

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Saad Dahlab Blida



Faculté des sciences

Département informatique

Mémoire de fin d'études

En vue d'obtenir le diplôme de master

Filière : Informatique

Spécialité : Informatique

Option : Ingénierie des logiciels

THEME

Développement d'une plateforme Web-CAD pour l'analyse et le diagnostic d'images cliniques.

Présenté par :

M.CHAHRAT DJAMEL EDINNE

M. NADIR SEDDIK

Soutenu le : 06 juillet 2020, devant le jury composé de :

M^{me}. ABED H

Présidente

M^{me}. CHIKHI I

Examinatrice

M^r. AOUACHE M

Encadreur

M^{me}. GUESSOUM D

Promotrice

Année universitaire : 2019/2020

Remerciements

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
وَمَا تَوْفِيقِي إِلَّا بِاللَّهِ عَلَيْهِ تَوَكَّلْتُ وَإِلَيْهِ أُنِيبُ

سورة هود – الآية (88)

Nous remercions ALLAH de nous avoir donné la santé et le courage afin de pouvoir réussir ce travail en temps de corona virus.

Nous remercions très chaleureusement Mr. AOUACHE Mustapha pour son soutien, ses précieux conseils et tous le temps qui nous a consacré même pendant la nuit des fois jusqu'à le lever du soleil. Il nous a aidé du début jusqu'à la fin et il nous a fait rentrer dans ce vaste monde l'intelligence artificiel et de la recherche .Nous avons eu le privilège de travailler avec vous et apprécier vos qualités et vos valeurs, votre sérieux et votre compétence.

Nous exprimons notre grande gratitude à notre promotrice Mme. GUESSOUM Dalila, d'avoir accepté de nous encadrer pour notre travail et pour ses précieux conseils et ses orientations et pour nous avoir supporté durant cette dur période.

Nous désirons exprimer notre reconnaissance envers Mr FELOUAT Hisham qui nous a apporté son aide et a répondu à de nombreuses questions et à son support moral et intellectuel.

Nous désirons exprimer toute notre reconnaissance à toutes les personnes des forums de stackoverflow pour nous avoir orientés plusieurs fois.

Nous voudraient remercié aussi Mlle MOUSSERATI Fatma Zohra et Mlle OULD SETTI Hiba pour avoir dit du bon de nous a Mr AOUACHE avant même qu'on commence le travail avec lui.

Enfin nous tenons à remercier l'ensemble du corps enseignant et administrative qui travail du fond du Cœur pour données toutes les bonnes conditions de travail au étudiants.

ملخص

المستندة إلى الويب دورًا حاسمًا في المجال الطبي الذي يوفر أدوات وصول سريعة وسهلة للخبراء (Web-CAD) الكمبيوتر الطبيين لإجراء عملية معالجة وتحليل سهل الاستخدام للصورة.

كمنصة لتحليل الصور ومنصة تشخيص عن طريق التصوير الحراري WebCAD-IAD وبالتالي ، نحن ملتزمون بتطوير لتحليل وتشخيص سرطان الثدي.

لتحقيق ذلك بموضوعية ، قمنا بتقسيم العمل إلى مرحلتين مختلفتين: مراحل الكشف والتشخيص.

في مرحلة الكشف عن المنطقة المرغوبة ، تم ذلك عن طارق عملية تكيفية أوتوماتيكية باستخدام طريقتين لاستخراج منطقة الثدي.

النتائج إلى مرحلة التعرف التالية إلى تحليله ، مع / بدون مرض ، باستخدام نهجين مختلفين تحت (Brest-ROI) يهدف تمرير الإشراف و بدين إشراف.

أثناء العمل ، تم استخدام قاعدة بيانات صور متاحة للجمهور لتقييم أداء الخوارزميات ، بما في ذلك مصفوفات الاختلاف وقياس التشابه المطابق.

كانت دقة الخوارزميات القائمة على طرق خاضعة للإشراف وغير خاضعة للإشراف مثالية لتحقيق التصنيف و المطابقة الصحيحة لصورة المنطقة المرغوبة المقدمة (مع / بدون) المرض.

بناءً على النتائج التي تم الحصول عليها ، قد توصلنا إلى أن الخوارزميات لديها القدرة على القبول للاستخدام العملي ، ويمكن المطور كأداة فحص سريعة ودقيقة للصورة الحرارية السريرية بناءً على تشخيص سرطان Web CAD-IAD استخدام الثدي لمساعدة أخصائي الأشعة مع أنشطتهم الطبية.

Résumé

De nos jours, le développement de plates-formes de diagnostic assisté par ordinateur (Web-CAD) sur le Web joue un rôle crucial dans le domaine médical qui fournit des outils d'accès rapides et faciles aux experts médicaux pour effectuer un traitement et une analyse conviviaux des images. Ainsi, nous nous engageons à développer un CAD-IAD Web en tant que plateforme d'analyse d'images et de diagnostic appliquée à l'imagerie thermique pour l'analyse et le diagnostic du cancer du sein. Pour y parvenir objectivement, deux phases différentes sont impliquées: les phases de détection et de diagnostic. Dans la phase de détection, un processus de segmentation à travers un processus adaptatif utilisant deux méthodes, l'adaptative de transformée de Hough(Adp-HT) et l' adaptative de transformée de Hough Booster(Adp-BHT), a été exploré pour extraire la région mammaire d'intérêt (Brest- ROI). Le résultat qui passe à la prochaine phase de reconnaissance vise à l'analyser, avec / sans lésions, en utilisant deux approches différentes supervisées (classification) et non supervisées (appariement). Au cours des travaux, une base de données d'images accessible au public a été utilisée pour évaluer les performances des algorithmes, y compris les matrices de confusion et la mesure de similarité de correspondance. La précision des algorithmes basés sur des méthodes supervisées et non supervisées était parfaite pour obtenir une (classification / correspondance) correcte de l'image de Brest-ROI présentée (avec / sans) lésions. Sur la base des résultats obtenus, nous avons conclu que les algorithmes ont le potentiel d'être acceptés pour une utilisation pratique, et le Web CAD-IAD développé peut être utilisé comme un outil de dépistage rapide et précis pour diagnostic du cancer du sein avec l'image thermique clinique pour aider les radiologues avec leurs activités médicales.

Mots clés:

Cancer du sein, détection, techniques supervisées et non supervisées, réseau neuronal, apprentissage profond, mesure des performances; Web-CAD

ABSTRACT

Nowadays, the development of Web-based Computer-aided Diagnostic (Web-CAD) platforms plays a crucial role in the medical field that provides rapid and easy access tools for medical experts to perform a user-friendly image treatment and analysis process. Thus, we are committed to developing a Web CAD-IAD as image analysis and diagnostic platform applied to thermal imaging for breast cancer analysis and diagnosis. To achieve this objectively, two different phases are involved: the detection and diagnostic phases. In the detection phase, a segmentation process through an adaptive process using two methods, adaptive based Hough transform (Adp-HT) and adaptive based Boost Hough transform (Adp-BHT), was explored to extract the breast-region of interest (Brest-ROI). The resulting (Brest-ROI) pass to the next recognition phase aims to analyze it, with/without lesions, using two different supervised (classification) and unsupervised (matching) approaches. During the work, a publicly accessible image database was used to assess the performance of the algorithms, including confusion matrices and matching similarity measure. The accuracies of the algorithms based on supervised and unsupervised methods were perfect to achieve a correct (classification/matching) of Brest-ROI image presented (with/ without) lesions. Based on the results obtained, concluded that the algorithms have the potential to be accepted for practical use, and the CAD-IAD Web developed can be used as a rapid and precise screening tool for the clinical thermal image based on breast cancer diagnosis to help the radiologist with their medical activities.

Keywords:

Breast, cancer, detection, supervised and unsupervised techniques, deep learning neural network, similarity and matching, performance measure, Web-CAD

Table des matières

Remerciements.....	2
ملخص.....	3
Résumé.....	4
ABSTRACT.....	5
Listes des figures.....	8
Listes des tableaux.....	10
Liste des abréviations:.....	11
Chapitre 1.....	12
Contexte.....	12
Motivations.....	12
Problématiques.....	13
Questions de recherche.....	13
Objéctive de la recherche.....	14
Organisation du mémoire.....	15
Chapitre 2 TRAVAUX ET TECHNIQUES CONNEXES ANTÉRIEURS.....	16
2.1 Introduction.....	16
2.2 Structure du sein de la femme.....	16
2.3 Cancer du sein.....	17
2.4 Type de Cancer du sein [7].....	18
2.4.1 Les cancers in situ.....	18
2.4.2 les cancers infiltrant.....	19
2.5 Travaux Connexes.....	21
2.6 Comparaison entre les méthodes utilisées dans les travaux.....	24
2.7 Conclusion.....	27
Chapitre 3 DESIGN ET MÉTHODOLOGIE.....	29
3.1 Introduction.....	29
3.2 Schéma proposé.....	29
3.3 Base de données.....	30
3.4 Phase de detection:.....	34
3.4.1 Prétraitement :.....	34
3.4.2 Procédure de la segmentation.....	37
3.5 Phase de diagnostic.....	41
3.5.2 Procédure supervisée.....	43
3.5.3 Procédure non supervisée.....	49
3.6 Segmentation des couleurs:.....	51
3.7 Conception de la plateforme web.....	53
3.7.1 Diagramme de classe.....	53

5.3.2 Diagramme de cas d'utilisation général	55
5.3.3 Diagramme de cas d'utilisation patient.....	55
5.3.4 Diagramme de cas d'utilisation médecin	56
3.6 Conclusion	57
Chapitre 4 RÉSULTATS ET DISCUSSION.....	58
4.1 Introduction.....	58
4.2 Résultats de segmentation	58
4.2.1 Résultats utilisant le processus manuel	58
4.2.2 Résultats utilisant l'algorithme adaptatif 1	58
4.2.3 Résultats utilisant l'algorithme adaptatif 2	60
4.2.4 Comparaison entre les algorithmes :.....	61
4.3 Résultat de la méthode supervisée	62
4.3.1 Resultat du CNN	62
4.3.2 Results from VGG16	63
4.3.3 Resultat du MobilNetV2	64
4.3.4 Resultat du ResNet50	65
4.4.5 Comparaison entre les méthodes supervisées.....	66
4.5 Résultat de la méthode non supervisée	66
4.5.1 VGG16+K-ppv	66
4.5.2 MobileNetV2+K-ppv	69
4.5.3 Comparaison entre les méthodes non supervisées	72
4.6 Comparaison entre méthodes superviséé et non supervisée	73
4.7 Résultat visuel du CBIR	74
4.7.1 Distance Euclidienne.....	74
4.7.2 Distance Euclidienne avec SKLearn	75
4.7.3 Distance MSE	75
4.7.4 Distance SSIM	76
4.8 Conclusion	76
Chapitre 5 Implémentation de la plateforme WebCAD	77
1.1 Introduction.....	77
1.2 Schéma fonctionnel de la plate-forme WebCAD.....	77
5.3 Ressources matérielles et logicielles utilisées lors du developpement:	78
5.4 Système développé :	80
5.5 conclusion.....	96
Chapitre 6 CONCLUSION.....	97
Bibliographie.....	99

Listes des figures

Figure 12.1 Structure du sein féminin [6].....	17
Figure 2.2 cancer canalaire in situ et infiltrant[7]	18
Figure 2.3 Carcinome canalaire in situ (CCIS)[7]	19
Figure 2.4 Carcinome lobulaire in situ (CLIS)[7].	19
Figure 2.5 Carcinome canalaire infiltrant (CCI)[7]	20
Figure 2.6 Carcinome lobulaire infiltrant (CLI)[7].....	21
Figure 3.1 Principales phases impliquées dans un système de détection / diagnostic thermique du cancer du sein.....	30
Figure 3.2 échantillons de la base de données d'images, a) image infrarouge avec une résolution de 680 × 480, b) latéral droit 45 °, c) latéral droit 90 °, d) latéral gauche 45 ° et e) latéral gauche 90 ° [5].....	31
Figure 3.3 illustre des échantillons de la base de données d'images, a) image infrarouge avec une résolution de 680 × 480, b) latéral droit 45 °, c) latéral droit 90 °, d) latéral gauche 45 ° et e) latéral gauche 90 ° [5].....	32
Figure 3.4 illustre des échantillons de la base de données d'images, a) image infrarouge avec une résolution de 680 × 480, b) côté droit 45 °, c) côté droit 90 °, d) côté gauche 45 ° et e) côté gauche 90 [5].....	33
Figure 3.5, schéma du plan de travail.....	34
Figure 3.6, l'image d'entrée.....	35
Figure 3.7, images après redimensionnement.....	35
Figure 3.8, l'avant et l'après exécution de l'algorithme d'extraction d'arrière-plan, (a) c'est l'image original, (b) l'image après l'exécution	36
Figure 3.9, l'élimination des 100 premiers pixels	36
Figure 3.10, illustre le découpage en deux parties, (droite et gauche)	37
Figure 3.11, plan de détection du ROI.....	38
Figure 3.12, illustration de l'importance de brouillage d'image, (a) image d'entrée, (b) image après brouillage, l'image est plus claire, (c) image d'entrée avec des grains, (d) image claire et nette.....	40
Figure 3.13 illustre l'effet de normalisation, (a) l'image originale au niveau de gris, (b) l'image aux niveaux de gris après normalisation.....	41
Figure 3.14, Exemple de vérité des données examinées (a), (b), (c) des cas normaux.....	42
Figure 3.15, Exemple de vérité des données examinées (a), (b), (c) des cas malades	42
Figure 3.16, trois exemples qui démontrent les différentes tailles	43
Figure 3.17 représentation du CNN [1]	44
Figure 3.18 Architecture de VGG16.....	45
Figure 3.19 Architecture de MobileNetV2 [29]	46
Figure 3.20 Architecture de ResNet50 [34]	47
Figure 3.21 Matrice de Confusion [32]	48
Figure 3.22 Diagramme explicatif du CBIR.....	49
Figure 3.23 L'application de la méthode sur une image	52
Figure 3.24, l'application de segmentation sur le ROI	52
Figure 3.25 Diagramme de classe.....	53
Figure 3.26 Diagramme de cas d'utilisation	55
Figure 3.27 Diagramme de cas d'utilisation patient.....	55
Figure 3.28 Diagramme de cas d'utilisation médecin	56
Figure 3.29 Diagramme de cas d'utilisation administrateur	56
Figure 4.1, résultat de la Adp-HT cas corps normale, (a) l'image originale, (b) ROI gauche, (c) ROI droite.....	59
Figure 4.2 Matrice de confusion du CNN	62
Figure 4.3 Matrice de confusion de VGG16.....	63
Figure 4.4 Matrice de confusion de ResNet50	65
Figure 4.5 Matrice de confusion de VGG16+K-ppv (K=2)	67
Figure 4.6 Matrice de confusion pour VGG16+K-ppv (K=3)	68

Figure 4.7 Matrice de Confusion pour VGG16+K-ppv (K=5).....	69
Figure 4.8 Matrice de confusion de MobileNetV2+K-ppv (K=2).....	70
Figure 4.9 Matrice de confusion de MobileNetV2+K-ppv (K=3).....	71
Figure 4.10 Matrice de confusion de MobileNetV2+K-ppv (K=5).....	72
Figure 5.1 schéma fonctionnel du system webCAD	77
Figure 5.2, logo de solution	80
Figure 5.3, page d'accueil	81
Figure 5.4 page principale pour faire les diagnostics.....	82
Figure 5.5, Identification	83
Figure 5.6, registrement	83
Figure 5.7 une barre de travail d'un admin simple avec moins d'action à faire	84
Figure 5.8, choisir une image.....	85
Figure 5.9, cliquer sur prediction (predict) pour faire une diagnostique	85
Figure 5.10, les résultats pour un cas positif.....	86
Figure 5.11, plus de détails après un clic sur plus de détails.	87
Figure 5.12 autre diagnostic anormal.....	88
Figure 5.13 autre diagnostic anormal avec plus de détails.	89
Figure 5.14 autre diagnostic anormale très voyant	90
Figure 5.15 image diagnostiquer anormale très voyante	91
Figure 5.16, un résultat pour un cas négative	92
Figure 5.17 autre résultat de cas négative	93
Figure 5.18 autre diagnostic négative	94
Figure 5.19, page principale pour les admins/super admins	95
Figure 5.20, exemple des actions possible	95
Figure 5.21, notification d'un nouveau inscrit de type doctor, et les actions possible	96

Listes des tableaux

Table 2.2.1 Comparaison entre les différents travaux connexes.....	27
Table 3.1 Dictionnaire du diagramme de classe.....	54
Table 4.1 Comparaison entre les deux algorithmes du ROI.....	61
Table 4.2 Résultats du CNN.	62
Table 4.3 Résultats du modèle VGG16.	63
Table 4.4 Résultats du modèle MobileNetV2.....	64
Table 4.5 Matrice de confusion de MobileNetV2	64
Table 4.6 Résultats du modèle ResNet50.....	65
Table 4.7 Résultats de tous les modèles.	66
Table 4.8 Résultats du VGG16+K-ppv (K=2).	66
Table 4.9 Résultats du VGG16+K-ppv (K=3).	67
Table 4.10 Résultats du VGG16+K-ppv (K=5).	68
Table 4.11 Résultats du MobileNetV2+K-ppv (K=2).	69
Table 4.12 Résultats du MobileNetV2+K-ppv (K=3).	70
Table 4.13 Résultats du MobileNetV2+K-ppv (K=5).	71
Table 4.14 Comparaison entre tous les résultats obtenus.....	72
Table 4.15 Comparaison entre les résultats de l'apprentissage supervisé et non supervisé	73
Table 4.16 résultat visuel de la distance euclidienne.....	74
Table 4.17 résultat visuel de la distance euclidienne avec SKLearn.....	75
Table 4.18 résultat visuel de la distance MSE	75
Table 4.19 résultat visuel de la distance SSIM	76
Table 5.1 Outils utilisés	79

Liste des abréviations:

Acc	Accuracy
FPR	False positive rate
GLCM	Gray-level co-occurrence matrix
KNN	K-nearest neighbors algorithm
RGB	Red, green and blue color space
MF	Matching filter
NN	Neural network
NPV	Negative predictive value
STARE	Structured Analysis of the REtina (retinal image database)
Sp	Specificity
UI	User interface
CNN	Réseau neuronal convolutif
K-ppv	k plus proches voisins
ROI	Region of interest / region d'intéret
CBIR	content-based image retrieval / La recherche d'image par le contenu
FDA	Food and Drug Administration
CAD	Computer aided diagnosis
WebCAD	Web Computer aided diagnosis
HSV	hue, saturation, value / Teinte saturation lumière
Csv	Comma-separated values / Valeurs séparées par des virgules
Adam	
SGD	Stochastic gradient descent / gradient stochastique
RMSprop	
SQL	Structured Query Language / langage de requête structurée
ANN	Reseaux Neuronal artificiel
SVM	support vector machine
Ux	User experience
MSE	Mean Squared Error
SSIM	Structural Similarity Index Measure

Chapitre 1

Contexte

Dans le domaine médical, de nombreuses méthodes d'imagerie sont utilisées pour aider à diagnostiquer et à surveiller dans différentes conditions. La plupart des méthodes d'imagerie, telles que la mammographie et la thermographie, sont non invasives mais fournissent un aperçu du corps d'une manière que les yeux humains ne peuvent pas voir. Par exemple, l'imagerie thermique peut montrer des différences de température approximatives dans les différentes régions du corps. Des différences peuvent être adoptées pour le débit sanguin et les taux métaboliques du tissu sous-cutané. L'imagerie thermique a été utilisée pour étudier de nombreuses maladies telles que le diabète, le cancer du sein, les maladies vasculaires et les maladies inflammatoires [12]. Dans le cas du cancer du sein, l'imagerie thermique peut être utilisée comme méthode d'imagerie supplémentaire pour diagnostiquer et caractériser le cancer du sein.

Motivations

Jusqu'à présent, le cancer du sein reste le type de cancer le plus courant chez les femmes et est la deuxième cause de décès par cancer. Il représente environ 30% de tous les cancers [27] et on estime que les décès par cancer du sein atteindront 2,5 millions en 2021 [7]. La détection précoce du cancer du sein est un moyen majeur de guérison et il s'agit d'une maladie hautement traitable avec 97% de chances de survie grâce à un diagnostic et un traitement précoce. Par conséquent, la motivation de cette étude est de développer un système d'outils de dépistage pour la détection précoce du cancer du sein, basé sur la technologie d'imagerie thermique infrarouge basée sur le cancer du sein, qui est un

complément à la mammographie, approuvé par la FDA (Food and Drug Administration) en 1982, offrant une nouvelle option pour la détection précoce du cancer du sein.

L'imagerie thermique infrarouge se concentre davantage sur les caractéristiques physiologiques que sur les anomalies pathologiques pour identifier les défauts dans les régions d'intérêt (ROI). L'avantage de l'imagerie thermique est que le tissu tumoral s'accompagne souvent d'une température anormale en raison de l'augmentation de l'apport sanguin et de la formation vasculaire. Par conséquent, la technologie d'imagerie thermique non invasive, indolore et non radiographique l'imagerie thermique est très appropriée pour détecter les anomalies tissulaires, offrant une performance supérieure pour les méthodes conventionnelles.

Problématiques

Questions de recherche

La détection et le diagnostic rapides et précis du cancer du sein restent une tâche difficile. Certaines méthodes ont relativement plus de capacités de localisation et de détection pour la zone d'intérêt pour le sein, tandis que d'autres n'ont que des capacités très réduite. Les différentes catégories d'algorithmes, les méthodes supervisés et non supervisés, devrait être examiné pour une étude détaillée et mis en œuvre. La sélection et la mise en œuvre des algorithmes dépendront de la meilleure précision obtenue par rapport à l'autre. Par conséquent, le choix de ces méthodes soulève les questions de recherche suivantes.

- Quelle est la précision des algorithmes de segmentation du cancer du sein examinés pour extraire la région d'intérêt du sein (ROI) des images d'entrée ?
- Les méthodes d'extraction automatisée de la région mammaire sont-elles fiables par rapport à la segmentation manuelle effectuée par des experts ?
- Quelle est la précision des approches supervisées et non supervisées du cancer du sein examinées dans le diagnostic du cancer du sein ?
- Quelle est l'utilité de la plateforme WebCAD dans le processus de dépistage et de diagnostic du cancer du sein ?

Objétive de la recherche

L'objectif principal de la thèse est de développer une plate-forme WebCAD en tant qu'outil de support pour le dépistage du cancer du sein et l'identification des anomalies du sein qui aident les médecins spécialistes du diagnostic rapide et précis. En mammographie, identifier de petits morceaux dans les tissus denses, rechercher une distorsion architecturale et prédire un type de masse bénigne ou maligne en fonction de leur forme et de leur taille, etc. Une telle stratégie et ses méthodes de détection traditionnelles qui présentent les inconvénients de l'exposition aux rayonnements, d'un coût élevé et le manque de ressources médicales, ce qui limite la popularité de la détection précoce du cancer du sein. Par conséquent, il est urgent de trouver une méthode de détection bon marché, peu coûteuse et facile à utiliser. L'imagerie thermique infrarouge, Peut être considérée comme une méthode moderne de détection du cancer du sein, est très sensible aux anomalies tissulaires causées par l'inflammation et la formation vasculaire. Pour y parvenir, les sous-objectifs impliqués peuvent être énumérés comme suit :

- Examiner une base de données publique d'imagerie thermique infrarouge du sein avec / sans lésions.
- Étudier un processus de segmentation pour la localisation et la détection de la région mammaire.
- Étudier et valider les méthodes supervisées de diagnostic du cancer du sein (module de classification).
- Étudier et valider les méthodes non supervisées de diagnostic du cancer du sein (module de récupération).
- Développer l'outil de dépistage diagnostique WebCAD en intégrant les modules développés au sein d'une application web.

Organisation du mémoire

Le chapitre 2 présente le contexte physiologique lié à la structure du sein de la femme et à la fonctionnalité de ses composants. Elle est suivie d'une brève introduction des différents types de segmentation du cancer du sein et des méthodes de classification proposées par différents auteurs.

Le chapitre 3 contient les processus détaillés qui ont été suivis lors de la mise en œuvre des différentes méthodes de segmentation du cancer du sein. En outre, le chapitre explique les différentes catégories de méthodes de diagnostic du cancer du sein: supervisé et non supervisé long avec une explication détaillée d'une méthode de chaque catégorie. Il comprend le flux de travail des algorithmes suivi du calcul de leurs mesures de performance.

Le chapitre 4 décrit les résultats obtenus à partir des algorithmes mis en œuvre et compare également les mesures de performance correspondantes avec les résultats des auteurs originaux.

Le chapitre 5 présente la conception et L'implantation du WebCAD avec les différents outils et matériels utilisés.

Le chapitre 6 présente les conclusions de la thèse sur la base de la problématique et des résultats expérimentaux.

Chapitre 2 TRAVAUX ET TECHNIQUES CONNEXES ANTÉRIEURS

2.1 Introduction

Dans ce chapitre, nous définissons la structure générale des seins féminins, y compris le contexte biologique et naturel du cancer du sein, ainsi que les types d'images médicales du cancer du sein, en particulier l'accent sur l'imagerie thermique comme source de données de ce projet de recherche. De plus un examen des travaux liés au cancer du sein sera examiné à la fin de ce chapitre.

2.2 Structure du sein de la femme

Le sein est une glande mammaire. Les principaux composants d'un sein sont les lobules qui produisent du lait et sont entourés de tissu glandulaire et adipeux. Les lobules mammaires se vident dans le canal lactifère et dans le mamelon. La figure 1.1 montre la structure et les parties susmentionnées d'un sein. Les lobules et les conduits entourent le tissu adipeux ou gras. La densité mammaire décrit le rapport approximatif du tissu adipeux au tissu glandulaire et fibreux. Les tissus denses ont moins de tissu adipeux que les seins gras et peuvent être étiquetés comme hétérogènes ou extrêmement denses. Le tissu mammaire gras peut être étiqueté comme fibroglandulaire gras ou dispersé et est communément trouvé chez les femmes plus âgées. La densité du sein peut également être liée à la croissance tumorale [27].

La plupart des tumeurs, définies comme une croissance anormale de cellules mutées, commencent soit dans le canal, soit dans le lobule. Les carcinomes canaux, y compris le carcinome canalaire in situ (CCIS) et le carcinome canalaire invasif (IDC), représentent environ 90% des cancers du sein. Les carcinomes lobulaires, y compris le carcinome lobulaire in situ (LCIS) et le carcinome lobulaire invasif (LIC), constituent.

Les seins sont appariés et les organes glandulaires symétriques jouent un rôle important chez les femmes dans la forme de son corps et l'exercice de sa féminité dans les domaines émotionnel et social. La principale fonction naturelle du sein est la lactation, c'est-à-dire la production de lait: il est destiné à nourrir le nouveau-né. En surface, le sein est formé: 1) le teint, 2) Aréole et 3) le mamelon. En interne, le sein est principalement composé de 1) lobules, 2) du flux de lait et 3) du tissu adipeux conjonctif.

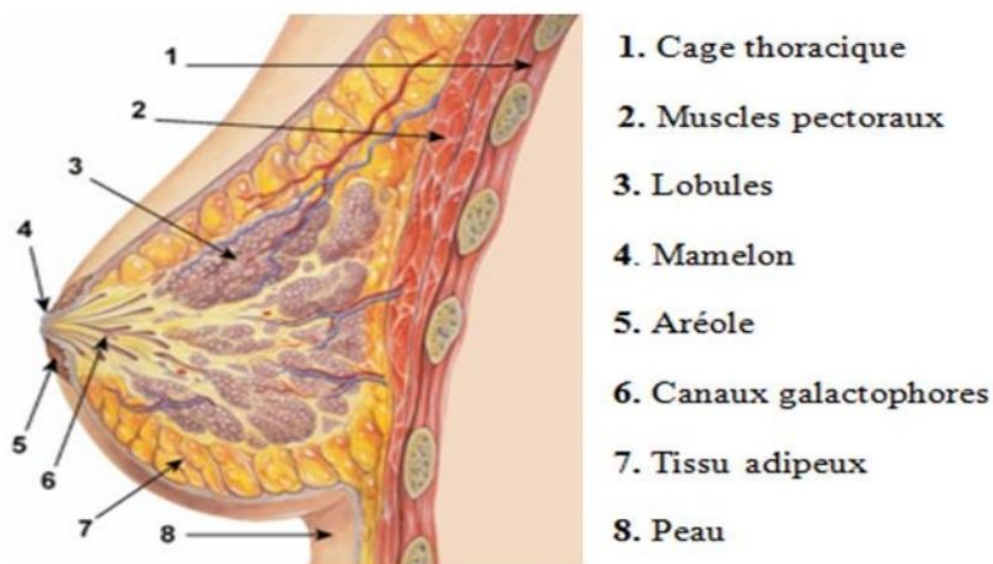


Figure12.1 Structure du sein féminin [6]

2.3 Cancer du sein

Le cancer du sein est une tumeur maligne de la glande mammaire. Autrement dit, c'est un cancer qui naît dans les unités cellulaires dont la fonction est de sécréter le lait, les unités ducto-lobulaires du sein, essentiellement chez la femme. 8 cancers du sein sur 10 se déclarent après 50 ans.

5 à 10 % de ces cancers ont une origine génétique héréditaire ; 85 à 90 % des cas (forme dite sporadique ou non-héréditaire) ont des origines environnementales ou inconnues. Une

proportion importante des cancers du sein sporadiques est induite par la prise d'hormones, œstrogènes et la progestérone contenue dans les contraceptifs ou les traitements contre la ménopause. Les facteurs de risque sont également la consommation d'alcool, de sucre, de produits laitiers, de graisses hydrogénées, l'obésité, le manque d'activité physique. Une première grossesse tardive et l'absence d'allaitement favoriseraient aussi ce cancer [6].

2.4 Type de Cancer du sein [7]

Il existe d'autres types rares de cancer du sein, On distingue : les cancers in situ et les cancers infiltrant :

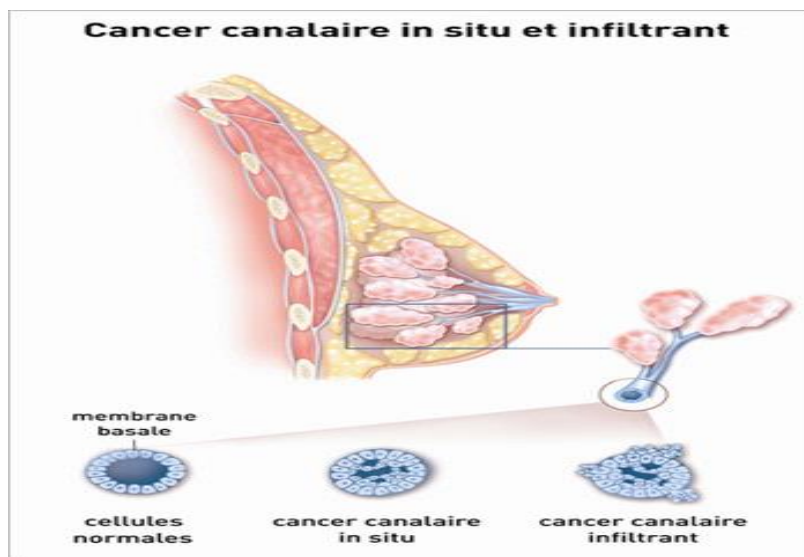


Figure 2.2 cancer canalaire in situ et infiltrant[7]

2.4.1 Les cancers in situ

- **Carcinome canalaire in situ (CCIS) :**

C'est le type le plus fréquent de cancer du sein non-invasif chez la femme. C'est une forme de cancer très précoce. Comme son nom l'indique, cette forme de cancer se développe à l'intérieur des canaux de lactation du sein. Les cellules anormales ne se sont pas dispersées à l'extérieur des canaux de lactation.

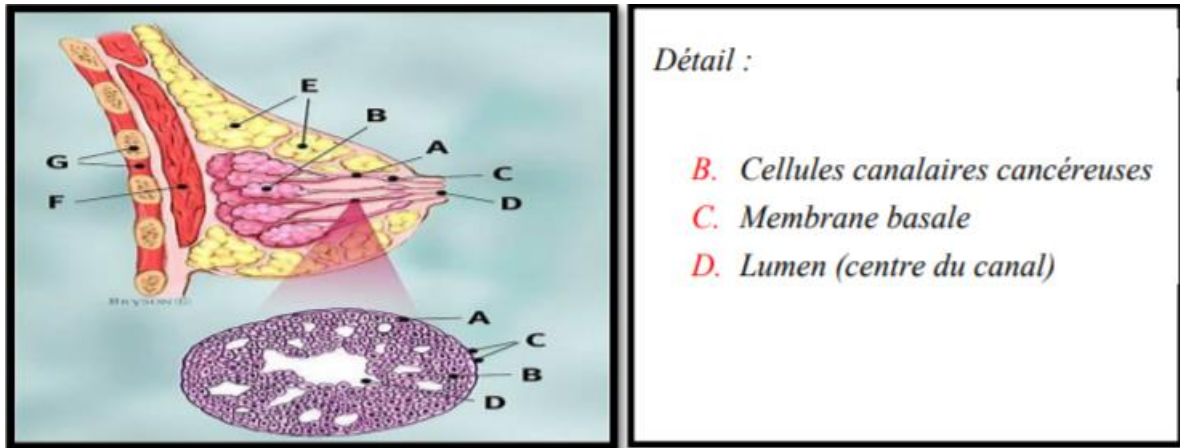


Figure 2.3 Carcinome canalaire in situ (CCIS)[7]

- **Carcinome Lobulaire in situ (CLIS):**

Le carcinome lobulaire in situ (CLIS) n'est pas vraiment un état précancéreux ou un vrai cancer du sein. Le CLIS est un marqueur révélant qu'une femme risque davantage d'être un jour atteinte d'un cancer du sein lobulaire ou canalaire infiltrant. Dans le cas du CLIS, des cellules anormales s'accumulent dans les lobules du sein, mais elles ne se propagent pas hors des lobules jusque dans le tissu mammaire voisin. Le CLIS apparaît souvent dans bien des parties différentes du sein et il est plus susceptible d'affecter les deux seins.

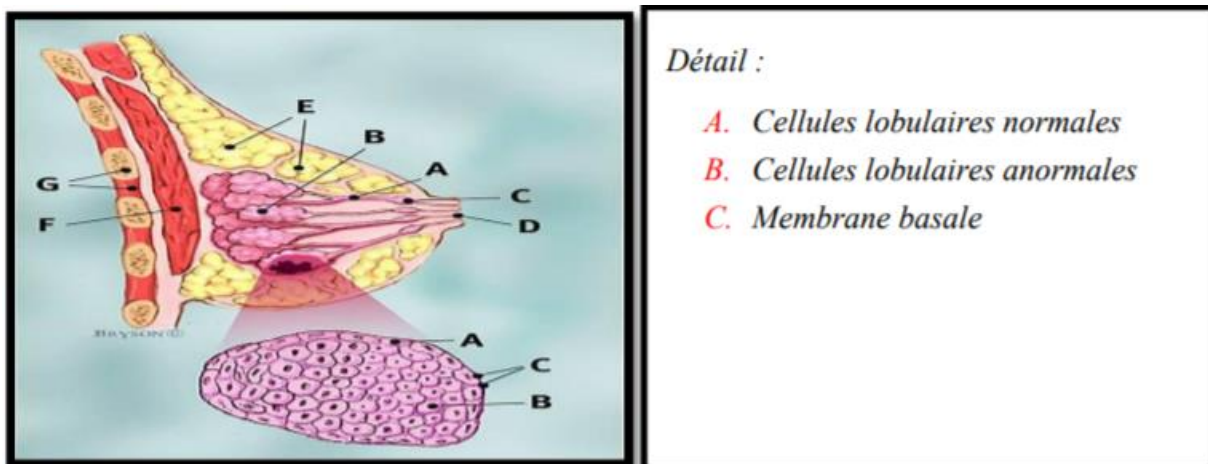


Figure 2.4 Carcinome lobulaire in situ (CLIS)[7].

2.4.2 les cancers infiltrant

Il existe différents types de cancers infiltrants. Certains sont plus fréquents et d'autres plus rares. Parfois, une femme peut présenter une combinaison de différents cancers infiltrants. Le

carcinome canalaire in situ (CCIS) accompagne souvent le carcinome infiltrant. Nous vous parlerons des 2 types les plus fréquents, le carcinome canalaire infiltrant et le carcinome lobulaire infiltrant.

- **Carcinome canalaire infiltrant (CCI) :**

Est le type de cancer le plus commun. Il représente 76 % des cancers infiltrants. Le cancer provient des canaux mais a envahi le reste du sein. Il peut se présenter comme une masse palpable ou être trouvé lors d'une mammographie de dépistage.

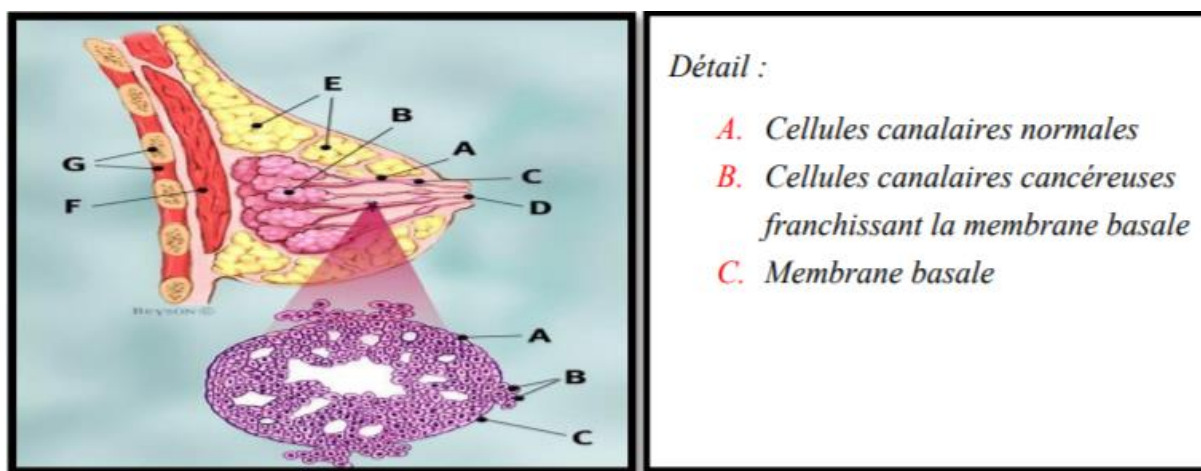


Figure 2.5 Carcinome canalaire infiltrant (CCI)[7]

- **Le carcinome lobulaire infiltrant (CLI) :**

Est le second plus fréquent type de cancer infiltrant du sein, représentant 8 % des cancers infiltrants. Il est souvent difficile à palper et à identifier à la mammographie. L'étendue réelle de ce cancer sera souvent beaucoup plus grande que ce qui était soupçonné à l'examen clinique du sein ou à la mammographie. Le carcinome canalaire in situ peut accompagner le carcinome lobulaire infiltrant. Le carcinome lobulaire infiltrant peut se retrouver plus fréquemment dans les 2 seins.

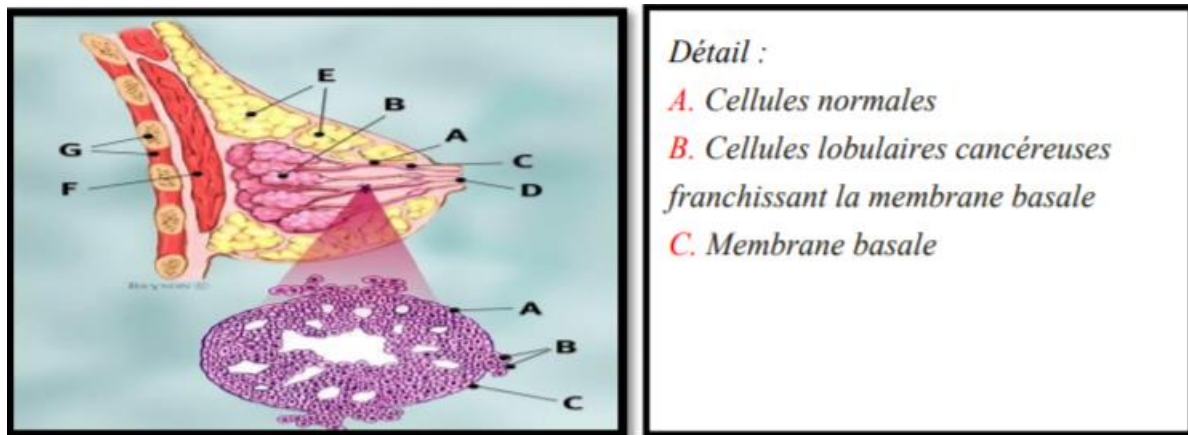


Figure 2.6 Carcinome lobulaire infiltrant (CLI)[7]

2.5 Travaux Connexes

Plusieurs travaux et recherches ont été effectués tout le long de ses dernières années au sujet du cancer du sein en utilisant la mammographie ou la thermographie pour améliorer et aider le dépistage et la prévention du cancer bien avant que le patient ait contracté la maladie, Voici quelques travaux effectués sur ce sujet :

Maitra et al. (2011) ont proposé une technique pour identifier la croissance anormale de masses dans le sein en utilisant des algorithmes très simples. Le diagnostic est l'une des meilleures technologies actuellement utilisées pour le diagnostic du cancer du sein. Dans leur article, une méthode a été développée pour faire un outil de support, cela rendra l'identification des masses anormales en mammographie numérique les images plus faciles et moins longues. La technique d'identification est divisée en deux parties distinctes. Formation de blocs et de couleurs homogènes quantification après prétraitement. La forme et la répartition des masses, la taille des masses, type de masses, orientation des masses et symétrie entre deux paires sont clairement localisés après l'exécution de la méthode proposée sur mammographie peut faciliter la détection des anomalies à un stade plus précoce [18].

Naveed et al (2011) ont proposé une nouvelle technique pour améliorer la classification des mammographies malignes et bénignes à l'aide de la classification des mammographies malignes en six classes d'anomalies. Les caractéristiques sont extraites de l'image et passées à travers un classificateur différent [22].

Zadeh, et al (2012) ont utilisés dans leurs recherches sur les algorithmes génétiques et réseaux neuronaux artificiel pour améliorer le diagnostic du cancer du sein, et ils ont tenté de diagnostiquer le cancer en traitant les données quantitatives et qualitatives des informations obtenues par imagerie infrarouge médicale. Après avoir analyser ses informations. Les meilleurs paramètres de diagnostic parmi les paramètres disponibles sont sélectionnés dans le diagnostic du cancer en utilisant l'algorithme génétique et réseau neuronal artificiel [38].

Narang, et al (2012) présente un aperçu de la classification du cancer du sein à partir de la mammographie en utilisant un réseau neuronal à résonance adaptative (ARNN), les performances du réseau sont évaluées en utilisant Ensemble de données sur le cancer du sein de l'université du Wisconsin [21].

Raad, et al (2012) le but de leur étude est de proposer une approche pour le cancer du sein en distinguant les différentes classes du cancer. Cette approche est basée sur les ensembles de données de diagnostic des caractéristiques et la classification des différents types de cancer du sein .L'ensemble de données se compose de neuf fonctionnalités qui représentent la couche d'entrée au réseau neuronal. Le réseau neuronal classera l'entrée caractéristique en deux classes de type de cancer (bénigne et maligne) [28].

Yasmine et al (2013) ont utilisé des techniques de traitement d'image pour la détection de la tumeur du sein et pour interpréter son stade dans certains cas afin qu'un traitement approprié puisse être administré au patient atteint de cancer et améliorer sa qualité de vie. Ils ont utilisé la mammographie comme images et information d'entré [37].

Gayathri et al (2013) ont utilisé divers algorithmes d'apprentissage automatique (supervisé Apprentissage, apprentissage non supervisé, apprentissage semi-supervisé, transduction et des méthodes pour améliorer la précision de la prédiction du cancer du sein [9] .

Francis et al (2014) ont effectué une première étape de prétraitement d'image thermique pour extraire la région d'intérêt après ils ont utilisé cela pour trouver et extraire la tumeur et enfin pour savoir si cette tumeur et bénigne ou maligne ils ont travaillé avec le classifieur svm [8].

Villalobos-Montiel A. J et al (2015), l'algorithme proposé commence par la correction d'inclinaison, ensuite il détecte le ROI depuis l'image thermal de façon approximative avec l'algorithme de « Hough circle Transform », ensuite il élimine les régions indésirables, la prochaine étape et l'utilisation de contours actifs pour avoir des contours gras et en fin l'adoucissement des contours [36].

Mona A. S. A et al (2015) dans le prétraitement, Ils ont commencé à lire l'image thermique en niveaux de gris, ensuite prendre la hauteur et la largeur de l'image, pour les utiliser dans la fonction qui va calculer les coordonnées de début (x_1, y_1) et de fin (x_2, y_2) de la région d'intérêt (ROI), avant de la convertir en image binaire utilisent la technique de « thresholding (seuillage) » pour différencier le corps de l'arrière-plan, ensuite l'extraction de caractéristiques et la classification [20].

Sedong M et al (2017), ont entamé leur travail avec l'élimination de l'arrière-plan de l'image thermique en utilisant le filtre gaussien et les algorithmes de « Canny edge detection » et « Hough Transform », ensuite ils récupèrent la région d'intérêt en appliquant l'algorithme de « Hough circle Transform » pour avoir les parties ressemblent à un cercle, et pour avoir un bon résultat ils ajoutent l'algorithme de « Gaussian filter », et en fin l'étape d'Extraction de caractéristique et d'analyse [31].

Kandilkar et al (2017) ont effectué une première étape de prétraitement d'image thermique pour extraire la région d'intérêt puis ils ont détecté la tumeur en utilisant un réseau neuronal artificiel (ANN) et enfin pour savoir si cette tumeur est bénigne ou maligne ils ont travaillé avec le classifieur svm [14].

Hany M. H et al (2018) ils ont commencé leur solution par le prétraitement de l'image thermal, d'abord ils détectent les bords et les limites du sein, ensuite ils ont utilisé l'algorithme de « Hough transform », et pour avoir un meilleur ROI, ils améliorent le contraste des limites et en fin l'étape de extraction de caractéristiques et classification [16].

Ponraj, et al (2018) ont utilisé des opérations morphologiques afin d'améliorer le contraste de l'image de la mammographie. La morphologie a diverses opérations, lorsque ils sont appliqués à la mammographie, ils produisent une image à contraste élevé. Image l'amélioration se fait comme une étape de prétraitement. L'étape de prétraitement est nécessaire pour chaque image de la mammographie. Cette image prétraitée sert comme

entrée pour d'autres étapes de segmentation, ce qui permet d'identifier facilement la portion cancéreuse [26].

Mambou et al (2018) ont effectué une première étape de prétraitement d'image thermique pour extraire la région d'intérêt puis ils ont détecté la tumeur en utilisant un réseau neuronal profond (DNN) et enfin pour savoir si cette tumeur est bénigne ou maligne ils ont travaillé avec le classifieur svm [19].

Tello-Mijares et al (2019) ont effectué une première étape de prétraitement d'image thermique avec l'algorithme d'apprentissage supervisé pour extraire la région d'intérêt puis ils ont détecté la tumeur en utilisant un réseau neuronal convolutionnel (CNN) et enfin pour savoir si cette tumeur est bénigne ou maligne ils ont travaillé aussi avec le réseau neuronal convolutionnel (CNN) [33].

2.6 Comparaison entre les méthodes utilisées dans les travaux

Francis et al (2014) ont utilisé dans leur travail les images thermiques et Ponraj, et al (2018) ont utilisé pour leur part la mammographie tous les deux sont passés par une 1ère étape de prétraitement c'était l'extraction de la région d'intérêt (ROI) pour les 1er et les autres on utilise des opérations morphologiques pour améliorer le contraste ensuite pour localiser la tumeur Francis et al (2014) ont continué avec la région d'intérêt et après le classifieur svm pour détecter si c'est malin ou bénin cependant Ponraj, et al (2018) sont passés par plusieurs autres étapes de segmentation pour arriver à leur fin et identifier la portion cancéreuse ou non.

Mambou et al (2018) ont utilisé pour leur travail les images thermiques et Narang, et al (2012) ont utilisé pour leur part la mammographie les 1er sont passés par une étape de prétraitement c'était l'extraction de la région d'intérêt (ROI) puis ils ont utilisé un réseau neuronal profond (DNN) pour détecter la tumeur et enfin pour la classification ils ont utilisé le classifieur svm, Narang, et al (2012) ont utilisé un réseau neuronal à résonance adaptative (ARNN) tout le long de leur travail. Les données de l'Université du Wisconsin leur servent de dataset quand à Mambou et al (2018) ont utilisé la seule base de données pour les images thermiques qui est disponible sur le site web d'une université brésilienne.

Tello-Mijares et al (2019) et Hany M. H et al (2018) ont tous deux utilisé les images thermiques, ils ont extrait la région d'intérêt comme étape de prétraitement mais les 1er ont utilisé un algorithme d'apprentissage supervisé et Hany M. H et al (2018) on préféré détecter d'abord les bords et les limites du sein, ensuite l'algorithme de « Hough transform » après ils ont continué avec plusieurs étapes pour arriver à la classification de la tumeur avec le SVM « Support Vector Machine » plus le ANN, les réseaux de neurone artificiel cependant Tello-Mijares et al (2019) ont travaillé avec un réseau neuronal convolutionnel (CNN) pour tout le travail qui reste jusqu'à la classification de la tumeur .

Chaque un des deux travaux basés sur l'image thermique de Mona A. S. A et al (2015) et de Hany M. H et al (2018), on fait leur propre prétraitement, pour la première solution, ils ont commencé par lire l'image au niveaux de gris en suite ils ont utilisé une fonction pour calculer la position de la région d'intérêt, contrairement à la deuxième solution, où ils ont utilisé aussi le « thresholding » et la méthode « Canny edge detection », cela a permis d'avoir la forme du corps, Mona A. S. A et al (2015), construisent des fonctions pour découpé l'image et avoir le ROI, pour Hany M. H et al (2018), de plus les fonctions mathématiques, ils ont ajouté une autre étape sel de « Hough transform algorithm », et pour les deux, à la fin ils ont extrait les caractéristiques et on fait une classification, en suite ils passent à l'apprentissage, Hany M. H et al (2018), préfère le SVM « Support Vector Machine » plus le ANN, les réseaux de neurone artificiel, par contre Mona A. S. A et al (2015)) on choisit que le SVM.

Sedong M et al (2017) et Mona A. S. A et al (2015), tous les deux ont ouvert l'image au niveaux de gris, Sedong M et al (2017), trouve d'abord le contour en utilisent « Canny edge algorithm and Hough transform », ensuite relèvent le fond d'image en couleur par extraction du fond en utilisent le contour déjà détecté. Pour Mona A. S. A et al (2015) par leur tour, passent directement à la détection du ROI en utilisent des formules mathématiques, Sedong M et al (2017) par son coté, l'extraction de ROI ce fait par dessiné un cercle sur chaque sein, et ça se fait par le « Hough Circle Transform », et en fin l'étape d'Extraction des caractéristiques et d'analyse pour les deux solutions. Enfin les deux on utiliser le SVM « Support Vector Machine » comme méthode de classification.

Villalobos-Montiel A. J et al (2015), trouvent qu'il faut corriger l'inclinaison d'image ou le corps est incliné, ensuite ils ajoutent les deux algorithmes « Gaussian filtre et Hough Circle Detection » après que l'image soit coupée en 2 partie gauche et droite, contrairement à Hany M. H et al (2018) ils cherchent à trouver directement le ROI par les algorithmes « Canny edge detection » et « thresholding algorithm » ensuite il faut passer par des fonctions mathématiques pour avoir les bonnes coordonnées de la poitrine, ensuite couper l'image en deux parties gauche et droite, après avoir le ROI, les deux algorithmes de Villalobos-Montiel A. J et al (2015) et de Hany M. H et al (2018), passent à l'étape de classification.

Mona A. S. A et al (2015) et Villalobos-Montiel A. J et al (2015), les deux cherchent à trouver d'abord le ROI, la première méthode par des formules mathématiques et la deuxième en utilisant l'algorithme de « Hough Circle Transform », en suite ils passent à l'apprentissage, Mona A. S. A et al (2015) préfère le SVM « Support Vector Machine » par contre Villalobos-Montiel A. J et al (2015) n'ont pas mentionné la technique utilisée dans la recherche.

Hany M. H et al (2018) et Sedong M et al (2017), les deux dans la phase de prétraitement, ont choisi les algorithmes « Hough Transform » et « Canny edge detection », de plus Hany M. H et al (2018) ajoutent « thresholding algorithm », Sedong M et al (2017) de son tour utilisent l'algorithme « Hough Circle Transform » après la restriction du fond d'image. La dernière phase après le « ROI detection » c'est la classification et l'apprentissage, Hany M. H et al (2018) ainsi que Sedong M et al (2017), leur système sont basé sur le SVM « Support Vector Machine » et les réseaux de neurone artificiel ANN.

Les deux, Sedong M et al (2017) et Villalobos-Montiel A. J et al (2015) dans le prétraitement ont passé par « Gaussian Filter algorithm » ensuite ils ont choisi le « Hough Circle Transform » pour la détection du ROI. A la différence que Villalobos-Montiel A. J et al (2015). Nous avons pensé à corriger la position du corps lorsqu'il est incliné, Sedong M et al (2017) et leur travail est basé sur l'élimination du fond en gardant que le corps.

Le tableau ci-dessous résume la comparaison qui a été faite entre quelques travaux connexes on comparant les types d'images utilisés (thermique ou mammographie) et l'étape de prétraitement puis l'extraction des caractéristiques et enfin la classification.

Travaux	Types D'images	Prétraitements	Extraction des caractéristiques	Classification
Narang, et al (2012)	Mammographie	réseau neuronal à résonance adaptative (ARNN)	réseau neuronal à résonance adaptative (ARNN)	réseau neuronal à résonance adaptative (ARNN)
Francis et al (2014)	Thermique	Extraction de la zone d'intérêt (ROI) avec diverses étapes	Utilisation du ROI	Classifieur SVM
Mona A. S. A et al (2015)	Thermique	Extraction de la zone d'intérêt (ROI) avec diverses étapes	Utilisation du ROI	Classifieur SVM
Villalobos-Montiel A. J et al (2015)	Thermique	Extraction de la zone d'intérêt (ROI) avec diverses étapes	Utilisation du ROI	Autre étapes de segmentation et de classification
Sedong M et al (2017)	Thermique	Extraction de la zone d'intérêt (ROI) avec diverses étapes	Utilisation du ROI	Classifieur SVM
Hany M. H et al (2018)	Thermique	Extraction de la zone d'intérêt (ROI) avec diverses étapes	détection des bords et limites du sein	Classifieur SVM + réseaux de neurone artificiel (ANN)
Ponraj, et al (2018)	Mammographie	Amélioration du contraste	Segmentation de l'image prétraitée	Autre étapes de segmentation et de classification
Tello-Mijares et al (2019)	Thermique	Extraction de la zone d'intérêt (ROI) avec diverses étapes	algorithme d'apprentissage supervisé	avec un réseau neuronal convolutionnel (CNN)

Table 2.2.1 Comparaison entre les différents travaux connexes

2.7 Conclusion

Dans ce chapitre on a abordé quelques travaux connexes qui ont été faits sur la détection du cancer du sein en utilisant la mammographie et la thermographie puis nous avons fait

une comparaison entre quelques travaux. Dans le prochain chapitre nous aborderons toutes les méthodes et les algorithmes utiliser pour la phase de détection et la phase de diagnostic.

Chapitre 3 DESIGN ET MÉTHODOLOGIE

3.1 Introduction

Dans ce chapitre nous allons aborder en détails toutes les méthodes et algorithmes que nous avons utilisés lors de la phase de diagnostic mettant en évidence un ensemble d'illustrations issu de notre travail.

3.2 Schéma proposé

À travers la revue des articles discutés précédemment, on a remarqué l'importance du traitement d'image, qui est actuellement bien effectué par un être humain mais qui n'est pas encore adéquat lorsqu'il est effectué par des méthodes d'intelligence artificielle. Dans ce contexte, un CAD sera un réseau neuronal profond avec différents modèles en tant que classificateur des images de sein ROI résultantes de la phase de détection (processus de segmentation et extraction de caractéristiques) comme le montre la figure 3.1 (en supposant qu'il est déjà formé) qui prendra les images thermiques dans et en sortie classent les images comme contenant un cancer ou non. Nous devons préciser que le module d'apprentissage en profondeur produira la probabilité que l'image thermique d'un sein soit classée comme malade (ayant un cancer) ou saine (sans cancer). Si la sortie de confiance (malade) de notre modèle de classificateur / processus d'appariement est supérieure au seuil, il appartiendra à notre classification de distinguer l'état du sein du patient en fonction de la matrice des caractéristiques de l'image thermique du sein. L'architecture proposée, comme indiqué ci-dessous, peut être liée à un organigramme où chaque phase représente un composant spécifique.

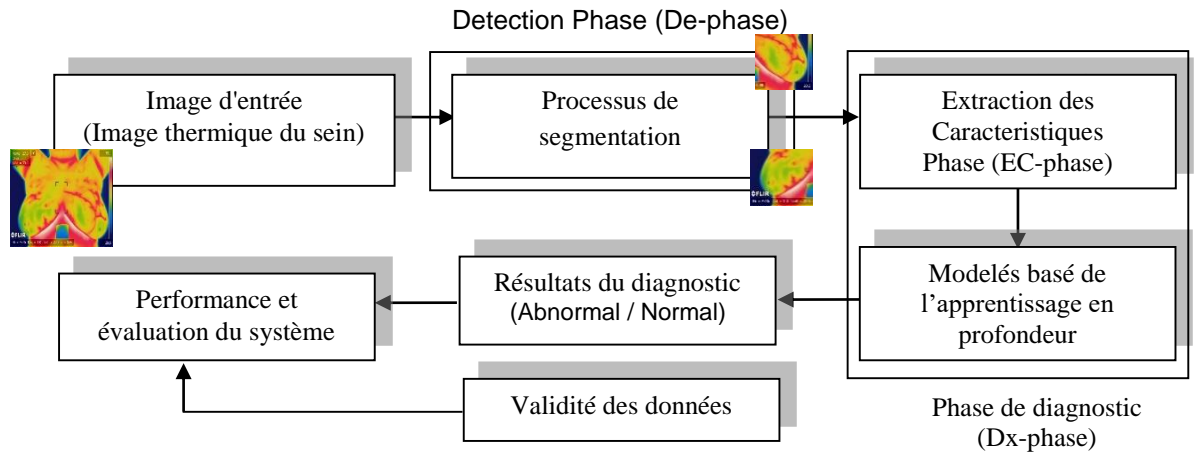


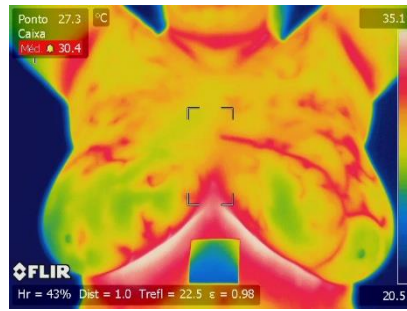
Figure 3.1 Principales phases impliquées dans un système de détection / diagnostic thermique du cancer du sein

3.3 Base de données

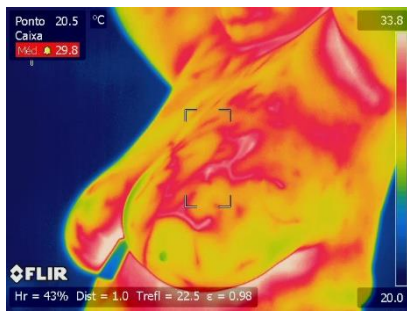
Dans ce travail , la base de données d'images a été prise à partir d'une base de données de recherche (DMR) contenant des images thermiques frontales, acquises avec une caméra IR FLIR SC-620 avec une résolution de 640 x 480 pixels. La source de données est un ensemble de données public obtenu selon des protocoles spécifiques d'acquisition d'images et contient des images de personnes âgées de 29 à 85 ans. Ces images incluent des seins de différentes formes et tailles, tels que des seins moyens, larges et asymétriques. Les données démographiques des sujets sont présentées dans le tableau 2, avec:

- Nombre total de sujets (N) = 67.
- Nombre total de sujets sains / normaux (NH) = 43.
- Nombre total de sujets malades / anormaux (NS) = 24.
- Chaque participant est capturé en moyenne 27 images.
- Le protocole statique fournit une position de prise de vue différente (par exemple, avant, latéral droit 45 °, latéral droit 90 °, latéral gauche 45 ° et latéral gauche 90 °) et le protocole dynamique génère la position avant des 20 images séquentielles et de l'autre 2 images latérales (par exemple, latéral droit 90 °, gauche latéral 90 °) [5].

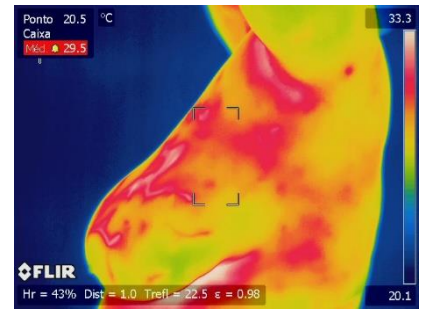
Voici dans les figures ci dessus (3.2 , 3.3 et 3.4) illustrent des exemples d'images de notre base de données.



(a)



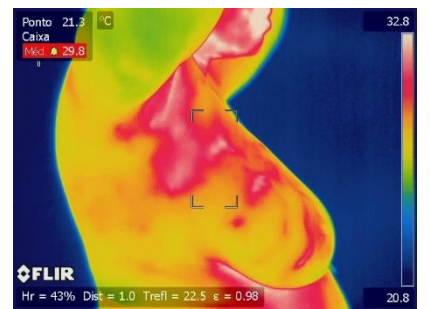
(b)



(c)

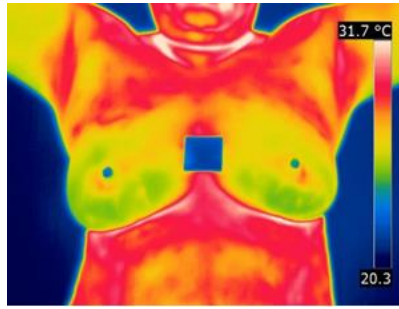


(d)

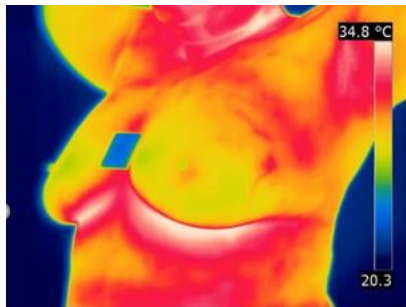


(e)

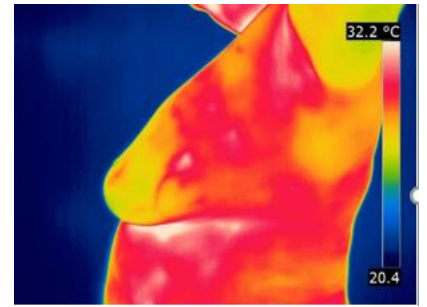
Figure 3.2 échantillons de la base de données d'images, a) image infrarouge avec une résolution de 680×480 , b) latéral droit 45° , c) latéral droit 90° , d) latéral gauche 45° et e) latéral gauche 90° [5].



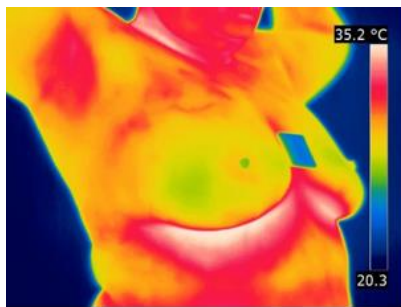
(a)



(b)



(c)



(d)

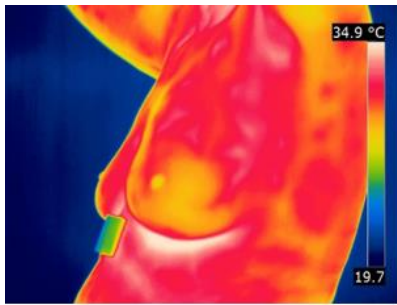


(e)

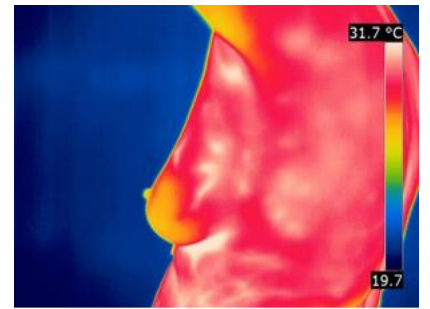
Figure 3.3 illustre des échantillons de la base de données d'images, a) image infrarouge avec une résolution de 680×480 , b) latéral droit 45° , c) latéral droit 90° , d) latéral gauche 45° et e) latéral gauche 90° [5].



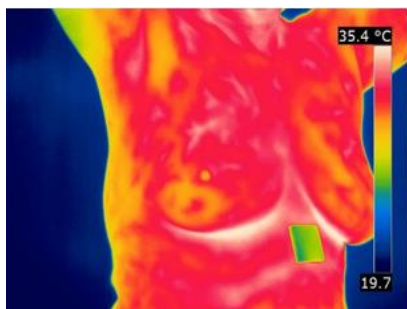
(a)



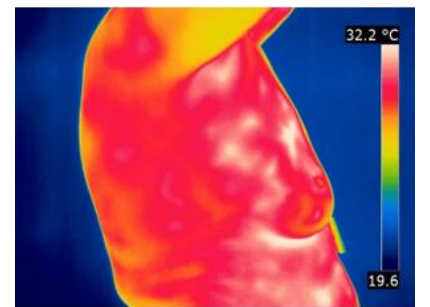
(b)



(c)



(d)



(e)

Figure 3.4 illustre des échantillons de la base de données d'images, a) image infrarouge avec une résolution de 680×480 , b) côté droit 45° , c) côté droit 90° , d) côté gauche 45° et e) côté gauche 90° [5].

3.4 Phase de detection:

Cette phase se base sur deux principales procedures nommées le prétraitement et la segmentation qui sont détaillées dans les sections suivantes :

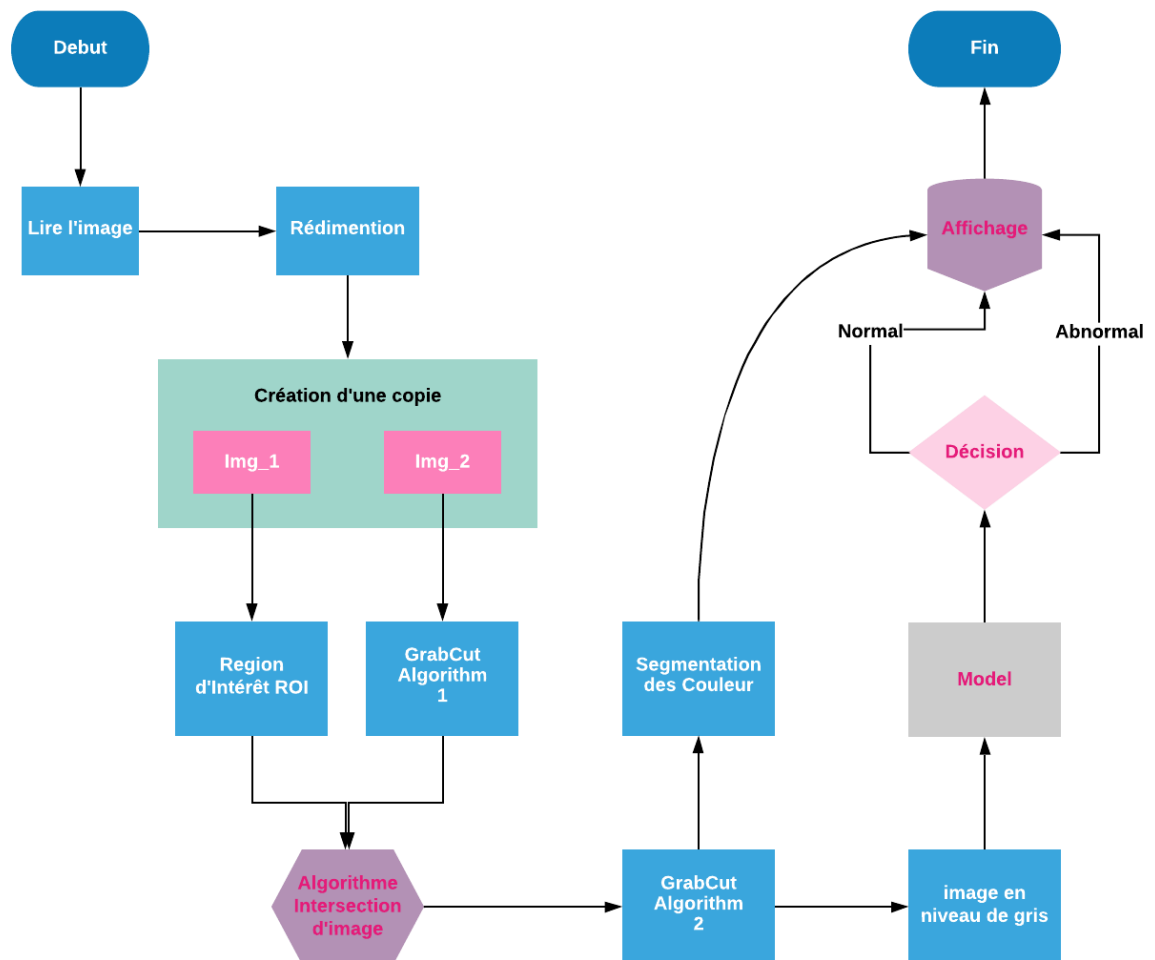


Figure 3.5, schéma du plan de travail

3.4.1 Prétraitement :

Le traitement d'image est la première phase dans tous systèmes de diagnostic assisté par ordinateur (CAD : Computer Aided Diagnosis), et c'est l'ensemble des méthodes et fonctions utilisé pour avoir la bonne forme d'image pour la solution donner.

Après la lecture d'image sous format RGB (trois canal, Red, Green & Blue pour respectivement Rouge, Vert et Bleu), dans la figure [3.5] on a utilisé ensuite pour la segmentation de couleur, et une autre sous format de niveau de gris et c'est l'entrée de la phase de détection.

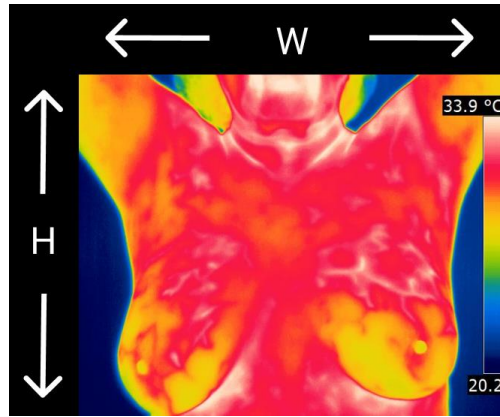


Figure 3.6, l'image d'entrée

Nous avons appliqué les étapes suivantes :

- Le redimensionnement d'image :

Les images ne sont pas toujours avec la même dimension, donc à l'entrée d'image on doit la redimensionner, tel que la taille soit comme ça largeur = 400px et l'hauteur = 300px comme il est montré dans la figure [3.6].

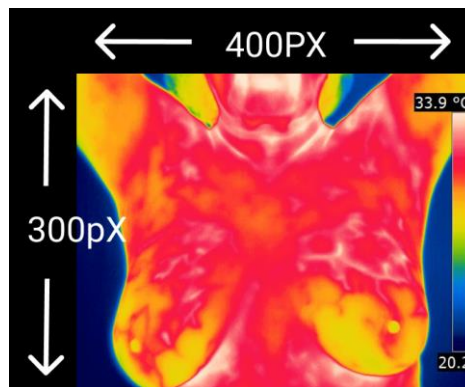


Figure 3.7, images après redimensionnement

- Extraction de premier-plan (arrière-plan) :

Cette étape concerne que l'image en RGB, pour que la précision soit plus grande, il faut éliminer tous les objets dans l'arrière-plan, pour que l'algorithme se concentre sur les objets nécessaires.

Les images fournies ont plusieurs couleurs et objets en arrière-plan, comme la figure [3.7] le montre.

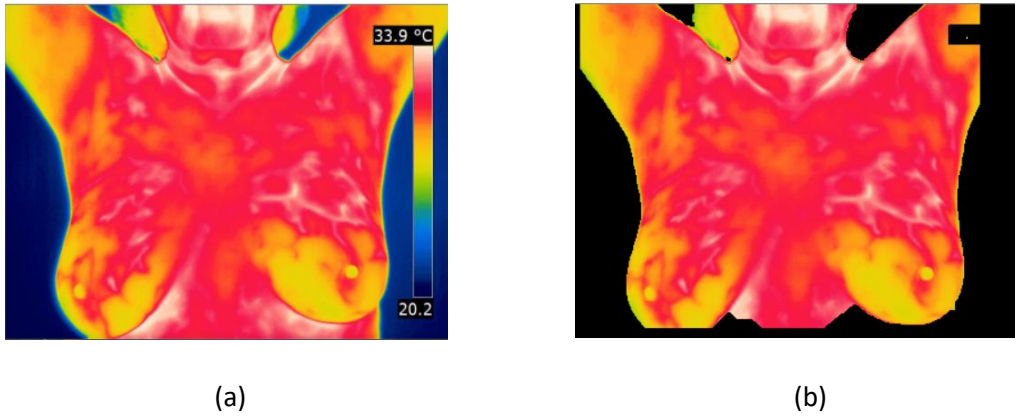


Figure 3.8, l'avant et l'après exécution de l'algorithme d'extraction d'arrière-plan, (a) c'est l'image originale, (b) l'image après l'exécution

- Découpage d'image :

Nous nous intéressons au sein, donc on va éliminé les parties non importants avec ses étapes :

- Apres des centaines de test on a trouvé que si on pas éliminé les 100 premiers pixel le resultats de detection du seain avec notre algorithm sera plus facile et plus efficace. Tel que : le premier pixel se trouve aux point $(x, y) = (0, 0)$, et le dernier pixel c'est (l'originale largeur, 100). Voir figure [3.8]

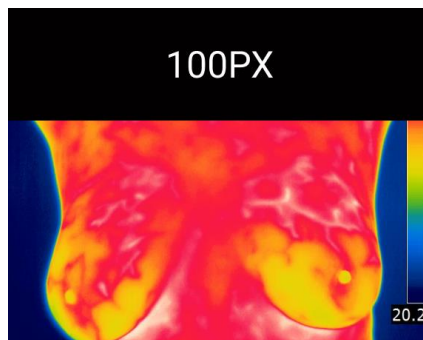


Figure 3.9, l'élimination des 100 premier pixels

- Le Découpage au milieu en deux parties gauche et droite, cette etape reduite le temps de detection et les erreurs de detection d'objet, la recherche dans une petite region et plus efficace. Le découpage est comme suit :

Pour la partie gauche, du pixel (0, 0) jusqu'a le pixel (200, l'hauteur d'image)

Pour la partie droite, du pixel (200, 0) jusqu'a le pixel (400, l'hauteur d'imSage)

Voir figure [3.9]

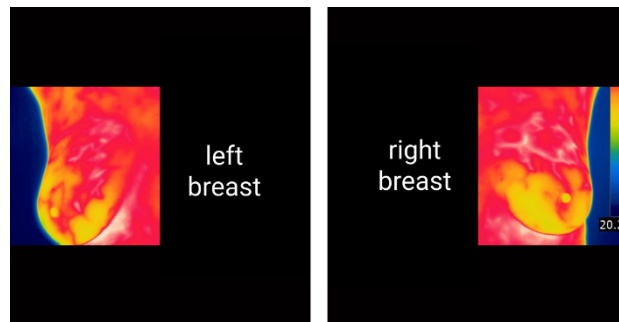


Figure 3.10, illustre le découpage en deux partie, (droite et gauche)

3.4.2 Procédure de la segmentation

C'est la partie qui va détecter la région d'intérêt voir figure [3.11] (ROI en anglais Region Of Interest) pour les deux seins, les résultats de cette procédure influencent directement le résultat final de la recherche.

A ce stade, on a implémenté trois méthodes, une manuel et deux automatique.

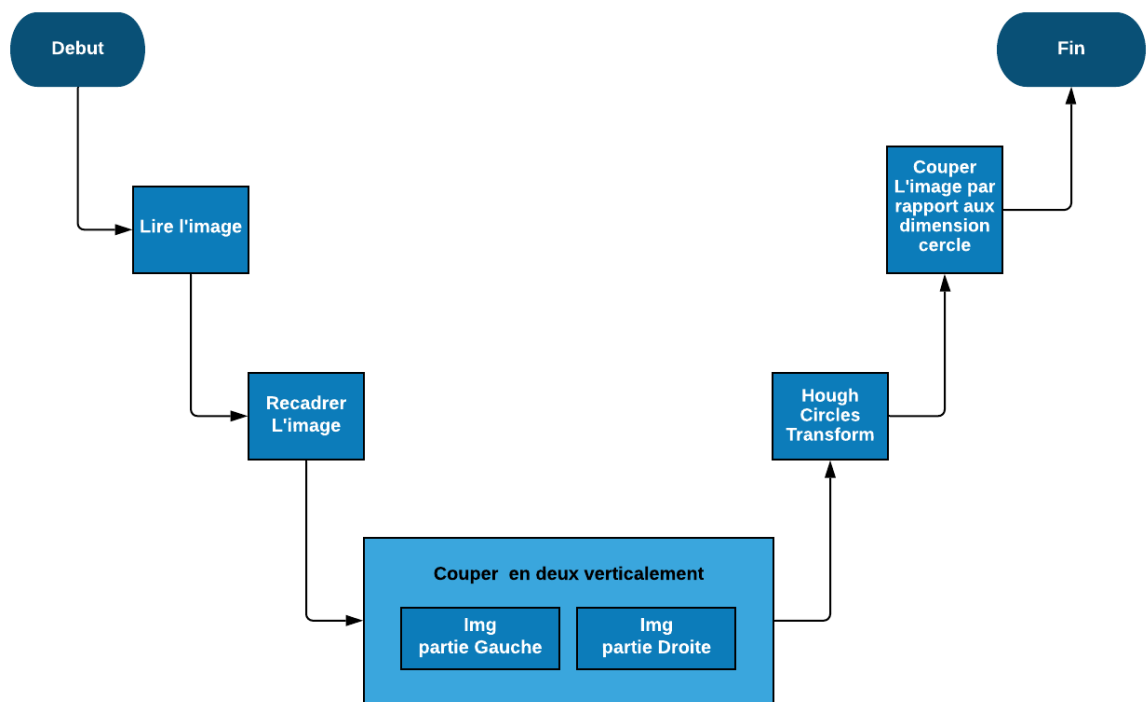


Figure 3.11, plan de détection du ROI

Manuel

Dans cette méthode, c'est l'utilisateur qui va choisir les coordonnées du ROI gauche en suite droite, s'il estime que la position est correcte (par rapport au cercle dessiné sur l'image) alors il peut la sauvegarder, sinon il peut réinitialiser l'image (image sans aucun cercle). Les coordonnées seront stockées dans un fichier CSV afin de les comparer avec les méthodes Adp-HT et Adp-BHT (les deux méthodes adaptative).

La méthode Adp-HT (adaptative de transformée de Hough)

Dans cette méthode on a pas mis tous étapes de prétraitement mentionné en haut, tel que l'extraction d'arrière-plan et le brouiller l'image. On a construit notre algorithme en suivant ces étapes:

- Hough Circles Transform Algorithm

Après l'élimination de fond d'image, on cherche la position de sein, généralement le sein a une forme ovale, donc grâce à cette algorithme on peut détecter des parties d'image ou ça forme rassemble à un cercle ou une partie de cercle [11].

Un cercle est représenté par « $(X - X_{center})^2 + (Y - Y_{center})^2 = r^2$ » ou (X_{center}, Y_{center}) est le centre du cercle, r est le rayon, ici « opencv » utilise la méthode « Hough Gradient »

C'est un algorithme efficace avec des objets plus ou moins ronds ou avec une forme circulaire, mais ce n'est pas le cas dans un corps humain il peut donner beaucoup de faux cercles, donc il faut l'automatiser pour qu'il détecte que le sein et à chaque fois que l'on utilise. Pour cela nous avons ajouté d'autres étapes.

- Vérifier que l'algorithme a détecté et dessiné un cercle dans chaque image (coté) et pas plus qu'un.
- Pour éliminer les faux cercles, il faut que le centre du cercle ne soit pas entre les deux points $(150, y)$ et $(250, y)$ pixels par rapport à l'image d'entrée pour éviter la détection d'un cercle au milieu du corps.
- Le rayon aux premiers temps donné par l'algorithme, mais des fois il est un peu petit et d'autre plus grand que le sein, et ça aussi peut fausser les résultats, donc après plusieurs tests on a mis le rayon entre 70px et 80px.
- Dernière étape est d'enregistrer l'image dans un dossier avec un nom qui mentionne l'original image et le coté (gauche / droite).

La méthode Adp-BHT (adaptative de transformée de Hough Booster)

L'algorithme passe par tous les étapes de prétraitements, et dans la phase de détection nous avons rajoutés quelques étapes par rapport la méthode (Adp-HT) tel que:

- Brouiller l'image:

On a utilisé cette phase pour éliminer les grains d'image (avoir une image plus clair et plus nette), avec une taille linéaire d'ouverture égale a (5).

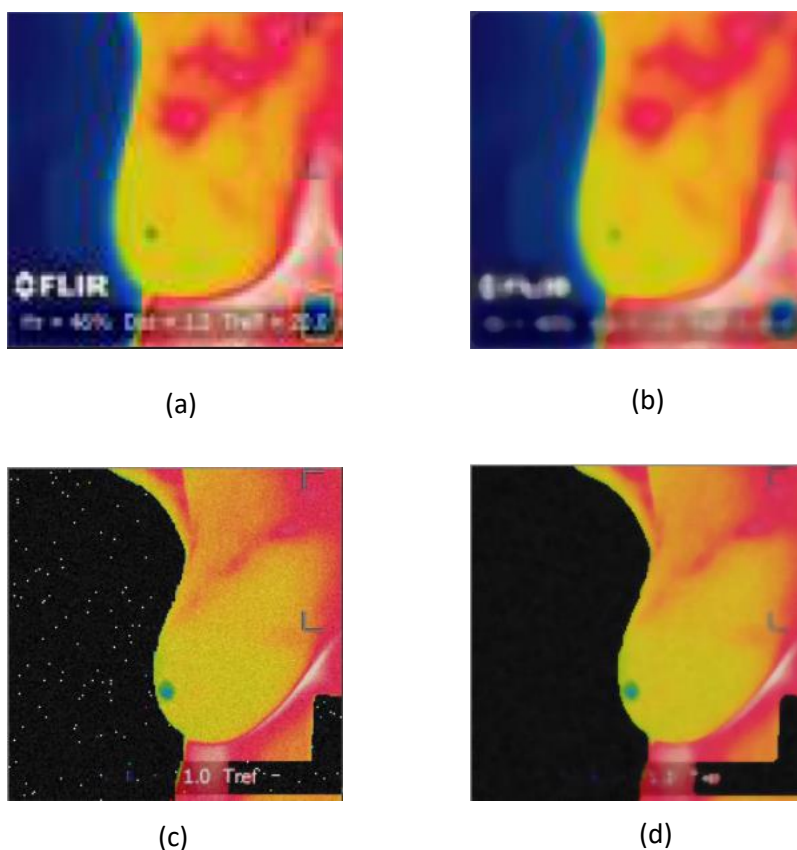


Figure 3.12, illustration de l'importance de brouillage d'image, (a) image d'entrée, (b) image après brouillage, l'image et plus Claire, (c) image d'entrée avec des grains, (d) image claire et nette

- Après le changement du rayon (c'est le faux), et pour avoir une bon résultat, on a trouvé qu'il faut aussi changer la position du centre. $(X \text{ ou } Y) > 100\text{px} : [X / Y = X / Y - | \text{rayon} - \text{nouveaux rayon} |]$

$$(X \text{ ou } Y) < 100\text{px} : [X / Y = X / Y + | \text{rayon} - \text{nouveaux rayon} |]$$

- Ensuite, on va éliminer les morceaux en excès avec le même algorithme d'extraction de premier-plan pour garder que la région du sein selon le type l'image d'entré (coté droite / gauche).

Si (coté gauche) : effectuer l'algorithme comme il est.

Sinon : changer la direction d'image avant l'algorithme .

- Normalisation D'image :

Pour avoir plus de détails et d'informations sur l'image, l'image passe par la technique de normalisation de la plage normes ou de valeurs d'un vecteur de sort que :

$$\|dst\|_{L^p} = \alpha$$

Où « dst » est la sortie, et « alpha » est la valeur normale à normaliser ou limite inférieure de la plage en cas de normalisation de la plage, « beta » est la limite supérieure de la plage en cas de normalisation de la plage, il n'est pas utilisé pour la normalisation de la norme, et « p » change selon la norme utilisée en entrant [23].

$$\min Idst(I) = \alpha, \max Idst(I) = \beta$$

La figure X illustre le résultat de la normalisation.

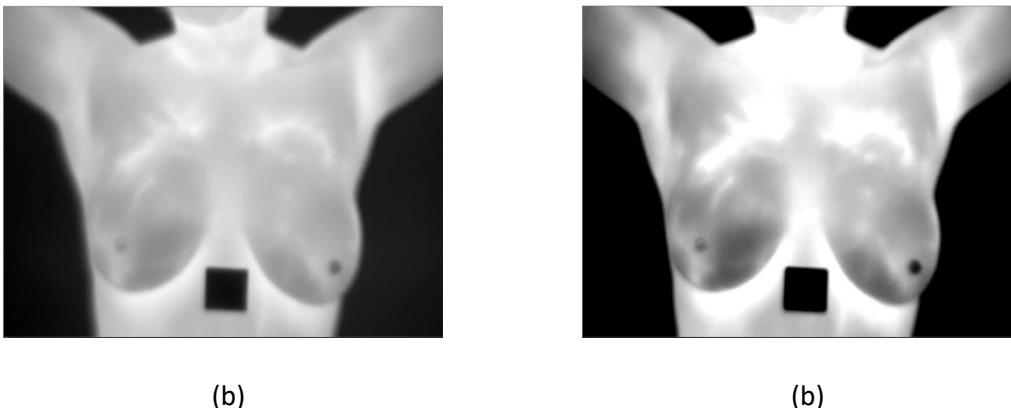


Figure 3.13 illustrer l'effet de normalisation, (a) l'image original aux niveaux de gris, (b) l'image aux niveaux de gris après normalisation

NB: Pour les trois, méthode si l'images est au niveaux de gris (cas de prédiction et non pas la segmentation des couleurs), l'algorithme ne passe pas sûr l'étape d'extraction d'arrière-plan, car le corps est déjà bien esolé de l'arrière paln, et si on fait cette etape, peut etre les résultats serons faussé.

3.5 Phase de diagnostic

Nous examinons deux stratégies, des algorithmes supervisés et non supervisés

3.5.1 Validité des données

Il faut mentionner que les images fourni par “Visual Lab” indique si les images sont des cas normal ou abnormal, aussi les images on une grande variété, des poitrines symétrique

et d'autre asymétrique par rapport au volume ou à la position, aussi des corps mince, d'autre sont obese.

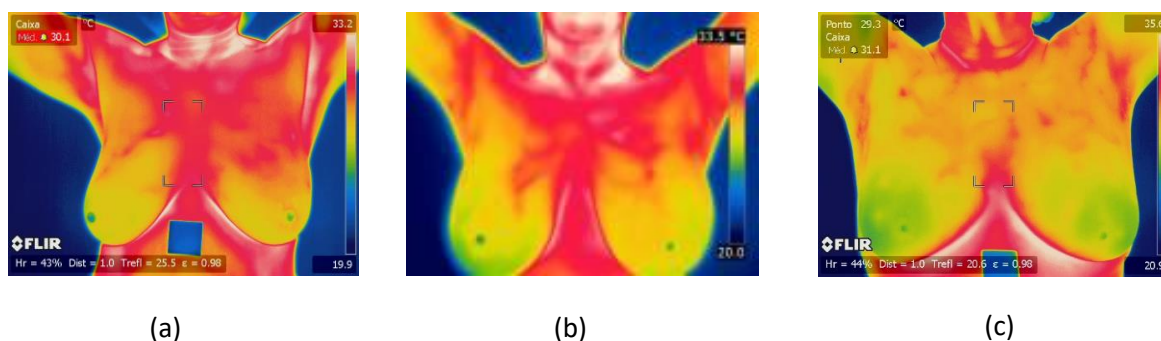


Figure 3.14, Exemple de vérité des données examinées (a), (b), (c) des cas normal

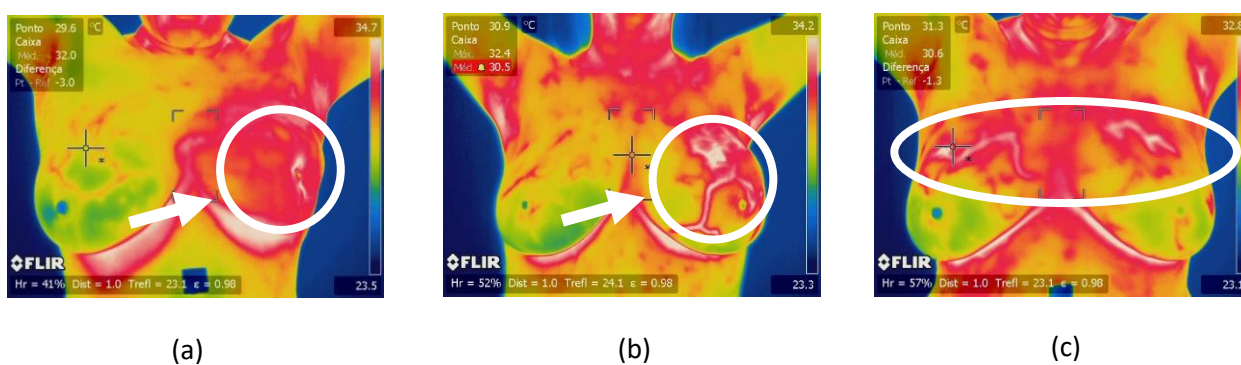


Figure 3.15, Exemple de vérité des données examinées (a), (b), (c) des cas malade

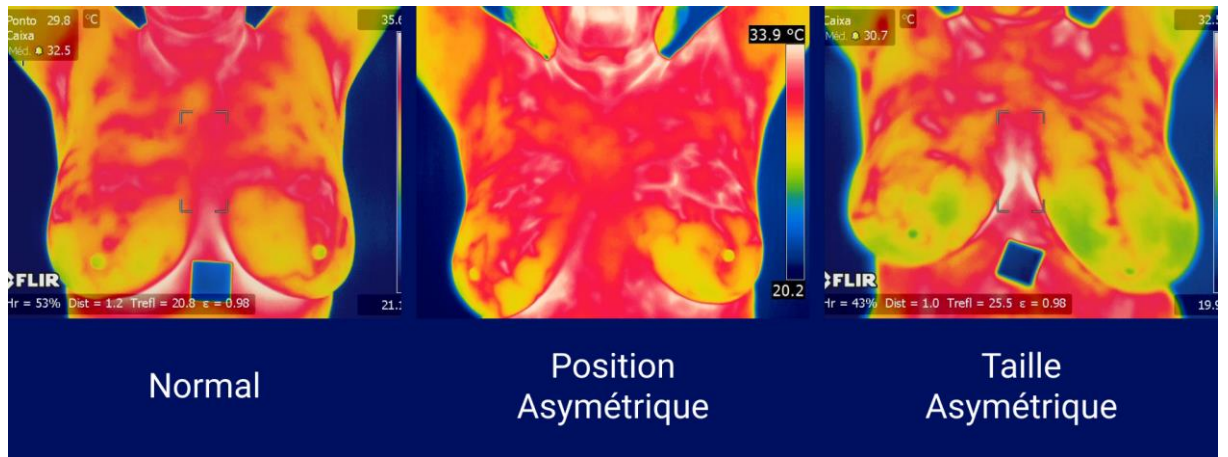


Figure 3.16, trois exemples qui démontrent les différentes tailles

3.5.2 Procédure supervisée

Cette section discute de l'approche de supervisée en utilisant les différents modes basés sur l'apprentissage profond examinés dans cette étude pour obtenir un diagnostic correct à partir des résultats obtenus résultant de la phase de détection comme suit:

- **CNN**

-Les réseaux neuronaux convolutionnels (CNN : Convolutional Neural Networks) sont l'une des variantes des réseaux neuronaux largement utilisés dans le domaine de la vision par ordinateur. Ils dérivent leur nom du type de calques cachés dont ils se composent. Les couches cachées d'un CNN sont généralement constituées de couches convolutionnelles, de couches de regroupement, de couches entièrement connectées et de couches de normalisation. Ici, cela signifie simplement qu'au lieu d'utiliser les fonctions d'activation normales définies ci-dessus, les fonctions de convolution et de regroupement sont utilisées comme fonctions d'activation. [1]

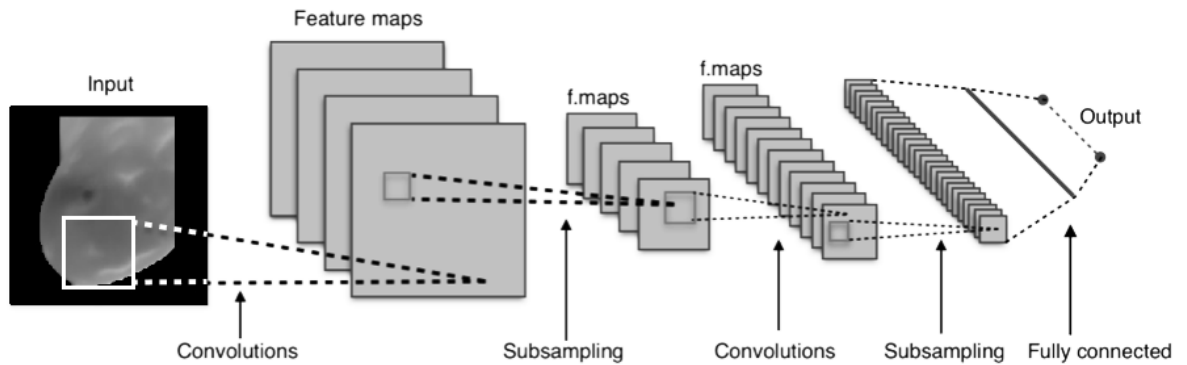


Figure 3.17 représentation du CNN [1]

Nous avons développé un modèle CNN de base suivant l'architecture de la figure [3.18]. Nous avons divisé notre base de données comme suit, 80% pour l'entraînement, 10% pour la validation et 10% pour l'ensemble des tests. Toutes les expériences avaient une normalisation par couches, fonction d'activation ReLU et nous avons ajouté des couches de dropout pour limiter le <<Overfitting>> puis nous avons testé plusieurs optimiseurs tels qu'Adam, RMSprop et SGD. Tout au long de les expériences mais le meilleur résultat qu'on a pu avoir était avec l'optimiseur Adam il a été testé plusieurs architectures variant le nombre de couches de au niveau de la convolution, le taux de décrochage et le nombre d'unités dans les dernières couches connectées. L'image d'entrée provient de l'algorithme du roi d'extraction de région d'intérêt et sa taille est de 250x250 «pixels». L'entraînement du system a été effectué dans le cadre des lots de formation mini-augmentés (batch size) (32 images augmentées par étape), avec 53 étapes par époque (53 évaluations de 32 instances, par epochs) et 40 époques au total. Nous avons fait de multiple test jusqu'à enfin trouver le meilleur modèle en fonction des mesures de performance et du temps d'exécution.

- **VGG16**

VGG16 est un modèle de réseau de neurones convolutionnel pré entraîné proposé par [35]. Son architecture est illustrée dans la figure ci-dessous :

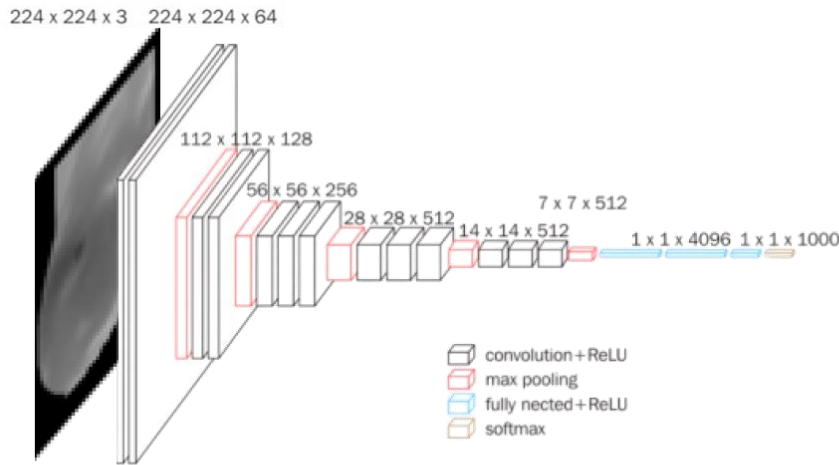


Figure 3.18 Architecture de VGG16

Nous avons utilisé VGG16 pour sa grande communauté dans le monde et car il existe beaucoup de système fait à partir de son model pré entrainé et aussi pour sa grande performance et ses résultat données dans les diverses système où il est utilisé. Ici nous avons ajouté une étape d'extraction des caractéristiques pour passer à l'entraînement de celui-là , on a pris comme taille d'images d'entré 224x224 «pixels» car c'est la taille standard et recommander pour travailler avec VGG16 après on a fait une augmentation de données <<Data Augmentation>> qui consiste a multiplié les images qu'on possède dans notre base de données en plusieurs autre mêmes images en appliquant certaines fonction, Après plusieurs essaye nous avons aboutie que le meilleur résultat donné était en appliquant un retournement horizontal des images et en faisant un zoom de 20% et comme sa on a pu multiplié les images de notre base de donné par 5 , puis on a fait ce qu'on appelle le fine tunning qui consiste a gelé les derniers couches connecté pour ensuite ajouté une nouvelle couche connecté qui vas résoudre notre problème car le VGG16 est déjà entrainé pour pouvoir faire une prédiction de 1000 classes et ici dans notre cas nous avons besoins de prédire que seulement deux classes (normal ou anormal),Pour cela et après de nombreux test nous avons choisie de travailler avec un batch size de 32 (images augmentées par étape) et avec l'optimiseur Adam et 120 epochs au total.

- **MobileNetV2**

MobileNetV2 est un modèle de réseau de neurones convolutionnel pré entrainé proposé par Google qui peut classifié et prédire jusqu'à 1000 classes ou objets et il a était consue

pour être utilisé sur des machines à puissance de calcul limitée comme les appareils mobiles [29]. Son architecture est illustrée dans la figure ci-dessous :

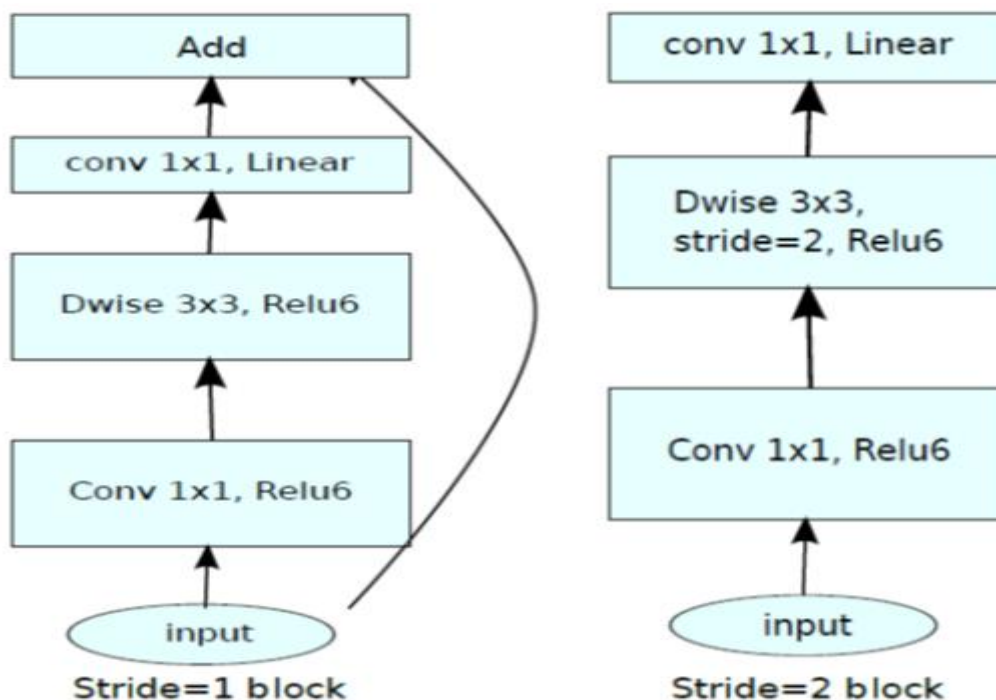


Figure 3.19 Architecture de MobileNetV2 [29]

Nous avons utilisé MobileNetV2 pour sa grande communauté dans le monde comme VGG16 et aussi pour sa grande performance et vitesse et ses résultats donnés dans tous les systèmes où il a été sollicité. On a pris comme taille d'images d'entrée 224x224 «pixels» à l'encontre de VGG16 où il a une taille d'images standards. MobileNetV2 a différentes tailles qu'on peut utiliser selon notre désir puis on a fait une augmentation des données comme pour VGG16 et avec les mêmes fonctions un retournement horizontal des images et en faisant un zoom de 20%. On est passé au fine tuning et cette fois après plusieurs tests on a choisi de travailler avec un batch size de 32 (images augmentées par étape) et avec l'optimiseur RMSprop et 200 époques au total pour au final avoir le meilleur résultat possible.

- **ResNet**

ResNet-50 est un réseau résiduel profond. Le «50» fait référence au nombre de couches qu'il a. Il s'agit d'une sous-classe de réseaux de neurones convolutifs, ResNet étant le plus couramment utilisé pour la classification d'images [34]. Son architecture est illustrée dans la figure ci-dessous

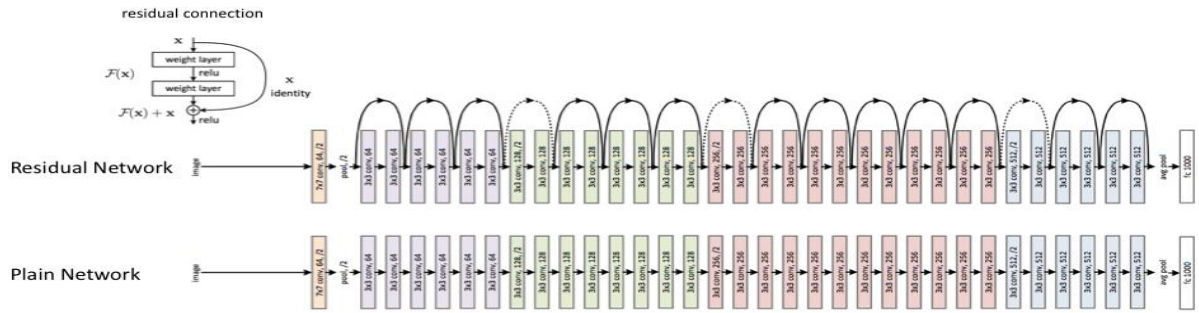


Figure 3.20 Architecture de ResNet50 [34]

Nous avons utilisé ResNet50 dans notre projet pour sa grande popularité dans le monde de la classification d'images et sa grande performance dans ce domaine-là. On a pris 299x99 «pixels» comme taille d'images d'entrée car c'est la taille standard et recommander pour travailler avec ResNet50 puis on a fait une extraction des caractéristiques avec ce même model la pour passer à la augmentation de données <<Data Augmentation>> après plusieurs essaye nous avons appliqué les mêmes fonctions qu'on a utilisé avec le VGG16 et MobileNetV2 qui sont un retournement horizontal des images et en faisant un zoom de 20%.Puis on est passé au fine tuning et cette fois après plusieurs teste on a choisie de travailler avec un batch size de 32 (images augmentées par étape) et avec l'optimiseur Adam et 120 epochs au total pour au final avoir le meilleur résultat possible.

- **Evaluation matrices [32]**

Dans un test de classification, nous classons les éléments selon deux classes : positive ou négative, le résultat peut être correcte ou incorrecte. Pour cela, nous distinguons quatre combinaisons de résultats possibles :

- Vrais positifs « TP» (True Positive) : Un résultat est dit vrai positif lorsqu'un élément positif est correctement classé.
- Faux positifs « FP» (False Positive) : Un résultat est dit faux positif lorsqu'un élément est classé positif, alors qu'il ne l'est pas.
- Vrais négatifs « TN» (True Negative) : Un résultat est dit vrai négatif lorsqu'un élément négatif est correctement classé.
- Faux négatifs « FN» (False Negative) : Un résultat est dit faux positif lorsqu'un élément est classé négatif, alors qu'il est positif.

Nous présentons dans la suite les mesures les plus utilisées dans le domaine de RI :

(a) La précision : La proportion de prédictions correctes parmi les points que l'on a prédits positifs

$$précision = \frac{True\ positive}{True\ positive + False\ positive} \quad (3.1)$$

(b) Le rappel : (sensibilité) C'est le taux de vrais positifs, c'est à dire la proportion de positifs que l'on a correctement identifiés.

$$rappel = \frac{True\ positive}{True\ positive + False\ négative} \quad (3.2)$$

(c) F1-score : Pour évaluer un compromis entre rappel et précision, on peut calculer la "F-mesure", qui est leur moyenne harmonique.

$$F1\text{-score} = 2 * \frac{Précision + Rappel}{Précision * Rappel} \quad (3.3)$$

- **Matrice de confusion [32]**

Chaque colonne de la matrice représente le nombre d'occurrences d'une classe estimée, tandis que chaque ligne représente le nombre d'occurrences d'une classe réelle (ou de référence). Un des intérêts de la matrice de confusion est qu'elle montre rapidement si un système de classification parvient à classifier correctement

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

Figure 3.21 Matrice de Confusion [32]

3.5.3 Procédure non supervisée

- **La recherche d'image par le contenu (CBIR)**

La recherche d'image par le contenu est une technique permettant de rechercher des images à partir de ses caractéristiques visuelles, c'est-à-dire induite de leurs pixels. Les images sont classiquement décrites comme rendant compte de leur texture, couleur, forme. Un cas typique d'utilisation est la recherche par l'exemple où l'on souhaite retrouver des images visuellement similaires à un exemple donné en requête. Dans notre travail on a utilisé le K-ppv [4].

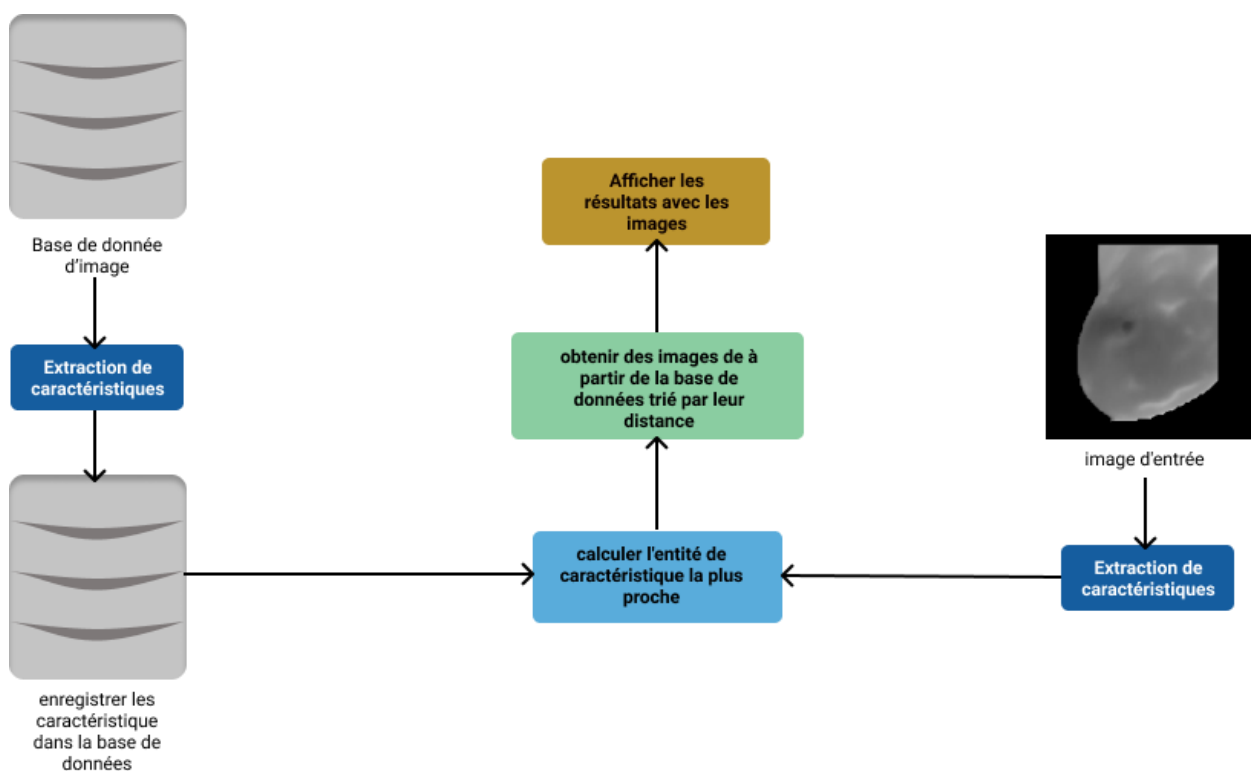


Figure 3.22 Diagramme explicative du CBIR

- **K-ppv**

L'algorithme k-voisins les plus proches (KNN) est un algorithme d'apprentissage automatique simple et facile à implémenter qui peut être utilisé pour résoudre des problèmes de classification et de régression. Il suppose la similitude entre le nouveau cas de données et les cas disponibles et place le nouveau cas dans la catégorie qui est la plus similaire aux catégories disponibles [17].

Tous d'abord dans notre travail et avant de mettre en œuvre l'algorithme du k-ppv on a du faire une extraction des caractéristiques on a essayé plusieurs model comme VGG16, MobileNetV2 et ResNet50 et enfin InceptionV3 mais le on a gardé que les modèles qui nous on donner les meilleurs résultat qui sont VGG16 et MobileNetV2 puis on est passé à l'exécution de l'algorithme du k-ppv on commençant par sélectionner le nombre K de voisins, Nous avons remarqué après plusieurs test que en utilisant K=5 jusqu'à k=17 la marge d'erreurs était réduite et plus proche de zéro mais on travailler et tester avec K=2, K=3 et K=5 puis pour calculer la distance des K nombre des voisins nous avons travaillé avec les distance suivante :

a) Euclidienne :

La distance euclidienne ou métrique euclidienne est le « ordinaire » linéaire la distance entre deux points dans l'espace euclidien. Sa formule est ci-dessous :

$$d(X, Y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (3.4)$$

b) Cosinus :

La similarité cosinus (ou mesure cosinus) permet de calculer la similarité entre deux vecteurs à n dimensions en déterminant le cosinus de l'angle entre eux. Sa formule est ci-dessous :

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|} \quad (3.5)$$

c) Minkowski :

La distance de Minkowski est une mesure dans un espace vectoriel normé qui peut être considéré comme une généralisation à la fois la distance euclidienne et la distance de Manhattan. Sa formule est ci-dessous :

$$D(X, Y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p} \quad (3.6)$$

Après avoir choisie le Nombre de K voisins et calculer la distances, on a lancé l'algorithme du K-ppv.

3.5.1 Evaluation métriques

La Précision moyenne ou Précision au sommet k.

3.6 Segmentation des couleurs:

La dernière phase, où l'algorithme utilise le HSV colorspace pour cette étape. La segmentation des couleurs dans une image thermique signifie une segmentation par température, du bleu foncé jusqu'au rouge clair (vers le blanc), la partie bleu signifie une basse température, la partie rouge signifie une température élevée.

- Cette fonction permet de vérifier si les éléments du tableau (d'image) se trouvent entre les éléments des deux autres tableaux.
- Le premier tableau entre les deux derniers tableaux contient les valeurs de premier seuil (valeur minimum) de couleur de la classe et le deuxième contient les valeurs du second seuil (valeur maximum).
- Le résultat sera une autre image (masque / mask). Figure [3.23], cette image contient que des pixels noirs et blancs, où les pixels noirs indiquent que sa couleur n'est pas entre les deux seuils, et les pixels blancs indiquent le contraire.

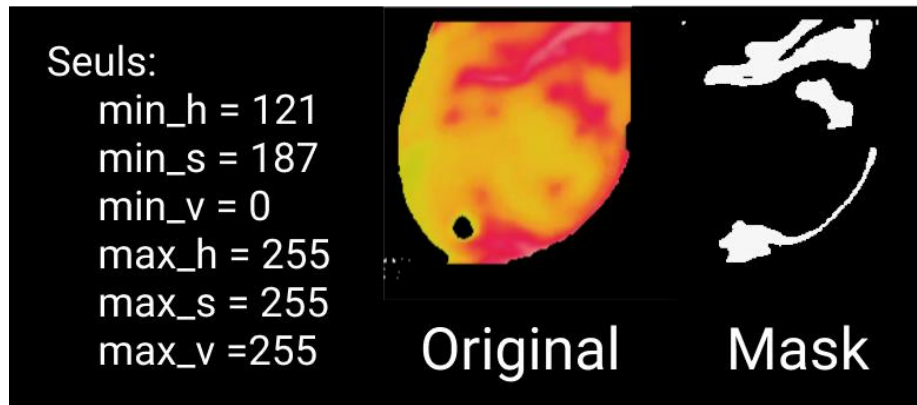


Figure 3.23 L'application de la méthode sur une image

La fonction vérifie la plage comme suit (DST, c'est la destination ou la sortie (output), il est défini sur 255 bit) :

- Pour chaque élément d'un tableau d'entrée à canal unique:
 $dst(I) = \text{lowerb}(I)_0 \leq src(I)_0 \leq \text{upperb}(I)_0$
- Pour les tableaux à deux canaux
 $dst(I) = \text{lowerb}(I)_0 \leq src(I)_0 \leq \text{upperb}(I)_0 \wedge \text{lowerb}(I)_1 \leq src(I)_1 \leq \text{upperb}(I)_1$

Et ainsi de suite pour plus de canaux.

Pour avoir juste la partie du masque en couleur, il suffit retourner la différence entre l'image original et le masque.

La figure 3.24 suivante indique le résultat de cette méthode sur deux masques (Chaud et Normal).

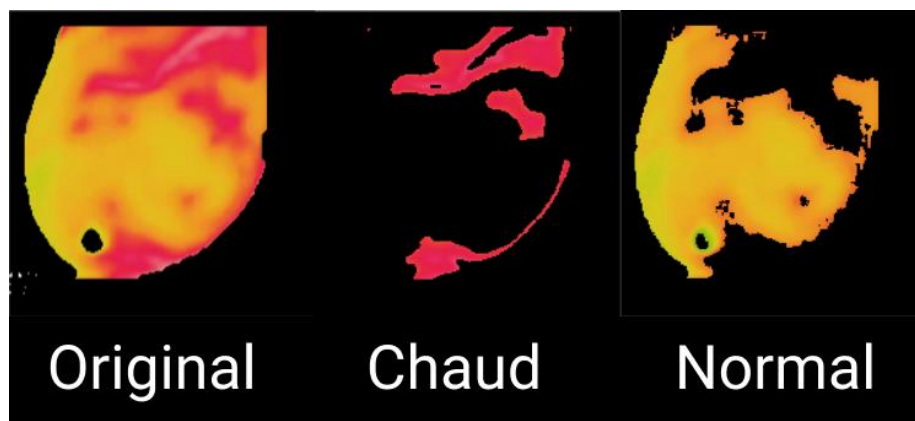


Figure 3.24, l'application de segmentation sur le ROI

La formule mathématique est comme suit :

$$dst(I) = src1(I) \wedge src2(I) \text{ if } mask(I) \neq 0$$

3.7 Conception de la plateforme web

3.7.1 Diagramme de classe

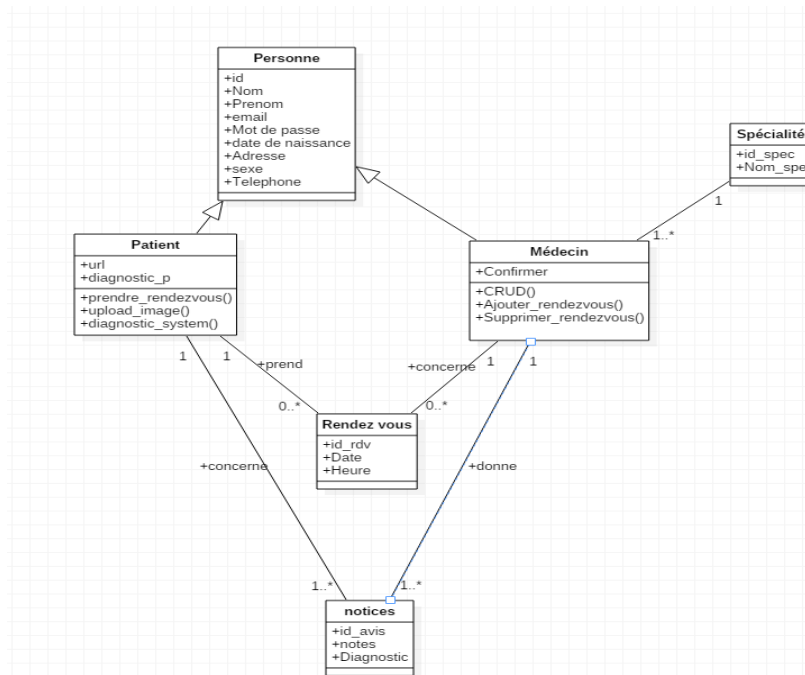


Figure 3.25 Diagramme de classe

Dans la figure 3.25 au-dessus on peut voir le diagramme de classe qu'on a conçu pour le système d'information de la plateforme et voici son dictionnaire en dessous dans la table 5.2.

- **Dictionnaire du diagramme de classe:**

classe	Attribute	signification
Personne	Id	Id de la personne
	Nom	Nom de la personne
	Prénom	Prénom de la personne
	Mot de Passe	Mot de passe de la personne
	Email	Email de la personne
	Adresse	Adresse de la personne
	Date de de naissance	Date de naissance de la personne
	Sexe	Sexe de la personne
	Téléphone	Téléphone de la personne
Rendez-vous	Id_rdv Date Heure	Id du rendez-vous La date du rendez vous Heure précise du rendez vous
Spécialité	Id_spec Nom_spec	Id de la spécialité Nom de la spécialité
Notices	Id_avis	Id de l'avis

	Notes diagnostic	Avis du médecin écrit Diagnostic donné par le médecin
Patient	url diagnostic_p	url de l'image diagnostic du patient donné par le systeme
Médecin	confirmer	Définie un médecin confirmer ou pas encore

Table 3.1 Dictionnaire du diagramme de classe

- **Model Relationel**

Patient (Id_patient, email_patient, mot de passe_patient, nom_patient, prenom_patient, date_naissance_patient, adresse_patient, telephone, sexe_patient, date de naissance, diagnostic_p,url)

Médecin (Id_médecin, email_médecin, mot_de_passe_médecin, nom_médecin, prenom_médecin, date_naissance_médecin, adresse_médecin, telephone_médecin, sexe_médecin,confirmer,#id_spécialité)

Rendez-vous (Id_rdv, date, heure, # Id_patient, # Id_médecin)

Spécialité (id_spécialité, nom_spécialité)

Avis (id_avis, notes, #id_médecin, #id_patient, diagnostic)

5.3.2 Diagramme de cas d'utilisation général

Afin d'illustrer les fonctionnalités de notre plateforme, la figure 3.26 illustre le diagramme de cas d'utilisation général de la Plateforme qui présente les acteurs et leurs fonctions

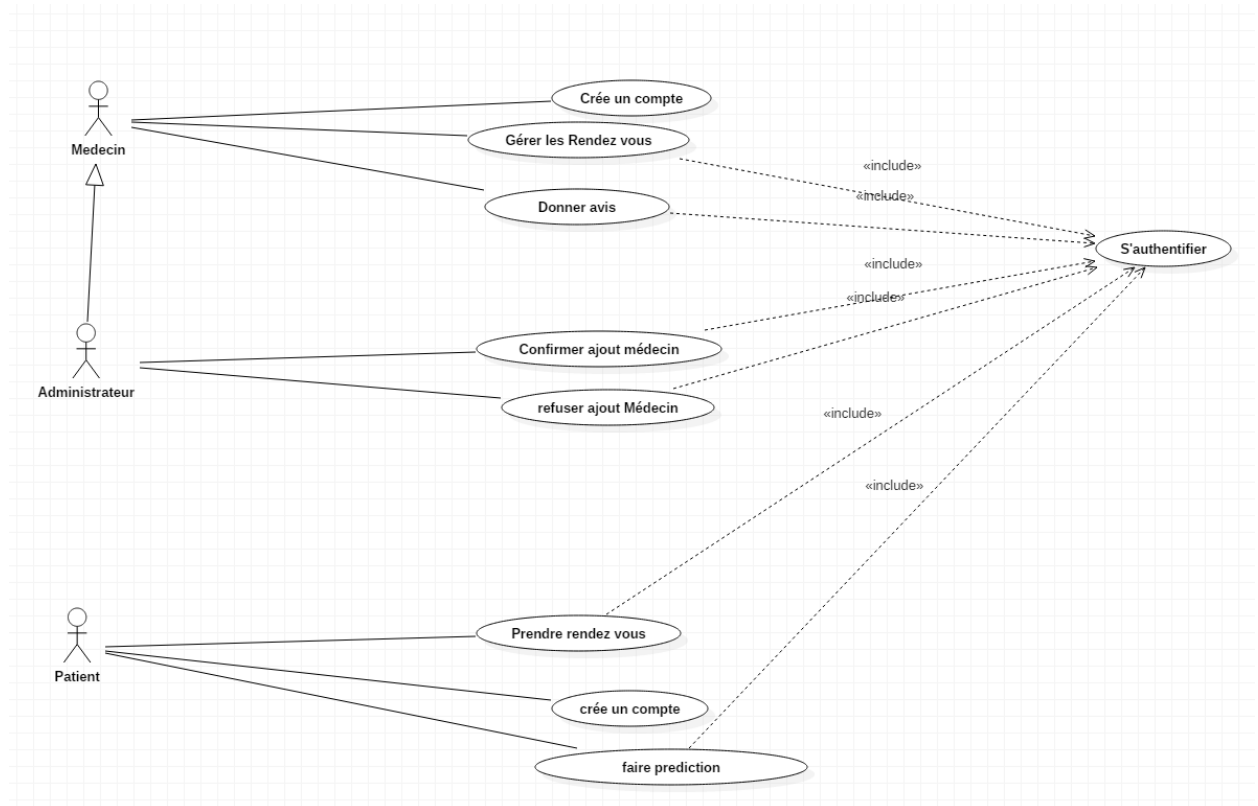


Figure 3.26 Diagramme de cas d'utilisation

5.3.3 Diagramme de cas d'utilisation patient

Dans la figure 3.27 présente le diagramme de cas d'utilisation du patient et ses différentes fonctions.

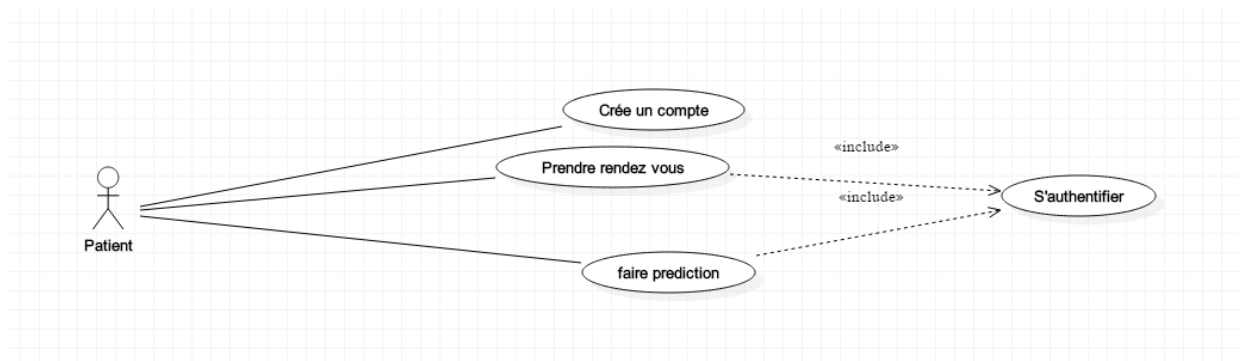


Figure 3.27 Diagramme de cas d'utilisation patient

5.3.4 Diagramme de cas d'utilisation médecin

Dans la figure 3.28 on peut voir le diagramme de cas d'utilisation du médecin et ses différentes fonctions.

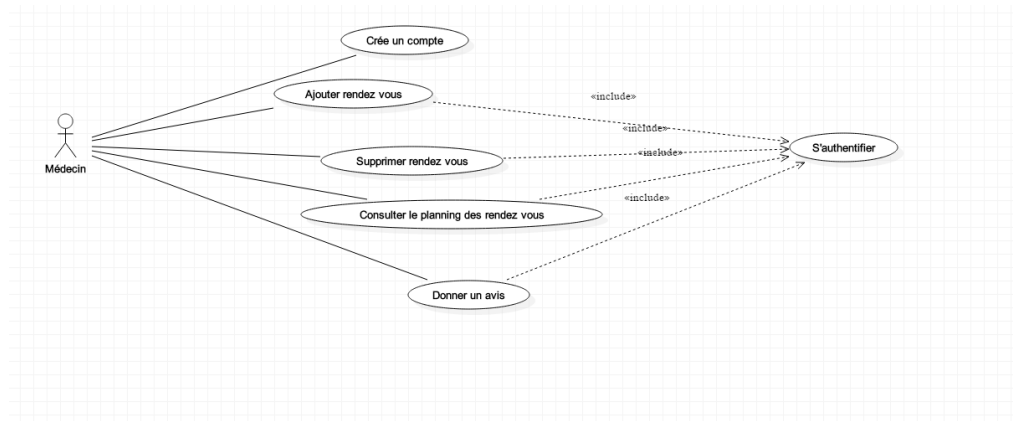


Figure 3.28 Diagramme de cas d'utilisation médecin

5.3.5 Diagramme de cas d'utilisation administrateur

Dans la figure 3.29 on peut voir le diagramme de cas d'utilisation de l'administrateur et ses différentes fonctions.

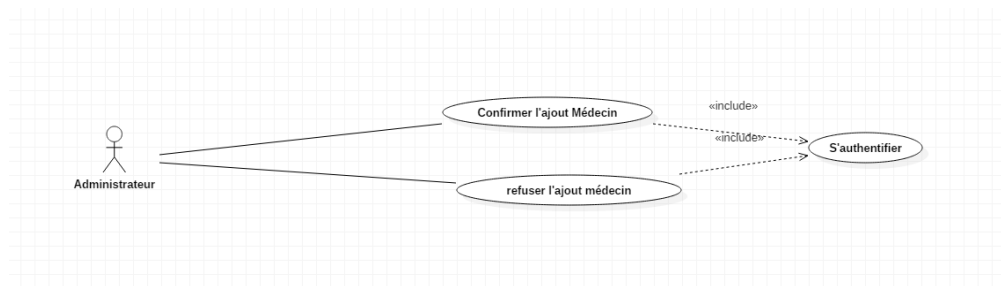


Figure 3.29 Diagramme de cas d'utilisation administrateur

3.6 Conclusion

Au cours de ce chapitre nous avons présenté les différentes solutions proposées pour l'extraction de la zone d'intérêt (ROI) avec ses différentes méthodes ainsi que les solutions proposées pour le diagnostic avec la classification à l'aide de l'apprentissage supervisé et le calcul de similarité avec l'apprentissage non supervisé. Dans le prochain chapitre nous présenteront les résultats obtenus.

Chapitre 4 RÉSULTATS ET DISCUSSION

4.1 Introduction

Dans ce chapitre, nous allons présenter tous les résultats obtenus par les différents algorithmes pour l'extraction de la zone d'intérêt (ROI) ainsi que les résultats obtenus pour le diagnostic, à la fin nous allons effectuer une comparaison entre ces algorithmes pour étudier leur efficacité.

4.2 Résultats de segmentation

Les différentes méthodes de segmentations ont donné des résultats avec des avantages et des inconvénients. Pour cela et pour chaque'une d'elles, on a créé des fichiers csv pour stocké les positions du sein après qu'on a lancé les méthodes (Adp-HT) et (Adp-BHT) en boucle sur 230 images.

Les résultats des trois méthodes sont présentés si-dessous :

4.2.1 Résultats utilisant le processus manuel

Pour la méthode manuelle, les résultats du ROI sont les meilleurs au niveau de la précision, car l'utilisateur lui-même clique sur le bon positionnement du sein, mais il reste un peu lourd à manipuler et à chaque fois l'utilisateur clique et confirme le stockage où réinitialise l'image s'il y a trop de clique et trop de cercle sur l'image.

4.2.2 Résultats utilisant l'algorithme adaptatif 1

- **Résultat du Adp-HT:**

La première méthode est bonne avec les femmes maigre et normale, aussi avec les images aux seins sont symétriques, facile a détecté, mais elle a du problème avec les femmes obèse (une personne obese implique plus d'affaissement du corps et donc plus de faux cercle détecté.

La figure[4.1] , illustre les différent résultats avec cette methode:

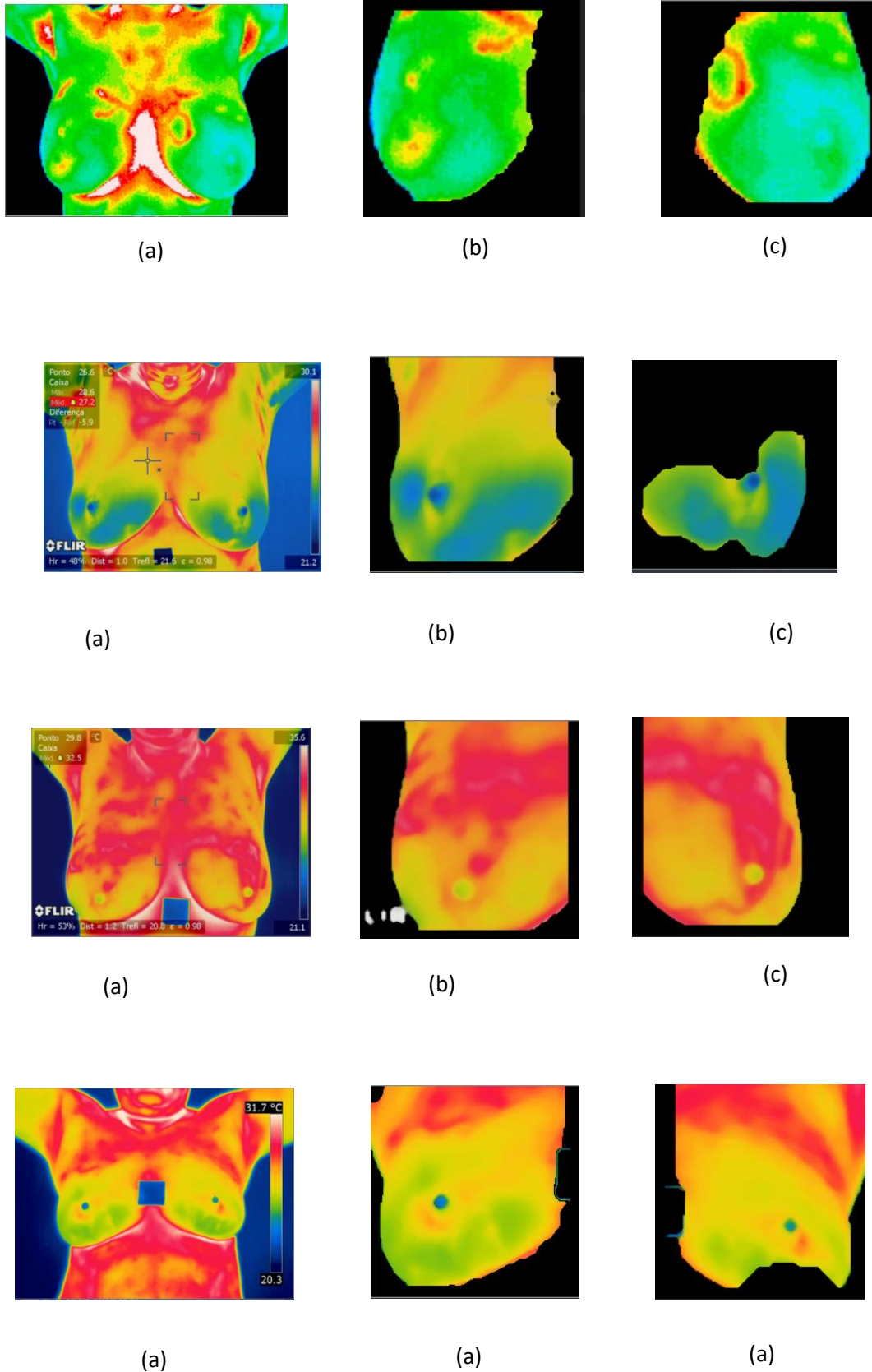


Figure 4.1, résultat de la Adp-HT cas corps normale, (a) l'image original, (b) ROI gauche, (c) ROI droite

4.2.3 Résultats utilisant l'algorithme adaptatif 2

- Résultat du Adp-BHT :

La méthode (Adp-BHT), est compatible avec tous les type d'images et tous les type du corps grâce au traitements supplémentaires, et qu'elle prendre que la moitié du corps (droite / gauche) la méthode a réussi à détecter le ROI dans la majorité des images même avec un corps asymétrique ou avec une mauvaise qualité.

La figure[4.5] , illustre les différente résultats avec cette méthode:

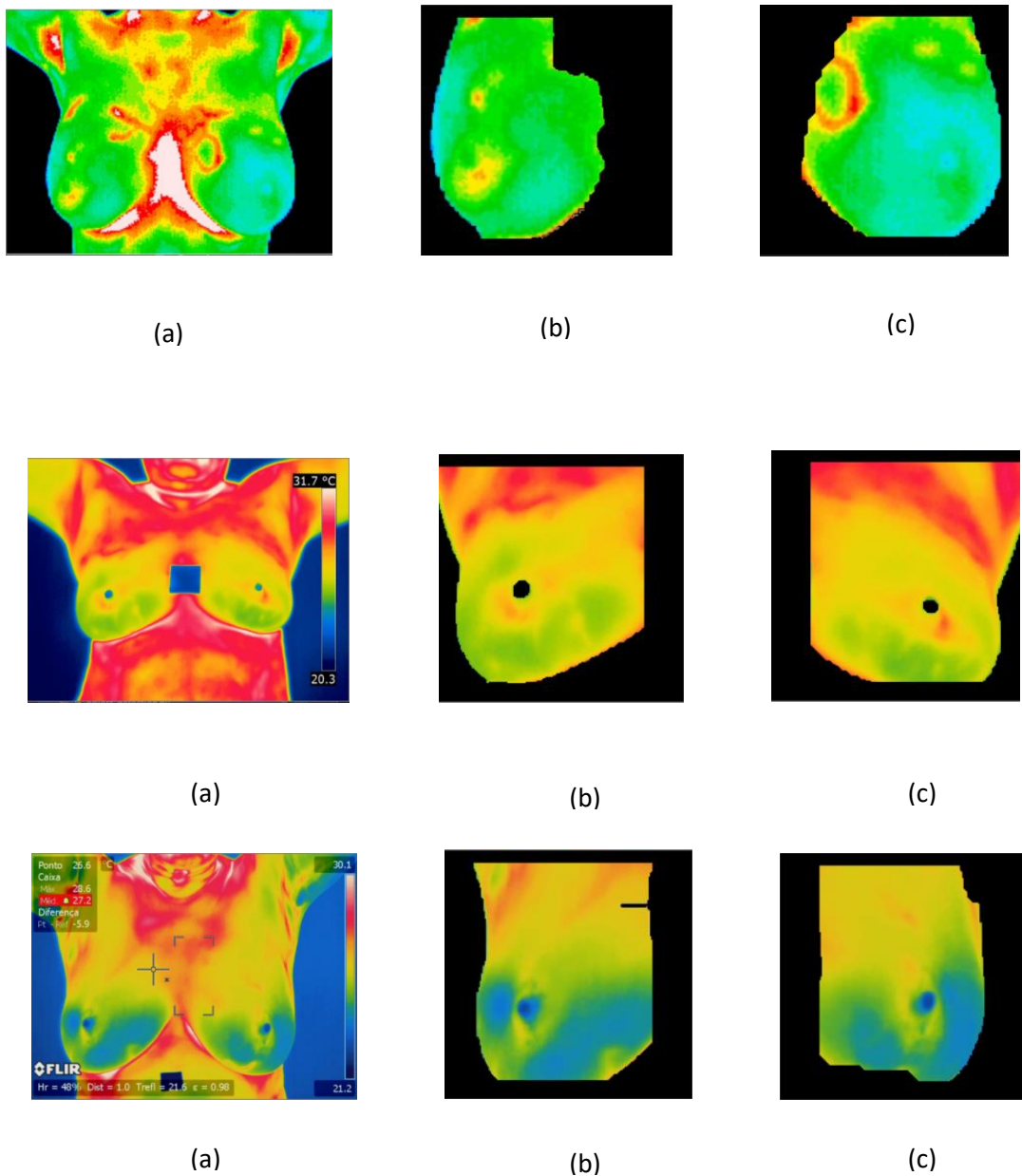


Figure 4.5, résultat de la Adp-BHT cas corps normale, (a) l'image original, (b) ROI gauche, (c) ROI droite

4.2.4 Comparaison entre les algorithmes :

Le but de cette recherche est de construire un système automatique, donc la méthode manuelle n'est pas favorable malgré que cette méthode à la meilleure précision entre les trois méthodes, sans cité qu'elle n'est pas vraiment pratique. Comme il arrive que l'utilisateur tombe dans un cas où la machine ne peut pas déterminer le bon positionnement du ROI, dans ce cas la méthode manuelle aide beaucoup pour dépasser ces problèmes.

Aussi comme elle est optimale à la précision, nous avons utilisé ses résultats comme base de comparaison entre les deux autres méthodes (Adp-HT et Adp-BHT).

Comme le nombre d'images est trop grand, on ne peut pas visualiser toutes les images et choisir une des deux méthodes pour notre projet. C'est pour cette raison que les deux méthodes ont la possibilité de stocker les coordonnées du ROI de chaque image dans un fichiers CSV, est avec les calculs de distance euclidienne entre les CSV de Adp-HT et Adp-BHT avec le fichier CSV de la méthode manuelle, pour chaque image on prend la petite valeur entre eux, et on augmente le score de Adp-HT et Adp-BHT initialisé au départ à zéro, on obtient à la fin les deux valeurs et grâce à ces deux valeurs on a choisi la méthode adaptée par notre système (Adp-BHT) comme il est montré dans le tbaleau [4.1]. Le programme développé qui calcule la distance euclidienne entre le ROI de la méthode manuelle et les deux autres méthodes adaptatives est basé sur l'utilisation de 230 images (460 ROI).

Algorithmes	Nombre de Distances
Adp-HT	203
Adp-BHT	210

Egalité	47
Total	460

Table 4.1 Comparaison entre les deux algorithmes du ROI

4.3 Résultat de la méthode supervisée

Nous allons montrer tous le résultat obtenue par toutes les méthodes proposées , nous précisons que ces tests ont étaient effectué sur 10% de la base de données.

4.3.1 Resultat du CNN

Le tableau ci-dessous présente les résultats obtenus après l'exécution du CNN, on peut voir qu'il donne assez bon résultats avec un F1-score de 83.23%.

Modèle	Précision	Rappel	F1score
CNN	92.56%	76.12%	83.23%

Table 4.2 Resultats du CNN.

Voici la matrice de confusion du CNN

(Anormalité=0, Normal=1)

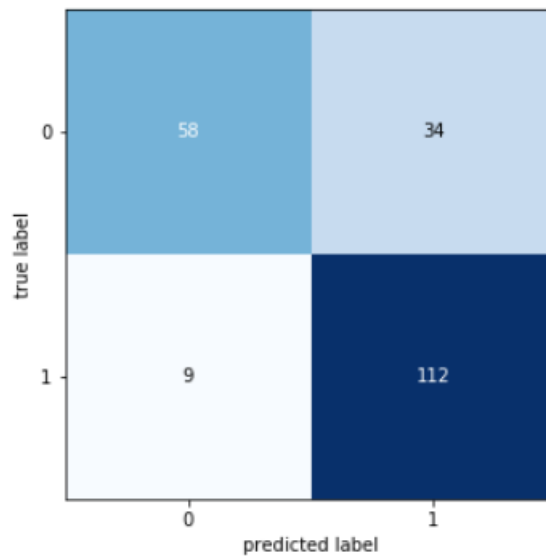


Figure 4.2 Matrice de confusion du CNN

On peut voir qu'il y a eu confusion entre les deux classes normal et anormal une fois par mal classée 9 images et l'autre fois 34 images

4.3.2 Results from VGG16

Dans le tableau ci-dessous , nous présentons les résultats obtenue après l'exécution de VGG16 on peut voir qu'il donne d'assez bon résultats avec un F1-score de 92.08%.

Modèle	Précision	Rappel	F1score
VGG16	96.69%	90%	92.08%

Table 4.3 Résultat du Modelé VGG16.

Voici la matrice de confusion de VGG16

(Anormalité=0, Normal=1)

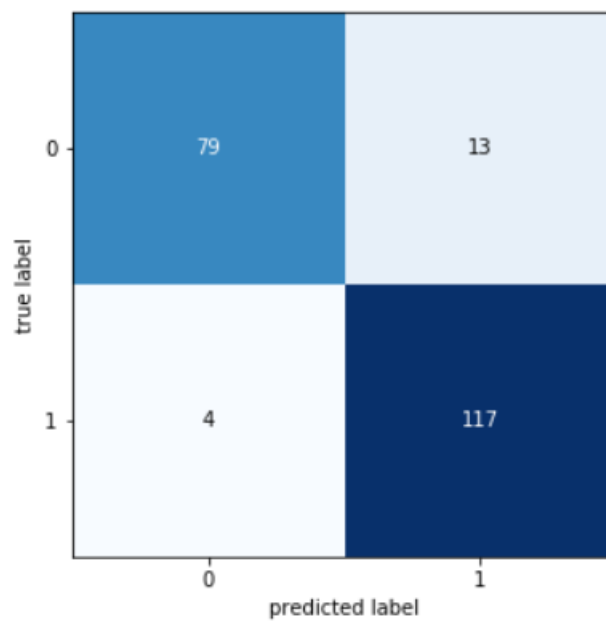


Figure 4.3 Matrice de confusion de VGG16

On peut voir qu'il y a eu confusion entre les deux classes normal et anormal une fois par mal classée 13 images et l'autre fois 4 images

4.3.3 Resultat du MobilNetV2

Dans le tableau ci-dessous vous pouvez voir les résultats obtenue après l'exécution de MobileNetV2 on peut voir qu'il donne d'assez bon résultats avec un F1-score de 94.90%.

Modèle	Précision	Rappel	F1score
MobileNetV2	98.34%	92.96%	94.90%

Table 4.4 Résultats du modelé MobileNetV2

Voici la matrice de confusion de VGG16

(Anormalité=0, Normal=1)

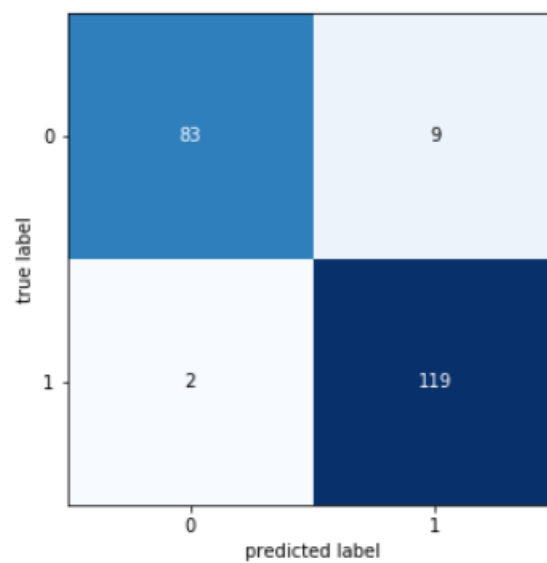


Table 4.5 Matrice de confusion de MobileNetV2

On peut voir qu'il y a eu confusion deux fois la première pour mal classée 9 images et la deuxième fois pour mal classée 2 images.

4.3.4 Resultat du ResNet50

Dans le tableau ci-dessous vous pouvez voir les résultats obtenue après l'exécution de ResNet50 on peut voir qu'il donne d'assez bon résultats avec un F1-score de 87.48%.

Modèle	Précision	Rappel	F1score
Resnet50	98.43%	79.86%	87.48%

Table 4.6 Résultats du modèle ResNet50.

Voici la matrice de confusion de ResNet50

(Anormalité=0, Normal=1)

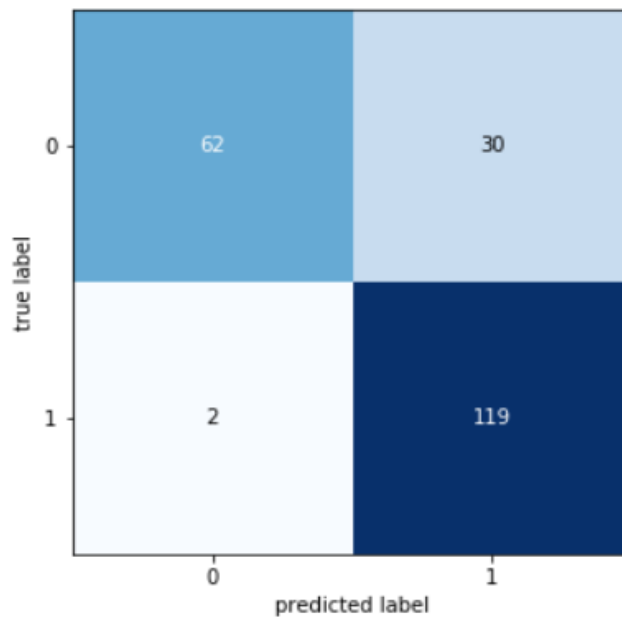


Figure 4.4 Matrice de confusion de ResNet50

On peut voir qu'il y a eu confusion deux fois la première pour mal classée 30 images et la deuxième fois pour mal classée 2 images.

4.4.5 Comparaison entre les méthodes supervisées

Voici dans le tableau ci-dessous une comparaison entre les quatres méthodes supervisée.

Modèle	Précision	Rappel	F1score
MobileNetV2	98.34%	92.96%	94.90%
VGG16	96.69%	90%	92.08%
Resnet50	98.43%	79.86%	87.48%
CNN	92.56%	76.12%	83.23%

Table 4.7 Résultats de tous les modèles.

On peut voir dans le tableau en dessus que le MobileNetV2 à donner de meilleurs résultats que tous les autres avec son F1-score de 94.90% malgré que même les autres méthodes on donnait de très bon résultats.

4.5 Résultat de la méthode non supervisée

Nous allons montrer tous les résultats obtenus par toutes les méthodes proposées, nous précisons que les tests ont étaient effectués sur et 10% de la base de données.

4.51 VGG16+K-ppv

Voici les résultats obtenue pour K=2 :

Modèle	Précision	Rappel	F1score
VGG16+ K-ppv	95.04%	77.18%	85.01%

Table 4.8 Résultats du VGG16+K-ppv (K=2).

On peut voir que VGG16+k-ppv pour K=2 nous donne un F1-score de 85.01% qui est assez bon.

Voici la matrice de confusion de VGG16+K-ppv

(Anormalité=0, Normal=1)

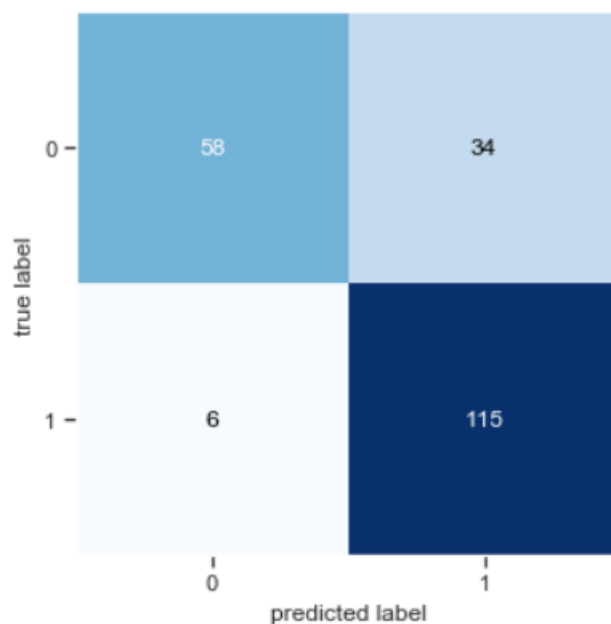


Figure 4.5 Matrice de confusion de VGG16+K-ppv (K=2)

Nous remarquons que il y a eu confusion une fois pour mal classé 6 images et l'autre fois pour mal classé 34 images.

Voici les résultats obtenue pour K=3 :

Modèle	Précision	Rappel	F1score
VGG16+ K-ppv	99.17%	75.74%	85.34%

Table 4.9 Résultats du VGG16+K-ppv (K=3).

On peut voir que VGG16+k-ppv pour K=3 nous donne un F1-score de 85.34% qui est assez bon et qui surpasse de peu celui de K=2.

Voici la matrice de confusion pour K=3

(Anormalité=0, Normal=1)

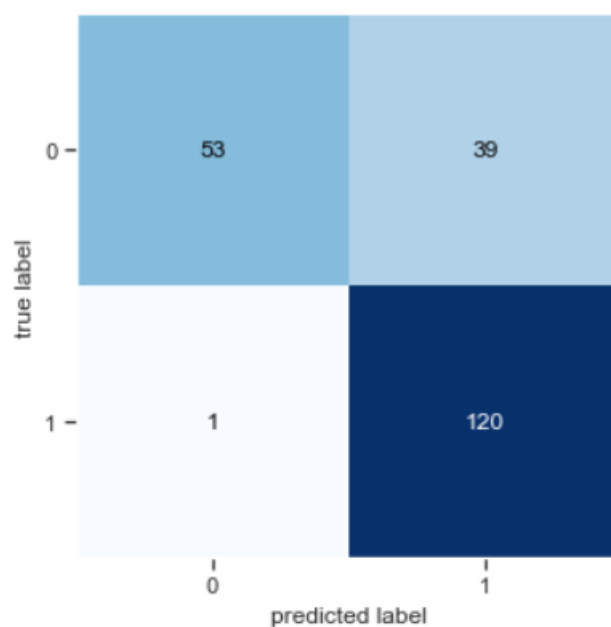


Figure 4.6 Matrice de confusion pour VGG16+K-ppv (K=3)

Nous remarquons que il y a eu confusion une fois pour mal classé une image et l'autre fois pour mal classé 39 images.

Voici les résultats obtenue pour K=5 :

Modèle		Précision	Rappel	F1score
VGG16+	K-	100%	76.58%	86.36%
ppv				

Table 4.10 Résultats du VGG16+K-ppv (K=5).

On peut voir que VGG16+k-ppv pour K=5 nous donne un F1-score de 86.36% qui est très bon et qui surpasse celui de K=2 et de K=3.

Voici la matrice de confusion pour K=5

(Anormalité=0, Normal=1)

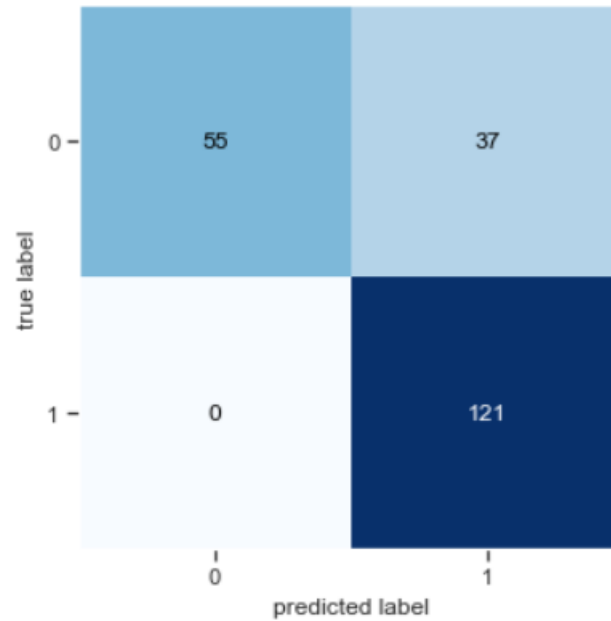


Figure 4.7 Matrice de Confusion pour VGG16+K-ppv (K=5)

Nous remarquons que il y a eu confusion une pour mal classé 37 images.

4.5.2 MobileNetV2+K-ppv

Voici les résultats obtenue pour K=2 :

Modèle	Précision	Rappel	F1score
MobileNetV2+K-ppv	90.90%	81.48%	85.26%

Table 4.11 Résultats du MobileNetV2+K-ppv (K=2).

On peut voir que MobileNetV2+k-ppv pour K=2 nous donne un F1-score de 85.26% qui est assez bon.

Voici la matrice de confusion de MobileNetV2+K-ppv

(Anormalité=0, Normal=1)

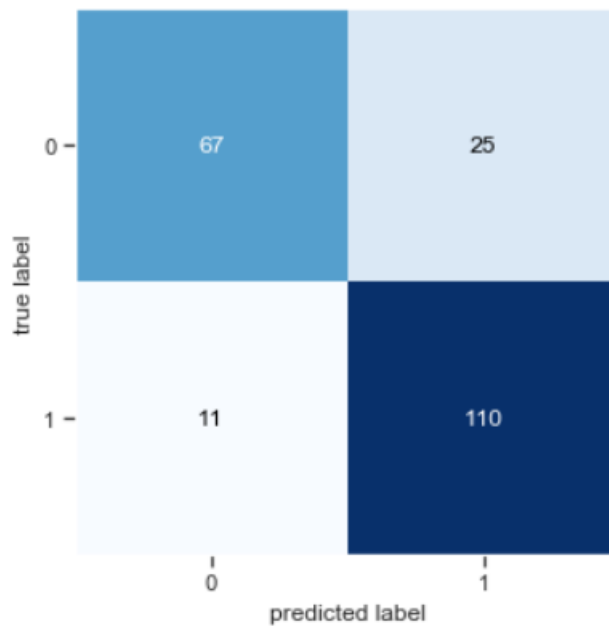


Figure 4.8 Matrice de confusion de MobileNetV2+K-ppv (K=2)

Nous remarquons que il y a eu confusion une fois pour mal classé 11 images et l'autre fois pour mal classé 25 images.

Voici les résultats pour K=3

Modèle	Précision	Rappel	F1score
MobileNetV2+K-ppv	95.08%	76.82%	84.84%

Table 4.12 Résultats du MobileNetV2+K-ppv (K=3).

On remarque que pour K=3 nous donne un F1-score de 84.84% qui est assez bon mais qui est inférieur à celui de K=2.

Voici la matrice de confusion pour K=3

(Anormalité=0, Normal=1)

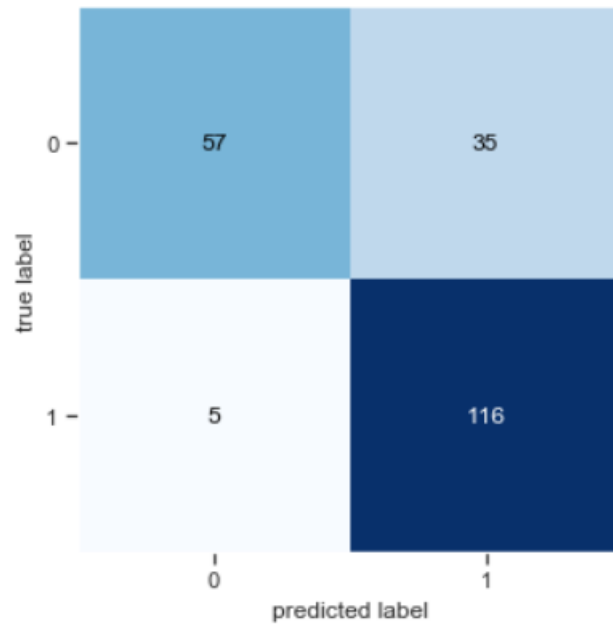


Figure 4.9 Matrice de confusion de MobileNetV2+K-ppv (K=3)

Nous remarquons que il y a eu confusion une fois pour mal classé 5 images et l'autre fois pour mal classé 35 images.

Voici les résultats pour K=5 :

Modèle	Précision	Rappel	F1score
MobileNetV2+K-ppv	100%	76.58%	86.36%

Table 4.13 Résultats du MobileNetV2+K-ppv (K=5).

On remarque que pour K=5 nous donne un F1-score de 86.36% est très bon et qui est supérieur à celui de K=2 et K=3.

Voici la matrice de confusion de K=5
(Anormalité=0, Normal=1)

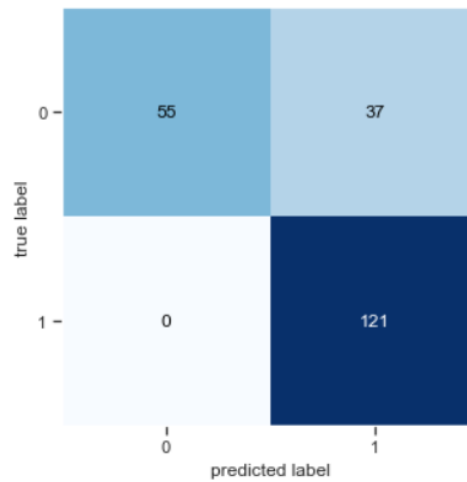


Figure 4.10 Matrice de confusion de MobileNetV2+K-ppv (K=5)

Nous remarquons que il y a eu confusion une pour mal classé 37 images.

4.5.3 Comparaison entre les méthodes non supervisées

Voici dans le tableau ci-dessous une comparaison entre les méthodes non supervisées:

Modèle	Précision	Rappel	F1 score
MobileNetV2+K-ppv (K=5)	100%	76.58%	86.36%
VGG16+ K-ppv (K=5)	100%	76.58%	86.36%
VGG16+ K-ppv (K=3)	99.17%	75.74%	85.34%
MobileNetV2+K-ppv (K=2)	90.90%	81.48%	85.26%
VGG16+ K-ppv (K=2)	95.04%	77.18%	85.01%
MobileNetV2+K-ppv (K=3)	95.08%	76.82%	84.84%

Table 4.14 Comparaison entre tous les Résultats obtenue

On peut remarquer dans le tableau au-dessus que MobileNetV2+K-ppv (K=5) et VGG16+K-ppv (K=5) ont eu exactement le même F1-score donc tous les deux sont les meilleurs mais tous les autres ont eu aussi des assez résultats qui on était très proche l'un de l'autre

4.6 Comparaison entre méthodes supervisé et non supervisé

Voici dans le tableau ci-dessous une comparaison entre la méthode d'apprentissage supervisé et celle non supervisée :

Modèle	Précision	Rappel	F1score
MobileNetV2	98.34%	92.96%	94.90%
VGG16	96.69%	90%	92.08%
Resnet50	98.43%	79.86%	87.48%
MobileNetV2+K-ppv (K=5)	100%	76.58%	86.36%
VGG16+ K-ppv (K=5)	100%	76.58%	86.36%
VGG16+ K-ppv (K=3)	99.17%	75.74%	85.34%
MobileNetV2+K-ppv (K=2)	90.90%	81.48%	85.26%
VGG16+ K-ppv (K=2)	95.04%	77.18%	85.01%
MobileNetV2+K-ppv (K=3)	95.08%	76.82%	84.84%
CNN	92.56%	76.12%	83.23%

Table 4.15 Comparaison entre les résultats de l'apprentissage supervisé et non supervisé

On peut remarquer que les résultats de l'apprentissage supervisé ont été nettement meilleurs que ceux de l'apprentissage non supervisé sauf au niveau du CNN qui a été le plus mauvais d'entre tous. MobileNetV2 qui a donné le meilleur résultat obtenu avec un F-score de 94.90%.

4.7 Résultat visuel du CBIR

On va vous présenter les résultats visuels obtenus par l'application de l'apprentissage non supervisé

4.7.1 Distance Euclidienne








Image	Distance
	0
	0.8398643467100408
	0.8532702761193823
	0.8610130831382509
	0.8864670826321336
	0.8961630723091719
	0.9068880270592262

Table 4.16 résultat visuel de la distance euclidienne

4.7.2 Distance Euclidienne avec SKLearn








Image	Distance
	0
	1.1696007373734947e-16
	1.586492155888884e-09
	1.8035696816197302e-14
	2.511771476344479e-18
	3.6234414410496463e-20
	6.146814618858573e-11

Table 4.17 résultat visuel de la distance euclidienne avec SKLearn

4.7.3 Distance MSE








Image	Distance
	0
	10293.896483333334
	11752.550991666667
	14014.379183333333
	3878.753016666666
	6060.192733333333
	6264.24175

Table 4.18 résultat visuel de la distance MSE

4.7.4 Distance SSIM








Image	Distance
	1
	0.7885991266707788
	0.6962291258958772
	0.6895236681319875
	0.5157994682087458
	0.5013116161694309
	0.4756108840468693

Table 4.19 résultat visuel de la distance SSIM

4.8 Conclusion

Dans ce chapitre nous vous avons montré tous les résultats obtenus par les trois algorithmes d'extraction de la zone d'intérêt (manuel, Adp-HT et Adp-BHT) et puis tous les résultats obtenus pour la classification et la similarité par les algorithmes d'apprentissage supervisé et non supervisé en utilisant les mesures de performance (précision, rappel, f1-score). dans le prochain chapitre nous aborderons les différents outils et matériels utilisés pour la conception et le développement de la Platform web.

Chapitre 5 Implémentation de la plateforme WebCAD

1.1 Introduction

Dans ce chapitre, nous allons présenter l'application web développée avec toutes les interactions possibles ainsi que les Ressources matérielles et logicielles utilisées lors du développement de ce projet.

1.2 Schéma fonctionnel de la plate-forme WebCAD

L'architecture de la plateforme développées est présentée dans la figure ci-dessous:

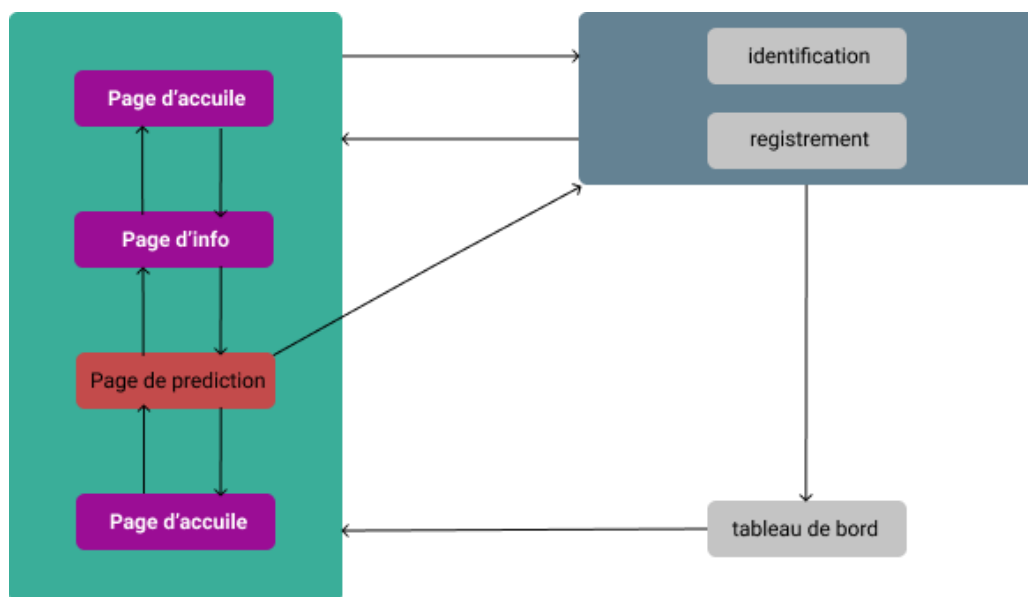


Figure 5.1 schéma fonctionnel du system webCAD

L'utilisateur peut accéder et se déplacer entre les différentes pages sans besoins d'identifications. Il peut s'identifier de n'importe quelle page, dans le cas où l'utilisateur veut faire un diagnostic il faut qu'il s'identifie (faire un registrement si c'est la première fois dans le site et puis une identification). Pour accéder au tableau de bord, il faut avoir l'autorisation (une identification de type ADMIN).

5.3 Ressources matérielles et logicielles utilisées lors du développement:

Afin de réaliser ce projet de recherche, nous avons utilisé les ressources suivantes:

Ressources Matérielles:

Dans le développement de site web et l'algorithme de ROI, nous avons travaillé sur un ordinateur personnel (HP envy-15) avec une configuration de 12Go de RAM, un processeur intel (I7) 6eme generation, et windows 10 comme un système d'exploitation.

Outils/matérielles	Fonctions
Google colabraty	Dans l'apprentissage, on n'utilise google colab pour accélérer les test, il forne 20Go de RAM et deferentes type de processus graphique (CPU, GPU, TPU).
Python	Est le langage principale de développement, c'est un langage fort et modern, qui faurné plusieurs fonctionnalités et bibliothèques aident dans notre developement.
opencCV	Une bibliothèques de python, spécialisé dans le domaine de traitement d'images, nous avons l'utilisé dans le notre pre-traitment, et la detection de ROI.
Tensorflow	TensorFlow est une plateforme open source de bout en bout pour l'apprentissage automatique. Il nous aide a créer et de déployer facilement des applications propulsées par machine/deep learning.
Keras	Keras est une API d'apprentissage profond écrite en Python, fonctionnant au-dessus de la plate-forme d'apprentissage automatique TensorFlow. Nous avons l'utilisé grace a son prototypage simple et rapide, il prendre en charge les notre réseau de norone convolutionnel

Flask	Flask est un Mini Framework du langage pythons, nous l'avons utilisé pour développer le côté back-end (non-visuel) de ce site web puisque il est conçu pour faciliter et accélérer la mise en route des pages.
HTML	Un langage de balisage, utilise pour restructuré, structurer et de mettre en forme le contenu de nos pages web.
CSS	Les feuilles de style, et un langage informatique de stylage qui décrit la présentation des pages html /xml, et il est utilisé dans la conception de notre pages web et le rendre plus beau, pour donner une bonne expérience aux utilisateurs.
jQuery	une bibliothèque de JavaScript très forte, il rend le code JavaScript plus facile, toutes les animations dans notre site sont développées grâce à cette bibliothèque.
Bootstrap	Le Framework Bootstrap, est une collection d'outils qui facilite la création du design pour les pages web. Il contient des codes HTML, CSS (Formulaire, bouton, des menus et autres éléments de page web) et JS. Donc grâce à ce framework on a facilité et accéléré les tâches de développement des pages web côté utilisateur.
Ajax	Est une architecture informatique qui nous a permis de construire notre site web dynamique interactif, les requêtes et les mises à jour se font sans recharger la page web.
SQLite	Pour gérer les données des utilisateurs, nous avons choisi SQLite, une base de données accessible en langage SQL, elle est intégrée directement dans le projet donc on n'a pas besoin d'un mini serveur.

Table 5.1 Outils utilisés

5.4 Système développé :



Figure 5.2, logo de solution

La figure 5.2 montre le logo qu'on a conçu et qu'on appelé <<PinkSfifa>> qui veut dire rubon rose quand on traduit les deux mots en français, On a choisi la couleur rose et le rubbon en relation avec le octobre rose qui est le mois conçacrait a la maladie du cancer du sein.


[Home](#)
[Prediction](#)
[Datasets](#)
[About](#)
[Contacts](#)

[SignIn](#)
[SignUp](#)

PinkSfifa Diagnostic


The best and the fastest way to detect or diagnose a breast abnormality
Also the first pure algerian Cad on the web.

[Get Started !](#)



**Breast
CANCER
AWARENESS**

THERMOGRAPHY



Thermography


thermography detects temperature differences on the surface of the body.

This property could be used to detect breast cancer.

Why choosing thermography instead of mammography ?

Because its painless and more reliable.


CONTRIBUTOR



AOUACHE
Mustapha

Senior Researcher at Center for Development of Advanced Technologies (CDTA)
Doctor of Philosophy (PhD)
Domaine d'études Electrical, Electronic & Systems Engineering


[f](#) [e](#) [in](#)



CHAHRAAT
Djamel Eddine

Master 2 computer science student specializing in IT at Blida University and young ambitious and talented developer


[f](#) [e](#) [in](#)



NADIR
Seddik

Master 2 computer science student specializing in IT at Blida University and young ambitious and talented developer

[f](#) [e](#) [in](#)




GUESSOUM
Dalila


Assistant professor (A). Researcher at the university of BLIDA 1 since 2012. And she has participated in numerous international conferences

[f](#) [e](#) [in](#)


ABOUT




1. Taking a thermal photo and putting it on the site




4. Exit with two images of the right and left breast




2. Image resizing




5. Features extraction




3. extraction of the region of interest



6. Classification



CONTACTS



Your Name *

Your Email *

Your Phone Number *

[Send Message](#)

Your Message *

+213-XXX-XXX
City

[f](#) [e](#) [in](#)
 © 2020 Copyright: PinkSfifa

Figure 5.3, page d'accueil

81

Dans la figure 5.3 au-dessus on peut voir la page d'accueil du site ou on y trouve en haut la barre du menu et il y a quelques information au sujet de la thermographie, En haut à droite on trouve un bouton pour se connecter si on a déjà un compte et un autre bouton pour créer un compte on peut pas accéder au bouton prédiction et faire une prediction, Il faut d'abord avoir un compte de patient et être connecté, On y trouve aussi des information sur les contributeur a la création du plateforme ou en peut accéder à leur compte sur les réseaux sociaux de cette plateforme et un coin ou n'importe quelle personne accédant au site peut nous envoyer un message et toute en bas il y a les réseaux sociaux du site et le compte github.

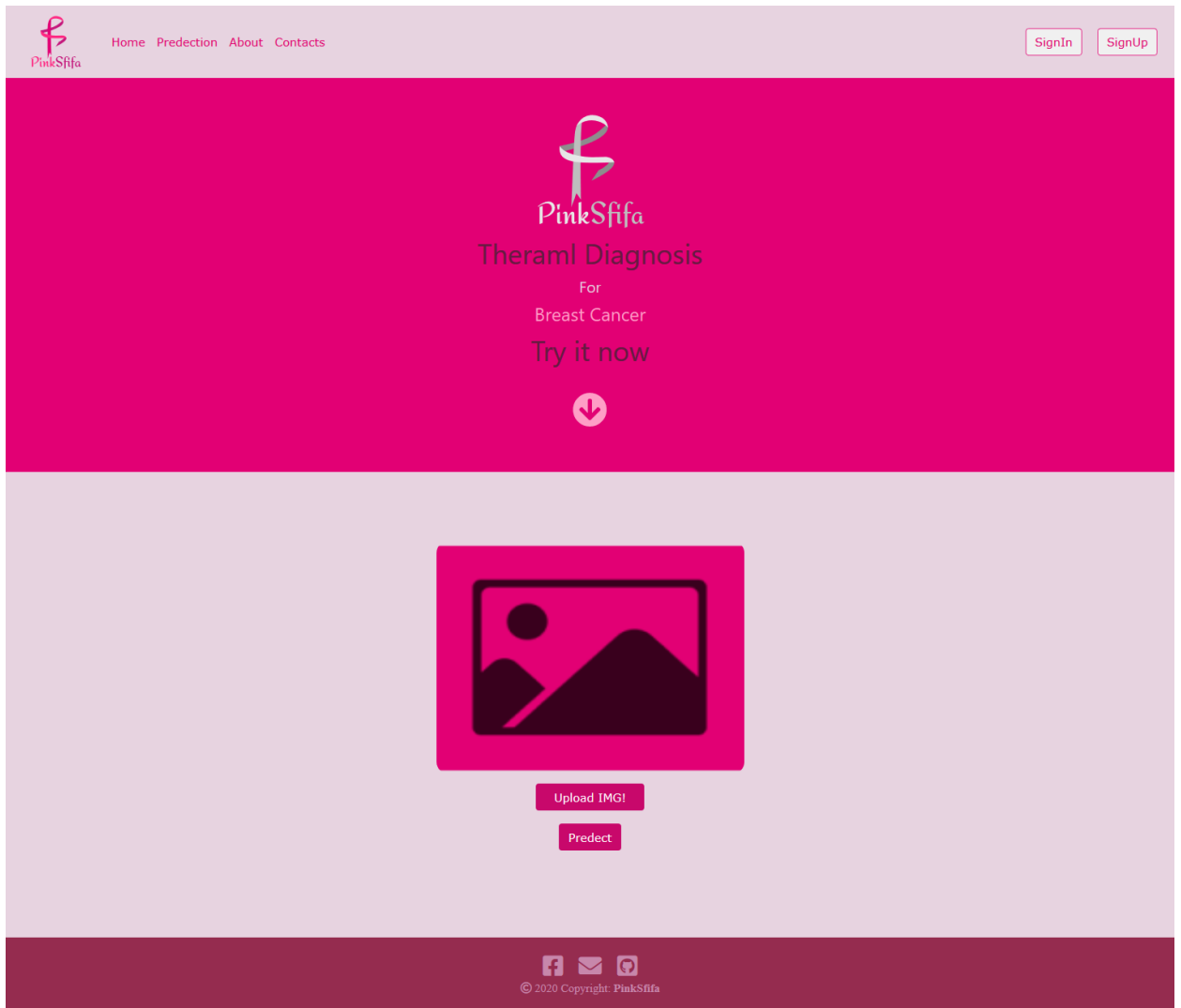


Figure 5.4 page principale pour faire les diagnostiques

Ici dans la figure 5.4 au-dessus c'est la page où le patient ayant un compte peut y accéder pour faire un diagnostic en cliquant sur le bouton upload pour choisir une image et puis sur predict pour avoir son diagnostic si un patient n'as pas de compte et entre dans cette page et il essaye de faire un diagnostic il ne pourra pas et un message d'erreur apparaîtra pour lui faire comprendre.

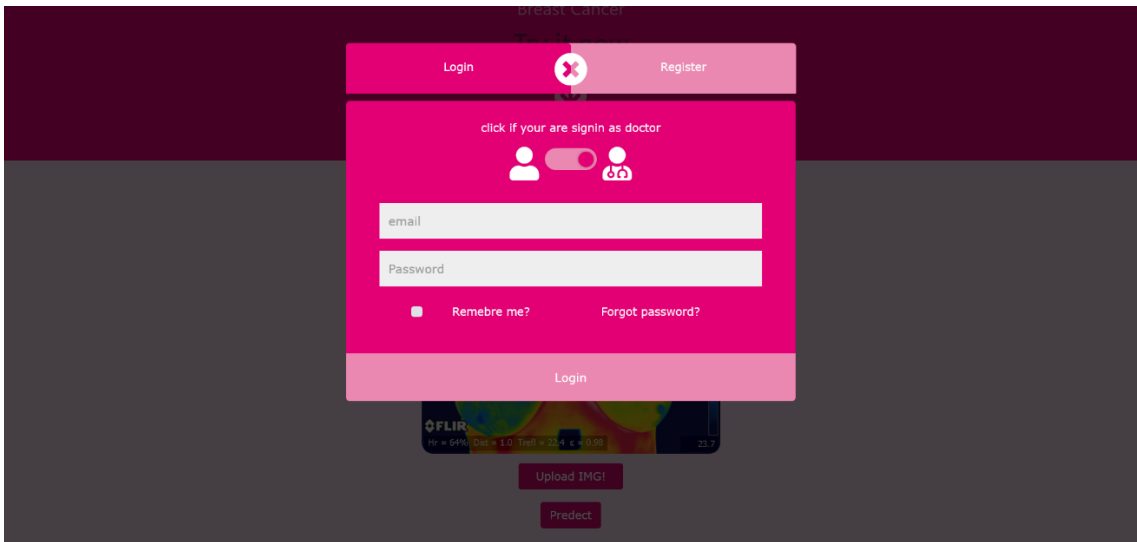


Figure 5.5, Identification

La figure 5.5 au-dessus montre la page d'identification ou le patient ou le médecin ayant déjà un compte pourront se connecté en introduisant leur email et leur mot de passe, dans cette figure en vois que c un médecin qui vas se connecté car la case médecin est cocher.



Figure 5.6, registrement

La figure 5.6 au-dessus nous montre comment un patient ou un docteur pourront crée leur compte au entrant les informations qui sont montré sur la page et en choisissant le type de personne sois patient ou médecin au niveau de (Type of registrant) puis en cliquant sur le bouton register .

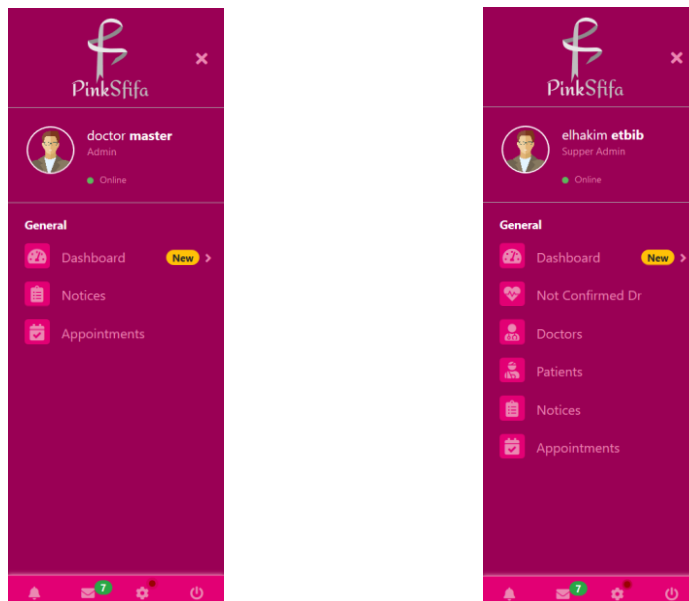


Figure 5.7 une barre de travail d'un admin simple avec moins d'action à faire

Dans la figure 5.7 on peut voir la différence entre la barre de travail d'un simple admin et d'un super admin ou il y a une différence entre les fonctions comme pour un super admin il peut ajouter des médecin en confirmant leur demande ou supprimer leur demande alors que le simple admin peut que voir son planning de rendez-vous et maitre un avis sur un diagnostic d'un patient.

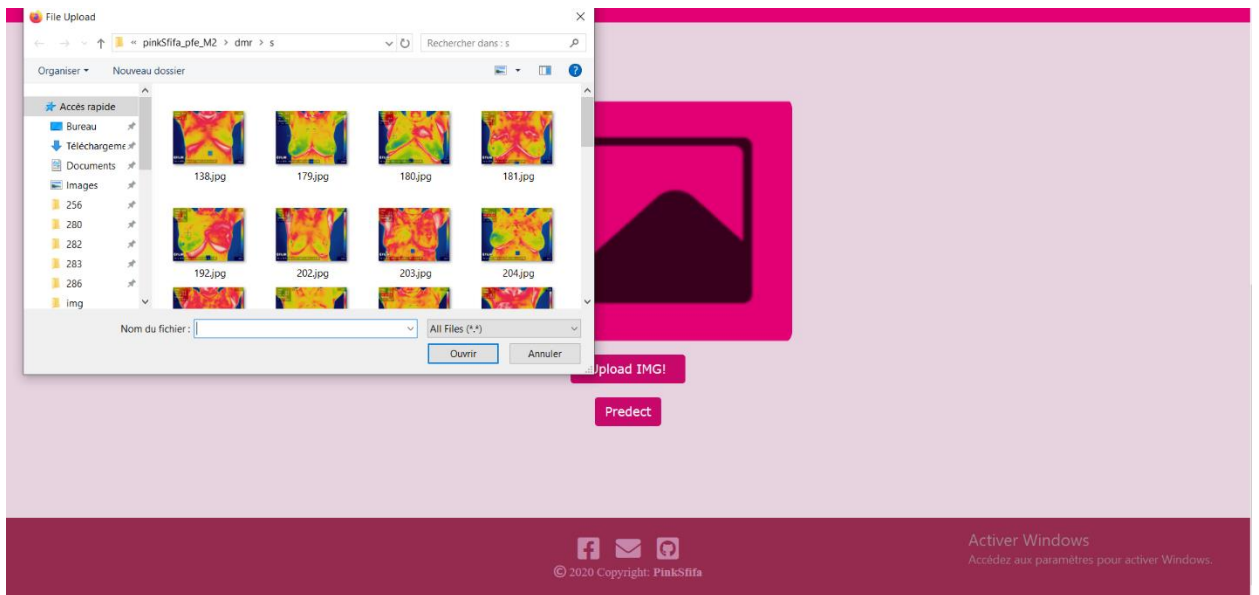


Figure 5.8, choisir une image

Dans la figure 5.8 on peut voir comment un patient peut choisir une image en cliquant sur le bouton upload.

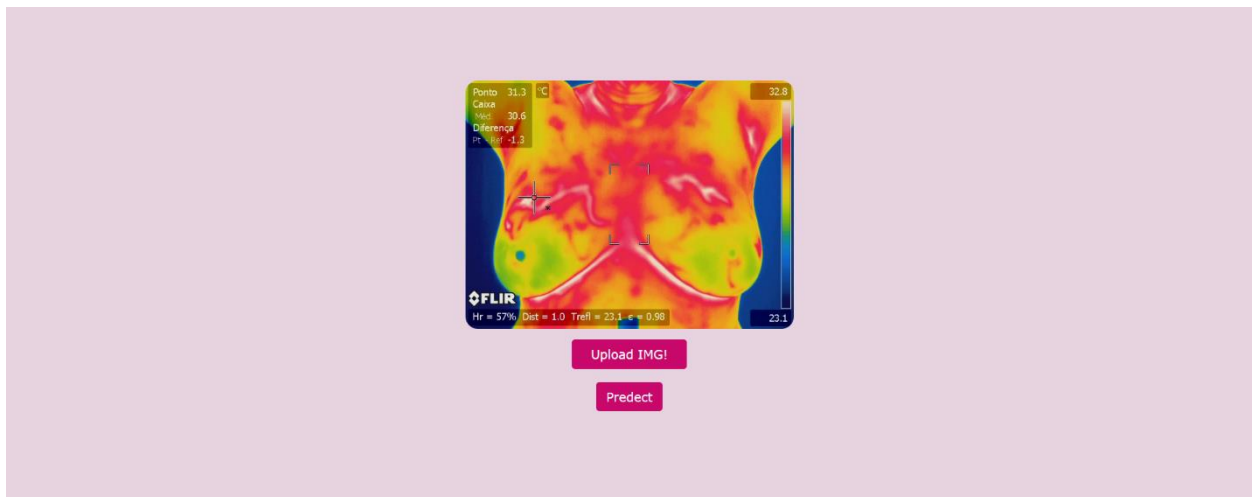


Figure 5.9, clique sure prediction (predict) pour faire une diagnostique

Dans la figure 5.9 l'image a bien été choisie et elle est affichée, le patient doit cliquer sur predict pour avoir son diagnostic si il est connecté avec son compte bien sûr.



Figure 5.10, les résultats pour un cas positif

Dans la figure 5.10 on peut voir un diagnostic fait par un patient et on peut voir ses informations affichées à droite de la page et son diagnostic en dessous, On voit aussi les deux parties de la poitrine gauche et droite être afficher et en dessous il y a un autre résultat afficher pour l'image introduite par le patient, Ce résultat est fait par l'algorithme d'apprentissage non supervisé en calculant la distance entre l'image introduite et les autres images de la base de donnés.

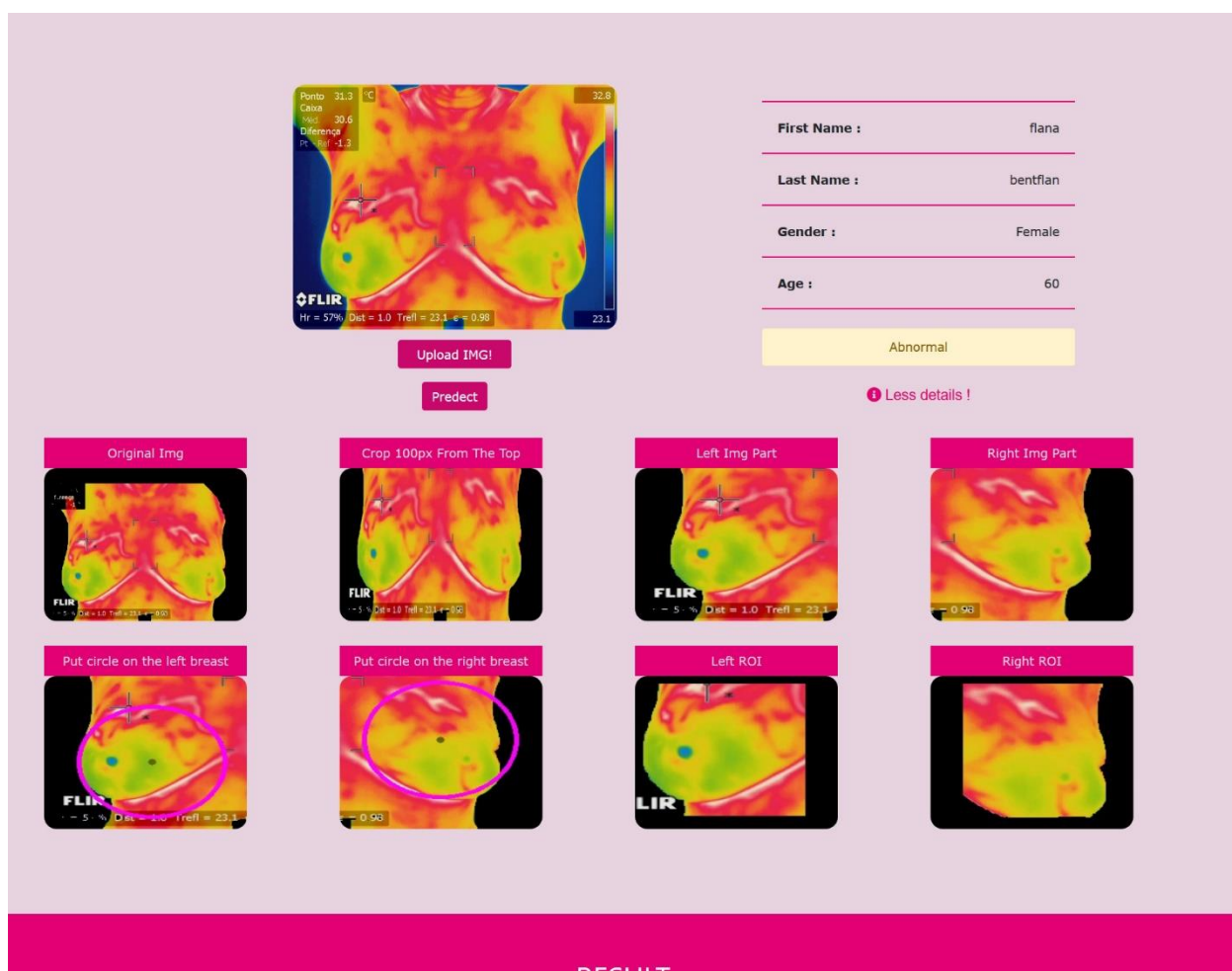


Figure 5.11, plus de details après un clique sur plus de details.

Ici dans la figure 5.11 on peut voir en détails les images extrait par notre algorithme d'extraction de la région d'intérêt jusqu'à avoir extrait les deux partie de la poitrine droit et gauche si le diagnostic vas être (Abnormal) une anormalité au niveau de la poitrine.

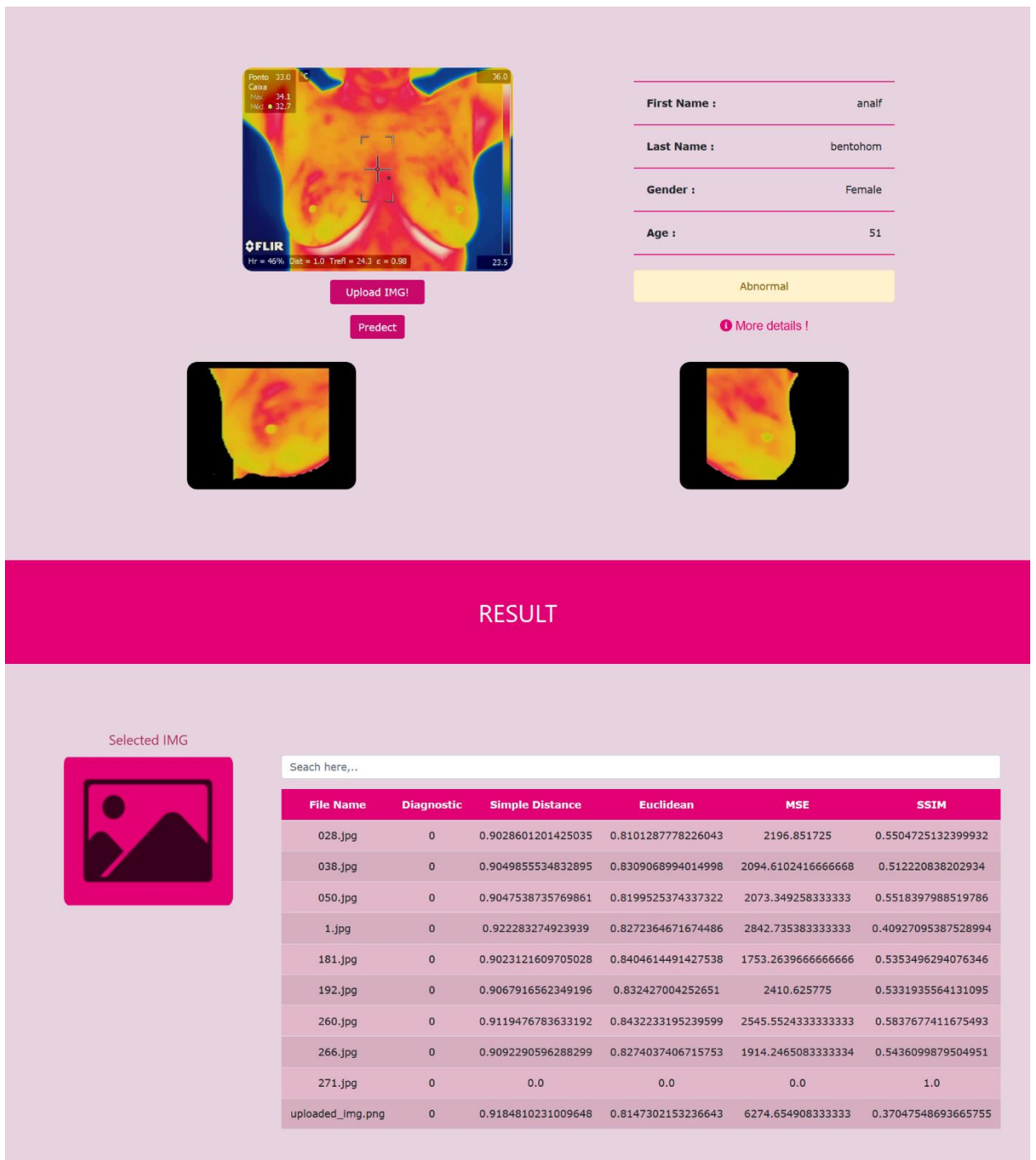


Figure 5.12 autre diagnostic anormal

Ici dans la figure 5.12 on peut voir un autre diagnostic anormal pour une femme qui a 31 ans et on voit aussi en dessous un autre résultat avec l'algorithme d'apprentissage non supervisée avec le calcul de la similarité.

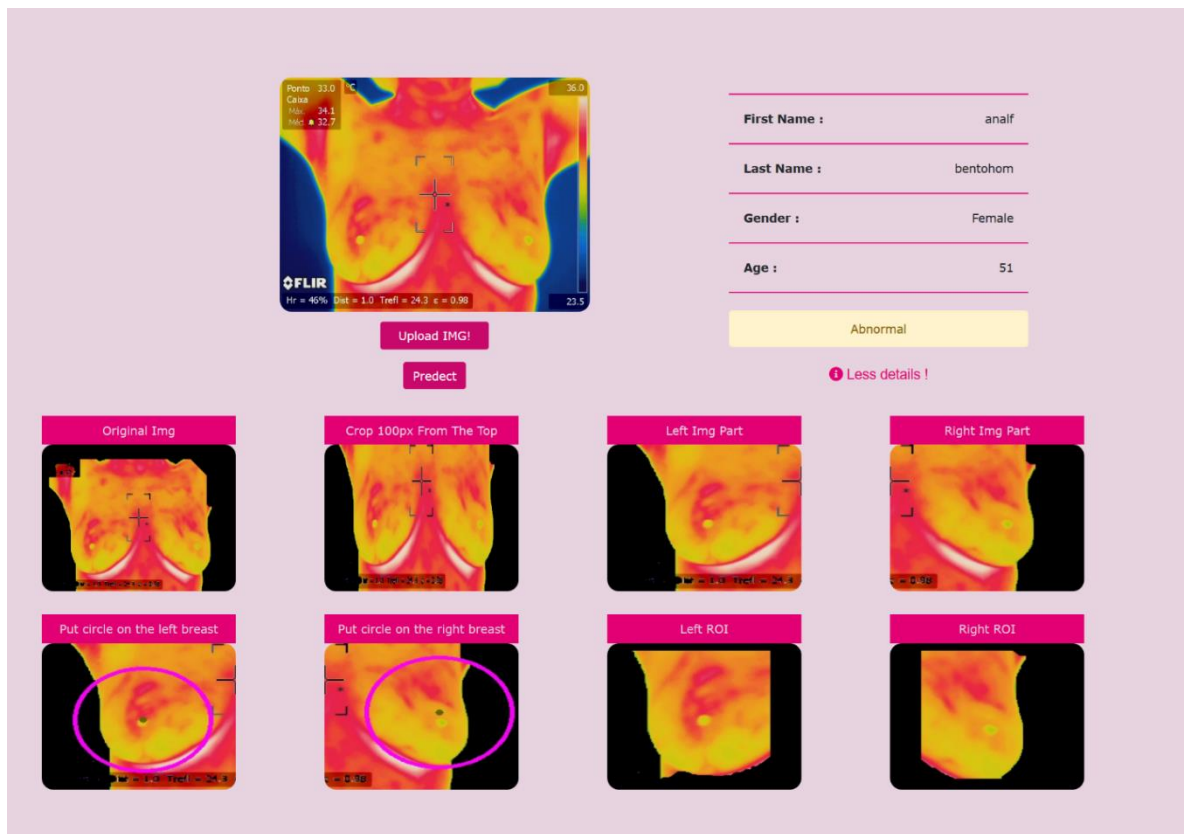
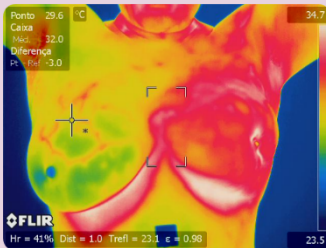


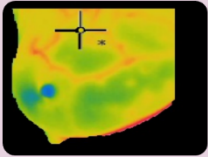
Figure 5.13 autre diagnostic anormal avec plus de détails.

Dans la figure 5.13 on voit une autre femme de 51 diagnostiquer anormal avec plus de détails.



Upload IMG!

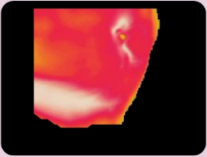
Predict



First Name :	analf
Last Name :	bentohom
Gender :	Female
Age :	51


Abnormal

More details !



RESULT

Selected IMG



Search here,...

File Name	Diagnostic	Simple Distance	Euclidean	MSE	SSIM
028.jpg	0	0.9082282748120813	0.8209845938960625	3364.1309833333335	0.4840756802340216
038.jpg	0	0.9071457279927733	0.821620758077856	2633.1664333333333	0.49582484785451625
050.jpg	0	0.9062286843874804	0.8278786615400601	2734.39365	0.5196867211733847
1.jpg	0	0.9204288556227948	0.8268714730999454	3671.8254416666667	0.3637971541338213
181.jpg	0	0.9043235632599519	0.8305033911333672	2069.5079583333333	0.5504331271635107
192.jpg	0	0.0	0.0	0.0	1.0
260.jpg	0	0.8964412169174337	0.8339746157903895	2406.9448916666665	0.563426391719248
uploaded_img.png	0	0.9260633846168227	0.8150049884106149	6490.4198166666665	0.3592537587353223

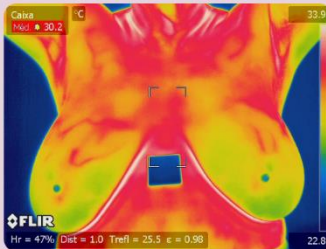
Figure 5.14 autre diagnostic anormale très voyant

Dans la figure 5.14 on voit bien qu'il y a une anomalie au niveau du sein droit juste pas les yeux et sa était confirmé par l'algorithme de classification du site.



Figure 5.15 image diagnostique anormale très voyante

Dans la figure 5.23 on a la même image que dans la figure 5.14 mais avec plus de détails sur la poitrine



FLIR
Hr = 47% Dist = 1.0 Trefl = 25.5 ε = 0.98

Upload IMG!

Predict

First Name : flana

Last Name : bentflan


Gender : Female

Age : 60

Normal

RESULT

Selected IMG

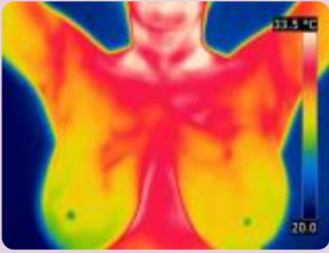


Search here,...

File Name	Diagnostic	Simple Distance	Euclidean	MSE	SSIM
028.jpg	0	0.9080139704934613	0.8189077231386884	2919.605666666667	0.5115938649851013
038.jpg	0	0.0	0.0	0.0	1.0
181.jpg	0	0.9024971517563262	0.8403836430457217	1345.6243583333332	0.5225860684920557
260.jpg	0	0.9093970340245278	0.8404080259689126	2800.8339083333335	0.5564414245966223
uploaded_img.png	0	0.9229091892089712	0.8288295204523968	6715.57445	0.351621125984186

Figure 5.16, un résultat pour un cas negative

Ici dans la figure 5.16 on peut voir le diagnostic d'un patient qui est normal et les autres images détaillé du processus ne seront pas afficher car le patient se trouve en bonne santé et qu'il y a aucune anormalité.



Upload IMG!

Predict

First Name : analf

Last Name : bentohom


Gender : Female

Age : 51

Normal

RESULT

Selected IMG



Search here, ..

File Name	Diagnostic	Simple Distance	Euclidean	MSE	SSIM
028.jpg	0	0.9196117687311129	0.8295558983457353	3515.9891416666665	0.4071582995577483
038.jpg	0	0.9220301784334208	0.833490435774157	3535.7460583333333	0.3686835225639607
050.jpg	0	0.9203963167143372	0.8364072087691693	3451.0914083333333	0.4244723369199536
1.jpg	0	0.0	0.0	0.0	1.0
181.jpg	0	0.9209299366980246	0.8339794004498747	3228.6303	0.3565918603589235
260.jpg	0	0.9235228512108179	0.8376564830090194	4433.8695	0.39939049559456474
uploaded_img.png	0	0.922742199068805	0.8215081569250352	7641.450541666667	0.3681148519406164

Figure 5.17 autre résultat de cas négative

Ici dans la figure 5.17 on peut voir le diagnostic d'un autre patient qui a été diagnostiqué négative et la confirmation en dessous avec le tableau de similarité.

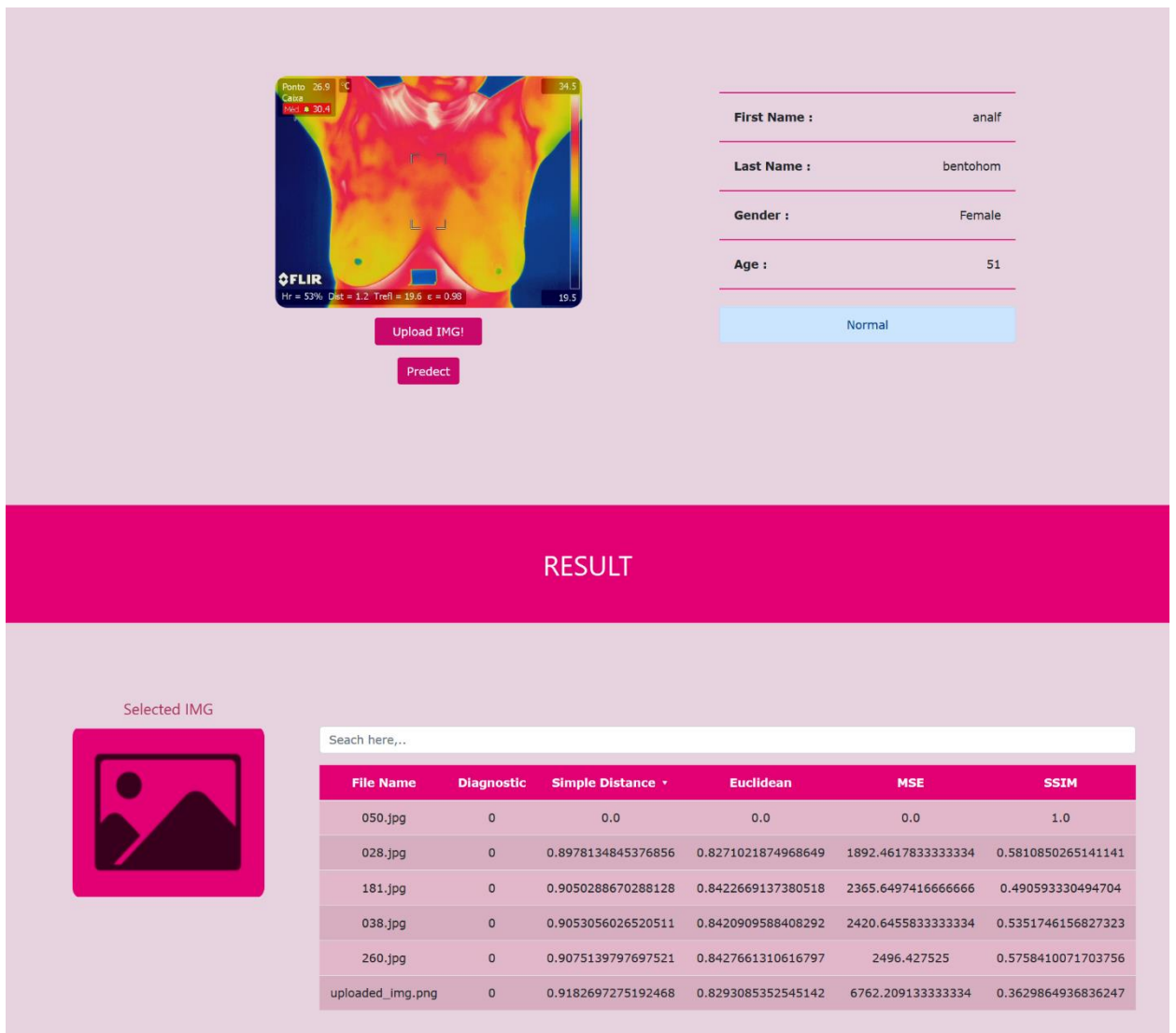


Figure 5.18 autre diagnostic négative

Ici dans la figure 5.18 on peut voir le diagnostic d'un patient qui est normal et les autres images détaillées du processus ne seront pas afficher car le patient se trouve en bonne santé et qu'il y a aucune anomalie avec la confirmation du tableau de similarité en dessous.

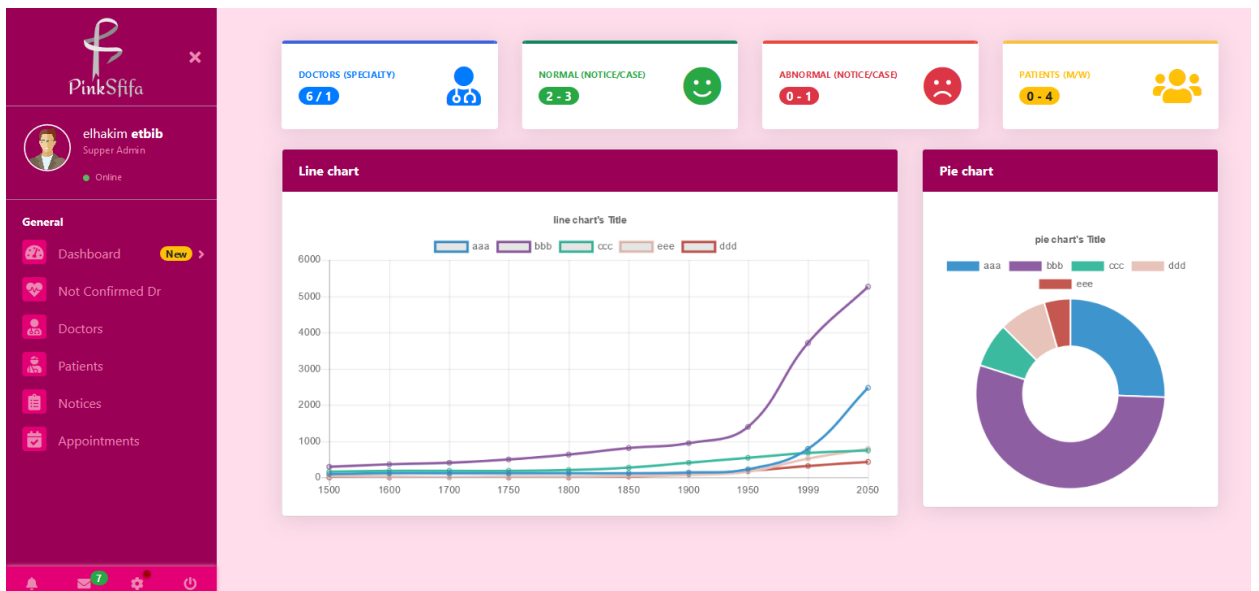


Figure 5.19, page principale pour les admins/super admins

Dans la figure 5.19 dessus on peut voir le page des super admins qui sont des médecins, On peut apercevoir tous les informations en ce qui concerne les patient et leur diagnostic et à gauche il y a une colonne qui comporte des bouton pour accède à tous les informations sur les patient, les médecins non confirmé qui sont des médecin ayant créé un compte mais qui non pas encore était validé par le super admin quand il seront validé ils seront automatiquement défini comme médecin au sein du site.

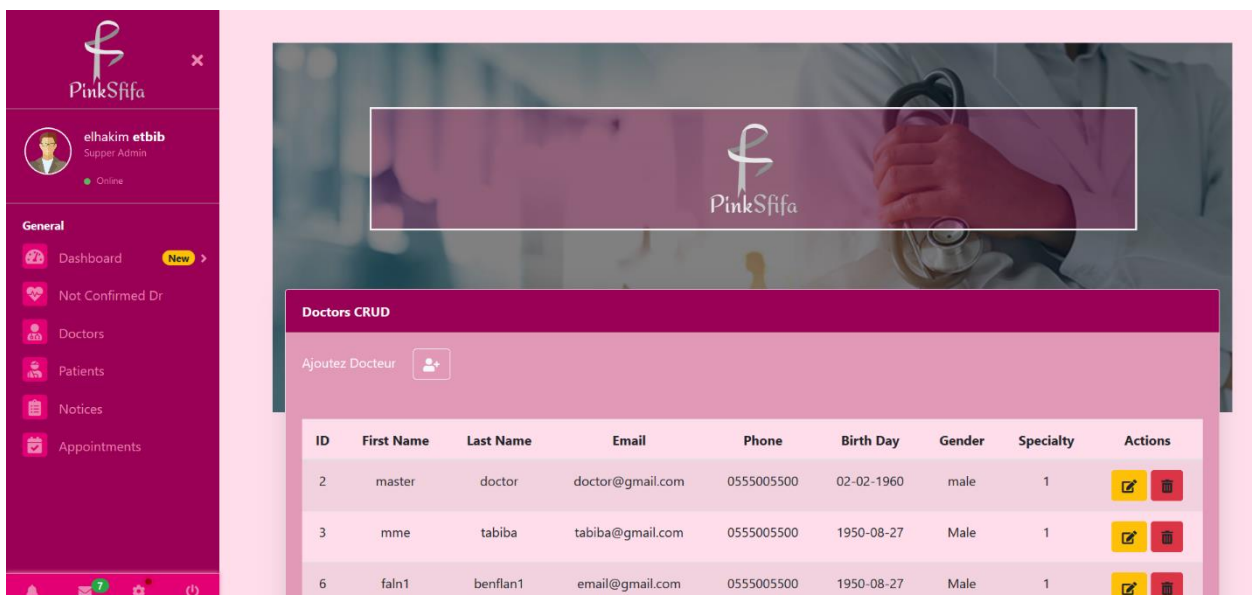


Figure 5.20, exemple des actions possible

Dans la figure 5.20 au-dessus il y a un tableau afficher qui contient tous les informations concernant les médecins et on peut voir à droite pour chaque médecin le super admin peut sois modifié ses information ou le supprimer complètement de la base de données.

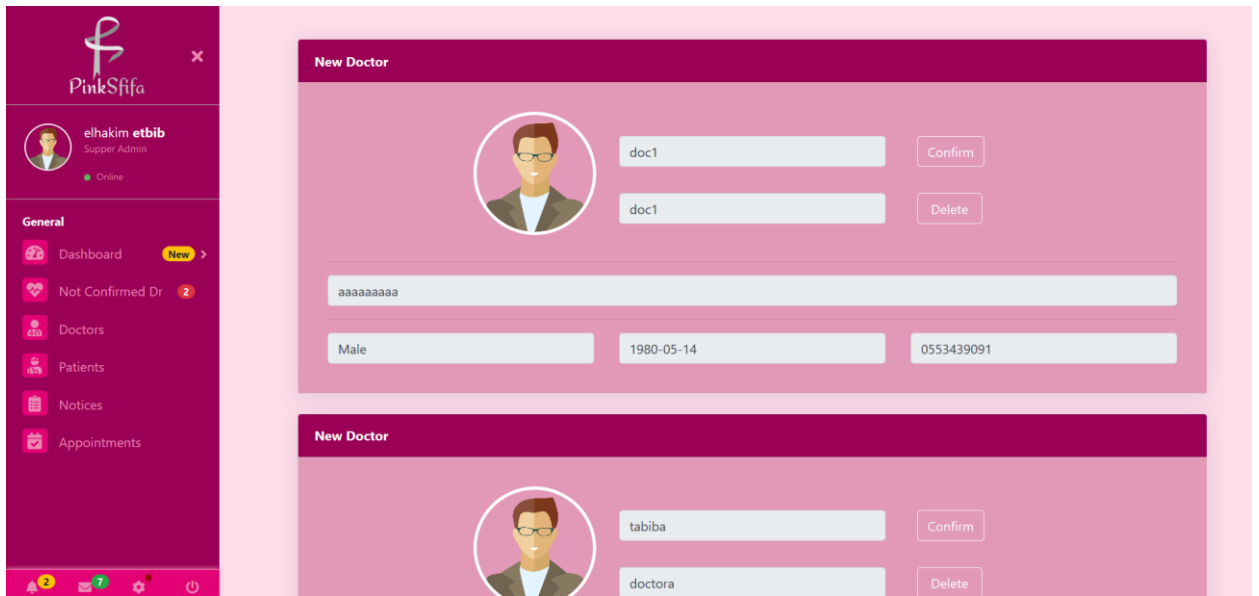


Figure 5.21, notification d'un nouveau inscri de type doctor, et les actions possible

Dans la figure 5.21 au-dessus voit deux notification arrive au super admin que deux médecin ont créé un compte et veulent accédé au site, Le super admin peut sois confirmer leur demande en cliquant sur le bouton confirm ou soit la refuser en cliquant sur le bouton delete.

5.5 conclusion

De nos jours, le développement de plates-formes de diagnostic assisté par ordinateur (Web-CAD) joue un rôle crucial dans le domaine médical qui fournit des outils d'accès rapide et facile aux experts médicaux pour effectuer un traitement et une analyse conviviaux des images. Ainsi, pour y parvenir, nous avons développé un WebCAD-IAD en tant que plate-forme d'analyse d'image et de diagnostic du cancer du sein appliquée sur des imageries thermiques. Le WebCAD-IAD est une plate-forme intelligente pour l'analyse d'images et un système de diagnostic qui récupère des données similaires dans une base de données médicales pour aider le radiologue à diagnostiquer les lésions.

Chapitre 6 CONCLUSION

Ce travail de recherche s'inscrit dans le cadre de la détection du cancer du sein à l'aide d'images thermiques. Dans le cadre de ce mémoire, nous avons traité une problématique qui touche l'un des cancers les plus connues et inquiétants à travers le monde.

La technologie d'imagerie thermique dans le cadre de la détection et le diagnostic du cancer du sein reste toujours difficile à réaliser, C'est pour ça que de nombreux travaux et recherches ont été effectués dans ce domaine, mais sans avoir pour autant trouvé une solution optimale et prête pour être utilisée dans le secteur médical. Pour le moment ces travaux restent des suggestions pour un deuxième avis au médecin.

Pour faire face à cette problématique, nous avons adopté une méthodologie de recherche qui consiste à concevoir et implémenter un système pour l'extraction de la zone d'intérêt et puis faire la classification et le calcul de similarité entre les images avec l'apprentissage supervisé et non supervisé. En effet, Ce système comporte plusieurs modules à savoir :

- Le premier module qui consiste à prendre une image thermique et à extraire de celle-ci les deux parties de la poitrine avec notre algorithme d'extraction de la zone d'intérêt(ROI).
- Le deuxième module consiste à prendre les images extraites du ROI et de faire avec une classification pour faire le diagnostic à l'aide de l'apprentissage supervisé.
- Le troisième module consiste à prendre les mêmes images extraites du ROI et calculer la similarité entre eux et les images de la base de données avec l'apprentissage non supervisé.

Le système développé dans notre projet a été évalué par plusieurs critères d'évaluations et plusieurs variantes, les résultats obtenus par le système sont encourageants par rapport aux systèmes développés antérieurement.

Les résultats obtenus par l'algorithme de segmentation (extraction de la région d'intérêt) ont été examinés par rapport à la base de données et les résultats sont très bons.

Les méthodes d'extraction automatisée de la région mammaire sont-elles fiables par rapport à la segmentation manuelle effectuée par des experts ?

Les résultats obtenus par l'algorithme de comparaison basé sur la distance euclidienne entre les coordonnées de la région d'intérêt de la méthode manuelle et les coordonnées de la méthode adaptative ont été très satisfaisants.

- Les résultats obtenus par les approches supervisées ont été très satisfaisants par rapport aux résultats obtenus dans les travaux antérieurs et pour les résultats de l'approche non supervisée ont été très encourageants.
- L'utilité de la Plateforme est qu'elle est très simple à manipuler et donne une très bonne expérience aux utilisateurs et les résultats affichés sont identiques à notre processus de dépistage. Cette plateforme règle tous les problèmes liés au déplacement des patients et le temps d'obtenir leurs résultats.

Notre travail ouvre des perspectives scientifiques à court et à long terme. Nous soulignons dans ce qui suit les perspectives qui nous semblent pertinentes pour l'évolution des systèmes développés dans ce projet.

- La première perspective sera de travailler avec une base de données plus volumineuse et de créer notre propre base de données basée sur notre algorithme de ROI.
- La deuxième perspective est d'améliorer notre algorithme d'extraction de la région d'intérêt pour les images de position latéral droit 45 °, latéral droit 90 °, latéral gauche 45 ° et latéral gauche 90 °.
- La troisième perspective est d'étendre notre travail pour d'autres types d'images comme la mammographie et l'élastographie.
- La quatrième perspective est d'améliorer les résultats obtenus dans l'approche supervisée et non supervisée et fournir aux utilisateurs plus de variétés en termes de résultats pour consolider l'avis du médecin.

Bibliographie

- [1] (2018) A Comprehensive Guide to Convolutional Neural Networks. [Online]. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- [2] A.Ali, Detection of Breast Abnormalities of Thermograms based on a New Segmentation Method. Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, 2015.
- [3] Bandyopadhyaya, Significance of PELP1 in ER-Negative Breast Cancer Metastasis, 2011.
- [4] (2017) Content Based Image Retrieval Using a Convolutional Denoising Autoencoder. [Online]. <https://medium.com/sicara/keras-tutorial-content-based-image-retrieval-convolutional-denoising-autoencoder-dc91450cc511>
- [5] (2015) Database of mastologic images. [Online]. <http://visual.ic.uff.br/dmi/>
- [6] e-cancer, les maladies du sein. Récupéré sur e-cancer:, 2018.
- [7] Strasser Weippl K., Li J. J., Louis J. S., Finkelstein D. M., Yu K. D., Chen W. Q., Shao Z. M., Goss P. E. Fan L., *Breast Cancer in China. Lancet Oncology Vol 15 Issue 7 p. 279-289.*, 2014.
- [8] S. V., Sasikala, M., Bhavani Bharathi, G., & Jaipurkar, S. D. Francis, Breast cancer detection in rotational thermography images using texture features, 2014.
- [9] M. A. S., Sayed, G. I., Gaber, T., Hassanien, A. E., Snasel, V., & Silva, L. F. Gaythri. (2012) Breast Cancer Diagnosis Using Machine Learning Algorithms.
- [10] A. Hossam, AUTOMATIC IMAGE SEGMENTATION METHOD FOR BREAST CANCER ANALYSIS USING HERMOGRAPHY, *Journal of Engineering Sciences*, Vol. 46, No. 1, pp. 14-29, 2018.
- [11] (2018) hough circle algorithm. [Online]. https://docs.opencv.org/2.4/doc/tutorials/imgproc/imgtrans/hough_circle/hough_circle.html
- [12] Mohammed S, Ali H. A. ibrahim A, "Breast cancer detection and classification using thermography," in *a review. international conferance on advanced machina learning technologies and Applicattions*, Vol.723 p. 159-164, 2018.
- [13] SINIEWICZ K, S: L'importance de la thermo vision chez les enfants souffrant de troubles cardiovasculaires orthostatiques avant et après les exercices physiques dirigés, un recueil de l'imagerie infrarouge en médecine clinique, medpress , pp 87 – 93., 2003.
- [14] S. G., Perez-Raya, I., Raghupathi, P. A., Gonzalez-Hernandez, J.-L., Dabydeen, D., Medeiros, L., & Phatak, P Kandlikar, Infrared imaging technology for breast cancer detection – Current status, 2017.
- [15] Silva L, A New Database for Breast Research with Infrared Image, 2014.
- [16] Hany M, "Breast cancer detection in rotational thermography images using texture features. *Infrared Physics & Technology*, 67, 490–496. doi:10.1016/j.infrared.2014.08.019 ," 2018.

- [17] (2018) Machine Learning Basics with the K-Nearest Neighbors Algorithm. [Online]. <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>
- [18] Maitra, Detection of Abnormal Masses using Divide and Conquer Algorithm in Digital Mammogram, 2011.
- [19] S., Maresova, P., Krejcar, O., Selamat, A., & Kuca, K Mambou, Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model, 2018.
- [20] Mona, "Perspectives of Cosmesis following Breast Conservation for Multifocal and Multicentric Breast Cancers. International Journal of Breast Cancer, 2015, 1–9. doi:10.1155/2015/126793," 2015.
- [21] Narang, A Review of Breast Cancer Detection using ART Model of Neural Networks, 2012.
- [22] Naved, Fenugreek, a Naturally Occurring Edible Spice, 2011.
- [23] (2017) Operations on arrays. [Online]. https://docs.opencv.org/3.4/d2/de8/group_core_array.html
- [24] ponce, Corrélation entre les performances physiques mesurées des détecteurs et la qualité diagnostique de l'image en mammographie numérique, 2016.
- [25] Ponra, Mitochondrial dysfunction-induced apoptosis in breast carcinoma cells through a pH-dependent intracellular quercetin NDDS of PVPylated-TiO₂NPs.
- [26] M Ponraj, *Detection of Breast Abnormalities of Thermograms based on a New Segmentation Method. Proceedings of the 2015 Federated Conference on Computer Science and Information Systems*, 2018.
- [27] Zemouri R, "Breast cancer diagnosis based on joint variable selection and constructive deep neural network.IEEE," in *4h Middle East Conference on Biomedical Engineering*, p 496-505, 2018.
- [28] U. R., Ng, E. Y. K., Tan, J.-H., & Sree, S. V. Raad, Pathologic Response and Long-Term Follow-up in Breast Cancer Patients Treated With Neoadjuvant Chemotherapy: A Comparison Between Classifications and Their Practical Application., 2012.
- [29] (2019) Review: MobileNetV2 — Light Weight Model (Image Classification). [Online]. <https://towardsdatascience.com/review-mobilenetv2-light-weight-model-image-classification-8febb490e61c>
- [30] CHEF JF ET ELLIOTT RL., Imagerie infrarouge: progresser dans la réalisation de sa promesse médicale. IEEE Eng Med Biol Mag 21: 80-85., 2002.
- [31] Sedong M. Jiyoung H. Youngsun K. Yunyoung N. P. L. Bong-Keun J. Dongik O. & Wonhan Shin, Thermal Infrared Image Analysis for Breast Cancer Detection, KSII Transactions on Internet and Information Systems, V.11, No.2, pp.1134-1147., 2017.
- [32] (2018) taking the confusion out of confusion matrices. [Online]. <https://towardsdatascience.com/taking-the-confusion-out-of-confusion-matrices-c1ce054b3d3e>.
- [33] S., Woo, F., & Flores, F. Tello-Mijares, "Breast Cancer Identification via Thermography Image Segmentation with a Gradient Vector Flow and a Convolutional Neural Network. Journal of Healthcare Engineering, 2019, 1–13. doi:10.1155/2019/9807619," 2019.

- [34] (2019) Understanding and Coding a ResNet in Keras. [Online].
<https://towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33>
- [35] (2019) VGG16 – Convolutional Network for Classification and Detection. [Online].
<https://neurohive.io/en/popular-networks/vgg16/>
- [36] Villalobos-Montiel, Automatic Segmentation of Regions of Interest in Breast Thermographic Images. Pattern Recognition, 135–144, 2015.
- [37] Yasmine, Survey Paper on Diagnosis of Breast Cancer Using Image Processing, 2013.
- [38] Zadeh, Assessing the Potential of Thermal Imaging in Recognition of Breast Cancer, 2012.