

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université de 8 Mai 1945 – Guelma -
Faculté des Mathématiques, d'Informatique et des Sciences de la matière
Département d'Informatique



Mémoire de Fin d'études Master

Filière : Informatique

Option : Science et technologie de l'information et de la communication

Thème

**Combinaison de classifieurs pour la reconnaissance
des anomalies mammaires**

Présenté par:

FADEL Ouissal

Membres du jury :

N	Nom et prénom	Qualité
1	BENHAMDI Soulef	Président
2	FERKOUS Chokri	Superviseur
3	HALLACI Samir	Examineur

JUIN 2023

Résumé

Le cancer du sein est le cancer le plus fréquent chez la femme, où la détection précoce joue un rôle important et augmente les chances de guérison complète et permet dans certains cas un traitement moins invasif. Ces dernières années, l'intégration des techniques d'apprentissage automatique (ML) et spécialement l'apprentissage profond (DL) ont montré des résultats prometteurs dans la classification du cancer du sein.

Les expérimentations que nous avons réalisées dans ce projet combinent trois modèles d'apprentissage par transfert (VGG16, Resnet50 et efficient_B0) et cinq techniques classiques d'apprentissage supervisé (MLP, SVM, les forêts aléatoires, les arbres de décision et la régression logistique) en se basant sur les caractéristiques d'haralick ainsi que l'âge du patient et la densité de la mammographie. Dans ce travail nous avons utilisé quatre bases de données complètes. Ces expérimentations nous permettront de choisir la meilleure combinaison qui inclut efficient_B0 et SVM. Cette combinaison atteint un taux de reconnaissance de 96,70%, une précision de 99,6%, un rappel de 93,8% et un score FI de 96,6%.

Mots-clefs : Apprentissage automatique, apprentissage profond, apprentissage par transfert, Mammographie, Extraction de caractéristiques.

ملخص

سرطان الثدي هو أكثر أنواع السرطانات شيوعاً عند النساء، حيث يلعب الاكتشاف المبكر دوراً مهماً، حيث يزيد من فرص الشفاء التام، وفي بعض الحالات يُتيح علاجاً أقل تدخلاً. في السنوات الأخيرة أظهر ادماج تقنيات التعلم الآلي (ML) والتعلم العميق (DL) بصفة خاصة نتائج واعدة في تصنيف سرطان الثدي.

تجمع التجارب التي أجريتها في هذا المشروع بين ثلاثة نماذج لنقل التدريب (VGG16, Resnet50 و Efficient_B0) وخمسة تقنيات تعلم كلاسيكية (MLP, SVM, Les forets aléatoires, Les arbres de décision, La régression logique) وذلك بالاعتماد على خصائص Haralick بالإضافة الى عمر المريض وكثافة التصوير الشعاعي للثدي. في هذا العمل، استخدمنا أربع قواعد بيانات كاملة. ستمكنا هذه التجارب من اختيار أفضل تركيبة تشتمل على (efficient_B0 و SVM)، حيث حقق هذا المزيج معدل تعرف يبلغ 96.70٪ ودقة 99.6٪ واسترجاع 93.8٪ ودرجة F1 96.6٪.

الكلمات المفتاحية: التعلم العميق، التعلم الآلي، نقل التدريب، التصوير الشعاعي للثدي، استخراج الميزات.

Abstract

Breast cancer is the most common cancer in women, where early detection plays an important role, increasing the chances of a complete cure and in some cases, enabling less intrusive treatment. In recent years, the integration of machine learning (ML) techniques and especially deep learning (DL) have shown promising results in breast cancer classification.

The experiments we have carried out in this project combine three transfer learning models (VGG16, Resnet50 and efficient_B0) and five classical supervised learning techniques (MLP, SVM, random forests, decision trees and logistic regression) using haralick features as well as patient age and mammography density. In this work, we used four complete databases. These experiments will enable us to choose the best combination that includes efficient_B0 and SVM. This combination achieves a recognition rate of 96.70%, a precision of 99.6%, a recall of 93.8% and an F1 score of 96.6%.

Keywords: Machine Learning, Deep Learning, Transfer Learning, Mammography, Feature Extraction.

REMERCIEMENTS

Je voudrais exprimer ma profonde gratitude à Dieu pour m'avoir donné la force et la détermination de mener à bien cette mémoire.

Je tiens également à exprimer ma reconnaissance sincère envers mes parents. Je suis profondément redevable pour leurs encouragements constants et leur confiance en mes capacités.

Je souhaite également exprimer ma profonde gratitude à mon superviseur Dr. FERKOUS Chokri, pour ses conseils inébranlables, son expertise et son engagement en faveur dans ce projet. Son mentorat et son approche disciplinée ont joué un rôle crucial dans l'élaboration de cette mémoire.

Je tiens à remercier les membres du jury qui ont accepté la responsabilité d'évaluer mon travail. Je vous remercie pour votre dévouement, vos conseils et votre considération tout au long de ce processus.

Je voudrais également exprimer ma sincère gratitude à mes amis et à mes camarades de classe et surtout à ma meilleure amie BOURAZI Roukia pour son soutien et son amitié indéfectibles. Merci d'avoir toujours été là pour moi.

Enfin, je voudrais me remercier moi-même d'avoir entrepris ce projet. J'ai été mon plus grand supporter, me poussant à surmonter les obstacles et à persévérer pendant cette période.

Une fois encore, j'exprime ma profonde gratitude à tous ceux qui ont joué un rôle dans mon parcours académique.

DEDICACE

Je dédie ce travail

À mon père et ma mère, qui m'a tout donné,

À ma sœur et à mon frère,

À ma chère grand-mère, que Dieu la protège,

Et à ma charmante tante maternelle,

À mes amis, mes précieux compagnons de vie, ma source de rires et de souvenirs sans

fin.

FADEL Ouissal

Table des matières

RESUME	I
ملخص	II
ABSTRACT	III
REMERCIEMENTS	IV
DEDICACE	V
TABLE DES MATIERES	1
LISTE DES FIGURES	6
LISTE DES TABLEAUX	8
ABREVIATIONS ET ACRONYMES	9
INTRODUCTION GENERALE	10
CHAPITRE I: INTELLIGENCE & VISION ARTIFICIELLE	13
I.1. INTRODUCTION	14
I.2. L'INTELLIGENCE ARTIFICIELLE	14
I.3. L'APPRENTISSAGE AUTOMATIQUE	15
I.3.1. <i>L'apprentissage supervisé</i>	16
I.3.1.1. Les machines à vecteur de support SVM	16
I.3.1.2. L'Arbre de décision	17
I.3.1.3. Forêt aléatoire	18
I.3.1.4. Régression logistique	18
I.3.1.5. Les réseaux de neurones	19
I.3.1.6. L'apprentissage profond (Deep Learning)	20
I.3.2. <i>L'apprentissage non supervisé</i>	20
I.3.2.1. L'algorithme des k plus proches voisins	21
I.3.2.2. Le clustering K-moyenne	22
I.3.3. <i>L'apprentissage par renforcement</i>	23
I.4. DOMAINES D'UTILISATION DE L'APPRENTISSAGE AUTOMATIQUE	23
I.5. CONCLUSION	24
CHAPITRE II: LES RESEAUX DE NEURONES ARTIFICIELS ET A CONVOLUTION	25
II.1. INTRODUCTION	26
II.2. NEURONE BIOLOGIQUE	26
II.3. NEURONE ARTIFICIEL	27

II.4.	ARCHITECTURE DES RESEAUX DE NEURONES.....	27
II.4.1.	<i>Les fonctions d'activations</i>	28
II.4.1.1.	La fonction d'étape binaire	28
II.4.1.2.	Fonction d'activation linéaire.....	29
II.4.1.3.	Fonction d'activation sigmoïde	29
II.4.1.4.	Fonction d'activation Tanh.....	30
II.4.1.5.	Fonction d'activation ReLU.....	30
II.4.1.6.	Fonction d'activation Leaky ReLU	31
II.4.1.7.	Fonction PReLU : (Paramétric ReLU)	31
II.4.1.8.	Fonction ELU :	32
II.4.1.9.	Fonction d'activation swish.....	32
II.4.1.10.	Fonction d'activation softmax	32
II.4.2.	<i>Les algorithmes d'optimisation</i>	33
II.4.2.1.	La descente de gradient (GD).....	33
II.4.2.1.1.	La descente du gradient par batch.....	33
II.4.2.1.2.	Descente du gradient stochastique.....	34
II.4.2.1.3.	Descente du gradient avec mini-batch.....	34
II.4.2.2.	Momentum	35
II.4.2.3.	Gradient accéléré de Nesterov	35
II.4.2.4.	Adagrad	36
II.4.2.5.	RMSprop.....	36
II.4.2.6.	Adadelta	37
II.4.2.7.	Adam.....	37
II.4.2.8.	Nadam	38
II.5.	APPRENTISSAGE PROFOND (DEEP LEARNING OU DL)	38
II.5.1.	<i>Les réseaux de neurones convolutifs</i>	38
II.5.1.1.	Les couches d'un CNN.....	38
II.5.1.1.1.	Couche de convolution	39
II.5.1.1.2.	Couche de pooling.....	40
II.5.1.1.3.	La couche fully connected (FC).....	40
II.6.	QUELQUES ARCHITECTURE BASEES CNN	41
II.6.1.	<i>VGG</i>	41
II.6.2.	<i>GoogLeNet</i>	41
II.6.3.	<i>ResNet</i>	42
II.6.4.	<i>EfficientNet</i>	42
II.6.5.	<i>EfficientNet-BO</i> :	43
II.7.	APPRENTISSAGE PAR TRANSFERT	44
II.8.	CONCLUSION.....	45
CHAPITRE III: EXTRACTION DES CARACTERISTIQUES		46
III.1.	INTRODUCTION.....	47

III.2.	EXTRACTION DES CARACTERISTIQUES LINEAIRES ET NON LINEAIRES	47
III.3.	TYPES D'EXTRACTION DE CARACTERISTIQUES.....	48
III.3.1.	<i>Caractéristiques de forme</i>	48
III.3.1.1.	Les algorithmes d'image binaire.....	48
III.3.1.2.	Segmentation horizontale et verticale.....	49
III.4.	LES CARACTERISTIQUE DE COULEUR.....	49
III.4.1.	<i>Moments de couleur</i>	50
III.4.2.	<i>Histogramme de couleur</i>	50
III.4.3.	<i>Corrélogramme de couleur</i>	50
III.4.4.	<i>La matrice de cooccurrence des couleurs (MCC)</i>	51
III.5.	CARACTERISTIQUES DE TEXTURE	52
III.5.1.	<i>SGLDM (Spatial gray level dependence matrices)</i>	52
III.5.1.1.	Le Contraste	53
III.5.1.2.	L'Entropie	54
III.5.1.3.	La Variance	54
III.5.1.4.	La Corrélacion	54
III.5.1.5.	La moyenne de la somme	54
III.5.1.6.	Le second moment angulaire(ASM).....	55
III.5.1.7.	IDM.....	55
III.5.1.8.	La somme de variances	55
III.5.1.9.	La somme des entropies	55
III.5.1.10.	La variance de différence	55
III.5.1.11.	L'entropie de différence	56
III.5.1.12.	IMC1	56
III.5.1.13.	IMC2	56
III.5.2.	<i>Local binary pattern (LBP)</i>	56
III.5.3.	<i>Histogramme du gradient orienté (HOG)</i>	57
III.5.4.	<i>Filtres de Gabor</i>	57
III.6.	CONCLUSION.....	57
CHAPITRE IV: MAMMOGRAPHIE ET DIAGNOSTIC ASSISTE PAR ORDINATEUR.....		58
IV.1.	INTRODUCTION.....	59
IV.2.	LES ANOMALIES MAMMAIRES	59
IV.2.1.	<i>Anomalies mammaires</i>	59
IV.2.1.1.	Masse	59
IV.2.1.2.	Les Microcalcifications	60
IV.2.1.3.	Distorsions architecturales.....	61
IV.3.	L'IMAGERIE MEDICALE DEDIEE AU DEPISTAGE DES ANOMALIES MAMMAIRES.....	61
IV.3.1.	<i>Echographie</i>	61
IV.3.2.	<i>Imagerie par résonance magnétique (IRM)</i>	61

IV.3.3.	<i>La Mammographie</i>	61
IV.4.	DETECTION ET DIAGNOSTIC ASSISTE PAR ORDINATEUR	63
IV.4.1.	<i>Évolution de la DAO</i>	64
IV.4.1.1.	DAO en seconde lecture.....	64
IV.4.1.2.	DAO à lecture simultanée	64
IV.4.1.3.	DAO en première lecture.....	65
IV.4.2.	<i>Diagnostic à apprentissage continu (CAO)</i>	65
IV.4.3.	<i>Mesures de performance des CAD</i>	65
IV.4.3.1.	Accuracy	66
IV.4.3.2.	Précision.....	66
IV.4.3.3.	Rappel ou sensibilité	66
IV.4.3.4.	Spécificité.....	66
IV.4.3.5.	Score F1	66
IV.5.	REVUE DE LA LITTERATURE	66
IV.5.1.	<i>Les systèmes d'aide au diagnostic basé sur les caractéristiques</i>	67
IV.5.2.	<i>Les systèmes d'aide au diagnostic basés sur l'apprentissage profond</i>	70
IV.6.	CONCLUSION :	72
CHAPITRE V: COMBINAISON DE CLASSIFIEURS POUR LE DIAGNOSTIC DES ANOMALIES MAMMAIRES		73
V.1.	INTRODUCTION.....	74
V.2.	PLATEFORME DE DEVELOPPEMENT.....	74
V.3.	ENVIRONNEMENT LOGICIEL	74
V.3.1.	<i>Langage de programmation Python</i>	74
V.3.2.	<i>Tensorflow</i>	75
V.3.3.	<i>Keras</i>	75
V.3.4.	<i>SKlearn</i>	75
V.3.5.	<i>Mahotas</i>	76
V.3.6.	<i>Numpy</i>	76
V.3.7.	<i>Pandas</i>	76
V.3.8.	<i>CV2</i>	76
V.3.9.	<i>Matplotlib</i>	76
V.3.10.	<i>PIL</i>	76
V.3.11.	<i>Gradio</i>	77
V.4.	ARCHITECTURE DU MODELE.....	77
V.4.1.	<i>Collections des mammographies</i>	78
V.4.1.1.	RSNA dépistage du cancer du sein par mammographie	78
V.4.1.2.	Base de données chinoise sur la mammographie (CMMD).....	78
V.4.1.3.	Mini-DDSM	79

Table des matières

V.4.1.4.	INbreast	79
V.4.2.	<i>Prétraitement et préparation des mammographies</i>	80
V.4.2.1.	Recadrer et redimensionner	81
V.4.2.2.	Augmentation et répartition des données.....	82
V.4.3.	<i>Sélection du premier classifieur</i>	83
V.4.3.1.	Apprentissage par transfert.....	83
V.4.3.2.	Modèle élu.....	86
V.4.4.	<i>Sélection du deuxième classifieur</i>	88
V.4.4.1.	Extraction des caractéristiques	88
V.4.4.2.	Apprentissages des modèles classiques	89
V.4.4.3.	Modèle classique élu	91
V.5.	DEPLOIEMENT DE L'APPLICATION :	92
V.6.	CONCLUSION :	93
CONCLUSION GENERALE		95
REFERENCES BIBLIOGRAPHIQUES		97
WEBOGRAHIE		104

Liste des Figures

FIGURE I-1: LES RELATIONS ENTRE L'INTELLIGENCE ARTIFICIELLE, L'APPRENTISSAGE AUTOMATIQUE ET L'APPRENTISSAGE PROFOND..	14
FIGURE I-2: TYPES D'APPRENTISSAGE EN APPRENTISSAGE PAR MACHINE.	16
FIGURE I-3: SCHEMA D'APPRENTISSAGE SUPERVISE "SUPERVISED LEARNING" (RUSSELL, 2020).....	16
FIGURE I-4: EXEMPLE SIMPLE REPRESENTANT LA CLASSIFICATION PAR SVM.....	17
FIGURE I-5: EXEMPLE DE LA STRUCTURE D'UN ARBRE DE DECISION	17
FIGURE I-6: UN EXEMPLE DE LA STRUCTURE D'UNE FORET ALEATOIRE PRENANT EN COMPTE PLUSIEURS ARBRES DE DECISION.....	18
FIGURE I-7: ARCHITECTURE D'UN MLP A DEUX COUCHES CACHEES. (RAMKUMAR ET AL., 2021).....	20
FIGURE I-8: SCHEMA D'APPRENTISSAGE NON SUPERVISE	21
FIGURE I-9: EXEMPLE DE K-PLUS PROCHE VOISINS (KPPV).....	21
FIGURE II-1: NEURONE BIOLOGIQUE	26
FIGURE II-2: REPRESENTATION SCHEMATIQUE D'UN NEURONE ARTIFICIEL (SARKER, 2021A)	27
FIGURE II-3: RESEAU NEURONAL AVEC 2 COUCHES CACHEES	27
FIGURE II-4: FONCTION D'ETAPE BINAIRE	28
FIGURE II-5 : GRAPH DE LA FONCTION D'ACTIVATION LINEAIRE.....	29
FIGURE II-6 : GRAPH DE LA FONCTION D'ACTIVATION SIGMOÏDE.....	29
FIGURE II-7 : GRAPH DE LA FONCTION D'ACTIVATION TANH	30
FIGURE II-8 : GRAPH DE LA FONCTION D'ACTIVATION RELU	30
FIGURE II-9 : FONCTION D'ACTIVATION LEAKY RELU	31
FIGURE II-10 : GRAPH DE LA FONCTION D'ACTIVATION PRELU.....	31
FIGURE II-11 : FONCTION D'ACTIVATION ELU	32
FIGURE II-12: UNE ARCHITECTURE TYPIQUE D'UN CNN (ALZUBAIDI ET AL., 2021)	39
FIGURE II-13: EXEMPLE D'OPERATION DE CONVOLUTION AVEC UN NOYAU DE TAILLE 3×3 SANS PADDING.....	39
FIGURE II-14: EXEMPLE D'OPERATION DE CONVOLUTION AVEC PADDING, UN TENSEUR D'ENTREE DE TAILLE 5×5, NOYAU DE 3×3 ET UN STRIDE DE 1. (YAMASHITA ET AL., 2018)	40
FIGURE II-15: TYPES D'OPERATION DE POOLING .(ALZUBAIDI ET AL., 2021)	40
FIGURE II-16: L'ARCHITECTURE VGG16.(KIM ET AL., 2021).....	41
FIGURE II-17: L'ARCHITECTURE GOOGLNET.(ALZUBAIDI ET AL., 2021)	42
FIGURE II-18: L'ARCHITECTURE RESNET.(KIM ET AL., 2021)	42
FIGURE II-19 : L'ARCHITECTURE DE BASE DU MODELE EFFICIENTNET-B0. (SHAH ET AL., 2022).....	44
FIGURE II-20: BLOC MBCONV	44
FIGURE II-21 : APPRENTISSAGE PAR TRANSFERT(ZHUANG ET AL., 2021)	45

FIGURE III-1 : CALCUL DES DISTANCES HORIZONTALES ET VERTICALES. (KALEL ET AL., 2016)	49
FIGURE III-2 : LA FENETRE 3×3 DIVISEE EN GRILLE 2×2.....	51
FIGURE III-3 : SEPT MODELES DE BALAYAGE. (PATEL & GAMIT, 2016).....	51
FIGURE III-4 : PRINCIPE DE CONSTRUCTION DE LA MATRICE DE COOCCURRENCE.....	53
FIGURE IV-1 : EXEMPLE D'UNE MASSE MALIGNE (A GAUCHE) ET UNE MASSE BENIGNE (A DROITE). (MAHMOOD ET AL., 2020)	60
FIGURE IV-2 : EXEMPLE D'UNE MICROCALCIFICATION .(ALSHEH ALI ET AL., 2019)	60
FIGURE IV-3 : UN EXEMPLE D'ETIQUETTES D'IMAGES DE MAMMOGRAPHIE.(WANG ET AL., 2020)	62
FIGURE IV-4 : DAO DE DEUXIEME LECTEUR	64
FIGURE IV-5 : DAO DE LECTEUR SIMULTANE	64
FIGURE IV-6 : DAO DE PREMIER LECTEUR.	65
FIGURE IV-7 :MATRICE DE CONFUSION	65
FIGURE IV-8 : SYSTEMES DAO BASES SUR L'APPRENTISSAGE AUTOMATIQUE	67
FIGURE IV-9 : SYSTEMES DAO BASES SUR L'APPRENTISSAGE PROFOND.....	70
FIGURE V-1 : L'ARCHITECTURE GENERALE DE NOTRE TRAVAIL.	77
FIGURE V-2 : LA DISTRIBUTION DES IMAGES DANS LA BASE RSNA	78
FIGURE V-3 : LA DISTRIBUTION DES IMAGES DANS LA BASE CMMD.....	79
FIGURE V-4 : LA DISTRIBUTION DES IMAGES DANS LA BASE DDSM	79
FIGURE V-5 : LA DISTRIBUTION DES IMAGES DANS LA BASE INBREAST.....	80
FIGURE V-6 : EXEMPLE DES MAMMOGRAPHIES BENIGNES ET MALIGNES DES 4 BASES	80
FIGURE V-7 : PROCESSUS DE RECADRAGE DE LA REGION D'INTERET	81
FIGURE V-8 : EXEMPLE DES 3 TECHNIQUES D'AUGMENTATION SUR UNE IMAGE MALIGNE.....	82
FIGURE V-9 : ÉVOLUTION DE L'APPRENTISSAGE ET LA PERTE DURANT L'APPRENTISSAGE DE VGG16	84
FIGURE V-10 : ÉVOLUTION DE L'APPRENTISSAGE ET LA PERTE DURANT L'ENTRAINEMENT D'EFFICIENTNET_B0	85
FIGURE V-11 : ÉVOLUTION DE L'APPRENTISSAGE ET LA PERTE DURANT L'ENTRAINEMENT DE RESNET50	86
FIGURE V-12 : MODELE ELU BASE SUR L'APPRENTISSAGE PAR TRANSFERT	86
FIGURE V-13 : MATRICE DE CONFUSION D'EFFICIENTNET_B0	87
FIGURE V-14 : MATRICE DE CONFUSION DE SVM EN UTILISANT LES DEGRES DE MALIGNITE ET DE BENIGNITE	91
FIGURE V-15 : INTERFACE GRAPHIQUE (RESULTAT DE CLASSIFICATION D'UNE IMAGE MALIGNE)	93

Liste des tableaux

TABLEAU III-1: FACTEUR DE FORME	49
TABLEAU III-2 : CONTRASTE DE DIFFERENTS DESCRIPTEURS DE COULEUR. (TIAN, 2013)	52
TABLEAU III-3 : DIFFERENCES ENTRE LES METHODES D'EXTRACTION DES CARACTERISTIQUES DE TEXTURE (KAVYA, 2015).....	52
TABLEAU V-1: RECAPITULATIF DES IMAGES D'ENTRAINEMENT, VALIDATION ET DE TEST.....	83
TABLEAU V-2 : RESULTAT DE VGG16.....	84
TABLEAU V-3 : RESULTAT D'EFFICIENTNET_B0.....	84
TABLEAU V-4 : RESULTAT DE RESNET50.....	84
TABLEAU V-5 : RESULTATS DES CLASSIFIEURS CLASSIQUES SANS L'UTILISATION DES DEGRES DE MALIGNITE ET DE BENIGNITE.....	89
TABLEAU V-6 : RESULTATS DES CLASSIFIEURS CLASSIQUES AVEC L'UTILISATION DES RESULTATS DE CLASSIFICATION D'EFFICIENTNET_B0	90
TABLEAU V-7 : RESULTATS DE LA COMBINAISON DES DEUX MODELES AVEC LA PREDICTION DES DEGRES DE MALIGNITE ET BENIGNITE	90
TABLEAU V-8 : EVALUATION DES METRIQUES SVM EN UTILISANT LES DEGRES DE MALIGNITE ET DE BENIGNITE.....	92

Abréviations et Acronymes

IA	Intelligence artificielle
ML	Apprentissage automatique
DL	Apprentissage en profondeur
CNN	Réseau neuronal convolutif
ANN	Artificial neural networks
MLP	Multilayer perceptron
KNN	K -nearest neighbor
SVM	Support vector machine
DT	Decision tree
RF	Random forest
LR	Logistic regression
RELU	Rectified linear unit
TANH	Hyperbolic tangent function
Elu :	Exponential linear unit
SGD	Stochastic gradient descent
VGG	Visual geometry group
SGLDM	Spatial grayscale dependency matrices
IRM	Magnetic resonance imaging
CC	Raniocaudal
MLO	medio-lateral oblique
BIRADS	Breast Imaging Data and Reporting System
DAOe	Computer-aided detection
DAOx	Computer-assisted diagnosis
GPU	Graphics Processing Unit
RSNA	North American Society of Radiology
CMMD	Chinese mammography database
ASM	Angular second moment
IDM	Inverse Difference Moment
IMC1	Information measure of correlation 1
IMC2	Correlation information measure 2
CSV	Comma-separated values

INTRODUCTION GENERALE

Le cancer du sein est une maladie répandue et potentiellement mortelle qui touche des millions de femmes dans le monde. La détection précoce et le diagnostic des anomalies mammaires jouent un rôle important dans l'amélioration de l'état des patientes et l'augmentation des taux de survie. Ces dernières années, les technologies d'imagerie médicale ont connu des progrès considérables, en particulier dans le domaine des images numériques et de la mammographie, qui ont révolutionné le dépistage et le diagnostic du cancer du sein.

Les mammographies numériques fournissent une représentation détaillée et à haute résolution du tissu mammaire, ce qui permet aux professionnels de la santé d'identifier et d'analyser les anomalies avec une plus grande précision. Ces images sont acquises à l'aide d'appareils d'imagerie spécialisés et peuvent être stockées, traitées et analysées à l'aide de techniques informatiques. La mammographie numérique, en particulier, est devenue la référence en matière de dépistage du cancer du sein, car elle offre une meilleure qualité d'image, un meilleur contraste et la possibilité de stocker et de transmettre les images par voie électronique.

La mammographie est une technique d'imagerie spécifique qui utilise des rayons X à faible dose pour capturer la composition du sein. Elle joue un rôle essentiel dans la détection précoce du cancer du sein, en permettant aux radiologues d'identifier les lésions suspectes, les calcifications et autres anomalies pouvant indiquer la présence de cellules cancéreuses. Cependant, l'interprétation des mammographies peut s'avérer difficile et subjective, car elle dépend fortement de l'expertise du radiologue.

Pour améliorer la précision et l'efficacité de la détection des anomalies mammaires, les chercheurs se sont tournés vers des technologies avancées telles que l'apprentissage profond et l'apprentissage automatique. Les modèles d'apprentissage profond, en particulier les réseaux neuronaux convolutifs (CNN), ont démontré des capacités remarquables dans les tâches de reconnaissance et de classification d'images. En exploitant les capacités d'extraction de caractéristiques hiérarchiques des CNN, les chercheurs ont obtenu des améliorations significatives dans la détection d'anomalies mammaires à partir d'images numériques et de mammographies.

Les techniques d'apprentissage automatique, quant à eux, fournissent un cadre pour la l'entraînement et l'optimisation des classificateurs sur la base de données d'apprentissage

étiquetées. Ces techniques peuvent apprendre des modèles et des relations à partir des caractéristiques et faire des prédictions ou des classifications sur de nouveaux échantillons non vus. La combinaison de la puissance de l'apprentissage profond et des techniques d'apprentissage automatique offre une approche prometteuse pour améliorer la précision, la sensibilité et la spécificité de la détection du cancer du sein.

Dans ce mémoire de fin d'étude, l'accent est mis sur la combinaison des forces de l'apprentissage profond et des approches traditionnelles d'apprentissage automatique pour la classification des anomalies mammaires. L'objectif est de tirer parti des capacités complémentaires de ces deux paradigmes pour améliorer la précision et la fiabilité de la classification du cancer du sein. Ce mémoire de fin d'études est structuré en cinq chapitres :

Dans le premier chapitre, nous présentons une vue d'ensemble de l'intelligence artificielle (IA), de l'apprentissage machine (ML) et de l'apprentissage profond (DL) dans le contexte des tâches de vision. Nous discutons des concepts fondamentaux et des algorithmes utilisés dans ces domaines. Le chapitre explore diverses techniques d'IA, y compris l'apprentissage supervisé et non supervisé, et leur application dans les tâches de vision par ordinateur.

Dans le deuxième chapitre, nous nous plongeons dans le monde des réseaux neuronaux artificiels (ANN) et nous nous concentrons sur les réseaux neuronaux convolutifs (CNN). Nous expliquons l'inspiration biologique qui sous-tend les réseaux neuronaux et expliquons le fonctionnement des neurones artificiels. Le chapitre couvre les différentes fonctions d'activation et les optimiseurs couramment utilisés dans les réseaux neuronaux. Nous discutons également de l'architecture et des variantes des réseaux neuronaux, en mettant l'accent sur leur efficacité dans l'analyse des images de mammographie numérique.

Le troisième chapitre est consacré aux techniques d'extraction de caractéristiques, et plus particulièrement aux méthodes basées sur la texture. Nous explorons diverses techniques d'extraction de caractéristiques pertinentes dans le domaine des mammographies. Nous discutons de l'importance de la capture des informations de texture pour améliorer la détection et la caractérisation des anomalies mammaires.

Le quatrième chapitre fournit une compréhension complète de la mammographie. Nous nous penchons sur le processus d'acquisition des images, les défis liés à l'interprétation des mammographies et le rôle des systèmes de diagnostic assisté par ordinateur (CAO). Nous discutons de l'intégration des techniques d'IA et de ML dans les systèmes CAD afin d'aider les radiologues dans la détection et le diagnostic du cancer du sein.

Le dernier chapitre est consacré à l'application pratique de l'approche de combinaison de classifieurs que nous avons proposée pour la détection d'anomalies dans les mammographies. Nous présentons la configuration expérimentale, y compris les bases de données et les mesures d'évaluation utilisées. En outre, nous discutons des résultats obtenus par notre modèle combiné d'apprentissage profond et d'apprentissage automatique et fournissons une analyse de ses performances.

Enfin, la conclusion générale présente une synthèse de ce travail et les perspectives qui en découlent.

Chapitre

1

INTELLIGENCE & VISION ARTIFICIELLE

I.1. Introduction

L'apprentissage est un processus naturel du comportement humain et devient également une partie importante des machines. Aujourd'hui, l'intelligence artificielle (IA), l'apprentissage automatique et l'apprentissage profond (Deep Learning) sont trois termes populaires qui sont parfois utilisés de manière interchangeable pour décrire des systèmes ou des logiciels qui se comportent intelligemment. Ces technologies apparaissent dans d'innombrables travaux scientifiques. A travers ces techniques plusieurs programmes sont développés pour automatiser des tâches spécifiques et qui sont utilisée pour un large éventail d'activités. Grâce à l'apparition de ces technologies ont permet de simplifier les taches dans divers domaines y compris le contrôle des robots des voitures autonomes et d'assistants virtuels, la télédétection, le commerce électronique et le diagnostic médical... ect.

Dans ce chapitre nous présentons les concepts essentiels pour comprendre les techniques de l'intelligence artificielle et de l'apprentissage automatique.

I.2. L'intelligence artificielle

Les origines de l'intelligence artificielle remontent aux années 1950, lorsque quelques pionniers de l'informatique naissante ont commencé à se demander si les ordinateurs pouvaient avoir des capacités de "pensée". Cette question, qui nous captive encore aujourd'hui, a ouvert la voie à l'exploration des possibilités de l'IA. En termes simples, selon Chollet (Chollet, 2021), l'intelligence artificielle peut être définie comme la tentative d'automatiser des tâches intellectuelles généralement effectuées par des humains. À l'origine, les premiers programmes d'échecs reposaient sur des règles codées en dur créées par des programmeurs, ce qui n'impliquait pas d'apprentissage automatique.

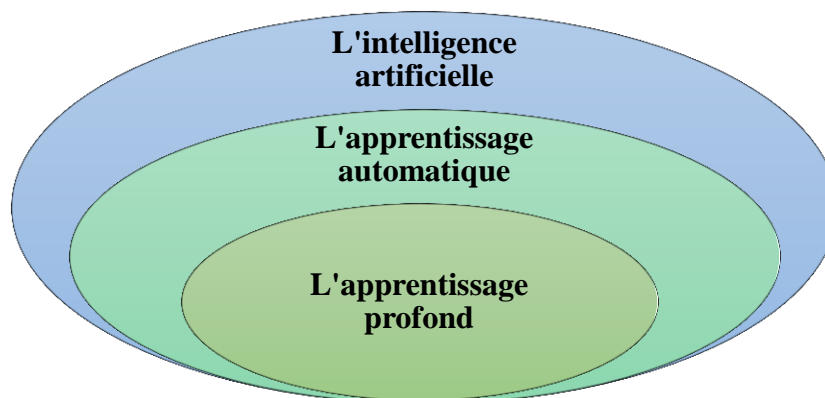


Figure I-1: Les relations entre l'intelligence artificielle, l'apprentissage automatique et l'apprentissage profond.

L'IA symbolique, également connue sous le nom d'IA classique, est née de la croyance dominante selon laquelle une intelligence artificielle de niveau humain pouvait être obtenue en construisant un vaste ensemble de règles explicites.(Chollet, 2021)

La relation entre l'intelligence artificielle (IA), l'apprentissage automatique, l'apprentissage profond et les liens qui les unissent est clairement illustrée dans la Figure I-1.

Dans les années 1980, l'intelligence artificielle a connu une avancée significative avec la montée en puissance des systèmes experts. L'intelligence artificielle symbolique a démontré sa capacité à résoudre des problèmes logiques bien définis, comme le jeu d'échecs. Cependant, elle s'est heurtée à des limites lorsqu'elle s'est attaquée à des problèmes complexes et ambigus tels que la classification d'images, la reconnaissance vocale et la traduction automatique, pour lesquels il était difficile de définir des règles explicites. Un nouveau paradigme, l'apprentissage automatique, est apparu pour remplacer l'IA symbolique.

L'apprentissage automatique offre une approche différente en s'appuyant sur des outils de programmation, des langages et des algorithmes pour traiter un large éventail d'applications au-delà de la robotique. De la traduction automatique à la sécurité, en passant par la reconnaissance vocale, la reconnaissance de l'écriture manuscrite, la médecine et bien d'autres encore, l'intelligence artificielle s'est infiltrée dans divers domaines informatiques. Passant souvent inaperçue, elle a le potentiel d'améliorer presque tous les types de logiciels. Le domaine de l'intelligence artificielle se développe rapidement, alimenté par les progrès des capacités informatiques qui permettent le développement d'algorithmes jusqu'alors inimaginables. Étant donné le rôle indéniable de l'intelligence artificielle dans la construction de notre avenir, il est essentiel pour les développeurs et les informaticiens de comprendre les mécanismes sous-jacents et de les appliquer efficacement.(Mathivet, 2017)

I.3. L'apprentissage automatique

L'apprentissage automatique, une sous-catégorie de l'intelligence artificielle. Il s'agit d'un auto-apprentissage basé sur un algorithme, ce qui signifie que le système apprend de son expérience. Par exemple, le type de données donné en entrée au système apprend le modèle et répond à partir de son apprentissage à la sortie. Dans ce cas, le système devient de plus en plus intelligent au fil du temps sans intervention humaine. Il utilise un algorithme d'apprentissage statistique qui apprend et s'améliore automatiquement sans aide humaine.

(N. Sharma et al., 2021)

Généralement, les approches d'apprentissage automatique peuvent être classées en 4 types principaux (Supervisé, Non supervisé, Semi-supervisé et l'apprentissage par renforcement)

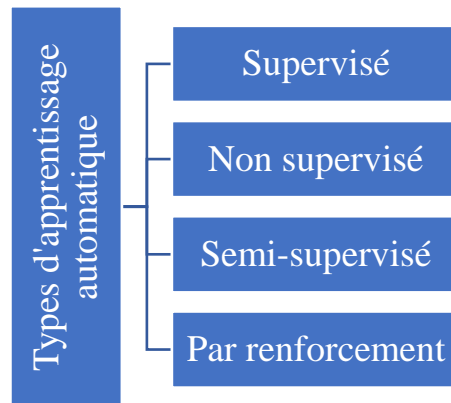


Figure I-2: Types d'apprentissage en apprentissage par machine.

I.3.1. L'apprentissage supervisé

L'objectif de l'apprentissage supervisé est d'utiliser que des données étiquetées pour apprendre. Il existe principalement deux types d'algorithmes d'apprentissage supervisé : les algorithmes de régression, qui peuvent convertir l'espace d'entrée en un domaine de valeurs réelles (par exemple, la régression bayésienne et la régression par vecteur de support), et les algorithmes de classification, qui peuvent convertir l'espace d'entrée en classes prédéfinies (comme les réseaux neuronaux artificiels, les arbres de décision, les forêts d'arbre et les machines à vecteur de support SVM). (H. Zhang & Moon, 2021)

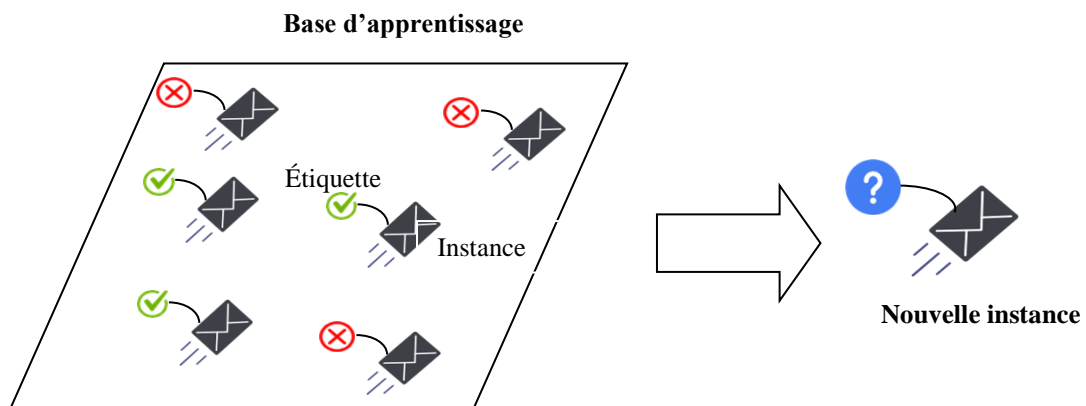


Figure I-3: Schéma d'apprentissage supervisé "supervised learning" (Russell, 2020)

I.3.1.1. Les machines à vecteur de support SVM

La machine à vecteurs de support (SVM) est une méthode d'apprentissage automatique très répandue pour résoudre les problèmes de classification et de régression. Les SVM

utilisent des limites de décision pour diviser les données en différentes classes dans un espace de caractéristiques à haute dimension, et les marges entre ces classes sont tracées pour optimiser la distance entre la marge et les classes. La capacité des SVM à traiter les problèmes de classification non linéaire en utilisant la méthode du noyau, qui leur permet de cartographier implicitement les données d'entrée dans un espace de caractéristiques à haute dimension, est l'un de leurs principaux avantages, et la sélection de la fonction noyau peut avoir un impact significatif sur les performances d'un SVM et peut nécessiter quelques essais et erreurs pour l'optimiser pour une base de données spécifique.

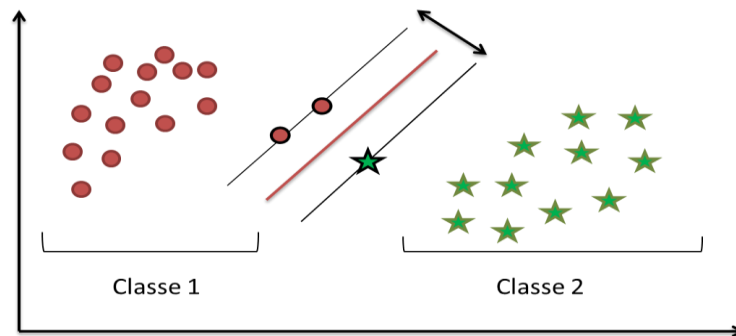


Figure I-4: Exemple simple représentant la classification par SVM.

I.3.1.2. L'Arbre de décision

L'arbre de décision est une illustration des choix et de leurs résultats sous la forme d'un arbre. Les arêtes du graphique représentent les critères ou les règles de décision, tandis que les nœuds du graphique représentent un événement ou une option. Les nœuds et les branches font partie de chaque arbre. Chaque nœud représente un ensemble de caractéristiques dans un groupe de classification et chaque branche indique une valeur possible pour le nœud. (Mahesh, 2019).

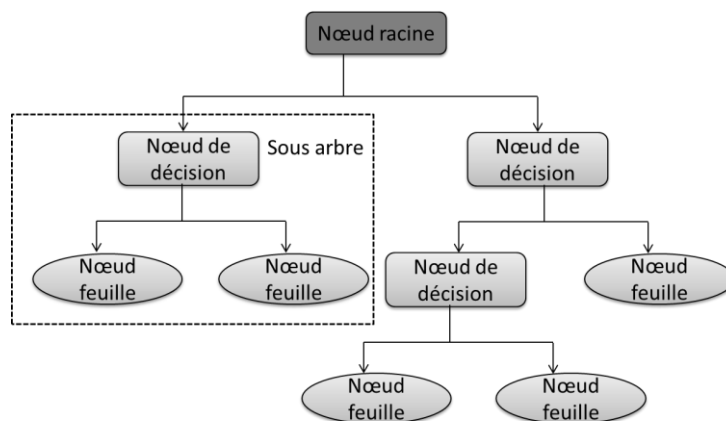


Figure I-5: Exemple de la structure d'un arbre de décision

Lors de la catégorisation d'instances à l'aide d'arbres de décision, l'algorithme commence au nœud racine et descend dans l'arbre en examinant l'attribut défini par chaque nœud et en adhérant à la branche qui correspond à la valeur de l'attribut de l'instance à catégoriser. Lorsqu'un nœud feuille, qui reflète la classification finale de l'instance, est atteint, la procédure est terminée. Deux des mesures les plus souvent utilisées pour diviser les nœuds d'un arbre de décision sont le "gini" pour l'impureté de Gini et l'entropie pour le gain d'information. Afin d'optimiser le gain d'information ou de réduire l'impureté, ces mesures aident à déterminer l'attribut optimal pour diviser les données à chaque nœud.

I.3.1.3. Forêt aléatoire

La forêt aléatoire est une technique d'apprentissage d'ensemble qui ajuste simultanément plusieurs arbres de décision et établit une moyenne ou vote sur les résultats pour les fusionner. Cela permet de réduire le sur-apprentissage et d'augmenter la précision des prédictions. Elle fonctionne en créant un certain nombre d'arbres de décision avec une variance contrôlée en utilisant l'agrégation bootstrap (Bagging) et la sélection aléatoire des caractéristiques.

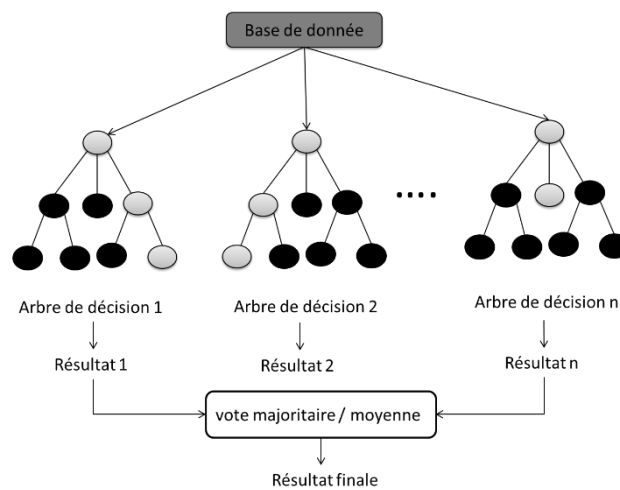


Figure I-6: Un exemple de la structure d'une forêt aléatoire prenant en compte plusieurs arbres de décision.

Les variables catégorielles et continues peuvent être utilisées avec succès avec la forêt aléatoire, ce qui la rend appropriée pour les questions de classification et de régression.

I.3.1.4. Régression logistique

La régression logistique est une approche statistique probabiliste très répandue pour traiter les problèmes de classification dans le cadre de l'apprentissage automatique. La fonction logistique, communément appelée fonction sigmoïde comme montré l'Équation I-1, est utilisée pour calculer la probabilité qu'une entrée entre dans une catégorie spécifique.

$$g(z) = \frac{1}{1 + \exp(-z)} \quad \text{Équation I-1}$$

La simplicité et l'efficacité de la régression logistique, qui en font une option populaire pour de nombreuses applications, est l'un de ses avantages. Mais comme lorsqu'elle fonctionne avec des bases de données de haute dimension, peut entraîner un surapprentissage. Il est possible d'éviter le surapprentissage en utilisant des techniques de régularisation telles que la régularisation L1 et L2. La régression logistique part du principe que les variables dépendantes et indépendantes sont linéairement liées. Cela signifie que la régression logistique peut ne pas fonctionner correctement si la connexion entre les variables n'est pas linéaire. Malgré ces inconvénients, la régression logistique peut être appliquée aux problèmes de classification et de régression, mais les problèmes de classification sont les plus typiques.(Sarker, 2021b)

I.3.1.5. Les réseaux de neurones

Les réseaux neuronaux artificiels classiques (RNA) sont un type de technique d'apprentissage automatique développé depuis les années 1940, ils sont des programmes informatiques capables de reconnaître les corrélations linéaires et non linéaires entre les variables d'entrée et de sortie. Les réseaux neuronaux artificiels (RNA) sont capables de résoudre des problèmes complexes qui ne peuvent être formulés mathématiquement. (Hashemi Fath et al., 2020)

Le réseau neuronal le plus généralisé est appelé perceptron multicouche (MLP). L'objectif fondamental du perceptron multicouche est de développer un modèle qui exécute avec précision la mise en correspondance des données d'entrée avec les données de sortie en utilisant les données antérieures ; par conséquent, le modèle peut être utilisé pour générer des données de sortie même lorsque les données de sortie requises sont inconnues. L'algorithme de rétro-propagation est utilisé avec le MLP pour apprendre le comportement informationnel de l'entrée. La figure ci-dessous illustre l'architecture du MLP avec deux couches cachées.

Un réseau neuronal feed-forward, ou MLP, propage l'information de manière unidirectionnelle de la couche d'entrée à la couche de sortie tout en traversant une ou plusieurs couches cachées. L'efficacité d'un modèle MLP est influencée par un certain nombre de variables, notamment le nombre de couches cachées, de nœuds et de données d'apprentissage, ainsi que par la sélection des paramètres d'apprentissage, tels que le taux d'apprentissage, l'élan et les itérations. (Taud & Mas, 2018)

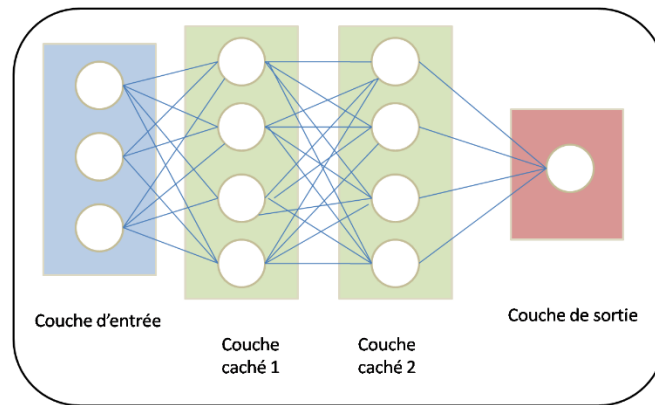


Figure I-7: Architecture d'un MLP à deux couches cachées. (Ramkumar et al., 2021).

I.3.1.6. L'apprentissage profond (Deep Learning)

L'apprentissage profond est une branche avancée de l'apprentissage automatique qui excelle dans les tâches d'analyse complexes et trouve des applications dans la vision par ordinateur, la bio-informatique et le traitement du langage naturel. Il est reconnu pour ses performances supérieures à celles des algorithmes traditionnels d'apprentissage automatique. L'apprentissage profond a été utilisé avec succès dans divers domaines, démontrant ainsi son efficacité. L'un des avantages notables de l'apprentissage profond est sa capacité à traiter efficacement de grandes quantités de données, ce qui le rend bien adapté aux applications qui impliquent le transfert de données importantes sur des réseaux. En outre, les modèles d'apprentissage profond peuvent découvrir et utiliser de nouvelles caractéristiques de manière autonome, ce qui permet de résoudre des problèmes sans intervention humaine explicite. Cet aspect de l'apprentissage profond met en évidence sa capacité à s'adapter et à relever des défis avec un minimum d'intervention manuelle..(Otoum et al., 2022)

I.3.2. L'apprentissage non supervisé

L'apprentissage non supervisé est un type d'apprentissage automatique dans lequel un modèle est formé à travers une base de données non étiquetée et peut reconnaître des modèles sans l'aide d'un superviseur. Les algorithmes d'apprentissage non supervisé peuvent être classés en deux types de problèmes :

Clustering : tels que K-means clustering, K-plus proches voisins et clustering hiérarchique, le rôle du clustering est de diviser automatiquement les données non étiquetées en différents groupes en fonction de la similarité.

Les algorithmes d'association : tels que l'analyse en composantes principales et la décomposition en valeurs singulières sont utilisés pour identifier les corrélations cachées entre différentes variables dans les grandes bases de données.

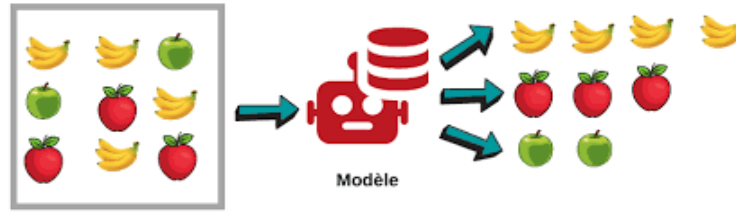


Figure I-8: Schéma d'apprentissage non supervisé.

I.3.2.1. L'algorithme des k plus proches voisins

L'algorithme du plus proche voisin (KNN) est un type d'apprentissage basé sur les instances, qui ne comporte pas de phase d'apprentissage explicite. Au lieu de cela, les données d'entraînement sont stockées pour référence ultérieure. Lorsqu'un nouvel exemple doit être classé, il est comparé aux exemples de la base d'entraînement stockés (appelés prototypes) à l'aide d'une mesure de similarité. Si la phase d'apprentissage est minimale puisque les données sont simplement stockées en mémoire, la classification d'un nouvel exemple peut s'avérer coûteuse en termes de calcul puisqu'elle nécessite de le comparer à tous les exemples précédemment classés.

Dans une approche KNN de base, toutes les variables descriptives d'un exemple sont prises en compte de la même manière lors du calcul de la similarité avec un nouvel exemple. Toutefois, dans les bases de données à haute dimension que l'on trouve couramment dans l'exploration de données, les variables moins pertinentes peuvent avoir un effet préjudiciable sur la précision de la classification. Pour résoudre ce problème, les variables peuvent être pondérées en fonction de leur pertinence. La validation croisée est une technique couramment utilisée pour déterminer ces poids, où les performances de différentes combinaisons de poids sont évaluées (Mathieu-Dupas, 2010).

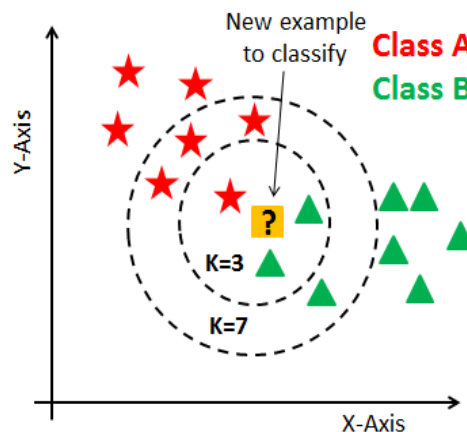


Figure I-9: Exemple de K-plus proche voisins (KPPV).

L'efficacité de la méthode KNN dépend de plusieurs facteurs, notamment le nombre de voisins à prendre en compte, le choix de la métrique de distance et la méthode utilisée pour combiner les réponses des voisins. Le réglage précis de ces paramètres peut avoir un impact significatif sur les performances et la précision de l'algorithme.(Hafida, 2015)

L'approche du K-voisin le plus proche (KNN) est en effet simple et directe. Elle élimine la nécessité d'une phase d'apprentissage traditionnelle en stockant les données d'apprentissage telles quelles. Le processus de classification suit les étapes suivantes :

1. Lorsqu'un nouvel élément de données d'une classe inconnue doit être classé, il est comparé à tous les éléments de données stockés dans l'ensemble de formation.
2. Les K plus proches voisins du nouvel élément de données sont identifiés sur la base d'une mesure de distance choisie, généralement la distance euclidienne.
3. La classe majoritaire parmi les K voisins les plus proches est attribuée au nouvel élément de données en tant que classe prédite (Chamroukhi, 2013)

Pour classer une nouvelle donnée, ses K plus proches voisins sont déterminés sur la base d'une mesure de distance choisie (telle que la distance euclidienne). Par exemple, si nous considérons deux données représentées par les vecteurs x_i et x_j , la distance euclidienne entre elles peut être calculée à l'aide de la formule suivante :

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{k=1}^d (x_{ik} - x_{jk})^2} \quad \text{Équation I-2}$$

I.3.2.2. Le clustering K-moyenne

Développée par MacQueen en 1967, l'approche des K-moyennes (k-means en anglais) est un algorithme populaire d'apprentissage automatique non supervisé largement utilisé pour le regroupement de données. Elle tente de diviser une base de données donnée en k groupes ou clusters, l'analyste prédéterminant la valeur de k . L'idée est de regrouper les éléments en groupe de manière à ce qu'ils soient très similaires les uns aux autres (présentant une forte similarité intra-classe) et dissemblables les uns des autres (présentant une faible similarité inter-classe) entre eux.

Dans l'algorithme de regroupement k-means, chaque groupe est représenté par son centre, appelé centroïde. Le centroïde est calculé comme la moyenne de tous les points de données assignés à ce groupe. Il sert de point représentatif pour la grappe, dont il résume les caractéristiques.

I.3.3. L'apprentissage par renforcement

L'apprentissage automatique qui se base sur les récompenses, il est connu sous le nom d'apprentissage par renforcement, dans lequel un agent interagit avec l'environnement en prenant des mesures et apprend ensuite à mieux exécuter les tâches en fonction des résultats de ces activités.

Afin de recevoir les récompenses les plus positives possibles tout au long du processus d'apprentissage par renforcement, l'agent modifiera ses états en fonction du retour d'information des actions antérieures, puisqu'il sera récompensé pour un bon comportement et pénalisé pour un mauvais.(H. Zhang & Moon, 2021). Il s'agit d'un outil puissant pour construire des modèles d'IA qui peuvent stimuler l'automatisation ou améliorer l'efficacité opérationnelle de systèmes complexes comme les robots, la conduite autonome, la fabrication et la logistique de la chaîne d'approvisionnement, mais il n'est pas recommandé de l'utiliser pour résoudre des problèmes simples ou élémentaires.

I.4. Domaines d'utilisation de l'apprentissage automatique

Ces dernières années, l'apprentissage automatique a été utilisé pour développer des applications dans divers domaines, comme le soulignent Ahuja et al (Ahuja et al., 2020), parmi les domaines notables, on peut citer :

- **Reconnaissance des objets:** Les techniques d'apprentissage automatique sont utilisées pour distinguer et reconnaître des objets sous diverses formes, telles que des visages, des images, des lettres, des mots, des textes avec images, des sons et de la musique.
- **Découverte de médicaments:** L'apprentissage automatique joue un rôle crucial dans la recherche de nouveaux médicaments et de traitements pour des maladies auparavant incurables. Il facilite l'analyse de vastes quantités de données biologiques, l'identification de candidats médicaments potentiels et l'optimisation de la conception des médicaments.
- **Moteurs de recherche, sécurité de l'information et marketing :** Les algorithmes d'apprentissage automatique sont utilisés dans les moteurs de recherche pour améliorer les résultats de recherche et l'expérience des utilisateurs. Ils contribuent également à la sécurité de l'information en détectant les anomalies et les menaces potentielles. Dans le domaine du marketing, l'apprentissage automatique contribue à la segmentation de la clientèle, aux recommandations personnalisées et à la publicité ciblée.
- **Systemes de réponse aux questions :** Les modèles d'apprentissage automatique peuvent être utilisés pour construire des systèmes de réponse aux questions qui fournissent

des réponses basées sur des informations stockées dans des bases de données. Cela permet d'extraire efficacement des informations pertinentes pour répondre aux questions des utilisateurs.

➤ **Prise de décision et recommandations** : Les techniques d'apprentissage automatique, telles que les réseaux neuronaux, simulent les processus de prise de décision du cerveau humain. Cela permet de développer des systèmes capables de prendre des décisions et de faire des recommandations en connaissance de cause, sur la base de modèles de données complexes et des préférences des utilisateurs.

➤ **Robotique**: L'apprentissage automatique contribue au progrès de la robotique, en permettant aux robots d'effectuer un large éventail de tâches, y compris celles qui sont difficiles ou délicates pour les humains. Les algorithmes d'apprentissage automatique aident les robots à comprendre leur environnement et à interagir avec lui, à s'adapter à des conditions changeantes et à effectuer des actions complexes.

Il ne s'agit là que de quelques exemples de l'application de l'apprentissage automatique dans différents domaines. La polyvalence et la puissance des techniques d'apprentissage automatique continuent de stimuler l'innovation et d'améliorer divers aspects de notre vie.

I.5. Conclusion

L'intelligence artificielle est un domaine d'étude dynamique qui progresse rapidement et qui englobe divers domaines de recherche, notamment la vision artificielle où la classification des images joue un rôle important. L'objectif de l'intelligence artificielle est de développer des méthodes de classification qui ressemblent étroitement à l'interprétation humaine, ce qui permet d'obtenir des représentations robustes et fidèles de la scène originale. Le choix des méthodes de classification dépend de l'objectif spécifique, comme la description globale ou locale, ainsi que du type d'objet analysé, comme les caractères, les sons, les visages, les images médicales, etc. Ce chapitre présente une introduction aux principes de l'intelligence artificielle et de l'apprentissage automatique. Dans le chapitre suivant nous présenterons une exploration détaillée des approches basées sur les réseaux de neurones artificiels (ANN) et les convolutifs (CNN). Ces techniques se sont avérées particulièrement efficaces pour relever les défis de la classification des images dans le domaine de l'intelligence artificielle.

Chapitre

2

Les réseaux de neurones artificiels et à convolution

II.1. Introduction

Les réseaux de neurones artificiels et à convolution sont des méthodes d'apprentissage supervisé. Un neurone formel a été créé pour modéliser un neurone biologique et imiter par analogie ses capacités. Néanmoins, ces premiers modèles ne sont même pas capables d'apprendre et il faudra attendre l'apparition du perceptron multicouche. De plus, il faudra attendre plusieurs innovations telles que la technique de la rétro propagation du gradient et les fonctions d'activation. Tous ces techniques sont couplés avec le développement des capacités matériel; cette révolution a permis la réalisation des premières utilisations des réseaux de neurones artificiel.

II.2. Neurone biologique

Le système nerveux humain contient des cellules, appelées neurones. Les neurones sont connectés les uns aux autres à l'aide d'axons et de dendrites, et les régions de connexion entre les axones et les dendrites sont appelées synapses. Ces connexions sont illustrées sur la figure ci-dessous.

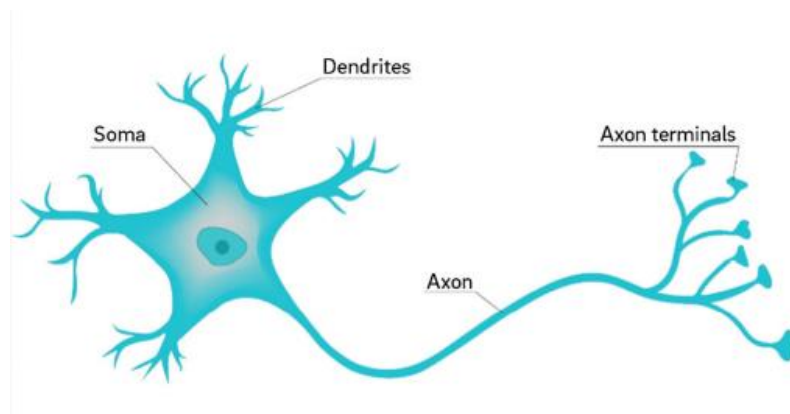


Figure II-1: Neurone biologique.

Le modèle de neurone se compose de trois parties fonctionnelles (Filippov et al., 2020):

- **Des dendrites** : sont des parties d'entrée d'un neurone
- **Un soma** : corps de neurone compare la somme des signaux des dendrites à une certaine valeur seuil.
- **Un axone** : qui correspond aux parties neuronales vivantes c'est est une partie de sortie.

II.3. Neurone Artificiel

Un réseau neuronal typique se compose principalement de nombreux éléments de traitement connectés appelés neurones, chacun produisant une séquence d'activations à valeur réelle pour un résultat cible. La figure ci-dessous montre un diagramme schématique du modèle mathématique d'un neurone artificiel, c'est-à-dire les éléments de traitement, les entrées mises en évidence (X_i), les poids (w), les biais (b), les fonctions de combinaison (Σ), les fonctions d'activation (f) et le signal de sortie correspondant (y).

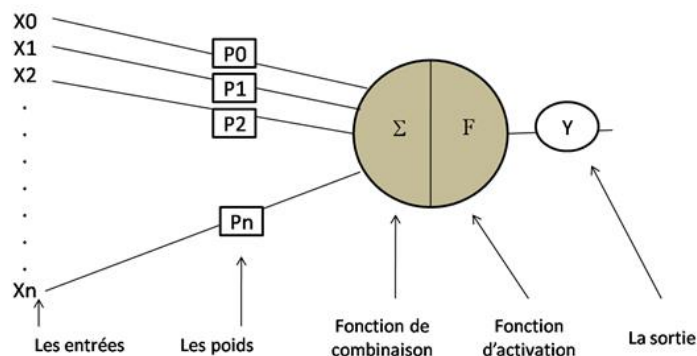


Figure II-2: Représentation schématique d'un neurone artificiel (Sarker, 2021a)

II.4. Architecture des réseaux de neurones

Comme le montre la Figure II-3, un réseau neurone artificiel est constitué des composants de couches et de poids de connexion. Le comportement du réseau dépend également de la communication entre les neurones.

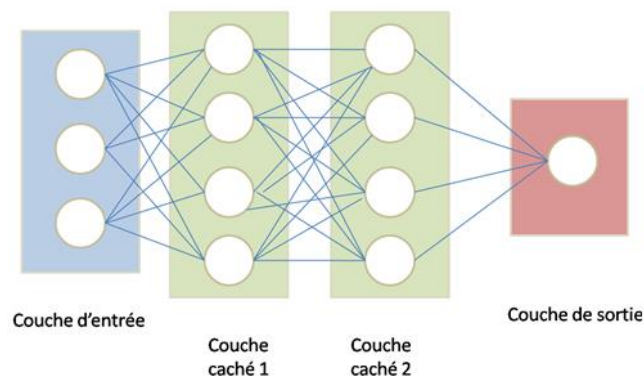


Figure II-3: Réseau neuronal avec 2 couches cachées

Il existe généralement trois types de couches neuronales:

- **La couche d'entrée** : cette couche recevoir les données du programme pour le réseau à partir d'une source externe, cette couche n'a pas de fonctions d'activation (pas de calcul).
- **La couche cachée** : cette couche intermédiaire reçoit les entrées des neurones précédents et les transmet aux neurones suivants, elle effectue tous les calculs et extraient les caractéristiques des données. Il peut y avoir plusieurs couches cachées interconnectées.

➤ **La couche de sortie** : Cette couche fournit les résultats du traitement effectuée par les neurones, il prend l'entrée des couches cachées précédentes et arrive à une prédiction finale basée sur les apprentissages du modèle. (Iglesias et al., 2021)

Les réseaux de neurone se composent d'une seule couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. D'après la Figure II-2, chaque entrée X_i est associée à un poids P_i et la fonction de combinaison va calculer la somme de multiplication du $X_i * P_i$, et enfin appliquer une fonction d'activation pour obtenir la sortie Y .

$$y = \sigma \left(w_0 + \sum_{j=0}^p w_j x_j \right) \quad \text{Équation II-1}$$

II.4.1. Les fonctions d'activations

Une fonction mathématique appelée fonction d'activation convertit un signal d'entrée en un signal de sortie qui est ensuite transféré comme entrée à la couche suivante. Un RNA sans fonction d'activation créera une fonction linéaire simple (régression linéaire), qui n'est qu'un polynôme de 1^{er} degré, dont les performances et la puissance sont limitées, et il ne sera pas en mesure d'effectuer des tâches complexes. C'est pour cela nous utilisons les fonctions d'activation pour rendre le réseau dynamique et non linéaire et lui ajouter la capacité d'extraire des informations complexes à partir d'une base de données. (S. Sharma et al., 2020)

II.4.1.1. La fonction d'étape binaire

C'est la fonction d'activation la plus simple, elle ne peut pas être utilisée en cas de classification multi classe dans la cible transportable. De plus, le gradient de la fonction d'étape binaire est nul, c'est-à-dire que si nous calculons la dérivée de $f(x)$ par rapport à x , elle est égale à zéro. Mathématiquement la fonction d'étape binaire peut être définie comme suit :

$$\begin{cases} f(x) = 1, x \geq 0 \\ f(x) = 0, x < 0 \end{cases} \quad \text{Équation II-2}$$

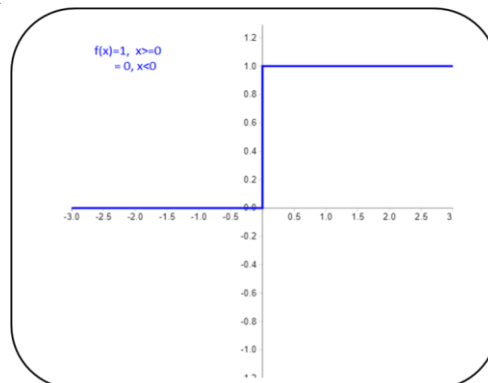


Figure II-4: Fonction d'étape binaire

II.4.1.2. Fonction d'activation linéaire

Il peut être défini par la fonction :

$$f(x) = ax \quad \text{Équation II-3}$$

Avec 'a' une valeur constante. Contrairement à la fonction d'étape binaire ici la fonction $f(x)$ n'est pas nulle mais elle est égale à la valeur de la constante utilisée, le gradient aussi n'est pas nul, mais une valeur constante qui est indépendante de la valeur d'entrée x , ce qui implique que les poids et les biais seront mis à jour lors de l'étape de rétro propagation bien que le facteur de mise à jour soit le même, et le réseau ne pourra pas identifier les modèles complexes à partir des données. Les fonctions linéaires sont idéales pour des tâches simples.

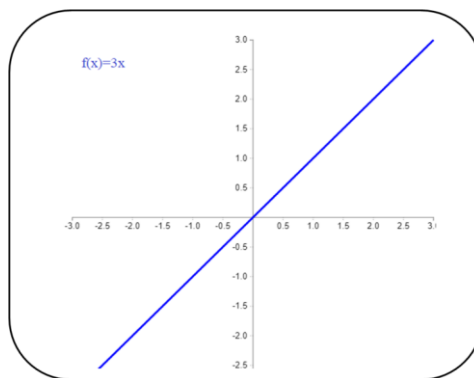


Figure II-5 : Graph de la fonction d'activation linéaire

II.4.1.3. Fonction d'activation sigmoïde

C'est la fonction d'activation la plus utilisée car c'est une fonction non linéaire. La fonction sigmoïde transforme les valeurs entre 0 et 1, et elle n'est pas symétrique. Elle peut être définie comme suit :

$$f(x) = 1/e^{(-x)} \quad \text{Équation II-4}$$

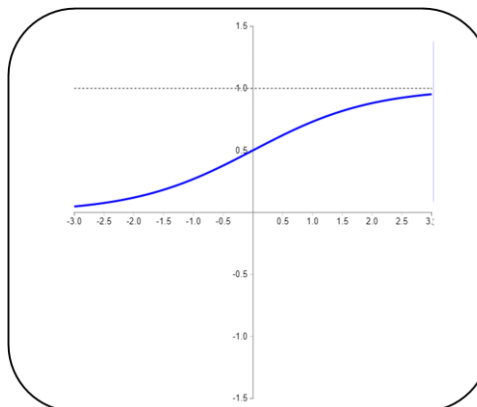


Figure II-6 : Graph de la fonction d'activation Sigmoid

II.4.1.4. Fonction d'activation Tanh

C'est la fonction tangente hyperbolique. La fonction Tanh est similaire à la fonction sigmoïde mais elle est symétrique autour de l'origine. Il peut être défini comme :

$$f(x) = 2\text{sigmoïde}(2x) - 1 \quad \text{Équation II-5}$$

La fonction Tanh est continue et différentiable, les valeurs sont comprises entre -1 et 1.

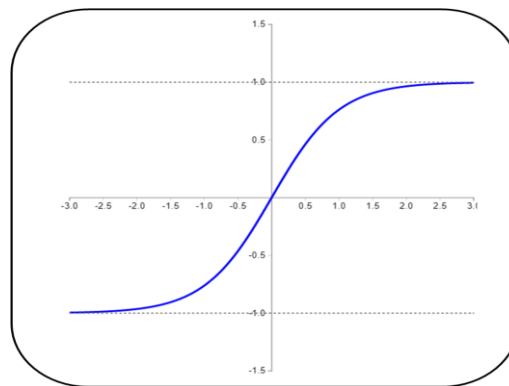


Figure II-7 : Graph de la fonction d'activation Tanh

II.4.1.5. Fonction d'activation ReLU

ReLU signifie (rectified liner unit), elle est largement utilisée dans les réseaux de neurones à convolution. Elle est utilisée parce que tous les neurones ne sont pas activés en même temps. Cela implique qu'un neurone ne sera désactivé que lorsque la sortie de la transformation linéaire est nulle. Il peut être écrit mathématiquement comme :

$$f(x) = \max(0, x) \quad \text{Équation II-6}$$

Les valeurs sont comprises entre 0 et $+\infty$.

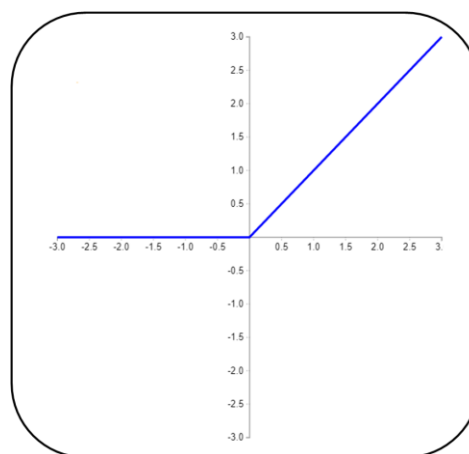


Figure II-8 : Graph de la fonction d'activation ReLU

II.4.1.6. Fonction d'activation Leaky ReLU

Est une version modifiée de la fonction ReLU. Elle définit la valeur de la fonction pour les valeurs négatives de x comme une composante linéaire incroyablement petite de x plutôt que zéro. Les valeurs sont comprises entre $-\infty$ et $+\infty$, elle peut être exprimé mathématiquement par la fonction suivante :

$$\begin{cases} f(x) = x, x \geq 0 \\ f(x) = 0.01x, x < 0 \end{cases} \quad \text{Équation II-7}$$

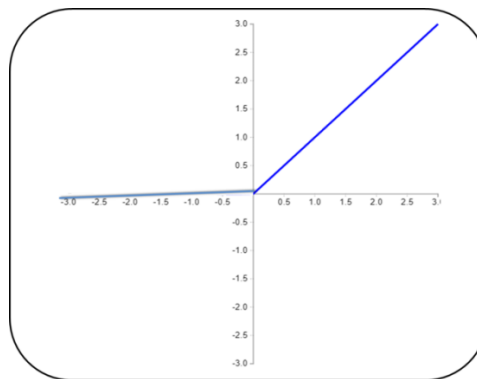


Figure II-9 : Fonction d'activation Leaky ReLU

II.4.1.7. Fonction PReLU : (Paramétric ReLU)

Il s'agit également d'une version quelque peu différente de ReLU, elle est plus performante. En ajoutant un nouveau paramètre à la composante négative de la fonction, elle résout le problème où le gradient de ReLU devient 0 pour les valeurs négatives de x , elle s'exprime comme suit :

$$\begin{cases} f(x) = x, x \geq 0 \\ f(x) = ax, x < 0 \end{cases} \quad \text{Équation II-8}$$

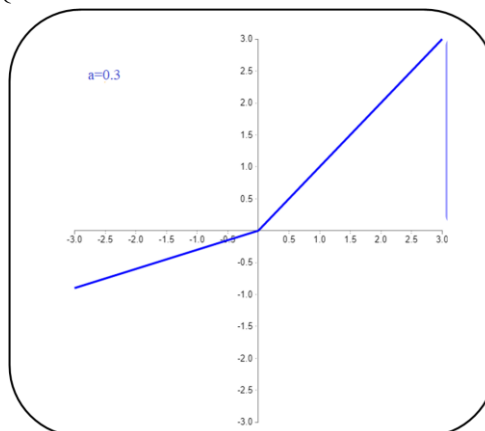


Figure II-10 : Graph de la fonction d'activation PReLU

II.4.1.8. Fonction ELU :

Une autre variation de ReLU est l'unité linéaire exponentielle, ou ELU. Pour les valeurs x négatives, l'ELU ajoute une pente de paramètre. Les valeurs négatives sont définies à l'aide d'une courbe logarithmique. Elle est exprimée comme suit :

$$\begin{cases} f(x) = x, x \geq 0 \\ f(x) = a(e^x - 1), x < 0 \end{cases} \quad \text{Équation II-9}$$

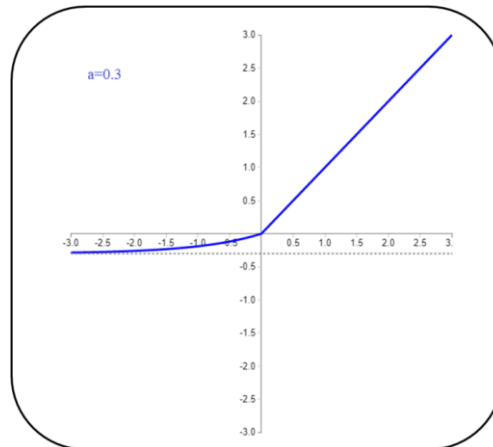


Figure II-11 : Fonction d'activation ELU

II.4.1.9. Fonction d'activation swish

La fonction Swish se distingue par son caractère non monotone, ce qui implique que même si les valeurs d'entrée augmentent, la valeur de la fonction peut baisser. Dans certaines circonstances, la fonction Swish est même plus efficace que la fonction ReLU. Elle s'exprime mathématiquement par :

$$\begin{cases} f(x) = x * \text{sigmoïde}(x) \\ f(x) = x / (1 - (e^{-x})) \end{cases} \quad \text{Équation II-10}$$

II.4.1.10. Fonction d'activation softmax

C'est une combinaison de plusieurs fonctions sigmoïdes. Ces valeurs peuvent être considérées comme des probabilités d'appartenance des points de données à une classe spécifique. Elle est utilisée pour les problèmes de classification multi classe.

L'un des facteurs les plus cruciaux à prendre en compte dans les réseaux neuronaux est la fonction d'activation. Le type de problème et le travail à effectuer déterminent la fonction d'activation à utiliser. La fonction sigmoïde est plus performante pour les problèmes de classification. Tanh et sigmoïde ne doivent pas être utilisées lorsque le gradient est égal à 0. ReLU est utilisé dans les couches cachées et donne de meilleurs résultats que sigmoïde et Tanh. (S. Sharma et al., 2020)

II.4.2. Les algorithmes d'optimisation

L'optimisation est l'un des composants essentiels de l'apprentissage automatique et profond, la plupart des modèles d'apprentissage en profondeur sont basés sur les algorithmes de rétro propagation, qui propagent les erreurs de la couche de sortie vers l'arrière et met à jour les paramètres couche par couche. Les algorithmes d'optimisation permettent de modifier les paramètres d'un réseau de neurone pour minimiser la fonction de perte, il y a plusieurs fonctions de perte qui mesurent les erreurs du modèle ceci quelques exemples :

- Erreur quadratique moyenne (MSE)
- Erreur absolue moyenne (MAE)
- Perte de charnière
- Perte d'entropie croisée
- Perte huber

Les algorithmes d'optimisation sont basés sur la descente de gradient qui joue un rôle important dans le processus d'apprentissage, une série de méthodes d'optimisation ont été proposées pour améliorer les performances et l'efficacité des modèles. (J. Zhang, 2019)

II.4.2.1. La descente de gradient (GD)

Avec l'approche de descente de gradient, les paramètres du réseau sont mis à jour à plusieurs reprises après chaque période d'apprentissage afin de réduire les erreurs. Plus précisément, il doit appliquer une dérivée du premier ordre par rapport aux paramètres du réseau afin de calculer le gradient de la fonction objectif (pente) et mettre à jour les paramètres de manière appropriée.

Le paramètre est ensuite modifié dans la direction opposée du gradient pour minimiser l'erreur (Alzubaidi et al., 2021). L'opération est représentée mathématiquement comme suit :

$$\theta_{ij}^{(\tau)} = \theta_{ij}^{(\tau-1)} - \nabla \theta_{ij}^{(\tau)} \text{ Avec } \nabla \theta^{(\tau)} = \frac{\partial E}{\partial \theta_{ij}} \quad \text{Équation II-11}$$

Différentes variantes de l'algorithme d'apprentissage par gradient sont disponibles et couramment utilisées, notamment les suivantes :

II.4.2.1.1. La descente du gradient par batch

Il est également appelée la descente de gradient de vanille, il optimise de manière itérative les paramètres du modèle à l'aide de l'équation suivantes :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \eta \cdot \nabla_{\theta} L(\theta^{(\tau-1)}, T) \quad \text{Équation II-12}$$

- Le terme θ (τ) désigne le vecteur variable du modèle mis à jour via l'itération τ .
- Le terme η désigne le taux d'apprentissage dans l'algorithme de descente de gradient.

- Le terme $\nabla_{\theta} L(\theta^{(\tau-1)}; T)$ désigne la dérivée de la fonction de perte $L(\theta; T)$ concernant la variable θ basée sur l'ensemble d'apprentissage **complet** T et la valeur vectorielle variable $\theta^{(\tau-1)}$ atteint à l'itération $\tau - 1$

L'inconvénient de cet algorithme est qu'il prend beaucoup de temps pour les grands bases d'apprentissage car il calcule le gradient de la fonction de perte par rapport aux paramètres pour met à jour les paramètres du modèle.

II.4.2.1.2. Descente du gradient stochastique

L'algorithme de descente de gradient stochastique (SGD) mettra à jour les paramètres du modèle avec l'équation suivante de manière itérative :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \eta \cdot \nabla_{\theta} L(\theta^{(\tau-1)}, (x_i, y_i)) \quad \text{Équation II-13}$$

Contrairement à l'algorithme de descente de gradient par batch, SGD met à jour les paramètres du modèle en calculant le gradient de la fonction de perte pour chaque instance et n'utilise pas l'ensemble d'apprentissage complet.

- ✓ SGD est beaucoup plus rapide que la descente de gradient par batch.
- ✓ La SGD peut presque certainement converger vers un optimum. et le processus de mise à jour peut introduire de grandes fluctuations dans la fonction de perte.

II.4.2.1.3. Descente du gradient avec mini-batch

Cet algorithme permet de faire l'équilibre entre les deux algorithmes précédents. Il propose de met à jour les paramètres du modèle avec un mini-batch (mini-lot) d'instances d'entraînement. Soit $B \subset T$ le mini-batch d'instances d'apprentissage échantillonnées à partir de la base d'apprentissage T , l'équation de mise à jour des paramètres du modèle d'apprentissage avec le mini-batch est définie par l'équation suivante :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \eta \cdot \nabla_{\theta} L(\theta^{(\tau-1)}, B) \quad \text{Équation II-14}$$

Tel que $L(\theta; B)$ désigne le terme de perte introduit par le modèle sur le mini-batch B . Les mini-batch sont échantillonnés séquentiellement à partir de la base d'apprentissage T et leur taille ne doit pas être trop grande. La taille du batch B est fournie en tant que paramètre.

Par rapport à GD par batch, cet algorithme est beaucoup plus efficace, en particulier pour les bases d'apprentissage volumineux. Par rapport à la descente de gradient stochastique, cet algorithme réalise une convergence beaucoup plus stable. Bien que ces trois algorithmes convergent bien vers des optima locaux ou globaux, ils ont aussi quelques problèmes :

- Taux d'apprentissage : le taux d'apprentissage η a une grande influence sur les algorithmes de descente de gradient :

Taux d'apprentissage élevé → diverger le processus d'apprentissage

Taux d'apprentissage faible → rend la convergence trop lente

- Ajustement du taux d'apprentissage : pour atteindre à un bon optimum dans les étapes initiales on utilise un taux d'apprentissage plus élevé, tandis que dans les étapes suivantes on peut ajuster le taux d'apprentissage avec une valeur petite.
- ❖ Taux d'apprentissage individuels des paramètres: si les paramètres sont différents, leur mise à jour peut nécessiter des taux d'apprentissage différents dans le processus de mise à jour.
- ❖ Éviter le point de selle: Si le gradient est nul dans toutes les dimensions, cela s'appelle un point de selle, cependant le point de selle peut être un minima ou maxima local

Pour résoudre les problèmes précédents, plusieurs algorithmes sont proposés :

II.4.2.2. Momentum

Elle est proposée pour accélérer la convergence de mise à jour des paramètres.

Momentum met à jour les paramètres avec l'équation suivante :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \eta \cdot \Delta_v^{(\tau)} \quad \text{Équation II-15}$$

Avec $\Delta_v^{(\tau)} = \rho \cdot \Delta_v^{(\tau-1)} + (1 - \rho) \cdot \nabla_{\theta} L(\theta^{(\tau-1)})$

$v(\tau)$ désigne le terme d'impulsion introduit pour enregistrer les gradients historiques jusqu'à l'itération τ . Paramètre $\rho \in [0, 1]$ désigne le poids du terme de quantité de mouvement

II.4.2.3. Gradient accéléré de Nesterov

Quant aux algorithmes de descente de gradient actuellement proposés, ils mettent tous les paramètres à jour en fonction du gradient du point historique ou du point courant sans connaître le futur. On ne sait pas quel point futur ils atteindront au cours du processus d'apprentissage, soit sur le gradient de l'histoire, soit sur le point actuel. Les points qu'ils atteindront au cours de leurs études. Cela rend le processus d'apprentissage aveugle et les performances d'apprentissage hautement imprévisibles.

La méthode du gradient accéléré (NAG) de Nesterov propose de résoudre ce problème en mettant à jour les paramètres avec des gradients à des points futurs approximatifs :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \eta \cdot \Delta_v^{(\tau)} \quad \text{Équation II-16}$$

Avec $\widehat{\theta}^{(\tau)} = \theta^{(\tau-1)} - \eta \cdot \rho \cdot \Delta_v^{(\tau-1)}$, et $\Delta_v^{(\tau)} = \rho \cdot \Delta_v^{(\tau-1)} + (1 - \rho) \cdot \nabla_{\theta} L(\widehat{\theta}^{(\tau)})$

L'algorithme NAG est capable de mettre à jour les paramètres du modèle de manière beaucoup plus efficace.

II.4.2.4. Adagrad

Comme pour les algorithmes d'apprentissage introduits précédemment, le taux d'apprentissage dans ces méthodes sont pour la plupart fixes et identiques pour tous les paramètres du vecteur θ . Cependant, dans le processus de mise à jour des paramètres, le taux d'apprentissage requis peut être différent pour différentes paramètres. Pour les paramètres qui atteignent la valeur optimale, un petit taux d'apprentissage est requis, pour les paramètres qui sont loin de la valeur optimale, nous pouvons avoir besoin d'utiliser un taux d'apprentissage relativement important afin d'atteindre la valeur optimale plus rapidement. Pour résoudre ces deux problèmes, un nouvel algorithme de gradient adaptatif a été introduit, appelé Adagrad. L'équation d'apprentissage d'Adagrad peut être exprimée par:

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \eta \frac{\eta}{\sqrt{\text{diag}(G^{(\tau)} + \varepsilon \cdot I)}} \text{ avec } \begin{cases} g^{(\tau-1)} = \nabla_{\theta} L(\theta^{(\tau-1)}) \\ G^{(\tau)} = \sum_{t=0}^{\tau-1} g^{(t)} \cdot (g^{(t)})^T \end{cases} \quad \text{Équation II-17}$$

$\text{diag}(G)$ définit une matrice diagonale de mêmes dimensions que G mais avec seulement les éléments sur la diagonale de G . $G(\tau)$ la matrice qui conserve les enregistrements des gradients historiques calculés depuis le début jusqu'à l'itération actuelle.

Dans Adagrad l'apprentissage permet à la fois un taux d'apprentissage adaptatif qui n'a pas besoin d'un réglage manuel, et un taux d'apprentissage différent pour différentes paramètres. Dans Adagrad les taux d'apprentissage continueront à diminuer et les paramètres ne peuvent plus être mises à jour avec les données d'apprentissage

II.4.2.5. RMSprop

RMSprop introduit pour résoudre le problème avec un taux d'apprentissage décroissant dans Adagrad. RMSprop décompose correctement le poids du gradient accumulé historique dans la matrice G définie dans Adagrad, et peut également ajuster le taux d'apprentissage dans le processus de mise à jour. Les équations de mise à jour des paramètres dans RMSprop peuvent être représentées par les équations suivantes :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \frac{\eta}{\sqrt{\text{diag}(G^{(\tau)} + \varepsilon \cdot I)}} \cdot g^{(\tau-1)} \quad \text{Équation II-18}$$

$$\text{avec } \begin{cases} g^{(\tau-1)} = \nabla_{\theta} L(\theta^{(\tau-1)}) \\ G^{(\tau)} = \rho \cdot G^{(\tau-1)} + (1 - \rho) \cdot g^{(\tau-1)} \cdot (g^{(\tau-1)})^T \end{cases}$$

Le paramètre ρ désigne le poids des gradients historiquement accumulés, ρ est généralement fixé à 0,9. Dans l'équation ci-dessus, le terme dénominateur calculé par :

$$\text{RMS}(g^{(\tau-1)}) = \sqrt{\text{diag}(G^{(\tau)} + \varepsilon \cdot I)} \quad \text{Équation II-19}$$

Par conséquent, l'équation de mise à jour s'écrit généralement comme suit :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \frac{\eta}{RMS(g^{(\tau-1)})} \cdot g^{(\tau-1)} \quad \text{Équation II-20}$$

II.4.2.6. Adadelta

Semblable à RMSprop, Adadelta se rapproche du dénominateur dans l'équation ci-dessus en tant que RMS du gradient précédent. Adadelta a proposé d'approximer le numérateur dans la formule ci-dessus et de remplacer $\Delta\theta$ par $RMS(\Delta\theta)$ de la même manière. Formellement, l'équation de mise à jour des paramètres dans Adadelta donnée par :

$$\theta^{(\tau)} = \theta^{(\tau-1)} - \frac{RMS(\Delta\theta^{(\tau-1)})}{RMS(g^{(\tau-1)})} \cdot g^{(\tau-1)} \quad \text{Équation II-21}$$

$RMS(\Delta\theta^{(\tau-1)})$ conserve les enregistrements des $\Delta\theta$ dans les itérations précédentes jusqu'à l'itération précédente $\tau - 1$.

Adadelta n'utilise pas le taux d'apprentissage dans l'équation de mise à jour, ce qui corrige efficacement deux faiblesses d'Adagrad précédemment introduites.

II.4.2.7. Adam

Adam peut être considéré comme une intégration de l'algorithme RMSprop et de l'algorithme Momentum. Adam enregistre le carré du gradient du premier ordre passé, qui décroît de façon exponentielle au cours du processus d'apprentissage.

$$\begin{aligned} m^{(\tau)} &= \beta_1 \cdot m^{(\tau-1)} + (1 - \beta_1) \cdot g^{(\tau-1)}, \\ v^{(\tau)} &= \beta_2 \cdot v^{(\tau-1)} + (1 - \beta_2) \cdot g^{(\tau-1)} \circ g^{(\tau-1)} \end{aligned} \quad \text{Équation II-22}$$

$m^{(\tau)}$ et $v^{(\tau)}$ sont les vecteurs qui désignent les termes stockant respectivement les gradients de premier ordre et les gradients de premier ordre au carré.

Lorsque β_1 et β_2 sont proches de 1, les vecteurs $m^{(\tau)}$ et $v^{(\tau)}$ sont biaisés vers zéro. Pour résoudre ce problème, Adam propose de redimensionner les termes comme suit :

$$\begin{aligned} \hat{m}^{(\tau)} &= \frac{m^{(\tau)}}{1 - \beta_1^{(\tau)}} \\ \hat{v}^{(\tau)} &= \frac{v^{(\tau)}}{1 - \beta_2^{(\tau)}} \end{aligned} \quad \text{Équation II-23}$$

En se basant sur le vecteur redimensionné $\hat{m}^{(\tau)}$ et la matrice $\hat{v}^{(\tau)}$, Adam mettra à jour la variable de modèle avec l'équation suivante :

$$\hat{\theta}^{(\tau)} = \theta^{(\tau-1)} - \frac{\eta}{\sqrt{\hat{v}^{(\tau)} + \epsilon}} \Theta \hat{m}^{(\tau)} \quad \text{Équation II-24}$$

Elle permet simultanément une convergence plus rapide et un taux d'apprentissage adaptatif.

II.4.2.8. Nadam

L'intégration des 2 algorithmes Adam et Gradient accéléré de Nesterov s'appelle Nadam (J. Zhang, 2019). Cet algorithme mettra à jour la variable de modèle avec l'équation:

$$\hat{\theta}^{(\tau)} = \theta^{(\tau-1)} - \frac{\eta}{\sqrt{\hat{v}^{(\tau)} + \varepsilon}} \Theta(\beta_1 \hat{m}^{(\tau)} + \frac{(1-\beta_1) \cdot g^{(\tau)}}{1-\beta_1^{(\tau)}}) \quad \text{Équation II-25}$$

II.5. Apprentissage profond (Deep learning ou DL)

En 2006, l'apprentissage profond a été introduit, il est basé sur le concept de réseau de neurones artificiel (ANN), L'apprentissage en profondeur est devenu une technique d'actualité dans tous les domaines en vision par ordinateur, entraînant une renaissance de la recherche sur les réseaux de neurones. (Sarker, 2021a)

On applique DL quand les êtres humains sont incapable de prendre des décision à l'aide de leurs expertise (compréhension de la langue, décisions médicales et reconnaissance de la parole), aussi dans les cas où il y a la notion de temps (prévision des prix, préférence des actions, prévision météorologique et suivi), cas où la taille du problème est extrêmement importante et dépasse nos capacités de raisonnement inadéquates (analyse des sentiments, correspondance des publicités avec Facebook...etc.).

II.5.1. Les réseaux de neurones convolutifs

Le réseau de neurones à convolution (convolutional neural network ou CNN) également connu sous les noms (ConvNet ou apprentissage en profondeur), c'est technique qui se compose de plusieurs couches. Chaque couche de convolution est connue sous le nom de carte de caractéristiques. Les réseaux de neurones à convolution sont le plus populaire algorithme du Deep Learning, ils apprennent directement à partir de l'entrée sans avoir besoin d'extraction de caractéristiques humaines. La structure des CNN a été inspirée par les neurones des cerveaux humains et animaux. Les CNN utilisent très peu de paramètres, ce qui simplifie le processus d'apprentissage et accélère le réseau.

II.5.1.1. Les couches d'un CNN

CNN à trois couches principales : couche de convolutions, couches de regroupement (pooling) et couches entièrement connectées (fully connected). Une architecture typique consiste en des itérations d'empilements de plusieurs couches de convolution et de pooling, suivies d'une ou plusieurs couches entièrement connectées. Ces couches sont introduites comme suit :

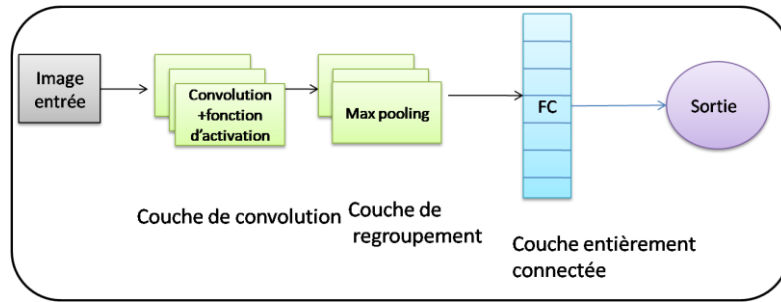


Figure II–12: Une architecture typique d'un CNN (Alzubaidi et al., 2021)

II.5.1.1.1. Couche de convolution

Cette couche distingue les CNN des autres réseaux de neurones et effectue une extraction de caractéristiques consistant à travers les opérations de convolution et des fonctions d'activation. La convolution est une opération linéaire qui applique un tableau de nombres appelé noyau à un tableau d'entrée appelé tenseur. Le produit élément par élément entre chaque élément du noyau et le tenseur est calculé à chaque position du tenseur et additionné pour obtenir la valeur de sortie à la position correspondante du tenseur de sortie, appelée la carte de caractéristique. La convolution est répétée plusieurs fois sur plusieurs noyaux et possède deux hypermètres : la taille du noyau (3×3, 5×5 ou 7×7) et le nombre de noyaux qui déterminent la profondeur des cartes de caractéristiques de sortie.

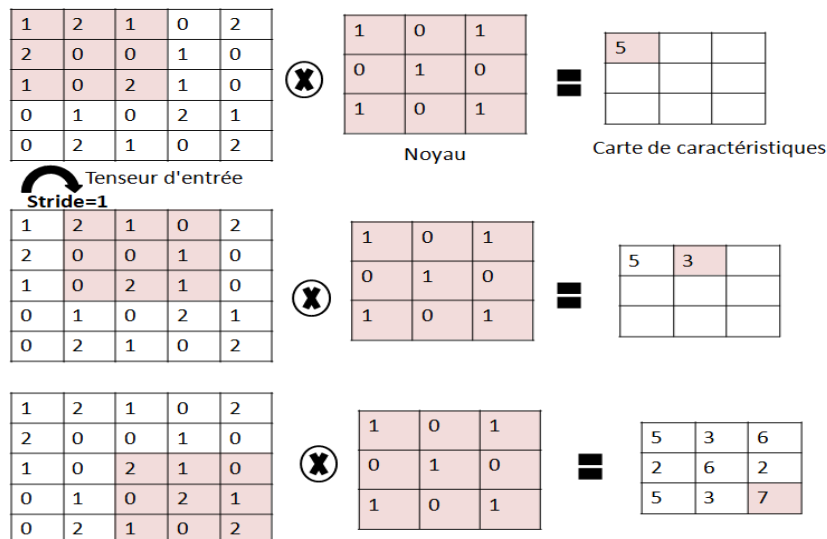


Figure II–13: Exemple d'opération de convolution avec un noyau de taille 3×3 sans padding

La distance entre deux positions consécutives d'un noyau s'appelle stride, et le choix habituel pour le stride est 1.

Padding, est une technique qui consiste à ajouter des lignes et des colonnes de zéros de chaque côté du tenseur d'entrée, de manière à faire tenir le centre d'un noyau sur l'élément le plus extérieur et à conserver la même dimension dans le plan tout au long de l'opération de convolution comme illustré la figure suivante :

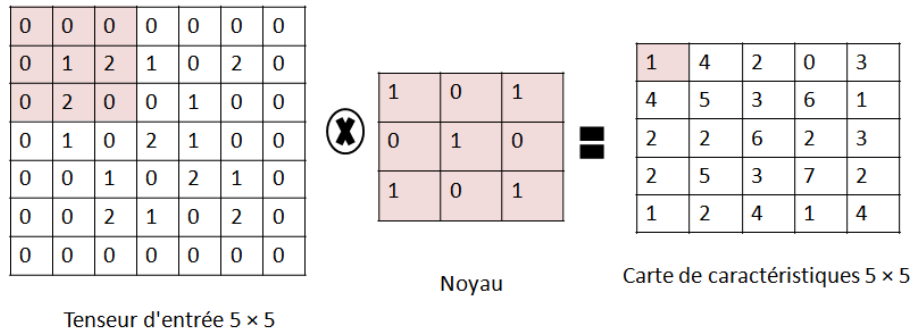


Figure II–14: Exemple d'opération de convolution avec padding, un tenseur d'entrée de taille 5×5, noyau de 3×3 et un stride de 1. (Yamashita et al., 2018)

La sortie de la convolution est ensuite passée à travers une fonction d'activation non linéaire, la fonction d'activation la plus couramment utilisée aujourd'hui étant ReLU.

II.5.1.1.2. Couche de pooling

La tâche principale de la couche de pooling est le sous-échantillonnage des cartes de caractéristiques. Le pooling réduit les cartes de caractéristiques de grandes taille pour créer des cartes plus petites, en conservant la majorité des informations (ou caractéristiques) dominantes à chaque étape. La taille du filtre, stride et le padding sont des hyper paramètres dans les opérations de pooling, similaires aux opérations de convolution. Plusieurs types de méthodes de pooling sont disponibles : tree pooling, gated pooling, average pooling, min pooling, max pooling, global average pooling (GAP), et global max pooling.

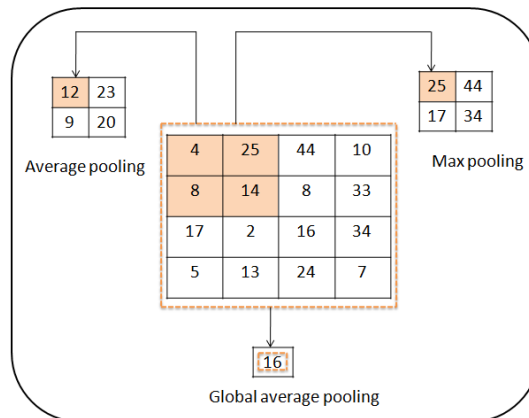


Figure II–15: Types d'opération de pooling .(Alzubaidi et al., 2021)

II.5.1.1.3. La couche fully connected (FC)

Cette couche est la dernière couche du CNN, elle est appelée fully connected car chaque neurone est connecté à chaque neurone de la couche précédente. Une couche fully connected prend la carte de caractéristique de la sortie et la convertit en un vecteur. Chaque couche fully connected est suivie d'une fonction d'activation. (Yamashita et al., 2018)

II.6. Quelques Architecture basées CNN

Plusieurs architectures de réseaux ont été utilisées dans la littérature pour les tâches de classification d'images. VGG16, Inception Net et Resnet sont parmi les plus notables.

II.6.1. VGG

Visual Geometry Group (VGG) est proposé par Simonyan et Zisserman, est un modèle qui utilise les filtres 3×3 plutôt que 5×5 ou 7×7 pour montrer que les filtres de petits taille produire la même influence que les filtres de grande taille. En général, VGG a obtenu des résultats significatifs pour les problèmes de localisation et de classification. (Alzubaidi et al., 2021). Il y a plusieurs variation de VGG dépend du nombre du couches : VGG16, VGG19...

VGG16 est l'architecture la plus couramment utilisée dans les CNN et comporte 16 couches. La principale contribution de l'architecture VGG16 est l'introduction d'une topologie homogène et d'une petite taille de noyau ($2 \times 2 \times 3$). VGG16 est configuré avec 138 millions de paramètres et occupe environ 500 Mo de mémoire. L'architecture de base de VGG16 est illustrée à la Figure II–16.

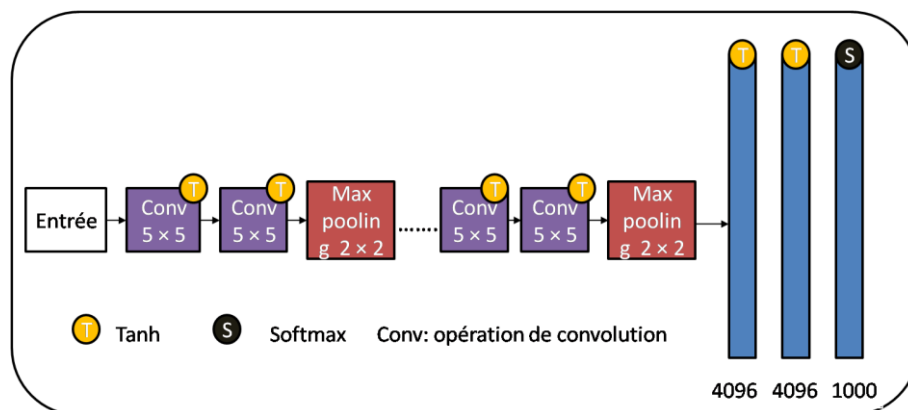


Figure II–16: L'architecture VGG16.(Kim et al., 2021)

II.6.2. GoogLeNet

L'architecture GoogLeNet, également connue sous le nom de Inception-V1, est une architecture CNN qui propose un nouveau concept de blocs de démarrage (modules) combinant des transformations convolutionnelles multi-échelles en utilisant des fonctions de fusion, de transformation et de division pour l'extraction de caractéristiques. Dans cette architecture, les couches convolutives sont remplacées par des petits blocs et les couches FC par des GAP. Il a utilisé des connexions clairsemées pour surmonter le problème des informations redondantes. Il utilise environ 5 millions de paramètres. GoogLeNet améliore l'efficacité des paramètres et de l'apprentissage.

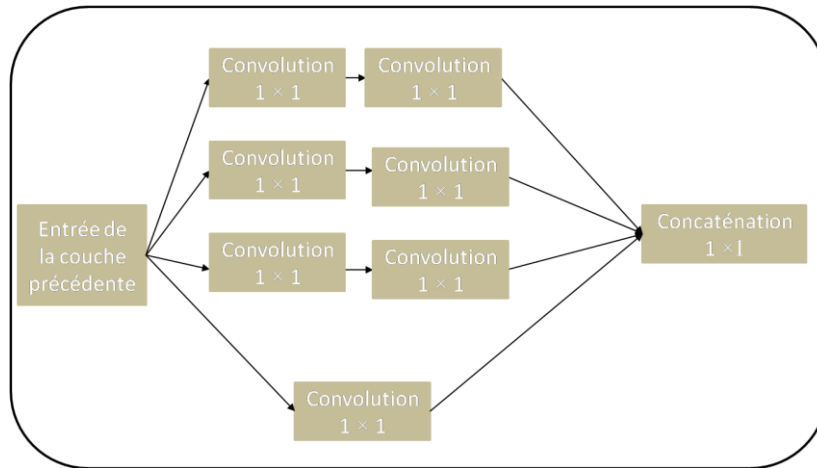


Figure II–17: L' architecture GoogLeNet.(Alzubaidi et al., 2021)

II.6.3. ResNet

ResNet utilise jusqu'à 152 couches, qui peuvent être très profondes Il se compose de 25 millions de paramètres .ResNet prend en charge des centaines de couches convolutives et il est l'un des premiers qui utilisent la normalisation par lots. Les blocs de construction de base de ResNet sont les blocs de convolution et d'identité. Le bloc convolutif repose sur la fonction d'activation ReLU et la normalisation des lots.

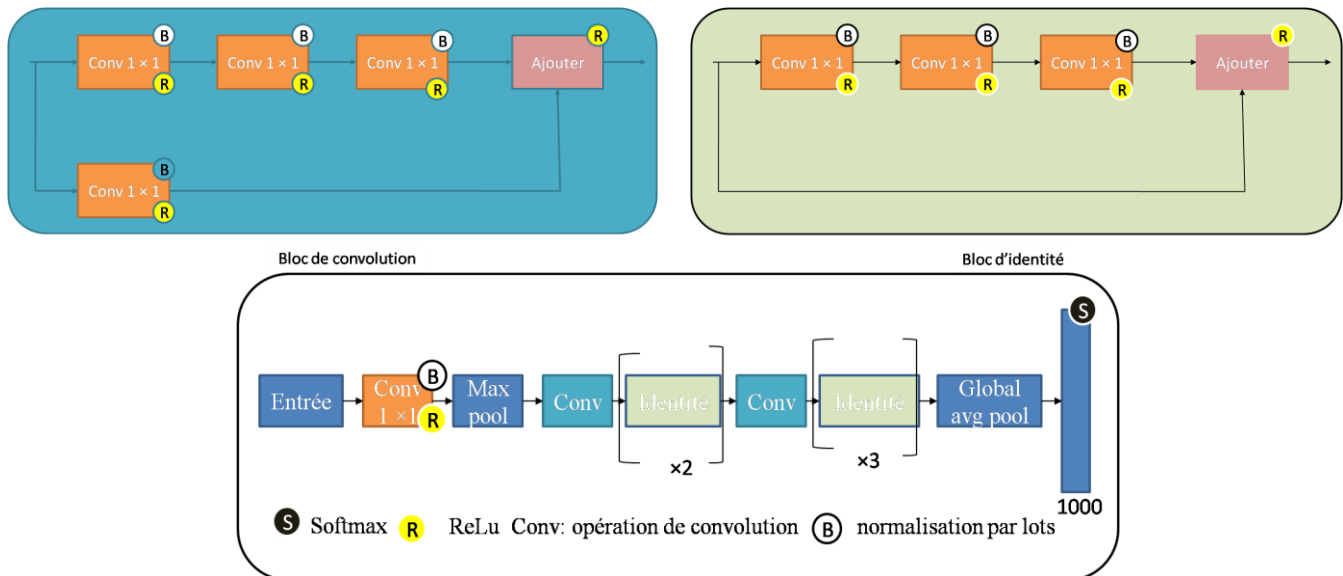


Figure II–18: L'architecture ResNet.(Kim et al., 2021)

II.6.4. EfficientNet

EfficientNet est une collection de modèles CNN qui ont été développés pour atteindre une grande précision tout en réduisant la puissance informatique nécessaire à l'apprentissage et à l'exécution. Avec les millions de photos contenues dans la base de données ImageNet, ces modèles sont particulièrement utiles pour les problèmes de classification d'images.

En fonction de leur niveau de complexité et de précision, les modèles EfficientNet sont classés de B0 à B7. Bien que la précision du modèle augmente avec le nombre de modèles, il n'y a pas de relation linéaire entre le nombre de paramètres nécessaires pour l'inférence et l'apprentissage. Par conséquent, la précision et la charge de calcul sont mieux gérées par les modèles les plus complexes. L'utilisation d'une toute nouvelle fonction d'activation appelée Swish est l'une des caractéristiques distinctives d'EfficientNet. Il a été démontré que cette fonction d'activation est à la fois plus précise et plus efficace en termes de calcul que la fonction d'activation ReLU (Rectifier Linear Unit) souvent utilisée. Swish permet à EfficientNet d'atteindre une excellente précision avec un minimum de ressources informatiques. (Atila et al., 2021)

EfficientNet est une architecture d'apprentissage profond qui réduit la taille du modèle tout en augmentant de manière cohérente la profondeur, la largeur et la résolution afin de produire des résultats plus efficaces avec moins de modèles.

Le nombre de couches du réseau détermine sa profondeur, tandis que le nombre de filtres dans chaque couche convolutive détermine sa largeur. La hauteur et la largeur de l'image d'entrée définissent la résolution. Ces trois éléments sont optimisés par EfficientNet en fonction d'un coefficient d'échelle, abrégé en ϕ , afin de maximiser l'efficacité. Les auteurs suggèrent d'utiliser ce coefficient pour construire une formule de mise à l'échelle pour la profondeur, la largeur et la résolution. EfficientNet atteint des résultats de pointe tout en utilisant moins de paramètres que d'autres modèles, car il met ces facteurs à l'échelle de manière cohérente. Les équations de ces éléments sont :

$$\begin{aligned}d &= \alpha^\phi \\w &= \beta^\phi \\r &= \gamma^\phi\end{aligned}\tag{Equation II-26}$$

De telle sorte que : $\alpha \geq 1, \beta \geq 1, \gamma \geq 1$ et $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$

Où d , w et r représentent respectivement la profondeur, la largeur et la résolution du réseau, et les termes constants α , β et γ ont été déterminés à l'aide de la technique de réglage de l'hyper paramètre par recherche sur la grille. Le coefficient est une variable définie par l'utilisateur qui régit toutes les ressources de mise à l'échelle du modèle.

II.6.5. EfficientNet-B0 :

EfficientNet-B0 a surpassé d'autres modèles de pointe formés sur la base de données ImageNet, contrairement à d'autres CNN profonds, qui ont ajusté chaque dimension à l'aide d'un ensemble prédéfini de coefficients d'échelle.

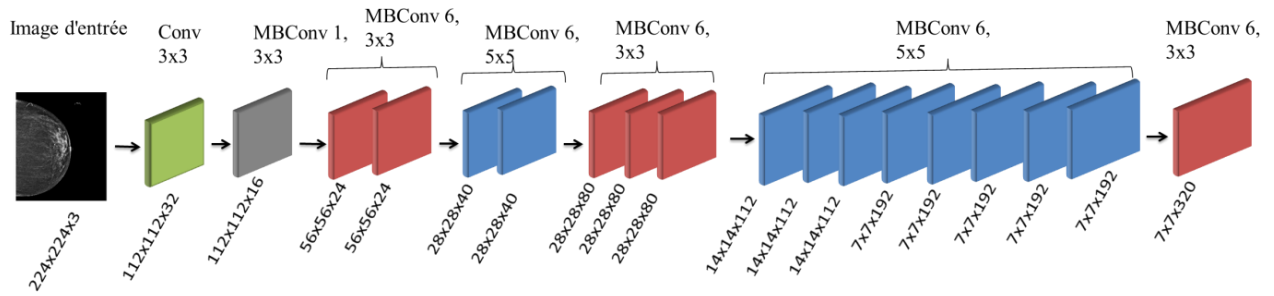


Figure II–19 : L'Architecture de base du modèle EfficientNet-B0. (Shah et al., 2022)

Le modèle EfficientNet-B0 utilise de nombreuses couches de convolution avec un champ réceptif de 3x3 et la couche de convolution mobile à goulot d'étranglement inversé (MBconv) pour capturer les caractéristiques à travers les couches. Il accepte une image d'entrée de 224x224x3. La Figure II–19 présente le dernier modèle de base EfficientNet-B0.

Le principal composant d'EfficientNet, introduit pour la première fois dans MobileNetV2, est le goulot d'étranglement inversé MBConv. Néanmoins, en raison de son budget FLOPS (opérations en virgule flottante par seconde) plus élevé, il est utilisé un peu plus fréquemment que MobileNetV2. Des connexions directes sont utilisées entre les goulots d'étranglement qui connectent beaucoup moins de canaux que les couches d'expansion dans les blocs MBConv, qui sont constitués d'une couche qui étend et comprime les canaux dans l'ordre inverse. Par rapport aux couches conventionnelles, les convolutions séparables en profondeur de cette architecture réduisent presque les calculs du facteur k^2 (k est la taille du noyau, qui indique la largeur et la hauteur de la fenêtre de convolution 2D). (Atila et al., 2021)

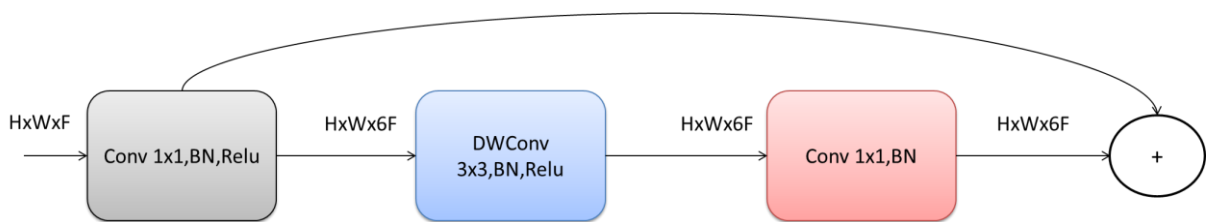


Figure II–20: Bloc MBConv

Cette figure représente l'architecture du bloc MBConv qui contient : DWConv signifie convolution en profondeur, (3x3/5x5 définit la taille du noyau), BN est la norme du lot, HxWxF désigne la forme du tenseur (hauteur, largeur, profondeur). (Diallo & Ju, 2020)

II.7. Apprentissage par transfert

L'apprentissage par transfert est une technique d'apprentissage automatique dans laquelle un modèle formé sur une tâche ou un domaine (domaine source) est affiné sur une tâche ou un domaine différent mais apparenté (domaine cible) afin d'améliorer ses performances comme

le montré la Figure II–21. Pour réduire la quantité de données étiquetées nécessaires à l'apprentissage et augmenter la généralisation, l'objectif est d'appliquer les informations et les représentations apprises dans le domaine source au domaine cible. Pour que l'apprentissage par transfert réussisse, les deux domaines doivent être liés. La probabilité d'un transfert de connaissances réussi augmente avec le degré de similitude entre les deux domaines. La transmission d'informations importantes susceptibles d'améliorer les performances du modèle dans le domaine cible peut être possible même si les domaines ne sont pas similaires.

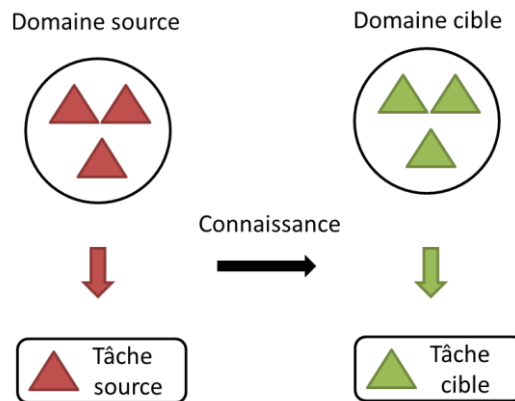


Figure II–21 : Apprentissage par transfert(Zhuang et al., 2021)

II.8. Conclusion

Dans ce chapitre, nous avons présenté les concepts fondamentaux des réseaux ed neurones et de l'apprentissage profond. Nous avons commencé par présenter le neurone artificiel et biologique. Nous avons présenté les fonctions d'activation telles que Sigmoid, tanh, ReLU et ReLU ...etc., qui jouent un rôle crucial dans le processus de classification. Nous avons également abordé les algorithmes d'optimisation, notamment la descente de gradient, la descente de gradient stochastique élastique, ADAGRAD et ADAM...etc., qui permettent d'ajuster les poids du réseau neuronal au cours du processus de formation. Nous nous sommes ensuite concentrés sur les CNN. Nous avons expliqué le concept de l'opération de convolution, qui est un élément clé des réseaux neuronaux à convolution pour l'extraction de caractéristiques pertinentes. Nous avons exploré d'autres concepts tels que le padding, le pooling et la couche entièrement connectée, qui contribuent à l'efficacité des CNN.

Enfin, nous avons mis en évidence les applications pratiques des CNN, telles que VGG16, Inception et ResNet. Ces modèles ont démontré la puissance et le potentiel des CNN dans les applications du monde réel.

Chapitre

3

Extraction des caractéristiques

III.1. Introduction

Sur la base de ce qu'ils observent et de leurs connaissances préalables, les humains sont capables de déduire une histoire à partir d'une image. De plus, à l'aide des méthodes d'extraction des caractéristiques les programmes deviennent capables d'extraire des idées sémantiques à partir de cette image. Néanmoins, plutôt que d'utiliser les connaissances préalables de l'homme, l'étape initiale d'un programme de compréhension sémantique consiste à extraire les caractéristiques pour développer des modèles de classification efficaces.

Le traitement des images peut être effectué en extrayant des caractéristiques pour l'identification, la classification, le diagnostic, le regroupement, la reconnaissance et la détection. La méthode d'extraction de caractéristiques est utilisée pour extraire autant de données de l'image que possible. Actuellement, une difficulté importante est de savoir comment choisir les caractéristiques et les extraire efficacement.

III.2. Extraction des caractéristiques linéaires et non linéaires

Une transformation linéaire peut toujours être exprimée comme un vecteur dans l'espace d'entrée multiplié par un point de données, et le vecteur de sortie correspond au point de données correspondant dans le sous-espace des caractéristiques. Cela signifie que décrire une méthode d'extraction de caractéristiques comme étant linéaire revient à affirmer que les caractéristiques récupérées dans l'espace d'entrée sont des lignes droites.

Afin d'extraire les caractéristiques des données d'entrée, les techniques d'extraction de caractéristiques linéaires telles que l'analyse en composantes principales (ACP) et l'analyse discriminante linéaire (ADL) utilisent une transformation linéaire. Alors que l'ADL détermine les directions qui augmentent la séparation entre les différentes classes dans une tâche de classification, l'ACP localise les directions de variance maximale pour projeter les données d'entrée dans un espace de dimension inférieure. Lorsque la structure sous-jacente des données peut être décrite par une combinaison linéaire des caractéristiques d'entrée, les méthodes d'extraction de caractéristiques linéaires sont avantageuses.

Le terme non linéaire décrit un type spécifique de transformation de points de données (projection) d'un espace d'entrée de haute dimension vers un sous-espace de caractéristiques de dimension inférieure. L'ACP à noyau est une illustration de l'extraction de caractéristiques non linéaires. Contrairement à l'ACP linéaire, la projection des points de données de l'espace d'entrée vers l'espace des caractéristiques ne peut pas être exprimée par une multiplication matricielle. Il est également important de garder à l'esprit que les techniques d'extraction de

caractéristiques non linéaires peuvent parfois nécessiter davantage de ressources informatiques que les techniques linéaires, car elles peuvent impliquer des calculs mathématiques plus complexes. Elles peuvent cependant être extrêmement efficaces pour capturer les corrélations complexes entre les caractéristiques et améliorer la fonctionnalité des modèles d'apprentissage automatique. (Liu, 2013)

III.3. Types d'extraction de caractéristiques

En général, les caractéristiques peuvent être divisées en caractéristiques générales (CG) et en caractéristiques spécifiques à un domaine (CFD). La couleur, la forme et la texture sont des exemples de CG qui sont indépendants de l'application, tandis que les aspects conceptuels et les caractéristiques du visage humain sont des exemples de CFD qui dépendent de l'application. Dans cette partie nous présentons les types de caractéristiques couramment utilisés dans les applications de traitement d'images et de vision par ordinateur

III.3.1. Caractéristiques de forme

Pour détecter et identifier des objets dans le monde réel, les caractéristiques de forme (CF) sont cruciales. Elles servent d'indice visuel principal pour la comparaison et la mise en correspondance des objets. Les caractéristiques de forme sont divisées en deux catégories : les caractéristiques basées sur les régions (BR) et les caractéristiques basées sur les contours (BC). Contrairement à la BC, qui extrait les CF de la frontière, la BR extrait les caractéristiques de l'objet complet. Elle est utilisée dans les domaines de la vision par ordinateur, du traitement des images et de la reconnaissance des formes. (Salau & Jain, 2019)

Sans forme, le contenu visuel ne permet pas d'identifier efficacement un objet et sans la reconnaissance de la forme, l'image est incomplète. Bien que les formes des deux objets ne puissent pas être exactement les mêmes, nous pouvons facilement identifier des formes comparables en utilisant une variété de méthodes.

Le principal problème de la reconnaissance des formes est que les différents endroits où un objet est placé affichent des formes différentes, ce qui rend difficile la détermination de la forme réelle de l'objet. Les méthodes d'extraction des caractéristiques de forme comprennent.

III.3.1.1. Les algorithmes d'image binaire

Qui convertissent une image en format noir et blanc avant de tracer la région extérieure de l'image et d'identifier la forme de l'image à l'aide de facteurs de forme.

$$\text{facteur de forme} = \frac{\text{zone}}{(\text{diametre})^2} \quad \text{Équation III-1}$$

La forme est déterminée par le facteur de forme, dont les valeurs sont prises en compte comme indiqué dans le tableau suivant.

Forme	Facteur de forme
Cercle	$0.7 \leq \text{Facteur} \leq 0.8$
Carré	$0.484 \leq \text{Facteur} \leq 0.55$
Rectangle	$0.25 \leq \text{Facteur} \leq 0.3$
Triangle	$0.445 \leq \text{Facteur} \leq 0.483$
Ovale	$0.32 \leq \text{Facteur} \leq 0.34$
Diamant	$0.36 \leq \text{Facteur} \leq 0.38$

Tableau III-1: Facteur de forme

III.3.1.2. Segmentation horizontale et verticale

Dans une segmentation horizontale, une image est divisée en segments horizontaux, les points de coordonnées sont tracés et la forme de l'objet est déterminée de sorte que le problème de l'emplacement de l'objet ne se pose pas. Dans la segmentation verticale, l'image est divisée en segments verticaux et les points de l'axe des coordonnées sont tracés conformément à la forme de l'axe des coordonnées de l'objet, reconnue avec précision.

La surface d'un élément indique le nombre de pixels inclus dans une limite fermée d'une image binaire, et les distances horizontales et verticales sont représentées en calculant la distance entre les lignes de délimitation, conformément à la méthode de code en chaîne pour le calcul des limites d'une image binaire.

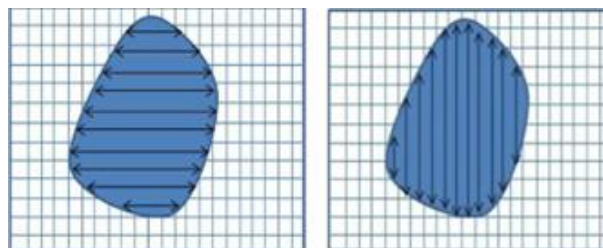


Figure III-1: Calcul des distances horizontales et verticales. (Kalel et al., 2016)

III.4. Les caractéristique de couleur

Basée sur la similarité des couleurs des images, la caractéristique de couleur est une caractéristique visuelle fondamentale utilisée dans la recherche d'images Elle est relativement résistante aux problèmes d'arrière-plan et indépendante des changements de taille, d'orientation ou d'échelle de l'image.(Patel & Gamit, 2016)

De nombreuses caractéristiques de couleur, telles que les moments de couleur (CM), les vecteurs de cohérence de couleur (VCC), les histogrammes de couleur et les corrélogrammes de couleur, sont proposées. (Salau & Jain, 2019)

III.4.1. Moments de couleur

Ces échelles permettent de différencier les photos en fonction de la caractéristique de couleur unique de chacune d'entre elles. La distribution de probabilité est représentée par les moments de couleur dans l'image. Les moments de couleur (Moyenne, Ecart-type et Skewness). (Mutlag et al., 2020)

III.4.2. Histogramme de couleur

La probabilité des intensités pour les trois canaux de couleur (Rouge, Vert, Blue) est définie par l'histogramme des couleurs. Le nombre de pixels de chaque couleur apparaissant dans une image est compté pour créer l'histogramme, et les pixels de chaque couleur sont ensuite placés dans des cases distinctes. Lorsque l'angle de vue change, l'histogramme de l'image change progressivement au lieu de tourner et de se déplacer avec le plan de l'image.

L'histogramme des couleurs peut être représenté par :

$$h_{A,B,C}(a, b, c) = N \cdot \text{prob}(A = a, B = b, C = c) \quad \text{Équation III-2}$$

N est le nombre de pixels dans l'image et A, B et C sont les trois canaux de couleur.

III.4.3. Corrélogramme de couleur

Les informations sur la manière dont les paires de couleurs varient en fonction de la distance sont fournies par un corrélogramme de couleurs. Les caractéristiques suivantes d'un corrélogramme de couleurs sont particulièrement remarquables :

- ✓ Il décrit la corrélation des couleurs dans un plan spatial.
- ✓ Il est utilisé pour décrire la distribution globale par la corrélation spatiale locale des couleurs.
- ✓ Il est relativement simple à calculer.
- ✓ La taille des caractéristiques est relativement petite.

Un corrélogramme de couleurs d'une image est un tableau obtenu en calculant le nombre de pixels de couleur j à une distance k de la couleur i, ce qui donne la k ème entrée à l'emplacement (i, j), divisé par le nombre total de pixels dans les images.

III.4.4. La matrice de cooccurrence des couleurs (MCC)

Est une technique très répandue pour capturer les changements de couleur dans l'image qui produisent la caractéristique de couleur. Chaque image peut être représentée par quatre motifs de balayage, qui peuvent ensuite être construits en quatre matrices bidimensionnelles, car chaque pixel d'une image correspond aux quatre pixels adjacents. Les motifs du modèle de balayage sont créés à partir de ces quatre matrices afin de capturer la variance des couleurs dans l'image. La Figure III-2 représente une fenêtre de 3x3 divisée en une grille de 2x2 ; elle est remplacée par les sept motifs de balayage de la figure 3.

Pour capturer la variation des couleurs d'une image, on utilise des motifs de balayage.

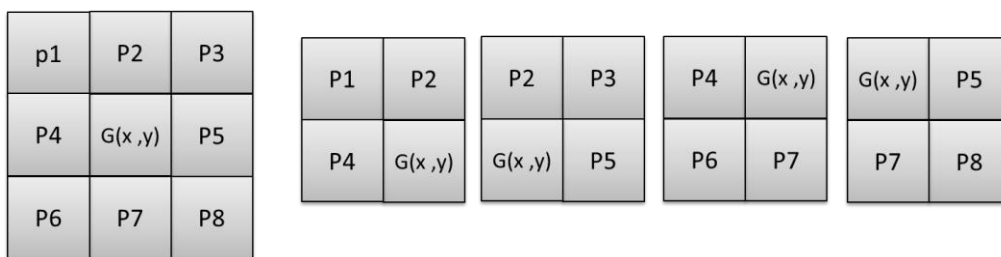


Figure III-2 : La fenêtre 3x3 divisée en grille 2x2

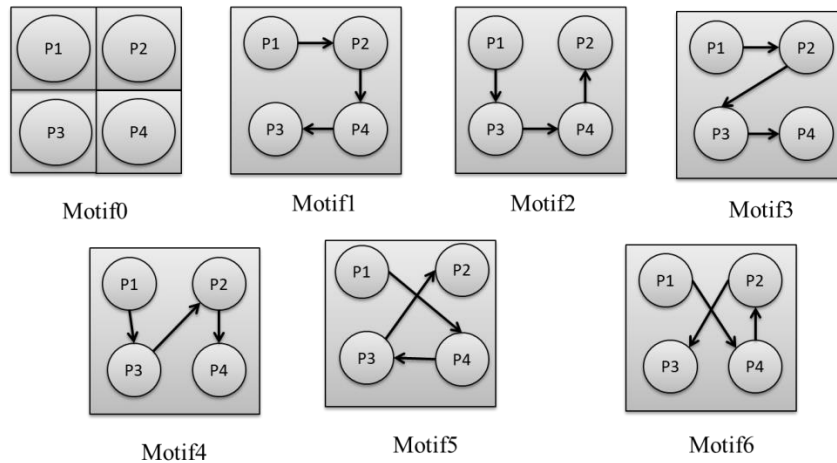


Figure III-3 : Sept modèles de balayage. (Patel & Gamit, 2016)

Le tableau suivant présente un résumé des différentes méthodes de coloration, y compris leurs forces et leurs faiblesses.

La méthode	Avantage	Inconvénients
CM	Compact et robuste.	Insuffisant pour décrire toutes les teintes, manque d'informations spatiales
Histogramme	Facile à calculer et intuitif.	Haute dimension, pas d'information spatiale et sensible au bruit

Corrélogramme	données spatiales.	Coût de calcul extrêmement élevé, sensible au bruit, à la rotation et à l'échelle
---------------	--------------------	---

Tableau III-2 : Contraste de différents descripteurs de couleur. (Tian, 2013)

III.5. Caractéristiques de texture

Pour de nombreuses formes différentes d'images que l'on trouve dans notre vie quotidienne, comme les images médicales, les images de capteurs et d'autres types, la texture est le composant le plus important. Elle est définie comme une représentation au niveau de la surface de la façon dont les humains perçoivent les objets naturels. (Salau & Jain, 2019)

La texture évalue des qualités telles que la douceur et la régularité en mesurant la variation d'intensité d'une surface. Les méthodes d'extraction de caractéristiques de texture peuvent être divisées en deux catégories : les méthodes d'extraction de caractéristiques de texture spatiales et les méthodes d'extraction de caractéristiques de texture spectrales, selon le domaine à partir duquel la caractéristique de texture est extraite. Dans la première méthode, les caractéristiques de texture sont extraites du domaine de l'image originale en calculant les statistiques des pixels ou en localisant les structures locales des pixels, mais dans la deuxième méthode, une image est transformée dans le domaine des fréquences avant que les caractéristiques soient calculées à partir de l'image transformée.

Les caractéristiques spatiales et spectrales présentent toutes deux des avantages et des inconvénients. Le tableau suivant résume leurs avantages et leurs inconvénients :

Méthode de texture	Avantages	Inconvénients
Texture spatiale	Significative, facile à comprendre, peut être extraite de n'importe quelle forme sans perdre d'informations.	Sensible au bruit et aux distorsions
Texture spectrale	Robuste, nécessite moins de calculs.	Aucune signification sémantique, besoin de régions d'image carrées avec une taille suffisante

Tableau III-3 : différences entre les méthodes d'extraction des caractéristiques de texture (Kavya, 2015)

De nombreuses techniques d'analyse de la texture, notamment :

III.5.1. SGLDM (Spatial gray level dependence matrices)

C'est une technique statistique qui consiste à construire des matrices de co-occurrence pour représenter la distribution spatiale des niveaux de gris dans la région d'intérêt. La densité de probabilité conditionnelle du second ordre, $g(i,j,d,\Theta)$ est la base de la méthode SGLDM. En

conséquence, un élément à la position (i, j) de la matrice SGLDM indique la probabilité que deux cellules de résolution distinctes, qui sont dans une orientation spécifique Θ (par rapport à l'horizontale) et sont séparées l'une de l'autre par une distance spécifiée (d) , aient des valeurs de niveau de gris (i, j) correspondantes.

- ✓ Lors de l'évaluation de la direction de la texture, l'angle ne peut généralement avoir que les valeurs $0^\circ, 45^\circ, 90^\circ$ et 135° .
- ✓ Une description significative de la taille de la texture de périodicité peut être obtenue en appliquant de nombreuses valeurs de distance, et la distance d est limitée à des valeurs qui sont des multiples entiers de la taille du pixel.
- ✓ Par conséquent, des matrices SGLDM distinctes sont produites pour des valeurs Θ et d différentes.

Cet exemple montre comment calculer la matrice SGLD avec $\Theta = 0$ et $d=1$.

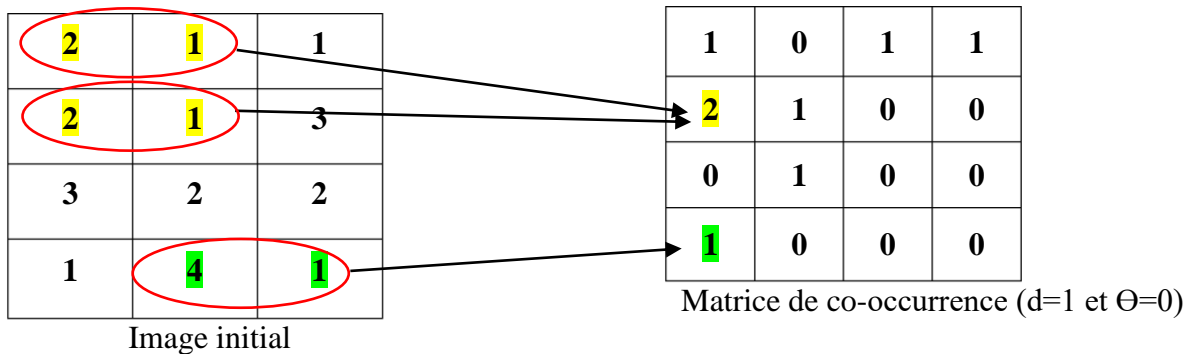


Figure III-4 : Principe de construction de la matrice de cooccurrence.

La taille de la matrice SGLD est $N_g * N_g$, où N_g est le niveau de gris maximal de la région d'intérêt. L'élément $(2, 1)$ de la SGLDM contient la valeur 2 car il y a deux instances dans l'image où deux pixels horizontalement adjacents ont les valeurs 2 et 1. L'élément $(4, 1)$ de la SGLDM contient la valeur 1 car il n'y a qu'une instance dans l'image où deux pixels horizontalement adjacents ont les valeurs 4 et 1. Les différentes caractéristiques de texture reflètent différentes propriétés de l'image, parmi ces caractéristiques :

III.5.1.1. Le Contraste

Il s'agit d'une mesure des changements locaux des niveaux de gris présents dans une image. Un contraste élevé est associé aux images qui présentent des différences significatives dans les niveaux de gris. La dispersion des valeurs de la matrice par rapport à sa diagonale principale peut également être décrite par ce paramètre. L'équation suivante permet de calculer le contraste dans l'image:

$$Cont = \sum_i \sum_j (i - j)^2 g(i, j) \quad \text{Équation III-3}$$

Où $g(i, j)$ correspond aux éléments de la matrice de cooccurrence, c'est-à-dire la probabilité de passer d'un pixel de niveau de gris i à un pixel de niveau de gris j .

III.5.1.2. L'Entropie

Une mesure de la non-uniformité de l'image ou de la zone d'intérêt est appelée entropie. De nombreux membres de la matrice de cooccurrence ont de petites valeurs lorsque l'image est hétérogène, ce qui suggère que l'entropie est assez élevée. L'expression suivante décrit l'entropie :

$$Ent = - \sum_i \sum_j g(i, j) \log(g(i, j)) \quad \text{Équation III-4}$$

III.5.1.3. La Variance

Elle est une mesure de l'hétérogénéité et présente une forte corrélation avec l'écart-type. Elle décrit la distribution des niveaux de gris entourant la valeur moyenne précédemment déterminée. Ainsi, lorsque les valeurs des niveaux de gris s'écartent de leur moyenne, la variance augmente. L'expression de la variance est la suivante : (Sassi & Slima, 2012)

$$Var = \sum_i \sum_j (i - mean)^2 g(i, j) \quad \text{Équation III-5}$$

III.5.1.4. La Corrélation

La caractéristique de corrélation d'une image est une mesure de l'association d'un pixel avec ses pixels voisins dans l'image. Les valeurs de corrélation vont de -1 à 1 lorsqu'une image est parfaitement corrélée positivement, et à l'infini lorsqu'une image est constante. L'équation suivant donne la définition de la caractéristique "Correl" d'une image :

$$Correl = \sum_i \sum_j \frac{(i - \mu_i)(j - \mu_j)g(i, j)}{\sigma_i \sigma_j} \quad \text{Équation III-6}$$

Où μ_i représente la valeur moyenne des intensités des pixels dans la i ème ligne de la matrice de cooccurrence des niveaux de gris (GLCM), et μ_j représente la valeur moyenne des intensités des pixels dans la j ième colonne de la GLCM.

III.5.1.5. La moyenne de la somme

La caractéristique de moyenne de la somme calcule la moyenne d'une image pour la distribution des niveaux de gris. L'équation suivant donne la définition de la caractéristique de la moyenne de la somme d'une image (Mall et al., 2019):

$$Aver = \sum_{i=2}^{2Ng} i g_{x+y}(i) \quad \text{Équation III-7}$$

III.5.1.6. Le second moment angulaire(ASM)

Appelé aussi uniformité, il représente la mesure de l'homogénéité de l'image. (Mazher et al., 2011)

$$ASM = \sum_{i=0}^{Ng} \sum_{j=0}^{Ng} g(i,j)^2 \quad \text{Équation III-8}$$

III.5.1.7. IDM

Le moment de différence inverse (IDM) peut être utilisé pour calculer l'homogénéité locale ou l'homogénéité de la texture d'une image. L'IDM est un indicateur statistique de l'homogénéité locale de la matrice de cooccurrence des niveaux de gris (GLCM). L'équation de l'IDM est la suivante :

$$IDM = \sum_i \sum_j \frac{g(i,j)}{1 + (i-j)^2} \quad \text{Équation III-9}$$

III.5.1.8. La somme de variances

L'addition des valeurs de variance de chaque pixel de l'image permet de déterminer la la somme de variances d'une image. La formule de calcul de la somme de variances est donnée dans l'équation suivante:

$$SV = \sum_{i=2}^{2N} (i - SE)^2 g_{x+y}(i) \quad \text{Équation III-10}$$

III.5.1.9. La somme des entropies

La somme des entropies est un indicateur statistique de la densité d'information d'une image. Selon l'équation au-dessous, elle est déterminée en utilisant la probabilité d'apparition de chaque niveau d'intensité dans l'image.

$$SE = - \sum_{i=2}^{2N} g_{x+y}(i) \log(g(i,j)) \quad \text{Équation III-11}$$

III.5.1.10. La variance de différence

La variance de différence d'une image est différenciée pour obtenir la variance de différence présentée dans l'équation suivante :

$$DV = \sum_{i=0}^N (i - mean)^2 g_{x-y}(i) \quad \text{Équation III-12}$$

III.5.1.11. L'entropie de différence

Les valeurs d'entropie d'une image sont différenciées pour obtenir l'entropie de différence comme indiqué dans l'équation :

$$DE = - \sum_{i=0}^N g_{x-y}(i) \log(g_{x-y}(i)) \quad \text{Équation III-13}$$

III.5.1.12. IMC1

Les mesures d'information de la corrélation¹ sont illustrées dans cette équation, où HX et HY sont les entropies de Px et Py.

$$ICM1 = \frac{HXY - HXY1}{\max(HX, HY)} \quad \text{Équation III-14}$$

$$\text{Avec } HXY = - \sum_i \sum_j g(i, j) \log(g(i, j)) \quad \text{Équation III-15}$$

$$\text{Et } HXY1 = - \sum_i \sum_j g(i, j) \log(g_x(i)g_y(j)) \quad \text{Équation III-16}$$

III.5.1.13. IMC2

Les entropies de Px et Py sont utilisées dans cette équation pour représenter les mesures d'information de la corrélation². (Perumal et al., 2021)

$$ICM2 = (1 - e^{-(HXY2-HXY)})^{1/2} \quad \text{Équation III-17}$$

$$\text{Avec } HXY2 = - \sum_i \sum_j g_x(i)g_y(j) \log(g_x(i)g_y(j)) \quad \text{Équation III-18}$$

III.5.2. Local binary pattern (LBP)

Dans les tâches de vision par ordinateur et de reconnaissance des visages, le LBP est un descripteur utile. La valeur du pixel central est substituée par le LBP codant les niveaux de gris d'une image en comparant le pixel central avec ses voisins et le résultat compte comme un nombre binaire traduit en nombre décimal. La formule pour calculer le LBP en utilisant un voisinage est la suivante :

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad \text{Équation III-19}$$

Où g_p est la valeur du voisin, g_c est la valeur du pixel actuel, R est le rayon du voisinage circulaire, et P est le nombre de voisins.

III.5.3. Histogramme du gradient orienté (HOG)

Le HOG (histogram of oriented gradient) est un descripteur d'objet développé par Dalal et Triggs qui se concentre sur la forme ou la présentation d'un objet dans une image. L'histogramme des gradients orientés donne des caractéristiques distinctives en présence de bruit de fond et d'éclairage variable, ce qui en fait un descripteur utile. L'approche HOG a déterminé les emplacements des axes de gradient dans le ROI. Cette méthode utilise le gradient des niveaux de gris locaux pour capturer la forme structurelle dans la région d'intérêt (ROI). (Farhan & Kamil, 2020)

III.5.4. Filtres de Gabor

La caractéristique de texture la plus étudiée est probablement l'ondelette de Gabor. Les statistiques de ces micro-caractéristiques à un endroit donné sont fréquemment utilisées pour définir l'information de texture sous-jacente, et les filtres de Gabor peuvent être considérés comme des détecteurs de bords et de lignes réglables en orientation et en échelle. En définissant les paramètres de fréquence centrale et d'orientation, le filtre de Gabor est spécifiquement conçu pour échantillonner l'ensemble du domaine de fréquence d'une image. (Kavya, 2015)

III.6. Conclusion

Dans de nombreuses applications de vision artificielle et de traitement d'images, l'extraction de caractéristiques est une étape cruciale. Les caractéristiques telles que la forme, la texture et la couleur peuvent être extraites des données brutes afin d'obtenir des informations pertinentes qui peuvent ensuite être utilisées pour améliorer les modèles d'apprentissage automatique.

Chapitre

4

Mammographie et diagnostic assisté par ordinateur

IV.1. Introduction

Le cancer du sein est une maladie répandue et dangereux qui touche les femmes. Selon l'Organisation mondiale de la santé, ce type de cancer est la principale cause de décès non accidentels; en particulier, 8,8 millions d'individus dans le monde sont décédés des suites d'un cancer en 2015. Le pronostic et le taux de survie des patients atteints de cancer du sein peuvent donc être grandement améliorés par un diagnostic précoce et précis.

La mammographie (radiographie du sein), les sonogrammes, la tomographie par ordinateur (CT), l'imagerie par résonance magnétique (IRM) et les images d'histopathologie (HP) sont toutes des modalités d'imagerie médicale courantes pour le diagnostic du cancer du sein. Ces modalités d'imagerie médicale sont très efficaces pour détecter le cancer du sein.

IV.2. Les anomalies mammaires

Le dépistage du cancer du sein est le processus qui consiste à utiliser des modalités médicales multi-image pour trouver des anomalies du sein afin de les diagnostiquer rapidement et de prévenir la propagation du cancer. Il est recommandé pour les femmes de 40-44 ans pour commencer le dépistage annuel du cancer du sein, de 45-54, il convient de procéder à un dépistage des seins chaque année, de 55-au-dessus si leur santé est bonne, elles devraient se faire dépister tous les ans ou tous les deux ans (au choix).(Ramadan, 2020)

Après le dépistage du cancer du sein, on a recours à la mammographie de dépistage. Une mammographie de diagnostic est généralement conseillée pour examiner de manière plus approfondie les zones préoccupantes si la mammographie de dépistage a révélé une anomalie dans les tissus mammaires.(Mahmood et al., 2020)

IV.2.1. Anomalies mammaires

Toute modification ou imperfection du tissu mammaire qui s'écarte de sa texture ou de son apparence habituelle est appelée anomalie mammaire. Le cancer du sein peut être lié ou non à ces anomalies. Parmi les anomalies mammaires typiques, on peut citer :

IV.2.1.1. Masse

La masse décrit par un gonflement localisé du sein, une protubérance, une saillie ou le développement d'une bosse, Ces éléments sont distincts du tissu mammaire voisin ou du tissu mammaire de la même région du sein opposé. Les masses mammaires peuvent être aussi bien

bénignes (non cancéreuses) que malignes (cancéreuses), ils sont décrites selon leur forme, leur marge et leur densité. (Singh & Choudhary, 2019)

Forme : Une masse peut être lobulaire, irrégulière, sphérique, ovale, ou toute combinaison de ces formes.

Les contours d'une masse peut être confinée, micro lobulée, dissimulée (par l'épaisseur des tissus voisins), indistincte (mal définie) ou hypothétique.

Densité : La densité d'une masse peut être inférieure, isodense ou supérieure à la densité du tissu glandulaire voisin. La masse peut également contenir de la graisse. (Balleyguier et al., 2007)

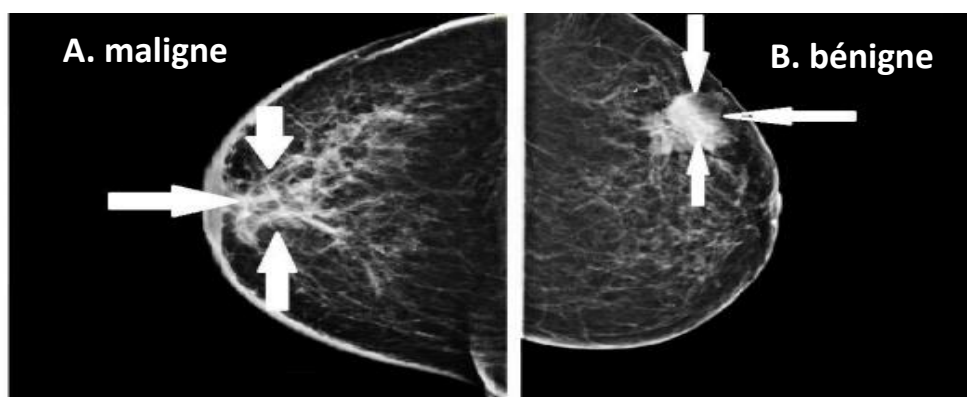


Figure IV-1 : Exemple d'une masse maligne (à gauche) et une masse bénigne (à droite). (Mahmood et al., 2020)

IV.2.1.2. Les Microcalcifications

La présence de minuscules dépôts de calcium dans les tissus du sein donne lieu à des microcalcifications. Elles peuvent être soit bénignes, soit suspectes, soit très probablement cancéreuses. Par rapport aux tissus environnants, les calcifications bénignes sont souvent plus grandes et plus visibles. L'agrandissement n'est pas nécessaire pour les étudier. L'agrandissement est nécessaire pour étudier les caractéristiques des calcifications suspectes car elles sont plus petites. (Singh & Choudhary, 2019)

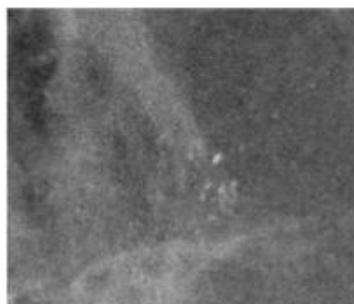


Figure IV-2 : Exemple d'une microcalcification. (Alsheh Ali et al., 2019)

IV.2.1.3. Distorsions architecturales

Lorsque l'architecture régulière du sein est déformée, aucune masse ne l'accompagne. Cette distorsion peut se présenter sous la forme d'un arrangement tissulaire irrégulier s'étendant vers l'extérieur à partir d'un point unique, d'une rétraction focale ou d'un motif désordonné. Le centre transparent (opacité centrale) ou le centre dense sont deux exemples de ce type de distorsion. (Singh & Choudhary, 2019)

IV.3. L'imagerie médicale dédiée au dépistage des anomalies mammaires

Le cancer est une affection, lorsqu'elle n'est pas traitée, peut entraîner la mort et des lésions indolentes. Une mutation génétique entraîne des cellules anormales qui se développent de manière incontrôlée et deviennent malignes à la suite de modifications de leur acide désoxyribonucléique. Alors que la malignité se propage dans nombreuses fonctions de l'organisme via le système lymphatique et puise ses ressources dans les tissus de l'organisme, la bénignité n'envahit pas les tissus voisins. (Mahmood et al., 2020)

IV.3.1. Echographie

Les sonogrammes sont un autre nom pour l'échographie. Contrairement aux mammographies et à l'IRM, l'échographie utilise des ondes sonores à haute fréquence pour scanner le sein sans utiliser de rayonnement. En plus du cancer du sein, l'échographie peut être utilisée pour identifier des anomalies telles que l'inconfort, l'œdème et les infections des organes internes du corps. De plus, l'échographie peut aider à la biopsie mammaire à l'aiguille pour l'évaluation intrinsèque des tissus mammaires. (Murtaza et al., 2020)

IV.3.2. Imagerie par résonance magnétique (IRM)

De fortes ondes radio et des champs magnétiques sont utilisés par l'IRM, une technique d'imagerie diagnostique, pour enregistrer et afficher une image en 3D du tissu mammaire. Après l'identification d'un cancer, les médecins utilisent fréquemment l'IRM pour obtenir des informations détaillées. (Mahmood et al., 2020)

IV.3.3. La Mammographie

La mammographie est une technique d'imagerie courante pour identifier et diagnostiquer le cancer du sein. La mammographie utilise des rayons X de faible énergie pour obtenir une image du sein. Les anomalies mammaires telles que les masses, les calcifications et les

distorsions architecturales peuvent être détectées et analysées à l'aide de la mammographie mammaire. Ces anomalies permettent aux médecins de calculer le risque de développer un cancer du sein chez une patiente.(Singh & Choudhary, 2019)

De nos jours, des appareils de mammographie numérique plein champ sont utilisés pour obtenir des mammographies, qui sont ensuite livrées sous forme de données d'imagerie brutes (pour traitement) et de données post-traitées (pour présentation). Les images utilisées pour les présentations sont destinées à l'évaluation visuelle des radiologues. Ces dernières années, des tentatives importantes ont été faites pour trouver des données pertinentes dans les mammographies afin de faciliter la prédiction du risque de cancer du sein à long terme. On connaît maintenant plusieurs facteurs qui augmentent le risque de cancer du sein à long terme, notamment la densité mammographique des seins, les antécédents familiaux de cancer du sein, l'âge tardif de la ménopause, les premières règles précoces et l'âge tardif de la première naissance.(Alsheh Ali et al., 2019)

Le dépistage par mammographie est une technique d'examen peu coûteuse, non invasive et confortable. Sur une mammographie, les calcifications apparaissent comme des taches blanches ou des zones d'ombre. Les mammographies à deux lectures ont une sensibilité et une spécificité plus élevées.

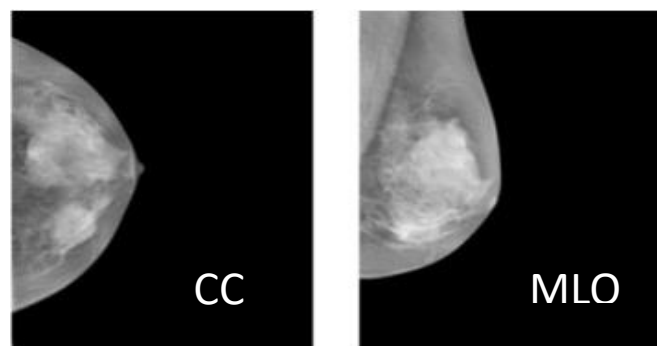


Figure IV–3 : Un exemple d'étiquettes d'images de mammographie.(Wang et al., 2020)

Les vues (CC) cranio-caudale et (MLO) oblique médio-latérale sont les deux projections standard utilisées pour visualiser les tissus mammaires dans différentes orientations. La vue CC est prise en regardant le sein de face et la vue MLO est prise en regardant le sein de côté. En utilisant ces deux vues, les radiologues peuvent mieux visualiser les différentes parties du tissu mammaire et détecter les anomalies qui peuvent être cachées dans une seule vue. En outre, les deux vues peuvent aider à confirmer la présence de l'anomalie en montrant la lésion dans des projections multiples. La Figure IV–3 illustre les étiquettes des clichés mammographiques (Les images et les étiquettes pour CC et MLO sont présentées dans les respectivement).

Le sein entier avec ses structures glandulaires est évalué en vue cranio-caudale. La bande noire montre les tissus graisseux, tandis que le contour montre le mamelon. De plus, l'oblique médio-latérale est une projection à 90 degrés dans laquelle on peut voir l'aisselle lorsque les tissus mammaires sont en dehors du quadrant du sein.(Mahmood et al., 2020)

IV.4. Détection et diagnostic assisté par ordinateur

L'erreur humaine est un facteur important qui peut affecter la qualité de l'analyse visuelle des images médicales, y compris les mammographies. En raison de la densité du tissu mammaire, de la qualité de l'image, de la fatigue oculaire, du manque d'expérience et d'autres facteurs, les radiologues peuvent manquer des anomalies dans la mammographie, conduisant à des faux négatifs. C'est pourquoi les systèmes de diagnostic assisté par ordinateur (DAO) ont été développés pour aider les médecins à détecter les anomalies à un stade précoce. (Ruchai et al., 2021)

La détection assistée par ordinateur (DAO) est un logiciel qui analyse les images médicales, telles que les mammographies ou les radiographies, afin d'identifier les zones anormales ou suspectes. Le logiciel utilise des algorithmes de reconnaissance des formes pour détecter les caractéristiques qui peuvent indiquer la présence d'une maladie ou d'une affection. L'objectif de la DAO est d'aider les radiologues à détecter les anomalies ou des signes de la maladie qui peuvent être difficiles à détecter à l'œil nu. En attirant l'attention sur les zones qui nécessitent un examen plus approfondi, la DAO peut contribuer à réduire le risque de résultats faussement négatifs, qui se produisent lorsqu'une anomalie est présente mais n'est pas détectée (Castellino, 2005). Les systèmes d'aide au diagnostic sont classés en deux catégories DAOe (détection assistée par ordinateur) et DAOx (diagnostic assisté par ordinateur):

Détection assistée par ordinateur désigne les systèmes de détection assistée par ordinateur, sont conçus pour aider les radiologues à détecter et à localiser les anomalies dans les images médicales. Ces anomalies peuvent inclure divers types d'anomalies, comme des nodules, des masses, des calcifications ou d'autres signes de maladie.

Diagnostic assisté par ordinateur désigne les systèmes de diagnostic assisté par ordinateur, sont conçus pour aider les radiologues à déterminer la classe d'anomalies identifiées par le système CADe et à les classer comme bénignes ou malignes. Ces systèmes utilisent des algorithmes avancés et des techniques d'apprentissage automatique pour analyser les images médicales et fournir des informations supplémentaires au radiologue, ce qui peut contribuer à améliorer la précision des diagnostics.

Les systèmes DAOx peuvent être utilisés conjointement avec les systèmes DAOe pour fournir des informations supplémentaires et aider à la prise de décision. En fournissant des informations plus détaillées sur les anomalies identifiées par le système CADe, les systèmes DAOx peuvent aider les radiologues à prendre des décisions plus éclairées sur les soins aux patients et les options de traitement.

IV.4.1. Évolution de la DAO

Les systèmes de DAO sont classés en fonction de la manière d'utiliser les circonstances entourant l'interprétation de l'image par le médecin :

IV.4.1.1. DAO en seconde lecture

Fait référence aux systèmes de DAO qui sont utilisés comme 2ème avis après l'interprétation initiale par un radiologue. Ces systèmes analysent les mêmes images que le radiologue a déjà examinées et fournissent une analyse supplémentaire et des résultats potentiels. Le radiologue peut alors utiliser ces informations pour établir un diagnostic final.

Ce type contient les DAO interactif est un autre type de système de diagnostic qui permet aux radiologues d'interagir avec l'ordinateur pendant le processus d'interprétation. Les systèmes DAO peuvent fournir un retour d'information en temps réel au radiologue pendant l'analyse de l'image.

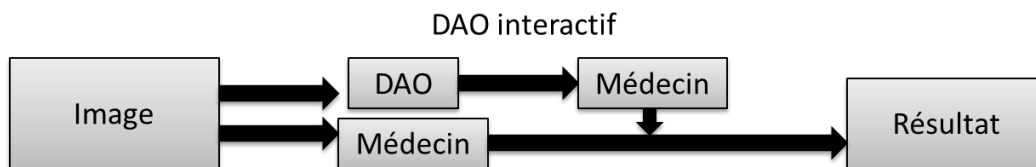


Figure IV-4 : DAO de deuxième lecteur

IV.4.1.2. DAO à lecture simultanée

La lecture simultanée fait référence aux systèmes de DAO qui fonctionnent simultanément avec le radiologue pendant le processus d'interprétation des images. Ces systèmes fournissent une analyse en temps réel et aident le radiologue à identifier les anomalies potentielles et à poser un diagnostic.



Figure IV-5 : DAO de lecteur simultané

IV.4.1.3. DAO en première lecture

Un système DAO en première lecture fait référence aux systèmes de DAO qui sont utilisés comme méthode principale d'interprétation des images, le radiologue examine les résultats du système DAO et introduit le diagnostic final. Ces systèmes sont encore relativement nouveaux et leur utilisation n'est pas encore très répandue, mais ils ont le potentiel d'améliorer la précision et l'efficacité de l'interprétation des images.

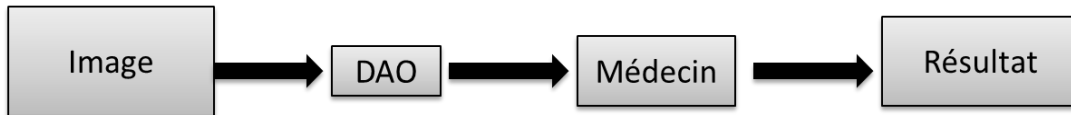


Figure IV-6 : DAO de premier lecteur.

IV.4.2. Diagnostic à apprentissage continu (CAO)

Est un nouveau type de système de CAO qui utilise des algorithmes d'apprentissage automatique pour améliorer continuellement ses performances au fil du temps. Ces systèmes sont conçus pour apprendre à partir de nouvelles données au fur et à mesure qu'elles sont disponibles, ce qui leur permet de s'adapter à des circonstances changeantes et d'améliorer leur précision. Les systèmes de CAO à apprentissage continu ont le potentiel de révolutionner l'imagerie médicale en fournissant des outils de diagnostic plus précis et plus fiables qui peuvent aider les médecins à prendre des décisions mieux informées sur les soins à apporter aux patients. (Fujita, 2020)

IV.4.3. Mesures de performance des CAD

La matrice de confusion est l'une des estimations utilisées pour évaluer la performance d'un modèle. Elle indique la relation entre les valeurs prédites et les valeurs réelles.

Vrai positif (TP) : patients diagnostiqués positifs et qui le sont réellement.

Vrai négatif (VN) :- patients diagnostiqués négatifs et réellement négatifs.

Faux positif (FP) : patients diagnostiqués positifs mais réellement négatifs.

Faux négatif (FN) : patients diagnostiqués négatifs et qui sont réellement positifs.

	Positif	Négatif
Positif	Vrai positif (VP)	Faux négatif
Négatif	Faux positif (FP)	Vrai négatif

Figure IV-7:Matrice de confusion

D'autres mesures permettant d'évaluer l'efficacité d'un modèle sont examinées ci-dessous :

IV.4.3.1. Accuracy

Il s'agit de la proportion de prédictions correctes (à la fois vraies positives et vraies négatives) par rapport à l'ensemble des prédictions effectuées.

$$Accuracy = \frac{VP + VN}{VP + FP + FN + VN} \quad \text{Équation IV-1}$$

IV.4.3.2. Précision

Elle mesure la proportion de vraies prédictions positives par rapport à l'ensemble des prédictions positives (vraies positives et fausses positives).

$$Précision = \frac{VP}{VP + FP} \quad \text{Équation IV-2}$$

IV.4.3.3. Rappel ou sensibilité

Il s'agit de la proportion de prédictions réellement positives par rapport à l'ensemble des cas positifs (vrais positifs et faux négatifs).

$$Rappel = \frac{VP}{VP + FN} \quad \text{Équation IV-3}$$

IV.4.3.4. Spécificité

Elle mesure la proportion de vraies prédictions négatives par rapport à tous les cas négatifs réels (vrais négatifs et faux positifs).

$$Spécificité = \frac{VN}{VN + FP} \quad \text{Équation IV-4}$$

IV.4.3.5. Score F1

Il s'agit de la moyenne harmonique de la précision et du rappel, qui permet d'équilibrer les deux mesures. (Chugh et al., 2021)

$$F1 = 2 * \frac{Précision * Rappel}{Précision + Rappel} \quad \text{Équation IV-5}$$

IV.5. Revue de la littérature

Les systèmes DAOe et DAOx sont des outils importants dans le domaine de l'imagerie médicale et peuvent contribuer à améliorer la précision des diagnostics et à réduire le risque de diagnostics manqués ou incorrects. Cependant, il est important de se rappeler que ces

systèmes sont conçus pour être utilisés comme des aides à l'expertise et au jugement du radiologue, et qu'ils ne peuvent pas s'y substituer.(Chugh et al., 2021)

IV.5.1. Les systèmes d'aide au diagnostic basé sur les caractéristiques

Système de DAO mettant en œuvre l'apprentissage automatique comprenant les étapes suivantes :

- **Prétraitement de l'image :** Cette étape consiste à éliminer le bruit et à nettoyer l'image pour la préparer à l'étape suivante. Cela peut inclure l'utilisation de filtres pour supprimer le bruit et améliorer la qualité de l'image.
- **Segmentation d'image :** Cette étape consiste à partitionner l'image en plusieurs parties, pour ne travailler que sur les régions d'intérêt (ROI) qui contiennent des anomalies ou des caractéristiques pertinentes. Cela permet de réduire la quantité de données à traiter et d'améliorer la précision de l'analyse.
- **Extraction et sélection de caractéristiques :** Cette étape consiste à extraire des caractéristiques pertinentes à partir des images prétraitées. Les caractéristiques peuvent inclure des mesures de forme, de texture, de couleur, etc. La sélection des caractéristiques est importante pour améliorer la précision et réduire les coûts et le temps d'apprentissage.
- **Classification:** Cette étape consiste à attribuer des étiquettes ou des classes à différents groupes à l'aide de plusieurs classificateurs tels que la machine à vecteurs de support (SVM), le K-plus proche voisin (KNN), Naïve Bayes (NB), l'arbre de décision (DT), la forêt aléatoire (RF), régression logistique (LR), etc. Le classificateur utilise les caractéristiques extraites pour identifier les anomalies et les classer comme bénignes ou malignes.



Figure IV–8: Systèmes DAO basés sur l'apprentissage automatique

Fayez et al.(AlFayez et al., 2020) ont proposé une stratégie basée sur le thermogramme avec quatre phases distinctes pour l'identification du cancer du sein. Ils ont effectué un prétraitement de l'image au cours de la première phase, y compris la réduction du bruit à l'aide de l'algorithme de filtrage homomorphe et des méthodes d'amélioration de l'image telles que la transformation top-hat et l'égalisation adaptative de l'histogramme.

La deuxième phase s'est concentrée sur la segmentation de la région d'intérêt à l'aide de techniques de masquage binaire et de regroupement K-means. Ensuite, l'extraction des caractéristiques a été effectuée via des approches de limites de signature dans la troisième phase. Les classificateurs (ELM) et (MLP) ont été utilisés dans la dernière phase pour effectuer la tâche de classification. Des comparaisons ont été effectuées pour évaluer l'efficacité de ces classificateurs. Ils ont utilisé un ensemble de données largement disponible appelé DMR-IR pour évaluer l'approche proposée. La fusion de l'extraction des caractéristiques géométriques et de l'extraction des caractéristiques texturales était l'un des nombreux scénarios d'expérience qu'ils ont créés en utilisant divers mécanismes d'activation. Diverses mesures, dont la précision, la sensibilité et la spécificité, ont été utilisées pour étayer l'évaluation.

Les résultats des tests ont montré que le classificateur MLP n'a produit qu'une précision de 82,20 %, alors que la détection du cancer du sein basée sur l'ELM a atteint une précision spectaculaire de 100 %. Les résultats ont montré que la stratégie basée sur l'ELM était plus précise que l'approche basée sur le MLP de plus de 19 %.

En utilisant la radiomique et les sous-types moléculaires, Sutton et al. (Sutton et al., 2020) ont créé et validé un modèle d'apprentissage automatique pour classifier les réponses pathologiques complètes (pCR) du cancer du sein après une chimiothérapie néoadjuvante (NAC) sur l'IRM. Le modèle comportait une analyse radiomique automatisée, qui comprenait la segmentation des images, l'extraction des caractéristiques, l'application de pré-filtrés et la création de classificateurs avec la forêt aléatoire à élimination récursive des caractéristiques (RFE-RF) de l'apprentissage automatique. Les caractéristiques radiomiques comprenaient des caractéristiques d'histogramme de premier ordre, des caractéristiques de texture de Haralick de second ordre, des caractéristiques de cartes de bords de Gabor et des mesures de texture de Haralick calculées à partir de cartes de bords de Gabor. En raison de leur association et de leur valeur démontrée dans des travaux antérieurs, cinq textures de Haralick - énergie, entropie, corrélation, contraste et homogénéité - ont été générées au lieu de 13. Le modèle a obtenu des valeurs d'aire sous la courbe caractéristique d'exploitation du récepteur (AUROC) prometteuses en validation croisée pour les modèles 1 (0,72) et modèle 2 (0,80), avec des valeurs AUROC similaires dans les ensembles d'apprentissage et de test.

Une étude de Soumya Roy et al. (Roy et al., 2021) a utilisé un ensemble de données IDC sur le cancer du sein de 277 524 images. Parmi ces photos, 198 738 étaient négatives et 78 786 positives. Dans leur étude, ils ont voulu classer les photos dans les catégories IDC(+) et IDC(-). Pour ce faire, ils ont utilisé divers extracteurs de caractéristiques, tels que le SIFT

(Scale-Invariant Feature Transform), le SURF (Speeded-Up Robust Features) et l'ORB (Oriented FAST and rotated BRIEF), qui permettent de capturer des informations texturales. Ils ont également utilisé des éléments statistiques tels que les caractéristiques de texture Haralick. Ces caractéristiques ont été combinées pour fournir un ensemble de 782 caractéristiques.

Ils ont utilisé des techniques d'apprentissage automatique pour améliorer les performances de classification en empilant les caractéristiques et en utilisant une variété de classifieurs d'apprentissage automatique, tels que Random Forest, Extra Trees, XGBoost, AdaBoost, CatBoost et Multi Layer Perceptron. Un ensemble de données plus petit ne comportant que quatre caractéristiques a été produit à la suite de la sélection des caractéristiques à l'aide du coefficient de corrélation de Pearson. La dernière catégorisation a été effectuée à l'aide de ces quatre caractéristiques.

Sur la base des résultats de leurs expériences, ils ont conclu que CatBoost avait la meilleure précision, 92,55 %. D'autres techniques de pointe, dont beaucoup utilisent des architectures d'apprentissage profond, étaient comparables en termes de précision à notre méthode.

Kadhim et al. (Kadhim & Kamil, 2022) se sont concentrés sur le carcinome canalaire invasif lors de la création d'un système de diagnostic assisté par ordinateur (CAO) pour les images histologiques. L'équipe a utilisé l'extracteur de caractéristiques Histogramme des gradients orientés (HOG) pour extraire des informations statistiques des images histopathologiques.

Une partie de l'ensemble de données, contenant des photos portant les numéros 100, 200, 400, 1000 et 2000, a été choisie au hasard afin d'entraîner le système. Ensuite, une variété d'algorithmes, y compris l'arbre de décision (DT), le perceptron multicouche (MLP), la machine à vecteur de support (SVM), les arbres randomisés supplémentaires, le gradient boosting (GB), le classificateur de processus gaussien (GPC), Bayes naïf (NB), le centroïde le plus proche (NC), et Bayes naïf (NB) ont été entraînés en utilisant ces caractéristiques statistiques. L'objectif était d'identifier le développement malin ou non cancéreux à partir des photographies. Les performances de ces algorithmes ont été évaluées sur la base d'un certain nombre de critères, notamment la spécificité, l'exactitude, la sensibilité, la précision, le score F1 et l'aire sous la courbe (AUC). D'après les résultats des tests, la précision la plus élevée (90 %) a été atteinte lorsque 100 images ont été utilisées avec l'algorithme NC. Les algorithmes GPC et NC ont permis d'atteindre la valeur de sensibilité la plus élevée de 100 % avec 100

photos. En outre, l'utilisation de 100 et 200 photos a permis à l'algorithme NB d'obtenir une sensibilité de 100 %. L'utilisation des algorithmes QDA et GB avec 100 photos a permis d'obtenir la plus grande spécificité (90 %). Toutefois, l'efficacité des algorithmes diminue à mesure que le nombre des images augmente. Ce résultat montre que l'apprentissage automatique peut être moins efficace lorsqu'il s'agit d'un grand nombre de photos de ce type.

IV.5.2. Les systèmes d'aide au diagnostic basés sur l'apprentissage profond

L'apprentissage en profondeur, et plus précisément les réseaux de neurones convolutifs (CNN), ont montré une grande efficacité dans la détection et le diagnostic de différentes pathologies, dont le cancer. Les CNN permettent d'extraire automatiquement des caractéristiques pertinentes à partir d'images médicales et de les utiliser pour la classification des anomalies détectées. Les CNN sont capables d'apprendre des motifs de texture et de forme complexes dans les images, ce qui permet une détection plus précise des lésions.

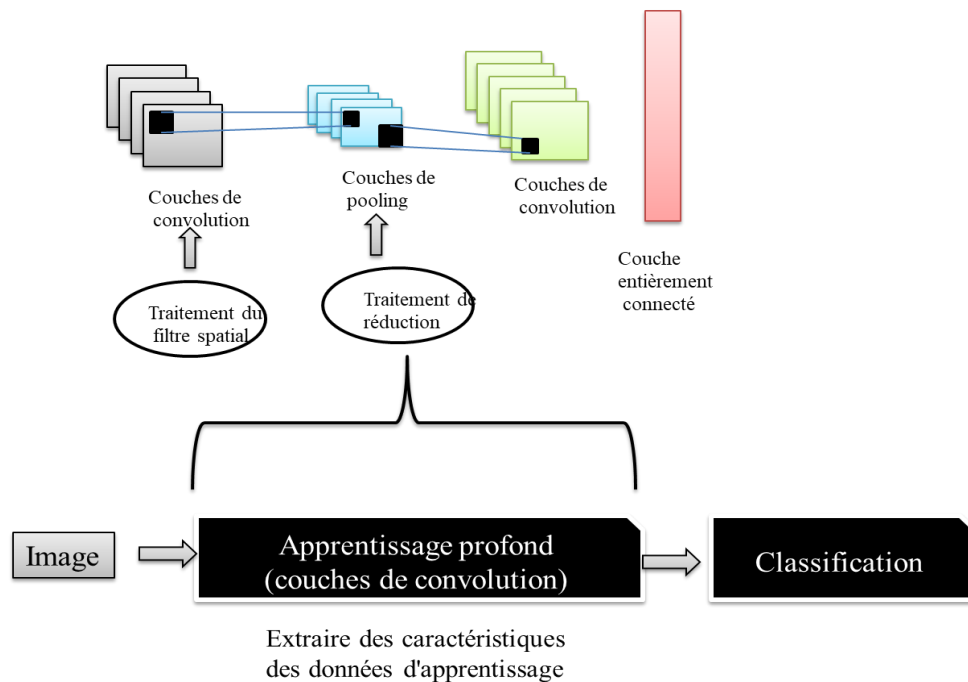


Figure IV-9 : Systèmes DAO basés sur l'apprentissage profond

Une méthode de classification du cancer du sein utilisant le modèle VGG16 et une méthodologie d'apprentissage par transfert a été proposée par Hossain (Hossain et al., 2023). Ils ont utilisé les couches de convolution et les couches maxpooling du modèle VGG16 pré-entraîné comme extracteurs de caractéristiques et ont appliqué un filtrage médian pour découper les images. Afin d'optimiser le modèle, ils ont créé un classificateur de réseau neuronal profond à deux couches et ont utilisé l'optimiseur Adam avec un taux

d'apprentissage de 0,001 et l'entropie croisée binaire comme fonction de perte. Le surajustement a été évité par la mise en œuvre de dropout. Un ensemble de 897 images d'échographie mammaire provenant de deux bases de données a servi de base à l'entraînement, à la validation et au test du classificateur. Outre la réduction de la complexité informatique, les résultats expérimentaux ont montré une accuracy d'apprentissage de 98,2 % et une accuracy de test de 91 % pour les données de test en aveugle. La méthode Grad-CAM, utilisée par les chercheurs, a produit des résultats remarquables lorsqu'il s'est agi de visualiser et d'évaluer la localisation de régions spécifiques au niveau de la couche convolutionnelle finale. Les résultats de l'étude peuvent être utilisés à des fins cliniques pour la détection du cancer du sein.

Li et al. (Li et al., 2021) ont proposé un modèle de classification du cancer du sein EfficientNet qui intègre le mécanisme d'attention CBAM et une méthode de fusion multi-échelle. Ils ont appliqué ce modèle à un ensemble de données d'imagerie médicale sur le cancer du sein et ont constaté qu'il était assez précis pour classer les différents stades du cancer du sein. Comparée aux algorithmes de catégorisation d'images existants, leur solution a atteint la plus grande précision. Selon les chercheurs, la combinaison d'EfficientNet_B5 avec CBAM et la fusion multi-échelle augmente la performance de la classification, fournissant des informations utiles pour les études futures dans le traitement des images médicales et la stadification du cancer du sein. Les résultats indiquent que la précision du modèle EfficientNet est passée de 64,71 % (base) à 70,58 % (EfficientNet+CBAM) et 76,47 % (EfficientNet+CBAM+fusion multi-échelle). L'accuracy atteinte pour l'étape 0 était de 75,52 %, l'étape 1 de 80,91 %, l'étape 2 de 74,21 % et l'étape 3 de 68,95 %. Dans l'ensemble, le modèle EfficientNet proposé avec CBAM et fusion multi-échelle a surpassé le modèle de base et les modèles optimaux actuels en termes de précision et d'interopérabilité.

Hammed et al. (Hameed et al., 2020) ont présenté une stratégie d'apprentissage profond d'ensemble pour la classification des images histopathologiques du cancer du sein, en se concentrant sur la classification définitive des images non carcinomateuses et carcinomateuses. Ils ont formé quatre modèles alternatifs basés sur des architectures VGG16 et VGG19 pré-entraînées, en utilisant à la fois des versions entièrement entraînées et des versions affinées. Ils ont évalué les modèles individuels à l'aide d'une validation croisée 5 fois avant d'employer une technique d'ensemble basée sur la probabilité prédite moyenne. L'ensemble de modèles VGG16 et VGG19 affinés a donné de bons résultats, en particulier dans la classe des carcinomes. Il a obtenu une sensibilité de 97,73 % et une accuracy globale

de 95,29 %, avec un score F1 de 95,29 %. Les résultats confirment l'efficacité de leur approche d'apprentissage profond pour la catégorisation automatique d'images histopathologiques de nature complexe, notamment les images de cancer. L'accuracy était de 93,53 % pour le modèle non carcinomateux entièrement entraîné et de 93,99 % pour le modèle carcinomateux VGG16+VGG19. Après ajustement, la précision est passée à 95,56 % pour le carcinome et à 95,29 % pour le non-carcinome. Il convient de noter que leur ensemble de données ne contient que des photos de deux classes et qu'il est quelque peu limité par rapport à d'autres études.

Samala et al. (Samala et al., 2019), ont développé un réseau neuronal convolutionnel profond (CNN) pour la classification des masses malignes et bénignes dans la tomosynthèse numérique du sein (DBT) en utilisant une approche d'apprentissage par transfert en plusieurs étapes qui a utilisé des données provenant de domaines auxiliaires similaires pour un réglage fin intermédiaire. L'AUC basée sur la vue sur l'ensemble de test pour l'apprentissage par transfert en une étape était de $0,85 \pm 0,05$ et s'est améliorée de manière significative ($p < 0,05$) à $0,91 \pm 0,03$ pour l'apprentissage en plusieurs étapes.

AL et al. (Al-antari et al., 2018), ont proposé un nouveau réseau neuronal convolutionnel profond (CNN) pour la détection, la segmentation et la classification des masses des mammographies comme étant bénignes ou malignes. Les masses détectées et segmentées ont été classées par CNN et ont atteint une précision globale de 95,64 %, une AUC de 94,78 %, une MCC de 89,91 % et un score F1 de 96,84 %, respectivement.

IV.6. Conclusion :

L'utilisation du diagnostic assisté par ordinateur (DAO) en mammographie a pris de l'importance en tant qu'outil permettant aux radiologues de contribuer à la détection précoce du cancer du sein. Les systèmes DAO peuvent aider à identifier les zones inquiétantes qui peuvent nécessiter une étude supplémentaire en analysant les mammographies et en extrayant des données telles que la forme, la densité et la texture de la masse. Le DAO a montré des résultats encourageants dans l'amélioration de la précision et de l'efficacité de la détection du cancer du sein et bien qu'elle ne puisse pas remplacer les compétences d'un radiologue qualifié, elle devrait jouer un rôle plus important dans le dépistage et le diagnostic du cancer du sein à l'avenir.

Chapitre

5

Combinaison de classifieurs pour le diagnostic des anomalies mammaires

V.1. Introduction

Dans la catégorisation du cancer du sein, la combinaison de modèles d'apprentissage profond et de modèles d'apprentissage automatique s'est révélée prometteuse. Les réseaux de neurones convolutifs (CNN) par exemple, excellent dans l'apprentissage de modèles et de caractéristiques détaillés à partir d'images médicales, tandis que les modèles ML offrent une capacité d'interprétation et de traitement d'un large éventail de sources de données. Ce chapitre présente un modèle combiné basé sur `efficient_b0` et SVM. Ce modèle fait des prédictions basées sur la texture des images afin d'automatiser la tâche de classification des tissus mammographiques.

V.2. Plateforme de développement

Les expérimentations que nous avons effectuées ont été exécutées sur la plateforme Kaggle qui se considère comme la plus grande communauté de science de données en ligne au monde ; où plus de 1000000 développeurs inscrits sur cette plateforme (Tauchert et al., 2020). Sa principale fonction est d'organiser des compétitions d'apprentissage automatique pour diverses organisations, tout en jouant le rôle d'un immense entrepôt de bases de données ouvertes et en encourageant la communication entre les scientifiques des données par le biais de forums de discussion et de carnets Jupyter ouverts.

La plateforme Kaggle offre des unités de traitement graphique connue sous le nom de GPU. Il est plus souvent appelé processeur graphique. Il s'agit d'un système informatique distinct qui fonctionne en parallèle avec l'unité centrale pour traiter les données graphiques (Futura, s. d.). Dans notre cas, nous avons utilisé le GPU NVIDIA T4, qui est une ressource précieuse pour toute une série de charges de travail et qui convient aux scientifiques des données et aux praticiens de l'apprentissage automatique grâce à son architecture Turing, son traitement efficace de l'inférence, son efficacité énergétique, sa mémoire suffisante et sa vitesse plus élevée. (GPU, s. d.)

V.3. Environnement logiciel

V.3.1. Langage de programmation Python

Python est un langage de programmation solide, simple à lire et à comprendre. En outre, il est utile pour les applications du monde réel tout en affichant la majorité de ces capacités qui sont partagées par de nombreux autres langages. En outre, il est entouré d'une

communauté de hackers importante et accueillante, il dispose d'une implémentation standard et est un logiciel libre.(Python, 2021)

V.3.2. Tensorflow

TensorFlow est un cadre logiciel qui a été initialement développé par des chercheurs de Google pour les calculs numériques utilisant des graphes de flux de données. Il est incroyablement adaptable et peut être utilisé pour développer, former et employer efficacement en production des réseaux neuronaux et d'autres modèles d'apprentissage automatique. Outre la prise en charge d'interfaces de langage supplémentaires en dehors de Python, TensorFlow fournit des API de bas niveau et de haut niveau comme Keras et Estimators. Les utilisateurs peuvent rapidement et à moindre coût combiner des techniques probabilistes avec des réseaux neuronaux profonds en utilisant la bibliothèque TFP et TensorBoard pour visualiser les métriques du modèle et de l'entraînement.(Pang et al., 2020)

V.3.3. Keras

La bibliothèque d'apprentissage profond la plus appréciée des professionnels est Keras, est une interface de programmation d'applications (API) basée sur TensorFlow qui facilite la configuration, l'entraînement et l'analyse des réseaux neuronaux. Elle simplifie la construction de réseaux profonds en faisant abstraction de nombreuses fonctionnalités sophistiquées de TensorFlow. Grâce à sa polyvalence et à sa capacité de personnalisation, Keras offre une variété de composants de réseaux neuronaux, notamment des couches épaisses, des couches convolutives, des couches récurrentes et des couches dropout. Pour une utilisation optimale, il gère dynamiquement les ressources CPU et GPU. Afin d'améliorer la convivialité, Keras propose en outre des implémentations de fonctions d'activation, d'optimiseurs, de formules métriques et de techniques de traitement des sessions d'entraînement.(Chicho & Sallow, 2021)

V.3.4. SKlearn

Scikit-learn est une boîte à outils d'apprentissage automatique complète, gratuite et libre écrite en Python, reconnue pour la variété des applications logicielles qu'elle prend en charge et la simplicité de son interface. Il se distingue des autres outils d'apprentissage automatique par sa couverture complète des approches d'apprentissage automatique, ses implémentations d'algorithmes optimisées pour les performances de calcul et le soutien important de la communauté pour la documentation et l'assurance qualité. Bien qu'il existe de nombreux algorithmes d'apprentissage automatique bien implémentés dans Scikit-learn, ils ne peuvent

pas tous être utilisés dans une application particulière. Scikit-learn doit être envisagé si les algorithmes pertinents sont déjà accessibles ; dans le cas contraire, un logiciel supplémentaire pourrait être nécessaire (Hao & Ho, 2019).

V.3.5. Mahotas

Mahotas est un module Python de vision par ordinateur qui fonctionne avec des tableaux Numpy .Pour le stockage des informations et les calculs simples, il utilise l'infrastructure fournie par Numpy. Mahotas évite de créer une nouvelle structure de données d'image en utilisant les tableaux Numpy. Plus de 100 fonctions sont disponibles pour une variété d'applications, telles que les décompositions en ondelettes, les calculs de caractéristiques locales, le filtrage d'images et les procédures morphologiques. Chaque fonction de la bibliothèque fonctionne séparément et adhère à une interface procédurale sans état global (Coelho, 2013).

V.3.6. Numpy

Numpy est une bibliothèque Python libre et gratuite, elle est la bibliothèque la plus fondamentale dans Python, elle se considère comme la plus puissante pour le calcul mathématique et la manipulation de données.

V.3.7. Pandas

Pandas est une bibliothèque de manipulation et d'analyse de données. Elle peut lire et écrire des feuilles de calcul, des fichiers CVS et effectuer toute une série d'autres opérations.

V.3.8. CV2

OpenCV est une boîte à outils performante, libre et gratuite, pour le traitement d'images numériques et la vision par ordinateur.

V.3.9. Matplotlib

Matplotlib est un logiciel de visualisation des tableaux 2D en Python. Matplotlib est une bibliothèque de visualisation de données multiplateforme construite sur les tableaux NumPy et conçue pour fonctionner avec la pile SciPy.

V.3.10. PIL

La bibliothèque Python PIL est utilisée pour la manipulation d'images. Elle permet d'ouvrir et de redimensionner les images des bases de données. (Kumar, 2020)

V.3.11. Gradio

La bibliothèque open-source Python Gradio est utilisée pour créer des composants d'interface utilisateur personnalisés basés sur des algorithmes d'apprentissage automatique. Gradio permet aux utilisateurs de tester des modèles en glissant et déposant des images, en écrivant du texte et en enregistrant des sons pour visualiser les résultats du modèle dans le navigateur. Gradio permet de : créer une démonstration de base basée sur un pipeline d'apprentissage automatique entraîné, obtenir un retour d'information en temps réel sur les performances du modèle, débogage interactif du modèle.(pycaret, s. d.)

V.4. Architecture du modèle

La Figure V-1 représente la méthodologie de combinaison de classifieurs pour la classification des anomalies mammaires. Il comprend les méthodes de classification classiques et à convolution.

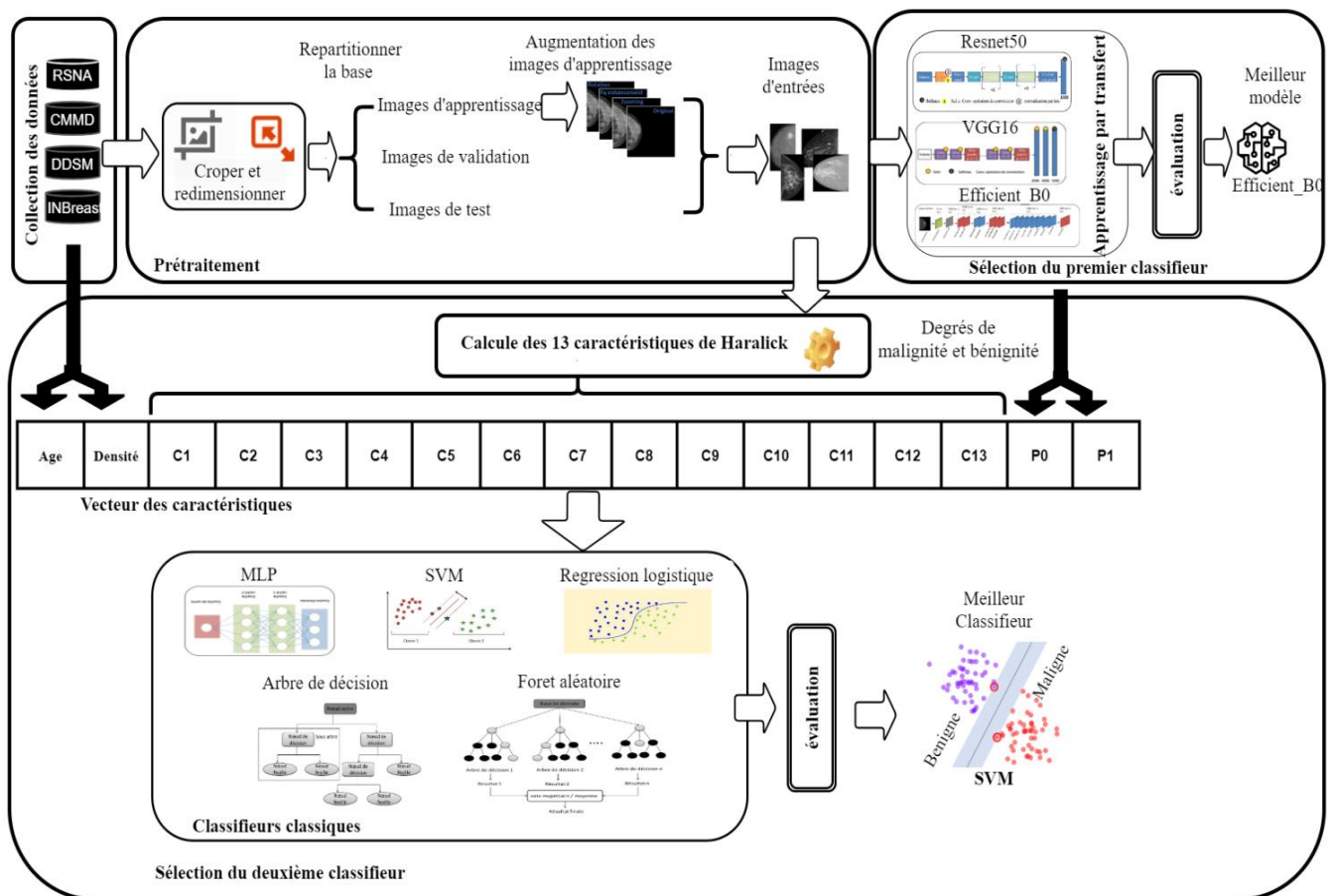


Figure V-1 : L'architecture générale de notre travail.

V.4.1. Collections des mammographies

Les images mammographiques utilisées dans les expérimentations de ce travail ont été collectées à partir de quatre bases de données distinctes. Ces bases de données nous fournissent des données vastes et étendues que nous pouvons utiliser pour entraîner et évaluer nos modèles.

V.4.1.1. RSNA dépistage du cancer du sein par mammographie

RSNA (Société de radiologie d'Amérique du Nord)(Carr et al., s. d.) est une base de données de mammographie qui comprend des images radiographiques de seins de femmes présentant des masses bénignes et malignes dans les deux incidences. Cette collection, contient des mammographies de 11 913 patientes. Cette base contient 53548 images bénignes et 1158 malignes comme montré la figure suivantes :

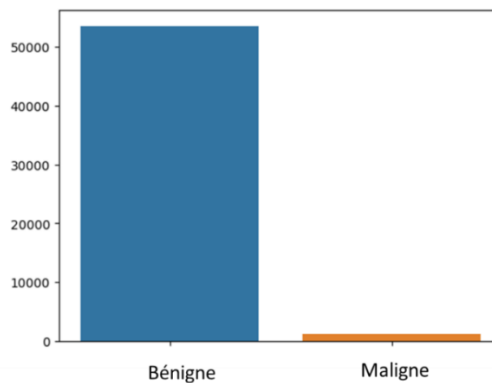


Figure V-2 : La distribution des images dans la base RSNA

V.4.1.2. Base de données chinoise sur la mammographie (CMMD)

CMMD est une base mammographique de 1 775 femmes chinoises atteintes d'une maladie mammaire bénigne ou maligne qui ont subi une mammographie entre juillet 2012 et janvier 2016. Les images mammographies de cette base sont réparties sur les deux classes (lésions bénignes ou malignes) confirmées par biopsie.

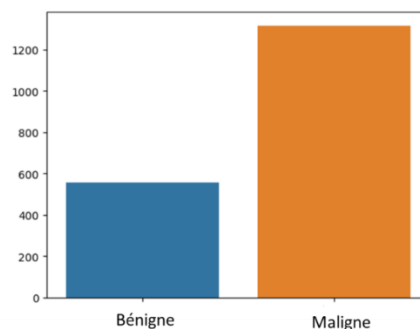


Figure V-3 : La distribution des images dans la base CMMD

Les données d'image ont été collectées à l'aide d'un appareil de mammographie GE Senographe DS.(Cui et al., 2021). Elle contient 4150 mammographies malignes et 1172 mammographies bénignes.

V.4.1.3. Mini-DDSM

Mini-DDSM est la version allégée de la célèbre base de données DDSM (Base de données numérique pour la mammographie de dépistage). Cependant, les fichiers d'images sont des fichiers d'extension jpeg et png qui sont générés à l'aide d'un logiciel spéciale. Cette base comprend 2 cas :2716 maligne , 2684 bénigne.(Lekamlage et al., 2021)

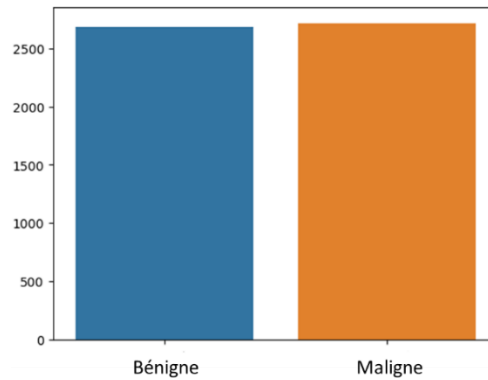


Figure V-4 : La distribution des images dans la base DDSM

V.4.1.4. INbreast

La base de données - INbreast - contient un total de 115 cas (410 images), dont 90 cas de femmes ayant les deux seins touchés (quatre images par cas) et 25 cas de patientes ayant subi une mastectomie (deux images par cas). Plusieurs types de lésions (masses, calcifications, asymétries et distorsions) ont été inclus. Les images de cette base sont réparties sur les deux classes (343 malignes et 76 bénignes).(Moreira et al., 2012)

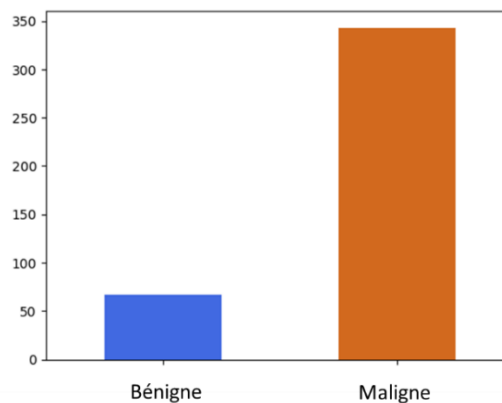


Figure V-5 : La distribution des images dans la base INbreast

La figure suivante représente des images bénigne et maligne des 4 bases de données utilisées

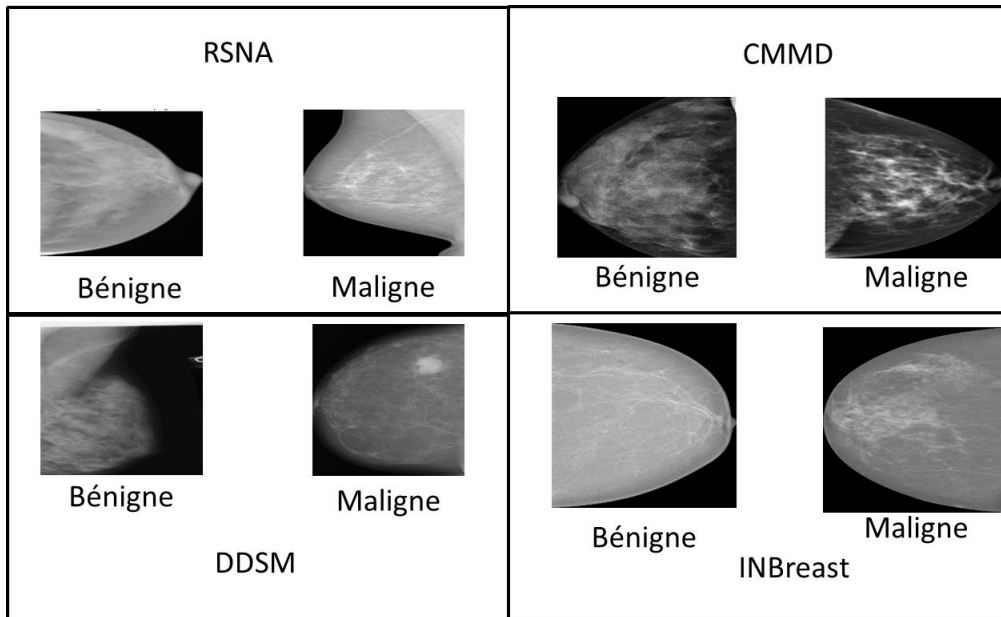


Figure V-6: Exemple des mammographies bénignes et malignes des 4 bases

En premier lieu la base de données RSNA choisi pour l'étude présente des classes déséquilibrées, ce qui signifie que la classe bénigne contenait beaucoup plus d'échantillons que la classe maligne. Pour résoudre ce problème, d'autres images cancéreuses ont été rassemblées et ajoutées à travers les bases de données DDSM, CMMD et INBreast. L'ensemble de données final comportait deux classes - maligne et bénigne et incluait des images provenant des quatre bases de données. En particulier, 8676 images étaient présentes pour la classe maligne et 8670 images étaient présentes pour la classe bénigne. Le problème de sur apprentissage constaté lors de l'utilisation d'une base de données déséquilibrées devait être réduit par l'utilisation des images provenant d'autres bases de données.

V.4.2. Prétraitement et préparation des mammographies

Dans cette étape deux méthodes de prétraitements (recadrage et le redimensionnement) ont été appliquées sur les quatre bases utilisées, par la suite la répartition des mammographies sur les différentes bases (apprentissage, validation et test). L'augmentation de données sera appliquée seulement sur la base d'apprentissage.

V.4.2.1. Recadrer et redimensionner

Recadrer une image consiste à extraire la zone d'intérêt qui s'appelle aussi région d'intérêt (ROI). La région d'intérêt est une notion essentielle dans l'analyse des images médicales. Il s'agit des zones importantes ou remarquables d'une image. L'objectif de la ROI est de rationaliser le traitement des images en se concentrant sur les zones importantes et en éliminant celles qui ne le sont pas.

- ❖ Tout d'abord, un filtrage gaussien est appliqué à l'image pour réduire le bruit. Ensuite, la technique de seuillage d'Otsu crée un masque binaire du sein en distinguant le premier plan (région du sein) de l'arrière-plan.
- ❖ Ensuite, les contours sont identifiés dans le masque binaire. Le contour ayant la plus grande surface, supposé correspondre à la région du sein, est sélectionné.
- ❖ Pour définir la région d'intérêt (ROI), un rectangle de délimitation est calculé pour entourer le contour sélectionné. Ce rectangle fournit les coordonnées (x, y) du coin supérieur gauche, ainsi que la largeur (w) et la hauteur (h) du rectangle.
- ❖ Enfin, la fonction renvoie les coordonnées (x, y, w, h) qui représentent les coordonnées et la taille du rectangle de délimitation, qui définit la région d'intérêt (ROI) dans l'image.

Cette figure montre ces étapes sur une image bénigne :

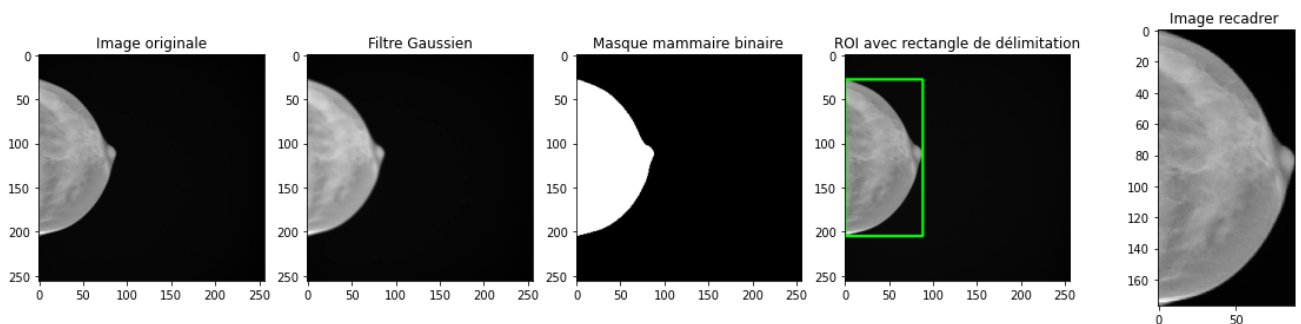


Figure V-7: Processus de recadrage de la région d'intérêt

Les images ont généralement des hauteurs et des largeurs différentes après le recadrage. Lorsque des images sont fournies à un réseau neuronal profond en vue d'une analyse ou d'une classification supplémentaire, cette variance dans la taille des images peut s'avérer problématique. Il est conseillé de redimensionner les images à une taille standard pour résoudre ce problème.

En redimensionnant les images à une taille uniforme, on s'assure que le réseau neuronal profond peut les traiter efficacement. La définition d'une taille prédéterminée, comme 256×256 pixels, permet d'obtenir un format d'entrée uniforme qui permet au réseau

d'apprendre et d'extraire des informations utiles des images. Cette étape est cruciale lors de l'utilisation de modèles d'apprentissage profond qui exigent des dimensions d'entrée fixes.

V.4.2.2. Augmentation et répartition des données

La base de données finale utilisée se compose d'un total de 16727 images, après avoir divisé nos données en ensembles d'apprentissage, de validation et de test avec des pourcentages de 70 %, 20 % et 10 %, nous avons constaté que la base d'apprentissage était composée de 5856 images bénignes et de 5856 images malignes.

Nous avons utilisé des techniques d'augmentation des données afin d'améliorer les performances et la généralisation du modèle. Nous avons utilisé deux méthodes en particulier l'inversion horizontale et la rotation aléatoire.

➤ **Rotation aléatoire** : En utilisant une gamme d'angles qu'on a fixée à 30 degrés, nous avons appliqué une rotation aléatoire aux images. Le modèle est plus résistant aux images tournés ou inclinés en raison de l'ajout, qui introduit des différences dans l'orientation des mammographies.

➤ **Retournement horizontal** : Pour retourner les images horizontalement, nous avons utilisé la fonction `horizontal_flip`. Cette méthode retourne les images sources horizontalement et les fait miroiter le long de l'axe vertical pour produire des variations. En particulier, lorsque l'orientation ou l'emplacement des éléments dans les images n'est pas essentiel pour la classification, cet ajout peut aider notre modèle à apprendre des caractéristiques à partir de mammographies retournées.

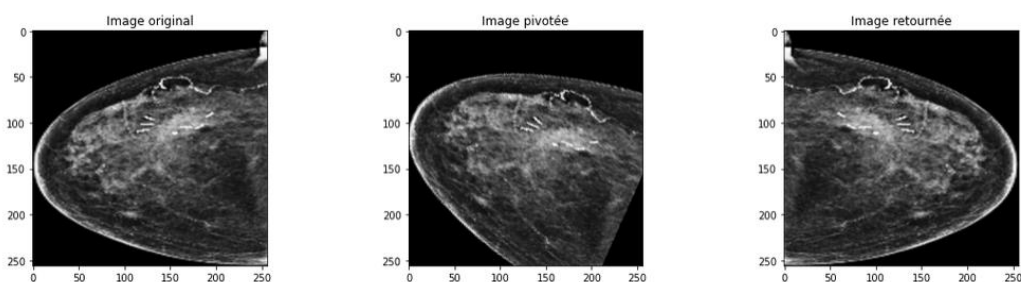


Figure V-8 : Exemple des 3 techniques d'augmentation sur une image maligne

L'application de ces techniques a augmenté le nombre des images dans la base d'apprentissage, voici le nombre d'images final dans chaque base :

Ensemble	Nombre des images
Apprentissage	23424
Validation	3348

Test	1674
------	------

Tableau V-1: Récapitulatif des images d'entraînement, validation et de test

Une base de données comprenant 23 424 échantillons a été utilisée tout au long de la procédure d'apprentissage. Cette base nous a permis d'affiner et d'entraîner nos modèles. La base de validation contient 3 348 échantillons, a également été mis de côté. Afin d'évaluer les performances du modèle et dans le cas échéant d'apporter des corrections, nous avons utilisé la base de validation, un sous-ensemble de données différent qui n'a pas été utilisé lors de l'apprentissage du modèle.

Une base de test distinct comprenant 1 674 échantillons a également été mise de côté. Cette base de test n'a pas été utilisée pour la création ou la validation du modèle, ce qui permet une évaluation équitable des performances de notre modèle. L'évaluation de la précision et de la généralisation sur cette base indépendante nous permettra d'en savoir plus sur les performances du modèle dans le monde réel et de prendre des décisions judicieuses quant à son application.

V.4.3. Sélection du premier classifieur

Cette étape nous permettra de sélectionner un modèle en se basant sur l'apprentissage par transfert, les modèles utilisées dans cette étape sont VGG16, ResNet50 et EfficientNet_B0.

V.4.3.1. Apprentissage par transfert

Nous avons introduit les images dans un réseau neuronal profond après avoir terminé les procédures de prétraitement et l'augmentation de données sur les images mammographiques. Trois modèles distincts - VGG16, ResNet50 et EfficientNet-B0 - ont été mis à l'épreuve dans ce cadre. Nous avons utilisée des versions pré-entraînées de ces modèles pour profiter des avantages de l'apprentissage par transfert.

Nous avons voulu comparer les performances de ces trois modèles lorsque on est les ai mis à l'épreuve pour voir lequel était le plus performant pour notre problème de classification. Grâce à l'apprentissage par transfert, on a pu utiliser les poids pré-entraînés de ces modèles pour bénéficier de leurs représentations apprises, ce qui a permis d'améliorer la précision et l'efficacité de ma tâche de classification. Les résultats sont présentés ci-dessus :

Métriques	Accuracy	Précision	Rappel
L'ensemble			
Apprentissage	92%	99%	88%
Validation	94%	99%	89%
Test	94%	99%	88%

Tableau V-2 : Résultat de VGG16

Métriques	Accuracy	Précision	Rappel
L'ensemble			
Apprentissage	93%	98%	87%
Validation	94%	99%	89%
Test	94%	99%	89%

Tableau V-3 : Résultat d'Efficientnet_b0

Métriques	Accuracy	Précision	Rappel
L'ensemble			
Apprentissage	89%	97%	81%
Validation	91%	98%	84%
Test	91%	98%	84%

Tableau V-4 : Résultat de Resnet50

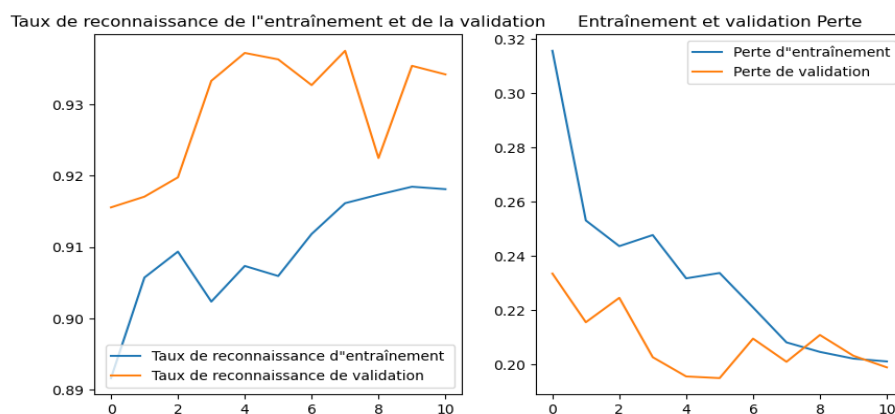


Figure V-9 : Évolution de l'apprentissage et la perte durant l'apprentissage de VGG16

Nous pouvons suivre l'évolution du taux de reconnaissance et de l'erreur des trois modèles sur un certain nombre d'époques en consultant le journal d'apprentissage du modèle

fourni. Les figures suivantes représentent le taux d'apprentissage et de l'erreur pour les bases d'apprentissage et de validation :

Dans VGG16 la perte a diminué durant l'entraînement, ce qui montre que le modèle s'améliore et apprend à faire des prédictions plus précises. L'amélioration progressive de la précision reflète la capacité du modèle à produire des prédictions précises.

Le modèle a atteint un taux de reconnaissance de 89,16% et une perte de 0,3157 dans les premières itérations. Au fur et à mesure de l'apprentissage, l'erreur est tombée à 0,2012 et le taux de reconnaissance est passé à 91,81%. Les performances du modèle sur la base de validation ont également été évaluées. Contrairement à l'augmentation du taux de reconnaissance sur la base validation de 93,42%, l'erreur de validation a diminué de 0,2336 à 0,19506. Ces résultats montrent que le modèle a continué à améliorer la précision de la prédiction et qu'il s'est bien adapté aux nouvelles données.

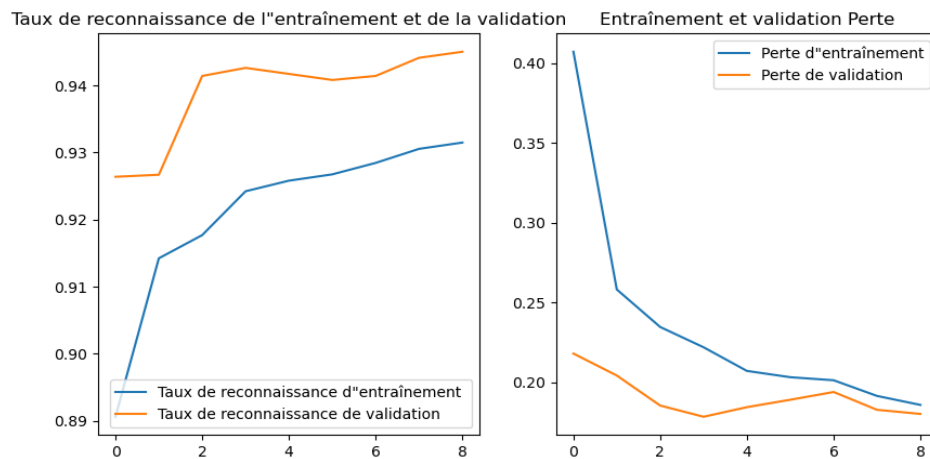


Figure V-10 : Évolution de l'apprentissage et la perte durant l'entraînement d'Efficientnet_b0

Les mesures d'apprentissage et de validation se sont toutes deux améliorées au fil des époques. L'erreur du modèle a continué à diminuer, ce qui suggère une meilleure convergence et des prévisions plus précises. Les performances sur la base d'apprentissage et de validation du modèle ont été suivies tout au long du processus d'apprentissage. Pour la phase d'apprentissage, la perte s'est améliorée, passant d'une valeur initiale de 0,4071 à 0,1858 dans la dernière époque, tandis que la perte de validation s'est améliorée de manière significative, passant d'une valeur initiale de 0,21802 à 0,1802. Après la neuvième époque, le processus de d'apprentissage a été interrompu en raison d'un arrêt prématuré. À ce moment-là, le modèle avait un taux de reconnaissance sur la base de validation de 94,50 % et un taux de reconnaissance sur la base d'apprentissage de 93,1 %.

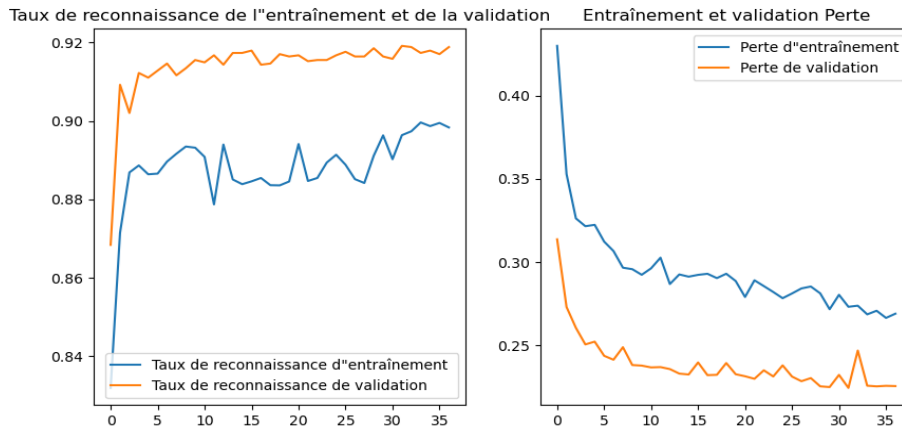


Figure V–11 : Évolution de l'apprentissage et la perte durant l'entraînement de Resnet50

Pour Resnet la perte est assez élevée (0,4298) et le taux de reconnaissance est moyenne (0,8319) au début de l'apprentissage (époque 1). La perte diminue progressivement au fur et à mesure de l'apprentissage et le taux de reconnaissance augmente. Par exemple, à l'époque 4, la perte est de 0,3216 et la précision de 88,87%. Cela montre que le modèle se développe et produit des prédictions plus précises. Au fur et à mesure de l'apprentissage, la perte du modèle diminue et sa précision augmente. Au cours de la quinzième époque, le taux de reconnaissance est de 88,39% et la perte de 0,2913. La perte et le taux de reconnaissance sur la base de validation présentent un schéma similaire à celui des mesures d'apprentissage, ce qui montre que le modèle se généralise bien aux nouvelles données.

V.4.3.2. Modèle élu

D'après les résultats, c'est le modèle EfficientNetB0 qui a obtenu les meilleurs résultats en matière de classification du cancer du sein. 18 couches composent le modèle pré-entraîné, puis d'autres couches ont été ajoutées. Pour préserver les caractéristiques apprises, les couches pré-entraînées ont été gelées. Une couche d'aplatissement, une couche dense avec 256 neurones et une fonction d'activation ReLU, une couche d'abandon pour éviter le sur-apprentissage avec un taux de 0,5, et une couche dense finale avec une fonction d'activation sigmoïde pour la classification binaire sont ajoutées comme la montre la Figure V–12.

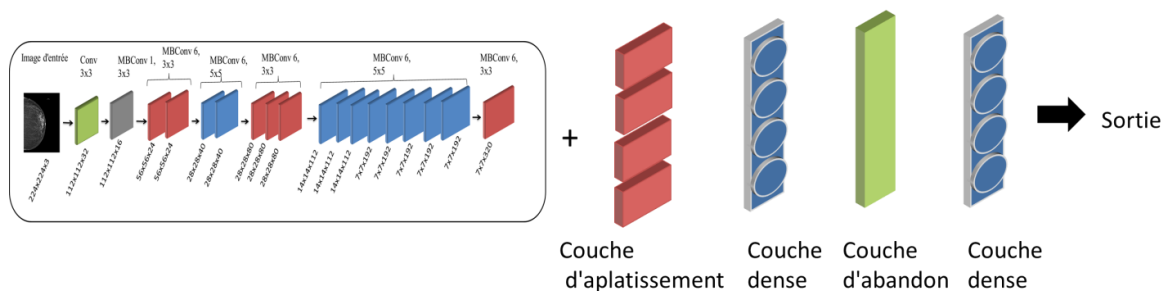


Figure V–12 : Modèle élu basé sur l'apprentissage par transfert

Le modèle a été compilé à l'aide d'un optimiseur Adam et d'un taux d'apprentissage de 0,001. L'entropie croisée binaire a été utilisée comme fonction de perte, et des critères d'évaluation tels que le rappel, l'accuracy et la précision ont également été inclus.

Des procédures d'arrêt précoce et de vérification du modèle ont été utilisées pour suivre la progression de l'apprentissage. Si la perte de validation ne diminuait pas, un arrêt anticipé avec une patience de 5 a été utilisé pour interrompre l'apprentissage. En se basant sur l'erreur de validation pendant l'apprentissage, la vérification du modèle a permis de conserver le meilleur modèle.

Le nombre d'étapes par époque a été déterminé en fonction de la taille du lot après l'apprentissage du modèle à l'aide d'un générateur. Les étapes de validation ont été établies de la même manière en utilisant les données de validation, qui ont également été fournies. Avec les rappels désignés, l'entraînement a été effectué sur 100 époques et le model arrête à la neuvième époque.

Sur la base de test, le taux de reconnaissance globale du modèle était de 94 %. Nous avons constaté que la précision, qui compte le pourcentage de cancers correctement identifiés parmi toutes les tumeurs anticipées, était de 99 %. La proportion de cancers classés avec précision parmi toutes les tumeurs réelles est connue sous le nom de rappel, souvent appelé sensibilité ou taux de vrais positifs. Il a été établi à 89 %.

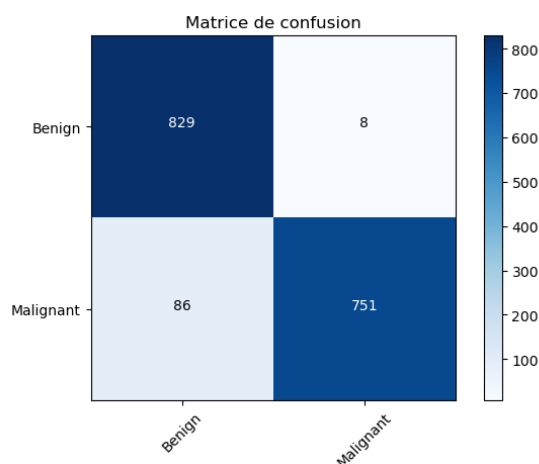


Figure V-13 : Matrice de confusion d'Efficientnet_b0

D'après la matrice de confusion représentée dans la Figure V-13, sur la base de test le modèle détecte correctement 829 anomalies bénignes et 751 malignes. Cependant, il a mal classé 86 mammographies malignes en tant que bénignes, et seulement 8 anomalies bénignes en tant que malignes.

V.4.4. Sélection du deuxième classifieur

Dans cette étape nous avons utilisé les 17 caractéristiques dont 13 d'haralick, l'Age et la densité provenant des étiquettes des quatre bases de données et les sorties de classification (P0 degré de bénignité et P1 degré de malignité) du classifieur élu à l'étape précédente. Toutes ces caractéristiques ont été fournies pour faire entraîner 5 classifieurs classiques

V.4.4.1. Extraction des caractéristiques

Les caractéristiques de texture Haralick des mammographies ont été calculées sur la base de la matrice de cooccurrence des niveaux de gris (SGLDM) pour quatre angles différents (0°, 45°, 90° et 135°), ce programme crée une base de caractéristiques de texture. Ces attributs fournissent des détails sur les motifs de texture visibles dans les mammographies.

Pour chaque angle, la fonction haralick génère un vecteur de 13 caractéristiques de texture. Ces caractéristiques sont les suivantes

C1 : Angular Second Moment (ASM)

C2 : Contrast

C3 : Correlation

C4: Sum of Squares: Variance

C5 : Inverse Difference Moment (IDM)

C6 : Sum Average

C7 : Sum Variance

C8 : Sum Entropy

C9 : Entropy

C10 : Difference Variance

C11 : Difference Entropy

C12: Information Measure of Correlation 1 (IMC1)

C13: Information Measure of Correlation 2 (IMC2)

Des caractéristiques telles que la douceur, le contraste et la complexité sont toutes capturées par ces caractéristiques. Nous pouvons en apprendre davantage sur les qualités texturales de l'image en identifiant et en évaluant ces aspects. Par conséquent, 13 caractéristiques de texture Haralick sont calculées pour chacun des quatre angles, soit un total de 52 valeurs distinctes. L'ensemble final de 13 caractéristiques de texture de Haralick moyennes est créé en calculant la moyenne de chacune des 13 caractéristiques.

V.4.4.2. Apprentissages des modèles classiques

Un fichier CSV est produit après l'extraction des caractéristiques de texture Haralick en leurs ajoutant :

1. L'Age du patient
2. La densité de la mammographie
3. P0 : degré de bénignité d'EfficientNet_B0
4. P1 ; degré de malignité d'EfficientNet_B0

L'utilisation des prédictions du modèle EfficientNetB0 comme entrée supplémentaire pour l'entraînement du deuxième classificateur est l'objectif de la fusion du premier et le deuxième classifieur. Cette combinaison permet d'obtenir une procédure de classification plus précise et peut-être plus efficace.

Ensuite le fichier CSV est divisé en une base d'apprentissage, de validation et de test. La base d'apprentissage est utilisée pour entraîner cinq modèles différents : Perceptron multicouche (MLP), Support Vector Machine (SVM), Régression logistique, Arbre de décision et Forêt aléatoire.

Dans un premier temps, nous avons expérimenté le second modèle seul sans le fusionner avec le premier classificateur (sans l'utilisation des degrés de bénignité et de malignité), et nous avons obtenu les résultats suivants :

	MLP	SVM	Regression logistique	Arbre de décision	Forêt aléatoire
Apprentissage	90%	95%	94%	95%	95%
Validation	90%	94%	94%	93%	94%
Test	90%	95%	95%	94%	94%
Matrice de confusion	$\begin{bmatrix} 790 & 38 \\ 115 & 714 \end{bmatrix}$	$\begin{bmatrix} 822 & 6 \\ 65 & 764 \end{bmatrix}$	$\begin{bmatrix} 823 & 5 \\ 72 & 757 \end{bmatrix}$	$\begin{bmatrix} 808 & 20 \\ 74 & 755 \end{bmatrix}$	$\begin{bmatrix} 825 & 3 \\ 80 & 749 \end{bmatrix}$

Tableau V-5 : Résultats des classifieurs classiques sans l'utilisation des degrés de malignité et de bénignité

Parmi ces résultats, nous observons que le modèle SVM a fourni le meilleur résultat.

Deuxièmement, nous avons combiné les deux modèles ensemble en ajoutant une colonne étiquette_prédite qui prend 1 si maligne 0 sinon, ensuite, et nous obtenons les résultats suivants :

	MLP	SVM	Regression logistique	Arbre de décision	Forêt aléatoire
Apprentissage	95%	95%	94%	95%	95%
Validation	95%	94%	94%	94%	94%
Test	94%	94%	94%	94%	94%
Matrice de confusion	$\begin{bmatrix} 824 & 5 \\ 79 & 750 \end{bmatrix}$	$\begin{bmatrix} 826 & 3 \\ 82 & 747 \end{bmatrix}$	$\begin{bmatrix} 819 & 10 \\ 80 & 749 \end{bmatrix}$	$\begin{bmatrix} 821 & 8 \\ 79 & 750 \end{bmatrix}$	$\begin{bmatrix} 824 & 5 \\ 78 & 751 \end{bmatrix}$

Tableau V-6 : Résultats des classifieurs classiques avec l'utilisation des résultats de classification d'EfficientNet_B0

Troisièmement, nous avons utilisé les deux modèles ensemble et au lieu d'utiliser des prédictions binaires (0 et 1), nous avons directement utilisé la sortie du premier modèle (degré de bénignité **P0** et degré de malignité **P1**) de chaque mammographie, et nous obtenons les résultats suivants :

	MLP	SVM	Regression logistique	Arbre de décision	Forêt aléatoire
Apprentissage	94%	95%	94%	94%	96%
Validation	95%	95%	95%	94%	95%
Test	96%	96%	96%	96%	96%
Matrice de confusion	$\begin{bmatrix} 825 & 4 \\ 57 & 772 \end{bmatrix}$	$\begin{bmatrix} 826 & 3 \\ 51 & 778 \end{bmatrix}$	$\begin{bmatrix} 823 & 6 \\ 55 & 774 \end{bmatrix}$	$\begin{bmatrix} 823 & 6 \\ 60 & 769 \end{bmatrix}$	$\begin{bmatrix} 825 & 3 \\ 66 & 763 \end{bmatrix}$

Tableau V-7 : Résultats de la combinaison des deux modèles avec la prédiction des degrés de malignité et bénignité

Les degrés de malignité et de bénignité sont calculés comme suit :

Si l'anomalie est classée Maligne : la sortie du premier classifieur (Efficientnet_B0) est considérée comme degré de malignité (P1), et le degré de bénignité P0 est calculée par la probabilité complémentaire qui = 1- P1

Si l'anomalie est classée bénigne : la sortie du premier classifieur (Efficientnet_B0) est considérée comme degré de bénignité (P0), et le degré de malignité P1 sera fixé à 0.

V.4.4.3. Modèle classique élu

Enfin, nous avons opté pour le modèle SVM, après l'avoir combiné avec le premier classificateur et avoir utilisé les degrés de malignité et de bénignité comme résultats du premier classifieur basé sur l'apprentissage par transfert. La méthode StandardScaler de SKlearn est utilisée pour prétraiter les données, en standardisant les caractéristiques par l'élimination de la moyenne et la mise à l'échelle de la variance unitaire.

Les paramètres suivants sont utilisés pour initialiser le classificateur SVM :

- ✓ Probabilité=True active les estimations de probabilité
- ✓ Random_state=0 établit une graine aléatoire reproductible
- ✓ C=100 fixe la valeur du paramètre de régularisation à 100,
- ✓ Kernel='rbf' désigne le kernel Fonction de base radiale
- ✓ Gamma='scale' établit la valeur gamma qui sera calculée sur la base de l'échelle des caractéristiques d'entrée.

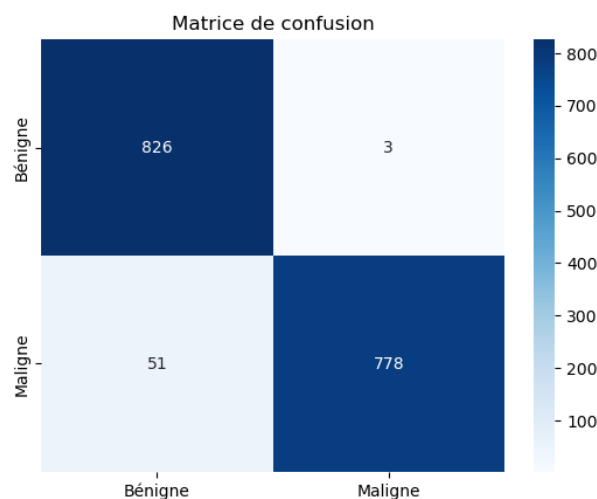


Figure V-14 : Matrice de confusion de SVM en utilisant les degrés de malignité et de bénignité

Sur la base de la matrice de confusion fournie :

- ✓ Vrai négatifs (**VN**) : Le nombre d'échantillons correctement classés comme négatifs (bénins) est représenté par la valeur 825. Dans ces situations, la prédiction du classificateur et l'étiquette réelle sont toutes deux négatives.
- ✓ Faux positifs (**FP**) : La valeur 3 indique le nombre d'échantillons pour lesquels un résultat positif (malin) a été prédit à tort. Il s'agit des cas où le classificateur prédit un résultat positif alors que l'étiquette réelle est négative.

- ✓ Faux positifs (**FP**):Le nombre d'échantillons classés par erreur comme négatifs (bénins) est de 57. Dans ces situations, même si l'étiquette réelle est positive, le classifieur prédit un résultat négatif.
- ✓ Vrai positifs (**VP**) : Le nombre d'échantillons correctement identifiés comme positifs (maligne) est représenté par la valeur 772. Dans ces situations, la prédiction du classificateur et l'étiquette réelle sont toutes deux positives.

Avec une précision d'apprentissage de 95,4 % et une précision de validation de 95,3 %, le modèle SVM a produit des résultats d'une grande précision à la fois sur les bases d'apprentissage et de validation. Cela suggère que le modèle a réussi à apprendre et à se généraliser aux nouvelles données.

Des mesures de performance sont ensuite calculées pour montrer la performance du modèle sur la base de test :

Métrique	Valeur
Taux d'apprentissage	96,7%
Précision	99,6%
Rappel	93,8%
F1 score	96,6%

Tableau V-8 : Evaluation des métriques SVM en utilisant les degrés de malignité et de bénignité

V.5. Déploiement de l'application :

Pour déployer notre application, nous avons opté pour la bibliothèque Gradio. A travers cette bibliothèque nous pouvons créer une interface conviviale où les utilisateurs peuvent entrer des données et obtenir des prédictions ou des résultats de votre modèle en temps réel.

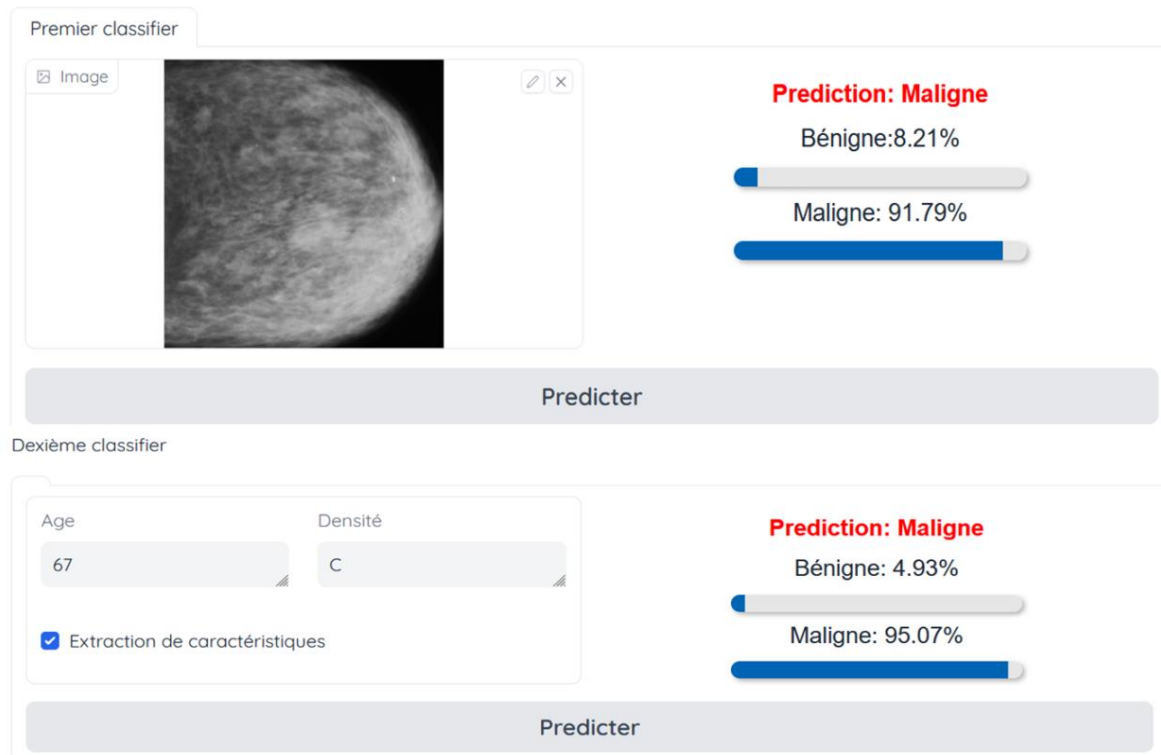


Figure V–15 : Interface graphique (Résultat de classification d’une image maligne)

Dans mon interface, il y a un espace pour parcourir une image, ainsi qu'un bouton en dessous pour lancer le processus de prédiction pour cette image. Sur le côté droit, l'interface affiche la visualisation du résultat de la prédiction.

En outre, il y a des champs de saisie pour l'âge, la densité et l'extraction de caractéristiques. Ces entrées, ainsi que le résultat du premier modèle, servent d'entrées au second modèle. Le résultat du second modèle sera affiché sur l'autre côté comme montré la Figure V–15.

V.6. Conclusion :

Notre travail a porté sur la classification du cancer du sein à l'aide du modèle EfficientNet_B0 combiné à un SVM. Tout au long de notre recherche, nous avons cherché à améliorer la robustesse et la généralisation de notre système de classification. Pour ce faire, nous avons intégré des données provenant de diverses sources, notamment les bases de données RSNA, DDSM, INBreast et CMMD. L'intégration de données provenant de diverses sources, a contribué à la richesse et à la diversité de notre ensemble de données. En incorporant des informations provenant de ces bases de données, nous avons pu entraîner notre modèle de classification sur un large éventail de cas de cancer du sein, augmentant ainsi

sa capacité à se généraliser et à classer avec précision des instances provenant de données inédites.

En utilisant le modèle EfficientNet_B0, nous avons tiré parti de ses capacités d'apprentissage en profondeur pour classer efficacement les cas de cancer du sein. La combinaison avec le SVM a encore amélioré la précision de la classification et a fourni des informations précieuses dans le processus de prise de décision.

En outre, l'inclusion de l'extraction de caractéristiques de texture, en particulier les caractéristiques de Haralick, nous a permis de capturer des informations importantes liées à l'arrangement spatial des régions tumorales. Cette technique d'extraction de caractéristiques a renforcé la capacité du système de classification à différencier les différents types de cancer du sein, ce qui a permis d'améliorer l'exactitude et la précision.

CONCLUSION GENERALE

Le cancer du sein est un problème de santé important dans le monde entier, qui touche de nombreuses personnes chaque année. La détection précoce et la classification précise du cancer du sein sont cruciales pour la réussite du traitement et l'amélioration des résultats pour les patientes. Ces dernières années, le domaine de l'intelligence artificielle (IA) a montré un grand potentiel dans l'aide au développement de systèmes intelligents pour la classification et le diagnostic du cancer du sein.

Le but de ce projet de fin d'étude est de proposer un système d'aide au diagnostic qui vise à tirer parti de la puissance de l'IA, en particulier des techniques d'apprentissage profond et d'apprentissage automatique, pour construire un système efficace de classification et de diagnostic du cancer du sein. Pour améliorer les performances du système, l'extraction de caractéristiques de texture, en particulier les caractéristiques de Haralick, a été incorporée en tant qu'intrant essentiel. Pour mener à bien cette recherche, une base de 16 727 mammographies a été utilisée en rassemblant des données provenant de quatre bases de données réputées : RSNA, DDSM, INBreast et CMMD. L'ensemble de données a fourni une gamme variée de cas de cancer du sein, permettant aux modèles d'apprendre et de se généraliser efficacement. Dans la première étape de ce projet de fin d'étude, trois modèles d'apprentissage profond ont été évalués. Après une analyse minutieuse, le modèle efficientnet_b0 a été sélectionné en raison de ses performances impressionnantes, atteignant un taux de précision de 93 %. Ce modèle a démontré sa capacité à classer avec précision les cas de cancer du sein en se basant sur les modèles et les caractéristiques appris, la deuxième phase de ce projet de recherche a intégré des techniques d'apprentissage automatique. Cinq modèles classiques ont été testés, et la machine à vecteur de support (SVM) s'est imposée comme le choix optimal. Avec un taux de précision de 95 %, le modèle SVM a démontré sa capacité à classer efficacement les cas de cancer du sein et à contribuer à un diagnostic précis.

La combinaison de l'apprentissage profond, de l'apprentissage automatique et de l'extraction de caractéristiques de texture s'est avérée être une approche puissante pour améliorer la précision et la fiabilité du système de classification et de diagnostic du cancer du sein. Les résultats de cette thèse contribuent au corpus croissant de connaissances dans le domaine des applications de l'IA à l'imagerie médicale, fournissant des indications précieuses pour de nouvelles avancées dans la détection et le traitement du cancer du sein.

En exploitant le potentiel de l'IA, nous pouvons continuer à améliorer le diagnostic du cancer du sein et, en fin de compte, avoir un impact positif sur les soins et les résultats pour les patients. Le système développé est prometteur pour aider les professionnels de la santé à classer avec précision les cas de cancer du sein, ce qui permet des interventions opportunes et des stratégies de traitement personnalisées. Avec la poursuite de la recherche et du développement, les systèmes basés sur l'IA ont le potentiel de révolutionner le domaine du diagnostic du cancer du sein et de contribuer à la lutte globale contre cette maladie prévalence.

Dans les travaux futurs, il est conseillé que les recherches portent sur l'intégration d'autres modalités d'imagerie, telles que la mammographie, l'échographie et l'IRM, pour améliorer la précision des systèmes de classification et de détection du cancer du sein. Envisager d'expérimenter différentes techniques d'extraction de caractéristiques ou d'incorporer des caractéristiques pertinentes supplémentaires pour améliorer les performances du modèle.

Références bibliographiques

- Ahuja, R., Chug, A., Gupta, S., Ahuja, P., & Kohli, S. (2020). Classification and Clustering Algorithms of Machine Learning with their Applications. In X.-S. Yang & X.-S. He (Éds.), *Nature-Inspired Computation in Data Mining and Machine Learning* (p. 225-248). Springer International Publishing. https://doi.org/10.1007/978-3-030-28553-1_11
- AlFayez, F., El-Soud, M. W. A., & Gaber, T. (2020). Thermogram Breast Cancer Detection : A Comparative Study of Two Machine Learning Techniques. *Applied Sciences*, *10*(2), Article 2. <https://doi.org/10.3390/app10020551>
- Alsheh Ali, M., Eriksson, M., Czene, K., Hall, P., & Humphreys, K. (2019). Detection of potential microcalcification clusters using multivendor for-presentation digital mammograms for short-term breast cancer risk estimation. *Medical Physics*, *46*(4), 1938-1946. <https://doi.org/10.1002/mp.13450>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning : Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, *61*, 101182. <https://doi.org/10.1016/j.ecoinf.2020.101182>
- Balleyguier, C., Ayadi, S., Van Nguyen, K., Vanel, D., Dromain, C., & Sigal, R. (2007). BIRADSTM classification in mammography. *European Journal of Radiology*, *61*(2), 192-194. <https://doi.org/10.1016/j.ejrad.2006.08.033>
- Castellino, R. A. (2005). Computer aided detection (CAD) : An overview. *Cancer Imaging*, *5*(1), 17. <https://doi.org/10.1102/1470-7330.2005.0018>
- Chamroukhi, F. (2013). *Classification supervisée : Les K-plus proches voisins*. Mémoire de fin d'étude, Université du Sud Toulon–Var.

Chicho, B. T., & Sallow, A. B. (2021). A Comprehensive Survey of Deep Learning Models Based on Keras Framework. *Journal of Soft Computing and Data Mining*, 2(2), Article 2.

Chollet, F. (2021). *Deep learning with Python* (Simon and Schuster).

Chugh, G., Kumar, S., & Singh, N. (2021). Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis. *Cognitive Computation*, 13(6), 1451-1470.

<https://doi.org/10.1007/s12559-020-09813-6>

Coelho, L. P. (2013). Mahotas : Open source software for scriptable computer vision. *Journal of Open Research Software*, 1(1), e3. <https://doi.org/10.5334/jors.ac>

Diallo, P. A. K. K., & Ju, Y. (2020). Accurate Detection of COVID-19 Using K-EfficientNet Deep Learning Image Classifier and K-COVID Chest X-Ray Images Dataset. *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, 1527-1531.

<https://doi.org/10.1109/ICCC51575.2020.9344949>

Farhan, A., & Kamil, M. (2020). Texture Analysis of Breast Cancer via LBP, HOG, and GLCM techniques. *IOP Conference Series Materials Science and Engineering*, 928, 072098.

<https://doi.org/10.1088/1757-899X/928/7/072098>

Filippov, V. A., Bobylev, A. N., Busygin, A. N., Pisarev, A. D., & Udovichenko, S. Yu. (2020). A biomorphic neuron model and principles of designing a neural network with memristor synapses for a biomorphic neuroprocessor. *Neural Computing and Applications*, 32(7), 2471-2485. <https://doi.org/10.1007/s00521-019-04383-7>

Fujita, H. (2020). AI-based computer-aided diagnosis (AI-CAD) : The latest review to read first. *Radiological Physics and Technology*, 13(1), 6-19. <https://doi.org/10.1007/s12194-019-00552-4>

Hafida, S. (2015). *Sélection de Données pour l'Apprentissage des Réseaux de Neurones, Arbres de Décision et les k-Plus Proches Voisins* [Universite mohamed boudiaf des sciences et de la technologie d'oran].

<https://www.ccdz.cerist.dz/admin/notice.php?id=00000000000000866306000903>

Hao, J., & Ho, T. K. (2019). Machine Learning Made Easy : A Review of *Scikit-learn* Package in Python Programming Language. *Journal of Educational and Behavioral Statistics*, 44(3), 348-361. <https://doi.org/10.3102/1076998619832248>

Hashemi Fath, A., Madanifar, F., & Abbasi, M. (2020). Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems. *Petroleum*, 6(1), 80-91. <https://doi.org/10.1016/j.petlm.2018.12.002>

Hossain, ABM, A., Nisha, Jannatul, K., Johora, & Fatematuj. (2023). Breast Cancer Classification from Ultrasound Images using VGG16 Model based Transfer Learning. *Modern Education and Computer Science Press*, 13(1), 12.

Iglesias, L. L., Bellón, P. S., del Barrio, A. P., Fernández-Miranda, P. M., González, D. R., Vega, J. A., Mandly, A. A. G., & Blanco, J. A. P. (2021). A primer on deep learning and convolutional neural networks for clinicians. *Insights into Imaging*, 12(1), 117. <https://doi.org/10.1186/s13244-021-01052-z>

Kadhim, R., & Kamil, M. (2022). Evaluation of Machine Learning Models for Breast Cancer Diagnosis Via Histogram of Oriented Gradients Method and Histopathology Images. *International Journal on Recent and Innovation Trends in Computing and Communication*, 10, 36-42. <https://doi.org/10.17762/ijritcc.v10i4.5532>

Kalel, M. D. S., Pisal, M. P. M., & Bagawade, M. R. P. (2016). *Color, Shape and Texture feature extraction for Content Based Image Retrieval System : A Study*. <https://www.semanticscholar.org/paper/Color%2C-Shape-and-Texture-feature-extraction-for-A-Kalel-Pisal/83a702dabdd394322a39a6c93cb1974093f53d23>

Kavya, R. (2015). *Feature Extraction Technique for Robust and Fast Visual Tracking : A Typical Review*. <https://www.semanticscholar.org/paper/Feature-Extraction-Technique-for-Robust-and-Fast-A-Kavya/2132b87331e6532bc7099ef63386750a96a73b2a>

Kim, B., Yuvaraj, N., Sri Preethaa, K. R., & Arun Pandian, R. (2021). Surface crack detection using deep learning with shallow CNN architecture for enhanced computation. *Neural Computing and Applications*, 33(15), 9289-9305. <https://doi.org/10.1007/s00521-021-05690-8>

- Kumar, V. (2020). Age Prediction using Image Dataset using Machine Learning. *International Journal of Innovative Technology and Exploring Engineering*, 8, 107-113. <https://doi.org/10.35940/ijitee.L1020.10812S319>
- Lekamlage, C. D., Afzal, F., Westerberg, E., & Cheddad, A. (2021). Mini-DDSM : Mammography-based Automatic Age Estimation. *2020 3rd International Conference on Digital Medicine and Image Processing*, 1-6. <https://doi.org/10.1145/3441369.3441370>
- Liu, R. (2013). *Feature extraction in classification*. <https://doi.org/10.25560/23634>
- Mahesh, B. (2019). *Machine Learning Algorithms -A Review*. <https://doi.org/10.21275/ART20203995>
- Mahmood, T., Li, J., Pei, Y., Akhtar, F., Imran, A., & Rehman, K. U. (2020). A Brief Survey on Breast Cancer Diagnostic With Deep Learning Schemes Using Multi-Image Modalities. *IEEE Access*, 8, 165779-165809. <https://doi.org/10.1109/ACCESS.2020.3021343>
- Mall, P. K., Singh, P. K., & Yadav, D. (2019). GLCM Based Feature Extraction and Medical X-RAY Image Classification using Machine Learning Techniques. *2019 IEEE Conference on Information and Communication Technology*, 1-6. <https://doi.org/10.1109/CICT48419.2019.9066263>
- Mathieu-Dupas, E. (2010). *Algorithme des k plus proches voisins pondérés et application en diagnostic*. 42èmes Journées de Statistique. <https://inria.hal.science/inria-00494814>
- Mathivet, V. (2017). *L'Intelligence Artificielle pour les développeurs—Concepts et implémentations en C#* (2e édition). Editions ENI.
- Mazher, A. N., Adwar, L., & Ali, A. H. (2011). Texture analysis using spatial gray level dependence matrix and the logical operators for Brodatz images. *Baghdad Science Journal*, 8(عدد خاص بمؤتمر الفيزياء 2011). <https://www.iasj.net/iasj/article/3679>
- Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., & Cardoso, J. S. (2012). INbreast : Toward a Full-field Digital Mammographic Database. *Academic Radiology*, 19(2), 236-248. <https://doi.org/10.1016/j.acra.2011.09.014>
- Murtaza, G., Shuib, L., Abdul Wahab, A. W., Mujtaba, G., Mujtaba, G., Nweke, H. F., Al-garadi, M. A., Zulfiqar, F., Raza, G., & Azmi, N. A. (2020). Deep learning-based breast

cancer classification through medical imaging modalities : State of the art and research challenges. *Artificial Intelligence Review*, 53(3), 1655-1720. <https://doi.org/10.1007/s10462-019-09716-5>

Mutlag, W. K., Ali, S. K., Aydam, Z. M., & Taher, B. H. (2020). Feature Extraction Methods : A Review. *Journal of Physics: Conference Series*, 1591(1), 012028. <https://doi.org/10.1088/1742-6596/1591/1/012028>

Otoum, Y., Liu, D., & Nayak, A. (2022). DL-IDS : A deep learning–based intrusion detection framework for securing IoT. *Transactions on Emerging Telecommunications Technologies*, 33(3), e3803. <https://doi.org/10.1002/ett.3803>

Pang, B., Nijkamp, E., & Wu, Y. N. (2020). Deep Learning With TensorFlow : A Review. *Journal of Educational and Behavioral Statistics*, 45(2), 227-248. <https://doi.org/10.3102/1076998619872761>

Patel, J. M., & Gamit, N. C. (2016). A review on feature extraction techniques in Content Based Image Retrieval. *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2259-2263. <https://doi.org/10.1109/WiSPNET.2016.7566544>

Perumal, V., Narayanan, V., & Rajasekar, S. J. S. (2021). Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. *Applied Intelligence*, 51(1), 341-358. <https://doi.org/10.1007/s10489-020-01831-z>

Python, W. (2021). *Python*. 24.

Ramadan, S. Z. (2020). Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms : A Review. *Journal of Healthcare Engineering*, 2020, e9162464. <https://doi.org/10.1155/2020/9162464>

Ramkumar, M., Babu, C. G., Kumar, K. V., Hepsiba, D., Manjunathan, A., & Kumar, R. S. (2021). ECG Cardiac arrhythmias Classification using DWT, ICA and MLP Neural Networks. *Journal of Physics: Conference Series*, 1831(1), 012015. <https://doi.org/10.1088/1742-6596/1831/1/012015>

- Roy, S. D., Das, S., Kar, D., Schwenker, F., & Sarkar, R. (2021). Computer Aided Breast Cancer Detection Using Ensembling of Texture and Statistical Image Features. *Sensors*, 21(11), Article 11. <https://doi.org/10.3390/s21113628>
- Ruchai, A. N., Kober, V. I., Dorofeev, K. A., Karnaukhov, V. N., & Mozerov, M. G. (2021). Classification of Breast Abnormalities Using a Deep Convolutional Neural Network and Transfer Learning. *Journal of Communications Technology and Electronics*, 66(6), 778-783. <https://doi.org/10.1134/S1064226921060206>
- Russell, R. (2020). *Machine Learning : Step-by-Step Guide To Implement Machine Learning Algorithms with Python*. (Knxb). http://repository.vnu.edu.vn/handle/VNU_123/92775
- Salau, A. O., & Jain, S. (2019). Feature Extraction : A Survey of the Types, Techniques, Applications. *2019 International Conference on Signal Processing and Communication (ICSC)*, 158-164. <https://doi.org/10.1109/ICSC45622.2019.8938371>
- Sarker, I. H. (2021a). Deep Learning : A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(6), 420. <https://doi.org/10.1007/s42979-021-00815-1>
- Sarker, I. H. (2021b). Machine Learning : Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Sassi, O., & Slima, M. (2012). Improved Spatial Gray Level Dependence Matrices for Texture Analysis. *International Journal of Computer Science and Information Technology*, 4, 209-219. <https://doi.org/10.5121/ijcsit.2012.4615>
- Shah, H. A., Saeed, F., Yun, S., Park, J.-H., Paul, A., & Kang, J.-M. (2022). A Robust Approach for Brain Tumor Detection in Magnetic Resonance Images Using Finetuned EfficientNet. *IEEE Access*, 10, 65426-65438. <https://doi.org/10.1109/ACCESS.2022.3184113>
- Sharma, N., Sharma, R., & Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings*, 2(1), 24-28. <https://doi.org/10.1016/j.gltp.2021.01.004>

- Sharma, S., Sharma, S., & Athaiya, A. (2020). ACTIVATION FUNCTIONS IN NEURAL NETWORKS. *International Journal of Engineering Applied Sciences and Technology*, 04(12), 310-316. <https://doi.org/10.33564/IJEAST.2020.v04i12.054>
- Singh, O. V., & Choudhary, P. (2019). A Study on Convolution Neural Network for Breast Cancer Detection. *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, 1-7. <https://doi.org/10.1109/ICACCP.2019.8882993>
- Tauchert, C., Buxmann, P., & Lambinus, J. (2020). Crowdsourcing Data Science : A Qualitative Analysis of Organizations' Usage of Kaggle Competitions. *Hawaii International Conference on System Sciences 2020 (HICSS-53)*. https://aisel.aisnet.org/hicss-53/cl/data_science_for_collaboration/3
- Taud, H., & Mas, J. F. (2018). Multilayer Perceptron (MLP). In M. T. Camacho Olmedo, M. Paegelow, J.-F. Mas, & F. Escobar (Éds.), *Geomatic Approaches for Modeling Land Change Scenarios* (p. 451-455). Springer International Publishing. https://doi.org/10.1007/978-3-319-60801-3_27
- Tian, D. (2013). *A Review on Image Feature Extraction and Representation Techniques*. <https://www.semanticscholar.org/paper/A-Review-on-Image-Feature-Extraction-and-Techniques-Tian/00d6fdf88581a11b24e4591848facd7fb28eb532>
- Wang, X., Li, J., & Qian, C. (2020). Semantic Label Prediction of Mammography Based on CC and MLO Views. *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, 1461-1465. <https://doi.org/10.1109/ICCC51575.2020.9345195>
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks : An overview and application in radiology. *Insights into Imaging*, 9(4), 611-629. <https://doi.org/10.1007/s13244-018-0639-9>
- Zhang, H., & Moon, S. K. (2021). Reviews on Machine Learning Approaches for Process Optimization in Noncontact Direct Ink Writing. *ACS Applied Materials & Interfaces*, 13(45), 53323-53345. <https://doi.org/10.1021/acsami.1c04544>
- Zhang, J. (2019). *Gradient Descent based Optimization Algorithms for Deep Learning Models Training* (arXiv:1903.03614). arXiv. <https://doi.org/10.48550/arXiv.1903.03614>

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, 109(1), 43-76. <https://doi.org/10.1109/JPROC.2020.3004555>

Webographie

Futura. (s. d.). *Définition | GPU - Graphical Processing Unit—Processeur graphique—Microprocesseur graphique | Futura Tech*. Futura. Consulté 5 juin 2023, à l'adresse <https://www.futura-sciences.com/tech/definitions/informatique-gpu-5739/>

GPU. (s. d.). *GPU T4 Vs GPU P100 | Kaggle | GPU | by Siddhartha | Medium*. Consulté 5 juin 2023, à l'adresse <https://siddhartha01writes.medium.com/gpu-t4-vs-gpu-p100-kaggle-gpu-cd852d56022c>

pycaret. (s. d.). *Supercharge Your ML with PyCaret and Gradio*. Consulté 5 juin 2023, à l'adresse <https://pycaret.gitbook.io/docs/learn-pycaret/official-blog/supercharge-your-ml-with-pycaret-and-gradio>

Carr, C., Kitamura, F., Kalpathy-Cramer, J., Mongan, J., Andriole, K., Vazirabad, M., Riopel, M., Ball, R., & Dane, S. (s. d.). *RSNA Screening Mammography Breast Cancer Detection AI Challenge (2023)*. Consulté 19 juin 2023, à l'adresse <https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/screening-mammography-breast-cancer-detection-ai-challenge>

Cui, C., Li, L., Cai, H., Fan, Z., Zhang, L., Dan, T., Li, J., & Wang, J. (2021). *The Chinese Mammography Database (CMMD) : An online mammography database with biopsy confirmed types for machine diagnosis of breast (Version 1)*. The Cancer Imaging Archive. Consulté 21 juin 2023, à l'adresse <https://doi.org/10.7937/TCIA.EQDE-4B16>