Résumé

Le cancer du sein est le cancer le plus répandu chez les femmes dans le monde. La technologie de dépistage la plus courante est la mammographie. Notre projet vise à proposer une approche d'apprentissage en profondeur pour la détection du cancer du sein afin de réduire le coût et la charge de travail des radiologues. Avec les progrès récents dans le domaine de l'apprentissage profond, l'utilisation des réseaux de neurones convolutifs (CNN) en imagerie médicale est devenue très encourageante. Travaillant sur le jeu de données CBIS-DDSM dans nos expériences, nous avons utilisé cinq modèles appronfondus pour une précision acceptable. L'un des modèles est le VGG-16 avec Deep Transfert Learning pour classer les images en masse/calcification, et un autre modèle d'un réseau scratch CNN à deux sorties pour une classification bénigne/maligne, ces deux derniers réseaux ont été composés pour élaborrer un nouveau réseau à 4 classses de sortie (masse bénigne, masse maligne, calcification bénigne, calcification maligne). Afin d'avoir une classification conjointe avec une segmentation, nous avons utilisé le réseau Deep celui de Mask R-CNN pour une classification avec localisation des masses mamairs. Une image mammographique a été fournie au masque R-CNN, et une boîte de délimitation et une image d'étiquette de la région du cancer du sein ont été obtenues. Suite aux résultats obtenus nous pouvons conclure qu'ils sont encourageants, mais le domaine de la détéction du cancer du sein reste un domaine qui necessite encore beaucoup d'effort de la part de la communauté des chercheurs du domaine.

Les Mots clés *: Cancer du sein, Apprentissage profond, mammographie, apprentissage par transfert, CNN, Mask-R-CNN.*

Abstract

Breast cancer is the most common cancer for women in the world. The most common screening technology is mammography. Our project aims to propose a deep learning approach for breast cancer detection to reduce the cost and the time for radiologists. With recent advances in the field of deep learning, the use of convolutional neural networks (CNNs) in medical imaging has become very encouraging. Working on the CBIS-DDSM dataset in our experiments, we used five deep models for acceptable accuracy. One of the models is the VGG-16 with Deep Transfer Learning to classify images in mass/calcification, and another model CNN from scratch with two-output for benign/malignant classification, these last two networks were composed to develop a new network with 4 output classes (benign mass, malignant mass, benign calcification, malignant calcification). In order to have a joint classification with a segmentation, we used the Deep network that of Mask R-CNN for a classification with localization of breast masses. A mammographic image was provided to the R-CNN mask, and a bounding box and a label image of the breast cancer region were obtained. From the results obtained we can conclude that they are encouraging, but the field of detection of breast cancer remains an area which still requires a lot of effort on the part of the research community in the field.

Keywords: *Breast cancer, Deep Learning , mammography, , Transfert Learning, CNN, Mask-R-CNN.*

مـلـخص

سرطان الثدي هو أكثر أنواع السرطانات شيوعًا بين النساء في جميع أنحاء العالم. أكثر تقنيات الفحص شيوعًا هي التصوير الشعاعي للثدي. يهدف مشروعنا إلى اقتراح نهج تعلم عميق للكشف عن سرطان الثدي لتقليل التكلفة وعبء العمل على اختصاصي الأشعة. مع التطورات الحديثة في مجال التعلم العميق ، أصبح استخدام CNN في التصوير الطبي أمرًا مشجعًا للغاية. من خلال العمل على مجموعة بيانات CBIS-DDSM في تجاربنا ، استخدمنا خمسة نماذج عميقة للحصول على دقة مقبولة. أحد النماذج هو VGG-16 مع Deep Transfer Learning لتصنيف الصور إلى كتلة/تكلس ، ونموذج آخر لشبكة CNN من الصفر ثنائية الإخراج لتصنيف حميد/خبيث ، تم تركيب هاتين الشبكتين الأخيرتين لتطوير شبكة جديدة مع 4 فئات من المخرجات (الكتلة الحميدة ، الكتلة الخبيثة ، التكلس الحميد ، التكلس الخبيث). من أجل الحصول على تصنيف مشترك مع التجزئة ، استخدمنا الشبكة العميقة الخاصة بـ Mask R-CNN لتصنيف مع توطين كتل الثدي. تم توفير صورة ماموجرافيك ل Mask R-CNN ، وتم الحصول على صندوق محيط و ملصق صورة لمنطقة سرطان الثدي. من النتائج التي تم الحصول عليها يمكننا أن نستنتج أنها مشجعة ، لكن مجال الكشف عن سرطان الثدي لا يزال مجالًا يتطلب الكثير من الجهد من جانب مجتمع البحث في هذا المجال.

<u>الكلمات المفتاحية</u>: *سرطان الثدي، التعلم العميق ،التصوير الشعاعي للثدي ، نقل التعلم ، CNN، Mask-R-CNN*

Remerciments

*Tout d'abord, c'est « **Allah** » qui nous a guidé sur le droit chemin tout au long du travail et nous a inspiré les bons pas et les justes reflexes. Sans sa miséricorde, ce travail n'aurait pas abouti.*

*L'encadrement scientifique de ce travail a été assuré par **Dr. Amroune Mohamed**, maitre de conférences classe A à la faculté des Sciences Exactes et des Sciences de la Nature et de la Vie, Université Tébessa. Nous tenons vivement à lui exprimer nos profonde reconnaissances gratitude pour sa disponibilité, sa patience, sa compréhension, ses qualités humaines et ses intérêts portés pour notre sujet de travail. Nous le remercions de nous avoir fait confiance et d'avoir été présent aussi souvent que possible malgré ses tâches pédagogiques. Son soutien permanent et son dynamisme nous ont permis d'avancer plus loin dans notre travail.*

Nos remerciements vont aussi à

*Dr. **Haouam. MY,** d'avoir ménagé son temps pour présider ce jury*

*Dr. **Gharbi. A,** pour avoir bien voulu siéger dans ce jury afin d'examiner et critiquer ce mémoire et nous éclairer par ces précieux conseils.*

*Dr. **Amroune. M***

Aucun remerciement ne saurait exprimer notre respect et considération pour les orientations que vous avez consenties pour notre étude de l'Université.

*Mr. **Gahmouss. A**, Mr. **Menassel. R**, Mr. **Zaguari. A**, Mr. **Laimech. A**, Mr. **Gattal. A**, Mr. **Bendib. I**, Mr. **Mahmoudi.R** tous les enseignants qui nous ont fait former durant ces 5 années merci pour votre encouragement et gentillesse.*

Tous nos amis pour leur solidarité

A nos parents

Pour l'enfance merveilleuse qu'ils nous ont offerte ainsi que pour leurs encouragements.

Pour leurs soutiens et leurs aides.

Avec tout notre amour.

A tous mes amis

Pour leurs bonnes humeurs, leurs gentillesses et pour tous nos fous rires partagés.

Pour tout ce qu'ils nous ont appris

Merci à tous ceux qui, d'une manière ou d'une autre qui ont contribué à la réalisation de ce travail, et que nous ne pouvons citer individuellement.

DEDICACE

A mes chers parents,

Que nulle dédicace ne puisse exprimer ce que je leur dois, pour leur bienveillance, leur affection et leur soutien…Trésors de bonté, de générosité et de tendresse, en témoignage de mon profond amour et ma grande reconnaissance « Que Dieu vous garde ».

À mes chers frères,

Mouhammed & Ibrahim, je leur dédie ce modeste travail en témoignage de mon grand amour et ma gratitude infinie.

À mes professeurs

Qui doivent voir dans ce travail la fierté d'un savoir bien acquis.

À tous mes amies

(Maya, Amel, Zaho, Chîmo,…)

Pour leur aide et leur soutien moral durant l'élaboration du travail de fin d'études.

À tous ma famille

À tous ceux dont l'oubli du nom n'est guère celui du cœur…



Table des matières

Introduction générale

1. Motivation 1
2. Problématique 1
3. Contribution 2
4. L'organisation du manuscrit 3

Chapitre 1 Contexte Médical

1. Introduction 4
2. L'anatomie du sein chez la femme 4
3. Cancer de sein 5
   3.1. Tumeurs bénignes 6
   3.2. Tumeurs malignes 6
      3.2.1. Les symptômes 6
      3.2.2. Les autres symptômes 8
4. Dépistage du cancer du sein 8
   4.1. L'imagerie par résonance magnétique (IRM) 8
   4.2. L'échographie 9
   4.3. La mammographie 10
      4.3.1. Acquisition des images mammographiques 10
      4.3.2. Corrélation entre la structure du sein et les images mammographiques 11
      4.3.3. Signes d'appel radiologique d'un cancer du sein 12
      4.3.4. Densité mammaire 17
      4.3.5. Classification des pathologies mammaires 18
   4.4. L'examen de confirmation diagnostic: l'examen histologique 20
   4.5. Nouveauté en imagerie du sein 21
   4.6. Système de détection assistée par ordinateur (DAOx) 22
5. Les bases de mammographie 22
   5.1. Mammographie Image Analysis Society Digital Mammogram Database (MIAS) 23
   5.2. Digital Database for Screening Mammography (DDSM) 24
   5.3. Banco Web LAPIMO Plus 24
   5.4. INbreast 24
6. Conclusion 25

Chapitre 2 Deep Learning

1. Introduction 26
2. L'apprentissage automatique « Machine Learning » 26

2.1. L’apprentissage supervisé 27
2.2. L’apprentissage non supervisé 27
2.3. Apprentissage semi-supervisé 28
2.4. Apprentissage par renforcement 28
3. L’apprentissage profond « Deep Learning » 28
3.1. Principes 28
3.2. Réseau de Neurones et l'architecture « Deep Learning » 29
3.3. Deep Learning vs Machine Learning? 30
3.4. Domaines d'application de l'apprentissage profond 31
3.4.1. Entreprises adoptant le « Deep Learning» 31
3.5. Les techniques de « Deep Learning » 32
3.5.1. Les réseaux des neurones convolutifs (CNN) les plus connus 32
3.5.1.1 AlexNet 34
3.5.1.2 GoogLeNet 34
3.5.1.3 VGGNet 35
3.5.1.4 ResNet 35
3.5.1.5 Inception-V3, V4 et Inception-ResNet 35
3.5.2. Les réseaux de neurones récurrent « RNN » 36
3.6. Les techniques avancées de « Deep Learning » 37
3.6.1. R-CNN 38
3.6.2. Fast R-CNN 38
3.6.3. Faster R-CNN 39
3.6.4. YOLO 40
3.6.5. Mask R-CNN 40
3.7. Deep transfert learning 42
3.7.1. Utilisation directe de modèles pré-entraînés 42
3.7.2. Extraction de caractéristiques à partir de modèles pré-entraînés 42
3.7.3. Fine-tuning les dernières couches des modèles pré-entraînés 43
4. Le « Deep Learning » pour le diagnostic du cancer du sein 43
4.1. CNN pour l'estimation de la densité mammaire 43
4.2. CNN pour la détection de la masse mammaire 44
4.3. « CNN features » pour la classification de masse 46
5. Conclusion 47

Chapitre 3 Une approche basée Deep Learning pour la détection du cancer du sein

1. Introduction 48

2. Ressources utilisées ... 48
2.1. Dataset ... 48
2.2. Prétraitement des données (Data pre-processing) ... 48
3. Présentation des outils utilisés ... 49
3.1. Le software ... 49
3.2. Le hardware ... 50
4. Solution proposée ... 50
5. Architecture des modèles proposés et résultats ... 51
5.1. Modèle 1 : Expérience 1 (du réseau VGG 2-classes Masse/Calcification) ... 51
5.1.1. Modèle 1.1 « VGG avec feature extraction » ... 52
5.1.2. Les résultats du modèle 1.1 ... 52
5.1.3. Modèle 1.2 « VGG avec fine tuning » ... 55
5.1.4. Les résultats du modèle 1.2 ... 56
5.2. Modèle 2 « scratch CNN 2-classes Benign/Malignant » ... 57
5.2.1. Les résultats du modèle 2 More ... 58
5.3. Modèle 3 « Scratch CNN 4-classes » ... 58
5.3.1. Modèle 3.1 ... 59
5.3.2. Les résultats du modèle 3.1 ... 59
5.3.3. Modèle 3.2 « more filters » ... 60
5.3.4. Les résultats du modèle 3.2 ... 61
5.4. Modèle 4 « Composite de Modèle 1 et Modèle 2 » ... 62
5.4.1. Les résultats du modèle 4 ... 62
5.5. Modèle 5 « Mask R-CNN » ... 63
5.5.1. Backbone ... 63
5.5.2. RPN ... 64
5.5.3. RoI Align et Classification ... 65
5.5.4. Mask ... 66
5.5.5. Évaluation du modèle mask R-CNN ... 66
6. Discussion des résultats ... 68
7. Conclusion ... 69
Conclusion générale
1. Conclusion ... 70
2. Perspectives futures ... 70
Références bibliographiques ... 72

Liste de Figures

Figure 1.1	L'anatomie du sein chez la femme.	**05**
Figure 1.2	Exemple d'un cancer de sein.	**05**
Figure 1.3	Symtômes du cancer de sein.	**07**
Figure 1.4	Exemple d'imagerie IRM mammaire.	**08**
Figure 1.5	Exemple d'IRM mammaire BI-SEIN.	**09**
Figure 1.6	Exemple d'echographie mammaire.	**09**
Figure 1.7	Exemples d'images echographiques du sein (A) kyste du sein, (B) lesion non visualise sur la mammographie, (C) presence d'une masse de 5 cm, contenant des calcifications.	**09**
Figure 1.8	Les composants d'un mammographe.	**10**
Figure 1.9	Deux vues de mammographie : (A) la vue mlo et (B) la vue cc.	**11**
Figure 1.10	Coupe du sein en profil.	**12**
Figure 1.11	Aspect radiologique du sein normal sur une incidence face/oblique.	**12**
Figure 1.12	Exemple de : (A) MCS canalaires et (B) MCS lobulaires.	**13**
Figure 1.13	La macrobiopsie par mammotome : A) la mammotome B) les pièces prélevées C) l'image radiographique.	**21**
Figure 1.14	(A) mammographie sein droit : opacité de 3 cm associée à des microcalcifications, (B) bilan tomosynthèse correspondant et le (C) bilan par angiomammographie.	**22**
Figure 2.1	Processus de machine learning.	**27**
Figure 2.2	Le contexte de deep learning.	**29**
Figure 2.3	Architecture de réseau de neurones.	**29**
Figure 2.4	ML vers DL.	**30**
Figure 2.5	L'architecture des réseaux de neurones convolutifs.	**32**
Figure 2.6	Réseau de neurones du deep learning.	**33**
Figure 2.7	Histoire évolutive des CNN profonds montrant les innovations architecturales de ConvNet aux architectures à ce jour.	**36**
Figure 2.8	L'architecture des réseaux de neurones récurrents.	**37**
Figure 2.9	L'architecture R-CNN.	**38**

Figure 2.10	L'architecture Fast R-CNN.	**39**
Figure 2.11	L'architecture Faster R-CNN.	**39**
Figure 2.12	YOLO.	**40**
Figure 2.13	L'architecture Mask R-CNN.	**41**
Figure 3.1	Exemple de (A) : Calcification (B) : Masse.	**48**
Figure 3.2	Architecture VGG-16.	**51**
Figure 3.3	Architecture VGG16 Feature Extraction.	**52**
Figure 3.4	VGG16 Feature Extraction (Vanilla).	**53**
Figure 3.5	VGG16 Feature Extraction avec dropout.	**53**
Figure 3.6	VGG16 Feature Extraction avec dropout et plus petite FC layer.	**54**
Figure 3.7	VGG16 FE avec dropout et stronger data augmentation.	**55**
Figure 3.8	Architecture VGG16 2-Class Fine Tuning (2 Layer).	**55**
Figure 3.9	VGG16 2-Class Fine Tuning (1 Layer).	**56**
Figure 3.10	VGG16 2-Class Fine Tuning (2 Layer).	**57**
Figure 3.11	Architecture CNN B/M (2-Classes).	**57**
Figure 3.12	CNN B/M (2-Classes).	**58**
Figure 3.13	4-Classes CNN 1 Confusion Matrix.	**59**
Figure 3.14	Scratch CNN 1 (4-Classes).	**60**
Figure 3.15	4-Classes CNN 2 Confusion Matrix.	**61**
Figure 3.16	Scratch CNN 2 (4-Classes).	**61**
Figure 3.17	Architecture du classifieur composite 4-Classes.	**62**
Figure 3.18	Composite 4-Classes classifieur Confusion Matrix.	**63**
Figure 3.19	Architecture Mask R-CNN proposée.	**63**
Figure 3.20	Régions D'intérets (ROI) Génerée Par Le (RPN) Avec Plusieurs « Anchor Boxes ».	**64**
Figure 3.21	Résultats de la détection et de la classification.	**65**
Figure 3.22	Résultat de segmentation de masse.	**66**
Figure 3.23	GT et prédictions finales des masses avec Mask R-CNN, GT=GREEN, PRED= RED, CAPTIONS: SCORE/IOU.	**67**
Figure 3.24	Training loss.	**67**

Liste de Tableaux

Tableau 1.1	Tableau récapitulatif de différents formes et contour des masses.	**14**
Tableau 1.2	Tableau de différentes distributions des MCS.	**15**
Tableau 1.3	Tableau récapitulatif de différents types de microcalcification.	**17**
Tableau 1.4	La classification BI-RADS de la densité mammaire.	**18**
Tableau 1.5	Classification des microcalcifications selon LE GAL.	**18**
Tableau 1.6	Classification BIRADS de l'american college of radiology (ACR).	**20**
Tableau 2.1	Différences entre l'apprentissage supervisé et non supervisé.	**27**
Tableau 2.2	ML vers DL.	**31**
Tableau 2.3	Résumé des methodes convolutives basées sur les reseaux de neurones pour l'estimation de la densité mammaire.	**44**
Tableau 2.4	Résumé des methodes convolutives basées sur les reseaux de neurones pour la détection de la masse mammaire.	**46**
Tableau 2.5	Résumé des methodes convolutives basées sur les reseaux de neurones pour la classification de la masse mammaire.	**47**
Tableau 3.1	Les performances de notre machine.	**50**
Tableau 3.2	Résultats VGG16 feature extraction (vanilla).	**52**
Tableau 3.3	Résultats VGG16 feature extraction avec dropout.	**53**
Tableau 3.4	Résultats VGG16 FE avec dropout et plus petite FC layer.	**54**
Tableau 3.5	Résultats VGG16 FE avec dropout et stronger data augmentation.	**54**
Tableau 3.6	Résultats VGG16 2-class Fine Tuning (1 layer).	**56**
Tableau 3.7	Résultats VGG16 2-class Fine Tuning (2 layer).	**56**
Tableau 3.8	Résultats CNN B-M (2-classes).	**58**
Tableau 3.9	Résultats scratch CNN 1 (4-classes).	**59**
Tableau 3.10	Résultats scratch CNN 2 (4-classes).	**61**
Tableau 3.11	Résultats Mask R-CNN.	**68**