimile à un filtrage dont voici le fonctionnement :

- Dans un premier temps, on définit la taille de la fenêtre de filtre située en haut à gauche.
- La fenêtre de filtre, représentant la feature, se déplace progressivement de la gauche vers la droite d'un certain nombre de cases défini au préalable

(le pas) jusqu'à arriver au bout de l'image.

- À chaque portion d'image rencontrée, un calcul de convolution s'effectue permettant d'obtenir en sortie une carte d'activation ou feature map qui indique où est localisées les features dans l'image : plus la feature map est élevée, plus la portion de l'image balayée ressemble à la feature

4.2 Le max pooling

Le Max-Pooling est un processus de discrétisation basé sur des échantillons. Son objectif est de sous-échantillonner une représentation d'entrée (image, matrice de sortie de couche cachée, etc.) en réduisant sa dimension. De plus, son intérêt est qu'il réduit le coût de calcul en réduisant le nombre de paramètres à apprendre et fournit une invariance par petites translations (si une petite translation ne modifie pas le maximum de la région balayée, le maximum de chaque région restera le même et donc la nouvelle matrice créée restera identique)

4.3 Couches entièrement connectées

Elles permettent d'effectuer la prédiction finale à partir des caractéristiques extraites. Bien que les CNN soient efficaces pour la classification d'images (par exemple déterminer si une image contient ou pas une tumeur) ils ne sont pas conçus pour prédire la localisation exacte des objets dans une image. Pour résoudre ce problème, des architectures spécifiques ont été mises en place pour réaliser la segmentation, c'est-à-dire attribuer une étiquette à chaque pixel

de l'image. Pour résoudre cette problématique, des architectures ont été développées et parmi celles-ci le modèle U-Net s'est avérée être la meilleure.

4.4 Le modèle U-Net

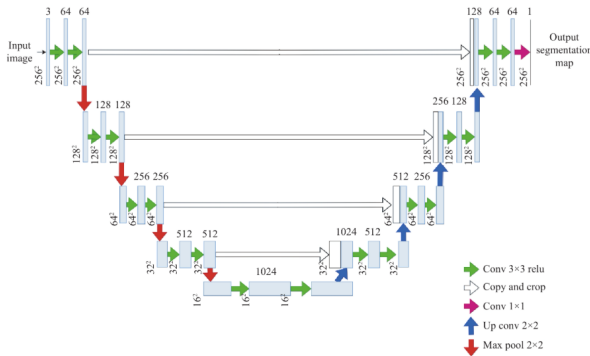


FIGURE 3 – Architecture du réseau [5]

Le U-Net est une architecture de réseau de neurones qui a été développé en 2015 par Olaf Ronneberger, Phillip Fischer et Thomas Brox en Allemagne. Il a été spécifiquement conçu pour la segmentation d'images biomédicales où les jeux de données sont généralement limités. Le nom U-Net vient de son architecture en forme de U avec un chemin descendant (encodeur) et un chemin ascendant (décodeur).

- **L'encodeur**, que l'on peut comparer une sorte d'analyse progressive de l'image, a pour objectif de repérer ce qui est important. À chaque niveau, le réseau applique des filtres pour extraire les caractéristiques essentielles, puis il réduit la taille de l'image grâce à une opération de regroupement appelée pooling. Ce processus est semblable à un zoom arrière qui permet au modèle de mieux comprendre le contexte général de l'image tout en gardant une trace des détails importants.

- Une fois que l'image a été fortement compressée et analysée, le **décodeur** entre en jeu. Cette partie du modèle fait le chemin inverse. Elle agrandit progressivement l'image tout en reconstruisant les détails. Le décodeur ne travaille pas seul, à chaque étape, il récupère des informations directement depuis l'encodeur ce qui lui permet de reconstruire l'image avec plus de précision. Ces connexions entre l'encodeur et le décodeur sont appelées connexions de saut et agissent comme des liens ou passerelles d'informations entre les deux parties du réseau.

- Dans notre projet, nous avons également utilisé un mécanisme appelé **attention Gate** qui vient affiner encore davantage ce processus. L'idée ici est simple mais très puissante. Plutôt que de transmettre toutes les informations de l'encodeur vers le décodeur, on aide le modèle à se concentrer uniquement sur ce qui est pertinent. C'est un peu comme si on ajoutait un filtre d'attention qui dit au modèle de se concentrer sur les zones qui sont probablement plus importantes que le reste. Grâce à ce mécanisme, le modèle peut ignorer le bruit inutile et se focaliser davantage sur les régions susceptibles de contenir des tumeurs. Cela améliore la précision de la segmentation en particulier lorsque les anomalies sont petites ou contrastées.

Finalement le modèle produit une carte binaire indiquant pour chaque pixel, s'il appartient à la tumeur ou pas. Grâce à cette approche, le U-Net combine une vision globale de l'image ce qui rend particulièrement

efficace pour des taches de segmentation comme la détection des tumeurs

III- Programmation a- Prétraitement des données

Comme mentionné précédemment, les données utilisées pour ce projet sont le BUSI (Breast Ultrasound Images), un ensemble de données publiques contenant des images échographiques du sein. Ce dataset a été téléchargé sur la plateforme Kaggle, où il est librement accessible. Une fois le fichier compressé téléchargé, il a été extrait puis exploité à l'aide de l'environnement Anaconda à travers Jupyter Notebook ce qui a permis une gestion des codes python pour le traitement d'image et l'entraînement du modèle.

La dataset comprend trois catégories d'images. La première regroupe les cas de tumeurs bénignes, la seconde les cas de tumeurs malignes et la dernière contient des images normales sans anomalies.

Désignation	Valeur
Nombre de données	1578
Nombre de classe	3 Malin Bénin Normal
Mask simples	780
Mask doubles	17
Type	RGB
Dimension moyenne	500*500
Redimensionne	256*256
Normalisation	0 à 1

TABLE 1 – Caractéristiques du dataset

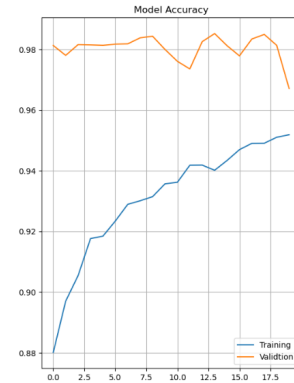


FIGURE 4 – Courbe d'évolution de la précision

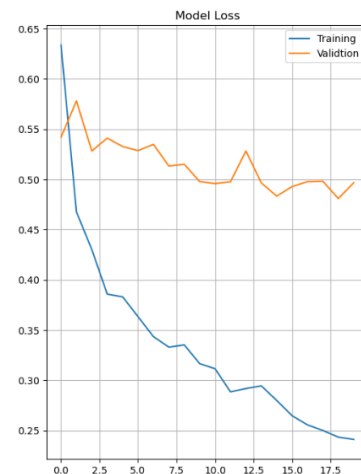


FIGURE 5 – Courbe d'évolution de la Perte



FIGURE 6 – Courbe de l'Intersection sur Union (IoU)

5 Résultats et interprétation

5.1 Interprétation des courbes

- La précision s’améliore progressivement sur l’ensemble d’entraînement, passant de 0.88 à 0.95, tandis que la précision de validation reste très élevée, atteignant jusqu’à 0.985. Cela montre que le modèle peut apprendre de manière efficace à différencier les images et à faire des prédictions fiables.

La justesse progresse de manière continue sur l’ensemble d’entraînement, passant de 0.88 à 0.95, tandis que l’exactitude de validation se maintient à un niveau élevé, atteignant même jusqu’à 0.985. Cela indique que le modèle réussit à saisir les traits distinctifs des images et à réaliser des prévisions précises.

- La fonction de coût diminue régulièrement sur les données d’entraînement, ce qui indique que le modèle se rapproche progressivement de l’optimum. En ce qui concerne la validation, la perte demeure stable et relativement faible, ce qui traduit un bon compromis entre apprentissage et performance de prédiction.
- L’IoU affiche une constance Impressionnante tout au long de l’apprentissage. Le score d’IoU avoisine les 0,455 sur l’ensemble d’entraînement et les 0,491 sur l’ensemble de validation, ce qui démontre une excellente aptitude du modèle à généraliser en termes de précision des contours segmentés.

5.2 Prédiction de quelques données

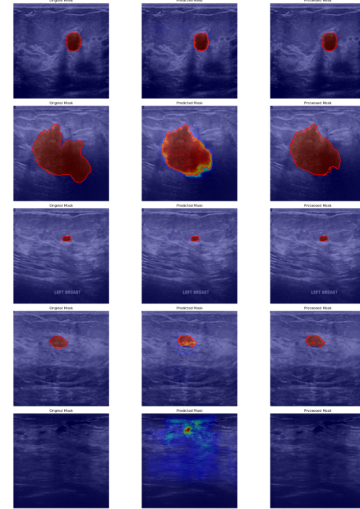


FIGURE 7 – Prédiction de quelques résultats

Après avoir établi le modèle de segmentation, nous avons effectué une évaluation visuelle de ses performances en comparant trois types d’images pour chaque échantillon : le masque original (référence), le masque prédit par le modèle, et le masque après le traitement post-segmentation.

Dans l’ensemble, les résultats montrent que le modèle réussit à identifier correctement les régions tumorales. La forme générale des masses est bien détectée, et les prédictions correspondent, dans la plupart des cas, à la structure réelle des lésions visibles sur les masques originaux. Cependant, on note quelques légers débordements autour des contours, reflétant une certaine incertitude du modèle sur les bords des tumeurs, un phénomène courant dans la segmentation d’images biomédicales.

Le traitement appliqué après la segmentation des masques prédits (Masque Traité) a permis d’améliorer considérablement ces défauts. Ce processus a principalement aidé à :

- Renforcer la netteté des contours en les rendant plus clairs,
- Éliminer les petites zones indésirables et les bruits résiduels,
- Affiner les formes pour mieux correspondre aux contours véritables des masses tumorales.

Dans des situations particulièrement complexes (zones à faible contraste ou anomalies très petites), le modèle a montré une bonne efficacité en réduisant les faux positifs après traitement.

Bien que le modèle de segmentation produise des résultats généralement satisfaisants, plusieurs limites significatives ont été identifiées lors de l'examen visuel des masques prédits.

Tout d'abord, on remarque que, sur certaines images, le modèle

- Ne segmente pas suffisamment, il échoue à identifier toutes les zones tumorales. Par exemple, certaines parties évidentes de certaines masses visibles dans les masques originaux ne sont pas prises en compte par la prédiction.
- Sursegmentation : à l'inverse, le modèle étend parfois ses prédictions au-delà des contours réels de la lésion, incluant des régions saines dans le masque segmenté.
- Produit des prédictions bruitées : dans certains cas, de petites zones parasites apparaissent dans l'image prédite, témoignant une confusion entre le bruit de fond et le véritable tissu tumoral.

6 Conclusion et Perspectives

Ce projet portait sur la segmentation des images mammaires pour la détection des tumeurs en utilisant un réseau de neurones convolutif, notamment le modèle U-Net, reconnu pour son efficacité dans la segmentation biomédicale. Ce choix s'est révélé adapté vu la nature du problème.

Après avoir entraîné et évalué le modèle sur notre base de données, nous avons obtenu des résultats assez satisfaisants, avec une précision atteignant 0,93. Cependant, le modèle a eu du mal à bien distinguer toutes les formes de tumeurs, ce qui est crucial dans le milieu médical, où l'exactitude est essentielle.

Notre approche U-Net, relativement simple, a montré de bonnes performances malgré la taille limitée de notre base de données. Essayer de rendre le modèle plus complexe n'a pas amélioré les résultats, ce qui met en lumière un aspect essentiel, la performance d'un modèle dépend beaucoup de la quantité et de la qualité des données disponibles. Un modèle plus complexe nécessite inévitablement un volume de données plus important pour donner des résultats optimaux.

Pour le futur, il serait judicieux d'élargir le volume de notre base de données pour entraîner un modèle plus robuste.

En bref, ce projet représente un début prometteur vers un outil pour aider à la détection automatique des tumeurs mammaires, tout en soulignant les défis liés à la généralisation et la compréhension dans le contexte médical.

7 Références

[1] Convolutional Neural Network : Tout ce qu'il y a à savoir

[2]<https://images.squarespacecdn.com/content/v1/519a7bc0e4b08ccdf8f31445/1524659432980->

[3] DataScientest. "Convolutional Neural Network." : <https://datascientest.com/convolutional-neuralnetwork>.

[4] Segmentation d'objets : les architectures référentes en imagerie médicale <https://www.imaios.com/fr/ressources/blog/introduction-aux-architectures-d-apprentissage-automatique-profond-les-plus-courantes-p>

[5] UNet et variantes pour la segmentation d'images médicales : <https://www.sciltp.com/journals/ijndi/article/view/411>

8 Lien code, base de données originale

Code : cliquez ici pour voir la base de données

BDOriginale : cliquez ici pour voir la base de données