Titre: Title:	Détection automatique du mamelon et du muscle pectoral pour l'évaluation du positionnement des seins en mammographie
Auteur: Author:	Lauriane Grondin-Reiher
Date:	2026
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Grondin-Reiher, L. (2026). Détection automatique du mamelon et du muscle pectoral pour l'évaluation du positionnement des seins en mammographie [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/76101/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/76101/
Directeurs de recherche: Advisors:	Lama Séoud, & François Guibault
Programme: Program:	Génie informatique

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Détection automatique du mamelon et du muscle pectoral pour l'évaluation du positionnement des seins en mammographie

LAURIANE GRONDIN-REIHER
Département de génie informatique et génie logiciel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie informatique

Mars 2026

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Ce mémoire intitulé :

Détection automatique du mamelon et du muscle pectoral pour l'évaluation du positionnement des seins en mammographie

présenté par **Lauriane GRONDIN-REIHER**
en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Farida CHERIET, présidente
Lama SÉOUD, membre et directrice de recherche
François GUIBAULT, membre et codirecteur de recherche
Tarek OULD-BACHIR, membre

DÉDICACE

À Valérie et Victor, mes bras droit et gauche,
nous ne faisons qu'un.

REMERCIEMENTS

Je tiens d'abord à remercier ma directrice de recherche, Lama Séoud. Je suis infinement reconnaissante pour ton soutien constant et ton implication sans pareille. Merci aux membres du VisionIC qui ont fait de ces deux années une expérience inoubliable. Je désire également remercier mon co-directeur François Guibault ainsi que Philippe Debanné pour leur support, leurs rétroactions et leur contribution au bon déroulement du projet. Je remercie les organismes ayant financé ce travail, soient le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), MEDTEQ+ et la compagnie iMD Research. Je remercie Karine Morency, Isabelle Jean et Ginette Goudreau, technologues expertes en mammographie, pour leur précieuse collaboration.

Enfin, merci Valérie et Victor. Je n'aurais pu imaginer de meilleurs partenaires.

RÉSUMÉ

Le cancer du sein est un enjeu majeur de santé publique à l'échelle mondiale. La mammographie est la principale modalité d'imagerie utilisée à ce jour pour le dépistage de la maladie, permettant de réduire le taux de mortalité associé. Toutefois, la performance diagnostique de la mammographie dépend fortement de la qualité technique de l'examen. En particulier, le positionnement inadéquat des seins lors de l'acquisition des images peut compromettre la visualisation complète du tissu mammaire, augmentant le risque de faux négatifs et entraînant des reprises d'examen. Malgré l'existence de lignes directrices cliniques, les erreurs de positionnement demeurent fréquentes en pratique.

Dans ce contexte, ce mémoire s'inscrit dans un projet plus large visant le développement d'un système d'évaluation automatique du positionnement des seins en mammographie. Plus spécifiquement, ce travail propose une approche géométrique fondée sur la détection automatique de repères anatomiques, soient le muscle pectoral et le mamelon. Ces structures jouent un rôle central dans l'évaluation du positionnement, de par leur géométrie et leur position. Nous proposons donc un modèle basé sur l'architecture DeepLabv3+ pour l'identification du muscle pectoral ainsi qu'un modèle U-Net gaussien permettant la localisation du mamelon par une régression de cartes de chaleur. Quatre algorithmes géométriques ont été conçus pour évaluer des critères de positionnement à partir des repères détectés.

L'approche proposée se distingue par son caractère explicable. Plutôt que de reposer sur une classification bout-enbout, elle met en évidence les repères anatomiques détectés et les relations géométriques utilisées pour établir l'évaluation, favorisant ainsi l'interprétabilité des résultats. Ce travail contribue ainsi au développement d'un système d'évaluation automatique du positionnement mammaire. Un tel outil permettrait de réduire les erreurs de diagnostique liées à un positionnement inadéquat des seins durant la mammographie et réduire le taux de rappel des patientes.

ABSTRACT

Breast cancer remains a major global public health challenge. Mammography is currently the primary imaging modality used for breast cancer screening and has been shown to significantly reduce associated mortality. However, the diagnostic performance of mammography strongly depends on the technical quality of the examination. For instance, inadequate breast positioning during image acquisition may compromise the full visualization of breast tissue, increasing the risk of false negatives and leading to repeat examinations. Despite the existence of established clinical guidelines, positioning errors remain common.

This work is part of a broader project aimed at developing a system for automatic assessment of breast positioning in mammography. Specifically, we propose a geometric approach based on the detection of key anatomical landmarks: the pectoral muscle and the nipple. These structures play a central role in positioning assessment due to their relative geometry and spatial configuration within standard mammographic views. A DeepLabv3+ based model was developed for pectoral muscle identification, while a Gaussian U-Net model was implemented to localize the nipple through heatmap regression. Based on the detected landmarks, four geometric algorithms were developed to assess established positioning criteria.

The proposed approach is designed with explainability in mind. Rather than relying on end-to-end classification, the system explicitly identifies anatomical landmarks and leverages their geometric relationships to derive the final assessment, thereby enhancing interpretability. This work contributes to the development of an automated breast positioning evaluation system that could help reduce diagnostic errors associated with inadequate positioning and ultimately decrease patient recall rates in mammographic screening.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

ACR	American College of Radiology
ASPP	Atrous Spatial Pyramid Pooling
BCDR	Breast Cancer Digital Repository
BCE	Binary Cross Entropy - Entropie croisée binaire
CADe	Computer Aided Detection - Détection assistée par ordinateur
CAD	Computer Aided Diagnosis - Diagnostic assisté par ordinateur
CBIS-DDSM	Curated Breast Imaging Subset of DDSM
CC	Craniocaudale
CFR	Conditional Random Fields - Champs aléatoires conditionnels
CMMD	Chinese Mammography Database
CNN	Convolutional Neural Network - Réseaux de neurones convolutifs
CVAT	Computer Vision Annotation Tool
DCNN	Deep Convolutional Neural Network - Réseau de neurones convolutif profond
DDSM	Digital Database for Screening Mammography
EUREF	European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services
FDA	Food and Drug Administration
FCN	Fully Convolutional Network - Réseau entièrement convolutif
FPN	Feature Pyramid Network - Modèle pyramidal de caractéristiques
GAN	Generative Adversarial Network - Modèle adversarial génératif
GRAD-CAM	Gradient-weighted Class Activation Mapping - Mappage d'activation de classe pondéré par gradient
HED	Holistically-nested Edge Detection
IMA	Inframammary Angle - Angle infra-mammaire
IoU	Intersection over Union - Intersection sur union
ISBI	International Symposium on Biomedical Imaging
MIAS	Mammographic Image Analysis Society Mammography Database
MLO	Médio-latérale oblique
MPA	Mean Pixel Accuracy - Précision moyenne au niveau du pixel
MSE	Mean Square Error - Erreur quadratique moyenne
OTIMROEPMQ	Ordre des technologues en imagerie médicale, en radio-oncologie et en électrophysiologie du Québec

PA	Pixel Accuracy - Précision au niveau du pixel
PL	Pectoral line - Ligne du muscle pectoral
PNL	Posterior Nipple Line - Ligne postérieure du mamelon
R-CNN	Regional Convolutional Neural Network - Réseau convolutif régional
RNN	Recursive Neural Network - Réseau de neurones récurrent
RSNA	Radiological Society of North America
RSNA - SMBC	RSNA Screening Mammography Breast Cancer Detection Dataset

CHAPITRE 1 INTRODUCTION

Le cancer du sein est le cancer le plus fréquemment diagnostiqué chez la femme à l'échelle mondiale, avec plus de deux millions de nouveaux cas rapportés en 2022 par l'Organisation mondiale de la santé, et demeure une cause majeure de mortalité liée au cancer [8–10]. Au Canada, il est estimé qu'une femme sur huit recevra un diagnostic de cancer du sein au cours de sa vie [11]. Le dépistage précoce du cancer du sein joue un rôle essentiel dans la réduction du taux de mortalité, puisqu'il permet la détection de la maladie avant l'apparition des premiers symptômes, favorisant ainsi une prise en charge et un traitement plus rapide. Dans cette optique, des programmes de dépistage ont été mis en place à travers le monde afin d'améliorer la survie des patientes et de réduire l'impact du cancer du sein sur les systèmes de santé. Ces derniers reposent principalement sur la mammographie comme modalité d'imagerie, un examen à faible dose de rayons X permettant de détecter des lésions à un stade pré-clinique.

Bien que la mammographie soit actuellement la méthode la plus fiable et la plus utilisée pour le dépistage du cancer du sein, elle n'est pas infaillible. En effet, entre 16 % et 30 % des cancers ne sont pas détectés. Plusieurs facteurs peuvent réduire la précision de l'examen, dont un âge plus jeune et une densité mammaire élevée. La qualité du positionnement des seins lors de l'acquisition des images constitue également un élément majeur influençant l'efficacité d'un examen de mammographie ainsi que la fiabilité du diagnostic. En effet, un positionnement inadéquat peut entraîner une visualisation incomplète du tissu mammaire, augmentant le risque de faux négatifs. De plus, des erreurs de positionnement sont une cause fréquente de reprises d'examen, exposant inutilement les patientes à une dose de radiation additionnelle, augmentant les coûts du système de santé et générant de l'anxiété chez les patientes. Selon un audit mené en 2017 dans plusieurs centres de dépistage du Québec, 47 % des examens de mammographie présentent au moins une erreur de positionnement, soulignant l'ampleur du problème en pratique clinique [12].

Dans ce contexte, un projet visant à développer un système d'évaluation automatique du positionnement des seins est né. Un tel outil, utilisé lors de l'acquisition de l'examen, pourrait offrir une rétroaction aux professionnels réalisant l'examen et permettre une reprise immédiate des images en cas d'erreur de positionnement, limitant ainsi le rappel des patientes et réduisant le taux de faux négatifs. Ce projet, en partenariat avec IMD Research, englobe plusieurs sous-projets menés par trois étudiants à la maîtrise explorant chacun une avenue différente de l'automatisation de l'évaluation du positionnement des seins.

1.1 Anatomie du sein et cancer du sein

Le sein est composé de graisse, de tissu conjonctif, de glandes et de canaux. Les ligaments soutiennent le sein, reliant la peau aux muscles thoraciques. Les lobules produisent le lait sous l'effet des hormones de grossesse, et les canaux le transportent des lobules vers le mamelon, qui est entouré par l'aréole. Les glandes constituent le tissu glandulaire, tandis que la graisse rétro-glandulaire fait référence au tissu graisseux situé derrière celles-ci. Le sein s'appuie sur le muscle pectoral, qui le sépare des côtes. Le sein contient également de nombreux vaisseaux sanguins et lymphatiques, et est entouré de ganglions [13].

Le cancer du sein prend principalement naissance dans les canaux et les lobules, résultant d'une croissance incontrôlée des cellules [10]. Le dépistage et le diagnostic du cancer du sein reposent sur diverses modalités d'imagerie, notamment la tomosynthèse, l'échographie, l'imagerie par résonnance magnétique, l'imagerie thermique, l'histopathologie et la mammographie [14]. Cette dernière constitue à ce jour la modalité employée dans la plupart des programmes de dépistage.

1.2 Mammographie de dépistage

La mammographie est un examen à faible dose de rayons X permettant de détecter des lésions dans le sein avant même qu'elles ne soient palpables [15]. Utilisée comme méthode de dépistage, elle contribue de manière significative à la réduction du taux de mortalité lié au cancer du sein [16]. Une mammographie standard repose sur l'acquisition de deux incidences complémentaires par sein : la vue craniocaudale (CC) et la vue médio-latérale oblique (MLO). L'incidence MLO correspond à une projection oblique d'environ 45°, allant du quadrant supéro-médial vers l'inféro-latéral, et permet la visualisation du quadrant supéro-externe, qui contient la plus grande proportion de tissu mammaire [1]. L'incidence CC est obtenue en orientant les rayons X verticalement depuis le haut du thorax, le détecteur étant placé sous le sein et la compression appliquée par le dessus. Correctement réalisée, elle permet une exploration plus complète du tissu glandulaire médial profond, souvent insuffisamment visualisé sur l'incidence MLO [1, 17]. L'association de ces deux vues est donc indispensable pour une visualisation complète du tissu mammaire [18]. Un exemple d'images composant l'examen complet est présenté sur la figure 1.1.

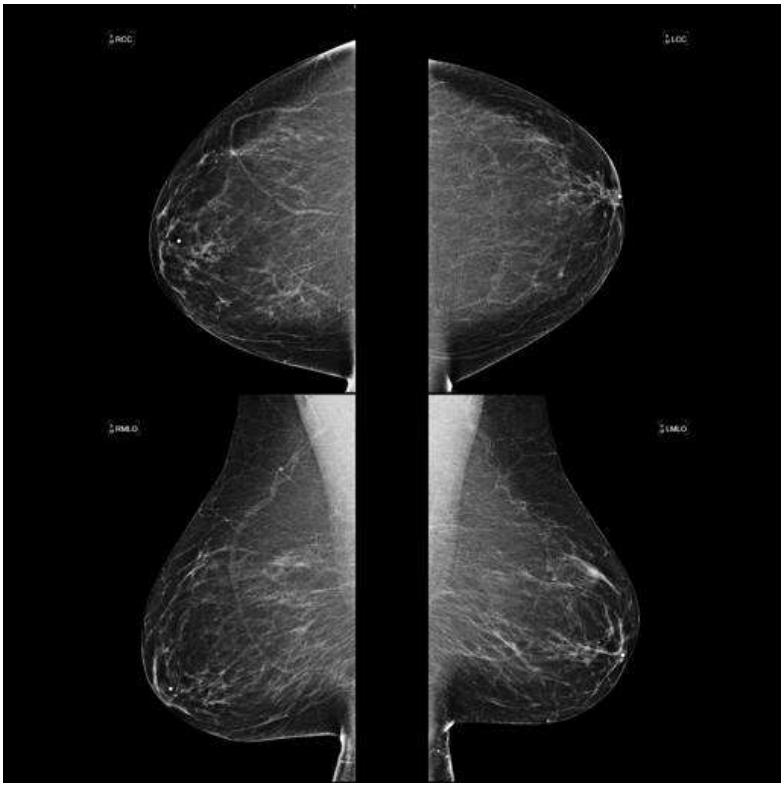


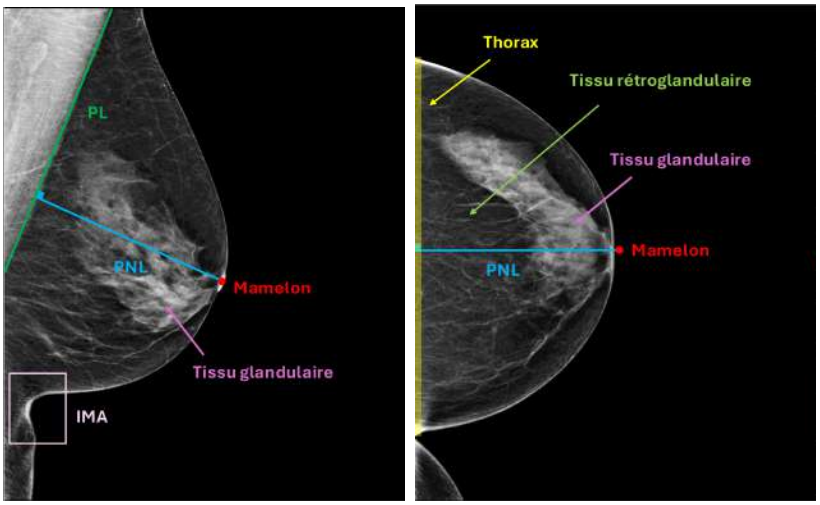
FIGURE 1.1 Les 4 images composant un examen mammographique : la vue CC (haut) du sein droit et du sein gauche ; la vue MLO (bas) du sein droit et du sein gauche.

Lors de l'examen, la patiente est positionnée par un ou une technologue en imagerie médicale, qui comprime le sein entre deux plaques afin de réduire l'épaisseur du tissu mammaire, permettant ainsi de diminuer la dose de rayonnement, de limiter le flou de mouvement et d'optimiser la qualité de l'image [19]. Cette compression, suivie d'une brève irradiation par rayons X, peut être source d'inconfort ou de douleur. Le rôle du ou de la technologue est central : il ou elle adapte le positionnement, la compression et les paramètres d'acquisition à la morphologie de chaque patiente tout en assurant au mieux son confort [19, 20]. Un marqueur de plomb est parfois apposé sur le mamelon durant l'examen, le rendant clairement visible sur la mammographie. Le déroulement de l'examen est fortement influencé par des facteurs propres à la patiente, notamment la densité mammaire, la taille et la forme du sein, l'indice de masse corporelle, la souplesse des tissus ainsi que la tolérance à la douleur. La mammographie est en effet un examen sensible à la morphologie, qui doit être prise en compte lors du positionnement des seins et l'acquisition des images.

1.3 Positionnement mammaire

Tel que mentionné précédemment, l'efficacité de l'examen de mammographie repose entre autres sur le positionnement adéquat des seins durant l'acquisition des images. Afin de standardiser la qualité des examens et de minimiser les erreurs de positionnement, plusieurs organismes ont établi des lignes directrices précises définissant des critères de positionnement adéquat des seins. En Europe, les recommandations sont émises par l'*European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services* (EUREF) [21]. Au Royaume-Uni, le *United Kingdom's National Health Service Breast Screening Programme* [22] propose également des critères d'évaluation de la qualité des mammographies. Aux États-Unis, l'*American College of Radiology* (ACR) [23] publie des standards détaillés dans le cadre de son programme d'accréditation en mammographie, tandis qu'au Canada, ces critères sont définis par la *Canadian Association of Radiologists* (CAR) [24]. Ces lignes directrices constituent une référence essentielle pour la formation des technologues et l'assurance de la qualité du positionnement en mammographie. Au Québec, une liste de critères dérivée des directives précédentes a été adaptée par l'Ordre des technologues en imagerie médicale, en radio-oncologie et en électrophysiologie (OTIMROEPMQ) dans le cadre de l'audit susmentionné. Ces derniers, introduits dans [12], sont ceux considérés dans ce travail. La compréhension de ces critères nécessite la connaissance de repères anatomiques importants, qui sont visibles dans une mammographie correctement réalisée. Ces repères sont illustrés sur la figure 1.2, adaptée d'un article rédigé par notre équipe et soumis au *Journal of Breast Imaging* [2]. Cet article, intitulé *Breast positioning in mammography : consolidated image criteria, their interdependence and trends in automatic evaluation*, aborde la définition des critères de positionnement d'un point de vue adapté à la vision par ordinateur, ainsi que la relation d'interdépendance entre ceux-ci.

Dans les deux vues, on y voit le mamelon, identifié par un point rouge. Le tissu glandulaire, correspondant aux glandes mammaires, est caractérisé par une zone d'intensité plus élevée dans l'image. Dans la vue MLO, l'angle infra-mammaire (IMA) est l'angle formé par la face inférieure du sein et la paroi thoracique, encadré en rose sur la figure 1.2. Le muscle pectoral se distingue par une zone claire de forme relativement triangulaire dans le coin supérieur de l'image (droit ou gauche selon le sein). Il est délimité par la ligne du muscle pectoral (PL), tracée ici en vert, qui s'étend du point inférieur du muscle au point supérieur, parallèlement aux fibres musculaires. La ligne postérieure du mamelon (PNL) relie le mammelon au muscle pectoral à angle droit. Dans la vue CC, le sein est imagé du mamelon au thorax, incluant le tissu glandulaire ainsi que le tissu rétro-glandulaire. La PNL relie dans cette vue le mamelon au thorax, situé au bord postérieur de l'image. Ces repères anatomiques sont d'une grande



(a) Vue MLO (b) Vue CC

FIGURE 1.2 Repères anatomiques importants devant être visibles sur une mammographie [2]

importance, puisque leur visibilité, leur position ou la relation géométrique entre certains de ceux-ci doivent être pris en compte dans l'évaluation de plusieurs critères de positionnement.

Les critères de positionnement considérés dans ce projet sont résumés dans le tableau 1.1 et décrits brièvement ici. La qualité globale de l'image, notamment en termes de contraste et de résolution, constitue un prérequis indispensable à l'évaluation des critères subséquents. Les deux vues doivent être exemptes de plis cutanés ou d'artéfacts susceptibles de masquer des structures d'intérêt. Le mamelon doit être visible de profil et aucun tissu mammaire ne doit être exclu de l'image. Les tissus profonds doivent également être correctement visualisés, sans quoi certaines lésions pourraient passer inaperçues. Dans la vue CC, la PNL doit être perpendiculaire au bord de l'image. Une déviation indique une désorientation du mamelon, pouvant entraîner l'exclusion d'une partie du sein. La profondeur suffisante du sein est évaluée en comparant la PNL en vue CC à celle en vue MLO, une différence supérieure à 1 cm traduisant un manque de profondeur. En vue MLO, deux critères permettent de vérifier si le muscle pectoral est suffisamment visible, un reposant sur la position de l'intersection entre la PL et la PNL, et l'autre relatif à l'angle de la PL. Enfin, l'IMA doit être visible et ouvert.

Le respect de l'ensemble de ces critères garantit une visualisation complète des tissus pertinents et favorise la détection d'éventuelles anomalies. À l'inverse, le non-respect de l'un de ces critères peut conduire à un cancer manqué ou à la nécessité d'un réexamen. Il apparaît donc pertinent de développer des systèmes d'évaluation automatique du positionnement, capables de détecter ces erreurs de manière fiable.

Tableau 1.1 Liste des critères de positionnement

Vue	**Critère**	**Description**
CC et MLO	Image évaluable	Évaluation de la qualité technique de l'image.
	Plis de peau dans le sein	Détection de plis cutanés marqués pouvant comprimer les tissus ou gêner la visibilité des masses, compromettant ainsi la détection des lésions.
	Artéfacts et artéfacts de superposition	Identification de tout artéfact (par exemple, traces d'antisudorifique, poils ou artéfacts liés à l'équipement d'imagerie) susceptible d'obscurcir le tissu mammaire ou de simuler une pathologie.
	Mamelon vu de profil	Vérification de la visibilité du mamelon de profil sans superposition du parenchyme.
	Partie du sein coupée	Évaluation de l'éventuelle exclusion involontaire de tissu mammaire de l'image, notamment aux bords médial, latéral ou postérieur.
	Bonne visualisation des tissus profonds	Vérification de l'inclusion complète des tissus rétroglandulaires et glandulaires.
CC	PNL perpendiculaire au bord de l'image	Vérification de l'orientation adéquate du mamelon, favorisant l'inclusion de l'ensemble des tissus mammaires dans l'image.
	Longueur de la PNL en CC à moins de 1 cm de celle en MLO	Comparaison de la longueur de la PNL entre les incidences CC et MLO.
MLO	Quantité adéquate de muscle pectoral	Évaluation de l'extension suffisante du muscle pectoral le long du bord proximal de l'image, indiquant une profondeur adéquate des tissus inclus.
	Vue de la largeur maximale du muscle pectoral	Évaluation de l'orientation du muscle pectoral (angle minimal) et de la croissance continue du muscle.
	Angle infra-mammaire (IMA) bien ouvert et montré	Vérification d'une ouverture suffisante de l'IMA sans contact peau à peau.

1.4 Problématique et objectif de recherche

Le positionnement inadéquat des seins durant l'examen mammographique s'est révélé être un problème majeur dans le domaine, non seulement au Québec et au Canada, mais également ailleurs dans le monde. En effet, la *Food and Drug Administration* (FDA) rapporte qu'un positionnement suboptimal est la cause première de reprises des examens [1]. De ce fait, il y a un intérêt grandissant au sein de la communauté scientifique pour le développement de systèmes d'évaluation automatique du positionnement mammaire. C'est dans ce contexte que le projet MammoPos a été lancé au laboratoire VisionIC, en collaboration avec la compagnie IMD Research. Ce projet de grande envergure vise à explorer différentes approches d'évaluation automatique du positionnement des seins. Une des approches constituant présentement l'état de l'art repose sur la détection de repères anatomiques pour l'évaluation de certains critères de positionnement. Celle-ci permet de mettre en évidence des éléments morphologiques et des relations géométriques, offrant une meilleure interprétabilité et favorisant potentiellement l'intégration clinique. Ainsi, le travail présenté dans ce mémoire porte sur la détection automatique de repères antomiques et l'évaluation subséquente du positionnement mammaire. Plus spécifiquement, l'objectif est de **développer des modèles de détection du muscle pectoral et du mamelon pour évaluer les critères** *Quantité adéquate de muscle pectoral*, *Vue de la largeur maximale du muscle pectoral*, *PNL perpendiculaire au bord de l'image* ainsi que *Longueur de la PNL en CC à moins de 1 cm de celle en MLO*. Ces quatre critères ont une définition géométrique permettant une évaluation directe à partir de la PL et du mamelon. Ce mémoire présente donc le travail réalisé pour atteindre cet objectif.

1.5 Contributions

Dans le cadre de ce travail, plusieurs articles scientifiques ont été rédigés. Un premier article, intitulé *Breast Positioning in Mammography : Consolidated Image Criteria, Their Interdependence and Trends in Automatic Evaluation*, propose une définition des critères de positionnement mammaire considérés dans ce projet, formulée sous un angle adapté à leur évaluation par des méthodes de vision par ordinateur [2]. Cet article analyse également l'interdépendance entre certains critères et dresse un portrait des tendances actuelles dans la littérature concernant l'évaluation automatique du positionnement mammaire. Un second article, intitulé *Breast Positioning Assessment in Mammography : An Inter-Rater Reliability Study*, décrit le protocole d'annotation mis en place ainsi que l'analyse de la variabilité inter-annotateur effectuée sur un sous-ensemble du jeu de données [25]. Ces deux articles ont été rédigés en collaboration avec les autres étudiants participant au projet et ont été soumis pour

publication au *Journal of Breast Imaging*. Enfin, une partie de la méthodologie développée et des résultats présentés dans ce mémoire font l'objet d'un troisième article, intitulé *Deep Learning Based Landmarks Detection for Positioning Assessment in Mammography*. Ce travail sera présenté à la conférence internationale ISBI (*International Symposium on Biomedical Imaging*) et constitue le chapitre 5 de ce mémoire.

1.6 Plan du mémoire

La suite de ce mémoire est organisée comme suit. Le **chapitre 2** est consacré à une revue de la littérature. Le **chapitre 3** expose la rationnelle de la recherche ainsi que les objectifs poursuivis. Le **chapitre 4** est dédié à la méthodologie ayant mené à la création du jeu de données et l'analyse préliminaire de celui-ci. Le **chapitre 5** correspond à un article de conférence portant sur une partie de la méthodologie et des résultats de ce travail. Le **chapitre 6** est consacré à des expérimentations et à des résultats supplémentaires, tandis que le **chapitre 7** est une discussion générale des résultats obtenus. Enfin, le **chapitre 8** est dédié à la conclusion du mémoire.

CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre vise à établir les fondements techniques relatifs au projet présenté dans ce mémoire. L'approche proposée reposant sur la détection du muscle pectoral et du mamelon, les tâches de segmentation et de régression de points clés seront d'abord abordées. Ces deux concepts serviront en effet de base technique pour la détection des repères anatomiques. Un portrait des principales applications de la vision par ordinateur en mammographie sera ensuite dressé : un survol des travaux portant sur la détection et la classification du cancer du sein sera d'abord proposé, suivi d'une présentation des méthodes dédiées à la détection de repères anatomiques. Les principales bases de données en mammographie seront également recensées. Enfin, l'état des connaissances actuelles en lien avec le domaine du positionnement mammaire sera exposé.

2.1 Principe de segmentation et de régression de points clés

Les tâches de segmentation et de régression de points clés constituent les bases techniques de ce travail. Dans cette section, les principales approches proposées dans la littérature sont présentées.

2.1.1 Segmentation

La segmentation d'images constitue une tâche fondamentale en vision par ordinateur. Elle vise à diviser une image en régions cohérentes partageant des caractéristiques communes, telles que l'intensité, le contraste ou la texture [26]. On distingue généralement trois types de segmentation, illustrés à la figure 2.1 :

- La segmentation sémantique, où chaque pixel se voit attribuer une étiquette correspondant à une classe d'objet ;
- La segmentation d'instances, qui permet d'identifier et de séparer chaque occurrence individuelle d'un même type d'objet ;
- La segmentation panoptique, qui offre une combinaison des deux précédentes approches [27].

La segmentation intervient dans de nombreuses applications, notamment en imagerie médicale, permettant d'isoler des structures anatomiques ou des zones pathologiques, contribuant ainsi au diagnostic et à la prise de décision clinique [28].

Les méthodes traditionnelles de segmentation, généralement basées sur des opérations rela-

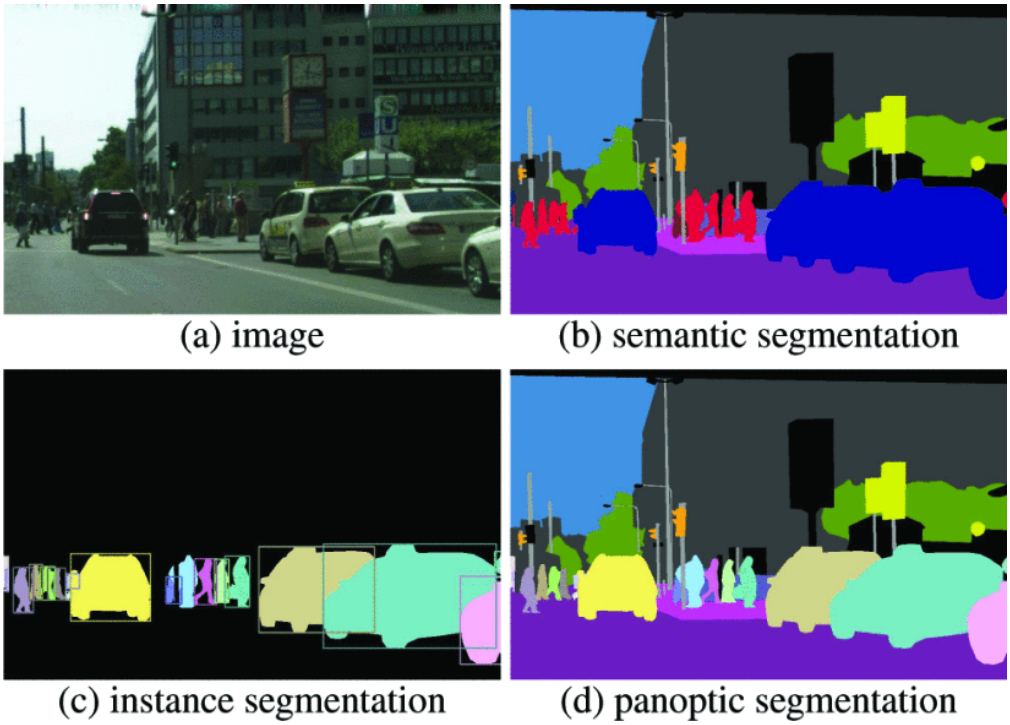


FIGURE 2.1 Les différents types de segmentation [3]

tives à l'intensité des pixels, ont longtemps dominé le domaine du traitement d'image [28]. On dénote parmi celles-ci les approches par seuillage, par contours et par régions. Avec l'émergence des techniques d'apprentissage automatique, d'autres familles sont apparues, incluant les méthodes de regroupement (k-means, fuzzy c-means) et les arbres de décision pour la classification de pixels. S'y ajoutent également les modèles probabilistes basés sur les graphes tels que les champs de Markov ou les champs aléatoires conditionnels (CRF). Malgré leur simplicité et leur interprétabilité, ces approches peinent à gérer la variabilité anatomique, le bruit et les artefacts, particulièrement présents en imagerie médicale [27–30].

L'essor de l'apprentissage profond et le développement des réseaux de neurones convolutifs (CNN) ont transformé le domaine, permettant l'extraction de caractéristiques visuelles et l'apprentissage de représentations complexes adaptées à la structure des images [28]. En effet, l'application successive de filtres convolutifs permet d'identifier des motifs visuels à différentes échelles. Les premières méthodes appliquées à la segmentation reposaient sur la classification de *patchs*, se voyant ainsi limitées en termes de précision spatiale [31]. Pour produire des prédictions pixel par pixel, ces architectures ont été adaptées. Shelhamer et al. introduisent les réseaux entièrement convolutionnels (FCN), remplaçant les couches entièrement connectées par des convolutions et employant des mécanismes de suréchantillonnage afin de générer des cartes de segmentation de même résolution que l'image d'entrée [27,28,32]. L'introduction des

connexions résiduelles, qui relient les cartes de caractéristiques profondes à celles des couches plus superficielles après le suréchantillonnage, permet de combiner l'information sémantique et visuelle et de gérer le problème de gradient qui s'annule [27, 31].

Les FCN inspirent l'architecture encodeur-décodeur, que Ronneberger et al. emploient dans leur modèle U-Net (Figure 2.2) [4]. Initialement conçu pour la segmentation biomédicale, U-Net est devenu l'architecture de référence en segmentation sémantique. Un encodeur permet l'extraction des caractéristiques de l'image d'entrée par un sous-échantillonnage successif, puis la résolution d'origine est retrouvée par un décodeur symétrique, conférant au modèle sa forme en U. L'utilisation de connexions résiduelles entre l'encodeur et le décodeur permet le transfert d'informations à différentes échelles, offrant une segmentation plus précise et détaillée en combinant les informations globales et locales. Ronnerberger et al. ont également démontré la pertinence d'appliquer des mécanismes d'augmentation pour entraîner un modèle à partir d'un nombre limité de données annotées. L'architecture U-Net a donné naissance à de nombreuses variantes, dont U-Net++ [33], qui améliore le transfert d'informations multi-échelle entre l'encodeur et le décodeur, ou encore l'Attention U-Net [34], qui intègre des mécanismes d'attention pour focaliser automatiquement le modèle sur les régions les plus pertinentes [28]. Plusieurs autres modèles proposés dans la littérature pour des tâches de segmentation sont conçus avec une architecture encodeur-décodeur : DeConvNet, SegNet, HRNet, etc. [27].

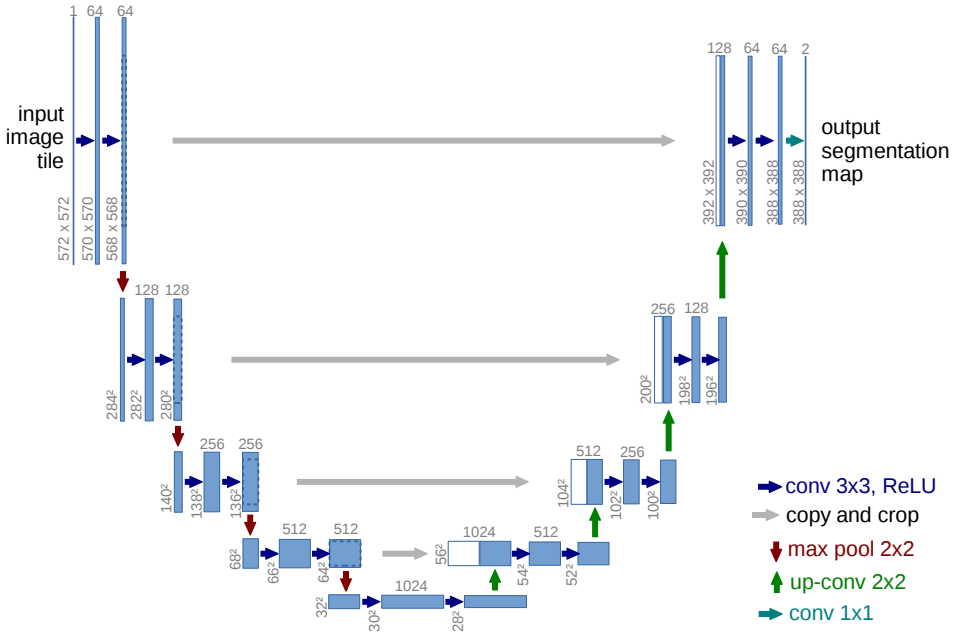


FIGURE 2.2 L'architecture U-Net [4]

Les méthodes DeepLab, proposées par Chen et al., introduisent des innovations majeures en segmentation sémantique [35]. L'usage de convolutions dilatées (ou *atrous convolutions*) permet d'élargir le champ réceptif des filtres convolutifs sans augmenter le nombre de paramètres ni le coût computationnel. Cela permet de contrôler la résolution spatiale de l'entrée à chaque couche et de capter des éléments de contexte global. DeepLab v2 introduit ensuite le module ASPP (*Atrous Spatial Pyramid Pooling*), qui applique des convolutions dilatées à plusieurs taux d'échantillonnage sur une même couche, pour une extraction dense de caractéristiques et une segmentation d'objets de taille variée [36]. DeepLab exploite également des CRF entièrement connectés pour améliorer la détection de contours. Avec DeepLab v3, les auteurs proposent une application cascadée ou en parallèle de convolutions dilatées ainsi qu'un module ASPP amélioré, éliminant le besoin de post-traitement par CRF [37]. Finalement, DeepLab v3+ (Figure 2.3) combine le module ASPP avec une structure encodeur–decodeur, permettant une meilleure récupération des détails fins tout en conservant la capacité d'extraction de contexte global [5]. Ces modèles, basés sur des backbones puissants (ResNet, Xception), atteignent des performances de pointe sur de nombreux benchmarks et sont devenus un standard dans les tâches de segmentation sémantique [27].

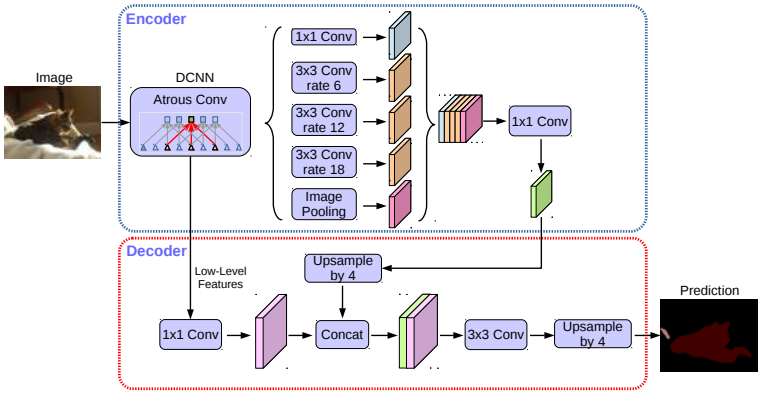


FIGURE 2.3 L'architecture DeepLabv3+ [5]

Plus récemment, les modèles transformeurs, comme SegFormer [38], ont redéfini l'état de l'art en exploitant des mécanismes d'auto-attention capables de capturer des relations à longue portée dans l'image. Ces approches offrent une excellente capacité de généralisation, mais nécessitent généralement de grands ensembles d'entraînement [39]. Cette évolution converge vers des modèles de plus en plus génériques, tels que Segment Anything [40], qui visent à fournir des capacités de segmentation *zero-shot* polyvalentes indépendantes des domaines d'application.

Tel que rapporté par Minae [27] et Xu [28], plusieurs autres architectures ont été proposées

dans la littérature au fil des années, comme les réseaux convolutifs régionaux (R-CNN) tels que MASK R-CNN pour la segmentation d'instances, les réseaux de neurones récurrents (RNN), les modèles adversariaux génératifs (GAN), les modèles pyramidaux multi-échelle (FPN), et les architectures auto-encodeurs, pour n'en citer quelques uns. Certains modèles combinent l'apprentissage profond à des méthodes traditionnelles ou d'apprentissage machine afin de tirer profit des avantages de chacun, donnant naissance à des approches hybrides. Par exemple, les CNN sont parfois employés de pair avec des modèles de contours actifs, des modèles probabilistes ou du post-traitement [27, 28].

Divers mécanismes peuvent être utilisés conjointement aux modèles pour améliorer les performances ou mitiger certaines difficultés, tels l'intégration de modules d'attention pour guider l'apprentissage, l'utilisation d'augmentations pour pallier à un manque de données ou encore la pondération par classe dans le cas de déséquilibre dans les jeux de données. Les fonctions de pertes utilisées couramment en segmentation d'images peuvent être basées sur les distributions, les régions, les frontières ou bien une combinaison de ces éléments. Dans la première catégorie, la *Cross-Entropy Loss* est la fonction de perte fondamentale, qui évalue les différences entre deux distributions et dont dérive la *Focal Loss* (avec pondération pour des classes déséquilibrées) ainsi que d'autres variantes. Les fonctions de pertes basées sur les régions visent à maximiser la superposition des régions segmentées et incluent la *Dice Loss*, la *Tsversky Loss* et la *Jaccard Loss*. Enfin, certaines fonctions permettent d'évaluer les distances entre des frontières ou combinent plusieurs approches selon l'application [41].

L'évaluation des modèles de segmentation repose principalement sur des métriques de précision quantitative. Parmi les métriques les plus utilisées, la *Pixel Accuracy* (PA) mesure la proportion de pixels correctement classés, tandis que la *Mean Pixel Accuracy* (MPA) en est une version moyennée par classe. L'*Intersection over Union* (IoU), ou indice de Jaccard, évalue le chevauchement entre la prédiction et la vérité terrain, et sa moyenne sur toutes les classes (Mean IoU) est devenue la référence pour évaluer la qualité d'une segmentation. La précision, le rappel et le F1-score permettent d'analyser les erreurs de classification. Enfin, le coefficient de Dice, très utilisé en imagerie médicale, mesure également le recouvrement entre prédiction et vérité terrain et est équivalent au F1-score dans le cas d'une segmentation binaire [27].

En mammographie, la segmentation occupe une place centrale dans les systèmes d'aide au diagnostic. Elle agit souvent à titre d'étape préliminaire en permettant d'isoler des structures d'intérêt ou l'extraction de caractéristiques, et intervient dans de nombreuses tâches dont la détection, la classification et le traitement du cancer du sein [42, 43].

2.1.2 Régression de points clés

Dans la plupart des applications modernes en vision par ordinateur, les réseaux de neurones se sont imposés comme outils de régression puissants. Leur capacité à modéliser des relations hautement non linéaires en fait des outils adaptés à des données complexes comme les images, les signaux ou les séquences. Dans le cadre de tâches de classification, les réseaux de neurones convolutifs sont généralement composés d'une succession de couches de convolution et de sous-échantillonnage, suivies de couches entièrement connectées et d'une couche de sortie de type Softmax, et sont optimisés par une fonction de perte entropique. Ces architectures peuvent être adaptées à des tâches de régression en remplaçant la couche de classification finale par une couche entièrement connectée produisant des sorties continues, associée à une fonction de perte adaptée, telle qu'une perte quadratique. Ces modèles se sont révélés particulièrement adaptés aux tâches de détection de points clés, telles que l'estimation de pose, la localisation de repères anatomiques ou la détection de points caractéristiques du visage, pour lesquelles la prédiction précise de coordonnées continues est essentielle [44].

Les méthodes de détection de points clés peuvent être basées sur la régression directe de coordonnées ou sur la régression de cartes de chaleur (*heatmaps*). Les premières estiment directement les coordonnées des points à partir de l'image, souvent à l'aide de réseaux de neurones convolutifs ou de cascades de CNN. Cependant, cette tâche reste difficile en raison de la forte non-linéarité du passage de l'espace image à l'espace des coordonnées, ainsi que des variations de vue, de pose ou d'illumination. Les approches basées sur les cartes de chaleur, de plus en plus utilisées, modélisent la probabilité d'emplacement des points clés sous forme de cartes, généralement une par point, permettant l'affinement progressif des estimations. Une architecture de base de régression de point clé par carte de chaleur est illustrée sur la figure 2.4. Dans le domaine médical, des architectures multi-étapes ou multi-tâches utilisent ces cartes pour localiser précisément des repères anatomiques. Comme pour la régression directe, ces modèles requièrent un nombre de points pré-déterminé, ce qui fixe la structure du tenseur de sortie [7].

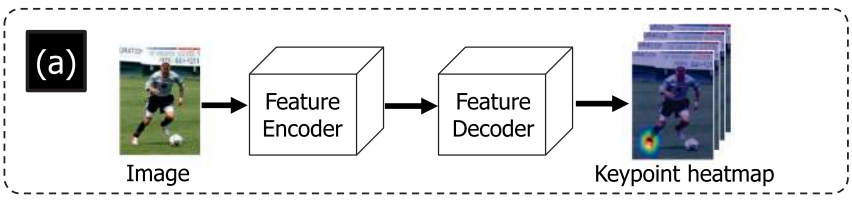


FIGURE 2.4 Cadre typique pour la régression de heatmaps [6]

Par exemple, Sharan et al. ont appliqué une méthode par régression de cartes de chaleur à la détection de points de suture [7]. Les auteurs proposent un modèle basé sur une architecture U-Net, avec des activations ReLU pour les couches convolutionnelles et des images RGB de taille 288×512 en entrée. Les positions des points de suture sont représentées par une carte de chaleur unique, obtenue en appliquant une fonction de distribution centrée sur chaque point annoté, suivie d'un filtre Gaussien différentiable pour lisser les prédictions. Le réseau est entraîné à minimiser une perte de similarité entre la carte de chaleur prédite et la vérité terrain, combinant l'erreur quadratique moyenne et le coefficient de Dice. L'architecture comprend également une couche Soft-Argmax convolutionnelle différentiable qui agit comme une suppression locale des maxima, permettant d'extraire les coordonnées finales des sutures tout en traitant plusieurs points sur un seul canal. Deux variantes sont explorées : une utilisant une carte de chaleur comme vérité terrain et l'autre un masque binaire. L'apprentissage est guidé par une fonction de perte composite qui contraint le réseau à produire des cartes de probabilité à la fois précises et spatialement cohérentes, en combinant une mesure d'erreur quadratique et des critères de similarité afin de gérer le fort déséquilibre entre pixels correspondant aux points de suture et l'arrière-plan. Le modèle final optimise conjointement ces pertes et les coordonnées prédites sont obtenues à partir du centre de masse des composantes connexes obtenues après seuillage. Un schéma de l'architecture proposée est présenté à la figure 2.5

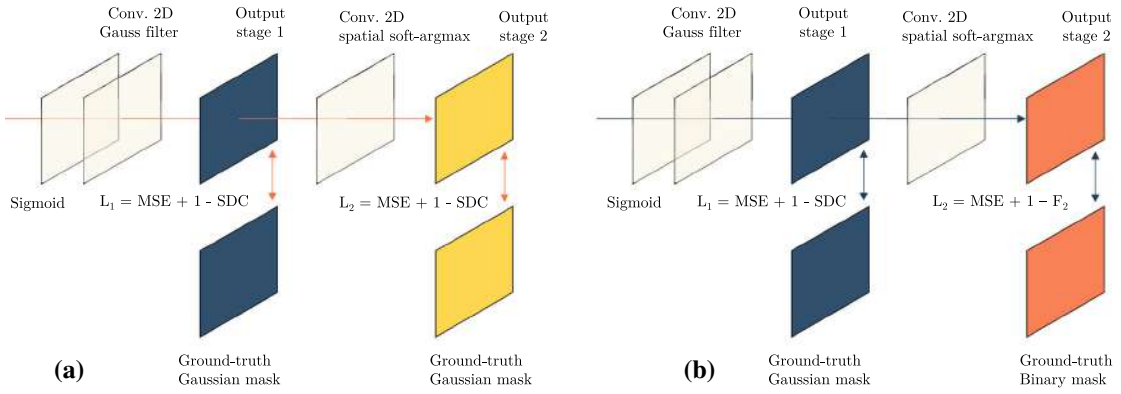


FIGURE 2.5 Les variantes d'architecture proposées dans [7]

Le modèle proposé atteint des performances solides sur des données intra-opératoires et simulées, avec des scores F1 allant jusqu'à 0,48 en contexte intra-opératoire et 0,77 sur des données de simulation, malgré la forte hétérogénéité visuelle et la présence d'artefacts. Ces résultats démontrent la capacité du cadre de régression par cartes de chaleur multi-instance à localiser de manière fiable un nombre variable de points dans des scènes complexes. La formulation générique du modèle, indépendante d'un nombre fixe de points, suggère par

ailleurs un potentiel de généralisation à d'autres tâches de détection de points multiples en vision par ordinateur et en imagerie médicale.

2.2 Vision par ordinateur en mammographie

La mammographie constitue la modalité de référence pour le dépistage du cancer du sein. Toutefois, son interprétation demeure complexe et sujette à plusieurs limitations relatives notamment à l'interprétation humaine [45]. Afin de répondre à ces enjeux, de nombreuses méthodes de vision par ordinateur ont été développées au fil du temps pour assister l'analyse des mammographies. Ces approches ont progressivement évolué de techniques classiques de traitement d'images et d'apprentissage automatique vers des modèles d'apprentissage profond capables d'extraire automatiquement des représentations pertinentes à partir des données. Après un recensement des bases de données en mammographie, cette section propose une revue des principales méthodes de vision par ordinateur appliquées à la mammographie, en traitant d'abord brièvement les approches dédiées à la détection et à la classification du cancer du sein, puis en abordant la détection de repères anatomiques.

2.2.1 Bases de données

Il existe plusieurs ensembles de données variés relatifs au cancer du sein, couvrant différentes modalités d'imagerie, notamment la mammographie, l'échographie, l'imagerie par résonance magnétique, l'histopathologie, et la thermographie [46]. Dans le cadre de ce travail, on s'intéresse aux données spécifiques à la mammographie. Pour cette modalité, un des premiers jeux de données publics est le *Mammographic Image Analysis Society Mammography Database* (MIAS), publié en 1994 au Royaume-Unis. En 1999, le *Digital Database for Screening Mammography* (DDSM) est rendu public aux États-Unis, avec plus de 2000 examens. Ces deux ensembles sont composés d'images numérisées de mammographies sur film. En 2012, les jeux de données InBreast et BCDR (*Breast Cancer Digital Repository*) sont développés au Portugal. Un sous-ensemble de DDSM n'incluant que des cas cancéreux, le *Curated Breast Imaging Subset of DDSM* (CBIS-DDSM) est proposé en 2017. En 2020, un ensemble d'images provenant d'un programme de dépistage en Suède est compilé pour former le jeu de données CSAW. Au Royaume-Unis, un des plus grands jeux de données en mammographie à ce jour est publié en 2021, OPTIMAM, comprenant plus de 2 millions d'images. La *Chinese Mammography Database* (CMMD) est rendue disponible la même année, compilant près de 2000 examens provenant de divers centres en Chine. En 2022, les jeux de données VinDr et ADMANI sont publiés au Vietnam et en Australie respectivement. Le *RSNA Screening Mammography Breast Cancer Detection Dataset* (RSNA - SMBC) est une base de données pu-

blique de mammographies numériques organisée par la Radiological Society of North America (RSNA), rendue disponible en 2023. Cette dernière comprend les images au format DICOM de près de 20 000 examens. Enfin, en 2025, un ensemble d'examens tirés d'un programme de dépistage canadien est compilé, dans l'objectif de refléter plus fidèlement la distribution des cas en contexte de dépistage (*NL-Breast Screaning*).

Il existe donc plusieurs jeux de donnée publics composés d'images tirées d'examens de mammographie. Cependant, aucun des jeux de données cités ci-dessus n'incluent d'annotations relatives au positionnement des seins [46–49].

Les principaux jeux de données publics relatifs à la mammographie sont présentés dans le tableau 2.1.

Tableau 2.1 Principaux jeux de données publics de mammographie

Jeu de données	**Pays, année**	**Nb. examens**	**Nb. images**	**Annotations**
MIAS [50]	Royaume-Uni, 1994	161	322	Masses, calcifications, asymétrie, distorsion
DDSM [51]	États-Unis, 1999	2620	10 480	Masses, calcifications
InBreast [52]	Portugal, 2012	115	410	Masses, calcifications, asymétrie, distorsion
BCDR [53]	Portugal, 2012	1734	7315	Masses, calcifications, asymétrie, distorsion
CBIS-DDSM [54]	États-Unis, 2017	1566	6264	Masses, calcifications
CSAW-CC [55]	Suède, 2020	24 788	98 788	Tumeurs
OPTIMAM [56]	Royaume-Unis, 2021	–	3 072 878	Lésions
CMMD [57]	Chine, 2021	1775	5202	Masses, calcifications
VinDr-Mammo [47]	Vietnam, 2022	5000	20 000	Masses, calcifications, asymétrie, distorsion, autres
ADMANI [58]	Australie, 2022	629 863	4 411 263	Classification (cancéreux / non-cancéreux)
RSNA-SMBC [59]	États-Unis, Canada, 2023	20 000	54 713	Classification (cancéreux / non-cancéreux)
NL-Breast-Screening [60]	Canada, 2025	5997	23 988	Classification (confirmée par biopsie)

2.2.2 Détection et classification du cancer du sein

Bien que la vision par ordinateur en mammographie ouvre la voie à un large éventail d'applications, la majorité des recherches se sont concentrées sur la détection et la classification des lésions mammaires [14]. Cette application a d'ailleurs fait l'objet de nombreux articles de revues [10, 14, 45, 46, 48, 61, 62].

Les systèmes de détection et de diagnostic assistés par ordinateur (CADe/CAD), basés sur des méthodes de traitement d'image traditionnelles et d'apprentissage machine, ont constitué les premiers outils développés pour automatiser l'analyse des mammographies [48]. De tels systèmes reposent généralement sur le prétraitement des images, la segmentation et la sélection des régions d'intérêt, l'extraction de caractéristiques, puis la classification à l'aide d'algorithmes d'apprentissage automatique [45,48]. Plusieurs facteurs limitent encore la généralisation de ces approches : les ensembles de données sont souvent restreints et peu diversifiés, les classes peuvent être déséquilibrées, la qualité des images varie selon le contraste, la résolution ou l'appareil utilisé, et des biais peuvent être introduits lors du prétraitement ou de la sélection des caractéristiques. De plus, l'évaluation de cas limites reste souvent insuffisante, limitant l'application clinique de ces modèles [45]. Enfin, les systèmes CAD conventionnels offrent souvent une faible spécificité et restent fortement dépendants de la qualité de la segmentation, du choix des caractéristiques et de la taille des bases de données, ce qui a motivé la transition vers des approches basées sur l'apprentissage profond [48, 61].

L'essor de l'apprentissage profond et l'accès grandissant à des données publiques a introduit de nombreuses avancées par rapport aux systèmes CAD conventionnels. Contrairement à ces derniers, les systèmes basés sur l'apprentissage profond intègrent automatiquement l'extraction des caractéristiques et la classification dans une seule phase d'apprentissage [48]. Plusieurs architectures ont été appliquées à la mammographie, tels des réseaux de neurones convolutifs (CNN), l'apprentissage par transfert, l'apprentissage ensembliste et des méthodes basées sur l'attention [62]. En effet, les CNN sont largement utilisés dans la littérature pour la détection et la classification de tumeurs mammaires [63–65]. Une étude publiée en 2022 a démontré l'intérêt d'affiner des modèles, employant un ResNet50 *fine-tuned* pour la classification de tumeurs et atteignant une précision de 99,96 % [66]. Selon Wang [62], les CNN permettent de réduire les taux de faux positifs, mais leur efficacité dépend largement de la quantité et de la qualité des données d'entraînement. Dans des contextes de données limitées, l'apprentissage par transfert (*transfer learning*) exploite des modèles pré-entraînés sur une tâche donnée, pour améliorer les performances sur une autre tâche, souvent avec peu de données. Cette approche a été explorée entre autres par Walayat et al. [67], qui ont testé plusieurs architectures de CNN, soient Inception-v3, ResNet-50, VGG-16, SqueezeNet et AlexNet, ce

dernier atteignant une précision de 96,7 % dans la détection du cancer du sein sur leur jeu de données. Par ailleurs, les méthodes d'apprentissage ensembliste (*ensemble learning*) combinent plusieurs modèles d'apprentissage profond pour accroître la robustesse et la précision des prédictions, via des stratégies telles que le *bagging*, le *boosting* ou le *stacking* [62]. Par exemple, Nemade et al. [68] ont obtenu une précision de 98,1% sur le jeu de données publiques DDSM dans la tâche de classification des lésions mammaires, démontrant le potentiel des modèles ensemblistes. Les mécanismes d'attention, quant à eux, permettent aux modèles de se concentrer sur les régions les plus pertinentes des images [46, 62]. Bobowicz et al. [69] emploient à ce titre des suggestions graphiques de cartes d'attention pour guider l'apprentissage d'un modèle de classification des lésions mammaires, basé sur un apprentissage faiblement supervisé à instances multiples.

Certains auteurs explorent également l'utilisation de modèles transformeurs et les modèles basés sur les graphes, qui peuvent se révéler très performants mais qui nécessitent un plus grand nombre de données d'entraînement [70]. Des techniques d'apprentissage semi-supervisé et non-supervisé sont par ailleurs de plus en plus employées [71, 72].

Malgré les gains de performances obtenus avec les modèles d'apprentissage profond, des limitations subsistent, dont l'accès limité à des données annotées, et la généralisation à d'autres bases de données ou d'autres appareils d'acquisition d'images. De plus, l'interprétabilité des modèles est un obstacle majeur à leur adoption clinique [46, 48].

2.2.3 Détection de repères anatomiques

La détection automatique de certains repères anatomiques constitue une tâche importante de l'analyse des mammographies. Des structures telles que le muscle pectoral, le mamelon ainsi que les tissus fibroglandulaire et adipeux jouent un rôle central dans de nombreuses étapes du traitement et de l'interprétation des images. En effet, celles-ci peuvent permettre l'évaluation de la qualité des images, la localisation de lésions, et l'estimation de la densité mammaire, contribuant à l'évaluation du risque de cancer du sein [73]. Cette section se concentre plus spécifiquement sur les méthodes dédiées à la détection du muscle pectoral et du mamelon.

Détection du muscle pectoral

Dans la littérature, la détection du muscle pectoral est largement abordée, motivée par divers objectifs comme le retrait de zones non pertinentes, le pré-traitement ou encore l'évaluation du positionnement. Le muscle pectoral étant généralement caractérisé par une zone de forte intensité et une géométrie relativement régulière, sa détection est souvent formulée comme une

tâche de segmentation. Les premières méthodes proposées pour la détection et la segmentation du muscle pectoral sont basées sur des méthodes classiques de traitement d'image, reposant sur des éléments géométriques et statistiques, tel que rapporté par Angelone et al. [74].

De nombreuses approches sont basées sur l'intensité, exploitant le contraste entre le muscle pectoral et le tissu mammaire. Ces méthodes s'appuient principalement sur des techniques de seuillage, parfois combinées à des heuristiques ou des opérations morphologiques [75–82]. Parallèlement, des méthodes fondées sur les régions, qui visent à regrouper des pixels homogènes selon des critères d'intensité, de texture ou de similarité statistique, ont également été proposées. À ce titre, Camilus et al. [83] exploitent l'algorithme de Watershed pour segmenter le muscle pectoral à partir de la topographie de l'image, alors que Chen et al. proposent une méthode par croissance de régions [84].

Plusieurs études utilisent les informations de textures et de formes [85–90]. D'autres travaux s'appuient sur la détection de contours et de lignes, approximant parfois la frontière du muscle par une fonction linéaire. Ture et al. [91] proposent un algorithme de détection de contours s'appuyant sur des informations topologiques, tandis que Chen et al. [92] utilisent l'algorithme de Canny pour extraire les contours avant d'appliquer des étapes de post-traitement. Zeng adapte l'algorithme Smith-Waterman pour la détection de frontière [93], tandis que Kinoshita et al. [94] détectent des lignes droites dans le domaine de Radon. La transformée de Hough est fréquemment employée pour détecter la ligne associée à la frontière pectorale, notamment dans les travaux de Kwok et al. [95], de Ferrari et al. [96], de Wei et al. [97] et de Firdi et al. [98]. D'autres combinent détection de contours et opérations morphologiques, telles que les cartes topographiques et l'enveloppe convexe, afin de raffiner la segmentation [99]. Shen et al. [100] estiment la frontière du muscle par une approximation polynomiale, en appliquant un algorithme génétique et des opérations morphologiques. Des méthodes statistiques ont également été appliquées à la détection du muscle pectoral [101–103]. Par exemple, Ma et al. [102] ont combiné la théorie des graphes à des modèles de contours actifs.

Certains travaux ont introduit des techniques d'apprentissage automatique classique pour la segmentation du muscle pectoral. Feudjio et al. [104] proposent une approche basée sur le regroupement fuzzy c-means, Pawar et al. [105] utilisent une stratégie de parcours en profondeur, et Priyanka et al. [106] explorent l'utilisation de divers algorithmes d'apprentissage automatique. Ces derniers atteignent 93,3% de précision sur le jeu de données MIAS. Ces méthodes classiques, bien que simples et interprétables, sont sensibles aux variations morphologiques et d'intensité ainsi qu'au bruit. Elles sont également peu généralisables. Par ailleurs, tel que souligné dans [74], ces méthodes présentent des performances limitées lorsque la frontière entre le muscle et le tissu mammaire est masquée par des tissus denses ou des plis, ou

dans le cas d'une forte courbure du muscle pectoral.

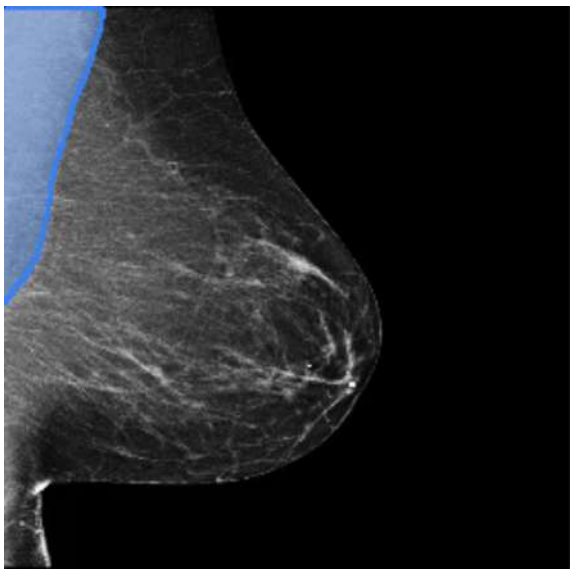
Afin de surmonter ces limites, certains chercheurs se penchent sur le développement de modèles d'apprentissage profond. L'usage de CNN pour la segmentation du muscle pectoral est fortement exploré dans la littérature. Dubrovina et al. [107] développent des réseaux de neurones convolutifs pour une segmentation multi-classe des tissus du sein, incluant le muscle pectoral. Guo et al. [108] proposent un CNN pour distinguer en premier lieu le muscle pectoral du tissu glandulaire, puis raffinent en second lieu la forme de la prédiction par un GAN. Ils obtiennent des scores Dice de 85.2 %, 94.8%, et 98.1% sur trois groupes d'un jeu de données privés, surpassant un U-Net à titre de référence. Soleimani et al. [109] combinent un CNN basé sur VGG16 avec l'algorithme de Dijkstra pour contraindre la segmentation le long de chemins optimaux. Rampun et al. [110] modifient un réseau de neurones convolutifs profond conçu pour la détection de contours, le Holistically-nested Edge Detection (HED) Network. Klaneck et al. utilisent pour leur part un ResNet18 et explorent l'utilisation d'un algorithme MonteCarlo pour évaluer l'incertitude des prédictions [111]. Le paradigme de distillation des connaissance est appliqué par Huemmer et al. [112] pour la segmentation du muscle pectoral, afin de transférer les performances de modèles lourds vers des architectures plus légères. Ils ont adapté l'architecture DenseNet-121 pour prédire la frontière du muscle pectoral par régression d'un vecteur colonne-index, de façon supervisée et semi-supervisée. Outre Dubrovina [107], d'autres chercheurs s'attaquent à la segmentation multi-classes des régions du sein [113–115]. Sierra-Franco et al. [116] testent plusieurs architectures, dont U-Net et FPN, pour segmenter les régions du sein dans les vues CC et MLO. Ils obtiennent dans le cas du muscle pectoral un IoU atteignant 0.85 et 0.96 pour ces deux vues avec leur modèle FPN. En ce qui concerne la segmentation du muscle pectoral, peu d'études ont inclu la vue CC, puisque le muscle n'y est pas toujours visible. Silva et al. [117] s'y sont cependant attardés, adoptant une approche centrée sur les données basée sur des FPN, mettant l'accent sur la qualité et la diversité des données d'entraînement.

Un grand nombre de travaux s'appuient sur des architectures de types U-Net. En effet, plusieurs études utilisent ou adaptent cette architecture pour la segmentation du muscle pectoral. Ma et al. [118] développent une architecture convolutive profonde (DCNN) inspirée du U-Net et montrent une amélioration notable par rapport à des méthodes antérieures basées sur les textures. Liu et al. [119] emploient un U-Net pré-entraîné et un post-traitement pour prédire le masque de segmentation. Dogan et al. [120] entraînent un réseau de type U-Net utilisant MobileNetV2 comme architecture de base. Dans une récente étude, Angelone et al. [74] implémentent un U-Net, en fournissant comme entrée des patches entourant la frontière du muscle pectoral, estimée au préalable par des méthodes classiques de traitement d'images. Des variantes architecturales du U-Net ont également été appliquées à la segmentation du

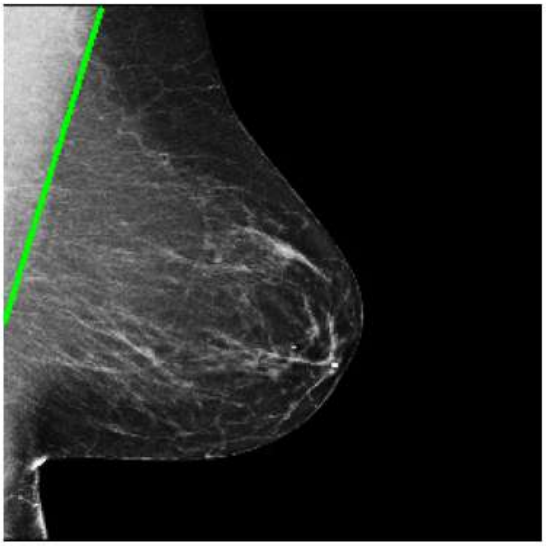
muscle pectoral, tels l'Attention U-Net [121], ou encore des U-Net intégrant des couches densément connectées [122]. Ali et al. [123] implémentent par ailleurs un réseau entièrement convolutif, s'apparentant à un ResUNet auquel ils intègrent des connexions résiduelles. La validation croisée menée sur trois ensembles de données publics (MIAS, INbreast et DDSM) a permis d'obtenir un IoU moyen de 97 %, un coefficient de similarité de Dice de 96 % et une précision de 98 %.

Les architectures de segmentation sémantique de type DeepLabv3 et DeepLabv3+, associées à différents backbones, ont également été appliquées à la tâche de segmentation du muscle pectoral [124, 125]. Yu et al. [126] proposent PeMNet, un modèle basé sur DeepLabv3+ intégrant un module d'attention spécifique pour améliorer la segmentation du muscle pectoral. Ils rapportent un IoU de 97,3% sur un ensemble de test tiré des jeux de données publics OPTITAM et InBreast.

Certaines études proposent non pas une segmentation de la région du muscle pectoral, mais bien la détection de la ligne droite délimitant celui-ci (voir figure 2.6). Cette ligne correspond en effet à un repère pertinent pour l'évaluation du positionnement mammaire, soit la PL, et permet l'obtention d'un autre élément important : la PNL. Il s'agit de l'objet de l'étude de Tran et al. [90], qui détectent la ligne du muscle pectoral et le mamelon dans le but d'identifier la PNL et d'en mesurer la longueur. Les auteurs emploient des filtres de Gabor pour identifier la frontière du muscle pectoral. De même, Gupta et al. [127] s'attardent à la prédiction de la PNL, employant d'abord un réseau Inception-V3 pour la détection de la ligne du muscle pectoral. La détection de la ligne du muscle pectoral est aussi une étape préliminaire à l'évaluation de certains critères de positionnement dans l'article de Tanyel et al. [128]. Ils implémentent pour ce faire un modèle permettant la détection simultanée de la ligne du muscle pectoral et du mamelon, CoordAttUNet, et rapportent des erreurs moyennes de 5,62 mm et 6,69 mm sur les points inférieurs et supérieurs de la ligne.



(a) Segmentation du muscle pectoral (SAM2)



(b) Détection de la ligne du muscle pectoral

FIGURE 2.6 Différence entre la segmentation du muscle pectoral et la détection de la ligne du muscle pectoral

Le pré-traitement des images constitue généralement une étape importante pour la segmentation du muscle pectoral par des modèles d'apprentissage profond. La détermination de régions d'intérêt, le rehaussement du contraste ainsi que la réduction du bruit sont souvent effectués [109, 123–125].

En somme, la segmentation du muscle pectoral a été un sujet très abordé dans la littérature, évoluant de méthodes traditionnelles de traitement d'images et d'apprentissage automatique vers des modèles d'apprentissage profond. On constate que les variantes de U-Net et de DeepLab constituent aujourd'hui l'état de l'art, offrant des performances très satisfaisantes. Toutefois, certains défis subsistent, notamment en termes de robustesse face à la variabilité morphologique, de dépendance aux annotations expertes et de reproductibilité [121, 126]. En outre, jusqu'à tout récemment, un problème majeur dans ce domaine était la quantité limitée de données publique comportant des annotations pour la segmentation du muscle pectoral, représentant un défi pour l'apprentissage de modèles supervisés. En 2024, Alynia et al. [121] ont rendu publiques les masques de segmentation du muscle pectoral pour les jeux de données INbreast, MIAS, CBIS-DDSM, ce qui constitue une contribution importante pour les travaux futurs. En ce qui concerne les annotations de la ligne du musque pectoral (par opposition au masque de segmentation), seul un sous-ensemble d'images MLO provenant du jeu de données VinDr est annoté, ayant été rendu publique par Tanyel et al. [128] en 2024. La majorité des

études présentées ci-dessus ayant été publiées plus tôt, elles ont souffert de ce manque de données publiquement accessibles.

Détection du mamelon

La détection automatique du mamelon constitue une étape clé de l'analyse des mammographies. En effet, le mamelon est un repère anatomique important, notamment pour l'alignement et entre les vues CC et MLO, l'évaluation de la qualité du positionnement, ainsi que la localisation relative des structures mammaires et des lésions [129]. En pratique, la détection du mamelon peut être formulée comme un problème de segmentation, de détection d'objet ou encore de régression d'un point. Il s'agit d'une tâche complexe, car les morphologies sont variées et le mamelon n'est pas toujours visible.

Les premières approches de détection du mamelon reposent majoritairement sur des caractéristiques conçues par l'humain, exploitant des propriétés géométriques, d'intensité ou de textures. Yin et al. [130] identifient le mamelon comme le point d'intensité moyenne maximale sur le contour du sein, atteignant une erreur moyenne inférieure ou égale à 10 mm sur un petit ensemble d'images. Mendez et al. [131] utilisent une approche fondée sur les variations de gradient, avec une erreur géométrique moyenne de 13,5 mm. Chandrasekhar et al. [132] proposent une méthode basée sur le calcul de la courbure le long du contour du sein, la position du mamelon correspondant au pixel de courbure maximale. Zhou et al. [133] et Kinoshita et al. [94] exploitent la convergence des structures glandulaires vers le mamelon, obtenant des précisions respectives de 0,85 et 0,793 sur des bases de plusieurs centaines à plus d'un millier d'images. Des approches basées sur des atlas multiples ont également été proposées, notamment par Iglesias et Karssemeijer [134]. L'analyse de l'intensité des pixels est combinée à des opérations morphologiques dans l'étude publiée par Mustra et al. [135]. Casti et al. proposent une méthode de détection du mamelon basée sur la matrice hessienne, combinant une analyse des champs de gradient à des contraintes géométriques [136]. Jas et al. [137] proposent une approche heuristique estimant la position du mamelon comme le point de la frontière le plus éloigné de la base du sein. Dans [93], une méthode est proposée pour détecter les mammelons visibles et invisibles (ainsi que le muscle pectoral) dans la vue MLO, en identifiant la zone convexe ou le point présentant la courbure maximale le long du contour du sein. Ces méthodes, basées sur des descripteurs conçus par l'humain, sont sensibles aux variations d'intensité et de qualité des images [73].

Certains travaux intègrent des éléments d'apprentissage automatique, comme Jiang et al. [138], qui proposent une détection du mamelon basée sur des caractéristiques radiomiques combinées à un Random Forest, atteignant une précision de 0,964 sur 1 442 images. Lin et

al. [129] introduisent ensuite un modèle d'apprentissage profond basé sur la classification de patchs pour localiser les mamelons visibles : la position prédite correspond au patch ayant le recouvrement détecté le plus élevé. Atteignant une précision de 98,0 % (une boîte englobante prédite dont le centre se situe à moins de 10 mm du centre réel du mammelon étant considérée acceptable) sur la base de données DDSM, cette approche ne fournit cependant pas de segmentation du mamelon, générant simplement une boîte englobante autour de celui-ci. Tel qu"abordé dans la section précédente, plusieurs travaux s'intéressent également à la détection conjointe du mamelon et du muscle pectoral. Tran et al. [90] emploient l'algorithme de watershed et la transformée de Hough circulaire pour détecter le mamelon ainsi qu'un filtre de Gabor pour la détecter la frontière du muscle pectoral. À partir de ces deux détections, la PNL, qui constitue un élément clé pour l'évaluation du positionnement, est obtenue. Les auteurs rapportent une erreur moyenne de 6.36 mm sur la longueur de la PNL prédite. Gupta et al. [127] se penchent également sur la prédiction de la PNL, employant un réseau Inception-V3 pour la détection de la ligne du muscle pectoral et un l'algorithme de détection de cercles par transformée de Hough pour localiser le mamelon. Leur approche de détection du mamelon repose toutefois sur la présence d'un marqueur de plomb identifiant la position du mamelon, ce qui n'est pas toujours le cas en pratique. Garcia et al. [139] proposent une approche itérative combinant la détection du muscle pectoral et du mamelon. Pour ce dernier. l'erreur moyenne obtenue est de moins de 10 mm. D'autres études visent l'évaluation de certains critères de positionnement, comme Zeng et al. [140], qui proposent un modèle CNN permettant de détecter simultanément la position du mamelon et d'évaluer s'il est en profil, en exploitant notamment des cartes d'attention (Grad-CAM) issues d'un modèle de classification. Tanyel et al. [128] s'intéressent également à l'évaluation du positionnement des seins en vues MLO, en passant par la détection du mamelon, du muscle pectoral et de la PNL. Ils développent à cette fin le modèle CoordAttUNet, qui permet une détection simultanée de la ligne du muscle pectoral et du mamelon. Il s'agit d'une architecture semblable à un U-Net, à laquelle est intégrée un module d'attention spécifique à la localisation de coordonnées. Ils obtiennent une erreur moyenne de 2.97 mm pour la détection du mamelon sur un sous-ensemble du jeu de données VinDr. Leur approche d'évaluation du positionnement sera expliquée en plus grand détail dans la section suivante.

Tel que mentionné dans la section précédente, certaines publications portent sur la segmentation multi-classe. Sierra-Franco et al. [116], par exemple, évaluent plusieurs modèles d'apprentissage profond et obtiennent un IoU moyen allant de 0.73 à 0.74 sur la vue CC et de 0.72 à 0.75 en vue MLO, selon le modèle. S'appuyant sur ce travail, Rogozinski et al. [73] ont récemment proposé une méthode de segmentation du mamelon, reposant sur une localisation initiale du mamelon par un modèle de détection d'objets, un recentrement puis une

segmentation. Leur approche surpasse la précédente, avec un IoU passant à 0,80 et 0,78 sur les vues CC et MLO respectivement.

Dans l'ensemble, l'évolution des méthodes de détection du mamelon reflète une transition progressive des approches traditionnelles basées sur l'intensité, les contours et la géométrie vers des modèles d'apprentissage profond. Ces derniers, bien qu'améliorant les performances par rapport aux approches basées sur des descripteurs conçus manuellement, sont cependant moins adaptés à des contextes de données limitées [90]. Comme le soulignent Rogozinski et al. [73] et Tanyel et al. [128], le mamelon demeure l'une des structures anatomiques les plus difficiles à détecter et segmenter de manière fiable, en raison de sa taille réduite, de son contraste variable et de sa variabilité anatomique. Une détection robuste du mamelon reste néanmoins essentielle pour de nombreuses applications en mammographie.

2.3 Évaluation automatique de la qualité du positionnement mammaire

Si un grand nombre de travaux ont été dédiés à la classification et à la détection du cancer du sein ainsi qu'à la détection du muscle pectoral et du mamelon, moins d'attention a été portée spécifiquement à l'évaluation du positionnement des seins en mammographie. Cette section résume les approches proposées, des méthodes de traitement d'image classiques aux méthodes d'apprentissage profond.

2.3.1 Approches basées sur le traitement d'image classique

Les premières méthodes développées pour l'évaluation automatique du positionnement mammaire reposaient principalement sur des techniques de traitement d'images classique et d'analyse statistique [141]. Une des premières approches a été proposée dans [142] et retravaillée en [143], articles dans lesquels les auteurs présentent divers algorithmes appliqués à la vue MLO. Plusieurs éléments de positionnement sont analysés, notamment l'exclusion du tissu mammaire, la présence du mamelon de profil, l'inclusion de l'angle infra-mammaire ainsi que le positionnement du muscle pectoral. D'abord, l'exclusion de tissus aux bords supérieur et inférieur de l'image est détectée par l'analyse de l'orientation locale du contour du sein. Cependant, l'exclusion potentielle de tissus au niveau du bord postérieur de l'image n'est pas évaluée, car aucune segmentation de la glande mammaire n'est obtenue. Puis, la méthode proposée pour déterminer si le mamelon est de profil repose sur l'analyse des variations du gradient moyen d'intensité le long de la normale à l'interface du sein. Le mamelon est considéré de profil si la variation du gradient dépasse un certain seuil. L'inclusion de l'angle infra-mammaire est vérifiée par l'analyse de la concavité du contour du sein le long du coin

inférieur. Enfin, le positionnement du muscle pectoral est évalué à partir de la segmentation obtenue dans [95]. La qualité de l'évaluation dépend donc de la segmentation adéquate du muscle. Les auteurs appliquent leurs algorithmes sur les images de vue MLO du jeu de données MIAS afin de déterminer la prévalence des cas de mauvais positionnement. Toutefois, aucune précision en termes d'évaluation du positionnement n'est présentée, faute de vérité terrain. Moran et al. [144] proposent également des algorithmes permettant d'évaluer le positionnement du muscle pectoral ainsi que la vue de profil du mamelon. La détection du muscle pectoral est effectuée à l'aide de la transformée de Hough, tandis que la détection du mamelon repose sur la squelettisation. Cette méthode permet de mettre en évidence la forme globale du sein et ses protubérances, la position du mamelon pouvant être identifiée à partir des ramifications du squelette. Les auteurs rapportent des précisions de 0,8 et 1 pour la détection du muscle pectoral et du mamelon respectivement. Toutefois, aucune précision quant à l'évaluation des critères de positionnement n'est rapportée.

Ces méthodes, bien que simples et rapides, sont peu généralisables et souffrent du manque de données annotées pour la tâche de l'évaluation du positionnement.

2.3.2 Approches basées sur l'apprentissage profond

Avec l'essor de l'apprentissage profond, la recherche concernant l'évaluation automatique du positionnement mammaire évolue. On voit naître deux principaux types d'approches : les approches basées sur des modèles de classification directe ainsi que les approches basées sur la détection de repères anatomiques.

D'une part, les réseaux de neurones convolutifs peuvent être utilisés pour effectuer directement la classification des mammographies selon certains critères de positionnement. C'est ce que proposent Brahim et al. [1], qui entraînent des réseaux convolutifs peu profonds pour effectuer la classification binaire des images selon six critères de positionnement. Parmi les critères considérés, quatre sont relatifs à la vue MLO et deux sont propres à la vue CC. Les auteurs ne spécifient pas l'origine exacte de ces critères, mais ils s'apparentent à ceux proposés par l'ACR. Ces critères ainsi que leur description telle que présentée dans l'article sont consignés dans le tableau 2.2.

Pour chacun de ces critères, un réseau spécifique est entraîné et optimisé. L'entraînement, la validation et le test des modèles a été faite à partir de 778 examens, issus de la base publique INbreast ainsi que de six hôpitaux allemands. Les images ont été classées selon les différents critères de positionnement par les auteurs, puis cette annotation a été validée par un radiologue disposant de cinq années d'expérience. Le nombre final d'examens a été obtenu suite à un tri des examens initiaux, les cas complexes ayant été retirés. En effet, les

Tableau 2.2 Critères de positionnement considérés par Brahim et al. [1]

Vue	**Critère**	**Description**
MLO	Angle du muscle pectoral	Le bord du muscle pectoral doit être bien visible, de forme triangulaire ou légèrement concave, avec un angle supérieur à 10 degrés.
	Niveau du muscle pectoral	Le bord inférieur du muscle pectoral doit se situer au niveau de la ligne PNL ou en dessous.
	Position du mamelon	Le mamelon doit être vu de profil
	Inclusion complète du tissu mammaire	Tout le tissu mammaire doit être visible, incluant la graisse rétro-glandulaire et la queue axillaire
CC	Position du mamelon	Le mamelon doit être vu de profil. Le mamelon ne doit pas être désorienté de plus de 20 degrés par rapport à l'horizontale.
	Inclusion complète du tissu mammaire	Tout le tissu mammaire doit être visible

examens de patientes ayant subi une opération ou possédant des implants mammaires n'ont pas été inclus dans le jeu de données. De plus, les images pour lesquelles le mamelon était difficile à identifier, pour des raisons morphologiques ou à cause d'un mauvais contraste, ont été retirées. Par ailleurs, la suppression des étiquettes indiquant la vue et le côté de l'image ont été retirés lors du prétraitement des images, et des techniques d'augmentation de données ont été employées. Une visualisation Grad-Cam a également été implémentée pour faciliter l'interprétation des prédictions. Dans la méthodologie proposée, deux architectures séquentielles sont employées avec, dans tous les cas, une fonction de perte entropique. Il est à noter que bien que le critère *Position du mamelon* en CC inclut deux éléments, soient la visibilité du mamelon de profil et son orientation, seul le premier élément a finalement été évalué, pour cause de manque de cas présentant une orientation inadéquate du mamelon. Chaque critère est évalué séparément par un modèle donné, puis le tout est combiné pour former un modèle global. Les auteurs rapportent une précision allant de 93.3 % à 98.2 % selon le critère, ainsi qu'une précision globale de 96,5 % dans la vue CC et de 93,3 % dans la vue MLO. Un schéma du modèle proposé pour l'évaluation de l'angle du muscle pectoral est présenté sur la figure 2.7.

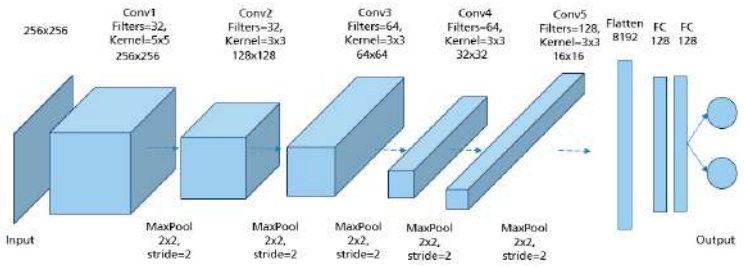


FIGURE 2.7 Architecture proposée dans [1]

Malgré les bons résultats obtenus par Brahim et al., cette étude présente plusieurs limites. Tout d'abord, certains cas difficiles ont été exclus du jeu de données, et une partie du jeu de données utilisé est privé, ce qui limite la reproductibilité et la généralisation des résultats. Par ailleurs, bien que plusieurs critères de qualité du positionnement soient évalués, certains critères cliniquement pertinents ne sont pas pris en compte, comme la présence de plis de peau et d'artéfacts. De plus, certains critères sont combinés en un seul. En outre, malgré l'utilisation de Grad-CAM pour l'interprétabilité, l'explicabilité du modèle demeure limitée. Bien que les régions de l'image sur lesquelles le réseau porte son attention soient mises en évidence, aucune information géométrique ou d'explication des décisions du modèles ne sont fournies. Enfin, l'évaluation a été réalisée sur un jeu de test artificiellement équilibré, qui ne reflète pas les forts déséquilibres de classes observés en pratique clinique, ce qui pourrait conduire à une surestimation des performances dans des conditions réelles.

Watanabe et al. [145] se penchent quant à eux sur l'utilisation de réseaux convolutifs profonds (DCNN) pour classifier des mammographies en termes de positionnement. Leur étude porte sur deux critères de positionnement tirés des recommandations japonaises, soient la visibilité de l'angle infra-mammaire et la position du mamelon. Un ensemble de 1 631 vues MLO issues du jeu de données CBIS-DDSM est utilisé pour entraîner et évaluer plusieurs architectures DCNN. La classification des images selon les classes *poor, average, excellent* est effectuée par deux technologues ayant discuté pour arriver à un consensus. Cinq modèles pré-entraînés sur Image-Net, soient VGG16, Xception, Inception-v3, Inception-ResNet-v2 et EfficientNet-B0, sont entraînés pour classifier les images selon les trois classes établies. Les meilleures performances de classification sont obtenues avec l'architecture VGG16 pour le critère relatif à l'angle infra-mammaire et avec Xception pour le critère du mamelon. Les auteurs rapportent une précision maximale de 0,7836 et de 0,7278 pour ces deux critères. Ils soulignent toutefois l'impact négatif du déséquilibre des classes sur les performances et identifient comme principale limite de leur approche le nombre restreint de critères évalués.

Quelques autres travaux ont appliqué des CNN pour la classification de mammographies, notamment pour la détection de pli de peau [146], l'évaluation du mamelon vu de profil [140] ainsi que l'évaluation de la visibilité de l'angle inframammaire [147].

Ces études montrent le potentiel de l'application de modèles de neurones convolutifs pour l'évaluation automatique du positionnement mammaire. Toutefois, ces approches demeurent limitées en termes d'explicabilité.

D'autre part, des modèles d'apprentissage profond peuvent être employés non pas pour effectuer une classification directe bout-en-bout, mais plutôt pour prédire la position de certains repères anatomiques, qui peuvent ensuite être utilisés pour évaluer la qualité du positionnement. De telles approches offrent des informations géométriques et morphologiques pertinentes, dont l'évaluation de certains critères de positionnement peut découler. Ces méthodes permettent ainsi une meilleure interprétation des prédictions et, par le fait même, une meilleure explicabilité.

Gupta et al. [127] emploient cette approche, basant l'évaluation du positionnement sur une détection préalable de la ligne du muscle pectoral, de la PNL, et du mamelon. Les auteurs utilisent un ensemble de données privé et y appliquent des techniques d'augmentation afin d'en accroître le nombre et la diversité. Les images sont annotées par un radiologue identifiant la ligne du muscle pectoral, la PNL et le mamelon dans les vues appropriées. 442 examens dont le positionnement est adéquat ont été utilisés pour entraîner et évaluer un modèle de détection de la ligne du muscle pectoral et de la PNL en MLO. Pour pallier la faible quantité de données, les auteurs ont appliqué le principe d'apprentissage par transfert, exploitant un plus gros modèle pré-entraîné sur un grand jeu de données et ensuite optimisé pour la tâche d'intérêt avec moins de données. Le réseau Inception-V3 pré-entraîné sur ImageNet est utilisé comme squelette, sa couche finale ayant été modifiée pour prédire les coordonnées de deux points, et l'apprentissage du modèle est optimisé à l'aide de la fonction de coût Log-Cosh. Pour la détection du mamelon, les auteurs exploitent la présence de marqueurs de plomb (apposés sur la patiente lors de l'examen) dans leur données et utilisent la transformée de Hough circulaire pour le localiser. À partir de ces prédictions, le positionnement est évalué. Dans la vue MLO, le critère considéré est la quantité adéquate de muscle pectoral, selon lequel la vue MLO est jugée adéquate si les deux lignes se croisent à l'intérieur de l'image. Le deuxième critère évalué est la comparaison de la profondeur du sein dans les vues CC et MLO. Pour ce faire, les longueurs de PNL en MLO et en CC sont comparées, et le critère est satisfait si la différence entre celles-ci est inférieure à 1 cm. La méthode proposée est très performante pour identifier des vues MLO correctement positionnées, avec un taux de vrais positifs de 91,35 %, mais beaucoup moins efficace pour détecter les vues MLO mal

positionnées, un taux de vrais négatifs de 45 % seulement étant rapporté. En ce qui concerne la vue CC, l'évaluation du positionnement a été effectuée sur les images pour lesquelles le mamelon était correctement détecté, soit 84,5 % des images, pour obtenir un taux de vrais positifs de 95,11 %. La méthode proposée par Gupta et al., [127], bien qu'obtenant de bons résultats en termes de sensibilité, est testée sur peu d'image et repose sur la présence d'un marqueur de plomb sur le mamelon, ce qui n'est pas toujours le cas en pratique.

Hejduk et al. [148] combinent les approches par classification et par régression de repères anatomiques. Ils implémentent cinq modèles de classification pour détecter la présence de repères ou caractéristiques anatomiques, et trois modèles de régression pour en localiser précisément la position. L'article évalue la qualité des images selon cinq critères principaux : la visualisation complète du tissu mammaire en vues CC et MLO, la visibilité de l'angle inframammaire en vue MLO, la présence du muscle pectoral en vue CC pour garantir l'inclusion des tissus proches de la paroi thoracique, et la bonne visualisation du mamelon, qui doit être visible de profil et centré sur la vue CC. Des DCNN on été entraînés pour effectuer la classification des images selon ces critères. Pour la localisation de repères, trois modèles sont proposés pour déterminer la position du mamelon dans les deux vues, le point cranial du muscle pectoral ainsi que le point caudal de celui-ci. Les auteurs proposent des modèles de classification composés de 13 couches de convolution entrecoupées de cinq couches de max-pooling, suivies de deux couches entièrement connectées avec activation ReLU. Ils intègrent la normalisation de *batch* et un *dropout* de 0.5 afin de limiter le surapprentissage, avec une couche softmax finale, une fonction de perte d'entropie croisée et une optimisation par descente de gradient stochastique. Les modèles de régression reposent sur une architecture similaire, mais utilisent une fonction de perte MSE (Mean Square Error - Erreur quadratique moyenne), une activation sigmoïde, et prennent en entrée à la fois l'image et les coordonnées d'un point correspondant à la position de la caractéristique à localiser. Les auteurs rapportent des précision de classification sur l'ensemble de test allant de 0,873 à 0,985 selon le critère, et des erreurs moyennes de localisation de 8,4 mm, 17,9 et 13,2 mm pour le mamelon ainsi que les points cranial et caudal du muscle respectivement.

Tanyel et al. [128] proposent un modèle permettant d'identifier simultanément la position du mamelon et les deux points délimitant la ligne du muscle pectoral en vue MLO, à partir desquels ils évaluent le positionnement mammaire. Le critère considéré est que la PNL, tracée du mamelon jusqu'au muscle pectoral à angle droit, croise le muscle pectoral dans les frontières de l'image. Le jeu de données est composé de 2000 images MLO tiré de l'ensemble public VinDr, qui ont été annotées par un radiologue expert. Une architecture U-Net, combinée à un module CoordConv (*coordinate convolution*) qui permet de mieux apprendre des informations de position spatiale et à des mécanismes d'attention, est utilisée pour la détection

de repères. L'apprentissage est optimisé avec une fonction de perte *Landmark-Aware Wing Loss*, qui permet d'améliorer la précision de prédiction des coordonnées de points de repère. Elle combine une forte sensibilité aux petites erreurs avec une réponse modérée aux grandes erreurs grâce à une fonction basée sur la *Wing Loss* :

$$L_{\text{Wing}}(y) = \begin{cases} w \cdot \ln\left(1 + \frac{|y|}{\epsilon}\right), & \text{si } |y| < w \\ |y| - C, & \text{si } |y| \geq w \end{cases} \tag{2.1}$$

où y est l'erreur absolue entre la prédiction et la vérité, w est la largeur de la zone non-linéaire, ϵ contrôle la courbure dans cette zone, et

$$C = w - w \cdot \ln\left(1 + \frac{w}{\epsilon}\right) \tag{2.2}$$

assure la continuité de la fonction.

Pour chaque point de repère, la perte est calculée sur ses coordonnées x et y, puis les pertes des différents points sont combinées avec des poids spécifiques :

$$L_{\text{LLAW}} = \alpha \cdot L_{\text{Wing}}(L_1) + \beta \cdot L_{\text{Wing}}(L_2) + \gamma \cdot L_{\text{Wing}}(L_3) \tag{2.3}$$

Cette approche permet au modèle de rester très précis pour les petites erreurs tout en limitant l'influence des grandes erreurs sur l'apprentissage [128]. L'architecture globale du modèle, nommé CoordAttUNet, est illustrée à la figure 2.8.

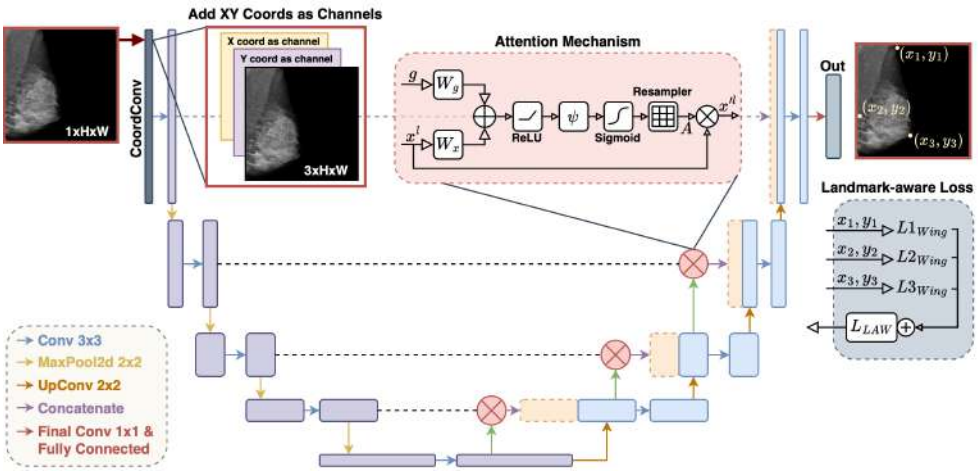


FIGURE 2.8 Architecture globale du modèle CoordAttUNet

Le modèle proposé permet une localisation du mamelon avec une erreur moyenne de 2,97 mm, et la localisation du muscle pectoral avec une erreur de 5,62 et 6,49 mm pour les points du bas et du haut de la ligne. En termes d'évaluation du positionnement, le modèle obtient une précision de 88,63 %, une spécificité de 90,25 % et une sensibilité de 86,04 % [128]. Le modèle CoordAttUNet surpasse les modèles de référence R-ResNeXt50, et UNet. Un AttentionUNet faisant aussi partie des modèles de référence rivalise quant à lui avec le modèle proposé, le devançant notamment en termes de sensibilité dans la classification. L'architecture implémentée par Tanyel et al. offre jusqu'ici parmi les meilleures performances quant à la détection du muscle pectoral et du mamelon, mais l'approche est limitée à la vue MLO, et un seul critère de positionnement est évalué.

Il existe également certaines solutions commerciales visant l'évaluation automatique du positionnement, telles Volpara Analytics, Intelli-Mammo par Densitas ou encore B-quality de B-rayZ [149–151]. Cependant, les publications au sujet de ces solutions concernent généralement leur évaluation, sans inclure d'éléments relatifs à l'architecture ou la méthodologie employée pour concevoir ces solutions [152].

En somme, l'évaluation du positionnement mammaire est un sujet d'un intérêt grandissant dans la littérature. Les études analysées mettent en exergue deux tendances principales à l'automatisation de cette évaluation : la classification bout-en-bout par critère et la détection de repères anatomiques à partir desquels les critères sont évalués. Malgré les résultats prometteurs, de nombreuses limites subsistent, soulignant la nécessité de poursuivre la recherche et le développement de solutions.

CHAPITRE 3 RATIONNELLE ET OBJECTIFS DE RECHERCHE

Les chapitres précédents ont permis de mettre en exergue les points suivants, formant la rationnelle du projet :

- La mammographie est la modalité de référence pour le dépistage du cancer du sein ;
- Le positionnement inadéquat des seins est une des causes principales de faux négatifs et de reprises d'examens ;
- Il n'existe présentement aucun jeu de données publique comportant d'annotations relatives à l'évaluation des critères de positionnement mammaire ;
- La segmentation du muscle pectoral fait l'objet de nombreux travaux, les architectures U-Net et DeepLab étant répandues ;
- La détection du mamelon a évolué des techniques basées sur les gradients d'intensité vers des méthodes d'apprentissage profond et demeure une tâche complexe en raison des variations morphologiques et des données limitées ;
- Dans la recherche concernant l'évaluation automatique du positionnement des seins, deux tendances principales émergent : une évaluation basée sur une classification bout-en-bout, et une détection préalable de repères anatomiques ;
- Quelque soit l'approche employée, l'utilisation de modèles d'apprentissage profond constitue aujourd'hui l'état de l'art.

Considérant ces éléments, le développement d'un système d'évaluation automatique du positionnement des seins en mammographie qui soit adapté aux normes québécoises se révèle nécessaire. De plus, il apparaît pertinent d'employer des modèles d'apprentissage profond et de tirer profit des capacités des réseaux de neurones à extraire des caractéristiques pertinentes des images. Pour ce faire, il est nécessaire d'avoir un jeu de données adapté à l'entraînement de ces modèles.

Par ailleurs, la revue de littérature a permis de faire ressortir la pertinence des approches par détection de repères anatomiques, qui, par opposition aux méthodes de classification bout-en-bout, bénéficient d'une explication géométrique des prédictions, favorisant l'interprétabilité des résultats. Ainsi, l'avenue explorée dans ce travail est l'évaluation du positionnement à partir de repères anatomiques, soient la ligne du muscle pectoral et la position du mamelon.

L'objectif et les sous-objectifs de ce projet sont donc définis comme suit :

Objectif : Développer une méthode d'évaluation automatique du positionnement mammaire basée sur la détection de repères anatomiques

Sous-objectifs :

1. Créer un jeu de données annoté selon les critères de positionnement proposés par [12] et les repères anatomiques, adapté à l'entraînement de modèles de classification et de détection
2. Analyser la variabilité inter-opérateur dans l'identification des repères anatomiques
3. Implémenter et valider un modèle de détection de la ligne du muscle pectoral
4. Implémenter et valider un modèle de détection du mamelon
5. Développer un algorithme d'évaluation du positionnement selon quatre critères géométriques

La méthodologie employée pour atteindre les objectifs 1 et 2 est présentée au **chapitre 4**, qui détaille le processus d'annotation et d'analyse des données. Les objectifs 3, 4 et 5 sont couverts en grande partie dans l'article de conférence constituant le **chapitre 5**. Cet article est accepté pour présentation à la conférence internationale ISBI (International Symposium on Biomedical Imaging) à Londres en Avril 2026. Le **chapitre 6** offre des éléments méthodologiques et des résultats complémentaires par rapport à ces objectifs. Les résultats obtenus sont analysés dans la discussion générale présentée au **chapitre 7**.

CHAPITRE 4 CRÉATION DU JEU DE DONNÉES MAMMO-POS

Dans ce chapitre, la création du jeu de données **Mammo-Pos**, menée en collaboration avec les deux autres étudiants participant au projet, est résumée. L'origine des données, le processus d'annotation, une analyse préliminaire ainsi que le pré-traitement des données y sont décrits. Plus de détails concernant le processus d'annotation ainsi que l'analyse de la variabilité inter-opérateur dans les annotations sont présentés dans l'article *Breast Positioning Assessment in Mammography : An Inter-Rater Reliability Study* [25]. Cet article, co-rédigé avec les deux autres étudiants travaillant sur le projet, est soumis pour publication à la revue *Journal of Breast Imaging*.

4.1 Origine des données

L'ensemble des données disponibles pour ce projet est constitué de 1923 examens issus d'un audit réalisé en 2017 à travers plusieurs cliniques de dépistage au Québec. Dans le cadre de cet audit, mandaté par l'OTIMROEPMQ et mené par notre partenaire IMD Research, les images de mammographie ont été évaluées selon les 17 critères de positionnement introduits précédement [12]. Les examens recueillis comprennent, sauf exception, les quatre images standard, soient les vues MLO et CC pour le sein droit et pour le sein gauche. Les examens de tomosynthèse, les mammographies de patientes ayant un implant mammaire ou ayant subi une double mastectomie ainsi que les mammographies de patients masculins ont été exclus. Les images constituant ces examens ainsi que les annotations associées aux évaluations des 17 critères (succès ou échec de chaque critère) nous ont été rendues accessibles. Les images ayant été acquises par différents systèmes et étant de résolution différentes, le ratio pixel à mm a également été fourni pour chaque image. L'utilisation de ces données pour les fins du projet a été approuvée par le comité éthique de la recherche de Polytechnique Montréal (CER-2425-26-D).

Malgré l'accès à ces données et aux annotations issues de l'évaluation du positionnement, un problème subsiste. En effet, les annotations initiales, ayant été réalisées par des technologues en contexte clinique, tiennent compte d'informations contextuelles telles que la morphologie des patientes, leurs antécédents chirurgicaux ou leurs examens antérieurs. Ainsi, un critère pouvant sembler non satisfait sur les images selon une définition stricte peut néanmoins être jugé acceptable en pratique pour une patiente donnée, considérant le contexte. Toutefois, nous ne disposons pas de ces informations dans le jeu de données. Par ailleurs, les annotations initiales ne comprennent aucune information concernant les repères anatomiques tels la ligne

du muscle pectoral et la position du mamelon.

Dans ce contexte, il a été décidé de lancer une nouvelle phase d'annotation sur un sous-ensemble de ces données, l'objectif étant d'obtenir des annotations plus complètes et mieux adaptées, c'est-à-dire reposant exclusivement sur l'information contenue dans l'image. Cette nouvelle phase inclut également des annotations visuelles des repères anatomiques, afin de permettre l'entraînement et l'évaluation de modèles de détection dédiés. Le temps et les ressources étant limités, seul un sous-ensemble du jeu de données initial a été ré-annoté. La sélection des examens constituant ce sous-ensemble a été faite sur la base d'expérimentations préliminaires, soient l'application de méthodes traditionnelles de traitement d'images pour la détection du mamelon et du muscle pectoral. Des méthodes de seuillage basées sur l'intensité, des approches par croissance de régions ainsi que la détection de points par la transformée de Hough ont notamment été utilisées. Celles-ci ont permis de mettre en évidence les cas les plus complexes du jeu de données, c'est-à-dire les images pour lesquelles le muscle pectoral ou le mamelon étaient difficilement identifiables par des algorithmes simples de traitement d'images. 1215 examens ont ainsi été sélectionnés et annotés, totalisant 4760 images dont 2386 vues CC et 2374 vues MLO. L'ensemble résultant de cette phase d'annotation constitue le jeu de données **Mammo-Pos**.

4.2 Processus d'annotation

Les images sélectionnées ont été annotées à l'aide de la plateforme Computer Vision Annotation Tool (CVAT) [153]. Chaque image a été annotée individuellement à l'aide de labels que nous avons manuellement définis via l'interface de CVAT. Cet outil offre la possibilité d'appliquer différents types d'annotations, tels que des boîtes englobantes, des lignes, des points ou des étiquettes, auxquels peuvent être associés divers attributs, comme des cases à cocher, des listes déroulantes ou des champs de texte.

Nous avons recruté trois technologues séniors expertes en mammographie faisant partie de l'OTIMROEPMQ pour participer au processus d'annotation. Celles-ci sont également formatrices de technologues en mammographie. Afin d'harmoniser les pratiques et d'établir un concensus sur certains aspects des annotations demandées, différant des éléments habituellement considérés en clinique, nous avons tenu une journée de calibration lors de laquelle les technologues ont pu se familiariser avec l'outil d'annotation. Pendant cette période de calibration, 10 examens ne faisant pas partie de l'ensemble final ont été annotés en guise de pratique. Les 1215 examens d'intérêt ont ensuite été répartis parmi les trois technologues, chacune ayant un nombre donné d'examens à annoter selon sa disponibilité. Un sous-ensemble de 250 examens a été annoté par chacune des trois technologues, constituant un ensemble

comparatif sur lequel une analyse de la variabilité inter-opérateur a pu être effectuée. Cet ensemble a également servi de jeu de test pour l'évaluation des modèles développés dans ce travail. En dehors des discussions initiales menées lors de la journée de calibration, visant à établir des normes communes pour le processus d'annotation, aucune communication relative aux détails des annotations n'a eu lieu entre les technologues. Cette organisation visait à garantir une évaluation non influencée par les avis des autres et à mesurer fidèlement la variabilité réelle entre opérateurs.

Pour chaque image, les annotatrices ont identifié les repères anatomiques pertinents, soient la position du mamelon, la PL, la PNL et l'IMA. Elles ont également encadré les artéfacts problématiques et les plis de peau. Le tableau 4.1, tiré de [25], comprend le type d'annotation apposé pour chaque structure. Outre les repères visuellement identifiés, les critères de positionnement prédéfinis [12] ont été évalués par les annotatrices, sans avoir accès à d'autres informations sur les patientes (âge, poids, antécédents médicaux ou examens antérieurs). Le type d'annotation employé pour chaque critère est consigné dans le tableau 4.2, adapté de [25].

Selon chaque critère, les technologues ont apposé les annotations appropriées à l'aide des étiquettes que nous avions prédéfinies sur la plateforme CVAT. Elles devaient d'abord indiquer, au moyen d'une étiquette, si l'image était de qualité suffisante, avec la possibilité d'ajouter un commentaire explicatif en cas de problème. Ensuite, pour chaque structure visuellement identifiée, les technologues ont évalué les critères de positionnement associés en modifiant les attributs correspondants.

La majorité des critères ont été évalués directement par les technologues, soit en cochant une case, soit en sélectionnant une option dans une liste déroulante afin d'indiquer la réussite ou l'échec du critère. Le critère *Longueur de la PNL en CC à moins de 1 cm de celle en MLO* n'a toutefois pas été évalué directement par les technologues ; il a été évalué a posteriori par notre équipe, en comparant la longueur des lignes tracées par les technologues.

Une fois l'annotation de l'ensemble des images complétée sur la plateforme CVAT, les données ont été exportées. Le travail d'extraction et de mise en forme a été réalisé par un membre de notre équipe, qui a compilé, pour chaque image, les coordonnées des repères identifiés ainsi que l'ensemble des labels de positionnement. Pour l'ensemble de test, annoté par les trois technologues, la classe attribuée par chacune ainsi que la classe majoritaire ont été consignées. Toutes ces informations ont ensuite été regroupées dans des fichiers utilisés pour le chargement de ces annotations, utilisées comme vérités terrain.

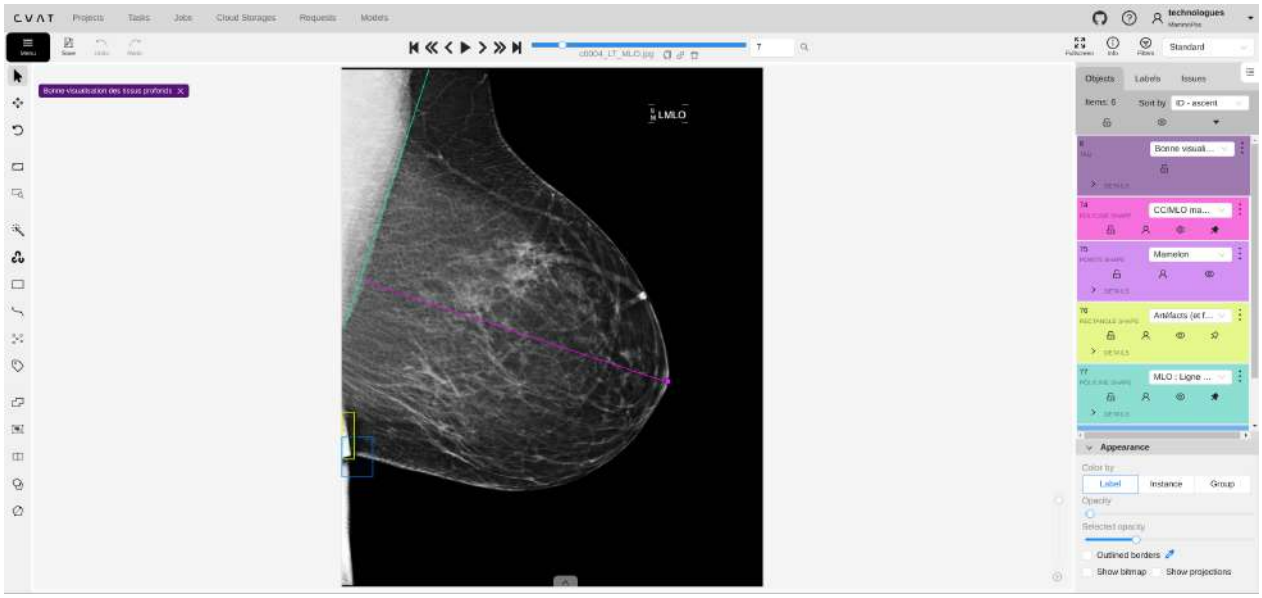
Tableau 4.1 Structures identifiées visuellement

Vue	**Structures**	**Type d'annotation**
CC et MLO	Plis de peau	Boîte englobante
	Artéfacts	Boîte englobante
	Position du mamelon	Point
MLO	PL	Ligne
	PNL	Ligne
	IMA	Boîte englobante
CC	PNL	Ligne
	Profondeur maximale	Ligne

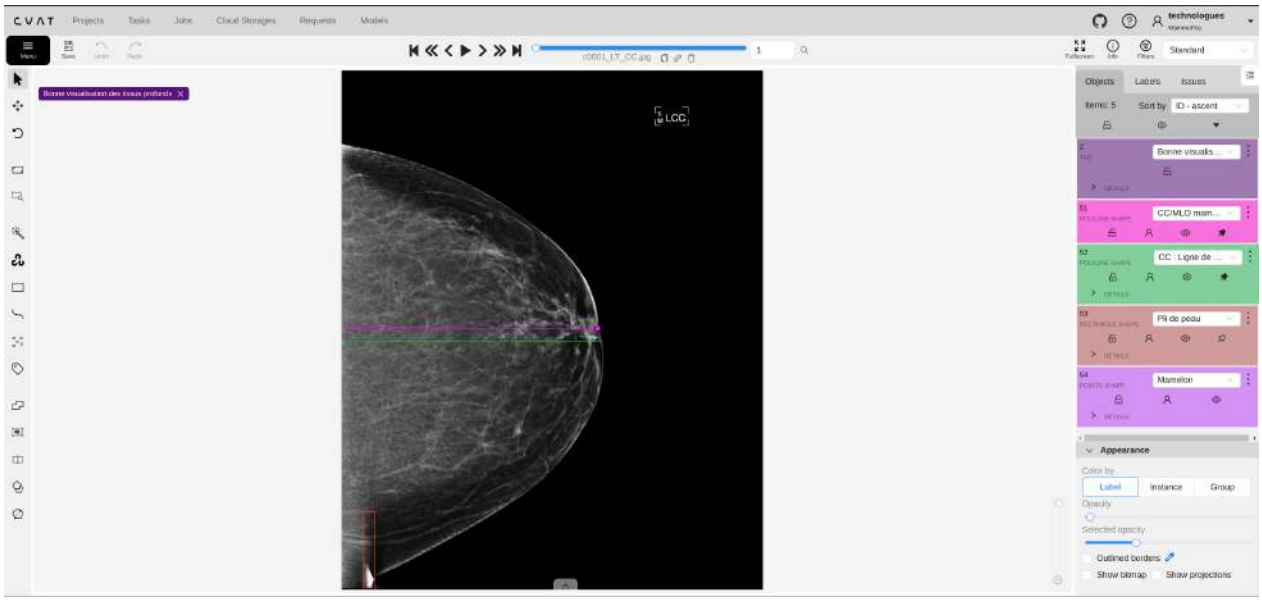
Tableau 4.2 Annotations associées à l'évaluation des critères de positionnement

Vue	**Critère**	**Type d'annotation**
CC et MLO	Image évaluable	Étiquette (Vrai/Faux), champs de texte pour commentaires
	Plis de peau	Cases à cocher
	Artéfacts et artéfacts de superposition	Cases à cocher
	Mamelon vu de profil	Case à cocher
	Partie du sein coupée	Cases à cocher
	Bonne visualisation des tissus profonds	Étiquette (Vrai/Faux), cases à cocher
MLO	Quantité adéquate de muscle pectoral	Case à cocher
	Vue de la largeur maximale du muscle pectoral	Case à cocher
	IMA bien ouvert et montré	Liste déroulante
CC	PNL perpendiculaire au bord de l'image	Case à cocher
	Longueur de la PNL en CC à moins de 1 cm de celle en MLO	Lignes (évaluation a posteriori)

L'interface de CVAT est montrée à la figure 4.1, où on peut voir des exemples d'annotations sur les deux vues, dont des lignes, des points et des boîtes englobantes. La barre de droite comprend toutes les annotations apposées, dont les attributs peuvent être modifiés.



(a) Vue MLO



(b) Vue CC

FIGURE 4.1 Interface CVAT avec exemples d'annotations

4.3 Analyse des données

Dans cette section, une analyse préliminaire des données est présentée. La distribution des classes et la variabilité dans les critères de positionnement traités dans ce travail est brièvement abordée. La variabilité inter-opérateur, spécifiquement en termes d'identification de la ligne du muscle pectoral et du mamelon, est analysée dans cette section. Tel que mentionné précédement, l'article fournit des explications détaillées de cette analyse ainsi que des résultats d'analyses supplémentaires, notamment sur la variabilité dans l'évaluation des critères de positionnement, effectuée par les autres membres de l'équipe.

4.3.1 Distribution des données et variabilité dans l'évaluation des critères de positionnement

Au terme de la phase d'annotation décrite ci-haut, les données ont été analysées afin d'en extraire des informations relatives à la distribution des classes ainsi qu'à la variabilité inter-opérateur. Dans le cadre de la rédaction de l'article *Breast Positioning Assessment in Mammography : An Inter-Rater Reliability Study* [25], la figure 4.2 a été générée, illustrant la distribution des cas de succès et d'échecs sur l'ensemble de test, pour tous les critères.

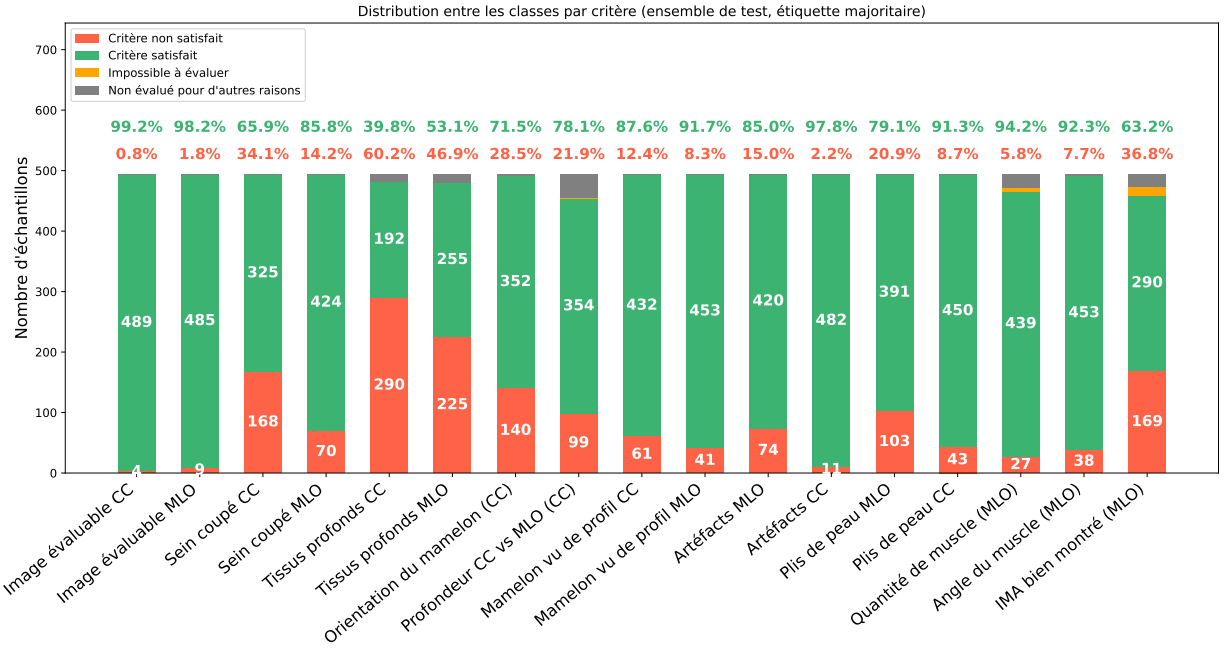


FIGURE 4.2 Distribution des classes (succès/échec) selon le critère

On remarque un fort déséquilibre de classe pour plusieurs critères, notamment *Quantité adéquate de muscle pectoral* et *Vue de la largeur maximale du muscle pectoral*, pour lesquels

seulement 5,8 % et 7,7 % d'échecs sont rapportés. Des résultats semblables sont rapportés dans la littérature, notamment par Rouette et al. [12] qui dénombrent 5,5 %, 13,8 % d'échecs pour ces critères. Brahim et al. [1] observent cependant des taux d'échecs plus élevés sur leur jeu de données, soient 94,4 % et 28,8 % pour les critères correspondant à *Quantité adéquate de muscle pectoral* et *Vue de la largeur maximale du muscle pectoral*.

On constate un déséquilibre moins prononcé dans nos données pour les deux autres critères d'intérêt dans ce mémoire, soient *PNL perpendiculaire au bord de l'image* (28,5 % d'échecs) et *Longueur de la PNL en CC à moins de 1 cm de celle en MLO* (21,9 % d'échecs). Rouette et al. [12] rapportent, quant à eux, seulement 14,9 % et 6,6 % d'échecs pour ces critères. Dans l'étude de Brahim et al. [1], 39 % d'échecs sont dénombrés pour un critère combinant la visibilité du mamelon de profil et la désorientation du mamelon, ce dernier élément s'apparentant à notre critère *PNL perpendiculaire au bord de l'image*, relatif à l'orientation du mamelon par rapport à l'horizontale.

Par ailleurs, la variabilité inter-opérateur quant à l'évaluation des critères de positionnement a été analysée par un membre de notre équipe, en calculant entre autres le taux de concordance (ou l'accord) entre les technologues pour chaque image de l'ensemble de test. Ce taux, calculé pour chaque critère, est défini comme le rapport du nombre d'annotations unanimes (les trois technologues sont d'accord pour dire qu'un critère donné échoue ou pas) et du nombre total d'images annotées. Les taux de concordance sont rapportés à la figure 4.3, tirée de l'article *Breast Positioning Assessment in Mammography : An Inter-Rater Reliability Study* [25].

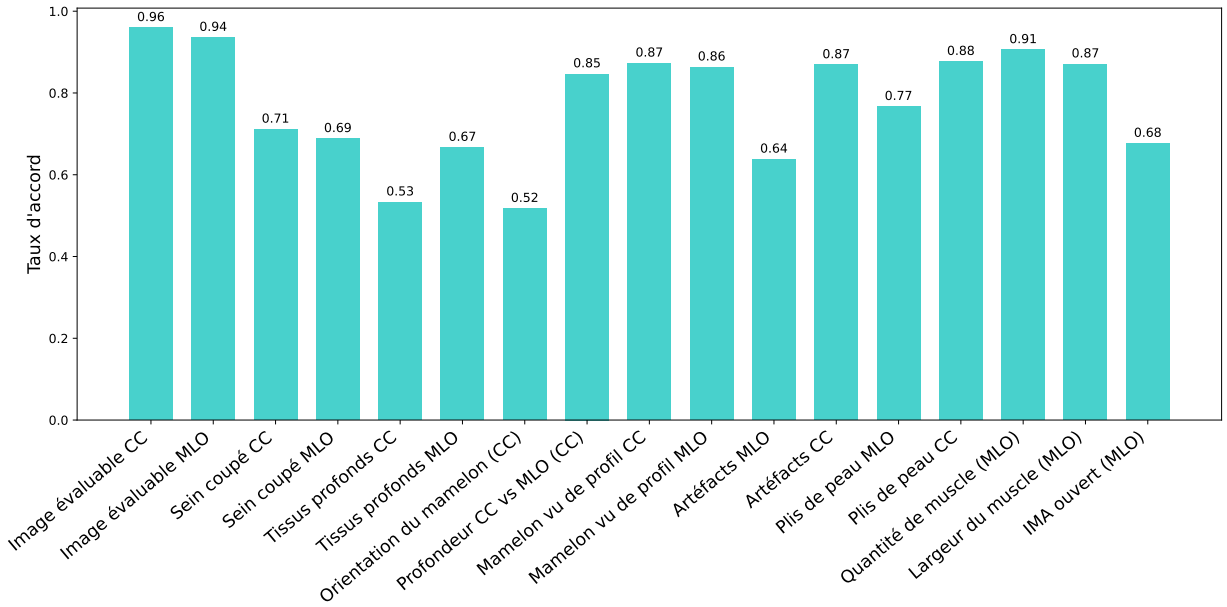
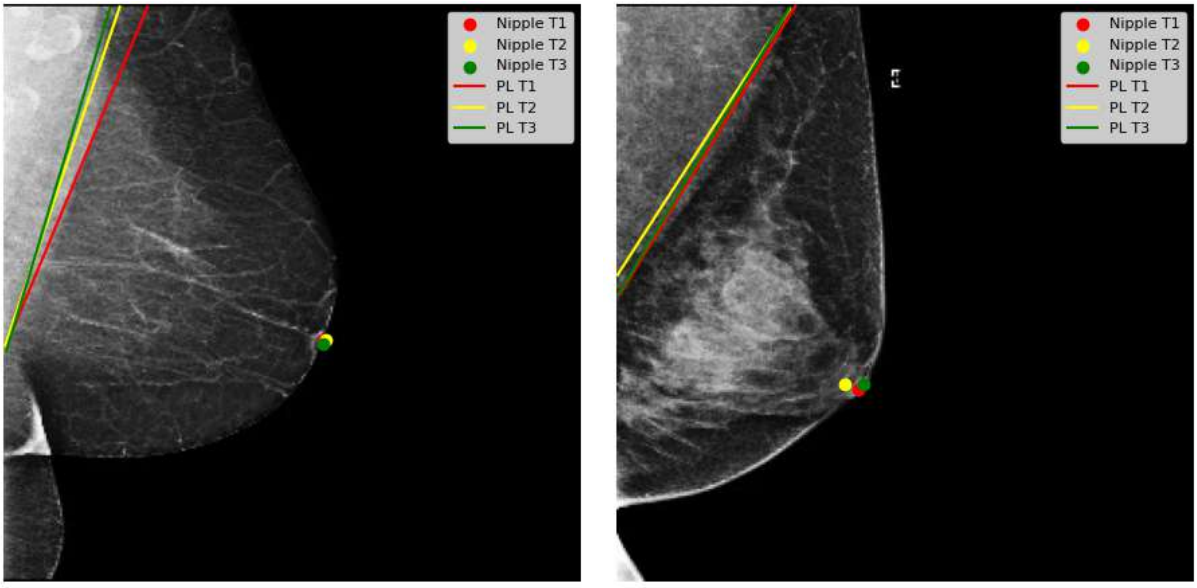


FIGURE 4.3 Taux d'accord entre les annotatrices sur l'ensemble de test

Ce graphique permet de mettre en évidence des taux de concordance de 91 %, 87 %, 52 % et 85 % pour les critères *Quantité adéquate de muscle pectoral*, *Vue de la largeur maximale du muscle pectoral*, *PNL perpendiculaire au bord de l'image* et *Longueur de la PNL en CC à moins de 1 cm de celle en MLO* respectivement.

4.3.2 Analyse de la variabilité dans l'identification des repères anatomiques

Le caractère qualitatif de la tâche d'identification du muscle pectoral et du mamelon, qui repose principalement sur une évaluation visuelle, entraine une certaine variabilité entre les trois annotatrices. En effet, pour une même image, les lignes tracées par les trois technologues pour délimiter le muscle pectoral ainsi que le point apposé pour identifier le mamelon sont rarement superposés. Bien entendu, une certaine variabilité inhérente à la précision de l'outil d'annotation, et plus largement à la subjectivité humaine, est tout à fait normale. Toutefois, dans certains cas, les différences observées sont importantes et méritent d'être analysées. La figure 4.4 présente un exemple de cas pour lesquels les annotations sont particulièrement variables.



(a) Variabilité au niveau de la PL (b) Variabilité au niveau du mamelon

FIGURE 4.4 Exemples d'annotations apposées par les trois technologues

Dans le but de quantifier cette variabilité, les annotations des 250 examens évalués par les trois technologues ont été comparées. Des mesures de distances ont été employées pour caractériser cette variabilité entre les points identifiés comme extrémités de la ligne du muscle pectoral

et la position du mamelon.

D'abord, la variabilité dans l'identification du muscle pectoral a été évaluée en prolongeant jusqu'au bord de l'image les lignes tracées par les annotatrices, puis en comparant les points supérieur et inférieur de ces lignes. Pour chaque image, la distance moyenne entre les trois points inférieurs, ainsi qu'entre les trois points supérieurs, a été calculée, puis ces valeurs ont été moyennées sur l'ensemble des images. La moyenne de la distance maximale entre les points a également été compilée. De plus, le centroïde de chaque annotation, soit la moyenne des coordonnées, a été identifié. La moyenne des distances entre les trois points annotés et ce centroïde sont rapportés dans le tableau 4.3, en millimètres.

La variabilité dans l'annotation du mamelon a été évaluée de la même manière. La distance moyenne entre les points, la moyenne de la distance maximale ainsi que la moyenne des distances du centroïde ont été calculées et sont consignées dans le tableau 4.3, en millimètres.

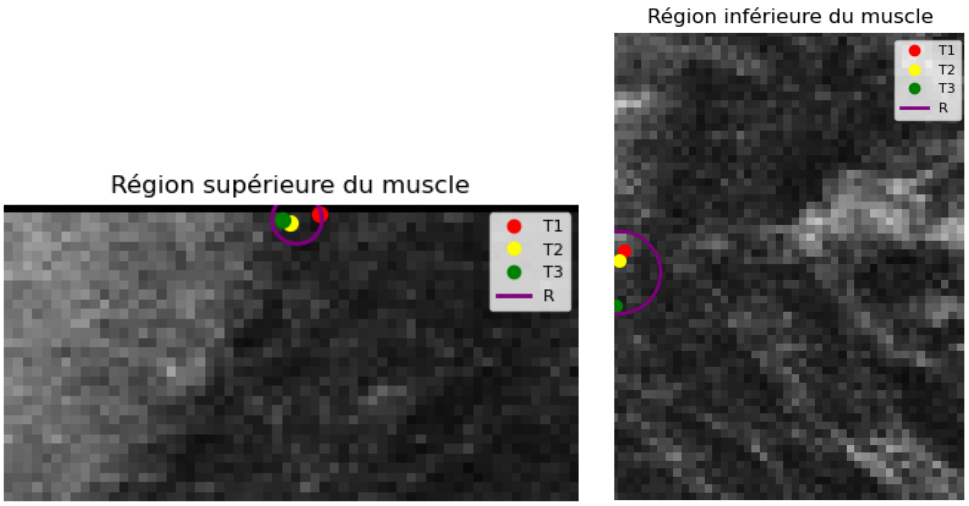
Tableau 4.3 Variabilité entre les annotatrices dans l'identification des repères anatomiques : distances moyennes (Dist. moy) et maximales (Dist. max) entre les annotations et distance moyenne du centroïde (Dist. centroïde)

	Dist. moy. (mm)			**Dist. max.(mm)**			**Dist. centroïde (mm)**		
	Moy.	É.T.	Med.	Moy.	É.T	Méd.	Moy.	É.T	Méd.
Point inférieur du muscle pectoral	3,55	5,02	2,30	5,32	7,53	3,45	2,07	3,10	1,33
Point supérieur du muscle pectoral	1,38	1,45	1,00	2,07	2,18	1,50	0,81	0,88	0,58
Point du mamelon	2,68	5,82	2,05	3,74	8,71	2,84	1,59	3,82	1,17

Pour le muscle pectoral, la distance moyenne entre les annotations des évaluatrices est de 3,55 mm pour le point inférieur de la ligne et de 1,38 mm pour le point supérieur. Dans les deux situations, la médiane est inférieure à la moyenne, tandis que l'écart-type est plus élevé, ce qui suggère la présence de quelques valeurs aberrantes, possiblement liées à des erreurs d'annotation ou à des cas plus ambigus. On constate par ailleurs une variabilité plus importante pour l'identification du point inférieur que pour celle du point supérieur, qui peut s'expliquer en partie par une définition visuelle moins nette de la région inférieure du muscle pectoral. Concernant la localisation du mamelon, les distances observées traduisent globalement une bonne concordance, malgré la présence de quelques annotations aberrantes.

Ces valeurs de variabilité ont été prises en compte dans l'évaluation des modèles de détection du muscle pectoral et du mamelon. En effet, si une distance non négligeable existe entre les points apposés par trois annotatrices pour identifier une même structure, une prédiction d'un modèle qui se situe environ à cette distance devrait être jugée acceptable. De ce fait, un

rayon de variabilité moyen, correspondant à la moyenne des distances maximales entre les trois annotations, a été défini et utilisé comme référence dans l'évaluation des prédictions. Ainsi, une prédiction se situant à une distance inférieure au rayon de variabilité par rapport à la vérité terrain (qui correspond au centroïde des trois points pour l'ensemble de test) sera jugée acceptable, au même titre qu'une annotation apposée par une technologue pourrait se situer à cette distance du centroïde. Ce rayon de variabilité est illustré à la figure 4.5 par un cercle violet centré autour du centroïde des trois annotations. Pour les points supérieur et inférieur du muscle pectoral, ce rayon est de 2,07 et 5,32 mm respectivement. Le rayon de variabilité moyen autour du mamelon, illustré à la figure 4.6 correspond à la distance maximale moyenne entre les annotations, soit 3,74 mm.



(a) Point supérieur (b) Point inférieur

FIGURE 4.5 Rayon de variabilité moyen (R) par rapport au centroïde des annotations des points supérieur et inférieur de la PL apposées par les technologues T1, T2 et T3

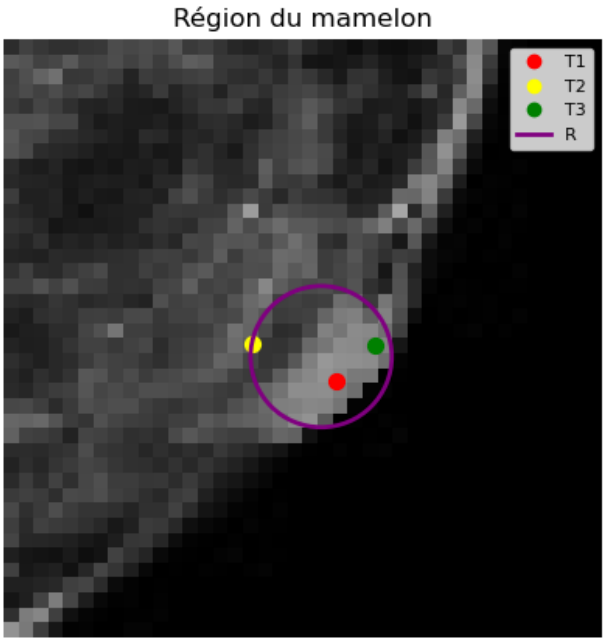
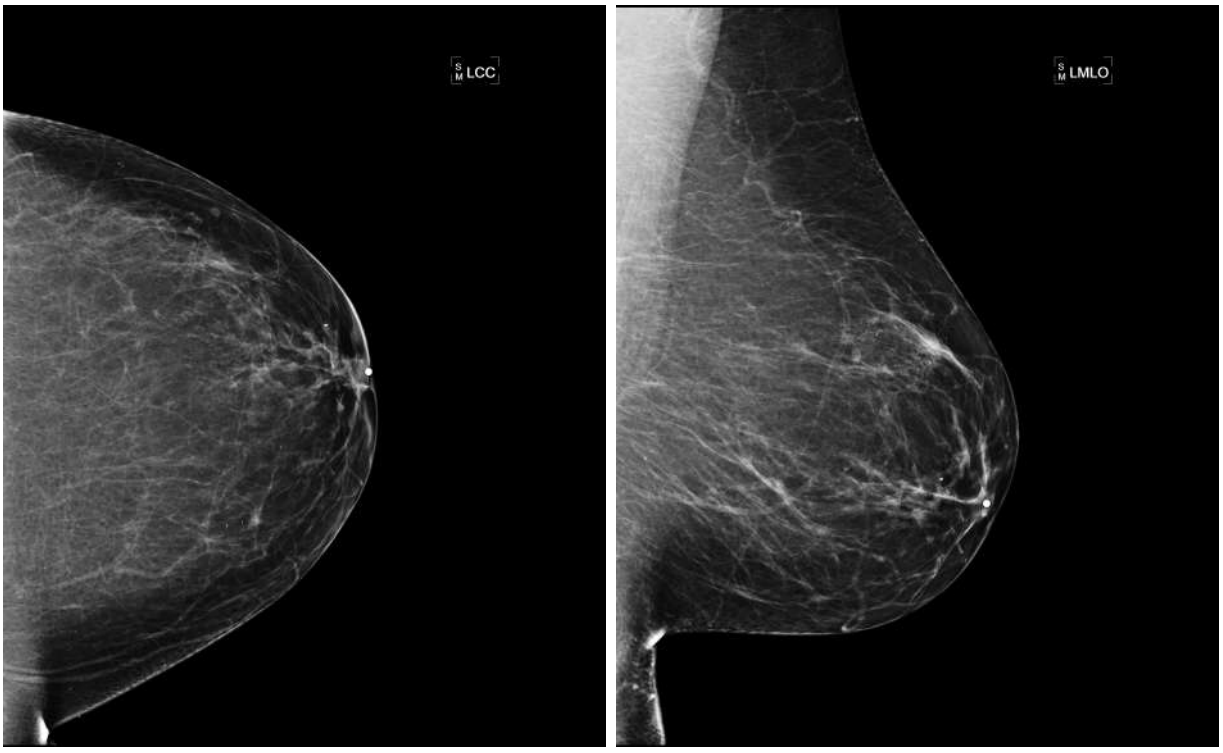


FIGURE 4.6 Rayon de variabilité moyen (R) par rapport au centroïde des annotations du mamelon apposées par les technologues T1, T2 et T3

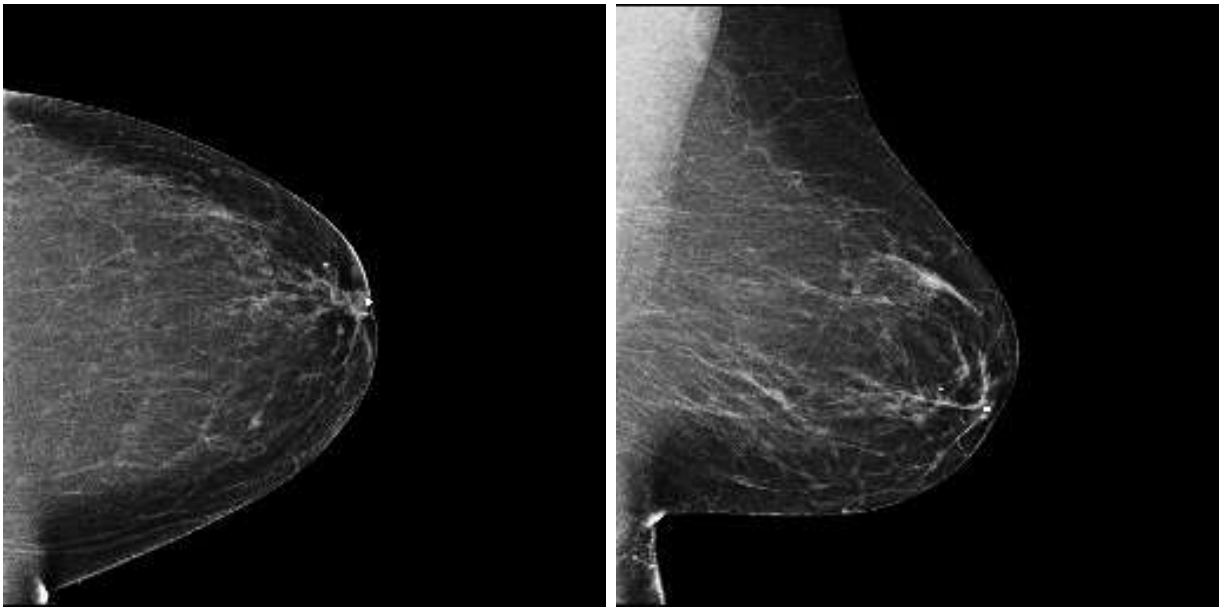
4.4 Pré-traitement des données

Afin d'uniformiser les données et de les adapter à l'entraînement de modèles d'apprentissage profond, des opérations de pré-traitement ont été effectuées sur les images. Les images originales sont de haute résolution, variant entre 2560 par 3328 pixels et 3328 par 4096 pixels. Sur chaque image, une étiquette indiquant le côté (sein droit ou sein gauche) et la vue (CC ou MLO) est visible dans un coin supérieur. Un algorithme de pré-traitement a été implémenté par un membre de l'équipe à l'aide de la librairie *OpenCV* de Python pour redimensionner les images et en retirer l'étiquette (masquage). Des exemples d'images en format original ainsi qu'après les étapes de pré-traitement sont illustrés dans les figures 4.7 et 4.8.



(a) Vue CC (b) Vue MLO

FIGURE 4.7 Exemple d'images d'un sein gauche dans leur format original



(a) Vue CC (b) Vue MLO

FIGURE 4.8 Exemple d'images après pré-traitement

CHAPITRE 5 ARTICLE 1 - DEEP LEARNING BASED LANDMARKS DETECTION FOR POSITIONING ASSESSMENT IN MAMMOGRAPHY

Dans ce chapitre, la méthodologie ainsi que les résultats relatifs à la détection du muscle pectoral, la détection du mamelon et l'évaluation de la quantité adéquate de muscle pectoral sont présentés sous forme d'un article publié dans les actes de la conférence *International Symposium of Biomedical Imaging* (ISBI) en date du 13 janvier 2026. Cet article, limité à 4 pages double colonnes sans références, repris ici dans un format adapté à ce mémoire, résume une partie des travaux effectués dans le cadre de cette maîtrise. La réalisation des objectifs 3, 4 et 5, soient la détection du muscle pectoral, la détection du mamelon et l'évaluation du positionnement y sont traités. Plus de détails concernant les sujets abordés dans l'article, des éléments de méthodologie ainsi que des résultats supplémentaires sont présentés dans le **chapitre 6**.

Auteurs : Lauriane Grondin-Reiher, Victor Patenaude, Valérie Gauvin, François Guibault, Lama Séoud

Ma contribution à cet article inclut la conception, le développement de la méthodologie et les expérimentations, la validation, la recherche bibliographique, la rédaction, ainsi que l'analyse, la visualisation et la présentation des données.

5.1 Abstract

Proper breast positioning is critical for reliable mammographic screening, yet remains a major source of diagnostic variability and exam failure. This study proposes an interpretable, landmark-based framework for the automated evaluation of the *Adequate amount of pectoral muscle* criterion in mediolateral oblique mammograms. Our approach combines a DeepLabv3+ model for pectoral muscle segmentation with a Gaussian U-Net for nipple localization, followed by a geometric decision algorithm assessing positioning adequacy from the relative locations of the predicted landmarks. The method was trained and evaluated on 2274 annotated MLO images derived from a multicenter audit, with a test subset independently labeled by three expert radiographers to account for inter-rater variability. The proposed system achieved a balanced accuracy of 0.88 (specificity of 0.92, sensitivity of 0.84), outperforming both classification-based CNN baselines and previous landmark detection models. Mean localization errors were within expert variability (1.5–3.0 mm for pectoral landmarks, 2.0 mm for nipple). These results highlight that anatomically grounded, explainable models

can deliver robust and clinically meaningful positioning assessments even in imbalanced and subjective datasets.

5.2 Introduction

Breast cancer remains the most prevalent cancer among women, with over two million new cases diagnosed in 2022 [8]. Mammography is the gold standard for breast cancer screening, enabling early detection and reducing mortality. However, diagnostic performance strongly depends on proper breast positioning during image acquisition. Suboptimal positioning is a leading cause of exam failure and can result in missed lesions, incomplete exams, and unnecessary recalls [154]. Positioning assessment in mammography follows standardized guidelines for the standard mediolateral oblique (MLO) and craniocaudal (CC) views, defining specific requirements such as adequate pectoral muscle (PM) visibility, full breast coverage, and correct anatomical alignment. In particular, the evaluation of the adequate amount of PM in MLO views is critical, as it reflects the inclusion of posterior breast tissue. These assessments are typically performed visually by radiographers, introducing subjectivity and inter-observer variability [155].

To address these limitations, recent studies have explored artificial intelligence–based systems for automated positioning assessment. Such tools could provide real-time feedback to radiographers, improving positioning quality. Most existing methods rely on convolutional neural networks (CNNs) trained to classify positioning adequacy for specific criteria [1, 148]. Although effective, these end-to-end approaches lack interpretability, as they do not explicitly leverage anatomical features. In contrast, some studies [127, 128] exploit geometric information derived from detected anatomical landmarks, including the nipple and the PM line, to evaluate positioning criteria based on their spatial relationship. The use of anatomical landmarks enhances explainability, since the algorithm's decisions can be directly attributed to observable anatomical features. However, the method proposed in [128] is limited in versatility, as it focuses solely on MLO views for simultaneous detection of the PM and nipple, and relies on annotations from a single expert, which may introduce bias. Many studies have also focused specifically on automatic landmark detection in mammography. Recent works on PM identification mainly rely on region segmentation using architectures such as DeepLabv3+ with Xception or ResNet backbones, or U-Net variants [108, 123–125]. These models generally segment the full pectoral region or trace a curved boundary, sometimes refined through pre-processing or GAN-based shape modeling. Fewer studies directly estimate the linear pectoral boundary [127, 128], which is clinically more relevant for positioning assessment. Indeed, positioning criteria related to PM visibility are typically evaluated based

on this linear boundary [1, 128]. Nipple detection has evolved from classical gradient-based approaches to U-Net–based frameworks, allowing for more robust and anatomically accurate localization [116, 128].

In this work, we propose an automated framework that detects the PM line and nipple position to evaluate the *Adequate amount of pectoral muscle* criterion. Our method combines CNN–based landmark localization with a geometric decision algorithm. Additionally, we benchmark our approach against existing baselines [1, 128] using a test set annotated by three radiographers, enabling comparison with inter-rater variability.

5.3 Methods

5.3.1 Dataset

In this study, we follow the positioning criteria published in [12], derived from the Canadian Association of Radiologists (CAR) guidelines. Our dataset consists of a subset of images collected during a 2017 audit of mammography positioning quality [12]. From this audit, 1,215 exams were annotated by three senior radiographers via a web-based platform (CVAT), including landmark coordinates (PM line and nipple) and positioning quality labels. MLO images with both nipple and PM line annotations were used for training, validation, and testing of our models. The dataset is highly imbalanced with respect to the *Adequate amount of pectoral muscle* criterion, with only 6% of cases labeled as failures. A total of 1,828 MLO images were used for training, with 10% reserved for validation. Among these, 22%, 45% and 33% of images were annotated by radiographers R1, R2, and R3, respectively. The test set consisted of 446 images, each independently annotated by all three radiographers, providing a hold-out set less prone to individual bias and enabling assessment of inter-rater variability. Ground truth landmark coordinates were defined as the mean positions across the three annotators, while for positioning labels, the majority class was used as the reference standard. The images originate from different mammography systems with varying pixel sizes ; this information was used to convert pixel-level errors into millimeters. All images were resized to 256×256 and masked to remove the tag indicating the side and view. Images of the right breast were horizontally flipped to ensure that all images were presented as left-breast views to the models for consistency.

5.3.2 Landmarks detection

Pectoral muscle detection

For the PM segmentation task, which we define as predicting a triangular region corresponding to the muscle area, we trained a DeepLabv3+ model with a ResNet-101 backbone [5]. The model was initialized with weights pre-trained on ImageNet, using the implementation provided by the Segmentation Models PyTorch library. Ground truth binary masks delineating the triangular pectoral region (see Figure 1) were derived from the PM lines annotated by the radiographers. The model was trained for binary segmentation to predict a mask matching the input size. Training was performed for up to 100 epochs using the Adam optimizer (learning rate 1×10^{-4}, batch size 8) and a combined Binary Cross-Entropy–Dice loss to balance pixel accuracy and region overlap. Early stopping (patience = 15) was applied to prevent overfitting, with a fixed seed ensuring reproducibility. On the test set, we extracted the PM line from the predicted masks using a RANSAC-based linear fitting.

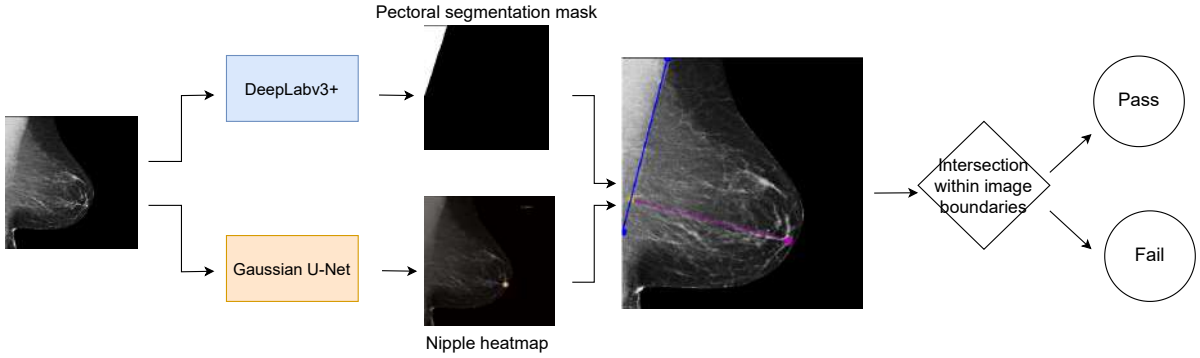


FIGURE 5.1 Overall architecture of our proposed framework to assess the *Adequate amount of pectoral muscle* criterion

Nipple detection

For nipple detection, we adapted a U-Net–based architecture originally proposed for suture point detection through heatmap regression [7]. Our network, named *Gaussian U-Net*, takes as input single-channel mammography images (which can be CC or MLO views) and outputs a heatmap encoding the likelihood of the nipple location. Ground-truth heatmaps were generated by applying a Gaussian kernel (std = 3) centered on the annotated nipple position, producing a smooth spatial distribution around the landmark. The U-Net is composed of DoubleConv blocks combining convolution, ReLU activation, batch normalization, and progressively increasing dropout rates. Feature maps are downsampled via max pooling in the

encoder and upsampled using bilinear interpolation in the decoder, allowing the model to capture both global breast context and fine local features necessary for precise nipple localization. The final feature map is smoothed with a Gaussian filter and processed through a differentiable 2D convolutional Soft-Argmax layer to produce a heatmap. To obtain the final nipple coordinates, we select the pixel with the maximum intensity value in the predicted heatmap. Unlike the original method from [7], which relied on a similarity loss, our model was trained using a combination of mean squared error (MSE) loss and an additional localization term based on the L1 distance to encourage accurate and stable convergence toward the annotated position. The loss function can be expressed as :

$$\mathcal{L}_{\text{total}} = \alpha \cdot \text{MSE}(\hat{H}, H) + \beta \cdot \|\hat{C} - C\|_1$$

where $\hat{H}$ and H are the predicted and ground-truth heatmaps respectively, and $\hat{C}$ and C the corresponding coordinates. The coefficients α and β control the relative weighting of the MSE and coordinate (L1) terms. The model was trained for up to 100 epochs using Adam (learning rate 1×10^{-3}, batch size 8, fixed seed), with $\alpha = 1$, $\beta = 0.1$, and early stopping (patience = 10).

5.3.3 Positioning assessment

We then evaluated the *Adequate amount of pectoral muscle* criterion on the test set using the predicted nipple coordinates and PM line. By definition, this criterion is considered satisfied when the intersection point between the PM line and the line drawn from the nipple at a 90° angle lies within the image boundaries. The evaluation is therefore based on a purely geometric computation derived from the identified anatomical landmarks. For each image in the test set, the intersection point between these two lines was computed. If it fell within the image boundaries, the criterion is met ; otherwise, the criterion is not satisfied. The resulting prediction was then compared against the consensus image-level annotation from the three radiographers, which corresponds to the predominant class among the three labels. The overall architecture of the framework, from landmark detection to positioning assessment, is illustrated in Figure 5.1.

5.4 Results and discussion

As previously mentioned, the test set used in this study was annotated independently by three radiographers, providing three sets of coordinates for each landmark (upper and lower PM points and the nipple). To quantify inter-operator variability, we computed, for each image,

the maximum distance between the three annotations of a given point and then averaged these distances across all images. The resulting mean maximum distances were 2.07 mm, 5.32 mm and 3.79 mm for the upper pectoral point, lower point, and nipple. These values reflect the inherent variability among expert annotations and should be considered when interpreting model errors, as predictions within this range can be regarded as comparable to human-level performance. See Figure 5.2 for examples of qualitative results.

5.4.1 Landmarks detection

We first evaluated the performance of our models on the localization of the PM line and the nipple coordinates. As a baseline, we used our data to train the CoordAttUNet proposed by [128], which predicts the two points of the PM line and the nipple position.

Table 1 reports the prediction errors for the upper and lower points of the PM line on the test set. The mean errors obtained with our model are slightly lower than those of the retrained CoordAttUNet, with a statistically significant difference for the lower point ($p = 0.004$) but not for the upper point ($p = 0.214$) according to the Wilcoxon test. Overall, the two models exhibit comparable and satisfying performances considering inter-annotator variability. The higher mean error observed for the lower point, consistent across models, can be partly explained by the greater inter-annotator variability, attributable to the general lack of clear and sharp gradients around the lower part of the PM.

Table 2 reports the mean and median localization errors for nipple detection. Our model outperforms the retrained CoordAttUNet, yielding a lower mean error (Wilcoxon test : $p = 1.41 \times 10^{-20}$). Although we report results on MLO images for direct comparison, we also trained our model using both CC and MLO views. When evaluated on MLO images only, the mean error was 2.00 mm, and it remained nearly unchanged (1.99 mm) on the combined CC + MLO test set, demonstrating strong versatility and generalization across views. Thus, our approach effectively doubles the available training images per dataset, unlike CoordAttUNet, which is limited to MLO views since it jointly detects the PM line and the nipple.

Overall, both of our proposed landmarks detection models achieve competitive results com-

Tableau 5.1 Error on lower pectoral point (ELP) and error on upper pectoral point (EUP)

Model	**ELP (mm)**		**EUP (mm)**	
	Mean	Median	Mean	Median
CoordAttUnet [128]	3.14	**1.65**	1.67	**1.20**
DeepLabv3+ (ours)	**3.01**	1.87	**1.49**	**1.19**

TABLEAU 5.2 Error on nipple coordinates (ENC) in mm.

Model	Mean	Median
CoordAttUnet [128]	3.42	2.16
Gaussian U-Net (ours)	2.40	**1.40**
Gaussian U-Net trained on MLO+CC	**2.00**	1.45

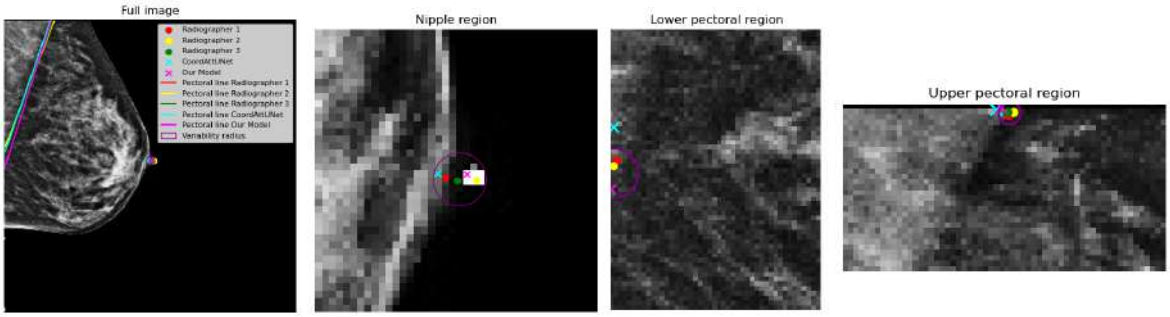


FIGURE 5.2 Example of qualitative results. For this image, the errors on the nipple and on the lower and upper pectoral points with our models are respectively 1.44 mm, 1.78 mm and 1.35 mm.

pared to [128]. Our DeepLabv3+ model provides reliable localization of the pectoral muscle points, achieving results comparable to those of the retrained CoordAttUNet. For nipple detection, our model outperforms CoordAttUNet and, in addition, CC and MLO views can be used to train our model with effectively twice the data using the same dataset, to reach even better performance. Our model also shows great generalization to both MLO and CC views when trained on both views.

TABLEAU 5.3 Classification metrics on positioning criterion *Adequate amount of pectoral muscle*. Bal.acc. : balanced accuracy

	Bal. acc.	**Specificity**	**Sensitivity**
CoordAttUnet [128]	0.79	0.91	0.68
CNN [1]	0.69	**0.95**	0.43
Our framework	**0.88**	0.92	**0.84**
R1 annotations	0.72	0.82	0.62
R2 annotations	0.69	0.71	0.68
R3 annotations	0.73	0.80	0.65

5.4.2 Positioning assessment

Our framework allows for the evaluation of criterion *Adequate amount of pectoral muscle* using a geometric approach based on the predicted anatomical landmarks. To establish performance baselines, we conducted two comparisons. First, we applied the same geometric algorithm to the landmarks predicted by the CoordAttUNet, in order to replicate the evaluation performed by [128], but on our dataset. Second, we implemented the model proposed by [1], who trained a separate CNN for many positioning criteria, including this one, and we trained this model on our dataset to obtain a classification-based baseline.

Table 5.3 summarizes the classification performance for the positioning adequacy criterion. Our proposed framework achieved the highest balanced accuracy (0.88), outperforming both baseline approaches. Compared to the CoordAttUnet [128], which reached a balanced accuracy of 0.79, our method demonstrates greater robustness across classes. The CNN baseline [1] obtained a high specificity (0.95) but a markedly low sensitivity (0.43), indicating that it tends to classify nearly all images as correctly positioned. This behavior is likely due to the strong class imbalance within our dataset. Such imbalance typically causes conventional CNNs to bias toward the majority class, achieving high specificity but poor detection of inadequate positioning. However, the accuracy of 0.968 reported in [1] was obtained on a perfectly balanced dataset, which does not accurately reflect the class distribution encountered in clinical practice. In contrast, our geometry-based model maintains a balanced performance (specificity = 0.92, sensitivity = 0.84), suggesting that integrating anatomical landmarks provides a more reliable representation of positioning adequacy. The difference in balanced accuracy between our method and CoordAttUnet may partly arise from our model's better nipple localization, which directly affects the geometric computation of the criterion. However, part of the observed difference may also stem from inter-rater variability in the manual annotations.

We additionally report the results obtained when the positioning criterion is evaluated geometrically using the ground-truth landmarks identified by each radiographer. Interestingly, the resulting balanced accuracies remain relatively low, indicating that the visual assessment of positioning adequacy by radiographers is subjective and does not always align with the outcome of a strict geometric evaluation, even when based on their own annotated landmarks. This suggests that the ground-truth labels may implicitly include an arbitrary tolerance margin, where some intersection points outside the expected region were still judged as acceptable, leading to minor inconsistencies between the visual ground-truth and the geometric computation performed by our framework.

5.5 Conclusion

In this work, we proposed a framework for automated landmarks detection in MLO mammographic views, using a DeepLabv3+ model for pectoral muscle segmentation and a Gaussian-based U-Net for nipple localization. These detected landmarks are then used in a geometric algorithm to assess the *Adequate amount of pectoral muscle* criterion. Experiments show higher balanced accuracy than existing baselines, with good sensitivity and specificity, demonstrating that interpretable landmark-based assessment provides a reliable basis for positioning evaluation even with imbalanced and subjective annotations. Future work will extend this approach to additional criteria, including nipple misorientation on CC views and pectoral muscle width estimation, further improving objective positioning assessment in mammography.

5.6 Compliance with ethical standards

This study was conducted retrospectively using human subjects data made available by iMD Research. Ethical approval was granted by the Ethics Committee of Polytechnique Montréal (CER-2425-26-D). All methods were carried out in accordance with the guidelines and regulations. The need to obtain the informed consent was waived by the Ethics Committee of Polytechnique Montréal.

5.7 Acknowledgments

The authors declare that they have no conflicts of interest. This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Medical Technology Consortium (MEDTEQ+), and iMD Research.

CHAPITRE 6 ASPECTS MÉTHODOLOGIQUES ET RÉSULTATS SUPPLÉMENTAIRES

Outre la méthodologie et les résultats présentés dans l'article du chapitre précédent, plusieurs autres expérimentations ont été effectuées dans le cadre de ce travail. Celles-ci sont décrites dans ce chapitre, qui offre également plus de détails concernant la méthodologie et la méthode proposée.

6.1 Détection du muscle pectoral

L'article 1 présenté au chapitre précédent résume la méthodologie employée pour l'implémentation d'un modèle de détection du muscle pectoral, qui repose sur une architecture DeepLabv3+. Dans cette section, les choix méthodologiques ayant mené à la solution proposée sont justifiés de manière plus détaillée.

6.1.1 Définition de la tâche et création des masques de vérité terrain

La tâche de détection du muscle pectoral peut être formulée de diverses manières, notamment comme une tâche de segmentation ou de régression. Dans le cadre de ce travail, l'objectif n'est pas de produire un masque de segmentation couvrant l'ensemble de la région du muscle, mais plutôt de prédire la droite correspondant à la PL, celle-ci constituant le repère utilisé pour l'évaluation des critères de positionnement relatifs à la visibilité du muscle pectoral. La sortie attendue est donc une droite s'étendant de la base du muscle jusqu'à son extrémité supérieure.

Une toute première approche envisagée a été d'appliquer des algorithmes de traitement d'images classiques afin d'isoler la région du muscle pectoral. L'analyse qualitative de ces résultats, obtenus avant même d'entamer la phase d'annotation décrite au chapitre 4, a mené à la décision de se tourner vers une approche par apprentissage profond. En effet, en raison de la variabilité anatomique, de la présence d'artéfacts et de la variabilité de contraste entre les images, ces méthodes de traitement d'images classiques se sont révélées peu généralisables. Dans ce contexte, une piste considérée a été de formuler ce problème comme une tâche de régression visant à prédire directement les points définissant cette droite. Les résultats de ces expérimentations sont présentés à la fin de cette section. Toutefois, conformément aux approches majoritairement rapportées dans la littérature pour la détection du muscle pectoral, une stratégie par segmentation a finalement été retenue.

Plus précisément, la méthode proposée repose sur l'entraînement d'un modèle de segmen-

tation chargé de prédire un masque triangulaire à partir duquel la droite recherchée peut ensuite être dérivée. Les annotations disponibles étant constituées de la droite correspondant à la PL, des masques de vérité terrain ont été générés à partir de cette information. La droite annotée a été prolongée jusqu'aux bords de l'image, puis la région triangulaire délimitée par cette droite et les frontières de l'image a été utilisée pour construire un masque binaire servant de vérité terrain. Le modèle de segmentation est ainsi entraîné à prédire ce masque, à partir duquel la PL est finalement extraite par une étape de post-traitement. Un exemple de masque de vérité terrain généré à partir de la ligne annotée est présenté à la figure 6.1.

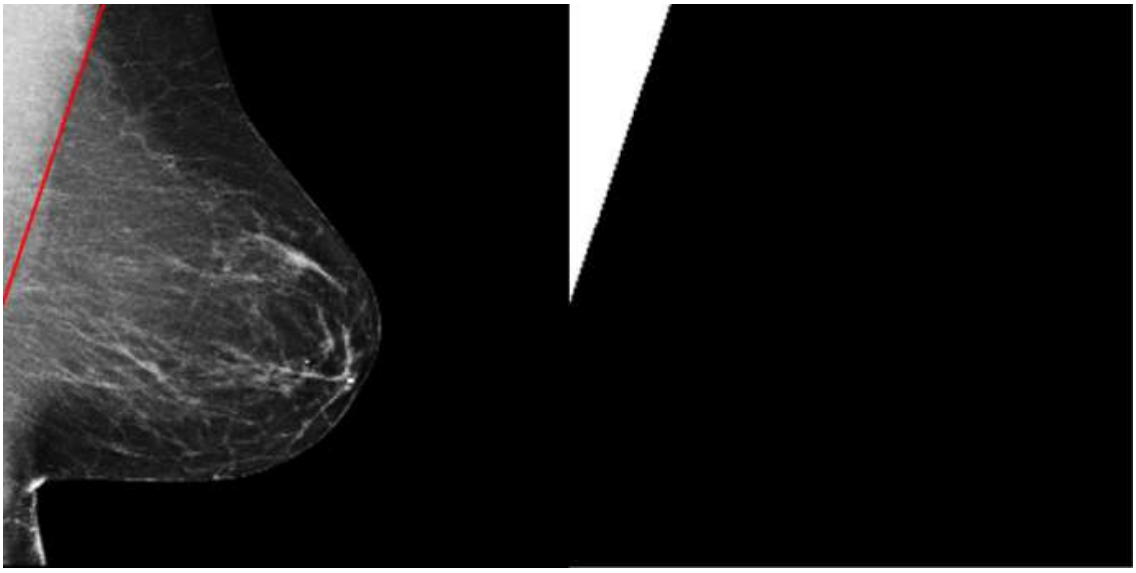


FIGURE 6.1 Vérité terrain : ligne en rouge à gauche, et masque binaire associé à droite.

6.1.2 Entraînement des modèles de segmentation

La revue de littérature a permis d'identifier les principales architectures utilisées pour la segmentation du muscle pectoral, soient des variantes de U-Net ainsi que DeepLab. L'intégration de modules d'attention est également rapportée [121]. Sur cette base, nous avons retenu pour nos expérimentations les architectures suivantes : U-Net [4], U-Net++ [33], Res-UNet [156], MA-Net (U-Net intégrant un module d'attention) [157] et DeepLabv3+ [5].

Les modèles ont été chargés via la librairie PyTorch. L'ensemble des modèles a été entraîné avec l'optimiseur Adam et la fonction d'activation ReLU. La fonction de perte utilisée est BCE-Dice, combinant l'entropie croisée binaire (BCE), qui pénalise les erreurs pixel par pixel, avec la perte Dice, qui mesure le recouvrement global entre la prédiction et la vérité terrain, afin d'optimiser à la fois la précision locale et la cohérence de la segmentation. Ce choix a été

fait aux termes d'expérimentations lors desquelles diverses fonctions de perte couramment utilisées en segmentation ont été testées. La fonction de perte est définie comme suit :

$$\mathcal{L}(y, \hat{y}) = \lambda \, \mathcal{L}_{\text{BCE}}(y, \hat{y}) + (1 - \lambda) \, \mathcal{L}_{\text{Dice}}(y, \hat{y}) \tag{6.1}$$

avec

$$\mathcal{L}_{\text{BCE}}(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \tag{6.2}$$

et

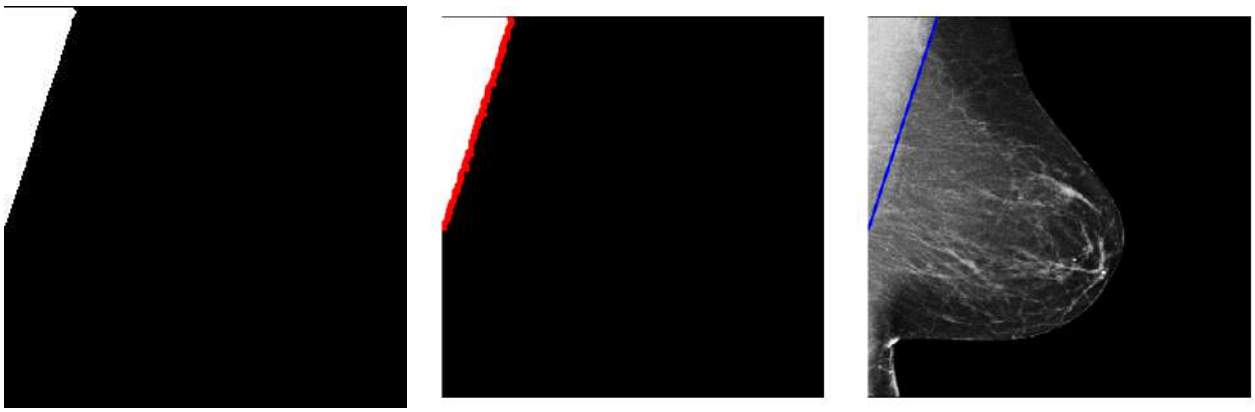
$$\mathcal{L}_{\text{Dice}}(y, \hat{y}) = 1 - \frac{2 \sum_{i=1}^{N} y_i \hat{y}_i + \varepsilon}{\sum_{i=1}^{N} \hat{y}_i + \sum_{i=1}^{N} y_i + \varepsilon} \tag{6.3}$$

en fixant $\lambda = 0,5$.

6.1.3 Extraction de la PL

À partir du masque de segmentation du muscle pectoral, une estimation de la PL est obtenue par une régression de type RANSAC. La méthode consiste d'abord à extraire, pour chaque ligne de l'image, le premier pixel appartenant au fond en partant du bord latéral correspondant au sein analysé (gauche ou droit), ce qui permet de constituer un ensemble de points situés sur la frontière du muscle pectoral. Ces points de contour sont ensuite filtrés afin d'éliminer ceux trop proches des bords de l'image, puis utilisés pour ajuster un modèle linéaire. L'algorithme RANSAC est employé afin de rendre l'estimation robuste aux points aberrants issus d'erreurs de segmentation locales ou de discontinuités du masque. La librairie *sklearn* de Python fournit une implémentation de l'algorithme de régression robuste, qui a été utilisée ici.

Une fois la droite estimée, ses paramètres (pente et ordonnée à l'origine) sont utilisés pour calculer les points d'intersection avec les bords de l'image, ce qui permet d'obtenir une représentation explicite de la ligne du muscle pectoral dans le repère de l'image. Cette approche combine ainsi une extraction simple de points de frontière et une régression robuste, garantissant une estimation stable de la ligne pectorale même en présence de bruit ou d'artefacts dans le masque de segmentation. Les étapes d'extraction de la PL sont illustrées à la figure 6.2.



(a) Masque prédit (b) Points du contour (c) Ligne extraite

FIGURE 6.2 Étapes d'extraction de la PL par l'algorithme de régression RANSAC

La détection de la PL étant réalisée en post-traitement, la comparaison des performances des modèles mentionnés à la section précédente a été effectuée sur le jeu de test. Les performances obtenues sur le jeu de validation (mesurées par la perte BCE-Dice) étaient d'un ordre de grandeur semblable pour l'ensemble des architectures. Le tableau 6.1 présente les erreurs moyennes obtenues sur les prédictions finales des points inférieur et supérieur de la PL, suite à son extraction à partir des masques prédits par les différentes modèles (*backbone* spécifié entre parenthèse).

Les architectures MANet et DeepLabv3+ sont celles offrant les meilleures performances de détection du muscle pectoral, minimisant les erreurs obtenues au niveau des points prédits. Le modèle DeepLabv3+ permet toutefois la meilleure détection du point inférieur du muscle, qui est de manière générale moins précise pour tous les modèles.

TABLEAU 6.1 Erreurs moyennes (mm) sur la prédiction du point supérieur (E.P.S.) et du point inférieur (E.P.I.) de la PL, selon le modèle

Modèle	**Pré-entraînement**	**E.P.S.**	**E.P.I.**
U-Net	Aucun	1,65	4,33
U-Net (ResNet50)	ImageNet	1,49	3,45
U-Net++ (EfficientNet)	ImageNet	1,49	3,47
MAnet (ResNet50)	ImageNet	**1,45**	3,25
DeepLabv3+ (ResNet101)	ImageNet	1,49	**3,01**

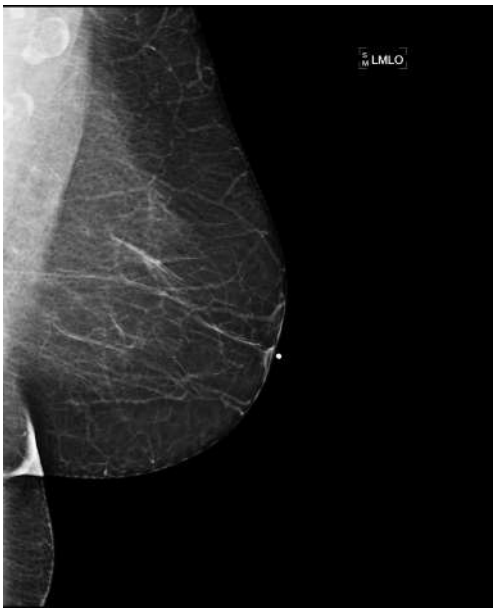
6.2 Détection du mamelon

La solution proposée pour la détection du mamelon, pour laquelle la méthodologie et les résultats sont présentés dans l'article 1, est un modèle U-Net gaussien, adapté de [7]. Certains aspects de la méthodologie sont approfondis ici, de même que quelques expérimentations additionnelles.

6.2.1 Choix de l'approche de régression

Le choix de l'approche par régression de cartes de chaleur a été motivée par divers éléments. D'abord, la revue de littérature a permis de faire ressortir les principales méthodes ayant été appliquées à la détection du mamelon. Au-delà des approches exploitant des propriétés géométriques et des caractéristiques extraites par l'humain, la majorité des solutions récentes reposent sur une segmentation de la région du mamelon ou sur la prédiction des coordonnées de celui-ci. Dans le cadre de ce travail, c'est cette dernière approche qui nous est pertinente, puisque l'évaluation des critères de positionnement repose sur la détection d'un point situé sur le mamelon. En outre, le mamelon peut apparaître sous diverses formes sur une image de mammographie, tel qu'illustré à la figure 6.3. En effet, un marqueur de plomb indique dans certains cas l'emplacement du mamelon. Dans d'autres, une protubérance ou une région d'intensité élevée caractérisent cette région. Le mamelon est parfois même invisible, rendant une segmentation impossible. Pour ces deux raisons, l'approche par segmentation a été écartée, la régression des coordonnées du mamelon ayant plutôt été privilégiée.

Dans la revue de littérature, deux approches de régression de points de repères ont été décrites : la régression directe et la régression de cartes de chaleur. Considérant les apparences très variables du mamelon, il nous a semblé logique de chercher à prédire une carte de chaleur autour de la région du mamelon. Par ailleurs, les travaux de Sharan et al. [7] ont démontré le potentiel d'une telle approche pour la tâche de détection de points de sutures. Le modèle proposé dans [7] a donc été repris et adapté à notre tâche de détection du mamelon. L'architecture du modèle, les modifications qui y ont été apportées et son entraînement sont décrits dans l'article 1. En complément, un schéma de l'architecture est présenté à la figure 6.4.



(a) Marqueur de plomb

(b) Protubérance

(c) Région d'intensité élevée

(d) Mamelon invisible

FIGURE 6.3 Différentes façons dont le mamelon peut se présenter sur une mammographie

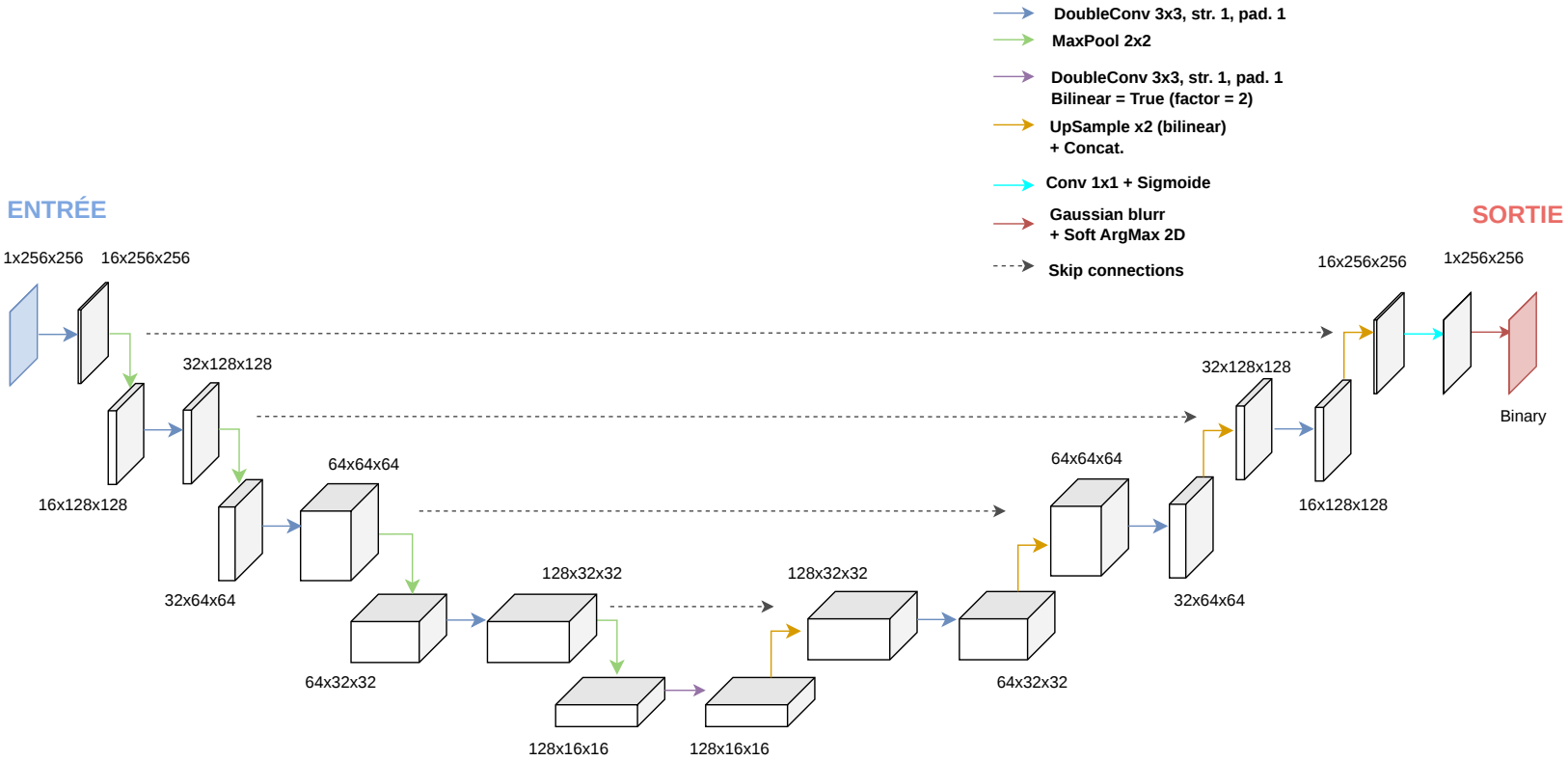
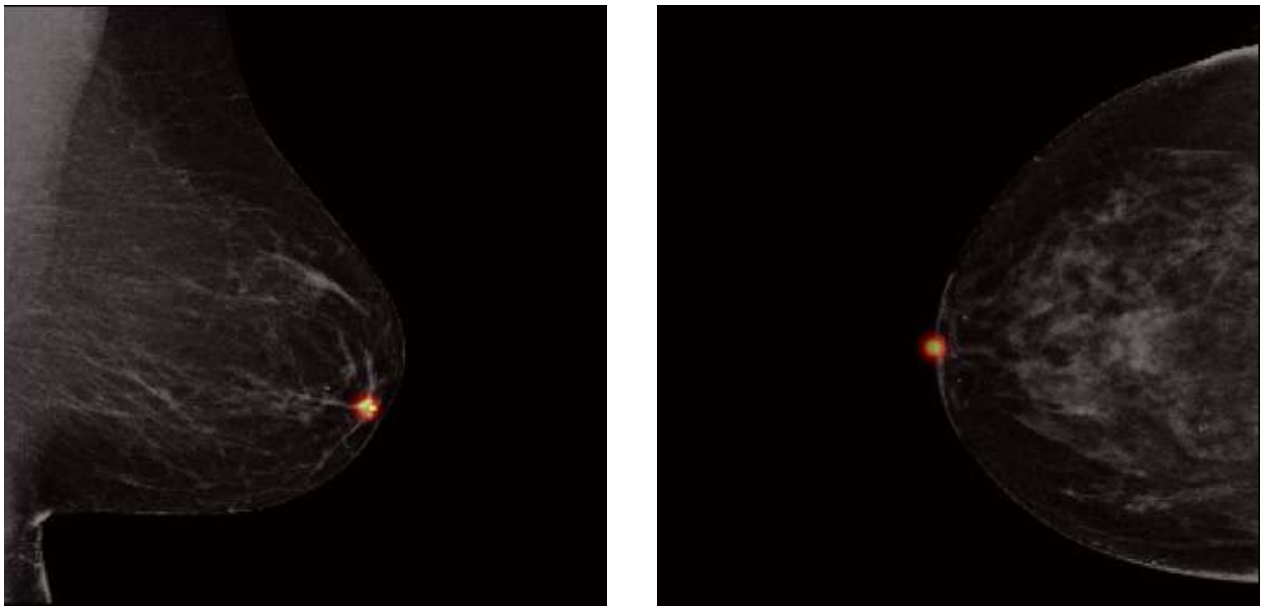


FIGURE 6.4 Architecture du U-Net gaussien

Tel qu'expliqué dans l'article 1, les cartes de chaleur utilisées comme vérités terrain ont été générées à partir des points annotés par les technologues, correspondant à la position du mamelon. Une carte de chaleur gaussienne centrée en ce point a été générée avec un écart-type $\sigma = 3$, contrôlant l'étalement spatial de la distribution. Cette valeur a été déterminée empiriquement. Un exemple de vérités terrain ainsi produites est présenté à la figure 6.5.



(a) MLO (b) CC

FIGURE 6.5 Cartes de chaleurs générées autour du mamelon

La fonction de perte retenue consiste en une combinaison de l'erreur quadratique moyenne (MSE) et d'un terme de localisation, telle que définie dans l'Article 1. Ce choix est le résultat d'une étude d'ablation menée afin d'analyser l'influence de différents termes de pénalisation sur les performances du modèle. Plus précisément, la MSE, la perte de Dice et le terme de localisation ont été évalués selon plusieurs configurations de combinaison. Les résultats de cette étude, compilés après 10 époch, sont consignés dans le tableau 6.2. L'erreur moyenne entre la prédiction et la vérité terrain, qui correspond à la moyenne des trois coordonnées annotées par les technologues (centroïde), est rapportée. En outre, la précision à 3,74 mm, c'est à dire le pourcentage de détections se situant à moins de 3,74 mm du centroïde, est présentée. Ceci correspond aux prédictions qui se situent à l'intérieur du rayon de variabilité inter-opérateur défini au chapitre 4, et qui sont donc considérées acceptables.

Tableau 6.2 Étude d'ablation sur 10 epochs : comparaison des performances de détection du mamelon (erreur moyenne, à minimiser, et précision à 3,74 mm, à maximiser) pour différentes fonctions de perte

	MSE	**Dice**	**Coord.**	**MSE + Dice**	**MSE + Coord.**	**Dice + Coord.**	**MSE + Dice + Coord.**
Erreur moyenne (mm)	3.66	3.93	111.23	4.21	**2.73**	4.55	3.14
Précision @ 3.74 mm	0.82	0.79	0.02	0.80	**0.87**	0.80	0.86

La combinaison minimisant l'erreur moyenne et maximisant la précision à l'intérieur du rayon de variabilité est l'erreur quadratique moyenne et le terme de localisation (Coord.). C'est donc cette configuration qui a été retenue pour la suite.

Finalement, afin d'évaluer la pertinence d'une approche par régression de cartes de chaleur par rapport à une régression directe pour la détection du mamelon, une étude d'ablation a été menée. Le modèle de prédiction des cartes de chaleur a été ré-entraîné en supprimant l'étape de floutage gaussien en sortie et en remplaçant les vérités terrain gaussiennes par des cibles ponctuelles (équivalent à fixer $\sigma = 0$). Les erreurs moyennes obtenues dans cette configuration se sont révélées être d'un ordre de grandeur très supérieur à celles observées avec l'approche par cartes de chaleur (résultats dans le tableau 6.3). En conséquence, cette piste n'a pas été approfondie.

Tableau 6.3 Comparaison des erreurs moyennes de détection du mamelon selon l'approche de régression employée

	Erreur moyenne (mm)
Régression de cartes de chaleur	2,59
Régression directe	74,90

Plusieurs éléments pourraient expliquer la hausse marquée de l'erreur moyenne observée avec l'approche par régression directe. D'abord, le modèle employé n'a pas été spécifiquement optimisé pour cette tâche, et la même fonction de perte, combinant MSE et un terme de localisation, a été utilisée dans les deux configurations. Bien que la MSE soit théoriquement appropriée pour une tâche de régression directe, cette formulation pourrait ne pas être optimale dans un contexte où il y a présence de valeurs aberrantes et où la variabilité anatomique est élevée. Il serait donc envisageable qu'une fonction de perte plus robuste améliore partiellement les performances.

Par ailleurs, nous émettons l'hypothèse que la forte variabilité d'apparence du mamelon, illustrée précédemment, pénalise davantage une approche par régression directe qu'une approche par cartes de chaleur. En effet, représenter la cible (mamelon parfois invisible ou caractérisé par une région allongée d'intensité élevée) sous forme de coordonnée ponctuelle impose une correspondance stricte entre une structure parfois ambiguë et une position unique, alors qu'une carte de chaleur permet de modéliser une distribution spatiale allouant une certaine incertitude morphologique.

Ces observations sont cohérentes avec la littérature en vision par ordinateur et en imagerie médicale. En estimation de pose humaine, les méthodes basées sur la prédiction de cartes de chaleur surpassent généralement les approches de régression directe des coordonnées, et constituent dorénavant l'état de l'art dans le domaine [158]. Des constats similaires ont été rapportés en imagerie médicale pour la localisation de repères anatomiques. Huang et al. [159] montrent que la régression directe nécessite davantage de données et demeure moins précise que les approches par cartes de chaleur, tandis que Payer et al. [160] soulignent que la modélisation d'un point sous forme de distribution spatiale facilite l'optimisation et exploite mieux le biais inductif convolutionnel des CNN. Plus précisément, la prédiction dense de cartes de chaleur fournit un signal de gradient spatialement distribué, alors que la régression directe condense l'information en un petit nombre de paramètres scalaires, rendant le problème d'optimisation plus difficile.

Ainsi, au-delà du choix de la fonction de perte, la dégradation observée semble également liée à une différence plus fondamentale de représentation et d'adéquation entre la tâche de localisation ponctuelle et les propriétés morphologiques du mamelon.

6.2.2 Identification du mamelon dans la vue CC

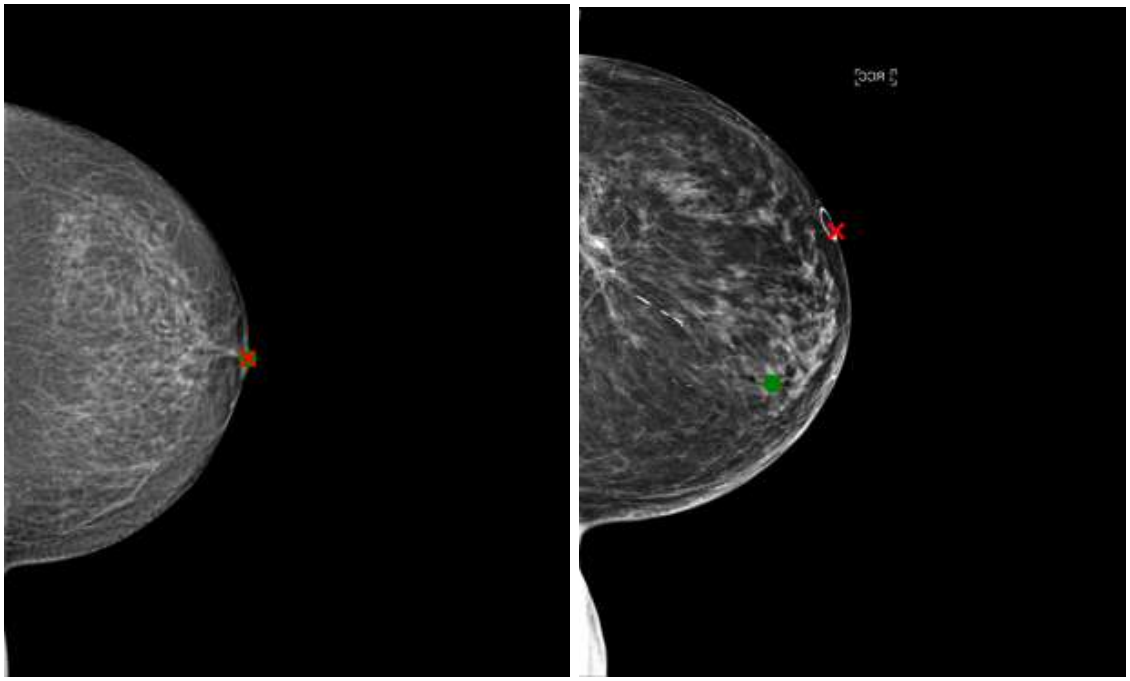
L'Article 1 présente les résultats de détection du mamelon sur un sous-ensemble de test n'incluant que la vue MLO afin de comparer les performances à celles du modèle CoordAttUNet [128], qui n'est applicable qu'à cette vue. Les résultats présentés ont d'ailleurs été obtenus suite à un entraînement uniquement sur les images MLO. Or, notre modèle U-Net gaussien peut être entraîné et évalué sur les vues MLO et CC. Lorsqu'entraîné sur l'ensemble des images (3 302 images dont 10 % pour la validation), notre modèle obtient les résultats présentées dans le tableau 6.4.

Le jeu de test MLO ne comprend que 446 images (au lieu de 454, ce qui aurait porté le total avec l'ensemble CC à 914) en raison du retrait de certaines images MLO pour lesquelles les annotations du mamelon et/ou de la ligne pectorale étaient manquantes (erreurs d'annotation ou images non évaluables). Ce choix a été motivé par la volonté d'utiliser un jeu de test

Tableau 6.4 Erreurs (en mm) de détection du mamelon par notre modèle, entraîné sur les vues MLO et CC

Jeu de test	**Moyenne**	**Médiane**
Sous-ensemble MLO (446 images)	**2,22**	2,00
Sous-ensemble CC (460 images)	6,42	**1,45**
Ensemble total (CC et MLO : 914 images)	2,59	2,4

strictement identique pour l'évaluation de CoordAttUNet et de notre modèle, afin de garantir une comparaison équitable. CoordAttUNet prédisant simultanément la position du mamelon et la ligne pectorale, la présence des deux vérités terrain était requise pour chaque image. À l'inverse, dans notre approche, la détection du mamelon et celle du muscle pectoral sont réalisées par des modèles distincts, ce qui a permis d'exploiter l'ensemble des images MLO pour l'entraînement du modèle de détection du mamelon, y compris lorsque l'annotation de la ligne pectorale était absente. En termes de précision, on constate que l'erreur moyenne de détection du mamelon est plus faible sur le sous-ensemble d'images MLO que sur celui ne contenant que des images CC. Toutefois, pour ce dernier, la médiane de l'erreur est inférieure, ce qui indique la présence de quelques valeurs aberrantes. Dans certains cas, ces valeurs sont dues à une morphologie particulière ou la présence d'artéfacts. Des exemples de prédictions obtenues en CC, incluant un cas où la distance entre la vérité terrain et la prédiction est très élevée, sont illutrées à la figure 6.6.



(a) Prédiction adéquate (b) Prédiction erronée

FIGURE 6.6 Exemple de prédiction du mamelon en CC (croix rouge) et vérité terrain associée (point vert)

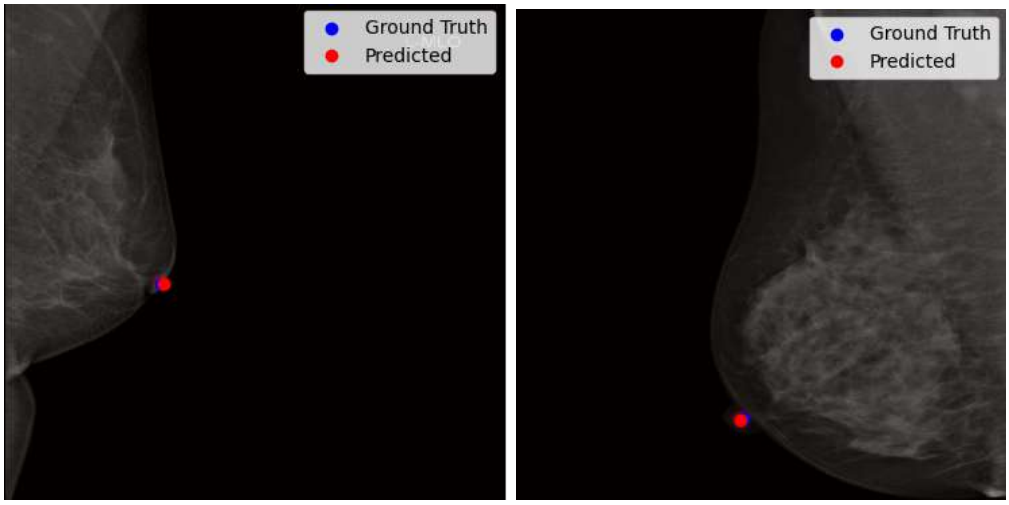
6.2.3 Identification du mamelon sur le jeu de données VinDr

À titre de comparaison supplémentaire, nous avons entraîné et évalué notre modèle de détection du mamelon sur un sous-ensemble du jeu de données public VinDr. Tanyel et al. [128], qui ont proposé le modèle CoordAttUNet utilisé comme référence pour la détection du mamelon et de la PL, ont rendu publiques les annotations du mamelon et de la PL pour 1 000 examens issus de ce jeu de données (2000 images MLO). Nous avons ainsi entraîné, validé et testé notre modèle U-Net gaussien sur ces données, en utilisant exactement la même répartition des images entre les ensembles d'entraînement, de validation et de test que celle employée par Tanyel et al. (80 % des images pour l'entraînement, 10 % pour la validation et 10 % pour le test). Les résultats de cette expérimentation sont consignés dans le tableau 6.5.

Tableau 6.5 Erreurs de détection du mamelon (mm) sur l'ensemble de test VinDr

Modèle	Moyenne	Médiane	Écart-type
CoordAttUnet [128]	2,97	2,45	2,46
U-Net gaussien	**1,82**	**1,26**	**2,08**

Ainsi, sur le même ensemble issu du jeu de données publique VinDr, notre modèle de détection du mamelon (U-Net gaussien) surpasse le modèle CoordAttUNet proposé par Tanyel et al. [128], illustrant la robustesse du modèle à différents jeux de données. Des exemples qualitatifs sont présentés à la figure 6.7.



(a) (b)

Figure 6.7 Exemples de prédiction (en bleu) et de vérité terrain (en rouge) sur le jeu de test VinDr

En ce qui concerne la détection de la PL, le format des vérités terrain fourni pour VinDr ne permettait pas d'entraîner directement notre modèle sans modifier la nature de la tâche. En effet, nos technologues tracent la PL du bas du muscle jusqu'à son extrémité supérieure en suivant l'orientation des fibres musculaires. Or, dans les données VinDr, la ligne est tracée de manière à être tangente à la base du muscle, sans nécessairement être parallèle aux fibres. Cette convention a pour effet d'inclure, dans la région délimitée par la ligne, une portion non négligeable de tissu mammaire ne correspondant pas au muscle pectoral. Notre modèle étant conçu pour segmenter la région triangulaire correspondant au muscle, nous n'avons pas réalisé de comparaison sur VinDr pour cette tâche, et nous nous sommes limités à appliquer CoordAttUNet à nos propres données. À titre d'exemple, la figure 6.8 illustre le genre d'annotations accessibles pour le dataset VinDr.

FIGURE 6.8 Exemple d'annotation de la ligne du muscle pectoral sur une image tirée du jeu de données VinDr

6.3 Évaluation du positionnement

L'article du chapitre précédent présente les résultats d'évaluation d'un seul critère de positionnement, soit *Quantité adéquate de muscle pectoral*. Trois autres critères peuvent être évalués à partir de nos prédictions de la PL et du mamelon : *Vue de la largeur maximale du muscle pectoral* ; *PNL perpendiculaire au bord de l'image* ; et *Longueur de la PNL en CC à moins de 1 cm de celle en MLO*. Les algorithmes développés afin d'évaluer ces critères sont

présentés dans cette section.

6.3.1 Métriques d'évaluation

Concernant les performances de nos algorithmes d'évaluation du positionnement, nous rapportons les métriques suivantes : la précision équilibrée (*balanced accuracy*), la spécificité (ou taux de vrais négatifs) et la sensibilité (ou taux de vrais positifs). Alors que la précision globale donne la proportion de prédictions correctes, la précision équilibrée permet de quantifier la moyenne des performances d'un modèle sur chaque classe. Dans le cas d'un déséquilibre entre les classes, la précision globale peut être trompeuse, dans le cas par exemple où la prédiction correspond presque toujours à la classe majoritaire. La précision équilibrée est définie en fonction de la sensibilité, qui représente le taux de cas réellement positifs parmi les cas identifiés comme positifs, et la spécificité, qui correspond au taux de cas réellement négatifs parmi ceux prédits comme étant négatifs. Les formules décrivant ces trois métriques sont les suivantes [161, 162] :

$$\text{Sensibilité} = \frac{VP}{VP + FN} \tag{6.4}$$

$$\text{Spécificité} = \frac{VN}{VN + FP} \tag{6.5}$$

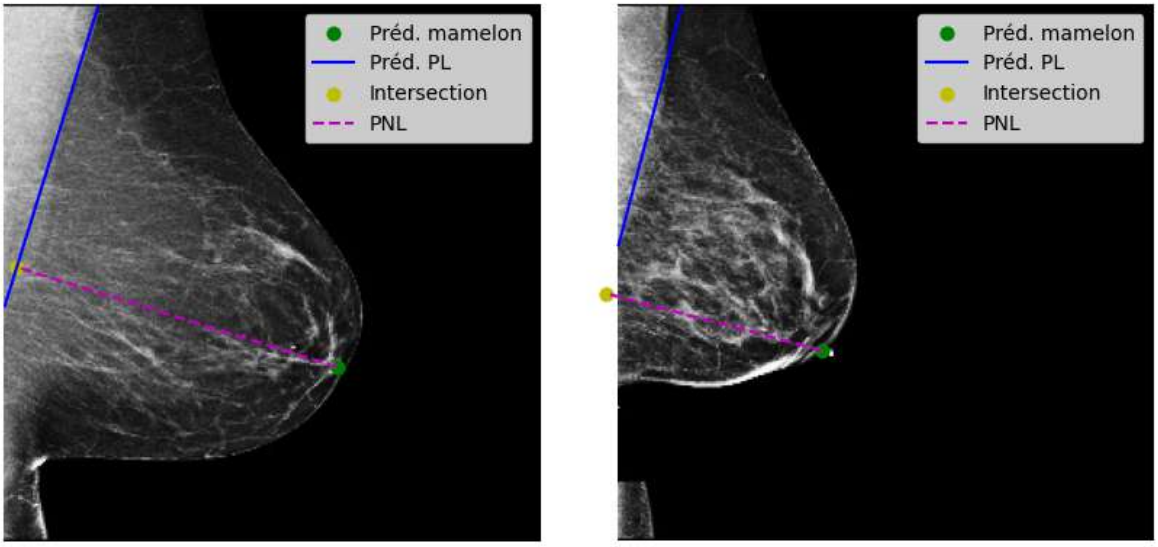
$$\text{Précision équilibrée} = \frac{\text{Sensibilité} + \text{Spécificité}}{2} = \frac{1}{2}\left(\frac{VP}{VP + FN} + \frac{VN}{VN + FP}\right) \tag{6.6}$$

où VP, VN, FP, et FN correspondent aux vrais positifs, vrais négatifs, faux positifs et faux négatifs respectivement. Dans le cas de l'évaluation des critères de positionnement, un cas positif correspond à un échec du critère. Dans un contexte médical, la détection des cas problématiques revêt une importance particulière. La sensibilité est donc une métrique importante à maximiser.

6.3.2 Quantité adéquate de muscle pectoral

Tel que décrit dans l'Article 1, l'évaluation de ce critère s'appuie sur un algorithme géométrique, prenant en entrée les prédictions de la PL et de la position du mamelon. La PNL, reliant le mamelon à la PL à angle droit, est tracée, puis l'intersection entre ces deux droites est calculée. Le critère est considéré satisfait lorsque cette intersection se situe dans les fron-

tières de l'image. Sur la figure 6.9, on peut voir un exemple de cas où le critère est satisfait ainsi qu'un exemple d'échec. Les résultats relatifs à l'évaluation de ce critère sont présentés dans l'Article 1.



(a) Succès du critère (intersection dans l'image)

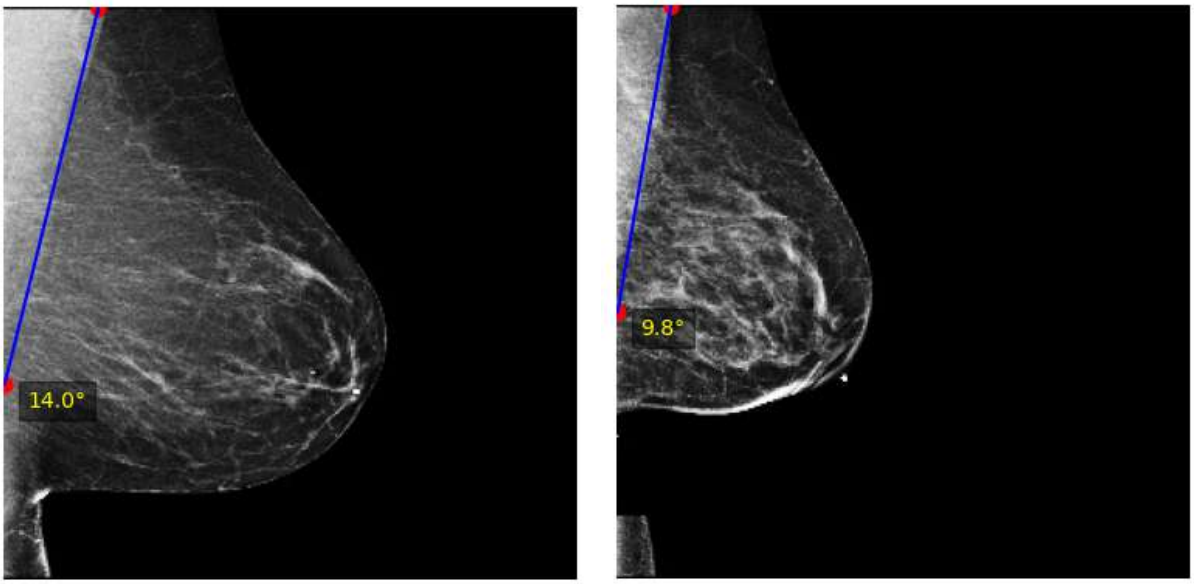
(b) Échec du critère (intersection hors de l'image)

FIGURE 6.9 Évaluation géométrique - *Quantité adéquate de muscle pectoral*

6.3.3 Vue de la largeur maximale du muscle pectoral

Ce critère vise à vérifier l'orientation adéquate du muscle sur la vue MLO. L'algorithme géométrique permettant de l'évaluer repose sur le calcul de l'angle formé entre la PL et le bord vertical de l'image. Si cet angle est inférieur à un seuil donné, le critère n'est pas satisfait ; dans le cas contraire, il est considéré comme respecté.

Bien que la définition du critère ne précise aucun seuil angulaire, une valeur de 10 degrés est rapportée dans la littérature pour l'évaluation d'un critère analogue [1]. Cette valeur a donc initialement été choisie comme seuil décisionel pour notre algorithme géométrique. Les technologues participant au projet nous ont par ailleurs indiqué que, quoique l'évaluation clinique ne repose pas sur l'application d'un seuil numérique stricte (mais plutôt sur leur jugement clinique et une appréciation visuelle), cet ordre de grandeur semble concorder avec la pratique clinique. La figure 6.10 illustre des exemples de succès et d'échecs du critère selon le seuil retenu de 10 degrés.



(a) Succès du critère (angle > 10 degrés) (b) Échec du critère (angle < 10 degrés)

FIGURE 6.10 Évaluation du critère d'angle du muscle

L'analyse des annotations fournies par les technologues a cependant révélé que la ligne tracée (PL) et l'évaluation binaire du critère ne sont pas systématiquement cohérentes avec un seuil fixe de 10 degrés. En d'autres termes, certaines images présentent une PL formant un angle inférieur à 10 degrés avec la verticale alors que le critère est jugé satisfait, et inversement.

Dans le cadre de l'étude de variabilité menée en collaboration avec les autres étudiants du projet [25], la relation entre l'angle mesuré à partir de la PL tracée par les technologues et le label binaire attribué à chaque image a été analysée. Cette étude a mis en évidence une variabilité importante, tant inter-annotatrice qu'intra-annotatrice. Pour un même angle, une technologue pouvait juger le critère satisfait dans un cas et non satisfait dans un autre. De plus, certaines technologues considéraient des images acceptables pour des angles variant de moins de 10 degrés à plus de 30 degrés, tandis que d'autres images étaient jugées inadéquates malgré des angles pouvant dépasser 20 degrés.

Afin de mieux représenter les annotations fournies et d'optimiser certaines métriques d'évaluation, un seuil angulaire a été estimé à partir des données annotées. Le seuil obtenu par maximisation de la précision équilibrée s'élève à 16,41 degrés, plutôt qu'aux 10 degrés initialement retenus. Ainsi, l'algorithme géométrique d'évaluation du positionnement a été appliqué aux images en utilisant ces deux seuils. Les performances résultantes sont consignées dans les tableaux 6.6 et 6.7.

À titre de comparaison, nous considérons les approches suivantes proposées dans la littérature :

— CoordAttUNet [128] : ce modèle permet la prédiction de la PL et du mamelon sur les images MLO. Nous avons entraîné le modèle proposé sur nos données, et appliquons à la PL prédite le même algorithme géométrique que celui utilisé avec nos propres prédictions. Cette stratégie permet une comparaison équitable au niveau de l'évaluation du critère, indépendamment de la méthode de détection des repères.

— Un CNN peu profond : Brahim et al. [1] ont proposé une architecture convolutive pour l'évaluation du critère relatif à l'angle du muscle pectoral (qui correspond à *Vue de la largeur maximale du muscle pectoral* dans notre liste de critères). L'architecture présentée dans [1] a été implémentée, et l'entraînement du modèle sur nos données a été effectué par un membre de notre équipe. Ceci permet la comparaison de notre approche géométrique avec une approche de classification bout-en-bout.

— Annotations T1, T2 et T3 : Chaque technologue (T) ayant annoté toutes les images du jeu de test, nous avons appliqué notre algorithme géométrique aux lignes tracées par les technologues et comparé la prédictions résultante au label majoritaire issu des annotations binaires.

Tableau 6.6 *Vue de la largeur maximale du muscle pectoral* : Métriques d'évaluations (seuil de 10 degrés)

	Précision équilibrée	**Spécificité**	**Sensibilité**
CoordAttUNet [128]	0,5	1,0	0,0
CNN [1]	**0,83**	0,79	0,86
Notre approche	0,62	0,99	0,24
Annotations T1	0,62	0,99	0,24
Annotations T2	0,62	0,99	0,24
Annotations T3	0,57	1,0	0,14

Tableau 6.7 *Vue de la largeur maximale du muscle pectoral* : Métriques d'évaluations (seuil de 16,41 degrés)

	Précision équilibrée	**Spécificité**	**Sensibilité**
CoordAttUNet	0,69	0,96	0,41
CNN	**0,83**	0,79	**0,86**
Notre approche	**0,83**	**0,83**	0,83
Annotations T1	0,79	0,84	0,74
Annotations T2	0,83	0,84	0,82
Annotations T3	0,80	0,87	0,73

Lorsque le seuil est fixé est 10 degrés (tableau 6.6), les performances obtenues avec les approches géométriques (CoordAttUNet et la nôtre) sont bien en deça de celles atteintes par le CNN. Pour ce dernier, aucun seuil d'angle n'est pris en compte ; la classe binaire attribuée en sortie du modèle est directement comparée à au label majoritaire apposé par les technologues. Par ailleurs, les précisions équilibrées obtenue en appliquant l'algorithme géométrique aux annotations visuelles (PL) identifiées par les technologues demeurent faibles, indiquant une différence entre l'évaluation binaire du critère et l'identification des repères anatomiques. Ces résultats semblent indiquer que le seuil de 10 degrés ne soit pas le choix idéal considérant les données et les annotations dont nous disposons.

Dans le cas où le seuil est plutôt fixé à 16,41 degrés (tableau 6.7), notre approche atteint 0,83 de précision équilibrée, égalisant celle obtenue par le réseau convolutif. Ce dernier offre cependant une meilleure sensibilité que notre méthode. Bien que le seuil de 16,41 degrés semble plus adapté aux annotations et permette d'obtenir de meilleurs résultats, la variabilité inter-annotatrice subsiste. En effet, chaque technologue (prise séparément) obtient une précision équilibrée ne dépassant pas 0,83 lorsqu'on compare la prédiction issue de l'application de l'algorithme géométrique à la PL tracée au label majoritaire. Cette variabilité implique donc qu'il serait difficile d'obtenir des performances plus élevées pour ce critère, considérant les annotations dont on dispose.

En outre, peu importe le seuil choisi, il faut considérer qu'un des éléments pris en compte par les technologues dans l'évaluation de ce critère, soit la croissance continue du muscle de son extrémité inférieure vers son extrémité supérieure, n'est pas intégré dans l'algorithme d'évaluation géométrique. Cette omission peut contribuer à certaines discordances observées entre la vérité terrain et la prédiction obtenue par notre approche.

Bref, notre approche géométrique permet d'obtenir une précision équilibrée égale à celle atteinte par le CNN, tout en offrant une identification visuelle de la PL et mesure quantitative de l'angle formé, favorisant l'interprétabilité de la prédiction.

6.3.4 PNL perpendiculaire au bord de l'image

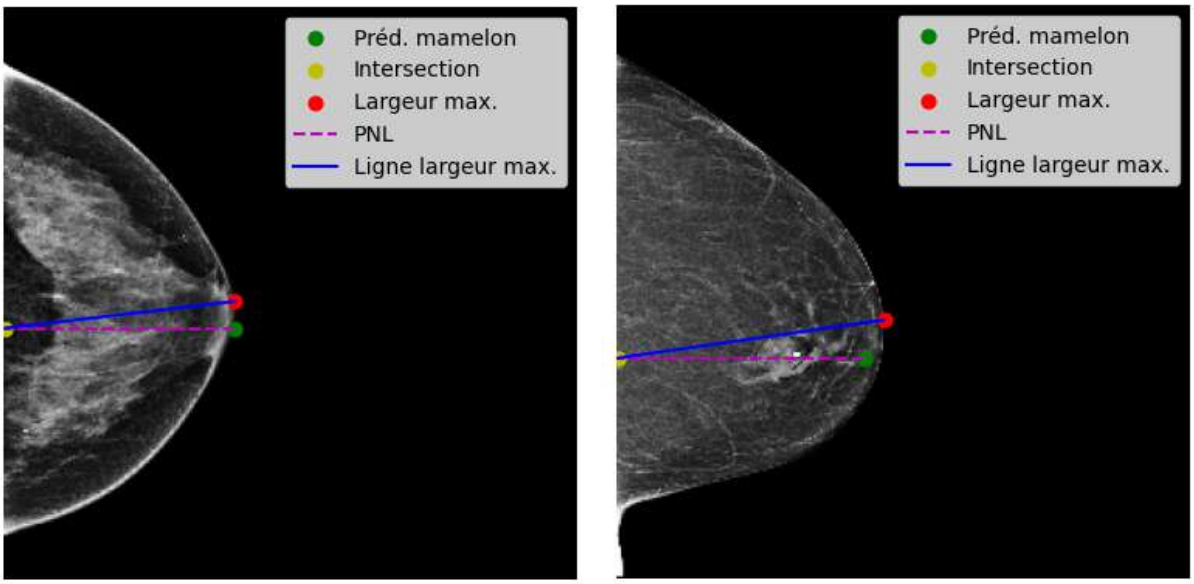
Ce critère vise à identifier une désorientation du mamelon par rapport à l'horizontale, dans la vue CC. Selon sa définition, le mamelon est considéré comme désorienté lorsque l'angle formé entre la PNL (ligne perpendiculaire au thorax passant par le mamelon) et la ligne de largeur maximale (définie comme la ligne reliant le point de largeur maximale du sein au bord de l'image, à son intersection avec la PL) dépasse un certain seuil angulaire. Un seuil de 5 degrés a été retenu pour l'évaluation de ce critère. Ce seuil n'est pas explicitement défini dans la littérature ni formalisé en pratique clinique ; il découle plutôt des discussions menées

avec les technologues participant au projet, qui ont indiqué qu'un ordre de grandeur de 5 degrés correspondait davantage à leur appréciation clinique. Contrairement au seuil utilisé pour le critère d'orientation du muscle pectoral, cette valeur diffère de celle rapportée par Brahim et al. [1], qui mentionnent plutôt un seuil de 20 degrés pour conclure à l'échec du critère.

L'algorithme géométrique d'évaluation du critère repose sur la détection préalable du mamelon, à partir de laquelle la PNL peut être tracée. Le point de largeur maximale est identifié au moyen d'un algorithme de détection de contour développé par un membre de notre équipe. La ligne reliant ce point au bord de l'image, à l'intersection avec la PNL, est obtenue, puis l'angle entre ces deux lignes est calculé. Si l'angle résultant est au delà du seuil de 5 degrés, le critère échoue. La figure 6.11 montre un cas de succès du critère et un cas d'échec de celui-ci.

Il convient toutefois de préciser que, en pratique, les technologues ne s'appuient pas sur une mesure angulaire numérique stricte, mais sur une appréciation visuelle. Cette absence de seuil formel est confirmée par notre analyse de variabilité [25]. Dans cette analyse, un membre de notre équipe a examiné la relation entre l'angle formé par les lignes tracées par les technologues (PNL et ligne de largeur maximale) et le label binaire qu'elles ont attribué pour évaluer le critère. Les résultats montrent une variabilité inter-annotatrice notable : certaines technologues tolèrent des angles pouvant atteindre environ 10 degrés, tandis que d'autres concluent à un échec pour des désorientations inférieures à 3 degrés. En outre, il existe une variabilité intra-annotatrice, c'est-à-dire qu'une même technologue n'appose pas toujours le même label pour un angle donné.

Ainsi, le seuil de 5 degrés correspond à la valeur qui permet d'obtenir la meilleure concordance globale avec les annotations disponibles (si on maximise la précision équilibrée). Il doit néanmoins être interprété comme une approximation opérationnelle plutôt que comme un seuil clinique formellement établi.



(a) Succès du critère (angle < 5 degrés) (b) Échec du critère (angle > 5 degrés)

FIGURE 6.11 Évaluation du critère relatif à l'orientation du mamelon

À notre connaissance, aucune méthode de référence n'a été rapportée dans la littérature pour l'évaluation de ce critère. En effet, Tanyel et al. [128] ne se sont pas intéressés aux critères spécifiques à la vue CC, et Brahim et al. [1] ont retiré ce critère de leur évaluation, estimant ne pas disposer d'un nombre suffisant de cas d'échec pour permettre une analyse robuste. Ils ont toutefois proposé un modèle pour évaluer le critère *Mamelon vu de profil*. En l'absence de méthode publiée, nous avons entraîné, à titre comparatif, le modèle présenté par Brahim et al. [1] pour la classification des images CC selon le critère *Mamelon vu de profil*. Cet entraînement a été réalisé par une autre étudiante impliquée dans le projet, dans le cadre de son propre travail de maîtrise. Les performances de cette approche de référence ainsi que celles de notre méthode sont présentées au tableau 6.8. Nous y intégrons également les métriques obtenues en comparant les annotations visuelles apposées par chaque technologue au label majoritaire.

TABLEAU 6.8 *PNL perpendiculaire au bord de l'image* : Métriques d'évaluations

	Précision équilibrée	**Spécificité**	**Sensibilité**
CNN [1]	0,5	0	1
Notre approche	**0,71**	**0,96**	0,46
Annotations T1	0,57	0,62	0,51
Annotations T2	0,54	0,59	0,48
Annotations T3	0,54	0,71	0,36

Dans le cas de ce critère, le CNN ne semble pas être en mesure de discriminer les cas d'échecs des cas corrects, classifiant toutes les images dans la classe négative. En outre, lorsqu'on calcule l'angle entre les lignes tracées par les technologues, soient la PNL et la ligne de largeur maximale, et qu'on y applique le seuil décisionnel de 5 degrés, les précisions obtenues sont faibles, tel que rapporté dans le tableau 6.8. Encore une fois, ceci témoigne des discordances entre les évaluations binaires du critères faites par les technologues et l'identification visuelle des repères.

6.3.5 Longueur de la PNL en CC à moins de 1 cm de celle en MLO

Ce critère permet de vérifier si la vue CC inclut une profondeur suffisante de tissu mammaire. Dans la vue MLO, cet élément est évalué en considérant la visibilité du muscle pectoral. Or, comme ce dernier n'apparaît pas (du moins presque jamais) dans la vue CC, une comparaison entre les deux vues est nécessaire. Ainsi, si la longueur de la PNL en CC est inférieure à celle de la MLO de plus de 1 cm, le critère échoue. À l'opposition des autres seuils décisionels utilisés dans ce travail, celui-ci est le seul qui soit clairement établi et fasse partie de la définition du critère. Une particularité relative à l'évaluation de ce critère est que la PNL, normalement définie comme la ligne reliant le mamelon au muscle pectoral (en vue MLO) ou au bord de l'image (en vue CC), doit dans ce cas-ci être mesurée jusqu'au contour réel du sein. Ainsi, si le mamelon n'est pas situé exactement sur le contour mammaire, la PNL doit être prolongée au-delà du mamelon jusqu'à son intersection avec le contour du sein. En outre, dans le cas où l'intersection entre la PL et la PNL en MLO est hors des frontières de l'image, la distance doit être calculée en s'arrêtant au bord de l'image.

L'algorithme géométrique développé pour évaluer ce critère repose sur trois étapes principales : la détection de la PL en vue MLO, la détection du mamelon dans les deux incidences et la détection du contour du sein. Pour cette dernière, le même algorithme de détection de contours que celui mentionné précédemment a été utilisé. Il faut mentionner que ce dernier est sensible à la présence d'artefacts, tel qu'illustré à la figure 6.12.

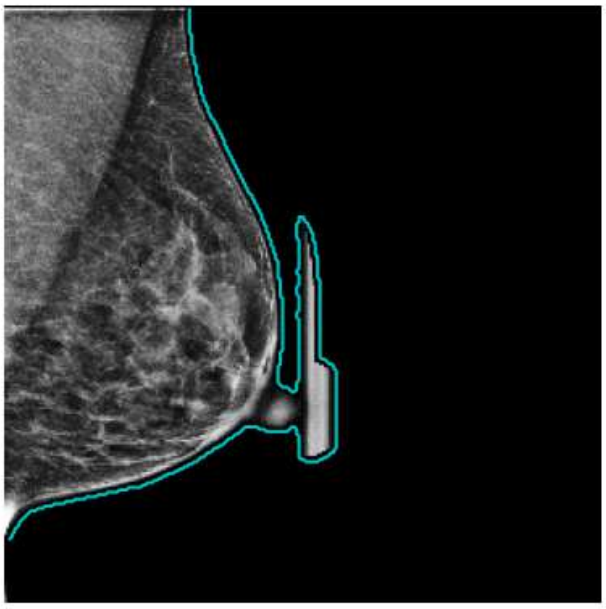


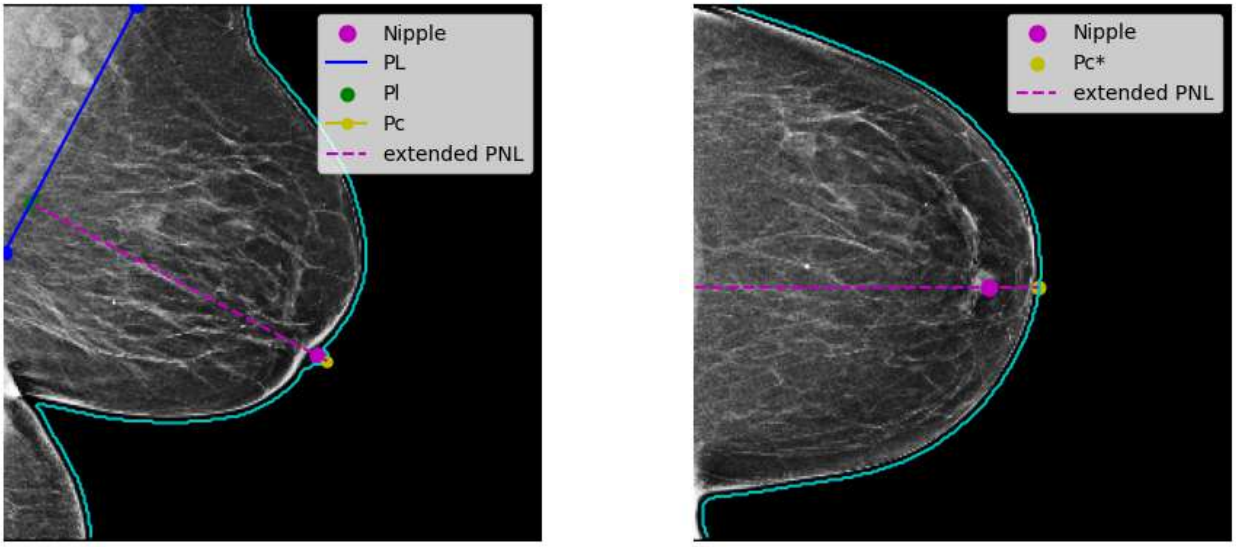
FIGURE 6.12 Exemple d'erreur de détection de contour due à la présence d'artefact

La figure 6.13 illustre les éléments géométriques pris en compte dans l'évaluation de ce critère.

En vue MLO, la PNL est définie à partir des prédictions de la PL et de la position du mamelon. Cette ligne est prolongée jusqu'à son intersection avec le contour du sein, et le point d'intersection obtenu est noté P_c. Le point d'intersection entre la PL et la PNL est également calculé et est noté P_l. Lorsque P_l se situe à l'extérieur des limites de l'image, l'intersection entre la PNL et le bord de l'image est déterminée (P_l'). La longueur de la PNL en MLO correspond ensuite à la distance entre les points P_c et P_l ou P_c et P_l' selon le cas.

En vue CC, la position prédite du mamelon ainsi que le contour détecté sont utilisés pour déterminer le point P_c*, correspondant au point du contour mammaire situé au même niveau vertical que le mamelon. La longueur de la PNL en CC est définie comme la distance horizontale entre le bord postérieur de l'image et ce point, ce qui revient à considérer la coordonnée x de P_c*. La différence entre les longueurs de la PNL mesurées en MLO et en CC est ensuite calculée. Si cette différence excède le seuil de 1 cm, le critère est considéré comme échoué.

La prédiction obtenue est comparée à la vérité terrain issue des annotations réalisées par les technologues. Ces dernières n'ont pas attribué de label binaire explicite ; elles ont plutôt tracé les lignes pertinentes sur les images. Les différences de longueur ont ensuite été calculées a posteriori par un membre de notre équipe, en appliquant le seuil de 1 cm pour déterminer la réussite ou l'échec du critère. Le label majoritaire résultant de ce calcul a été retenu comme vérité terrain. Ainsi, si le calcul des différences de longueur à partir des tracés de



(a) Vue MLO

(b) Vue CC

FIGURE 6.13 Éléments géométriques pris en compte pour évaluer le critère *PNL en CC à moins de 1 cm de celle en MLO*

deux technologues sur trois conduit à un échec du critère, celui-ci est comptabilisé comme un échec, et inversement.

À notre connaissance, aucune méthode de référence n'a été publiée dans la littérature pour évaluer ce critère. Nous rapportons donc les métriques d'évaluation obtenues en comparant nos prédictions avec le label majoritaire, ainsi qu'en comparant les annotations de chaque technologue au label majoritaire. Ces résultats sont consignés dans le tableau 6.9

TABLEAU 6.9 *Longueur de la PNL en CC à moins de 1 cm de celle en MLO* : Métriques d'évaluations

	Précision équilibrée	**Spécificité**	**Sensibilité**
Notre approche	0,84	0,82	0,85
Annotations T1	0,96	0,95	0,98
Annotations T2	0,97	0,96	0,97
Annotations T3	0,84	0,85	0,82

Notre approche permet d'atteindre une précision équilibrée de 0,84, une spécificité de 0,82 et une sensibilité de 0,85. Ces performances sont comparables à celles obtenues à partir des annotations de la technologue T3, dont l'évaluation semble différer de celle des deux autres technologues. Ceci implique qu'une certaine subjectivité subsiste dans l'évaluation de ce critère.

6.4 Expérimentations supplémentaires

Cette section survole quelques expérimentations supplémentaires effectuées au cours de ce projet. Celles-ci constituent des pistes explorées mais ne font pas partie de la solution finale proposée.

Méthodes classiques de traitement d'image

Avant même d'avoir accès aux annotations de la PL issues du processus décrit au chapitre 4, des algorithmes de détection du muscle pectoral, basées sur des méthodes classiques de traitement d'images, ont été appliquées au jeu de données initial de 1923 examens (vues MLO seulement). En collaboration avec un des autres étudiants participant au projet, deux approches ont été développées. La première, inspirée de [97], repose sur une chaîne de traitement combinant segmentation d'histogramme, opérations morphologiques et analyse de contours afin d'extraire la courbe du muscle pectoral. Dans un premier temps, la région d'intérêt est isolée par une méthode de croissance de région, ce qui permet de se focaliser sur la zone pertinente de l'image. La segmentation est ensuite réalisée à partir de l'histogramme des intensités, en identifiant automatiquement des seuils par élagage des extrema, puis en sélectionnant les minima les plus significatifs. Le nombre de seuils est ajusté dynamiquement en fonction de la densité mammaire, estimée au préalable, afin d'adapter la partition de l'image aux variations de contraste et de texture. Les régions obtenues sont ensuite raffinées par des opérations morphologiques, notamment une ouverture visant à corriger les discontinuités et les artefacts au sein de la région du muscle pectoral. Les contours sont alors extraits à l'aide d'un filtrage de Sobel pour mettre en évidence les gradients forts, suivi d'une détection des contours avec OpenCV, puis d'une sélection des contours pertinents selon des critères géométriques et de position (taille, localisation dans le quadrant haut gauche, etc.). Enfin, un ajustement polynomial (*polyfit*) est appliqué sur les points de contour retenus afin d'obtenir une approximation lisse et robuste de la courbe du muscle pectoral. Cette première approche est subséquemment référencée par le terme *Polyfit*.

La seconde approche, que nous avons surnomée *Drop Line*, vise à localiser le muscle pectoral en exploitant les variations brutales d'intensité le long de lignes horizontales partant du coin supérieur gauche de l'image (après symétrie si nécessaire pour les images droites). Pour un nombre fixé de lignes en partant du haut, on recherche, sur chaque ligne, le point correspondant à une chute significative d'intensité, détectée à l'aide d'un seuil prédéfini, ce qui permet d'obtenir un ensemble de points candidats appartenant à la frontière du muscle pectoral. À partir de ces points, deux modèles sont ajustés : une droite et une courbe quadratique,

chacune modélisant au mieux la distribution des points détectés. Les deux approximations sont ensuite comparées en analysant la différence d'aire entre les régions qu'elles délimitent ainsi que les variations d'intensité au sein de ces régions. Une forte divergence entre les deux modèles est interprétée comme une inclusion de tissus non pertinents ; dans ce cas, le modèle qui englobe le moins de tissu non musculaire est conservé. Cette stratégie permet ainsi de sélectionner une estimation plus robuste et plus spécifique de la frontière du muscle pectoral, tout en restant simple et efficace computationnellement.

Les deux méthodes, *Polyfit* et *Drop Line*, ont été toutes deux appliquées aux images, et un algorithme a été conçu pour déterminer automatiquement la méthode la mieux adaptée à l'image. Le figure 6.14 présente des exemples de résultats obtenus par ces méthodes.

(a) *Polyfit* (b) *Dropline* (linéaire) (c) *Dropline* (quadratique)

FIGURE 6.14 Exemples de détection - algorithmes de traitement d'image classique

Malgré certains cas où le muscle pectoral est bien détecté, une évaluation qualitative des résultats a tout de même permis de constater que pour près de 30% des images, la segmentation est erronée. Ainsi, l'approche par traitement d'image classique pour la détection du muscle pectoral n'a pas été explorée davantage.

Des méthodes de traitement d'image traditionnelles ont également été appliquées à la tâche de détection du mamelon. Ces expérimentations ont été menées par une autre membre de l'équipe, qui a développé un algorithme permettant d'identifier la position du mamelon dans 4 cas de figure, soient la présence d'un marqueur de plomb, d'une protubérance, d'une plaque d'intensité élevée ou encore un mamelon invisible. La méthodologie suivie dans le cadre de ces expérimentations n'est pas détaillée dans ce mémoire, mais les résultats de celles-ci ont été utilisés, c'est pourquoi nous en faisons mention.

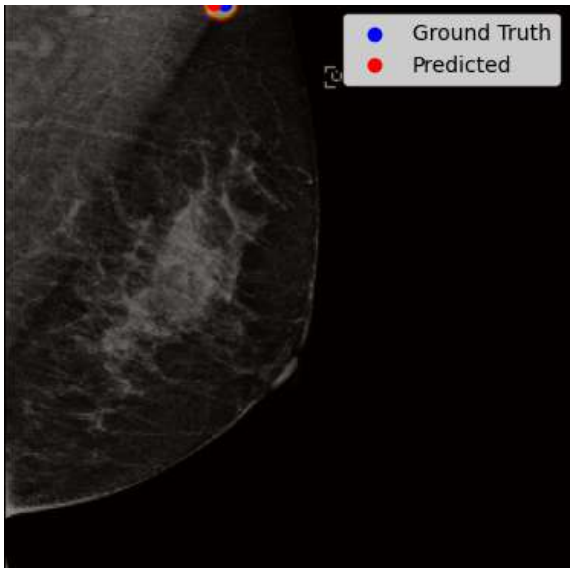
Régression des points formant la PL

L'approche de détection du mamelon par régression de cartes de chaleur montrant des performances intéressantes, nous avons tenté d'appliquer cette méthode à la détection des points formant la PL. Le modèle U-Net gaussien utilisé pour la détection du mamelon a donc été entraîné pour prédire, dans un premier temps, le point supérieur du muscle pectoral. Un entraînement distinct a été mené pour la prédiction du point inférieur. De la même manière que pour la détection du mamelon, des cartes de chaleur ont été générées autours des points correspondant aux annotations afin de former les vérité terrain. Les erreurs moyennes de prédiction sont consignées dans le tableau 6.10.

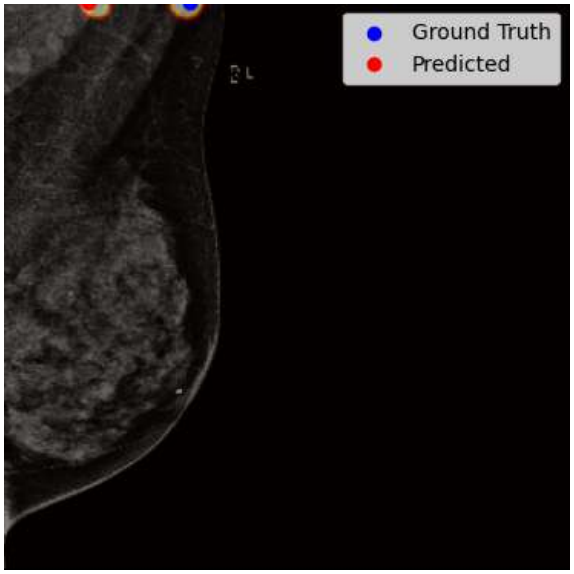
On constate que l'approche par régression de cartes de chaleur induit de très grandes erreurs, particulièrement pour la détection du point supérieur de la PL. Des exemples qualitatifs de prédictions sont présentés à la figure 6.15, où on voit la carte de chaleur prédite ainsi que la position obtenue en conservant la maximum de celle-ci. On remarque que dans la région inférieure du muscle, les cartes de chaleur prédites sont de forme allongée et s'étendent sur le bord de l'image. Ce phénomène est potentiellment attribuable au fait que le muscle pectoral soit généralement plus flou et sa frontière moins bien définie à cet endroit. Ceci est reflété dans les annotations des technologues, pour lesquelles une plus grande variabilité est observée au niveau du point inférieur de la PL. Nous pouvons également émettre l'hypothèse que, comme les points à prédire sont situés sur les bords de l'image et que la carte de chaleur créée comme vérité terrain est tronquée, le modèle peine à prédire une carte de chaleur précise.

Tableau 6.10 Erreurs moyennes de prédiction du point supérieur (E.P.S) et du point inférieur (E.P.I) du muscle pectoral selon l'approche employée

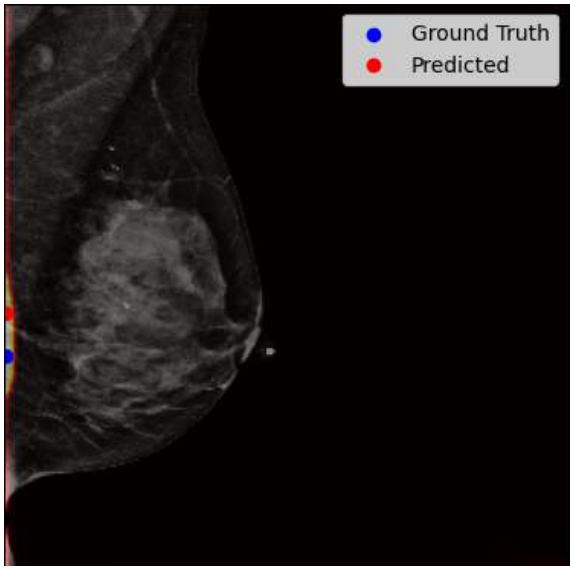
Approche	**E.P.S**	**E.P.I**
Régression de cartes de chaleur	4,24	27,8
Segmentation + RANSAC	**1,49**	**3,01**



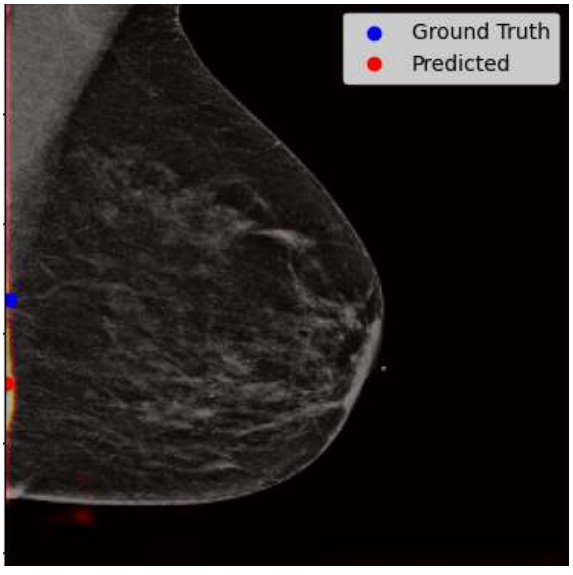
(a) Ex. 1 : Point supérieur



(b) Ex. 2 : Point supérieur



(c) Ex. 1 : Point inférieur



(d) Ex. 2 : Point inférieur

FIGURE 6.15 Exemples de cartes de chaleur prédites, coordonnées prédites et coordonnées réelles

CHAPITRE 7 DISCUSSION GÉNÉRALE

Le travail présenté dans ce mémoire s'inscrit dans un projet de grande envergure mené au sein du laboratoire VisionIC, ayant pour objectif le développement d'un système d'évaluation automatique du positionnement des seins en mammographie. En effet, un positionnement mammaire inadéquat lors de l'examen de dépistage peut entraîner des faux négatifs dans la détection du cancer du sein ainsi que le rappel des patientes pour un réexamen. Tel que présenté au début de ce mémoire, un ensemble de 17 critères a été proposé par l'OTIMROEPMQ afin d'évaluer la qualité du positionnement des seins sur les images de mammographie (vues CC et MLO).

Dans le cadre de ce projet, regroupant trois étudiants à la maîtrise, diverses avenues ont été explorées et différents critères ont été étudiés. Ultimement, les modèles développés pour évaluer chacun de ces critères seront combinés afin de former un outil intégré d'évaluation du positionnement, qui pourra être utilisé en clinique lors de l'acquisition des images pour offrir une rétroaction directe aux technologues en mammographie, leur permettant ainsi de reprendre l'image en cas de positionnement suboptimal.

Plus précisément dans ce mémoire, nous avons proposé une approche géométrique basée sur l'identification de repères anatomiques pour l'évaluation de quatre critères de positionnement. En effet, une définition géométrique (parfois simplifiée) reposant sur l'identification de la PL et du mamelon, peut être utilisée pour traiter les critères *Quantité adéquate de muscle pectoral*, *Vue de la largeur maximale du muscle pectoral*, *PNL perpendiculaire au bord de l'image* et *PNL en CC à moins de 1 cm de celle en MLO*. Par ailleurs, une approche géométrique reposant sur la détection préalable de ces repères anatomiques a été privilégiée afin de favoriser l'explicabilité du modèle, les prédictions étant fondées sur des structures anatomiques clairement identifiables. En outre, la visualisation explicite de ces repères constitue en elle-même une aide à la décision pour les technologues, qui peuvent s'appuyer sur des informations objectives, tout en mobilisant leur expertise, leur expérience clinique et les autres informations disponibles (morphologie des patientes, examens antérieurs) pour formuler un jugement éclairé.

Ainsi, dans le cadre de ce travail, nous avons proposé deux modèles de détection de repères anatomiques : DeepLabv3+ combiné à un algorithme de régression RANSAC pour la détection de la PL, et un U-Net gaussien pour la détection du mamelon par régression de cartes de chaleur. La méthodologie ainsi que les résultats associés ont fait l'objet de l'Article 1, qui sera présenté à la conférence internationale ISBI en avril 2026. Cet article, constituant le cha-

pitre 5 de ce mémoire, rapporte également les résultats de l'évaluation d'un premier critère de positionnement, *Quantité adéquate de muscle pectoral*. Tel que souligné dans cet article, nos modèles de détection de repères offrent des performances comparables voire suppérieures à celles obtenues par le modèle de référence CoordAttUNet. Entraîné sur le sous-ensemble de notre jeu de données composé des images en vue MLO, CoordAttUNet, qui permet une détection simultanée de la PL et du mamelon, atteint une erreur moyenne de 3,14 mm et 1,67 mm pour la détection des points inférieur et supérieur de la PL. Notre approche génère quant à elle des prédictions à 3,01 mm et 1,49 mm des vérités terrain, en moyenne. Des tests statistiques ont permis de démontrer une différence significative entre les erreurs obtenues par notre modèle et par CoordAttUNet pour le point inférieur de la PL. En ce qui a trait à la détection du mamelon, notre modèle U-Net gaussien surpasse de manière statistiquement significative CoordAttUNet, permettant une détection du mamelon avec une erreur moyenne de 2,40 mm sur le sous-ensemble MLO, alors que CoordAttUNet produit une erreur moyenne de 3,42 mm. Notre modèle de détection du mamelon est par ailleurs applicable en vue CC également, alors que CoordAttUNet ne s'applique qu'aux images MLO.

Concernant l'évaluation du critère *Quantité adéquate de muscle pectoral*, les résultats présentés dans l'Article 1 suggèrent qu'une approche géométrique est mieux adaptée qu'une approche de classification bout-en-bout. En effet, les deux méthodes basées sur la détection de repères anatomiques, soient CoordAttUNet et notre approche, offrent des performances supérieures à celles obtenues par un CNN, dont l'architecture est proposée dans [1]. Notre méthode atteint une précision équilibrée de 0,88 sur l'ensemble de test MLO, avec une spécificité de 0,92 et une sensibilité de 0,84, permettant une détection satisfaisante des cas problématiques. Il convient de rappeler que la fréquence de ces cas d'échecs pour ce critère demeure très faible, à moins de 6 % dans l'ensemble de test.

Dans le chapitre 6, nous avons présenté des résultats supplémentaires concernant la détection des repères anatomiques ainsi que l'évaluation d'autres critères de positionnement. Nous rapportons entre autres les résultats de détection du mamelon sur les images en vue CC. La détection est moins précise pour ces cas que dans la vue MLO, avec une erreur moyenne de détection de 6,42 mm. La médiane de cette erreur n'est toutefois que de 1,45 mm, suggérant la présence de quelques détections très erronnées faisant bondir l'erreur moyenne. L'analyse qualitative des résultats a confirmé qu'effectivement, certaines prédictions de la position du mamelon sont très éloignées de la vérité terrain. Il s'agit généralement de cas particuliers où la position réelle du mamelon déroge à la norme pour des raisons morphologiques, ou bien des cas pour lesquels un artefact est identifié à tort comme le mamelon par notre modèle. Dans un contexte clinique, ces erreurs pourraient facilement être repérées par les technologues, qui adaptent leur évaluation en fonction de la morphologie des patientes et d'autres facteurs.

Nous avons également entraîné notre modèle de détection du mamelon sur un sous-ensemble annoté du jeu de données public VinDr. Ces résultats montrent que notre modèle peut être entraîné sur des données représentant des morphologies différentes tout en offrant des performances surpassant celles obtenues par la référence dans la littérature, CoordAttUNet. Ainsi, notre approche atteint de meilleures performances sur notre jeu de données que le modèle CoordAttUNet, et est plus robuste que celui-ci puisque'elle permet une généralisation à d'autres sources de données.

Le chapitre 6 expose également les résultats d'évaluation du positionnement selon les critères *Vue de la largeur maximale du muscle pectoral*, *PNL perpendiculaire au bord de l'image* et *Longueur de la PNL en CC à moins de 1 cm de celle en MLO*. Pour le premier, les résultats semblent indiquer qu'une approche géométrique ne surpasse pas nécessairement un modèle de classification bout-en-bout. Pour le seuil optimal de 16,41 degrés, le CNN ainsi que notre approche obtiennent une précision équilibrée de 0,83. Le CNN atteint une sensibilité légèrement plus élevée que notre méthode, malgré un déséquilibre de classe (seulement 7,7 % des images de l'ensemble de test représentent un échec du critère). Notre approche géométrique bénédicie cependant d'une interprétabilité supérieure de par son aspect quantifiable et la mise en évidence d'un repère anatomique clé.

Les performances obtenues avec l'approche géométrique peuvent s'expliquer par plusieurs facteurs. Tout d'abord, la définition géométrique du critère, reposant uniquement sur la valeur de l'angle, demeure simplificatrice et ne reflète pas fidèlement la réalité clinique. En pratique, l'évaluation du critère *Vue de la largeur maximale du muscle pectoral* ne dépend pas exclusivement de l'orientation de la PL, mais également de l'aspect visuel du muscle, notamment sa croissance continue du bas vers le haut de l'image. L'évaluation du critère par les technologues impliquent ainsi une appréciation visuelle et parfois subjective. Par ailleurs, le seuil utilisé pour caractériser l'angle du muscle pectoral, quelque soit sa valeur, ne concorde pas nécessairement avec l'annotation binaire du critère. En effet, comme exposé dans l'aticle consacré à l'analyse de la variabilité inter-annotatrice [25], les annotations des technologues ne respectent pas systématiquement un seuil fixe. Des cas pour lesquels la PL présente un angle supérieur à 10 (ou 16,41 degrés, pouvant parfois dépasser 20 degrés) sont néanmoins jugés inadéquats par certaines technologues. De plus, une même technologue indique parfois un succès, parfois un échec pour une même valeur d'angle, démontrant qu'il existe même une variabilité intra-annotatrice. Ainsi, certaines annotations binaires associées au critère indiquent un échec bien que l'angle mesuré dépasse le seuil géométrique défini, alors que dans notre approche, tout angle supérieur à ce seuil est automatiquement classé comme un succès. Ceci explique en partie la capacité modérée du modèle à détecter les cas d'échec. La variabilité dans les annotations est également réflétée dans les résultats obtenus par une

approche géométrique basée sur les annotations de chaque technologue. Cette variabilité impose en quelques sortes une borne supérieure aux performances qu'il est possible d'atteindre par une approche géométrique reposant sur un seuil angulaire fixe et les vérités terrains utilisées.

Ces résultats suggèrent qu'il existe un intérêt à apprendre des représentations profondes directement à partir des images, celles-ci intégrant potentiellement des caractéristiques visuelles complexes qui échappent à une formalisation géométrique explicite. Il demeure toutefois que l'approche géométrique proposée présente un avantage important en termes d'explicabilité, puisqu'elle repose sur la détection et la visualisation explicite du muscle pectoral, ce qui constitue un atout dans un contexte clinique. Dans une perspective de travaux futurs, il serait donc pertinent d'explorer une approche consistant à entraîner un modèle de classification en intégrant la PL détectée comme information d'entrée additionnelle. Une telle stratégie pourrait permettre de concilier performance et interprétabilité.

Dans le cas du critère *PNL perpendiculaire au bord de l'image*, qui caractérise l'orientation du mamelon dans la vue CC, notre approche géométrique permet d'atteindre une précision équilibrée de 0,71. Le CNN entraîné à titre de référence ne semble pas apprendre de représentation pertinente à partir des images, plaçant toutes les images dans la classe d'échec. Bien que notre méthode surpasse ce dernier, la précision équilibrée obtenue demeure faible. Ce résultat peut s'expliquer, d'une part, par l'imprécision relative au seuil décisionnel de 5 degrés utilisé. Celui-ci, choisi aux termes de discussions avec les technologues expertes impliquées dans le projet, ne correpond pas au seuil retrouvé dans la littérature (20 degrés selon [1]). Pour refléter le contexte clinique québécois, nous avons toutefois retenu cette valeur de 5 degrés. Même si les technologues nous ont indiqué que 5 degrés soit un bon ordre de grandeur pour ce seuil, une variabilité inter et intra-annotatrice subsiste. Ainsi, l'annotation binaire apposée à une image pour ce critère ne concorde pas systématiquement au seuil choisi. La faible précision équilibrée obtenue en appliquant l'algorithme géométrique aux lignes tracées par les technologues reflète ce phénomène. D'autre part, certaines prédictions erronées découlent d'erreurs de détection du point de largeur maximale du sein, effectuée à partir d'un algorithme de détection de contours. En présence d'artefacts ou d'un contraste trop faible, les contours sont mal identifiés, générant une erreur au niveau du point de largeur maximale. Il serait pertinent de mener une évaluation de cet algorithme afin de pouvoir l'appliquer à l'évaluation de ce critère. En outre, des erreurs sur la position prédite du mamelon en vue CC peuvent expliquer la mauvaise classification finale d'une image. Enfin, certaines particularités morphologiques peuvent mener à l'échec du critère sans que cela se traduise par un angle trop élevé entre la PNL et la ligne de largeur maximale.

Le dernier critère évalué est *Longueur de la PNL en CC à moins de 1 cm de celle en MLO*. L'algorithme géométrique développé permet d'obtenir une précision équilibrée de 0,84, une spécificité de 0,82 et une sensibilité de 0,85. Ces performances peuvent être considérées satisfaisantes puisqu'elles sont comparables à celles d'au mois une technologue experte (T3). Néanmoins, plusieurs éléments peuvent être à la source d'une prédiction erronnée. Effectivement, l'évaluation géométrique de ce critère repose sur la détection de diverses structures ; une erreur dans l'identification d'une ou plusieurs de ces structures peut fausser la longueur des PNL et ainsi mener à la mauvaise classification d'une image. Ainsi, une erreur de détection du muscle pectoral, même minime, peut faire varier la longueur de la PNL en MLO de quelques millimètres, ce qui peut être suffisant pour mener à un faux négatif ou un faux positif. En outre, une erreur sur la détection du mamelon ou dans l'identification du contour peut également mener à des prédictions erronées. Il serait pertinent, dans la suite du projet, d'implémenter un modèle permettant de détecter la présence d'artefact, afin de filtrer automatiquement les images avant l'évaluation de ce critère.

Concernant l'intégration clinique d'un outil d'évaluation automatique du positionnement basé sur les modèles et algorithmes proposés dans ce mémoire, l'utilisation de cet outil en temps réel lors de l'examen de mammographie est tout à fait envisageable. En effet, le temps de calcul nécessaire à l'inférence pour les modèles de détection de repères anatomiques sont de 3,56 ms (PL) et 3,16 ms (mamelon) par image sur GPU (RTX-4090). En faisant plutôt rouler les modèles sur CPU, ces temps s'élèvent à 72,80 ms et 22,33 ms. L'application subséquente des algorithmes géométriques d'évaluation prend en moyenne moins de 0,5 s par image. Ainsi, un temps d'analyse de moins de une seconde par image est possible.

En somme, l'analyse des résultats a permis de mettre en exergue un fait important : bien que régie par une liste de critères définie, l'évaluation du positionnement des seins par les technologues demeure empreinte de subjectivité. En effet, les technologues n'utilisent pas l'information géométrique présente dans l'image au même titre qu'un ordinateur. Cette subjectivité se reflète dans les annotations du jeu de données, induisant une variabilité entre les annotatrices, et même une variabilité dans les annotations d'une même technologue. Dans ce travail, nous avons proposé des modèles de détection de repères anatomiques et des algorithmes d'évaluation de quatre critères de positionnement. Les performances obtenues par nos algorithmes géométriques sont bien entendu limitées lorsque comparés à des annotations issues d'une évaluation subjective. Toutefois, au delà de la classification des images selon les différents critères de positionnement, notre approche offre une mesure quantitative de certains aspects clés, laissant ainsi moins de place à la subjectivité dans l'évaluation. Ceci constitue donc un premier pas vers l'automatisation de l'évaluation du positionnement des seins en mammographie.

CHAPITRE 8 CONCLUSION

Ce chapitre conclut ce mémoire, en synthétisant les travaux effectués, soulignant les limites de la solution proposée et suggérant des pistes d'améliorations futures.

8.1 Synthèse des travaux

Dans le cadre de ce travail, plusieurs contributions au domaine de l'évaluation automatique du positionnement des seins en mammographie ont été réalisées, soient :

1. La création du jeu de données **Mammo-Pos**, un ensemble de 1215 examens mammographiques incluant des annotations relatives aux critères de positionnement mammaire et l'identification visuelle de repères anatomiques par trois technologues expertes.
2. L'analyse de la variabilité inter-opérateur dans l'identification de la PL et du mamelon (en plus des analyses de variabilité relatives aux critères de positionnement et les raisons d'échecs présentés dans l'Article 3 de l'annexe B).
3. L'implémentation d'un modèle de détection de la PL, reposant sur une architecture DeepLabv3+, offrant des performances comparables au modèle de référence issu de la littérature, CoordAttUNet.
4. L'implémentation d'un modèle U-Net gaussien pour la détection du mamelon par régression de cartes de chaleurs, surpassant la référence CoordAttUNet en termes de précision et de généralisation à une autre source de données.
5. Le développement de quatre algorithmes d'évaluation géométrique de critères de positionnement, basés sur les prédictions de la PL et du mamelon.

Ces éléments font l'objet de plusieurs articles rédigés au cours de ce projet, dont l'Article 1 présenté au chapitre 5, qui aborde la détection des repères anatomiques et l'évaluation du critère de positionnement *Quantité adéquate de muscle pectoral*. Ce dernier sera présenté lors de la conférence ISBI en avril 2026. L'article *Breast Positioning in Mammography : Consolidated Image Criteria, Their Interdependence and Trends in Automatic Evaluation* [2], rédigé en collaboration avec les autres étudiants participant au projet, définit les critères de positionnement d'un point de vue adapté à la vision par ordinateur et expose l'interdépendance entre ceux-ci. J'ai contribué à la recherche bibliographique, la génération des figures, et la rédaction de cet article. Enfin, l'article *Breast Positioning Assessment in Mammography : An Inter-Rater Reliability Study* [25], également écrit conjointement avec les autres membres du

projet, traite de la variabilité inter-opérateur et du processus d'annotation de notre jeu de données. Ma contribution à cet article comprend l'analyse de la variabilité des annotations des repères anatomiques, soient la PL et la position du mamelon, ainsi que la participation à la recherche bibliographique, l'analyse des résultats et la rédaction.

Globalement, la solution proposée est une approche explicable, basée sur l'identification de repères anatomiques et une évaluation géométrique du positionnement mammaire. La méthode offre une quantification d'éléments clés pris en compte dans l'évaluation du positionnement, favorisant une évaluation cohérente et objective.

8.2 Limitations de la solution proposée

La méthode présentée dans ce mémoire n'est pas sans lacunes. Nous identifions les limitations principales suivantes :

1. **Limites relatives à la base de données et aux annotations** Le jeu de données créé dans le cadre de ce projet souffre d'un grand déséquilibre de classe pour certains critères de positionnement, n'incluant parfois que très peu de cas d'échecs. Par ailleurs, les annotations du jeu de test sont sujettes à une variabilité inter et intra-annotatrice, et l'évaluation des nos algorithmes est basée sur une vérité terrain issue du label majoritaire et non d'un consensus.
2. **Limites inhérentes à l'approche géométrique** Les résultats suggèrent que l'approche géométrique retenue n'est pas adaptée à l'ensemble des critères évalués. En particulier, certains critères de positionnement prennent en compte plusieurs composantes, alors que l'algorithme d'évaluation géométrique ne considère qu'un seul de ces éléments. Par exemple, l'algorithme permettant d'évaluer le critère *Vue de la largeur maximale du muscle pectoral* repose uniquement sur la mesure d'angle entre la PL et le bord de l'image, tandis que la croissance continue du muscle n'est pas prise en compte. Cette simplification méthodologique introduit un biais structurel, limitant la capacité du modèle à refléter fidèlement la complexité de l'évaluation clinique réelle.
3. **Imprécision liée aux seuils de décision** Une autre limite concerne la définition des seuils utilisés pour évaluer les critères. En effet, il est difficile de traduire l'évaluation clinique, parfois subjective et dépendant de la morphologie, en règles décisionnelles strictes. Les seuils considérés ont été choisis afin d'être les plus représentatifs possible de la pratique clinique selon les discussions menées avec les technologues participantes. Toutefois, en réalité, ces seuils ne sont pas explicitement définis en clinique et reposent souvent sur une évaluation subjective et contextualisée plutôt que sur des valeurs fixes.

Ces seuils ne prennent par ailleurs pas en compte les variations morphologiques des patientes. Cette absence de définition objective rend l'évaluation de certains critères particulièrement complexe.

4. **Sensibilité limitée** Enfin, une limite importante de notre méthode réside dans sa sensibilité. Les performances obtenues montrent que l'approche géométrique identifie une proportion limitée des cas positifs (échecs du critère de positionnement), notamment dans le cas des critères *Vue de la largeur maximale du muscle pectoral* et *PNL perpendiculaire au bord de l'image*. Cette sensibilité réduite découle entre autres des deux limites précédentes, soit le caractère simplifié de l'approche géométrique et l'utilisation de seuils fixes.

Ces limites soulignent la nécessité d'explorer des pistes d'amélioration, certaines étant proposées à la section suivante. Malgré les limites identifiées, l'approche proposée conserve un intérêt significatif pour une intégration en contexte clinique. En effet, elle permet de mettre en évidence les repères géométriques clés que sont la ligne du muscle pectoral et le mamelon. Ces éléments visuels constituent un support pour les technologues et peuvent les aider dans leur évaluation du positionnement. Contrairement aux approches de classification bout-en-bout, qui fonctionnent souvent comme des « boîtes noires », notre méthode repose sur des relations géométriques interprétables entre des repères anatomiques identifiables. La prédiction produite par l'algorithme est donc explicable, ce qui favorise sa transparence, son acceptabilité et potentiellement son adoption en pratique clinique.

8.3 Améliorations et travaux futurs

Plusieurs pistes pourraient être explorées dans la suite de ce projet. D'abord, la bonification du jeu de données par l'annotation d'examens supplémentaires serait un premier pas vers l'amélioration des modèles proposés. Idéalement, l'inclusion de plus de cas d'échecs de certains critères permettrait de réduire le déséquilibre de classe présent dans le jeu de données actuel. De plus, un plus grand nombre de données ouvrirait la porte à diverses possibilités comme l'utilisation de modèles transformeurs visuels, qui nécessitent un grand volume d'images. Ce genre de modèle pourrait notamment être appliqué à la tâche de segmentation du muscle pectoral afin d'en améliorer les performances.

En outre, tel qu'énoncé au chapitre précédent, l'entraînement de modèles de classification prenant en entrée les repères anatomiques (PL et mamelon) pourrait être exploré, pour bénéficier à la fois des capacités d'extraction de caractéristiques des modèles d'apprentissage profond et de l'explicabilité fournie par la mise en évidence des repères anatomiques.

Il serait également pertinent d'employer des approches non-supervisées pour l'évaluation des critères de positionnement, dans l'optique de passer outre la subjectivité présente dans les annotations. En effet, le caractère subjectif de l'évaluation du positionnement se traduit par une variabilité inter-annotatrice et intra-annotatrice pouvant limiter la capacité de modèles supervisés ou d'approches géométriques basées sur des seuils décisionnels fixes à fournir une prédiction adéquate. Ainsi, des approches non-supervisées permettraient de s'émanciper de ces annotations.

Par ailleurs, la détection automatique des artefacts et l'évaluation de l'algorithme de détection de contours permettraient l'amélioration des performances de nos algorithmes géométriques.

Enfin, dans le cadre du projet global sur le positionnement mammaire, les algorithmes d'évaluation des quatre critères de positionnement abordés dans ce mémoire devront être combinés à des modèles permettant l'évaluation des autres critères, afin de former un outil complet d'évaluation du positionnement.

RÉFÉRENCES

[1] M. Brahim, K. Westerkamp, L. Hempel, R. Lehmann, D. Hempel et P. Philipp, "Automated assessment of breast positioning quality in screening mammography," *Cancers*, vol. 14, n°. 19, 2022. [En ligne]. Disponible : https://www.mdpi.com/2072-6694/14/19/4704

[2] D. Gurve, V. Gauvin, L. Grondin-Reiher, V. Patenaude, K. Morency, L. Nathaniel, F. Guibault et L. Séoud, "Breast positioning in mammography : consolidated image criteria, their interdependence and trends in automatic evaluation," 2026, soumis pour publication.

[3] A. Kirillov, K. He, R. Girshick, C. Rother et P. Dollár, "Panoptic segmentation," dans *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, p. 9396–9405. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8953237

[4] O. Ronneberger, P. Fischer et T. Brox, "U-net : Convolutional networks for biomedical image segmentation," dans *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells et A. F. Frangi, édit. Cham : Springer International Publishing, 2015, p. 234–241. [En ligne]. Disponible : https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28

[5] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff et H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," dans *Computer Vision – ECCV 2018*. Cham : Springer, 2018, p. 833–851. [En ligne]. Disponible : https://link.springer.com/chapter/10.1007/978-3-030-01234-2_49

[6] S. Du, Z. Zhang et T. Ikenaga, "Anatpose : Bidirectionally learning anatomy-aware heatmaps for human pose estimation," *Pattern Recognition*, vol. 155, p. 110654, 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0031320324004059

[7] L. Sharan, G. Romano, J. Brand, H. Kelm, M. Karck, R. De Simone et S. Engelhardt, "Point detection through multi-instance deep heatmap regression for sutures in endoscopy," *Int J Comput Assist Radiol Surg*, vol. 16, n°. 12, p. 2107–2117, 2021. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC8616891/

[8] World Health Organization (WHO), "Breast cancer," 2025, accessed : 2025-10-09. [En ligne]. Disponible : https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[9] J. Ferlay, M. Ervik, F. Lam, M. Laversanne, M. Colombet, L. Mery, M. Piñeros,

A. Znaor, I. Soerjomataram et F. Bray, “Global cancer observatory : Cancer today,” International Agency for Research on Cancer (IARC), Lyon, France, 2024, consulté le 25 février 2024. [En ligne]. Disponible : https://gco.iarc.who.int/today

[10] A. Amin, D. A. U, P. Koteshwara, S. P C et S. Mathew, “A systematic literature review on mammography : deep learning techniques for breast cancer detection with global and Asian perspectives,” *BMC Cancer*, vol. 25, p. 1627, oct. 2025. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC12542462/

[11] Public Health Agency of Canada, “Breast cancer in canada,” 2023, consulté le 20 janvier 2026. [En ligne]. Disponible : https://www.canada.ca/en/public-health/services/publications/diseases-conditions/breast-cancer.html

[12] J. Rouette, N. Elfassy, N. Bouganim, H. Yin, N. Lasry et L. Azoulay, “Evaluation of the quality of mammographic breast positioning : a quality improvement study,” *CMAJ Open*, vol. 9, nº. 2, p. E607–E612, avr. 2021. [En ligne]. Disponible : http://cmajopen.ca/lookup/doi/10.9778/cmajo.20200211

[13] Société canadienne du cancer. (2025) Les seins. Société canadienne du cancer. [En ligne]. Disponible : https://cancer.ca/fr/cancer-information/cancer-types/breast/what-is-breast-cancer/the-breasts

[14] N. Thakur, P. Kumar et A. Kumar, “A systematic review of machine and deep learning techniques for the identification and classification of breast cancer through medical image modalities,” *Multimedia Tools and Applications*, vol. 83, p. 35 849–35 942, 2024. [En ligne]. Disponible : https://doi.org/10.1007/s11042-023-16634-w

[15] L. Nicosia, G. Gnocchi, I. Gorini, M. Venturini, F. Fontana, F. Pesapane, I. Abiuso, A. C. Bozzini, M. Pizzamiglio, A. Latronico, F. Abbate, L. Meneghetti, O. Battaglia, G. Pellegrino et E. Cassano, “History of mammography : Analysis of breast imaging diagnostic achievements over the last century,” *Healthcare*, vol. 11, nº. 11, 2023. [En ligne]. Disponible : https://www.mdpi.com/2227-9032/11/11/1596

[16] S. W. Duffy, L. Tabár, H.-H. Chen, M. Holmqvist, M.-F. Yen, S. Abdsalah, B. Epstein, E. Frodis, E. Ljungberg, C. Hedborg-Melander, A. Sundbom, M. Tholin, M. Wiege, A. Åkerlund, H.-M. Wu, T.-S. Tung, Y.-H. Chiu, C.-P. Chiu, C.-C. Huang, R. A. Smith, M. Rosén, M. Stenbeck et L. Holmberg, “The impact of organized mammography service screening on breast carcinoma mortality in seven swedish counties,” *Cancer*, vol. 95, nº. 3, p. 458–469, 2002. [En ligne]. Disponible : https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.1002/cncr.10765

[17] C. Ju et L. Doepke, “Positioning and Technique - Radiology | UCLA Health,” 2026. [En ligne]. Disponible : https://www.uclahealth.org/departments/radiology/education/breast-imaging-teaching-resources/screening-mammogram/positioning-and-technique

[18] S. Lille et W. Marshall, *Mammographic Positioning*, 4e éd. Lippincott Williams & Wilkins, 2018, p. 100–159.

[19] S. Ding, T. Fontaine, M. Serex et C. Sá dos Reis, "Strategies enhancing the patient experience in mammography : A scoping review," *Radiography*, vol. 30, no. 1, p. 340–352, 2024. [En ligne]. Disponible : https://doi.org/10.1016/j.radi.2023.11.016

[20] U. Bick et F. Diekmann, "Digital mammography : what do we and what don't we know ?" *European Radiology*, vol. 17, no. 8, p. 1931–1942, 2007. [En ligne]. Disponible : https://doi.org/10.1007/s00330-007-0586-1

[21] N. Perry, M. Broeders, C. de Wolf, S. Törnberg, R. Holland et L. von Karsa, édit., *European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis*, 4e éd. Luxembourg : Office for Official Publications of the European Communities, 2006.

[22] GOV.UK. (2025) Guidance for breast screening mammographers. [En ligne]. Disponible : https://www.gov.uk/government/publications/breast-screening-quality-assurance-for-mammography-and-radiography/guidance-for-breast-screening-mammographers

[23] American College of Radiology, *Mammography quality control manual*, 1999. [En ligne]. Disponible : https://fr.scribd.com/document/639224239/Untitled

[24] C. Hamel, B. Avard, C. Flegg, V. Freitas, C. Hapgood, S. Kulkarni, P. Lenkov et M. Seidler, "Canadian association of radiologists breast disease imaging referral guideline," *Canadian Association of Radiologists Journal*, vol. 75, no. 2, p. 287–295, 2024, pMID : 37724018. [En ligne]. Disponible : https://doi.org/10.1177/08465371231192391

[25] V. Gauvin, V. Patenaude, L. Grondin-Reiher, F. Guibault et L. Séoud, "Breast positioning assessment in mammography : An inter-rater reliability study," 2026, À soumettre pour publication.

[26] N. Sharma et L. M. Aggarwal, "Automated medical image segmentation techniques," *Journal of Medical Physics*, vol. 35, no. 1, p. 3–14, 2010. [En ligne]. Disponible : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2825001/

[27] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz et D. Terzopoulos, "Image segmentation using deep learning : A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, p. 3523–3542, 2021. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9356353

[28] Y. Xu, R. Quan, W. Xu, Y. Huang, X. Chen et F. Liu, "Advances in medical image segmentation : A comprehensive review of traditional, deep learning and hybrid

approaches," *Bioengineering*, vol. 11, nº. 10, p. 1034, 2024. [En ligne]. Disponible : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11505408/

[29] J. Shi et J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, nº. 8, p. 888–905, 2000. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/868688

[30] M. Jena, S. Mishra et D. Mishra, "A survey on applications of machine learning techniques for medical image segmentation," *International Journal of Engineering and Technology*, vol. 7, p. 4489–4495, 01 2018. [En ligne]. Disponible : https://www.researchgate.net/publication/332220946_A_survey_on_applications_of_machine_learning_techniques_for_medical_image_segmentation

[31] A. Bou, "Deep learning models for semantic segmentation of mammography screenings," Master's thesis, KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, 2019, submitted September 25, 2019. [En ligne]. Disponible : https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-265652

[32] E. Shelhamer, J. Long et T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, nº. 4, p. 640–651, avr. 2017. [En ligne]. Disponible : https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7298965

[33] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh et J. Liang, "Unet++ : A nested u-net architecture for medical image segmentation," dans *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLMIA/ML-CDS 2018) – Lecture Notes in Computer Science, Vol. 11045*, sep 2018, p. 3–11, epub 20 Sep 2018. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC7329239/

[34] O. Oktay, J. Schlemper, L. L. Folgoc, M. J. Lee, M. P. Heinrich, K. Misawa, K. Mori, S. G. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker et D. Rueckert, "Attention u-net : Learning where to look for the pancreas," *ArXiv*, vol. abs/1804.03999, 2018. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:4861068

[35] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy et A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," dans *International Conference on Learning Representations (ICLR)*, 2015, poster. [En ligne]. Disponible : https://arxiv.org/abs/1412.7062

[36] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy et A. Yuille, "Deeplab : Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, nº. 4, p. 834–848, 2018. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/7913730

[37] L.-C. Chen, G. Papandreou, F. Schroff et H. Adam, “Rethinking atrous convolution for semantic image segmentation,” *arXiv*, 2017. [En ligne]. Disponible : https://arxiv.org/abs/1706.05587

[38] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez et P. Luo, “Segformer : Simple and efficient design for semantic segmentation with transformers,” dans *Advances in Neural Information Processing Systems (NeurIPS 34)*, 2021, poster / Conference Paper. [En ligne]. Disponible : https://dl.acm.org/doi/10.5555/3540261.3541185

[39] R. F. Khan, B. D. Lee et M. S. Lee, “Transformers in medical image segmentation : a narrative review,” *Quantitative Imaging in Medicine and Surgery*, vol. 13, nº. 12, p. 8747–8767, déc. 2023. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/38106306/

[40] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollar et R. Girshick, “Segment anything,” dans *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2023, p. 4015–4026. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:257952310

[41] J. Ma, J. Chen, M. Ng, R. Huang, Y. Li, C. Li, X. Yang et A. L. Martel, “Loss odyssey in medical image segmentation,” *Medical Image Analysis*, vol. 71, p. 102035, 2021. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1361841521000815

[42] N. Saidin, H. A. M. Sakim, U. K. Ngah et I. L. Shuaib, “Segmentation of breast regions in mammogram based on density : A review,” *ArXiv*, vol. abs/1209.5494, 2012. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:16816436

[43] M. E. Ma, H. Ma, H. Li, F. Kulwa et J. Li, “Breast cancer segmentation methods : Current status and future potentials,” *BioMed Research International*, vol. 2021, p. 9962109, 2021. [En ligne]. Disponible : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8321730/

[44] S. Lathuilière, P. Mesejo, X. Alameda-Pineda et R. Horaud, “A comprehensive analysis of deep regression,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, nº. 9, p. 2065–2081, 2020. [En ligne]. Disponible : https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8686063

[45] K. A. Manjusha et B. Naresh, “A study on mammogram image based breast cancer classification and detection using machine learning model,” dans *2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA)*, 2024, p. 1–9. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/11051984

[46] M. A. Rahman, M. S. H. Khan, Y. Watanobe, J. T. Prioty, T. T. Annita, S. Rahman, M. S. Hossain, S. A. Aitijjo, R. I. Taskin, V. Dhrubo, A. Hanip et T. Bhuiyan, "Advancements in breast cancer detection : A review of global trends, risk factors, imaging modalities, machine learning, and deep learning approaches," *BioMedInformatics*, vol. 5, nº. 3, 2025. [En ligne]. Disponible : https://www.mdpi.com/2673-7426/5/3/46

[47] H. T. Nguyen, H. Q. Nguyen, H. H. Pham *et al.*, "Vindr-mammo : A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography," *Scientific Data*, vol. 10, p. 277, 2023. [En ligne]. Disponible : https://doi.org/10.1038/s41597-023-02100-7

[48] N. M. Hassan, S. Hamad et K. Mahar, "Mammogram breast cancer cad systems for mass detection and classification : a review," *Multimedia Tools and Applications*, vol. 81, p. 20 043–20 075, 2022. [En ligne]. Disponible : https://doi.org/10.1007/s11042-022-12332-1

[49] A. Carriero, L. Groenhoff, E. Vologina, P. Basile et M. Albera, "Deep learning in breast cancer imaging : State of the art and recent advancements in early 2024," *Diagnostics*, vol. 14, nº. 8, p. 848, 2024. [En ligne]. Disponible : https://doi.org/10.3390/diagnostics14080848

[50] J. P. Suckling, "The mammographic image analysis society digital mammogram database," dans *Exerpta Medica International Congress Series*, vol. 1069. Amsterdam, The Netherlands : Elsevier Science, 1994, p. 375–378.

[51] M. Heath, K. Bowyer, D. Kopans, R. Moore et P. Kegelmeyer, "The digital database for screening mammography," dans *Proceedings of the Fifth International Workshop on Digital Mammography*, M. J. Yaffe, édit. Madison, WI, USA : Medical Physics Publishing, 2000, p. 212–218.

[52] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso et J. S. Cardoso, "Inbreast : Toward a full-field digital mammographic database," *Academic Radiology*, vol. 19, nº. 2, p. 236–248, 2012.

[53] M. A. G. Lopez, N. G. de Posada, D. C. Moura, R. R. Pollán, J. M. F. Valiente, C. S. Ortega, M. R. del Solar, G. D. Herrero, I. M. P. Ramos, J. Loureiro, T. C. Fernandes et B. M. F. de Araújo, "Bcdr : A breast cancer digital repository," 2012. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:16423079

[54] R. Sawyer-Lee, F. Gimenez, A. Hoogi et D. Rubin, "Curated breast imaging subset of digital database for screening mammography (cbis-ddsm) (version 1) [data set]," 2016. [En ligne]. Disponible : https://doi.org/10.7937/K9/TCIA.2016.7O02S9CY

[55] F. Strand, “Csaw-cc (mammography) – a dataset for ai research to improve screening, diagnostics and prognostics of breast cancer,” Karolinska Institutet Dataset, 2022. [En ligne]. Disponible : https://doi.org/10.5878/45vm-t798

[56] M. D. Halling-Brown, L. M. Warren, D. Ward, E. Lewis, A. Mackenzie, M. G. Wallis, L. S. Wilkinson, R. M. Given-Wilson, R. McAvinchey et K. C. Young, “Optimam mammography image database : A large-scale resource of mammography images and clinical data,” *Radiology : Artificial Intelligence*, vol. 3, n°. 1, p. e200103, 2021. [En ligne]. Disponible : https://doi.org/10.1148/ryai.2020200103

[57] C. Cui, L. Li, H. Cai, Z. Fan, L. Zhang, T. Dan, J. Li et J. Wang, “The chinese mammography database (cmmd) : An online mammography database with biopsy confirmed types for machine diagnosis of breast [data set],” 2021. [En ligne]. Disponible : https://doi.org/10.7937/tcia.eqde-4b16

[58] H. M. Frazer, E. Wheeler, C. Neil, D. Kontos *et al.*, “Admani : Annotated digital mammograms and associated non-image datasets,” *Radiology : Artificial Intelligence*, vol. 5, p. e220072, 2023. [En ligne]. Disponible : https://doi.org/10.1148/ryai.220072

[59] H. M. Trivedi, M. Vazirabad, F. C. Kitamura, Y. Chen et H. Frazer, “Open-source dataset for the rsna screening mammography cancer detection challenge,” *Radiology : Artificial Intelligence*, vol. 8, n°. 2, p. e250375, 2026, pMID : 41441444. [En ligne]. Disponible : https://doi.org/10.1148/ryai.250375

[60] E. Kendall, P. Hajishafiezahramini, M. Hamilton, G. Doyle, N. Wadden et O. Meruvia-Pastor, “Full field digital mammography dataset from a population screening program,” *Scientific Data*, vol. 12, n°. 1, p. 1479, 2025. [En ligne]. Disponible : https://doi.org/10.1038/s41597-025-05866-0

[61] I. Sechopoulos, J. Teuwen et R. Mann, “Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis : State of the art,” *Seminars in Cancer Biology*, vol. 72, p. 214–225, 2021, precision Medicine in Breast Cancer. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1044579X20301358

[62] L. Wang, “Mammography with deep learning for breast cancer detection,” *Frontiers in Oncology*, vol. 14, p. 1281922, 2024. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC10894909/

[63] M. Heenaye-Mamode Khan, N. Boodoo-Jahangeer, W. Dullull, S. Nathire, X. Gao, G. R. Sinha et K. K. Nagwanshi, “Multi- class classification of breast cancer abnormalities using deep convolutional neural network (cnn),” *PLOS ONE*, vol. 16, n°. 8, p. 1–15, 08 2021. [En ligne]. Disponible : https://doi.org/10.1371/journal.pone.0256500

[64] A. Yala, C. D. Lehman, T. Schuster, T. Portnoi et R. Barzilay, “A deep learning mammography-based model for improved breast cancer risk prediction,” *Radiology*, vol. 292, n°. 1, p. 60–66, 2019. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/31063083/

[65] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride et W. Sieh, “Deep learning to improve breast cancer detection on screening mammography,” *Scientific Reports*, vol. 9, n°. 1, p. 12495, 2019. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/31467326/

[66] V. Mudeng, J. woo Jeong et S. woon Choe, “Simply fine-tuned deep learning-based classification for breast cancer with mammograms,” *Computers, Materials and Continua*, vol. 73, n°. 3, p. 4677–4693, 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1546221822001023

[67] F. Walayat, A. Ditta, Z. I. Karmani, B. Shoaib, T. Mazhar, M. A. Khan et R. A. Saeed, “Deep learning–based analysis of mammographic images for breast cancer detection using transfer learning,” *Healthcare Technology Letters*, vol. 12, n°. 1, p. e70019, 2025. [En ligne]. Disponible : https://doi.org/10.1049/htl2.70019

[68] V. Nemade, S. Pathak et A. K. Dubey, “Deep learning-based ensemble model for classification of breast cancer,” *Microsystem Technologies*, vol. 30, n°. 5, p. 513–527, 2024. [En ligne]. Disponible : https://link.springer.com/article/10.1007/s00542-023-05469-y

[69] M. Bobowicz, M. Rygusik, J. Buler, R. Buler, M. Ferlin, A. Kwasigroch, E. Szurowska et M. Grochowski, “Attention-based deep learning system for classification of breast lesions—multimodal, weakly supervised approach,” *Cancers*, vol. 15, n°. 10, p. 2704, 2023. [En ligne]. Disponible : https://www.mdpi.com/2072-6694/15/10/2704

[70] F. Manigrasso, R. Milazzo, A. S. Russo, F. Lamberti, F. Strand, A. Pagnani et L. Morra, “Mammography classification with multi-view deep learning techniques : Investigating graph and transformer-based architectures,” *Medical Image Analysis*, vol. 99, p. 103320, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1361841524002457

[71] J.-N. Eckardt, M. Bornhäuser, K. Wendt et J. M. Middeke, “Semi-supervised learning in cancer diagnostics,” *Frontiers in Oncology*, vol. 12, p. 960984, 2022. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC9329803/

[72] S. Lee, C. Farley, S. Shim, W.-S. Yoo, Y. Zhao et W. Choi, “Unsupervised learning of deep-learned features from breast cancer images,” dans *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2020, p. 740–745. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9288023

[73] M. Rogozinski, J. Hurtado, C. A. Sierra-Franco, A. Ochoa, D. Cabrera, C. Valderrama, J. Betancur et E. Romero, "A novel deep learning framework for nipple segmentation in digital mammography," *Journal of Imaging Informatics in Medicine*, 2025. [En ligne]. Disponible : https://link.springer.com/article/10.1007/s10278-025-01567-7

[74] F. Angelone, C. D. Luca et A. Ragozini, "U-net based approach for pectoralis muscle segmentation," *Computer Methods and Programs in Biomedicine*, 2025. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:280468010

[75] C. Wang, "A novel and automatic pectoral muscle identification algorithm for mediolateral oblique (mlo) view mammograms using imagej," *ArXiv*, vol. abs/1603.01056, 2016. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:9116385

[76] S. I. Naz, M. Shah et M. I. H. Bhuiyan, "Automatic segmentation of pectoral muscle in mammogram images using global thresholding and weak boundary approximation," dans *2017 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2017, p. 199–202. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8468895

[77] J. C. Anusionwu, V. C. Chijindu, J. N. Eneh, T. I. Ozue, N. Ezukwoke, M. A. Ahaneku, E. C. Anoliefo et W. A. Ohagwu, "Pectoral muscle removal in digital mammograms using region based standard otsu technique," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 11, n°. 2, p. 337–348, 2023. [En ligne]. Disponible : https://section.iaesonline.com/index.php/IJEEI/article/download/4043/826

[78] C. Liu, C. Tsai, J. Liu, C. Yu et S. Yu, "A pectoral muscle segmentation algorithm for digital mammograms using otsu thresholding and multiple regression analysis," *Computers & Mathematics with Applications*, vol. 64, n°. 5, p. 1100–1107, 2012. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0898122112002337

[79] S. Sreedevi et E. Sherly, "A novel approach for removal of pectoral muscles in digital mammogram," *Procedia Computer Science*, vol. 46, p. 1724–1731, 2015. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1877050915001817

[80] P. Vikhe et V. Thool, "Intensity based automatic boundary identification of pectoral muscle in mammograms," *Procedia Computer Science*, vol. 79, p. 262–269, 2016. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1877050916001654

[81] K. Czaplicka et H. Wlodarczyk, "Automatic breastline and pectoral muscle segmentation," *Schedae Informaticae*, vol. 20, p. 195–209,

2012. [En ligne]. Disponible : https://ruj.uj.edu.pl/server/api/core/bitstreams/57c88898-ac69-433f-acbc-6717bcb348c7/content

[82] W. Yoon, J. Oh, E. Chae, H. Kim, S. Lee et K. Kim, "Automatic detection of pectoral muscle region for computer-aided diagnosis," *BioMed Research International*, p. 1–6, 2016. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/27847817/

[83] K. S. Camilus, V. K. Govindan et P. Sathidevi, "Pectoral muscle identification in mammograms," *Journal of Applied Clinical Medical Physics*, vol. 12, n°. 3, p. 215–230, 2011. [En ligne]. Disponible : https://aapm.onlinelibrary.wiley.com/doi/abs/10.1120/jacmp.v12i3.3285

[84] Z. Chen et R. Zwiggelaar, "A combined method for automatic identification of the breast boundary in mammograms," dans *2012 5th International Conference on BioMedical Engineering and Informatics*, 2012, p. 121–125. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6513144

[85] J. A. Ayala-Godoy, R. E. Lillo et J. Romo, "Automatic elimination of the pectoral muscle in mammograms based on anatomical features," 2020. [En ligne]. Disponible : https://arxiv.org/abs/2009.06357

[86] Y. Li, H. Chen, Y. Yang et N. Yang, "Pectoral muscle segmentation in mammograms based on homogenous texture and intensity deviation," *Pattern Recognition*, vol. 46, n°. 3, p. 681–691, 2013. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S003132031200427X

[87] C. Zhou, J. Wei, H. P. Chan, C. Paramagul, L. M. Hadjiiski, B. Sahiner et J. A. Douglas, "Computerized image analysis : Texture-field orientation method for pectoral muscle identification on mlo-view mammograms," *Medical Physics*, vol. 37, n°. 5, p. 2289–2299, May 2010. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/20527563/

[88] J. Chakraborty, S. Mukhopadhyay, V. Singla, N. Khandelwal et P. Bhattacharyya, "Automatic detection of pectoral muscle using average gradient and shape-based feature," *J. Digit. Imaging*, vol. 25, n°. 3, p. 387–399, 2012. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/22006275/

[89] C. Chen, G. Liu, J. Wang et G. Sudlow, "Shape-based automatic detection of pectoral muscle boundary in mammograms," *Journal of Medical and Biological Engineering*, vol. 35, n°. 3, p. 315–322, 2015. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC4491117/

[90] Q. Tran, T. Santner et A. Rodríguez-Sánchez, "Automatic computation of the posterior nipple line from mammographies," dans *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and*

Applications - Volume 4 : VISAPP, INSTICC. SciTePress, 2024, p. 629–636. [En ligne]. Disponible : https://www.scitepress.org/Papers/2024/125705/125705.pdf

[91] H. Ture et T. Kayikcioglu, "Accurate detection of distorted pectoral muscle in mammograms using specific patterned isocontours," *IEEE Access*, vol. 8, p. 147 370–147 386, 2020. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9163115

[92] S. Chen, D. L. Bennett, G. A. Colditz et S. Jiang, "Pectoral muscle removal in mammogram images : A novel approach for improved accuracy and efficiency," *Cancer Causes & Control*, vol. 35, p. 185–191, 2024. [En ligne]. Disponible : https://doi.org/10.1007/s10552-023-01781-0

[93] Y.-C. Zeng, "Automatic detections of nipple and pectoralis major in mammographic images," dans *2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, 2018, p. 1–2. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8448979

[94] S. K. Kinoshita, P. M. Azevedo-Marques, R. R. Pereira, J. A. Rodrigues et R. M. Rangayyan, "Radon-domain detection of the nipple and the pectoral muscle in mammograms," *Journal of Digital Imaging*, vol. 21, n°. 1, p. 37–49, 2008. [En ligne]. Disponible : https://link.springer.com/article/10.1007/s10278-007-9035-6

[95] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel et M. T. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, n°. 9, p. 1129–1140, 2004. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/1327692

[96] R. Ferrari, R. Rangayyan, J. Desautels, R. Borges et A. Frere, "Automatic identification of the pectoral muscle in mammograms," *IEEE Transactions on Medical Imaging*, vol. 23, n°. 2, p. 232–245, 2004. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/1263612

[97] C.-H. Wei, C.-Y. Gwo et P.-J. Huang, "Identification and segmentation of obscure pectoral muscle in mediolateral oblique mammograms," *British Journal of Radiology*, vol. 89, n°. 1062, p. 20150802, Jun 2016. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC5258151/

[98] N. P. Firdi, T. A. Sardjono et N. F. Hikmah, "Using pectoral muscle removers in mammographic image process to improve accuracy in breast cancer," *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, vol. 55, p. 131–142, 2022. [En ligne]. Disponible : https://doi.org/10.4028/p-35cy9o

[99] B. Mughal, N. Muhammad, M. Sharif, A. Rehman et T. Saba, "Removal of pectoral muscle based on topographic map and shape-shifting silhouette,"

BMC Cancer, vol. 18, n°. 1, p. 778, Aug 2018. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC6090971/

[100] R. Shen, K. Yan, F. Xiao, J. Chang, C. Jiang et K. Zhou, "Automatic pectoral muscle region segmentation in mammograms using genetic algorithm and morphological selection," *Journal of Digital Imaging*, vol. 31, n°. 5, p. 680–691, Oct 2018. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC6148808/

[101] L. Wang, M. Zhu, L. Deng et X. Yuan, "Automatic pectoral muscle boundary detection in mammograms based on markov chain and active contour model," *Journal of Zhejiang University SCIENCE*, vol. 11, n°. 2, p. 111–118, 2010. [En ligne]. Disponible : https://doi.org/10.1631/jzus.C0910025

[102] F. Ma, M. Bajger, J. Slavotinek et M. Bottema, "Two graph theory based methods for identifying the pectoral muscle in mammograms," *Pattern Recognition*, vol. 40, n°. 9, p. 2592–2602, 2007. [En ligne]. Disponible : https://doi.org/10.1016/j.patcog.2006.12.011

[103] L. Liu, Q. Liu et W. Lu, "Pectoral muscle detection in mammograms using local statistical features," *Journal of Digital Imaging*, vol. 27, n°. 5, p. 1633–1641, 2014. [En ligne]. Disponible : https://doi.org/10.1007/s10278-014-9676-1

[104] C. K. Feudjio, J. Klein, A. Tiedeu et O. Colot, "Automatic extraction of pectoral muscle in the mlo view of mammograms," *Physics in Medicine & Biology*, vol. 58, n°. 23, p. 8493, nov 2013. [En ligne]. Disponible : https://doi.org/10.1088/0031-9155/58/23/8493

[105] S. D. Pawar, K. K. Sharma, S. G. Sapate et G. Y. Yadav, "Segmentation of pectoral muscle from digital mammograms with depth-first search algorithm towards breast density classification," *Biocybernetics and Biomedical Engineering*, vol. 41, n°. 3, p. 1224–1241, 2021. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S020852162100098X

[106] S. Priyanka, S. Sivakumar et P. Selvam, "Optimizing breast cancer detection : Machine learning for pectoral muscle segmentation in mammograms," dans *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)*, 2024, p. 1–6. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/10498969

[107] A. Dubrovina, P. Kisilev, B. Ginsburg, S. Hashoul et R. Kimmel, "Computational mammography using deep neural networks," *Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization*, vol. 6, n°. 3, p. 243–247, 2018. [En ligne]. Disponible : https://doi.org/10.1080/21681163.2015.1131197

[108] Y. Guo, W. Zhao, S. Li, Y. Zhang et Y. Lu, "Automatic segmentation of the pectoral muscle based on boundary identification and shape prediction," *Physics in Medicine & Biology*, vol. 65, n°. 4, p. 045016, 2020. [En ligne]. Disponible : https://iopscience.iop.org/article/10.1088/1361-6560/ab652b

[109] H. Soleimani et O. V. Michailovich, “On segmentation of pectoral muscle in digital mammograms by means of deep learning,” *IEEE Access*, vol. 8, p. 204 173–204 182, 2020. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9252130

[110] A. Rampun, K. López-Linares, P. J. Morrow, B. W. Scotney, H. Wang, I. G. Ocaña, G. Maclair, R. Zwiggelaar, M. A. González Ballester et I. Macía, “Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network,” *Medical Image Analysis*, vol. 57, p. 1–17, 2019. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1361841518301129

[111] Z. Klanecek, T. Wagner, Y. K. Wang, L. Cockmartin, N. Marshall, B. Schott, A. Deatsch, A. Studen, K. Hertl, K. Jarm, M. Krajc, M. Vrhovec, H. Bosmans et R. Jeraj, “Uncertainty estimation for deep learning-based pectoral muscle segmentation via monte carlo dropout,” *Physics in Medicine & Biology*, vol. 68, n°. 11, May 2023. [En ligne]. Disponible : https://iopscience.iop.org/article/10.1088/1361-6560/acd221

[112] C. Huemmer, R. Biniazan, M. Datar, M. Kraus, A. Fieselmann et S. Kappler, “Improved pectoral muscle segmentation in mammograms through regression-based deep learning and knowledge distillation,” dans *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024, p. 1–5. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/10635448

[113] S. D. Verboom, M. Caballo, J. Peters, J. Gommers, D. van den Oever, M. J. M. Broeders, J. Teuwen et I. Sechopoulos, “Deep learning-based breast region segmentation in raw and processed digital mammograms : generalization across views and vendors,” *Journal of Medical Imaging*, vol. 11, n°. 1, p. 014001, Jan 2024. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC10753125/

[114] V. Tiryaki et V. Kaplanoğlu, “Deep learning-based multi-label tissue segmentation and density assessment from mammograms,” *IRBM*, vol. 43, n°. 6, p. 538–548, 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1959031822000562

[115] K. Wang, N. Khan, A. Chan, J. Dunne et R. Highnam, “Deep learning for breast region and pectoral muscle segmentation in digital mammography,” dans *Image and Video Technology*, C. Lee, Z. Su et A. Sugimoto, édit. Cham : Springer International Publishing, 2019, p. 78–91. [En ligne]. Disponible : https://doi.org/10.1007/978-3-030-34879-3_7

[116] C. A. Sierra-Franco, J. Hurtado, V. de A. Thomaz, L. C. da Cruz, S. V. Silva, G. F. M. Silva-Calpa et A. Raposo, “Towards automated semantic segmentation in mammography images for enhanced clinical applications,” *Journal*

of Digital Imaging, vol. 38, nº. 4, p. 2260–2280, 2025. [En ligne]. Disponible : https://doi.org/10.1007/s10278-024-01364-8

[117] S. V. Silva, C. A. Sierra-Franco, J. Hurtado, L. C. da Cruz, V. d. A. Thomaz, G. F. M. Silva-Calpa et A. B. Raposo, "A data-centric approach for pectoral muscle deep learning segmentation enhancements in mammography images," dans *Advances in Visual Computing*, G. Bebis, G. Ghiasi, Y. Fang, A. Sharf, Y. Dong, C. Weaver, Z. Leo, J. J. LaViola Jr. et L. Kohli, édit. Cham : Springer Nature Switzerland, 2023, p. 56–67. [En ligne]. Disponible : https://doi.org/10.1007/978-3-031-47969-4_5

[118] X. Ma, J. Wei, C. Zhou, M. A. Helvie, H.-P. Chan, L. M. Hadjiiski et Y. Lu, "Automated pectoral muscle identification on mlo-view mammograms : Comparison of deep neural network to conventional computer vision," *Medical Physics*, vol. 46, nº. 5, p. 2103–2114, 2019. [En ligne]. Disponible : https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13451

[119] W. Liu, C. Liu et Y. Wei, "Utilizing deep learning technology to segment pectoral muscle in mediolateral oblique view mammograms," dans *2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)*, 2020, p. 97–101. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9339411

[120] R. Doğan, H. Türe et T. Kayikçioğlu, "Segmentation of pectoral muscle region in mlo mammography images by backboned u-net," dans *2022 30th Signal Processing and Communications Applications Conference (SIU)*, 2022, p. 1–4. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9864865

[121] P. Aliniya, M. Nicolescu, M. Nicolescu et G. Bebis, "Towards robust supervised pectoral muscle segmentation in mammography images," *Journal of Imaging*, vol. 10, nº. 12, 2024. [En ligne]. Disponible : https://www.mdpi.com/2313-433X/10/12/331

[122] S. D. Deb et R. K. Jha, "Segmentation of pectoral muscle from mammograms using u-net having densely connected convolutional layers," *Multimedia Tools and Applications*, vol. 83, p. 24 505–24 526, 2024. [En ligne]. Disponible : https://doi.org/10.1007/s11042-023-16394-7

[123] M. J. Ali, B. Raza, A. R. Shahid, F. Mahmood, M. A. Yousuf, A. H. Dar et U. Iqbal, "Enhancing breast pectoral muscle segmentation performance by using skip connections in fully convolutional network," *International Journal of Imaging Systems and Technology*, vol. 30, nº. 4, p. 1108–1118, 2020. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.22410

[124] F. D. Cortes-Rojas, Y. M. Hernández-Rodríguez, R. Bayareh-Mancilla et O. E. Cigarroa-Mayorga, "An artificial intelligence-based tool for enhancing pectoral muscle segmentation in mammograms : Addressing class imbalance and validation challenges

in automated breast cancer diagnosis," *Diagnostics*, vol. 14, n°. 19, 2024. [En ligne]. Disponible : https://www.mdpi.com/2075-4418/14/19/2144

[125] K. Zhou, W. Li et D. Zhao, "Deep learning-based breast region extraction of mammographic images combining pre-processing methods and semantic segmentation supported by deeplab v3," *Technology and Health Care*, vol. 30, n°. S1, p. 173–190, 2022. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC9028646/

[126] X. Yu, S.-H. Wang, J. M. Górriz, X.-W. Jiang, D. S. Guttery et Y.-D. Zhang, "Pemnet for pectoral muscle segmentation," *Biology*, vol. 11, n°. 1, 2022. [En ligne]. Disponible : https://www.mdpi.com/2079-7737/11/1/134

[127] V. Gupta *et al.*, "Deep learning-based automatic detection of poorly positioned mammograms to minimize patient return visits," *arXiv preprint*, 2020. [En ligne]. Disponible : https://doi.org/10.48550/arXiv.2009.13580

[128] T. Tanyel, N. Denizoglu, M. E. Seker, D. Alis, E. Cerekci, E. Karaarslan, E. Aribal et I. Oksuz, "Mammographic breast positioning assessment via deep learning," dans *Deep Breast Workshop on AI and Imaging for Diagnostic and Treatment Challenges in Breast Care*. Springer, 2024, p. 107–116. [En ligne]. Disponible : https://doi.org/10.48550/arXiv.2407.10796

[129] Y. Lin, M. Li, S. Chen, L. Yu et F. Ma, "Nipple detection in mammogram using a new convolutional neural network architecture," dans *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2019, p. 1–6. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8966022

[130] F.-F. Yin, M. L. Giger, K. Doi, C. J. Vyborny et R. A. Schmidt, "Computerized detection of masses in digital mammograms : automated alignment of breast images and its effect on bilateral-subtraction technique," *Medical Physics*, vol. 21, n°. 3, p. 445–452, 1994. [En ligne]. Disponible : https://doi.org/10.1118/1.597307

[131] A. J. Mendez, P. G. Tahoces, M. J. Lado, M. Souto, J. Correa et J. J. Vidal, "Automatic detection of breast border and nipple in digital mammograms," *Computer Methods and Programs in Biomedicine*, vol. 49, n°. 3, p. 253–262, 1996. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/8800610/

[132] R. Chandrasekhar et Y. Attikiouzel, "A simple method for automatically locating the nipple on mammograms," *IEEE Transactions on Medical Imaging*, vol. 16, n°. 5, p. 483–494, 1997. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/640738

[133] C. Zhou, H. P. Chan, C. Paramagul, M. A. Roubidoux, B. Sahiner, L. M. Hadjiiski et N. Petrick, "Computerized nipple identification for multiple image analysis in computer-aided diagnosis," *Medical Physics*, vol. 31, n°. 10, p. 2871–2882, Oct 2004. [En ligne]. Disponible : https://pubmed.ncbi.nlm.nih.gov/15543797/

[134] J. E. Iglesias et N. Karssemeijer, “Robust initial detection of landmarks in film-screen mammograms using multiple ffdm atlases,” *IEEE Transactions on Medical Imaging*, vol. 28, n°. 11, p. 1815–1824, 2009. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/5071222

[135] M. Mustra, J. Bozek et M. Grgic, “Nipple detection in craniocaudal digital mammograms,” dans *2009 International Symposium ELMAR*. IEEE, 2009, p. 15–18. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/5342871

[136] P. Casti, A. Mencattini, M. Salmeri et M. Ancona, “Automatic detection of the nipple in screen-film and full-field digital mammograms using a novel hessian-based method,” *Journal of Digital Imaging*, 2013. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC3782608/

[137] M. Jas, S. Mukhopadhyay, J. Chakraborty et N. Khandelwal, “A heuristic approach to automated nipple detection in digital mammograms,” *Journal of Digital Imaging*, vol. 9, p. 932–940, 2013. [En ligne]. Disponible : https://doi.org/10.1007/s10278-013-9575-x

[138] J. Jiang, Y. Zhang, Y. Lu, Y. Guo et H. Chen, “A radiomic feature–based nipple detection algorithm on digital mammography,” *Medical Physics*, vol. 46, n°. 10, p. 4381–4391, 2019. [En ligne]. Disponible : https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13684

[139] E. Garcia, R. Martin, J. Martin, J. del Riego, C. Aynes, A. Oliver et O. Diaz, “Simultaneous pectoral muscle and nipple location in mlo mammograms, considering image quality assumptions,” dans *Proceedings of SPIE Medical Imaging — 16th International Workshop on Breast Imaging (IWBI2022)*, vol. 12286, 2022, p. 122860E. [En ligne]. Disponible : https://doi.org/10.1117/12.2625778

[140] Y.-C. Zeng, “Realizing nipple in profile recognition and nipple detection using a single classification,” dans *2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2023, p. 1500–1505. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/10317394

[141] P. Karnati, “Automatic assessment of mammographic images : positioning and quality assessment,” M. Eng. Thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2021. [En ligne]. Disponible : https://hdl.handle.net/1721.1/130694

[142] R. Chandrasekhar, M. K. Sze et Y. Attikiouzel, “Automatic evaluation of mammographic adequacy and quality on the mediolateral oblique view,” dans *Digital Mammography*, H.-O. Peitgen, édit. Berlin, Heidelberg : Springer Berlin Heidelberg, 2003, p. 182–186. [En ligne]. Disponible : https://doi.org/10.1007/978-3-642-59327-7_43

[143] S. M. Kwok, R. Chandrasekhar et Y. Attikiouzel, “Automatic assessment of mammographic positioning on the mediolateral oblique view,” dans *2004 International Conference on Image Processing, 2004. ICIP ’04.*, vol. 1, 2004, p. 151–154 Vol. 1. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/1418712

[144] M. Moran, A. Conci, S. Rêgo et C. A. P. Fontes, “Techniques for automated analysis of mammography positioning failures,” dans *European Congress of Radiology (ECR) 2018, Scientific Exhibit, Poster C-1089*, 2018. [En ligne]. Disponible : https://dx.doi.org/10.1594/ecr2018/C-1089

[145] H. Watanabe *et al.*, “Quality control system for mammographic breast positioning,” *Scientific Reports*, vol. 13, nº. 7066, 2023. [En ligne]. Disponible : https://doi.org/10.1038/s41598-023-34380-9

[146] Y.-C. Zeng et Y.-H. Chen, “Skinfold recognition for positioning quality evaluation on mammography,” dans *2022 IEEE 11th Global Conference on Consumer Electronics (GCCE)*, 2022, p. 862–864. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/10014258

[147] Y.-C. Zeng, Y.-C. Wu, C.-Y. Yeh, S.-C. Li, T.-H. Chou, Y.-W. Huang, G.-C. Hsu et H.-H. Hsu, “Mammography quality evaluation and model interpretation based on cnn-based inframammary fold classification,” dans *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2022, p. 1259–1264. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9980186

[148] P. Hejduk, R. Sexauer, C. Ruppert, K. Borkowski, J. Unkelbach et N. Schmidt, “Automatic and standardized quality assurance of digital mammography and tomosynthesis with deep convolutional neural networks,” *Insights into Imaging*, vol. 14, nº. 1, p. 90, May 2023. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC10195933/

[149] Volpara Health Limited. (2026) Analytics™ – Breast Health Software by Volpara Health. Volpara Health. [En ligne]. Disponible : https://www.volparahealth.com/breast-health-software/products/analytics/

[150] Densitas Inc. (2026) intellimammo® — continuous mammography quality improvement solution. Densitas. [En ligne]. Disponible : https://densitashealth.com/densitas-intellimammo/

[151] b-rayZ AG. (2026) b-rayZ Technology – AI-Powered Breast Imaging Solution. b-rayZ. [En ligne]. Disponible : https://b-rayz.com/technology/

[152] G. R. Kolb, K. Wang, A. Chan et R. P. Highnam, “Automating quality assurance metrics to assess adequate breast positioning in mammography,” 2015. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:7087088

[153] B. Sekachev, N. Manovich, M. Zhiltsov, A. Zhavoronkov, D. Kalinin, B. Hoff, TOsmanov, D. Kruchinin, A. Zankevich, DmitriySidnev, M. Markelov, Johannes222, M. Chenuet, a-andre, telenachos, A. Melnikov, J. Kim, L. Ilouz, N. Glazov, Priya4607, R. Tehrani, S. Jeong, V. Skubriev, S. Yonekura, v. truong, zliang7, lizhming et T. Truong, "opencv/cvat : v1.1.0," août 2020. [En ligne]. Disponible : https://doi.org/10.5281/zenodo.4009388

[154] E. Gourd, "Mammography deficiencies : the result of poor positioning," *The Lancet Oncology*, vol. 19, n°. 8, p. e385, 2018. [En ligne]. Disponible : https://doi.org/10.1016/S1470-2045(18)30489-3

[155] T. Santner, C. Ruppert, S. Gianolini, J.-G. Stalheim, S. Frei, M. Hondl, V. Fröhlich, S. Hofvind et G. Widmann, "Pgmi assessment in mammography : Ai software versus human readers," *Radiography*, vol. 31, n°. 5, p. 103017, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1078817425001610

[156] Z. Zhang, Q. Liu et Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, n°. 5, p. 749–753, 2018. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8309343

[157] Z. Yi, J. Chen, Z. Cai, J. Xiao, J. Liu, H. Liang, C. Quan, X. Zu, L. Wu et J. Hu, "Ma-net : A multi-attention segmentation network for anisotropic prostate magnetic resonance images," *Engineering Applications of Artificial Intelligence*, vol. 157, p. 111348, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0952197625013508

[158] M. M. Kappan, E. B. Sandoval, E. Meijering et F. Cruz, "A survey on deep learning for 2d and 3d human pose estimation," *Artificial Intelligence Review*, vol. 59, n°. 1, p. 32, 2025. [En ligne]. Disponible : https://link.springer.com/10.1007/s10462-025-11430-4

[159] W. Huang, C. Yang et T. Hou, "Spine landmark localization with combining of heatmap regression and direct coordinate regression," *ArXiv*, vol. abs/2007.05355, 2020. [En ligne]. Disponible : https://api.semanticscholar.org/CorpusID:220486819

[160] C. Payer, D. Štern, H. Bischof et M. Urschler, "Integrating spatial configuration into heatmap regression based cnns for landmark localization," *Medical Image Analysis*, vol. 54, p. 207–219, 2019. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1361841518305784

[161] K. H. Brodersen, C. S. Ong, K. E. Stephan et J. M. Buhmann, "The balanced accuracy and its posterior distribution," dans *2010 20th International Conference on Pattern Recognition*, 2010, p. 3121–3124. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/5597285

[162] O. Rainio, J. Teuho et R. Klén, “Evaluation metrics and statistical tests for machine learning,” *Scientific Reports*, vol. 14, n°. 1, p. 6086, mar 2024. [En ligne]. Disponible : https://doi.org/10.1038/s41598-024-56706-x