

ÉCOLE DOCTORALE MIPTIS
INSTITUT DENIS POISSON / LABORATOIRE D'INFORMATIQUE
MÉDICALE

THÈSE EN COTUTELLE INTERNATIONALE présentée par :

Tinhinane MEHDI

Soutenue le : 25 novembre 2025

Pour obtenir le grade de : Docteur de l'Université d'Orléans
et de l'Université de Béjaïa

Discipline/ Spécialité : Sciences et Technologies Industrielles – Informatique

Contributions aux techniques d'apprentissage
profond pour le diagnostic de la gonarthrose

THÈSE dirigée par :

M. JENNANE Rachid
Mme. ALOUI Soraya

PU, Université d'Orléans
MCF, Université de Béjaïa

RAPPORTEURS :

M. DJEMAL Khalifa
M. HARRAR Khaled

PU, Université d'Evry Paris-Saclay
PU, Université de Boumerdès

JURY :

M. AMROUN Kamel
M. DANIELLOU Richard
M. DJEMAL Khalifa
M. HARRAR Khaled
M. AMAD Mourad
M. NAIT-ALI Amine
Mme. ALOUI Soraya
M. JENNANE Rachid

PU, Université de Béjaïa, Président du jury
PU, Université d'Orléans
PU, Université d'Evry Paris-Saclay
PU, Université de Boumerdès
PU, Université de Bouira
PU, Université de Paris Est Créteil
MCF, Université de Béjaïa
PU, Université d'Orléans

Tinhinane MEHDI

Contributions aux techniques d'apprentissage profonds pour le diagnostic de la gonarthrose

Résumé :

L'arthrose est la forme d'arthrite la plus courante et touche les articulations, limitant la mobilité des personnes âgées. Ses causes et son évolution restent mal comprises. Il n'existe actuellement aucun traitement curatif.

Le diagnostic de l'arthrose repose sur la radiographie, une méthode accessible, mais peu sensible aux signes précoces de la maladie. En pratique clinique, cela rend le diagnostic précoce difficile.

Cette thèse vise à améliorer la détection précoce de l'arthrose du genou (gonarthrose) en utilisant l'apprentissage profond sur des radiographies. Son objectif est de dépasser les limites des méthodes traditionnelles de diagnostic.

La première contribution de cette thèse explore l'impact de l'emplacement des régions d'intérêt (ROIs) sur la classification des radiographies. L'utilisation d'autoencodeurs empilés parcimonieux montre qu'un placement optimal des ROIs améliore significativement la détection de la gonarthrose.

La deuxième contribution combine des informations de forme et de texture dans un modèle d'apprentissage profond afin d'améliorer la détection précoce de l'arthrose. En intégrant ces caractéristiques dans le modèle Compact Bilinear Pooling (CBP), une précision de 91,67 % a été atteinte, démontrant ainsi l'efficacité de cette approche.

Mots clés : Apprentissage automatique, Traitement d'image, Radiographies, Arthrose.

Contributions to deep learning techniques for the diagnosis of knee osteoarthritis

Summary :

Osteoarthritis is the most common form of arthritis, affecting the joints and limiting mobility in older people. Its causes and progression are still poorly understood and there is currently no cure.

The diagnosis of osteoarthritis is based on X-rays, which are accessible but not very sensitive to early signs of the disease. This makes early diagnosis difficult in clinical practice.

This work aims to improve the early detection of knee osteoarthritis using deep learning on plain radiographies. The aim is to overcome the limitations of traditional diagnostic methods.

The first contribution of this thesis investigates the impact of the location of regions of interest (ROIs) on the classification of radiographs. The use of parsimonious stacked autoencoders shows that optimal placement of ROIs significantly improves the detection of knee osteoarthritis.

The second contribution combines shape and texture information in a deep learning model to improve the early detection of osteoarthritis. By integrating these features into the Compact Bilinear Pooling (CBP) model, an accuracy of 91.67% was achieved, demonstrating the effectiveness of this approach.

Keywords: Deep Learning, Image Processing, Radiographs, Osteoarthritis.

REMERCIEMENTS

Je tiens à exprimer ma profonde gratitude à toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de cette thèse, menée en cotutelle entre l'Université d'Orléans (France) et l'Université de Béjaïa (Algérie).

Je souhaite adresser mes remerciements les plus sincères et respectueux à mon directeur de thèse, le Professeur JENNANE Rachid, de l'Université d'Orléans et du laboratoire IDP, pour la confiance qu'il m'a accordée et l'opportunité qu'il m'a offerte de réaliser ce travail en cotutelle. Je lui suis profondément reconnaissante pour la qualité de son encadrement scientifique, sa disponibilité, ses conseils avisés et son accompagnement tout au long de cette recherche. Au-delà de l'aspect scientifique, son soutien humain et professionnel m'a permis de progresser, de gagner en maturité et d'évoluer tant sur le plan académique que personnel. Son engagement, son exigence bienveillante et sa vision ont été déterminants dans l'aboutissement de cette thèse.

J'exprime également ma vive reconnaissance à ma co-directrice, le Docteur ALOUI Soraya, de l'Université de Béjaïa et du laboratoire LIMED, pour son encadrement attentif, son soutien constant et son aide précieuse, notamment sur le plan administratif. Étant la majeure partie du temps en France, sa disponibilité et son implication ont été essentielles au bon déroulement des démarches nécessitant ma présence en Algérie. Je la remercie sincèrement pour sa patience, son efficacité et son accompagnement.

Je remercie par ailleurs les membres du jury pour l'honneur qu'ils me font en acceptant d'évaluer ce travail, ainsi que pour l'intérêt qu'ils lui ont porté.

Mes remerciements s'adressent également à l'ensemble des enseignants, chercheurs et membres des laboratoires IDP et LIMED, pour l'environnement scientifique stimulant, les échanges enrichissants et le soutien apporté tout au long de cette recherche.

Enfin, je tiens à exprimer toute ma gratitude à ma famille et à mes proches pour leur soutien indéfectible, leur patience et leurs encouragements constants, qui m'ont permis de mener ce travail à son terme.

Abstract

Osteoarthritis is the most common form of arthritis, affecting the joints and limiting mobility in older people. Its causes and progression are still poorly understood and there is currently no cure. The diagnosis of osteoarthritis is based on X-rays, which are accessible but not very sensitive to early signs of the disease. This makes early diagnosis difficult in clinical practice. This work aims to improve the early detection of knee osteoarthritis using deep learning on plain radiographies. The aim is to overcome the limitations of traditional diagnostic methods. The first contribution of this thesis investigates the impact of the location of regions of interest (ROIs) on the classification of radiographs. The use of parsimonious stacked autoencoders shows that optimal placement of ROIs significantly improves the detection of knee osteoarthritis. The second contribution combines shape and texture information in a deep learning model to improve the early detection of osteoarthritis. By integrating these features into the Compact Bilinear Pooling (CBP) model, an accuracy of 91.67% was achieved, demonstrating the effectiveness of this approach.

Keywords

Deep Learning, Image Processing, Radiographs, Osteoarthritis.

Résumé

L'arthrose est la forme d'arthrite la plus courante et touche les articulations, limitant la mobilité des personnes âgées. Ses causes et son évolution restent mal comprises. Il n'existe actuellement aucun traitement curatif. Le diagnostic de l'arthrose repose sur la radiographie, une méthode accessible, mais peu sensible aux signes précoces de la maladie. En pratique clinique, cela rend le diagnostic précoce difficile. Cette thèse vise à améliorer la détection précoce de l'arthrose du genou (gonarthrose) en utilisant l'apprentissage profond sur des radiographies. Son objectif est de dépasser les limites des méthodes traditionnelles de diagnostic. La première contribution de cette thèse explore l'impact de l'emplacement des régions d'intérêt (ROIs) sur la classification des radiographies. L'utilisation d'autoencodeurs empilés parcimonieux montre qu'un placement optimal des ROIs améliore significativement la détection de la gonarthrose. La deuxième contribution combine des informations de forme et de texture dans un modèle d'apprentissage profond afin d'améliorer la détection précoce de l'arthrose. En intégrant ces caractéristiques dans le modèle Compact Bilinear Pooling (CBP), une précision de 91,67% a été atteinte, démontrant ainsi l'efficacité de cette approche.

Mots-clés

Apprentissage automatique, Traitement d'image, Radiographies, Arthrose.

Table des matières

1	Introduction	1
1.1	Contexte	3
1.2	Objectifs de la thèse	4
1.3	plan du Mémoire	5
2	La gonarthrose	7
2.1	Définition et caractéristiques de l'arthrose	9
2.1.1	Prévalence et épidémiologie	9
2.1.2	Localisation : Zones les plus touchées	10
2.2	Cas de la gonarthrose	11
2.2.1	Structure et composition de l'os	11
2.3	Diagnostic de la gonarthrose	13
2.3.1	Les critères de Kellgren and Lawrence (KL)	15
2.3.2	Les critères de Osteoarthritis Research Society International (OARSI)	15
2.4	Limitations	16
2.5	Base de données publiques sur la gonarthrose	17
2.5.1	Base de données Osteoarthritis Initiative (OAI)	18
2.5.1.A	Carctéristiques et Objectifs de l'étude OAI	19
2.5.1.B	Protocole d'acquisition des données OAI	20
2.5.2	Base de données Multicenter Osteoarthritis Study (MOST)	21
2.5.2.A	Caractéristiques et objectifs de l'étude MOST	22
2.6	Conclusion	23
3	État de l'Art sur les approches de detection de la gonarthrose	24
3.1	Intoduction	25
3.2	Méthodes assistées par ordinateur pour la classification de la gonarthrose	25
3.3	Le diagnostic automatique de la gravité de la gonarthrose à partir de radiographies	26
3.3.1	Segmentation des radiographies	26
3.3.2	Classification des grades de la gonarthrose	28

3.3.3	Méthodes classiques	29
3.3.3.A	Approches basées sur l'analyse de la texture	29
3.3.3.B	Approches basées sur les caractéristiques de forme	35
3.3.3.C	Approches basées sur les caractéristiques de forme et de texture	36
3.3.3.D	Approches basées sur des techniques d'apprentissage profond	39
4	Introduction à l'Apprentissage Profond	47
4.1	Introduction	49
4.2	Modèles de Base en Deep Learning (DL)	52
4.2.1	Autoencodeurs	52
4.2.1.A	Variantes des Autoencodeurs	53
4.2.2	Réseaux de Neurones Convolutifs (CNNs)	54
A –	Avancées récentes dans les CNNs	56
4.2.3	Recurrent Neural Network (RNN)	56
A –	Avancées récentes dans les RNNs	58
4.2.4	Transformers et Vision Transformers (ViT)	58
4.2.4.A	Vision Transformer (ViT)	60
4.2.5	Graph Neural Network (GNN)	61
4.3	Conclusion	64
5	Impact de la localisation des Région d'Intérêts (ROI) sur la détection précoce de la gonarthrose	65
5.1	Introduction et Motivation	67
5.2	Segmentation des radiographies et extraction des Regions Of Interest (ROI)s	68
5.2.1	Bonefinder pour la segmentation automatique du plateau tibio-fémoral.	68
5.2.2	Apprentissage automatique des caractéristiques à l'aide du modèle Stacked Sparse AutoEncoder (SSAE)	71
5.2.3	Extraction des caractéristiques	72
5.2.4	Classification de la sévérité de la gonarthrose	72
5.3	Données et paramètres expérimentaux	73
5.3.1	Ensemble de données	74
5.3.2	Application de l'approche	74
5.4	Résultats de la classification et discussion	74
5.5	Forces et limites de l'étude	76
5.6	Conclusion	76

6	Combinaison de caractéristiques de forme et de texture pour le diagnostic de la gonarthrose	79
6.1	Introduction	81
6.2	Approche Proposée	81
6.2.1	Extraction des caractéristiques de forme	81
6.2.2	Extraction des caractéristiques de texture	82
6.2.3	Compact Bilinear Pooling (Compact Bilinear Pooling (CBP))	85
6.3	Données et Paramètres Expérimentaux	87
6.3.1	Dataset	87
6.3.2	Extraction de Caractéristiques	87
6.3.3	Combinaison de Caractéristiques	88
6.4	Résultats expérimentaux et discussion	88
6.4.1	Sélection des modèles appropriés pour l’approche proposée	89
6.4.2	Effets des stratégies de combinaison sur les modèles DL de l’état de l’art	90
6.4.3	Étude ablative et évaluation des performances des stratégies proposées	91
6.4.4	Classification du grade douteux KL-1	92
6.4.5	Synthèse des résultats et discussion	92
6.5	Forces et limites de l’étude	95
6.6	Conclusion	96
7	Conclusion générale	97
7.1	Conclusion	99
7.2	Résumé des Contributions	100
7.3	Perspectives	101
	Bibliography	105

Table des figures

2.1	Illustration schématique montrant les principales caractéristiques pathologiques de l'arthrose du genou. Le côté gauche représente une articulation normale, tandis que le côté droit représente une articulation atteinte.(source [https://gideonburstein.co.il/])	12
2.2	Illustration de la région d'articulation du genou(source[https://www.cortho.org/knee/])	12
2.3	Différents grades de la gonarthrose selon Kellegren & Laurence. Les images sont extraites de la base de donnée OAI.	15
2.4	L'image présente des radiographies de deux genoux : un genou normal à gauche et un genou atteint d'arthrose sévère (OA) à droite. La comparaison met en évidence des caractéristiques distinctes, notamment les compartiments articulaires, le rétrécissement de l'espace articulaire (JSN) et la présence d'ostéophytes généralement associés à l'arthrose [Bayramoglu et al., 2020].	17
2.5	Caractéristiques des sujets de la cohorte OAI. Ces valeurs sont celles De visite 0 jusqu'au Visite 6, les modalités sont susceptible de varier d'une année a l'autre.	18
2.6	Disponibilité des images radiographiques et IRM aux différents points de suivi, accompagnée de captures d'écran des dossiers de stockage des données.	20
2.7	Disponibilité des données collectées par visite dans la base MOST	21
2.8	Caractéristiques des sujets de la base de donnée MOST. Ces valeurs sont celles De visite 0 jusqu'au Visite 5, les modalités sont susceptible de varier d'une année a l'autre.	22
3.1	Les ROIs de l'os trabéculaire latéral et médial utilisées dans [Woloszynski et al., 2010]. . .	30
3.2	Schéma décrivant l'emplacement des ROIs utilisées dans [Hirvasniemi et al., 2014]. Les rectangles noirs aux lignes continues représentent les ROIs de la plaque osseuse sous-chondrale. Les carrés noirs aux lignes pointillées représentent les ROIs de l'os trabéculaire sous-chondral tibial. LesROIs du fémur sont représentées par des rectangles noirs aux lignes mixtes (point-tiret)	30

3.3	Les ROIs ont été identifiées de manière cohérente sur des images avec différents niveaux de post-traitement. Deux ROIs ont été positionnées sous l'interface cartilage-os dans les plateaux tibiaux médial et latéral, tandis que seize autres ont été placées sous la zone dense de l'os trabéculaire sous-chondral. [Hirvasniemi et al., 2019]	31
3.4	Illustration du patchwork des ROIs extraites sur l'ensemble des radiographies de la base OAI [Janvier et al., 2017]	32
3.5	Ancrage des ROIs [Hladůvka et al., 2017].	33
3.6	ROIs sélectionnées à l'aide d'un algorithme semi-automatique. Les points rouges représentent les repères anatomiques définis manuellement et la ligne pointillée indique le bord tibial [Riad et al., 2018].	34
3.7	La zone sous-chondrale sélectionnée pour les analyses texturales comprenait 4 régions d'intérêt (ROI) dans le tibia proximal et une sur chaque condyle du fémur distal [Ljuhar et al., 2018]	34
3.8	Une ROI représentative du tissu trabéculaire tibial et des différents pré-traitements appliqués dans [Brahim et al., 2019]. ROI de (a) un sujet témoin et (b) un patient atteint d'arthrose (OA).	35
3.9	Les 105 points pris en compte par le modèle statistique de forme [Haverkamp et al., 2011] pour analyser la forme des radiographies du genou.	36
3.10	Paramètres mesurés sur les radiographies du genou, de gauche à droite : largeur de l'espace articulaire (JSW), angle de convergence de la ligne articulaire (JLCA), hauteurs des éminences et intensités de l'os sous-chondral, utilisé par [Rayegan et al., 2023]	37
3.11	Interface utilisateur du logiciel ODIA développé par [Rayegan et al., 2023]	37
3.12	Exemple d'échantillonnage d'une région circulaire pour calculer la signature fractale (à gauche). Exemple des 74 points de repère et des 4 points calculés à partir de ceux-ci, délimitant la région d'intérêt (ROI) utilisée pour évaluer la texture du tibia (à droite). où x est un vecteur représentant la forme dans un cadre de référence, $\hat{x}$ est la forme moyenne, P est l'ensemble des vecteurs propres correspondant aux t valeurs propres les plus élevées, qui décrivent différents modes de variation, et b est un ensemble de valeurs de paramètres. [Thomson et al., 2015]	38
3.13	Un exemple des 102 points de repère utilisés pour construire le mode de forme dans [Minicullo et al., 2017]	38
3.14	Pipeline proposé par Antony et al. pour la détection de la gonarthrose [Antony et al., 2017]	39
3.15	Le pipeline de classification proposé par [Tiulpin et al., 2018]	40

3.16	Illustration du pipeline d'évaluation de la gravité de l'articulation du genou proposé dans [Chen et al., 2019]. Il comprend la détection de l'articulation du genou et la classification du grade KL du genou.	41
3.17	(a) Les cinq ROIs extraites. (b) Architecture du modèle d'apprentissage DRAE proposée par [Nasser et al., 2020].	42
3.18	Le modèle Siamois utilisé par [Nguyen et al., 2020] pour prédire les grades KL correspondant aux côtés de l'articulation du genou (latéral et médial) des entrées.	43
3.19	Pipeline de prédiction de la gravité de la gonarthrose proposé par [Gu et al., 2022]	44
3.20	Architecture du modèle KOA-CCTNet proposé par [Jahan et al., 2024]	45
4.1	Représentation schématique d'un MLP avec deux couches cachées. [Chan et al., 2023] . .	51
4.2	Illustration du réseau AE traditionnel [Wang et al., 2022].	52
4.3	Architecture basic d'un CNN [Gu et al., 2019]	54
4.4	Architecture d'un RNN (source : https://zemdalk.github.io/2022/11/24/RNN,-Attention,-Transformer,-BERT.html)	57
4.5	Architecture des transformers et son mécanisme d'auto attention [Vaswani et al., 2017] . .	59
4.6	Graphe et le voisinage d'un nœud. L'état x_1 du nœud 1 dépend de l'information contenue dans son voisinage. [Scarselli et al., 2009]	62
5.1	Pipeline de l'approche proposée.	68
5.2	Image radiographique des deux genoux (A), Délimitation des contours des os fémoral et tibial à l'aide de l'outil BoneFinder (B), localisation de l'articulation du genou en rouge (C), détection de cinq contours mis en évidence respectivement en rouge, vert, bleu, jaune et noir, et extraction de 7 ROIs sur chaque contour (D).	69
5.3	Image radiographique des deux genoux (A), délimitation des contours des os fémoral et tibial à l'aide de l'outil BoneFinder (B).	70
5.4	Image radiographique des deux genoux (A), délimitation des contours des os fémoral et tibial à l'aide de l'outil BoneFinder (B).	71
6.1	Organigramme de l'approche proposée combinant les modèles Discriminative Regularized Auto-Encoder (DRAE) et siamois avec le CBP pour la classification des premiers stades de la gonarthrose.	82

6.2	L'architecture du modèle Deep Siamese. Le max-pooling (2×2) est représenté en bleu, les couches de convolution, la normalisation par lot (batch normalization) et le bloc ReLU en bleu clair. La couche softmax pour concaténer les sorties des deux réseaux est en orange. Les chiffres à l'intérieur des couches de convolution, de normalisation par lot et du bloc ReLU indiquent le nombre de filtres de convolution ; les chiffres au-dessus des couches de convolution correspondent à leurs paramètres ($K \times K, S$), où K représente la taille du filtre et S le pas de convolution.	83
6.3	Sélection des ROIs pour l'extraction des caractéristiques de texture. (a) Une image radiographique typique du genou issue de la base OAI. (b) Zone distale extraite avec les cinq ROIs récupérées.	83
6.4	Architecture du modèle d'apprentissage DRAE. X_1 et X_2 sont les ensembles d'entrées du réseau pour les deux classes (sujet normal et OAI minimale), X'_1 et X'_2 sont les entrées reconstruites, et Z_1 et Z_2 sont les vecteurs de caractéristiques appris en minimisant l'erreur de reconstruction et en maximisant la distance entre les caractéristiques des deux classes.	84
6.5	Structure du modèle d'apprentissage CBP.	86

Liste des tableaux

2.1	Épidémiologie des pathologies arthrosiques selon la zone concernée.	10
2.2	Résumé différentes des bases de données publique de la gonarthrose	19
5.1	Précisions de classification obtenues (%) avec chaque ROI à partir des différents contours.	73
5.2	Précisions obtenues (%) en utilisant la technique Max-voting.	73
6.1	Comparaison des modèles d'apprentissage sélectionnés basés sur la forme et la texture pour la classification (KL-0 vs. KL-2).	89
6.2	Effets du DRAE combiné à la stratégie CBP pour améliorer les modèles DL de State Of The Art (SOTA) sélectionnés.	90
6.3	Étude ablative et évaluation des performances obtenues en utilisant les caractéristiques de forme et de texture du genou provenant de (R1, R2, ... et R5) combinées à l'aide du CBP et et d'une simple concaténation.	93
6.4	Performances de classification obtenues pour (KL-0 vs. KL-1) et (KL-1 vs. KL-2) en combinant les caractéristiques du DRAE et du modèle Deep Siamese avec la stratégie CBP. .	94

Liste des algorithmes

1	Compact Bi-linear Pooling (CBP)	87
---	---	----

Liste des acronymes

Acc	Accuracy
AE	Auto-Encoder
AI	Intelligence Artificielle
AUC	Area Under the Curve
IMC	Indice de Masse Corporelle
BLSA	Baltimore Longitudinal Study of Aging, une base de données de radiographies du genou
CAD	Computer-Aided Diagnosis
CAE	Contractive AutoEncoder
CNN	Convolutional Neural Network
CR	Computed Radiography
DL	Deep Learning
DRAE	Discriminative Regularized Auto-Encoder
F1	F1-score
FSA	Fractal Signature Analysis
FCM	Fuzzy C-Means
HOG	Histogram of Oriented Gradients
IRM	Imagerie par Résonance Magnétique
ICA	Independent Component Analysis
KIDA	Knee Digital Imaging Data Analysis
JSW	Joint Space Width
JSA	Joint Shape Analysis
JSN	Joint Space Narrowing
KL	Kellgren and Lawrence
KOA	Knee Osteoarthritis
KOOS	Knee injury and Osteoarthritis Outcome Score

KOACAD	Knee Osteoarthritis Classification and Data Analysis
LBP	Local Binary Patterns
MAE	Mean Absolute Error
MLR	Multivariate Linear Regression
MLP	Multi-Layer Perceptron
MOST	Multicenter Osteoarthritis Study
NIH	National Institutes of Health
OADIA	Osteoarthritis Data and Imaging Analysis
OA	OsteoArthritis
OAI	Osteoarthritis Initiative
OARSI	Osteoarthritis Research Society International
OKOA	Objective Knee Osteoarthritis
OMS	Organisation Mondiale de la Santé
PA	PosteroAnterior
PCA	Principal Component Analysis
Pr	Precision
Re	Recall
RFCLM	Random Forest Combined with Landmark Modeling
RG	Radiographic
ROI	Regions Of Interest
SVM	Support Vector Machine
SSM	Statistical Shape Model
TBT	Trabecular Bone Texture
XR	X-ray Radiography
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index
KOACCTNet	Knee Osteoarthritis Classification and Clinical Treatment Network
CCT	Computed Clinical Treatment
AHE	Adaptive Histogram Equalization
CLAHE	Contrast Limited Adaptive Histogram Equalization
SDM	Digital Signature Measurement
DX	Diagnostic or Digital X-ray
DICOM	Digital Imaging and Communications in Medicine
SF-36	Short Form 36 Health Survey

IoU	Intersection over Union
ADN	Acide DésoxyriboNucléique
GCN	Graph Convolutional Network
MSE	Mean Squared Error
SAE	Sparse AutoEncoder
VAE	Variational AutoEncoder
RNN	Recurrent Neural Network
BPTT	Backpropagation Through Time
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
BERT	Bidirectional Encoder Representations from Transformers
GPT	Generative Pre-trained Transformer
GNN	Graph Neural Network
GAT	Graph Attention Network
BIRNN	Bidirectional Recurrent Neural Network
MPNN	Message Passing Neural Network
GCN	Graph Convolutional Network
SVM-RBF	Support Vector Machine with Radial Basis Function kernel
SSAE	Stacked Sparse AutoEncoder
CBP	Compact Bilinear Pooling
CSF	Count Sketch Function
SOTA	State Of The Art
NIA	National Institute on Aging
CHECK	Cohort Hip and Cohort Knee
GAN	Generative Adversarial Network
DCGAN	Deep Convolutional Generative Adversarial Network

1

Introduction

Sommaire

1.1	Contexte	3
1.2	Objectifs de la thèse	4
1.3	plan du Mémoire	5

1.1 Contexte

L'arthrose du genou, également connue sous le nom de gonarthrose, Knee Osteoarthritis (KOA) en anglais, est une maladie chronique dégénérative qui touche les articulations du genou. Elle est l'une des formes les plus courantes d'arthrose et représente un problème de santé publique majeur à l'échelle mondiale.

L'arthrose du genou est caractérisée par une dégradation progressive du cartilage articulaire, ce qui entraîne une inflammation, une douleur, une raideur et une réduction de la fonctionnalité du genou. Cette pathologie affecte principalement les personnes âgées, mais elle peut également toucher des individus plus jeunes en raison de facteurs tels que l'obésité, les traumatismes répétés, le surmenage ou des facteurs génétiques.

La prévalence de l'arthrose du genou est en constante augmentation en raison du vieillissement de la population et des modes de vie modernes. Elle entraîne une détérioration de la qualité de vie des individus touchés, limitant leurs capacités fonctionnelles, leur mobilité et leur participation aux activités quotidiennes. La douleur chronique associée à l'arthrose du genou peut également entraîner des problèmes psychologiques tels que la dépression et l'anxiété.

L'impact socio-économique de l'arthrose du genou est considérable. Elle entraîne une charge financière importante pour les systèmes de santé, en raison des coûts associés aux consultations médicales, aux traitements, aux chirurgies et aux rééducations nécessaires. De plus, elle entraîne des pertes de productivité économique dues à l'absentéisme au travail et à la diminution de la capacité de travail des personnes touchées.

Dans ce contexte, le diagnostic précoce et précis de l'arthrose du genou est essentiel pour une prise en charge appropriée des patients. Un diagnostic rapide permet de mettre en place des stratégies de traitement et de gestion de la maladie, d'atténuer les symptômes, de ralentir la progression de la maladie et d'améliorer la qualité de vie des patients. De plus, un diagnostic précis peut contribuer à éviter des interventions chirurgicales inutiles et coûteuses lorsque des traitements conservateurs sont suffisants.

Cependant, le diagnostic de l'arthrose du genou présente des défis. Les symptômes peuvent être similaires à d'autres affections du genou, rendant parfois difficile la distinction entre différentes pathologies. De plus, les méthodes de diagnostic traditionnelles, telles que l'examen clinique et les techniques d'imagerie médicale, peuvent manquer de sensibilité et de spécificité.

C'est dans ce contexte qu'émergent les techniques d'apprentissage profond, une branche de l'intelligence artificielle, pour améliorer le diagnostic de l'arthrose du genou. L'utilisation de ces techniques permet d'exploiter des ensembles de données d'imagerie médicale massives et d'entraîner des modèles d'apprentissage automatique capables de détecter et de classifier les signes caractéristiques de l'arthrose avec une précision accrue.

En améliorant les capacités de diagnostic, les techniques d'apprentissage profond peuvent contribuer

à une détection plus précoce de la maladie, à une meilleure évaluation de sa gravité et à des décisions de traitement plus éclairées. Elles offrent ainsi des perspectives prometteuses pour améliorer la prise en charge clinique des patients atteints d'arthrose du genou et réduire l'impact socio-économique de cette maladie.

En conclusion, l'arthrose du genou constitue un problème de santé publique majeur en raison de sa prévalence croissante et de son impact significatif sur la qualité de vie des individus. Le diagnostic précoce et précis de l'arthrose du genou est essentiel pour une prise en charge optimale des patients. Les techniques d'apprentissage profond offrent des possibilités prometteuses pour améliorer le diagnostic de l'arthrose du genou, contribuant ainsi à une meilleure gestion de la maladie et à une amélioration de la qualité de vie des patients.

1.2 Objectifs de la thèse

1. Développer des techniques d'apprentissage profond spécifiquement adaptées au diagnostic de l'arthrose du genou : Cette thèse vise à explorer et à développer des modèles d'apprentissage profond qui sont capables d'analyser les images d'imagerie médicale du genou et de détecter de manière précise les signes caractéristiques de l'arthrose. Il s'agit de concevoir des architectures de réseaux de neurones et des méthodes d'apprentissage appropriées pour améliorer la sensibilité, la spécificité et la fiabilité du diagnostic.
2. Collecter et prétraiter des ensembles de données d'imagerie médicale du genou : Pour former et évaluer les modèles d'apprentissage profond, il est nécessaire de collecter des ensembles de données cliniques comprenant des images d'imagerie médicales du genou provenant de patients atteints d'arthrose. Cette thèse se propose de rassembler ces données, de les prétraiter de manière adéquate et de les organiser dans un format approprié pour l'entraînement et l'évaluation des modèles.
3. Évaluer les performances des modèles d'apprentissage profond pour le diagnostic de l'arthrose du genou : L'évaluation des performances des modèles d'apprentissage profond est un aspect essentiel de cette thèse. Il s'agit de mesurer la précision, la sensibilité, la spécificité et d'autres métriques de performance des modèles par rapport aux méthodes de diagnostic traditionnelles. L'évaluation inclura également des comparaisons avec d'autres approches d'apprentissage automatique et des analyses de robustesse et de généralisation des modèles.

En résumé, cette thèse vise à développer des techniques d'apprentissage profond spécifiquement adaptées au diagnostic de l'arthrose du genou, en utilisant des ensembles de données cliniques et en évaluant les performances des modèles développés. Les objectifs ultimes sont de contribuer à une détection plus précoce et plus précise de l'arthrose du genou, ainsi qu'à une amélioration de la prise en charge clinique de cette maladie débilitante.

1.3 plan du Mémoire

Ce manuscrit est structuré en cinq chapitres, détaillant les concepts fondamentaux, les approches existantes et les contributions spécifiques de cette étude. Une vue d'ensemble du contenu de chaque chapitre est présentée ci-dessous :

Chapitre 1

Ce chapitre introduit l'arthrose, une maladie dégénérative des articulations, en détaillant ses caractéristiques, sa prévalence et son impact épidémiologique. Il aborde également les facteurs de risque, les zones les plus touchées, comme la gonarthrose (arthrose du genou), et les bases du diagnostic clinique et radiographique. Enfin, il explore les limites des systèmes de classification actuels et souligne l'importance des outils d'analyse assistés par ordinateur pour une détection et une évaluation objectives.

Chapitre 2

Ce chapitre examine les progrès récents dans l'utilisation des techniques d'apprentissage profond pour la détection précoce et l'évaluation de la gravité de l'arthrose du genou. Il met en lumière les méthodes basées sur les réseaux neuronaux, telles que les modèles transformeurs, réseaux siamois et autoencodeurs, ainsi que l'apprentissage semi-supervisé. Les aspects tels que le prétraitement d'images, la segmentation des articulations, l'intégration de mécanismes de régularisation et l'augmentation des données sont abordés. Les limites et défis persistants, notamment la gestion des données non étiquetées et la généralisation des modèles, sont également discutés.

Chapitre 3

Ce chapitre fournit une vue d'ensemble des principales architectures utilisées dans l'apprentissage profond, incluant les **Auto-Encoder (AE)s**, **Convolutional Neural Network (CNN)s**, **Recurrent Neural Network (RNN)s**, **Transformers**, **Vision Transformers (ViTs)** et **Graph Neural Network (GNN)s**. Chaque type de réseau est exploré en fonction de son domaine d'application, avec un focus sur leurs forces, leurs limites et les avancées qu'ils apportent dans des secteurs variés, tels que la vision par ordinateur, le traitement du langage naturel et l'analyse de données structurées.

Chapitre 4

Dans ce chapitre, une approche automatique basée sur des autoencodeurs empilés parcimonieux est décrite pour extraire des Regions Of Interest (ROI)s à partir d'images radiographiques du genou. L'impact de la taille et de l'emplacement des régions d'intérêts (ROIs) sur la classification de l'arthrose précoce est étudié. Bien que prometteuse, la méthode est examinée en vue d'améliorer sa robustesse et sa généralisation à d'autres modalités et bases de données.

Chapitre 5

Ce chapitre introduit une nouvelle approche combinant des caractéristiques de texture et de forme pour la détection précoce de l'ostéoarthrite du genou. Basée sur le Compact Bilinear Pooling (CBP), cette méthode montre que les symptômes précoces sont particulièrement visibles dans certaines régions spécifiques du genou. Les résultats obtenus soulignent l'importance de sélectionner des régions d'intérêt pertinentes pour améliorer la classification et fournir des outils d'aide à la décision clinique plus performants.

Ce mémoire se termine par une [conclusion](#), dans laquelle nous résumons les résultats obtenus, discutons de leurs implications et présentons les perspectives de recherches futures. Nous mettons également en lumière les limites de notre approche et suggérons des voies d'amélioration pour les travaux à venir. Les perspectives incluent des pistes pour surmonter les défis rencontrés, ainsi qu'une liste des travaux publiés dans le cadre de cette thèse.

2

La gonarthrose

Sommaire

2.1	Définition et caractéristiques de l'arthrose	9
2.2	Cas de la gonarthrose	11
2.3	Diagnostic de la gonarthrose	13
2.4	Limitations	16
2.5	Base de données publiques sur la gonarthrose	17
2.6	Conclusion	23

2.1 Définition et caractéristiques de l'arthrose

L'arthrose est une maladie articulaire dégénérative chronique caractérisée par la dégradation progressive du cartilage articulaire qui recouvre les extrémités des os au sein des articulations, entraînant douleur, raideur et limitations de la mobilité articulaire [Cross et al., 2014]. Ces caractéristiques principales sont :

1. Dégénération du cartilage : Dans l'arthrose, le cartilage articulaire se détériore progressivement, devenant plus mince et moins résistant. Cela entraîne une altération de sa fonction d'amortissement des chocs et de lubrification des articulations.
2. Douleur articulaire : La douleur est l'un des symptômes principaux de l'arthrose. Elle est généralement décrite comme une douleur profonde, lancinante ou brûlante dans l'articulation touchée. La douleur peut être présente lors des mouvements ou après des périodes d'activité prolongée.
3. Raideur articulaire : Les personnes atteintes d'arthrose peuvent ressentir une raideur articulaire, en particulier après une période de repos ou d'inactivité. La raideur peut se dissiper après avoir effectué des mouvements et une activité physique.
4. Limitations de la mobilité : L'arthrose peut entraîner une diminution de la mobilité articulaire. Les mouvements peuvent devenir difficiles et restreints, ce qui peut limiter les activités quotidiennes et réduire la qualité de vie.
5. Gonflement : Dans certains cas, l'arthrose peut provoquer un léger gonflement de l'articulation touchée. Cela est souvent dû à une inflammation et à une accumulation de liquide synovial dans l'articulation.
6. Déformations articulaires : À un stade avancé de l'arthrose, des déformations articulaires peuvent se produire. Par exemple, dans l'arthrose des mains, des nodules durs appelés nodosités de Heberden ou de Bouchard peuvent se former sur les articulations touchées.

Il est important de noter que les symptômes de l'arthrose peuvent varier d'une personne à l'autre et peuvent être influencés par plusieurs facteurs tels que l'âge, le poids, les prédispositions génétiques, les blessures antérieures et l'activité physique. Le diagnostic précis de l'arthrose est généralement basé sur une combinaison de symptômes, d'examen clinique et d'imagerie médicale telle que les radiographies.

2.1.1 Prévalence et épidémiologie

L'arthrose est l'une des maladies articulaires les plus répandues dans le monde. Selon l'Organisation Mondiale de la Santé (OMS) [Hunter and Felson,], environ 9,6% des hommes et 18% des femmes âgés de 60 ans et plus souffrent d'arthrose symptomatique au niveau des hanches et des genoux [Lim and Lau, 2011].

En France, elle touche 10 millions de Français, soit environ 17% de la population. Elle touche principalement les personnes âgées : 68 % des personnes atteintes d'arthrose ont plus de 50 ans, 50% des personnes

ont de l'arthrose après 65 ans, 85 % des personnes sont touchées après 70 ans. Les hommes et les femmes sont touchés de façon égale jusqu'à la cinquantaine. Après cet âge, les femmes sont plus nombreuses à être concernées, Cependant, l'arthrose peut apparaître chez des jeunes personnes, en particulier l'arthrose des doigts ou l'arthrose au genou.

2.1.2 Localisation : Zones les plus touchées

L'arthrose peut affecter différentes articulations du corps, quelques-unes des localisations les plus courantes de l'arthrose dans le tableau 2.1.

Zone Concernée	Pathologie	Épidémiologie
Colonne vertébrale	Arthrose cervicale, Arthrose lombaire	Touche 85 % des hommes et 75 % des femmes de plus de 50 ans, sans pour autant causer de symptômes dans bien des cas.
Mains	Arthrose des mains	La plus fréquente des arthroses douloureuses, touche surtout les femmes, a une forte composante familiale.
Genou	Gonarthrose	Touche environ 1/3 des personnes de plus de 60 ans. Douloureuse chez 6,6 % des femmes et 4,7 % des hommes dans la population générale.
Hanches	Coxarthrose	Très gênante. Touche environ 3 % des adultes. Sa prévalence varie de 1 % à 3,9 % chez les hommes et de 0,8 % à 5,1 % chez les femmes.

Table 2.1 – Épidémiologie des pathologies arthrosiques selon la zone concernée.

L'arthrose peut être classée en deux catégories principales :

1. Arthrose primitive : C'est la forme la plus courante d'arthrose, et elle se produit généralement en raison du vieillissement naturel des articulations. Elle est considérée comme "primitive" car elle n'est pas directement causée par une autre condition médicale ou un facteur externe spécifique. Les facteurs de risque pour développer une arthrose primitive incluent l'âge avancé, des antécédents familiaux d'arthrose, le surpoids, la sédentarité, des anomalies anatomiques (e.g. problèmes d'axe)
2. Arthrose secondaire : Contrairement à l'arthrose primitive, l'arthrose secondaire est causée par des facteurs spécifiques tels que des blessures articulaires, des malformations congénitales, des troubles métaboliques ou endocriniens, des infections articulaires passées, ou l'usage excessif des articulations dans certaines professions ou activités.

2.2 Cas de la gonarthrose

La gonarthrose (L'arthrose du(des) genou(x)), Knee Osteoarthritis (KOA) en anglais est l'une des formes d'arthrose les plus courantes. Environ 10% des hommes et 13% des femmes de plus de 60 ans souffrent de gonarthrose au niveau du genou. Le risque de développer une gonarthrose augmente avec l'âge. Environ 40% des personnes de plus de 70 ans présentent des signes de gonarthrose radiographique au niveau du genou. La gonarthrose peut avoir un impact significatif sur la qualité de vie des personnes touchées. Selon OMS, elle est classée parmi les dix principales causes d'incapacité chez les personnes âgées. Les femmes sont plus susceptibles de développer une gonarthrose que les hommes, en particulier après la ménopause. Cela peut être dû à des facteurs hormonaux et à des différences dans la composition du cartilage entre les sexes. Le surpoids et l'obésité sont des facteurs de risque importants pour le développement de la gonarthrose. Environ 66% des personnes atteintes de cette maladie sont en surpoids ou obèses [Guillemin et al., 2011]. Elle est sévère et invalidante, peut nécessiter une intervention chirurgicale telle que l'arthroplastie du genou, où l'articulation endommagée est remplacée par une prothèse. Environ 90% des arthroplasties du genou sont réalisées pour traiter la gonarthrose. La gonarthrose entraîne des coûts importants en matière de soins de santé. Ces coûts comprennent les consultations médicales, les médicaments, les traitements physiothérapeutiques et les éventuelles interventions chirurgicales.

Ces chiffres soulignent l'importance de la gonarthrose en tant que problème de santé publique, et ils mettent en évidence la nécessité de la prévention, du diagnostic précoce et d'une prise en charge appropriée pour améliorer la qualité de vie des personnes touchées par cette affection.

La gonarthrose présente des symptômes tels que douleur lancinante, raideur, craquements, léger gonflement et réduction de la mobilité. la douleur s'aggrave avec l'activité physique et s'améliore au repos, tandis que l'arthrite rhumatoïde peut provoquer une raideur matinale prolongée et des symptômes systémiques tels que fatigue et fièvre (Figure 2.1).

2.2.1 Structure et composition de l'os

Dans cette partie, nous allons présenter la structure et les composantes de l'articulation du genou, qui comprend les os spécifiques constituant cette articulation, à savoir le fémur, le tibia et la rotule (patella), Figure(2.2).

1. Le fémur : est l'os le plus long et le plus robuste du membre inférieur. Sa structure est essentielle pour supporter le poids du corps et participer à une grande variété de mouvements. Il est principalement composé de tissu osseux, qui contient une matrice extracellulaire solide et des cellules spécialisées, appelées ostéocytes, réparties dans cette matrice. Le fémur a une forme allongée et cylindrique. Sa partie supérieure est appelée la tête du fémur, qui s'articule avec l'os pelvien pour former l'articulation de la hanche. À partir de la tête, le col du fémur relie la tête à la partie

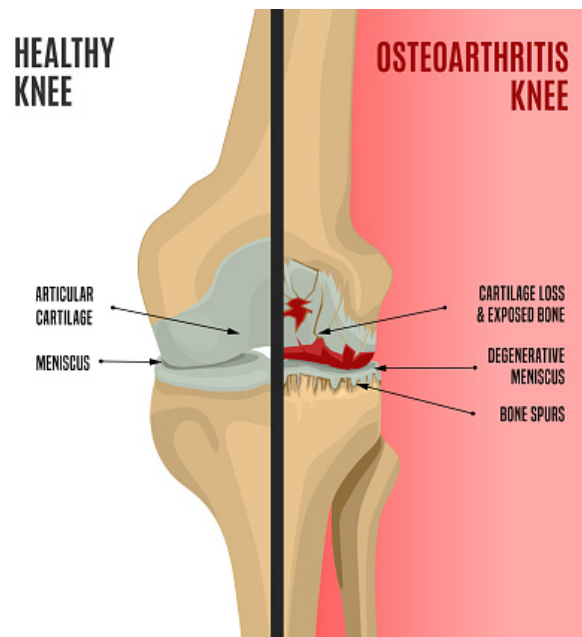


Figure 2.1 – Illustration schématique montrant les principales caractéristiques pathologiques de l'arthrose du genou. Le côté gauche représente une articulation normale, tandis que le côté droit représente une articulation atteinte.(source [<https://gideonburstein.co.il/>])



Figure 2.2 – Illustration de la région d'articulation du genou(source[<https://www.cortho.org/knee/>])

principale de l'os. La partie inférieure du fémur présente deux condyles arrondis, le condyle médial et le condyle latéral, qui s'articulent avec le tibia. À l'intérieur du fémur se trouve l'os compact, formant la couche externe dense de l'os, qui fournit sa résistance mécanique. Cette couche compacte entoure une zone centrale d'os spongieux (ou trabéculaire), qui contient des espaces poreux remplis de moelle osseuse rouge, responsable de la production de cellules sanguines.

2. Le tibia : est l'os principal de la jambe, situé médialement par rapport au fémur. Il joue un rôle crucial dans la stabilité du genou et la transmission de la force du fémur vers le pied. Comme le fémur, le tibia est principalement constitué de tissu osseux, avec une matrice extracellulaire solide contenant des ostéocytes. La partie supérieure du tibia a deux surfaces d'articulation appelées plateaux tibiaux, le plateau tibial médial et le plateau tibial latéral. Ces plateaux accueillent les condyles du fémur pour former l'articulation du genou. La partie inférieure du tibia s'élargit et forme la malléole médiale, qui est la protubérance osseuse sur le côté intérieur de la cheville. L'os compact entoure l'os spongieux à l'intérieur du tibia, tout comme dans le fémur. L'os compact assure la solidité de l'os, tandis que l'os spongieux contient la moelle osseuse rouge.
3. La rotule (patella) : est un petit os plat et triangulaire situé à l'avant du genou. Elle joue un rôle important en améliorant la mécanique du mouvement du genou et en protégeant les structures articulaires. : La rotule est également composée de tissu osseux avec une matrice extracellulaire et des ostéocytes. Elle est conçue pour s'insérer dans la rainure osseuse appelée trochlée du fémur. Cette forme permet à la rotule de se déplacer plus facilement pendant la flexion et l'extension du genou. La partie supérieure de la rotule est appelée la base, la partie inférieure est le sommet et le côté pointu est la pointe.

2.3 Diagnostic de la gonarthrose

Le processus de diagnostic commence par une anamnèse complète, où le médecin interroge le patient sur ses symptômes, leur évolution, les facteurs déclenchant et tout antécédent médical pertinent. Les principaux symptômes de la gonarthrose incluent une douleur progressive du genou, une raideur articulaire, une limitation des mouvements et des craquements lors des mouvements [Cibrián Uhalte et al., 2017]. L'examen clinique permet ensuite d'évaluer l'état du genou affecté. Le médecin observe la mobilité articulaire, recherche des signes d'inflammation ou d'épanchement, palpe la zone pour détecter d'éventuelles douleurs locales et analyse la stabilité de l'articulation. Pour confirmer le diagnostic et évaluer la gravité de la gonarthrose, des examens d'imagerie tels que les radiographies à rayons X sont utilisés. Les radiographies permettent de visualiser les altérations caractéristiques de l'arthrose, notamment l'usure du cartilage, la formation d'ostéophytes (excroissances osseuses) et la réduction de l'espace articulaire [Hun-

ter and Felson, , Sakellariou et al., 2017]. Dans certains cas, une imagerie plus avancée, comme l’Imagerie par Résonance Magnétique (IRM) ou l’échographie, peut être prescrite pour fournir une vue détaillée des tissus mous comme le cartilage, les ligaments, les tendons, la membrane synoviale, les bourses séreuses environnantes et évaluer l’étendue des lésions [Hayashi et al., 2016].

Cependant, les radiographies à rayons X restent l’outil de dépistage initial couramment utilisé pour le diagnostic de la gonarthrose, en raison de leur disponibilité, de leur coût moindre et de leur sensibilité aux changements osseux [Lee et al., 2021]. l’IRM peut être utilisée en complément ou en alternative pour un diagnostic plus précis.

Parmi les critères cliniques pour le diagnostic de la gonarthrose, on trouve le Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) est un questionnaire largement utilisé et bien établi pour évaluer l’état de santé et les limitations fonctionnelles des personnes atteintes d’ostéoarthrite OsteoArthritis (OA), développé dans les années 80 [Reginster, 1980] , il est conçu pour aider les professionnels de la santé et les chercheurs à mesurer la gravité de l’ OA et à suivre l’évolution de l’état d’un patient au fil du temps. Le questionnaire WOMAC se compose de 24 questions réparties en trois sous-échelles :

- Douleur (5 questions) : lors d’activités telles que la marche, la montée d’escaliers, la station debout, la position assise et la position couchée.
- Raideur (2 questions) : le matin et pendant la journée.
- Fonction physique (17 questions) : Cette section mesure la capacité du patient à effectuer diverses activités physiques, notamment monter et descendre des escaliers, se lever d’une position assise, monter et descendre d’une voiture et marcher.

Les patients notent leurs expériences sur une échelle de Likert, généralement de 0 à 4 ou de 0 à 10, les scores plus élevés indiquant une douleur, une raideur ou des limitations fonctionnelles plus importantes.

Les critères radiologiques consistent à visualiser les changements structurels associés à la maladie. Les signes radiographiques caractéristiques de la gonarthrose incluent le rétrécissement de l’espace articulaire, également connu sous le nom de diminution de l’interligne radiographique. La diminution de l’interligne radiographique est un signe majeur de perte de cartilage, et elle est évaluée par les critères cliniques de classification de la gonarthrose, tels que les critères Kellgren and Lawrence (KL) et Osteoarthritis Research Society International (OARSI). Cela permet d’établir un diagnostic précis et de classer la sévérité de la gonarthrose de manière plus détaillée. Ces critères jouent un rôle essentiel dans la standardisation de l’évaluation radiographique de la gonarthrose et permettent une approche plus objective et cohérente pour la classification de cette maladie débilitante du genou.



Figure 2.3 – Différents grades de la gonarthrose selon Kellgren & Lawrence. Les images sont extraites de la base de donnée Osteoarthritis Initiative (OAI).

2.3.1 Les critères de KL

Les critères KL, développés par Kellgren et Lawrence en 1957 [Kellgren and Lawrence, 1957], sont largement utilisés pour évaluer la sévérité de la gonarthrose à l’aide de radiographies du genou. Ce système de classification comporte cinq stades, qui sont les suivants (Figure 2.3) :

1. Grade 0 : Aucun signe radiographique d’arthrose visible. L’espace articulaire est normal, et aucun changement dégénératif n’est détecté.
2. Grade 1 : Un léger rétrécissement de l’espace articulaire peut être observé, indiquant une possible perte de cartilage. Cependant, les changements sont encore minimes.
3. Grade 2 : Un rétrécissement de l’espace articulaire plus évident est constaté, accompagné d’ostéophytes mineurs (éperons osseux) à la périphérie de l’articulation.
4. Grade 3 : À ce stade, le rétrécissement de l’espace articulaire est significatif, et des ostéophytes plus importants sont présents. Des signes d’une altération structurale plus avancée de l’articulation, tels que des déformations des contours osseux, peuvent également être observés.
5. Grade 4 : C’est le stade le plus avancé de la gonarthrose. Il montre un rétrécissement majeur de l’espace articulaire, des ostéophytes étendus et des signes évidents d’une arthrose sévère avec une altération importante de l’articulation.

2.3.2 Les critères de OARSI

l’Atlas OARSI (Osteoarthritis Research Society International) [Altman and Gold, 2007] propose des critères de classification spécifiques, des scores et une évaluation par zone du genou.

Les critères de classification de l’Atlas OARSI comprennent la perte de l’épaisseur du cartilage, la présence d’ostéophytes, l’écart articulaire, l’alignement du genou et d’autres caractéristiques radiologiques

pertinentes comme illustré dans la figure 2.4.

1. La perte de l'épaisseur du cartilage est évaluée en utilisant une échelle de notation allant de 0 à 4, où 0 n'indique aucune perte et 4 indique une perte complète du cartilage. Ce critère est appliqué à différentes zones du genou, notamment le compartiment médial, le compartiment latéral et la rotule.
2. Les ostéophytes, qui sont des excroissances osseuses autour des bords des surfaces articulaires, sont également évalués selon leur taille et leur nombre. Ils sont notés sur une échelle de 0 à 3 ou de 0 à 4.
3. Joint Space Narrowing (JSN), qui mesure l'espace entre les surfaces articulaires du genou, est évalué pour détecter des signes de réduction de l'espace, ce qui peut indiquer une perte de cartilage et une dégradation de l'articulation.
4. L'alignement du genou est évalué en utilisant des critères spécifiques pour déterminer s'il y a une malposition ou une déformation osseuse qui pourrait contribuer à l'arthrose.

OARSI évalue également différentes zones du genou pour une évaluation plus détaillée :

1. Le compartiment médial, qui correspond à la partie interne du genou, où l'articulation est formée entre le fémur et le tibia du côté médial de la jambe.
2. Le compartiment latéral, qui est la partie externe du genou, où l'articulation est formée entre le fémur et le tibia du côté latéral de la jambe.
3. Le compartiment fémoro-patellaire, qui concerne l'articulation entre la rotule et le fémur.
4. Le compartiment tibio-fémoral, qui englobe l'articulation principale entre le fémur et le tibia, divisée en compartiments médial et latéral.

L'Atlas OARSI fournit des critères, des scores et une évaluation par zone du genou pour la classification de la gonarthrose, ce qui permet une évaluation précise de la sévérité de la maladie et une adaptation du traitement en fonction des besoins individuels de chaque patient.

2.4 Limitations

Bien que l'indice de KL et OARSI soient des outils de classification couramment utilisés pour l'évaluation de l'arthrose du genou, ils présentent tous deux des limitations liées à la subjectivité. Le score KL repose sur l'évaluation visuelle des radiographies, ce qui peut introduire une certaine subjectivité dans les résultats en raison des différences d'interprétation entre les évaluateurs. Il peut manquer de sensibilité pour détecter les stades précoces de l'arthrose, rendant difficile la détection précoce de la maladie. Le grade KL se concentre principalement sur la présence d'ostéophytes (éperons osseux) et la réduction de l'espace articulaire, mais il ne prend pas en compte d'autres aspects importants de l'arthrose du genou,

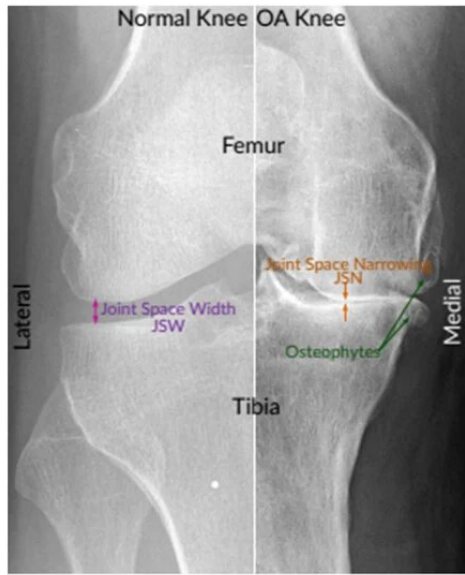


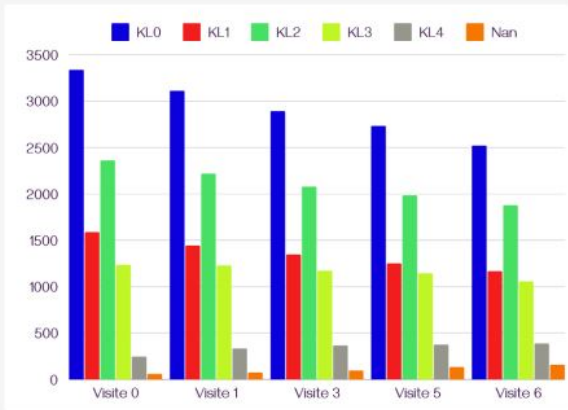
Figure 2.4 – L'image présente des radiographies de deux genoux : un genou normal à gauche et un genou atteint d'arthrose sévère (OA) à droite. La comparaison met en évidence des caractéristiques distinctes, notamment les compartiments articulaires, le rétrécissement de l'espace articulaire (JSN) et la présence d'ostéophytes généralement associés à l'arthrose [Bayramoglu et al., 2020].

tels que la déformation osseuse, les lésions méniscales ou la perte de cartilage dans des zones spécifiques. Concernant le grade OARSI, son utilisation peut être complexe et chronophage en raison de la prise en compte de multiples critères et mesures spécifiques sur les radiographies. Il peut encore présenter une certaine variabilité entre les évaluateurs, en particulier lorsque des mesures manuelles sont nécessaires, Il dépend de la qualité des radiographies, ce qui signifie que des radiographies de mauvaise qualité peuvent entraîner une évaluation moins précise et des erreurs.

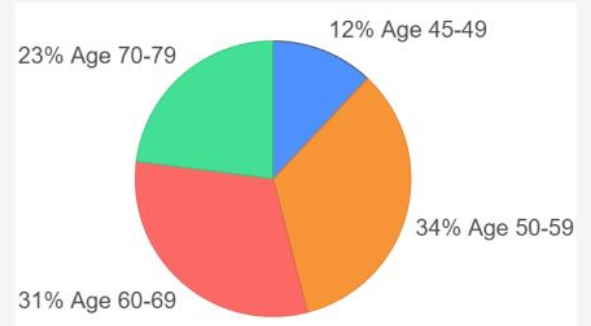
2.5 Base de données publiques sur la gonarthrose

Il existe plusieurs bases de données publiques dédiées au diagnostic de la gonarthrose, regroupant des informations cliniques et médicales essentielles, des données de recherche, ainsi que des renseignements spécifiques aux patients. Ces bases constituent des ressources précieuses pour l'étude de l'arthrose en fournissant des données variées, notamment des imageries médicales et des informations sur les caractéristiques des patients.

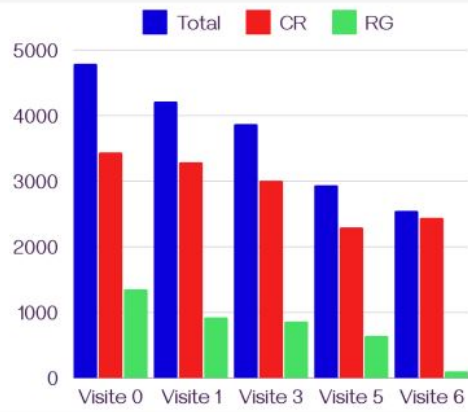
Le tableau 2.2 résume les principales bases de données publiques disponibles, en précisant le nombre de patients inclus, les niveaux de gravité de la pathologie représentés et les types d'imageries médicales disponibles. Parmi ces bases de données, les plus largement utilisées dans la recherche sont l'OsteoArthritis Initiative (OAI) [Peterfy et al., 2008], et l'étude Multicenter Osteoarthritis Study Multicenter Osteoarthritis Study (MOST) [Segal et al., 2013]. Dans cette thèse, la base de données utilisée pour les



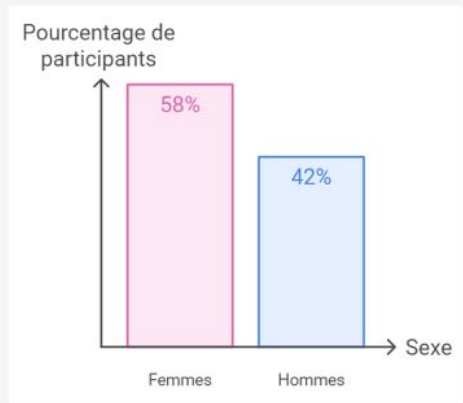
(a) Nombre d'images disponibles pour chaque garde et visite



(b) Age des sujets.



(c) Répartition par modalité



(d) Répartition hommes/femmes.

Figure 2.5 – Caractéristiques des sujets de la cohorte OAI. Ces valeurs sont celles De visite 0 jusqu'au Visite 6, les modalités sont susceptible de varier d'une année a l'autre.

expériences est l'OAI, qui fournit des images radiographiques du genou prises en position fixe et en flexion bilatérale de face. Les bases de données OAI et MOST sont présentées en détail dans les deux sections suivantes.

2.5.1 Base de données OAI

L'OAI a été lancée en 2002 en tant qu'étude de cohorte multicentrique, longitudinale et observationnelle, visant à recueillir des données sur l'évolution de l'arthrose du genou. Ce projet est le fruit d'un partenariat public-privé entre les National Institutes of Health (NIH) et des entreprises privées, et a pour objectif de développer une ressource de recherche accessible au public pour évaluer scientifiquement les biomarqueurs de l'arthrose [Almhdie-Imjabbar et al., 2021]

Base de données	Gravité (KL)	Modalités d'imagerie	Suivi (mois)
OAI	$0 \leq \text{KL} \leq 4$	XR, MRI	108
MOST	$0 \leq \text{KL} \leq 4$	XR, MRI	84
FNIH	$1 \leq \text{KL} \leq 3$	XR	48
SEKOIA	$1 \leq \text{KL} \leq 3$	XR	36

Table 2.2 – Résumé différentes des bases de données publique de la gonarthrose

2.5.1.A Caractéristiques et Objectifs de l'étude OAI

L'OAI vise à étudier la nature évolutive de la gonarthrose à travers diverses étapes de la maladie. Pour cela 4 796 participants ont été recrutés, avec une proportion de 58% de femmes, pour étudier l'évolution de la gonarthrose à travers différents stades de la maladie. Les critères d'inclusion stipulent que les participants doivent être âgés de 45 à 79 ans et présenter des symptômes d'arthrose du genou ou être à risque de développer cette condition. L'étude prend également en compte la diversité ethnique, avec une attention particulière à l'inclusion des Afro-Américains. Chaque participant est suivi annuellement pendant sept ans, ce qui permet de collecter des données longitudinales sur la progression de la maladie. Un tableau récapitulatif du suivi, ainsi que des informations sur la disponibilité des images, est présenté dans la figure 2.6. La couleur verte fait référence aux points de contrôle où les images sont disponibles en téléchargement. À huit ans de suivi, des images radiographiques (X-ray Radiography (XR)) et d'IRM ont été acquises, mais elles ne sont pas encore disponibles en téléchargement. Pour les autres points de suivi, représentés en noir, aucune image radiographique ni IRM n'a été acquise. En complément, la figure inclut des captures d'écran des dossiers contenant les images. Ces dossiers sont nommés selon un format spécifique : le premier élément correspond au numéro de la visite, le deuxième est une lettre indiquant le centre de référence, et le troisième chiffre représente la version des données. Après téléchargement de l'ensemble de la base de données, les images radiographiques seront accessibles du début de l'étude jusqu'au 72[U+1D49] mois de suivi. Dans chaque dossier, les images sont organisées sous des sous-dossiers portant l'identifiant du participant. Elles sont enregistrées au format Digital Imaging and Communications in Medicine (DICOM), accompagné d'un ensemble de métadonnées relatives à la procédure d'acquisition. La cohorte est stratifiée en fonction de la gravité de l'arthrose et des symptômes observés, ce qui permet d'analyser les trajectoires cliniques des individus. L'objectif de cette base de données est d'explorer les facteurs liés à la progression ou la régression de la gonarthrose, d'analyser les biomarqueurs biochimiques et génétiques associés, et de mieux comprendre l'impact des facteurs environnementaux et comportementaux sur l'évolution de la maladie. En outre, l'OAI sert de ressource pour la recherche collaborative en facilitant l'étude des données cliniques et d'imagerie, tout en contribuant à la découverte de nouveaux indicateurs précoces pour la détection et le traitement de la gonarthrose. La figure 2.5 reprend la répartition des caractéristiques principales des sujets de la cohorte.

qualité de vie des participants. Des évaluations similaires sont effectuées pour d'autres articulations, et des questionnaires comme le Short Form 36 Health Survey (SF-36) [TO and ALEXA, 1982] mesurent la performance physique générale et la qualité de vie.

Enfin, des biospecimens sont collectés, comprenant des échantillons de sang, d'urine et d'Acide DésoxyriboNucléique (ADN) pour des analyses biochimiques et génétiques. Le tout est effectué selon des protocoles stricts pour garantir la qualité des données et leur comparabilité au fil des années.

Le protocole de l'étude OAI, les mesures, les données cliniques et les images DICOM peuvent être téléchargés en ligne sur le site suivant : <https://nda.nih.gov/oai>. De plus, diverses évaluations d'images quantitatives et semi-quantitatives sont disponibles. Plus de 400 articles de recherche ont été produits à partir de ces données.

2.5.2 Base de données MOST

Ref.	V0	V1	V2	V3	V4	V5
Control visit	Baseline	15	30	60	72	84
XR Images	✓	✓	✓	✓		✓
MRI Images	✓	✓	✓	✓		✓
Clinical data	✓	✓	✓	✓	✓	✓
Outcomes data (TKR)	✓	✓	✓	✓	✓	✓

Figure 2.7 – Disponibilité des données collectées par visite dans la base MOST

MOST est une étude prospective de grande envergure qui examine les facteurs de risque et la progression de l'arthrose, en particulier de l'arthrose du genou, dans une population de 50 à 79 ans [Segal et al., 2013]. L'étude inclut 3 026 participants ayant des symptômes d'arthrose du genou ou à risque élevé de développer la maladie. Elle a été menée grâce à la collaboration de plusieurs centres cliniques et a reçu le soutien de la NIH et du National Institute on Aging (NIA).

Les modalités d'imagerie utilisées dans MOST comprennent les XR, pour observer les changements osseux (réduction de l'espace articulaire, ostéophytes) et l'IRM, qui permet d'analyser la dégradation du cartilage et des tissus mous. Ces données permettent de suivre l'évolution de la maladie, d'identifier les facteurs de risque (comme l'obésité et les blessures antérieures) et d'améliorer le diagnostic et le traitement.

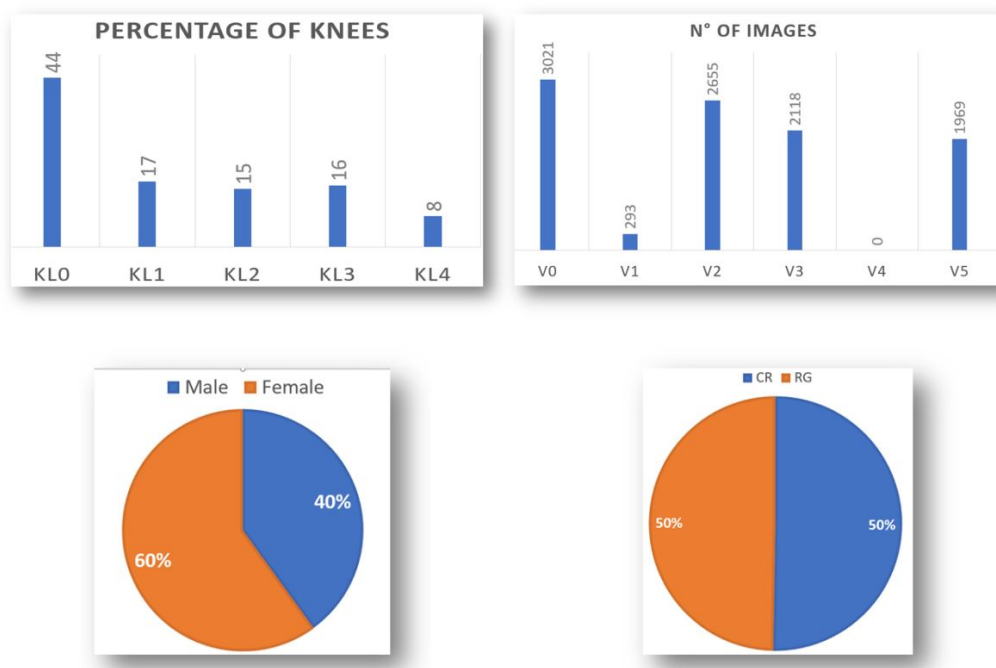


Figure 2.8 – Caractéristiques des sujets de la base de donnée MOST. Ces valeurs sont celles De visite 0 jusqu’au Visite 5, les modalités sont susceptible de varier d’une année a l’autre.

2.5.2.A Caractéristiques et objectifs de l’étude MOST

L’objectif principal de MOST est d’étudier les facteurs de risque, les mécanismes et la progression de l’arthrose du genou chez les personnes âgées à risque. MOST offre une vue d’ensemble détaillée et bien structurée des données longitudinales collectées sur les participants. La figure 2.7, sous forme de tableau, présente les données disponibles à travers six visites successives, échelonnées sur une période allant de l’inclusion initiale (baseline) jusqu’à 84 mois. Les lignes du tableau indiquent le moment de chaque visite en mois : V0 (baseline), suivie de 15, 30, 60, 72 et 84 mois, permettant d’identifier le suivi longitudinal des participants. Les colonnes renseignent sur les différentes catégories des données collectées pour chaque visite, à savoir : les images radiographiques (XR), les images IRM, les données cliniques, et les outcomes. Une coche dans le tableau indique la disponibilité de chaque type de données pour chaque visite. En complément, la figure 2.8 fournie une analyse statistique détaillée des données disponibles. Elle présente la répartition des images selon les visites, permettant ainsi d’apprécier la continuité du suivi longitudinal, ainsi que la répartition des images par grades KL, qui reflète la progression de l’arthrose chez les participants. Cette figure illustre aussi la distribution des âges des sujets, soulignant que la population étudiée est majoritairement âgée de 50 à 79 ans, en cohérence avec les objectifs de l’étude. Enfin, les modalités d’imagerie disponibles, distinguant les CR et les RG, sont également mises en lumière.

2.6 Conclusion

Ce chapitre débute par une définition de l'arthrose, une maladie dégénérative des articulations caractérisée par une dégradation progressive du cartilage articulaire, causant douleur, raideur et perte de mobilité. Contrairement à l'arthrite, l'arthrose résulte de l'usure naturelle des articulations plutôt que d'une inflammation systémique.

La prévalence et l'épidémiologie de l'arthrose sont ensuite abordées, soulignant son impact majeur chez les personnes âgées. Les facteurs de risque incluent l'âge, le sexe, l'obésité, les blessures articulaires et les prédispositions génétiques. Parmi les zones les plus touchées par l'arthrose, la gonarthrose, ou arthrose du genou, occupe une place centrale en raison des contraintes mécaniques élevées exercées sur cette articulation.

Ce chapitre traite également de la structure de l'os, composée de cellules spécifiques, de collagène et de minéraux tels que le calcium, indispensables à la santé articulaire. Enfin, le diagnostic de l'arthrose repose sur l'examen clinique et l'imagerie, avec une attention particulière portée à la gonarthrose.

Les systèmes de classification comme le grade KL, simple mais subjectif, et l'atlas OARSI, plus détaillé mais complexe, sont discutés. Pour pallier leurs limites, le développement de méthodes assistées par ordinateur pourrait améliorer la détection précoce et l'évaluation objective de la maladie, optimisant ainsi la prise en charge clinique. Dans la dernière partie du chapitre, nous avons introduit les principales bases de données publiques dédiées à l'étude de la gonarthrose. Ces bases de données constituent des ressources essentielles pour la recherche en fournissant des informations cliniques et radiographiques standardisées, permettant ainsi une meilleure compréhension de la progression de la maladie. Nous avons présenté en détail les bases OAI et MOST, qui offrent un suivi longitudinal des patients à travers différentes modalités d'imagerie, telles que les radiographies et l'IRM. Le protocole d'acquisition des données de ces bases a également été décrit, mettant en évidence l'importance de la standardisation pour garantir la fiabilité des analyses.

3

État de l'Art sur les approches de detection de la gonarthrose

Sommaire

3.1	Intoduction	25
3.2	Méthodes assistées par ordinateur pour la classification de la gonarthrose	25
3.3	Le diagnostic automatique de la gravité de la gonarthrose à partir de radiographies	26

3.1 Introduction

Le développement des systèmes d'aide au diagnostic (Computer-Aided Diagnosis, CAD) a transformé l'analyse des images médicales en permettant une évaluation plus rapide, plus objective et souvent plus précise des pathologies. Dans le cadre de la gonarthrose, ces outils jouent un rôle clé dans l'automatisation du diagnostic et l'estimation de la gravité de la maladie à partir de radiographies. En réduisant la subjectivité des évaluations cliniques et en exploitant la puissance des algorithmes d'intelligence artificielle, ces approches permettent d'améliorer la prise en charge des patients et d'optimiser les processus cliniques.

Ce chapitre explore les différentes approches assistées par ordinateur utilisées pour la classification de la gonarthrose. Il examine d'abord les méthodes de segmentation des radiographies, une étape essentielle pour isoler les structures anatomiques pertinentes à l'analyse. Ensuite, il présente les principales stratégies de classification des grades de la maladie, en distinguant les approches traditionnelles fondées sur l'analyse des caractéristiques de forme et de texture, et les méthodes plus récentes basées sur l'apprentissage profond.

3.2 Méthodes assistées par ordinateur pour la classification de la gonarthrose

Le Computer-Aided Diagnosis (CAD) pour l'arthrose du genou fait référence à l'utilisation de techniques et d'algorithmes informatiques pour aider les professionnels de la santé à diagnostiquer l'arthrose dans l'articulation du genou. Cette approche exploite des technologies avancées pour analyser des images médicales telles que des radiographies ou des IRM, et pour extraire des informations significatives qui aident à identifier les signes de l'arthrose.

Au fil des années, le CAD a joué un rôle important dans la recherche sur l'arthrose et la pratique clinique [Çağlar Cengizler and Kabakci, 2023, Shamir et al., 2009, Shamir et al., 2010]. Il implique le développement d'outils logiciels capables de détecter et de quantifier automatiquement les caractéristiques associées à l'arthrose, telles que le rétrécissement de l'espace articulaire, les changements de texture osseuse et la dégradation du cartilage. Ces outils aident les radiologues et les cliniciens en fournissant des mesures quantitatives et des repères visuels qui facilitent un diagnostic précis et la surveillance de la progression de la maladie.

Les systèmes de CAD pour l'arthrose du genou peuvent englober diverses méthodes, notamment celles qui sont basées sur l'analyse de la texture qui remonte à 1989 avec [Dacre et al., 1989]. et qui reposent sur l'analyse des motifs de texture osseuse, à l'aide de méthodes telles que Fractal Signature Analysis (FSA) introduite en 1991 par [Lynch et al., 1991].

On trouve aussi les méthodes d'apprentissage Automatique et Deep Learning (DL) qui est une formation d'algorithmes pour reconnaître les motifs dans les images médicales et effectuer des prédictions basées sur un grand ensemble de données d'images annotées. On les appelle méthodes de machine learning et deep learning. La suite de ce chapitre présente une revue complète de la littérature et de l'état de l'art dans l'évaluation de la gravité de l'arthrose du genou à l'aide des modèles de machine et deep learning.

3.3 Le diagnostic automatique de la gravité de la gonarthrose à partir de radiographies

Étant donné l'absence de traitement efficace pour l'arthrose, excepté en cas de stade avancé où une chirurgie de remplacement total de l'articulation peut être envisagée, le diagnostic précoce demeure la principale option pour prolonger la qualité de vie du patient. Sur le plan clinique, il est possible d'établir un diagnostic précoce de l'arthrose. Cependant, à l'heure actuelle, cela nécessite l'utilisation d'IRM coûteuse et disponible uniquement dans des centres spécialisés. De plus, cette méthode ne permet pas de détecter les changements dans la structure osseuse [Finnilä et al., 2017].

L'examen de radiographie simple est la méthode de référence pour diagnostiquer l'arthrose [Almhdie-Imjabbar et al., 2021]. Cette approche est sûre, abordable et largement disponible. Néanmoins, il est bien connu que la radiographie standard est peu sensible pour détecter les premiers signes d'arthrose. De ce fait, le diagnostic précoce de l'arthrose reste un défi en pratique clinique. En outre, la subjectivité du praticien joue un rôle important dans le diagnostic en raison de l'absence d'un système de notation précis.

Grâce aux avancées dans les ressources informatiques, il est maintenant possible de développer des modèles complexes en utilisant une grande quantité de données. L'apprentissage profond, en particulier les Convolutional Neural Network (CNN), a récemment montré des résultats remarquables dans diverses tâches de reconnaissance d'images et d'aide au diagnostic [Nasser et al., 2022, Tinhinane et al., 2023, Nasser et al., 2024, Mehdi et al., 2025]. Ces modèles puissants peuvent atteindre des performances comparables à celles des humains, ouvrant ainsi la voie à leur utilisation future en pratique clinique.

Le processus de diagnostic de l'arthrose peut être divisé en deux étapes principales : la localisation des régions d'intérêt (c'est-à-dire les articulations du genou) et la classification de ces régions en fonction de KL, qui évalue la gravité de l'arthrose.

3.3.1 Segmentation des radiographies

La localisation de la région d'articulation constitue un domaine crucial de la recherche en imagerie médicale [Pierri, 2023], visant à identifier et à isoler précisément les parties pertinentes de l'articulation du genou à partir d'images médicales, telles que les radiographies. Cette étape est fondamentale pour le

diagnostic et le suivi de la gonarthrose. Différentes approches ont été développées pour la localisation des régions d'intérêt. On trouve les méthodes traditionnelles de traitement d'image, comme le seuillage (une segmentation basée sur le pixel) [Yazid et al., 2012] et les contours actifs (une segmentation basée sur la forme) peuvent être utilisées pour isoler les régions d'intérêt à partir de radiographies du genou [Gornale et al., 2016]. Cependant, ces méthodes peuvent être sensibles aux variations de contraste et aux bruits, ce qui peut affecter la précision de la segmentation. Il existe aussi des méthodes hybrides telles que approches guidées par Atlas [Gubern-Mérida et al., 2012], les techniques clustering comme k-means [Ng et al., 2006], Fuzzy C-Means (FCM) [Zhang and Chen, 2004], champ aléatoire de Markov [Vittitoe et al., 1999] et les approches basées sur les réseaux de neurones artificiels [Zhou et al., 2018].

Cette thèse se penche sur l'exploration des méthodes automatiques pour détecter ou localiser les articulations du genou sur des radiographies. Les avantages des méthodes automatiques sont discutées. Dans la suite, nous proposons une revue de la littérature englobant diverses méthodes de segmentation.

Les approches antérieures, dans la littérature qui étudient les articulations du genou sur les radiographies, peuvent être classées en manuelles : il s'agit d'une définition manuelle des Regions Of Interest (ROI) [Mengko et al., 2005] des articulations du genou. Ces marquages servent de référence précieuse. Ils sont utilisés pour créer un modèle anatomique, aidant ainsi au développement de méthodes automatiques de segmentation [Pandey et al., 2015]. Cependant, ce processus manuel est sujet à la subjectivité et à la chronophagie, surtout avec de nombreux cas à analyser. Des méthodes semi-automatiques : ont été conçues pour automatiser des étapes clés de détection et de segmentation [Gan and Sayuti, 2017]. Ces méthodes cherchent à fusionner les bénéfices de l'automatisation avec l'expertise humaine en intégrant des processus automatisés aux évaluations manuelles, en vue d'améliorer la performance globale du processus. Ce modèle présente plusieurs atouts, incluant la flexibilité permise par l'incorporation de connaissances expertes, ainsi que l'application de techniques avancées de vision par ordinateur. L'une de ces méthodes, appelée Knee Osteoarthritis Classification and Data Analysis (KOACAD) (Diagnostic Assisté par Ordinateur de l'Arthrose du Genou), Proposé par [Oka et al., 2010] se base sur l'utilisation de filtres pour détecter les structures osseuses et localiser des points d'intérêt pertinents. Cependant, cette méthode nécessite une intervention humaine pour certaines étapes, ce qui limite son automatisation. De manière similaire, l'outil Knee Digital Imaging Data Analysis (KIDA) proposé par [Marijnissen et al., 2008] permet d'analyser quantitativement les radiographies du genou en mesurant différentes caractéristiques. Néanmoins, son utilisation requiert également l'expertise d'un professionnel pour une évaluation objective. Une autre approche, développée par [Duryea et al., 2000], repose sur un algorithme basé sur des règles, qui peut être formé pour mesurer l'espace articulaire dans les radiographies du genou. Cette méthode automatise la génération de contours des bords du fémur et du tibia, permettant ainsi le suivi de l'évolution de l'arthrose du genou. Toutefois, il est important de noter que ces approches rencontrent certaines limites. Les variations entre les observateurs ou les utilisateurs peuvent réduire la reproductibilité des résultats

obtenus. De plus, il existe un risque d'erreurs humaines au cours des évaluations manuelles effectuées. Malgré ces défis, les méthodes semi-automatiques et interactives continuent de jouer un rôle crucial dans le domaine de l'analyse d'images médicales. et en fin, on trouve les méthodes automatiques.

Une méthode dite "template matching" a été développée par [Shamir et al., 2008] pour détecter automatiquement les centres d'articulation du genou dans des images radiographiques. Les images sont prétraitées, puis des patches de 15x15 pixels avec le centre d'articulation du genou sont extraits comme modèles. En utilisant une fenêtre glissante de même taille, la distance euclidienne entre les patches et les modèles est calculée, permettant d'identifier le centre d'articulation. Une région de 250x200 pixels autour du centre est extraite comme zone d'intérêt. Les images proviennent de la base de données Baltimore Longitudinal Study of Aging, une base de données de radiographies du genou (BLSA), avec 55 images par niveau d'arthrose, divisées en 20 pour l'entraînement et le reste pour les tests. La méthode a réussi à détecter les centres d'articulation du genou avec succès dans toutes les images radiographiques examinées.

[Anifah et al., 2011] ont étudié l'utilisation de template matching et de Contrast Limited Adaptive Histogram Equalization (CLAHE) méthode pour détecter les articulations du genou et quantifier la zone de l'espace articulaire. Au total, 98 images radiographiques ont été utilisées dans cette étude. La précision de détection obtenue par cette méthode varie de 83,3 % à 100 % pour les genoux gauches et de 60,4 % à 100 % pour les genoux droits. Cependant, cette méthode est ad hoc, entièrement basée sur l'ensemble de modèles utilisés et a peu de chances de bien se généraliser pour des ensembles de données plus grands.

L'étude menée par [Tiulpin et al., 2017] présente une approche en deux étapes pour localiser automatiquement les articulations du genou dans des images radiographiques. La première étape exploite la structure anatomique des articulations en se basant sur les variations d'intensité. La seconde étape implique la formation d'un classifieur pour identifier les régions d'intérêt détectées dans la première étape. Cette classification utilise les Support Vector Machine (SVM) et Histogram of Oriented Gradients (HOG). Ils ont testé leur méthode sur diverses bases de données, comprenant 1574 images de la base de données MOST, 93 images de l'hôpital central de Finlande (Jyväskylä) et 77 images de Objective Knee Osteoarthritis (OKOA). Les résultats démontrent une amélioration significative dans la précision de la localisation des articulations, avec de un Intersection over Union (IoU) de 0.84%, 0.79% et 0.78% respectivement pour les bases de données MOST, Jyväskylä et OKOA.

[Chen et al., 2019]. ont personnalisé le réseau YOLOv2 Pour la détection d'objet a une phase. Ils ont utilisé la large base de données de OAI, avec 4130 image XR soit 8260 articulation, . Ils ont réussi à atteindre une moyenne d'index Jaccard de 0.85%

3.3.2 Classification des grades de la gonarthrose

Cette thèse se concentre principalement sur la classification de la gonarthrose à travers une analyse de la texture, de la forme ou d'une combinaison des deux. Pour cette raison, nous avons examiné les

travaux existants dans la littérature qui utilisent ces méthodes. Nous commençons par les méthodes classiques, basées sur l'extraction manuelle des caractéristiques pour identifier les changements associés aux différents stades de la gonarthrose. Ensuite, nous explorons les méthodes automatiques qui utilisent des techniques de deep learning, lesquelles connaissent ces dernières années un succès croissant dans le domaine du traitement d'images médicales.

3.3.3 Méthodes classiques

3.3.3.A Approches basées sur l'analyse de la texture

Au cours des dernières décennies, plusieurs méthodes ont été proposées pour analyser la texture osseuse à partir d'images radiographiques et ainsi détecter l'arthrose du genou. Ces approches reposent principalement sur l'extraction de caractéristiques manuelles, notamment la texture et la forme, afin d'identifier les altérations dues à la maladie.

[Woloszynski et al., 2010] ont développé une approche de classification des textures osseuses trabéculaires invariantes à la rotation, en utilisant une technique appelée Digital Signature Measurement (SDM), qui quantifie la rugosité, le degré d'anisotropie, et la direction de l'anisotropie des textures osseuses trabéculaires. Cette méthode a été appliquée à 203 images radiographiques : 68 images provenant de genoux avec OA et 135 images de genoux sans OA. L'approche proposée se compose de trois étapes successives.

La première étape consiste à utiliser une méthode automatisée de sélection de ROI sur les radiographies, qui identifie la zone osseuse sous-chondrale située sous la plaque corticale des compartiments tibial médial et latéral. Deux régions de taille 112×112 ont été extraites pour chaque compartiment tibial, comme illustré dans la figure 3.1. Ensuite, trois paramètres de texture osseuse trabéculaire – rugosité, degré d'anisotropie, et direction de l'anisotropie – ont été calculés à l'aide de la méthode de dissimilarité de signature. Enfin, les capacités prédictives des paramètres de texture ont été évaluées à l'aide d'un modèle de régression logistique binaire.

Les résultats ont montré que la texture osseuse trabéculaire du tibia médial est prédictive de la perte de l'espace articulaire médial dans les genoux présentant une arthrose. Cependant, une limitation importante de cette étude réside dans l'utilisation d'une base de données relativement petite et de deux protocoles radiographiques différents, ce qui pourrait influencer la généralisation des résultats.

L'étude [Hirvasniemi et al., 2014] visait à quantifier les différences de texture osseuse entre les sujets atteints d'arthrose du genou (OA) et des témoins appariés, à partir de radiographies standardisées, en utilisant des techniques avancées d'analyse d'images. Un total de 203 genoux a été radiographié et classé selon l'échelle de KL. Les paramètres de densité osseuse ont été dérivés des valeurs en niveaux de gris, tandis que les paramètres structuraux ont été calculés à partir d'images Laplaciennes et de Local Binary Patterns (LBP). La figure 3.2 montre l'emplacement des ROI utilisées dans l'étude. Les résultats ont



Figure 3.1 – Les ROIs de l’os trabéculaire latéral et médial utilisées dans [Woloszynski et al., 2010].

montré que les paramètres structuraux étaient plus reproductibles que les paramètres de densité osseuse. Les groupes KL2, KL3 et KL4 ont montré des valeurs significativement plus élevées de densité osseuse par rapport au groupe témoin (KL0), notamment au niveau de la plaque osseuse sous-chondrale médiale et de l’os trabéculaire. Les paramètres basés sur LBP ont révélé des différences significatives entre les groupes, en particulier dans les os trabéculaires et le fémur. Les paramètres Laplacien ont également montré des différences marquées entre les groupes KL0, KL2, KL3 et KL4.

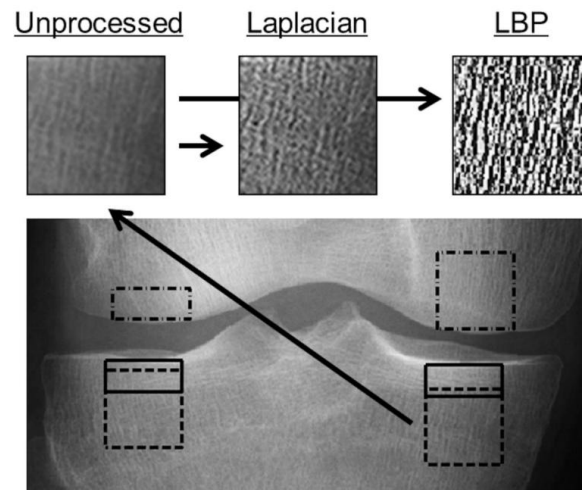


Figure 3.2 – Schéma décrivant l’emplacement des ROIs utilisées dans [Hirvasniemi et al., 2014]. Les rectangles noirs aux lignes continues représentent les ROIs de la plaque osseuse sous-chondrale. Les carrés noirs aux lignes pointillées représentent les ROIs de l’os trabéculaire sous-chondral tibial. Les ROIs du fémur sont représentées par des rectangles noirs aux lignes mixtes (point-tiret)

Le même auteur a également proposé une autre étude [Hirvasniemi et al., 2019] portant sur l’utilisation des radiographies du genou pour évaluer la densité et la texture osseuse chez des patients souffrant d’arthrose du genou, en se concentrant sur les changements dans la densité et la texture osseuse. Un

total de 213 radiographies de 109 sujets a été analysé à l'aide d'un système DigitalDiagnost, avec des radiographies postéro-antérieures bilatérales prises avec les genoux en semi-flexion. Les images ont été post-traitées en utilisant deux méthodes : un traitement minimal et un traitement clinique par défaut. Selon la figure 3.3 les ROIs ont été identifiées de manière cohérente sur des images post-traitées, avec deux placées sous l'interface cartilage-os et seize dans l'os trabéculaire sous-chondral. La densité osseuse a été évaluée à partir de la valeur moyenne des niveaux de gris dans des ROI et par mesure de l'épaisseur d'un étalon en aluminium, ajustée par un polynôme. La texture osseuse a été étudiée via la technique FSA, permettant de calculer la dimension fractale de la texture. Un modèle d'apprentissage automatique a ensuite classé les sujets selon leurs caractéristiques osseuses, afin de distinguer ceux atteints d'arthrose ou de lésions médullaires de la moelle osseuse. Les résultats ont révélé des différences significatives de densité et de texture osseuses entre les sujets sains et ceux atteints d'arthrose, suggérant que ces paramètres pourraient être utilisés pour évaluer la progression de l'arthrose du genou.

Cependant, l'étude présente deux limitations majeures. La première réside dans la taille relativement petite de l'échantillon, ce qui limite la généralisation des résultats à d'autres populations. La deuxième limitation concerne l'utilisation de deux méthodes de post-traitement des images, ce qui pourrait entraîner des variations dans la précision des mesures de densité et de textures osseuses. Ces aspects soulignent la nécessité de futures études avec des échantillons plus grands et une standardisation des techniques d'imagerie pour confirmer et étendre ces résultats.

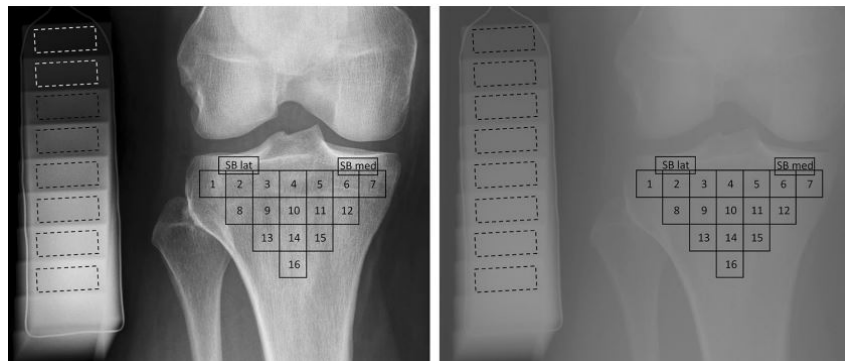


Figure 3.3 – Les ROIs ont été identifiées de manière cohérente sur des images avec différents niveaux de post-traitement. Deux ROIs ont été positionnées sous l'interface cartilage-os dans les plateaux tibiaux médial et latéral, tandis que seize autres ont été placées sous la zone dense de l'os trabéculaire sous-chondral. [Hirvasniemi et al., 2019]

[Janvier et al., 2017] ont proposé une méthode pour prédire l'incidence de la gonarthrose en analysant la Trabecular Bone Texture (TBT) dans les radiographies. À partir de 344 images issues de l'OAI, seize ROIs ont été extraites par image. Une analyse fractale a été utilisée pour calculer le paramètre fractal H , permettant de caractériser les variations locales dans chaque ROI. La figure 3.4 illustre le patchwork des ROIs sélectionnées. Un modèle de régression logistique a évalué la capacité prédictive de cette méthode. Les résultats ont montré que la combinaison des paramètres TBT avec des variables cliniques (âge, sexe et

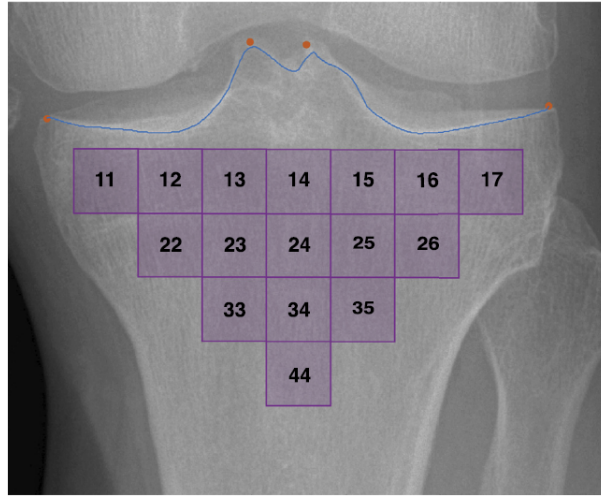


Figure 3.4 – Illustration du patchwork des ROIs extraites sur l'ensemble des radiographies de la base OAI [Janvier et al., 2017]

Indice de Masse Corporelle (IMC)) améliore significativement la précision des prédictions. L'Area Under the Curve (AUC) est passée de 0,57 avec des mesures traditionnelles à 0,69 en intégrant les données TBT, démontrant le potentiel de cette approche pour prédire l'apparition de la gonarthrose.

[Hladůvka et al., 2017] proposent une méthode avancée d'analyse des textures fémorales pour prédire la gonarthrose à partir de radiographies à haute résolution. Les auteurs ont utilisé un échantillon de 153 radiographies numériques de femmes caucasiennes, dont 67 cas d'arthrose et 86 témoins, avec un âge compris entre 43,5 et 81 ans. Des ROIs ont été semi-automatiquement placées sur les radiographies, ciblant principalement le tibia proximal et le fémur distal, des zones où les changements arthrosiques sont les plus apparents, comme montré dans la figure 3.5. L'analyse de texture a porté sur le coefficient de Hurst et l'entropie, qui caractérisent respectivement la dimension fractale de la texture osseuse et la variation de l'information.

Les auteurs ont appliqué des techniques statistiques et d'apprentissage automatique pour évaluer l'efficacité de différentes combinaisons de descripteurs de texture. Un modèle combinant des descripteurs fémoraux, entropiques et standard a permis d'obtenir une aire sous la courbe AUC de 0,85, indiquant une amélioration significative par rapport aux méthodes existantes pour prédire la gonarthrose. Cette approche a montré un potentiel prometteur pour améliorer la détection précoce de l'arthrose en utilisant des radiographies du genou.

Cependant, l'étude présente certaines limites qui pourraient influencer les résultats. Tout d'abord, la taille modérée de l'échantillon (153 radiographies) réduit la robustesse des conclusions et la capacité de généralisation des résultats à une population plus large. De plus, la population étudiée était limitée aux femmes de race blanche, ce qui restreint l'applicabilité des résultats à d'autres groupes démographiques, notamment les hommes et les autres ethnies. La segmentation de la population par d'autres facteurs

tels que l'âge, l'IMC ou l'alignement des genoux n'a pas été possible, ce qui limite la compréhension des influences potentielles de ces facteurs sur l'efficacité des méthodes proposées. Enfin, la subjectivité de la classification des radiographies par trois radiologues indépendants peut introduire des variations dans l'interprétation des résultats.

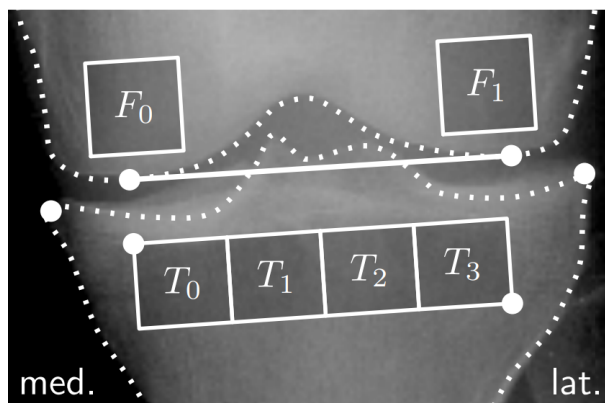


Figure 3.5 – Ancrage des ROIs [Hladůvka et al., 2017].

[Riad et al., 2018] ont aussi proposé une méthode basée sur la décomposition par ondelettes complexes pour analyser la texture des radiographies du genou dans le cadre de la détection de la gonarthrose. Dans cette étude, 688 radiographies provenant de la base de données OAI ont été utilisées. Trois ROIs carrées de 128×128 pixels situées dans différentes zones du tibia ont été extraites, comme illustré dans la figure 3.6. Le processus comprend plusieurs étapes : un prétraitement de l'image pour conserver les informations importantes sur la texture osseuse, suivi d'une transformation par ondelettes complexes. La phase relative est ensuite calculée et modélisée à l'aide de deux distributions circulaires. Les paramètres estimés sont utilisés dans une tâche de classification pour distinguer les genoux normaux des genoux présentant une arthrose minimale. Les résultats de la classification indiquent une spécificité de 85,47%, une sensibilité de 75,29% et une précision de 80,38%, montrant ainsi que cette méthode est un outil prometteur pour la détection de la gonarthrose par rapport aux méthodes traditionnelles d'extraction de caractéristiques.

[Ljuhar et al., 2018] ont utilisés des radiographies du genou de 574 participants (230 femmes et 344 hommes) issus de la base de données MOST, tous sans arthrose initiale (score KL0). L'objectif était de prédire l'incidence de l'arthrose au 84ème mois à l'aide du Joint Space Width (JSW), et Joint Shape Analysis (JSA). Parmi les participants, 41 femmes et 79 hommes ont développé une arthrose incidente (score $KL \geq 1$), tandis que les autres sont restés à KL0. La figure 3.7 représente la zone sous-chondrale sélectionnée pour les analyses texturales comprenant 4 ROI dans le tibia proximal et une sur chaque condyle du fémur distal. Les résultats ont montré que les mesures traditionnelles seules offraient une aire sous la courbe AUC de 0,67 pour les femmes et 0,61 pour les hommes. Cependant, lorsqu'elles étaient combinées avec l'analyse de texture, l'AUC s'améliorait significativement pour atteindre 0,80 pour les femmes et 0,69 pour les hommes, démontrant une précision prédictive accrue.

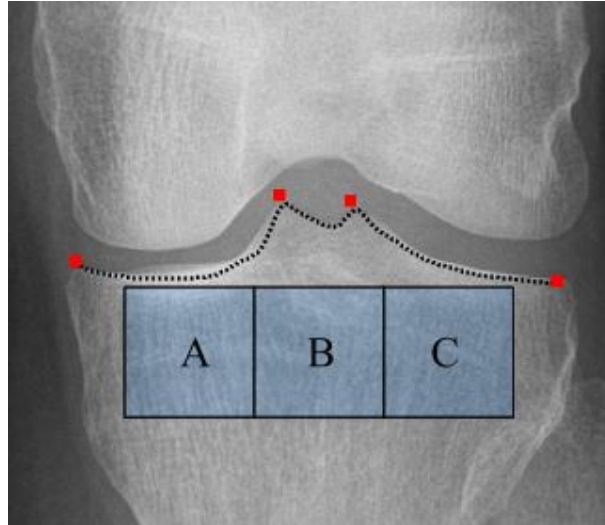


Figure 3.6 – ROIs sélectionnées à l’aide d’un algorithme semi-automatique. Les points rouges représentent les repères anatomiques définis manuellement et la ligne pointillée indique le bord tibial [Riad et al., 2018].

Cependant, une limite importante est liée à l’échantillon, qui était issu d’un seul centre, limitant potentiellement la généralisation des résultats à des populations plus diverses.

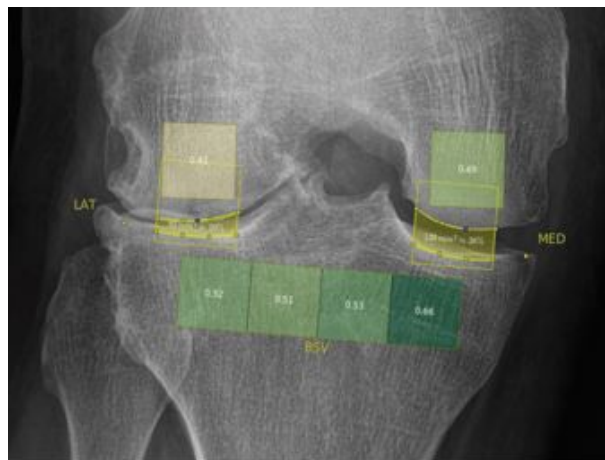


Figure 3.7 – La zone sous-chondrale sélectionnée pour les analyses texturales comprenait 4 régions d’intérêt (ROI) dans le tibia proximal et une sur chaque condyle du fémur distal [Ljuhar et al., 2018]

L’étude de [Brahim et al., 2019] présente un CAD pour la détection précoce de l’arthrose du genou à partir d’images radiographiques et d’apprentissage automatique. Le système a été testé sur 1024 radiographies de genoux issues de la base de données OAI, comprenant 514 genoux sains et 514 genoux atteints d’arthrose (OA). Trois régions d’intérêt (ROI) de 128×128 pixels ont été extraites de chaque radiographie, représentant les zones médiale, centrale et latérale du genou.

Le traitement des images a été réalisé en plusieurs étapes : un prétraitement dans le domaine de Fourier

pour stationnariser les images, une quantification des niveaux de gris et une normalisation d'intensité à l'aide de Multivariate Linear Regression (MLR) pour réduire la variabilité entre les groupes sains et OA. Une analyse par Independent Component Analysis (ICA) va ensuite permis de réduire la dimensionnalité des données. Enfin, des classifieurs Naive Bayes et Random Forest ont été utilisés pour la classification des genoux. La figure 3.8 représente une ROI du tissu trabéculaire tibial et des différents prétraitement appliqués.

Les résultats ont montré que la normalisation par MLR a amélioré la séparation entre les groupes témoins et OA, avec des performances de classification de 82,98 % de précision, 87,15% de sensibilité et 80,65% de spécificité. Cette étude met en évidence l'efficacité d'un système CAD basé sur l'apprentissage automatique pour détecter l'arthrose du genou de manière précise et fiable.

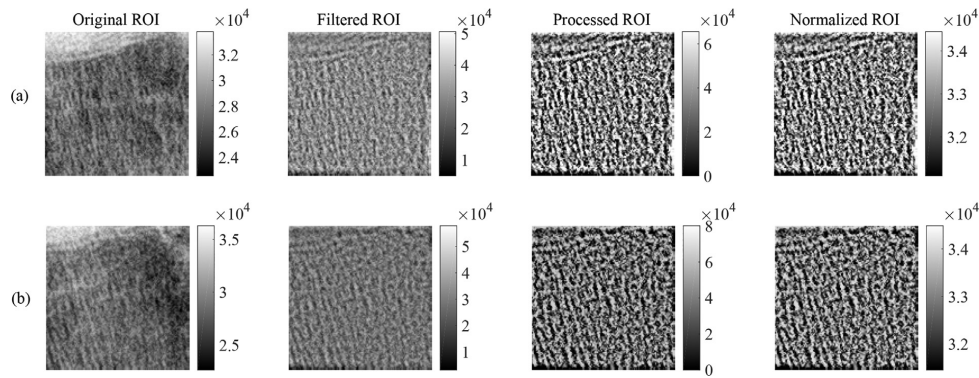


Figure 3.8 – Une ROI représentative du tissu trabéculaire tibial et des différents pré-traitements appliqués dans [Brahim et al., 2019]. ROI de (a) un sujet témoin et (b) un patient atteint d'arthrose (OA).

3.3.3.B Approches basées sur les caractéristiques de forme

En ce qui concerne les approches basées sur les caractéristiques de forme, certaines études ont trouvé des résultats significatifs. Parmi celles-ci, l'étude menée par [Haverkamp et al., 2011], dont l'objectif était d'examiner le rôle de la forme des articulations dans la gonarthrose en identifiant les aspects spécifiques de la forme osseuse qui diffèrent entre les genoux atteints d'arthrose et ceux des témoins. Pour ce faire, les chercheurs ont utilisé des Statistical Shape Model (SSM) pour analyser les radiographies de 609 femmes (1 218 genoux) extraites de l'étude Rotterdam. Afin de mieux comprendre les variations de forme associées à l'arthrose, l'approche proposée permet de tracer les contours du fémur, du tibia, de la patella et du condyle médial à travers 105 points comme illustré dans la figure 3.9. Les résultats obtenus ont montré que les genoux arthrosiques étaient caractérisés par une augmentation de la largeur des os fémoral et tibial (mode 2), ainsi qu'une réduction de la flexion du genou lors de la radiographie (mode 4), ce qui suggère que les genoux arthrosiques sont plus étendus. De plus, un plateau tibial latéral élevé (mode 15) a été observé chez les patients souffrant de douleurs, ce qui pourrait être lié à l'intensité des symptômes. En conclusion, ces résultats montrent que la forme des genoux joue un rôle crucial dans l'arthrose, ouvrant

la voie à des biomarqueurs d'imagerie pour surveiller ou prédire l'évolution de l'arthrose du genou.

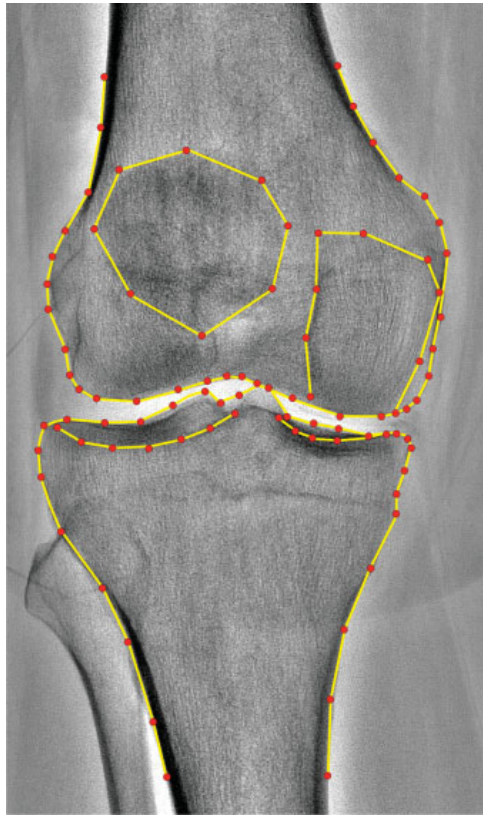


Figure 3.9 – Les 105 points pris en compte par le modèle statistique de forme [Haverkamp et al., 2011] pour analyser la forme des radiographies du genou.

Récemment, [Rayegan et al., 2023] ont développés un logiciel automatisé, Osteoarthritis Data and Imaging Analysis (OADIA) (Figure 3.11), pour mesurer les paramètres de l'arthrose du genou à partir de radiographies, en comparaison avec des méthodes manuelles KIDA. Basée sur les données de 1 002 participants de l'étude Cohort Hip and Cohort Knee (CHECK), elle a évalué des paramètres tels que la largeur de l'espace articulaire, l'angle de convergence articulaire et les intensités osseuses sous-chondrales comme illustré dans la figure 3.10 Avec 1 743 radiographies incluses, les résultats ont montré une forte fiabilité pour la plupart des mesures, bien que des limites aient été observées pour certains paramètres spécifiques, OADIA promet de simplifier l'analyse de grandes cohortes.

3.3.3.C Approches basées sur les caractéristiques de forme et de texture

Concernant les méthodes basées à la fois sur les caractéristiques de forme et de texture, on trouve [Thomson et al., 2015] qui a proposé Une segmentation automatisée extrait les caractéristiques de forme des os, tandis que des classificateurs Random Forest analysent séparément les caractéristiques de forme et de texture. Pour la texture, deux approches sont utilisées : la signature fractale, qui calcule les dimensions

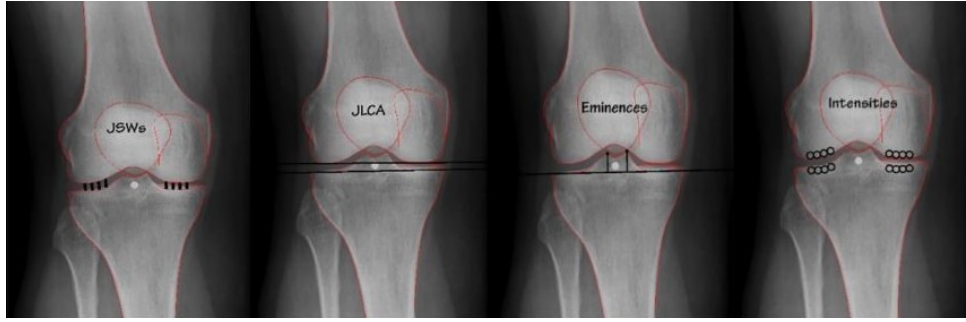


Figure 3.10 – Paramètres mesurés sur les radiographies du genou, de gauche à droite : largeur de l’espace articulaire (JSW), angle de convergence de la ligne articulaire (JLCA), hauteurs des éminences et intensités de l’os sous-chondral, utilisé par [Rayegan et al., 2023]



Figure 3.11 – Interface utilisateur du logiciel ODIA développé par [Rayegan et al., 2023]

fractales à partir des variations de pixels, et les rapports de pixels, une méthode plus simple basée sur les intensités des zones tibiales. Comme illustré dans la figre 3.12 un modèle statistique de forme (74 points) capte les variations osseuses, renforçant la classification. Ces modèles ont été testés sur des radiographies de la base OAI et évalués avec des courbes ROC, montrant des performances solides dans la détection de l’OA. L’étude est limitée par la diversité restreinte des données (500 radiographies), ce qui réduit sa généralisation à d’autres basess de données, et par une analyse texturale incomplète, qui ne tiet pas compte pleinement des variations complexes liées à l’arthrose.

On trouve aussi dans [Minciullo et al., 2017] une méthode entièrement automatisée pour détecter la gonarthrose à partir de radiographies latérales. Elle repose sur un modèle Random Forest Combined with Landmark Modeling (RFCLM) pour localiser précisément les points clés des structures osseuses et analyser leur forme. Un modèle de forme global, constitué de 102 points répartis sur la rotule, les

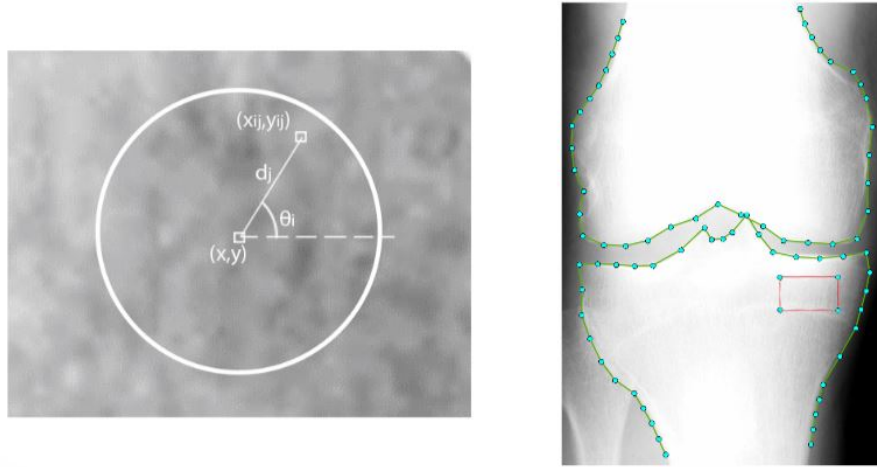


Figure 3.12 – Exemple d'échantillonnage d'une région circulaire pour calculer la signature fractale (à gauche). Exemple des 74 points de repère et des 4 points calculés à partir de ceux-ci, délimitant la région d'intérêt (ROI) utilisée pour évaluer la texture du tibia (à droite). où x est un vecteur représentant la forme dans un cadre de référence, $\hat{x}$ est la forme moyenne, P est l'ensemble des vecteurs propres correspondant aux t valeurs propres les plus élevées, qui décrivent différents modes de variation, et b est un ensemble de valeurs de paramètres. [Thomson et al., 2015]

condyles fémoraux, et le tibia comme illustré dans la figure 3.13. Ce model est utilisé pour caractériser la morphologie du genou. La méthode intègre une Principal Component Analysis (PCA) pour réduire les dimensions et identifier les variations essentielles de forme.

Les forêts aléatoires sont utilisées pour détecter et délimiter les os, tandis qu'une classification binaire (genoux sains ou arthrosiques) et multiclassée (grades de KL) est effectuée à partir des caractéristiques extraites. Les performances du système automatisé sont comparables, voire supérieures, à celles des annotations manuelles, démontrant l'efficacité de cette approche pour une détection précise et rapide de l'arthrose. Malgré cela, cette étude identifie des limites telles que la complexité des chevauchements osseux, la variabilité des performances selon les structures analysées et la dépendance aux paramètres de forme, soulignant des défis pour améliorer la détection automatisée de l'arthrose.

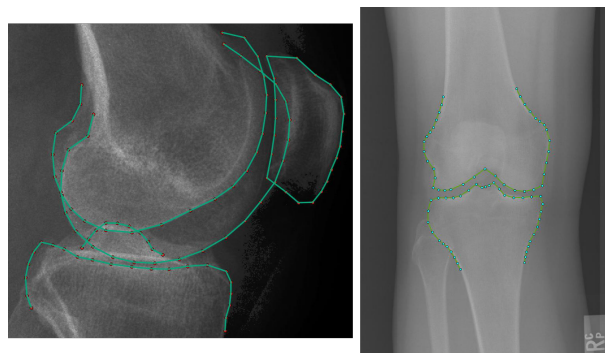


Figure 3.13 – Un exemple des 102 points de repère utilisés pour construire le mode de forme dans [Minciullo et al., 2017]

3.3.3.D Approches basées sur des techniques d'apprentissage profond

Au cours des dernières décennies, l'émergence de nouvelles technologies d'analyse d'image et de techniques d'apprentissage profond a révolutionné le domaine de l'imagerie médicale, notamment pour la détection, la segmentation et la classification. Plusieurs études ont été proposées pour le diagnostic de la gonarthrose à partir d'images radiographiques [Antony et al., 2017, Tiulpin et al., 2018, Chen et al., 2019, Norman et al., 2019, Nguyen et al., 2020, Yeoh et al., 2023, Singh et al., 2024, Sinha and Veemaraj, 2024].

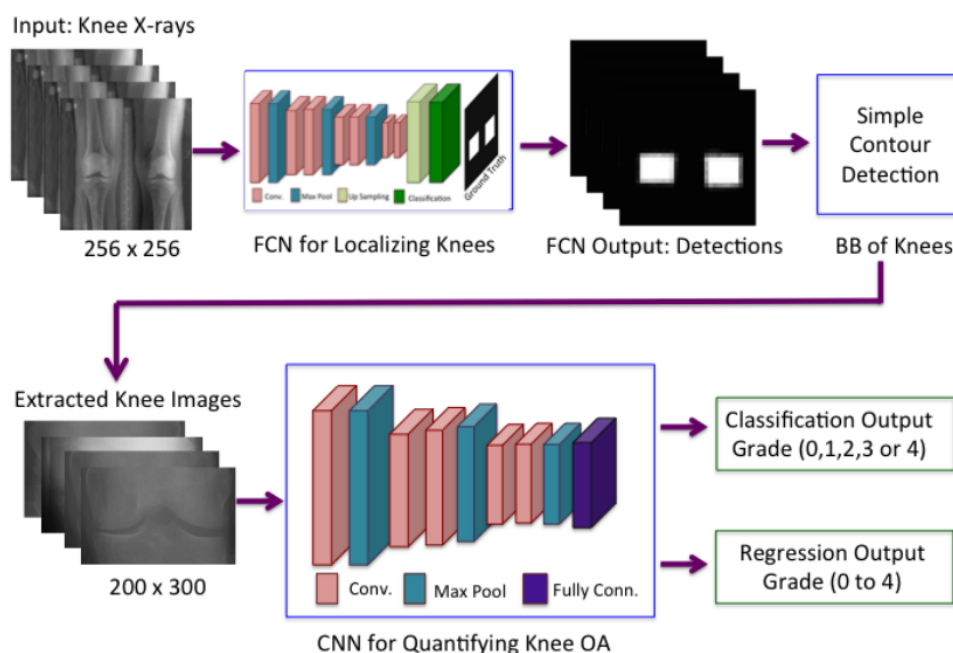


Figure 3.14 – Pipeline proposé par Antony et al. pour la détection de la gonarthrose [Antony et al., 2017]

Anthony et al. [Antony et al., 2017] ont appliqué des CNN pour évaluer la gravité de l'arthrose du genou à partir de radiographies issues de bases de données OAI et MOST. Un total de 8 892 images radiographiques ont été utilisées pour l'OAI, réparties selon les grades deKL comme suit : 3 433 images pour KL-G0, 1 589 pour KL-G1, 2 353 pour KL-G2, 1 222 pour KL-G3 et 295 pour KL-G4. Quant à la base MOST, elle a fourni 5 840 images, réparties comme suit : 2 498 images pour KL-G0, 1 018 pour KL-G1, 923 pour KL-G2, 971 pour KL-G3 et 430 pour KL-G4.

Le pipeline de l'approche proposée est illustré dans la figure 3.14. La méthode proposée commence par la détection de la ROI correspondant à l'articulation du genou, définie par une fenêtre de 300×300 pixels. Pour améliorer la précision de cette détection, les auteurs ont formé un CNN spécifiquement destiné à localiser la région du genou. La figure 3.14 illustre le pipeline proposé, incluant la détection de la région d'intérêt ainsi que les étapes suivantes de classification et de régression.

Pour l'évaluation de la gravité de l'arthrose, deux approches ont été utilisées. La première consiste à

entraîner le CNN uniquement pour la classification des images selon les grades KL. La deuxième approche, plus avancée, consiste à entraîner un CNN pour la classification et la régression conjointes, en traitant le grade KL comme une variable continue. Cette approche permet de quantifier plus finement la gravité de l'arthrose.

Les résultats expérimentaux ont montré que l'approche combinée (classification et régression) offrait une meilleure précision pour la classification des grades de l'arthrose et une réduction de l'erreur quadratique moyenne. En revanche, les grades inférieurs (KL-G0, KL-G1 et KL-G2) sont plus difficiles à distinguer en raison de leurs différences visuelles subtiles.

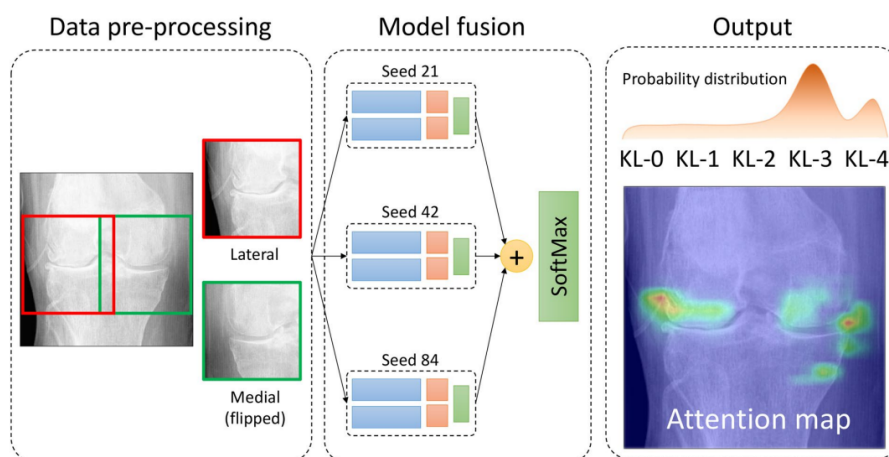


Figure 3.15 – Le pipeline de classification proposé par [Tiulpin et al., 2018]

Tiulpin et al. [Tiulpin et al., 2018] ont développé une méthode basée sur le réseau de neurones Siamese pour le diagnostic automatisé de l'arthrose du genou à partir de radiographies. Leur approche exploite la symétrie visuelle du genou et a été testée sur 18 376 images de l'étude MOST et 8 917 images de l'OAI. Selon le pipeline proposé 3.15, la méthode commence par la localisation et le recadrage de la région du genou, suivis de l'extraction de deux patches représentant les côtés médial et latéral de l'articulation. Le patch médial est retourné horizontalement pour renforcer la symétrie. Ces patches sont analysés par un réseau Siamese, composé de deux branches identiques, partageant les mêmes poids, qui extraient des vecteurs de caractéristiques combinés pour prédire le grade KL via une couche entièrement connectée. Trois modèles ont été entraînés avec des prédictions fusionnées pour améliorer la précision. Des cartes d'attention ont été utilisées pour visualiser les zones influençant les décisions du modèle, offrant une meilleure interprétabilité comme montré dans la figure 3.15. La méthode a montré une classification robuste des grades KL.

Dans la continuité de ces travaux, une autre étude a exploré des méthodologies alternatives, [Chen et al., 2019]. La première étape de l'étude consiste à détecter les articulations du genou à l'aide d'un modèle personnalisé de YOLOv2, qui est particulièrement adapté pour détecter les articulations sur des radiographies. Les performances mesurées ont atteint un indice de Jaccard moyen de 0,858 et un rappel de

92,2 % à un seuil de 0,75. Une fois l'articulation localisée, plusieurs architectures de CNN comme ResNet, VGG, DenseNet et InceptionV3 ont été utilisées pour classer les images selon le système de notation KL comme montré dans la figure 3.16. Pour améliorer la précision, l'étude introduit une fonction de perte ordinale ajustable, qui pénalise davantage les erreurs de classification éloignées de la vraie noteKL. Cette approche améliore les performances des CNN par rapport à la perte d'entropie croisée classique. Les résultats expérimentaux montrent que le modèle VGG-19, affiné avec cette fonction de perte ordinale, atteint une précision de classification de 69,7% et une erreur Mean Absolute Error (MAE) de 0,344. Les images radiographiques utilisées proviennent de la base OAI, qui offre une cohorte diversifiée d'individus âgés de 45 à 79 ans, avec un accent sur les images radiographiques bilatérales à flexion fixe PA pour cette analyse.

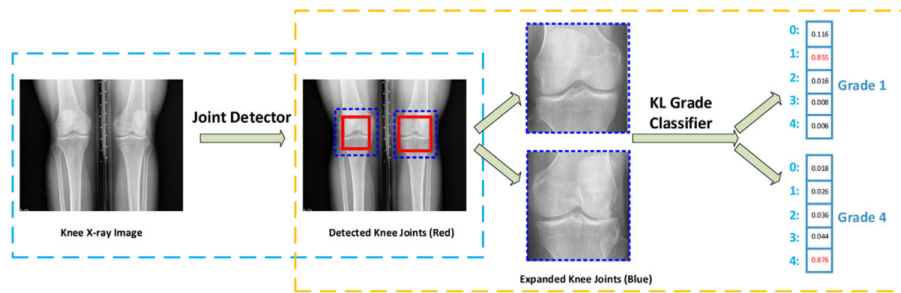


Figure 3.16 – Illustration du pipeline d'évaluation de la gravité de l'articulation du genou proposé dans [Chen et al., 2019]. Il comprend la détection de l'articulation du genou et la classification du grade KL du genou.

En poursuivant les efforts récents en apprentissage automatique pour améliorer la détection précoce de la gonarthrose, [Nasser et al., 2020] ont introduit le modèle Discriminative Regularized Auto-Encoder (DRAE). Ce modèle est conçu pour extraire des caractéristiques discriminantes pertinentes, en particulier à partir des textures présentes dans des ROIs sélectionnées, afin d'améliorer la classification des stades de la gonarthrose. Le modèle DRAE analyse cinq ROIs texturées de 32×32 pixels (R1, R2, R3, R4, R5), localisées sous le plateau tibial à partir des radiographies. Ces ROIs représentent les zones médiales (R1, R2), centrales (R3) et latérales (R4, R5) du genou.

Le processus commence par la segmentation des images radiographiques, où des repères anatomiques, tels que les épines tibiales et les extrémités médiale et latérale du tibia, sont identifiés manuellement (points rouge, figure 3.17). Ensuite, la région sous-chondrale est détectée automatiquement comme le chemin lumineux passant par ces repères anatomiques. Les ROIs sont extraites sous la bordure inférieure de l'os sous-chondral. Une fois les ROIs définies, elles sont utilisées comme entrée pour le modèle DRAE, qui capture les caractéristiques texturales complexes et discriminantes. Le modèle encode ces caractéristiques dans un espace latent tout en intégrant une perte discriminative, conçue pour maximiser la séparation entre les classes des stades de la gonarthrose. Enfin, les représentations apprises sont utilisées pour classer les ROIs en fonction des stades de gravité de l'arthrose, selon le système de notation KL.

Cette méthodologie innovante permet de coupler l'analyse texturale fine avec une classification précise, renforçant la capacité du DRAE à fournir des diagnostics fiables et objectifs. Validée sur les données de la base OAI, cette approche a démontré une amélioration notable des performances par rapport aux méthodes traditionnelles.

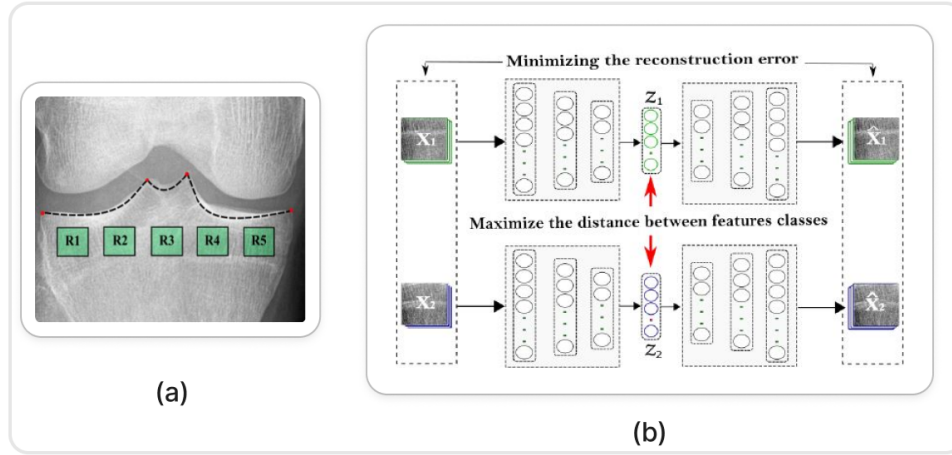


Figure 3.17 – (a) Les cinq ROIs extraites. (b) Architecture du modèle d'apprentissage DRAE proposée par [Nasser et al., 2020].

Dans [Nguyen et al., 2020], les auteurs proposent une méthode semi-supervisée innovante appelée Semixup, destinée à améliorer l'évaluation de la gravité de la gonarthrose à partir de radiographies. Cette approche s'appuie sur une architecture d'un réseau Siamois figure 3.18, adaptée des travaux de [Tiulpin et al., 2018] pour analyser simultanément des paires d'images (latérale et médiale) du genou. Des modifications ont été apportées au CNN, notamment l'utilisation de couches convolutives 3×3 optimisées, combinées à des fonctions de normalisation des instances et d'activation ReLU. Ces ajustements permettent d'améliorer l'extraction de caractéristiques pertinentes.

Au lieu des couches de max-pooling traditionnelles, des convolutions stridées sont utilisées pour préserver les dépendances essentielles entre les caractéristiques. Une stratégie de regroupement innovante, appelée SAM (Spatial Attention Mechanism) Pooling, est également introduite, offrant une alternative plus efficace au Global Average Pooling (GAP). Cette méthode conserve des informations plus riches issues des cartes de caractéristiques.

Semixup intègre également une régularisation de cohérence, exploitant des données étiquetées et non étiquetées pour renforcer la robustesse de l'apprentissage. Les évaluations expérimentales montrent que cette approche atteint des performances comparables à celles des modèles supervisés avec un nombre significativement réduit des données étiquetées, soulignant son efficacité dans un contexte semi-supervisé.

L'étude utilise la base de données OAI et MOST. En particulier, le modèle a surpassé le modèle supervisé correspondant lorsqu'au moins 100 échantillons par grade de KL étaient utilisés. Cependant, les auteurs notent que les performances se stabilisaient lorsque le nombre de données étiquetées dépassait 1 000 images par grade KL.

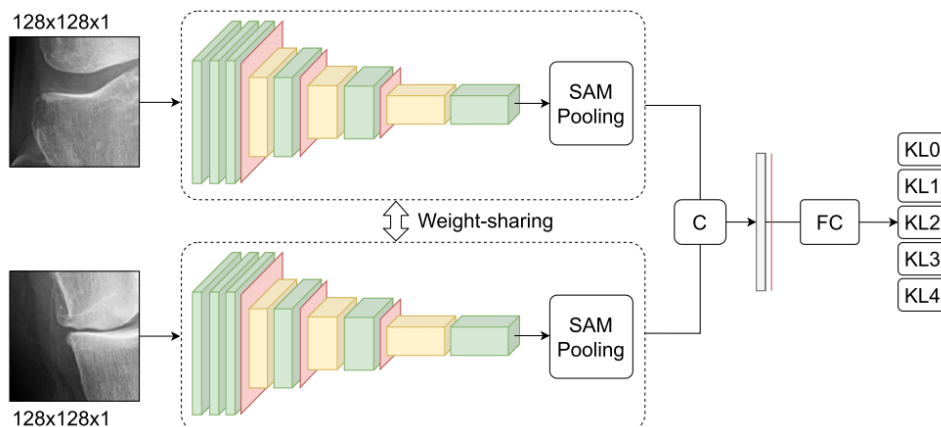


Figure 3.18 – Le modèle Siamois utilisé par [Nguyen et al., 2020] pour prédire les grades KL correspondant aux côtés de l'articulation du genou (latéral et médial) des entrées.

Dans [Gu et al., 2022], les auteurs ont aussi proposé une méthodologie approfondie pour évaluer la gravité de la gonarthrose, en utilisant un algorithme basé sur l'apprentissage profond. L'étude commence par la préparation minutieuse des ensembles de données et des annotations nécessaires à la formation et à la validation de l'algorithme, garantissant ainsi que l'algorithme apprenne efficacement à partir des données. L'algorithme proposé suit un processus en cinq étapes pour analyser les radiographies PosteroAnterior (PA) des genoux et déterminer la sévérité de la gonarthrose.

Le pipeline de la méthode proposée est illustré dans la figure 3.19. La première étape du processus consiste en un prétraitement des images pour améliorer leur qualité, notamment par normalisation et amélioration de la netteté. Ensuite, un modèle YOLO V3-tiny est utilisé pour localiser précisément les articulations du genou sur les radiographies, afin de focaliser l'analyse sur les zones pertinentes. La troisième étape implique l'utilisation d'un CNN pour une évaluation initiale de la sévérité de l'arthrose, en classifiant les images selon les caractéristiques apprises par l'algorithme.

Une fois les articulations localisées, l'algorithme segmente ces régions et mesure le rétrécissement de JSN, un indicateur clé pour déterminer la gravité de la gonarthrose. La dernière étape combine ces informations avec l'évaluation initiale pour déterminer un score final de KL, qui est utilisé pour catégoriser la sévérité de l'arthrose du genou.

Récemment, [Jahan et al., 2024] ont proposés plusieurs approches développées dans le cadre KOA-CCThet pour l'évaluation de la gravité de la gonarthrose. L'une des étapes clés de cette méthodologie est l'agrégation des données, où les auteurs ont combiné quatre ensembles de données distincts pour consti-

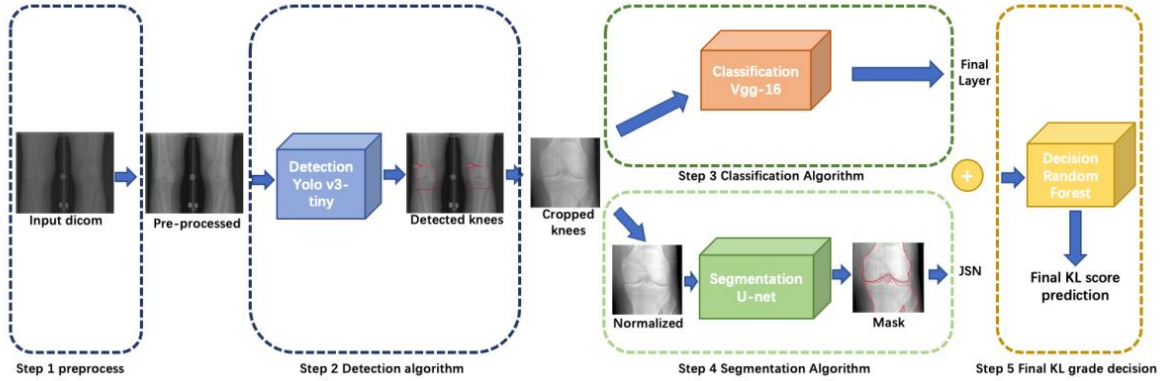


Figure 3.19 – Pipeline de prédiction de la gravité de la gonarthrose proposé par [Gu et al., 2022]

tuer un large ensemble de 110 232 images radiographiques brutes, essentiel pour entraîner un modèle robuste. Pour augmenter la taille de cet ensemble de données et améliorer sa diversité, un Deep Convolutional Generative Adversarial Network (DCGAN) a été utilisé pour générer des images supplémentaires, augmentant ainsi la capacité du modèle à généraliser.

Le prétraitement des images joue également un rôle crucial dans l'amélioration des performances. Adaptive Histogram Equalization (AHE) permet d'améliorer le contraste des images. De plus, les images sont redimensionnées, assurant ainsi la cohérence nécessaire pour le traitement dans le réseau.

Comme montré dans la figure 3.20, le modèle Knee Osteoarthritis Classification and Clinical Treatment Network (KOACCTNet) repose sur un Computed Clinical Treatment (CCT), conçu spécifiquement pour gérer de grands ensembles de données tout en optimisant les temps d'entraînement. L'optimisation des hyperparamètres a été un processus clé pour ajuster le modèle, en testant différentes configurations afin d'améliorer son efficacité.

Enfin, le modèle a été évalué en le comparant à d'autres modèles de pointe, tels que le Swin Transformer, le Vision Transformer et le Involutional Neural Network. KOACCTNet a montré des performances exceptionnelles, atteignant une précision de 94,58 %, surpassant les autres modèles évalués. Ces méthodes combinées permettent à KOA-CCTNet d'offrir une évaluation précise et efficace des grades d'arthrose du genou à partir de radiographies.

Conclusion

Ce chapitre a exploré les avancées récentes dans l'utilisation des techniques d'apprentissage profond pour la détection précoce et l'évaluation de la gravité de la gonarthrose, en se concentrant sur des méthodes innovantes et des modèles basés sur des réseaux de neurones. Plusieurs approches ont été présentées, mettant en évidence l'évolution des techniques utilisées pour améliorer la précision des diagnostics à

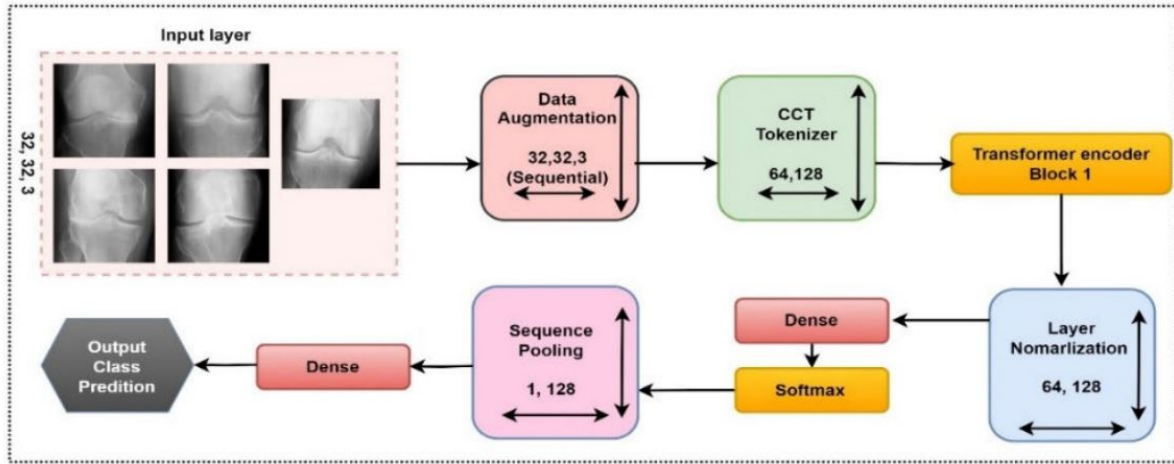


Figure 3.20 – Architecture du modèle KOA-CCTNet proposé par [Jahan et al., 2024]

partir de radiographies, en particulier celles qui exploitent les données non étiquetées via l'apprentissage semi-supervisé, ainsi que les modèles plus complexes et spécialisés comme les réseaux Siamois, les modèles transformeurs et les Auto-Encoder (AE).

Les méthodes de prétraitement d'images, la segmentation des articulations du genou et l'intégration de mécanismes de régularisation discriminative, telles que le DRAE, ont permis d'améliorer les performances des systèmes de diagnostic. En outre, l'augmentation des données à l'aide de techniques comme les réseaux génératifs antagonistes Generative Adversarial Network (GAN) a joué un rôle clé dans l'enrichissement des ensembles de données et l'amélioration de la robustesse des modèles.

Les approches récentes, telles que l'intégration du modèle KOA-CCTNet, illustrent l'importance de l'utilisation de grandes bases de données radiographiques, le développement de modèles d'optimisation des hyperparamètres, et la comparaison rigoureuse des performances à travers des études expérimentales. Les résultats obtenus montrent des améliorations notables en termes de précision des prédictions, surpassant parfois les évaluations traditionnelles par les radiologues.

Cependant, malgré ces progrès, plusieurs défis demeurent, notamment en ce qui concerne la gestion des données non étiquetées, la généralisation des modèles à différentes populations et la réduction de la subjectivité humaine dans l'interprétation des radiographies. Ces questions restent au cœur des préoccupations pour les recherches futures, visant à rendre les systèmes de diagnostic plus robustes et accessibles.

Dans l'ensemble, les contributions récentes offrent des perspectives prometteuses pour améliorer le diagnostic de l'arthrose du genou, ouvrant la voie à des outils d'Intelligence Artificielle (AI) plus performants et à une prise en charge clinique optimisée des patients.

4

Introduction à l'Apprentissage Profond

Sommaire

4.1	Introduction	49
4.2	Modèles de Base en DL	52
4.3	Conclusion	64

4.1 Introduction

Le DL, ou apprentissage profond, est une branche de l'intelligence artificielle qui repose sur l'utilisation de réseaux de neurones artificiels. Inspiré du fonctionnement du cerveau humain, il vise à extraire et apprendre des représentations hiérarchiques complexes des données. Grâce à des architectures profondes comportant plusieurs couches de traitement, les modèles de deep learning sont capables de reconnaître des motifs et d'analyser des données avec une précision remarquable.

Au cours des dernières décennies, le deep learning a connu une progression fulgurante grâce à trois facteurs principaux :

- Accès à de grandes quantités de données : Des bases de données massives, notamment dans le domaine médical (comme des images IRM ou des radiographies), permettent d'entraîner des modèles robustes.
- Amélioration des capacités de calcul : L'avènement des processeurs graphiques (GPUs) et des accélérateurs spécifiques (TPUs) a permis de traiter des réseaux de neurones complexes en des temps raisonnables.
- Avancées dans les architectures de modèles : Des réseaux tels que les CNNs, GANs et Transformers ont révolutionné des domaines comme la vision par ordinateur et le traitement du langage naturel.

Dans le domaine médical, le deep learning joue un rôle crucial, notamment dans l'analyse d'images médicales. Qu'il s'agisse de détecter des pathologies, de segmenter des tissus spécifiques ou de prédire l'évolution d'une maladie, ces modèles apportent des solutions précises, rapides et parfois même supérieures aux méthodes traditionnelles. Ils offrent ainsi un potentiel considérable pour améliorer le diagnostic, réduire les erreurs humaines et optimiser les traitements.

Dans ce chapitre, nous allons :

- Présenter les principaux modèles de deep learning, leurs concepts fondamentaux et leur fonctionnement.
- Explorer leurs applications dans l'analyse d'images médicales, comme la classification, la segmentation et la génération d'images.
- Mettre en lumière les techniques utilisées pour améliorer les performances des modèles, comme l'apprentissage par transfert, la régularisation et l'augmentation des données.

Perceptron Multi-Couche (MLP)

Multi-Layer Perceptron (MLP) est un réseau de neurones artificiels composé de plusieurs couches de neurones interconnectées. Chaque couche transforme ses entrées en sorties via une combinaison linéaire

suivie d'une fonction d'activation non linéaire. Cela permet au MLP de modéliser des relations non linéaires complexes.

L'architecture d'un MLP est composée des éléments suivants :

- **Couche d'entrée** : reçoit les données d'entrée. Chaque neurone correspond à une caractéristique (feature) dans les données.
- **Couches cachées** : une ou plusieurs couches intermédiaires effectuent des transformations non linéaires.
- **Couche de sortie** : produit la sortie finale. Pour une tâche de classification, une fonction Softmax est utilisée pour normaliser les sorties.

Les calculs dans une couche l du MLP se déroulent comme suit :

1. Les entrées $x^{(l)}$ subissent une transformation linéaire :

$$z^{(l)} = W^{(l)}x^{(l-1)} + b^{(l)},$$

où :

- $W^{(l)}$ est la matrice des poids de la couche l .
 - $b^{(l)}$ est le vecteur des biais de la couche l .
2. Une fonction d'activation f est appliquée pour introduire la non-linéarité :

$$x^{(l)} = f(z^{(l)}).$$

3. En sortie, pour une tâche de classification multi-classes, la probabilité de chaque classe est calculée avec Softmax :

$$y_k = \frac{\exp(z_k^{(L)})}{\sum_{j=1}^C \exp(z_j^{(L)})},$$

où $z_k^{(L)}$ est le score brut pour la classe k à la dernière couche L , et C est le nombre total de classes.

L'entraînement d'un MLP comprend les étapes suivantes :

1. **Propagation avant (Forward Propagation)** : les données traversent le réseau pour produire une prédiction.
2. **Calcul de la perte** : par exemple, pour une classification, la fonction d'entropie croisée est utilisée :

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^C y_{i,k} \log(\hat{y}_{i,k}),$$

où N est le nombre d'exemples, $y_{i,k}$ est la vraie étiquette, et $\hat{y}_{i,k}$ est la prédiction.

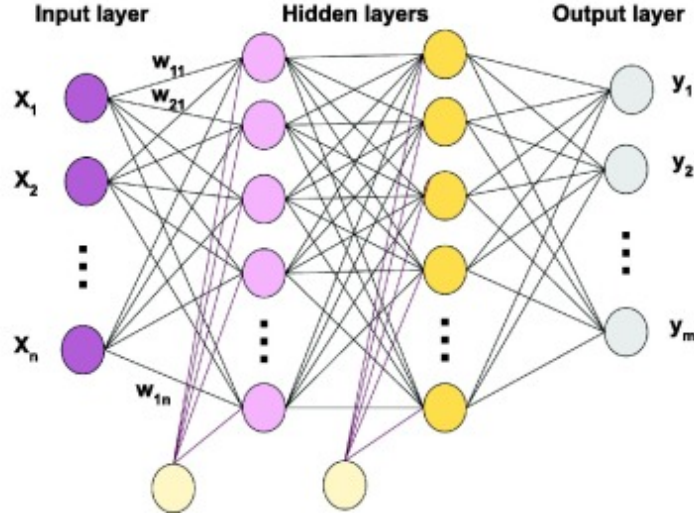


Figure 4.1 – Représentation schématique d'un MLP avec deux couches cachées. [Chan et al., 2023]

3. **Rétropropagation (Backpropagation)** : calcul des gradients des paramètres (poids et biais) en utilisant la méthode de descente de gradient :

$$W^{(l)} \leftarrow W^{(l)} - \eta \frac{\partial \mathcal{L}}{\partial W^{(l)}},$$

où η est le taux d'apprentissage.

4. **Mise à jour des paramètres** : les poids et biais sont ajustés pour minimiser la perte.

La figure 4.1 représente un MLP avec deux couches cachées, une couche d'entrée et une couche de sortie. Lors des interactions, chaque nœud affiche un certain niveau de biais. La couche d'entrée contient n variables d'entrée $X = \{x_1, x_2, \dots, x_n\}$, et la couche de sortie contient m variables de sortie $Y = \{y_1, y_2, \dots, y_m\}$.

Le nombre total de paramètres dans un MLP peut être déterminé par la formule suivante [Qaddoura et al., 2021] :

$$n \cdot h_1 + \sum_{k=1}^{N_h-1} h_k \cdot h_{k+1} + h_{N_h} \cdot m$$

où h_i représente le nombre de nœuds cachés dans la i -ième couche cachée, et N_h est le nombre total de couches cachées.

Des temps de calcul plus longs sont nécessaires pour optimiser un MLP lorsque N_h et h_k augmentent.

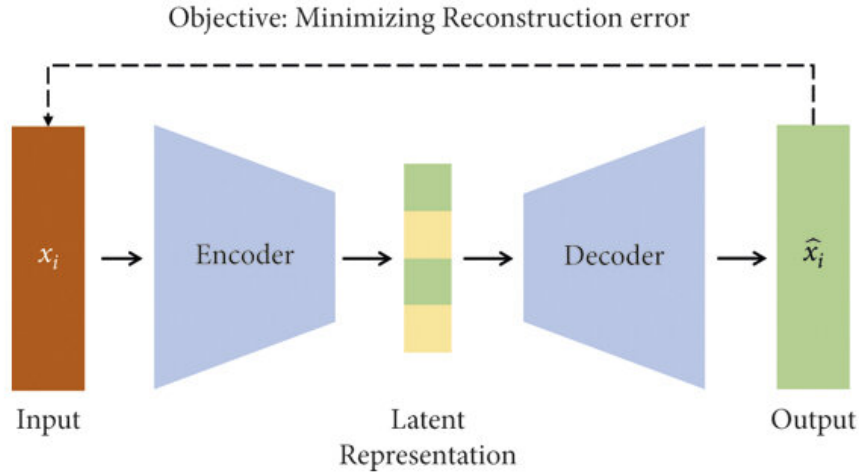


Figure 4.2 – Illustration du réseau AE traditionnel [Wang et al., 2022].

4.2 Modèles de Base en DL

Dans cette section, nous examinerons les modèles de deep learning les plus fondamentaux et leurs variantes avancées, tout en mettant en évidence leur fonctionnement et leurs applications. Les modèles présentés incluent : les Auto-Encoder (AE) utilisés pour l'apprentissage non supervisé et la réduction de dimensions, ils apprennent à encoder les données dans une représentation comprimée (cf. 4.2.1). Les Réseaux de Neurones Convolutifs (CNN) spécialisés dans le traitement des données spatiales comme les images, ils dominent les tâches de vision par ordinateur (cf. 4.2.2). Les Réseaux de neurones Récurrents (RNN) adaptés aux données séquentielles, comme le texte ou les séries temporelles, ils permettent de capturer les relations temporelles (cf. 4.2.3). Les Transformers (cf. 4.2.4) révolutionnant le traitement du langage naturel, ces modèles permettent une parallélisation efficace tout en capturant le contexte global. Et enfin les Graph Neural Networks (GNN) conçus pour manipuler des structures de données complexes telles que les graphes, ces réseaux trouvent des applications dans les réseaux sociaux et les molécules (cf. 4.2.5).

4.2.1 Autoencodeurs

Les AE sont des réseaux neuronaux non supervisés conçus pour apprendre une représentation latente (ou encodage) des données d'entrée. Leur objectif principal est de compresser les informations dans un espace de plus faible dimension, tout en préservant les caractéristiques les plus importantes. Les autoencodeurs ont trouvé des applications dans des domaines tels que la réduction de dimensionnalité, la détection d'anomalies, le pré-entraînement des réseaux profonds et la génération de données synthétiques [Hinton and Salakhutdinov, 2006].

Un autoencodeur se compose de deux parties principales comme illustré dans la figure 4.2 :

- **L’encodeur** : Cette partie compresse les données d’entrée dans une représentation latente de dimension inférieure.
- **Le décodeur** : Cette partie reconstruit les données d’origine à partir de la représentation latente.

Un AE peut être formalisé comme suit : étant donné une entrée $\mathbf{x} \in \mathbb{R}^n$, l’encodeur applique une transformation f_θ paramétrée par θ pour produire une représentation latente $\mathbf{z}$:

$$\mathbf{z} = f_\theta(\mathbf{x}), \quad (4.1)$$

où $\mathbf{z} \in \mathbb{R}^m$ avec $m \ll n$.

Le décodeur applique ensuite une transformation g_ϕ paramétrée par ϕ pour reconstruire une approximation $\hat{\mathbf{x}}$ de l’entrée initiale :

$$\hat{\mathbf{x}} = g_\phi(\mathbf{z}). \quad (4.2)$$

Le modèle est entraîné en minimisant une fonction de coût qui mesure la différence entre l’entrée $\mathbf{x}$ et la reconstruction $\hat{\mathbf{x}}$. La fonction de coût la plus couramment utilisée est l’erreur quadratique moyenne (Mean Squared Error (MSE)) :

$$\mathcal{L}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \sum_{i=1}^n \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2. \quad (4.3)$$

4.2.1.A Variantes des Autoencodeurs

Au fil des années, plusieurs variantes des autoencodeurs ont été développées pour répondre à des besoins spécifiques, chacune intégrant des mécanismes supplémentaires afin d’améliorer l’apprentissage des représentations latentes ou de renforcer la robustesse du modèle. Parmi ces variantes, on trouve :

Autoencodeur Sparse (SAE) : Ce modèle contraint les activations de la couche latente à être rares, ce qui favorise une meilleure compression et extraction des caractéristiques significatives [Ng et al., 2011].

Autoencodeur Denoising (DAE) : Cette variante ajoute du bruit à l’entrée et entraîne le réseau à reconstruire les données originales, augmentant ainsi la robustesse aux perturbations [Vincent et al., 2010].

Autoencodeur Contractif (CAE) : Le Contractive AutoEncoder (CAE) introduit une régularisation basée sur la dérivée de la fonction d’encodage par rapport à l’entrée, minimisant la sensibilité aux petites variations dans les données [Rifai et al., 2011]. La fonction de coût associée est :

$$\mathcal{L}_{CAE} = \mathcal{L}(\mathbf{x}, \hat{\mathbf{x}}) + \lambda \|\nabla_{\mathbf{x}} f_\theta(\mathbf{x})\|^2, \quad (4.4)$$

où λ contrôle l’importance de la régularisation.

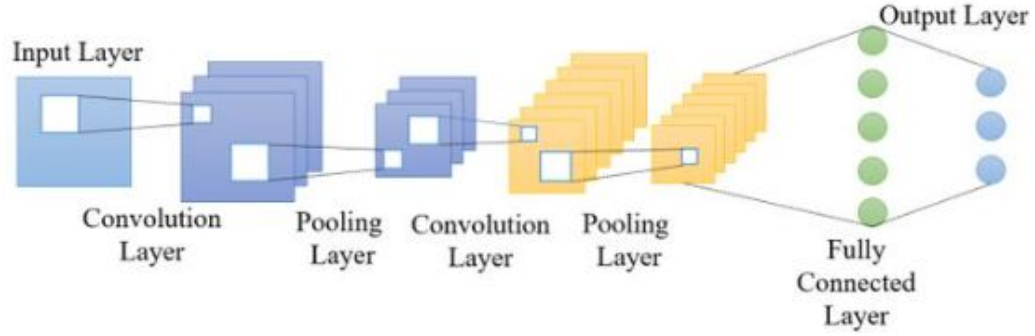


Figure 4.3 – Architecture basic d'un CNN [Gu et al., 2019]

Autoencodeur Variationnel (VAE) : Contrairement aux autoencodeurs classiques qui produisent une représentation latente déterministe $\mathbf{z}$, le Variational AutoEncoder (VAE) [Kingma, 2013] applique des principes bayésiens pour modéliser la distribution des données, générant une distribution probabiliste. La fonction de coût combine une erreur de reconstruction avec une régularisation basée sur la divergence de Kullback-Leibler (KL) :

$$\mathcal{L}_{VAE} = \mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})}[\log p_{\theta}(\mathbf{x}|\mathbf{z})] - \text{KL}[q_{\phi}(\mathbf{z}|\mathbf{x})\|p(\mathbf{z})]. \quad (4.5)$$

Ces variantes enrichissent considérablement les capacités des autoencodeurs, qu'il s'agisse d'apprendre des représentations robustes face au bruit ou de modéliser les données dans un cadre probabiliste, renforçant ainsi leur utilité dans diverses applications.

4.2.2 Réseaux de Neurones Convolutifs (CNNs)

Les CNNs représentent l'une des architectures les plus emblématiques de l'apprentissage profond, particulièrement conçue pour traiter des données avec une topologie en grille, telles que les images. Proposés initialement par LeCun et al. dans les années 1990, les CNNs ont révolutionné le domaine de la vision par ordinateur et trouvé des applications dans divers domaines, tels que le traitement du langage naturel et l'imagerie médicale [LeCun et al., 1998].

L'architecture d'un CNN se compose généralement de trois types de couches principales comme montré dans la figure 4.3 : les couches convolutionnelles, les couches de regroupement (pooling) et les couches entièrement connectées.

- **Couches Convolutionnelles** Ces couches appliquent une série de filtres (ou noyaux) appris sur les données d'entrée pour extraire des caractéristiques locales. Chaque filtre parcourt la matrice d'entrée en réalisant une multiplication élément par élément suivie d'une sommation, décrite mathématiquement par :

$$h_{ij} = \sum_m \sum_n x_{(i+m)(j+n)} w_{mn} + b$$

où x_{ij} représente l'entrée, w_{mn} les poids du filtre, et b le biais. Le résultat est une carte d'activation mettant en évidence des caractéristiques spécifiques, comme des contours ou des textures.

- **Couches de Pooling** Les couches de pooling réduisent les dimensions spatiales des cartes d'activation, rendant le modèle plus efficace et robuste face aux translations mineures dans l'entrée. Le pooling maximal (max-pooling), l'opération la plus courante, sélectionne la valeur maximale dans une fenêtre glissante :

$$p_{ij} = \max\{h_{(i+m)(j+n)} \mid 0 \leq m, n < k\}$$

où k représente la taille de la fenêtre de pooling.

- **Couches Entièrement Connectées** Après les couches convolutionnelles et de pooling, la sortie est aplatie et transmise aux couches entièrement connectées. Ces couches mappent les caractéristiques apprises aux sorties souhaitées, telles que les probabilités des classes. La computation est définie comme :

$$y = \sigma(Wx + b)$$

où W est la matrice des poids, x le vecteur d'entrée, b le vecteur de biais, et σ une fonction d'activation, comme ReLU ou sigmoid.

L'entraînement des CNNs consiste à optimiser les paramètres du réseau pour minimiser une fonction de perte, généralement à l'aide de l'algorithme de rétropropagation et de descente de gradient. Les fonctions de perte fréquemment utilisées incluent la perte par entropie croisée pour les tâches de classification et l'erreur quadratique moyenne (MSE) pour les tâches de régression.

La perte par entropie croisée est définie comme suit :

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(\hat{y}_{ic})$$

où y_{ic} est le label réel et $\hat{y}_{ic}$ la probabilité prédite pour la classe c .

Quant à la MSE, utilisée pour les tâches de régression, elle est définie par :

$$L = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

où y_i est la valeur réelle et $\hat{y}_i$ la valeur prédite.

Les CNNs ont trouvé des applications dans de nombreux domaines. En classification d'images, ils permettent d'attribuer des étiquettes aux images, comme l'a démontré Krizhevsky et al. avec leur modèle AlexNet [Krizhevsky et al., 2012]. En détection d'objets, ils servent à identifier et localiser les objets présents dans une image, comme le montrent des modèles tels que YOLO et Faster R-CNN [Redmon et al., 2016, Ren et al., 2015]. Enfin, dans le domaine de l'imagerie médicale, ils sont utilisés pour détecter des anomalies dans des radiographies ou des IRM, apportant une contribution significative au diagnostic médical [Litjens et al., 2017].

A – Avancées récentes dans les CNNs

- **Réseaux Résiduels (ResNets)** : Proposés par He et al., ces réseaux utilisent des connexions de saut pour résoudre le problème de gradient évanescent, permettant l'entraînement de réseaux très profonds [He et al., 2016].
- **Réseaux Denses (DenseNets)** : Introduits par Huang et al., les DenseNets connectent chaque couche à toutes les couches suivantes, favorisant la réutilisation des caractéristiques et réduisant le nombre de paramètres [Huang et al., 2017].
- **EfficientNets** : Ces modèles optimisent les compromis entre précision et efficacité en ajustant systématiquement la profondeur, la largeur et la résolution [Tan and Le, 2019].
- **Réseaux Siamois (Siamese Networks)** : Les réseaux siamois sont une architecture de CNN conçue pour apprendre des relations entre deux entrées en utilisant une architecture partagée. Cette approche est particulièrement utile pour des tâches de similarité ou de comparaison, telles que la reconnaissance faciale ou la vérification de signatures. Les réseaux siamois utilisent une perte de type contraste pour minimiser la distance entre les paires similaires et maximiser celle entre les paires dissemblables [Chopra et al., 2005].

4.2.3 Recurrent Neural Network (RNN)

Les RNNs sont une classe d'architectures de réseaux de neurones particulièrement adaptés aux données séquentielles, telles que les séries temporelles, le texte ou la parole. Contrairement aux réseaux de neurones classiques, les RNNs possèdent des connexions récurrentes, ce qui leur permet de conserver une mémoire des états passés et de capturer les dépendances temporelles. Cette architecture a été introduite dans les

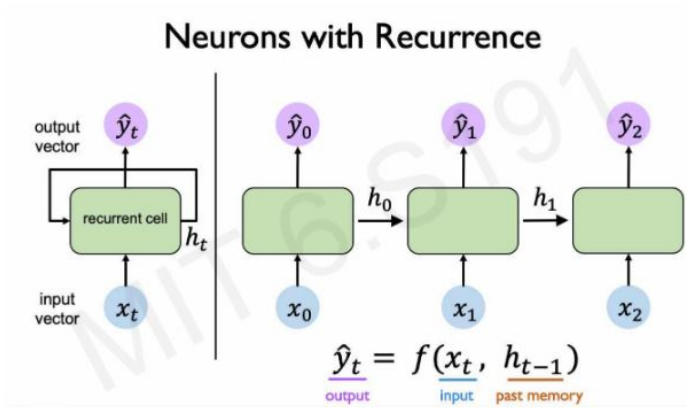


Figure 4.4 – Architecture d'un RNN (source : <https://zemdalk.github.io/2022/11/24/RNN,-Attention,-Transformer,-BERT.html>)

années 80 et est devenue essentielle dans les tâches de modélisation des séquences [Rumelhart et al., 1986].

Les RNNs sont caractérisés par la présence de boucles internes permettant aux informations de circuler d'un pas de temps à l'autre. Un RNN classique peut être vu comme un réseau de neurones feedforward à une seule couche, mais avec une connexion récurrente ajoutée à chaque unité comme illustré dans la figure 4.4. Mathématiquement, un RNN peut être formulé comme suit :

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b)$$

où h_t est l'état caché à l'instant t , h_{t-1} est l'état caché précédent, x_t est l'entrée à l'instant t , W_h et W_x sont des matrices de poids, et b est le biais. La fonction d'activation σ est généralement une fonction non linéaire comme tanh ou ReLU.

L'entraînement des RNNs consiste à minimiser une fonction de perte en ajustant les poids à l'aide de l'algorithme de rétropropagation à travers le temps (Backpropagation Through Time (BPTT)) . Cependant, un problème majeur des RNNs traditionnels est le problème du gradient évanescent, qui empêche les réseaux de capturer des dépendances à long terme dans les séquences.

La Perte par entropie croisée pour les tâches de classification est définie comme suit :

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(\hat{y}_{ic})$$

où y_{ic} est le label réel et $\hat{y}_{ic}$ la probabilité prédite pour la classe c .

Quand à la MSE pour les tâches de régression est définie comme suit :

$$L = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

où y_i est la valeur réelle et $\hat{y}_i$ la valeur prédite.

Les RNNs sont utilisés dans divers domaines où les données sont séquentielles ou temporelles, telles que la traduction automatique, la modélisation de la langue et la génération de texte [Sutskever et al., 2014].

A – Avancées récentes dans les RNNs

- **Long Short-Term Memory (LSTM)** : Proposé par Hochreiter et Schmidhuber, les LSTM sont une extension des RNNs classiques qui résolvent le problème du gradient évanescent en introduisant des portes de contrôle (input, forget, output gates), permettant aux réseaux de mieux capturer les dépendances à long terme [Hochreiter and Schmidhuber, 1997].
- **Gated Recurrent Unit (GRU)** : Les GRUs, introduits par Cho et al., sont une variante des LSTM, mais avec une structure simplifiée, ce qui les rend plus efficaces tout en maintenant des performances similaires dans de nombreuses tâches [Cho et al., 2014].
- **Attention Mechanism** : Le mécanisme d'attention permet aux RNNs d'accorder plus de poids aux parties pertinentes d'une séquence lors de la modélisation des dépendances temporelles, ce qui est particulièrement utile dans des tâches comme la traduction automatique [Bahdanau et al., 2015].
- **Bidirectional Recurrent Neural Network (BIRNN)s** : Ces réseaux traitent les séquences dans les deux directions (avant et arrière) et ont été largement utilisés pour améliorer les performances dans des tâches comme la reconnaissance de la parole ou l'étiquetage des séquences [Schuster and Paliwal, 1997].

4.2.4 Transformers et Vision Transformers (ViT)

Les transformers ont révolutionné le domaine du traitement du langage naturel (NLP) avec des modèles tels que Bidirectional Encoder Representations from Transformers (BERT), Generative Pre-trained Transformer (GPT), et T5. Introduits par [Vaswani et al., 2017], les transformers sont des modèles d'architecture d'attention capables de capturer des dépendances à longue portée dans des séquences, tout en permettant un parallélisme plus efficace par rapport aux RNNs et LSTMs. Leurs caractéristiques les ont rendus populaires pour une grande variété de tâches dans le NLP (Natural Language Processing) et au-delà, y compris dans la vision par ordinateur avec l'architecture Vision Transformer (ViT) [Vaswani et al., 2017, Dosovitskiy et al., 2020].

Selon la figure 4.5 L'architecture des transformers repose sur une composante clé : le *mécanisme d'auto attention*, qui permet au modèle de se concentrer sur différentes parties de la séquence d'entrée à chaque étape. Un transformer classique est constitué de plusieurs couches identiques, chacune comprenant

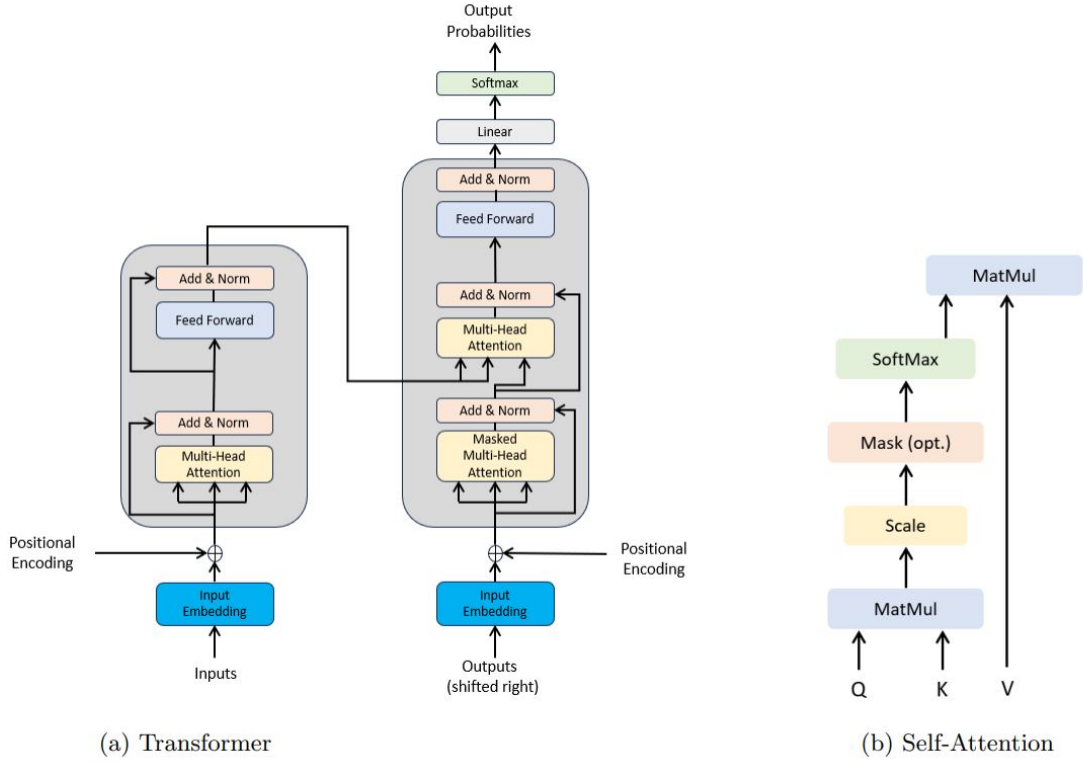


Figure 4.5 – Architecture des transformers et son mécanisme d'auto attention [Vaswani et al., 2017]

deux sous-couches principales : l'attention multi-tête et une couche de réseau de neurones feedforward positionnelle.

Le mécanisme d'auto attention permet de calculer des poids pour chaque élément de la séquence d'entrée en fonction de sa pertinence pour un élément cible à chaque étape du traitement. L'attention est formulée comme suit :

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

où Q , K , et V sont les matrices de requêtes, de clés et de valeurs, respectivement, et d_k est la dimension des clés. L'attention multi-tête consiste à appliquer cette opération en parallèle sur plusieurs « têtes » d'attention, permettant au modèle d'extraire différentes relations à chaque étape.

Chaque bloc Transformer comprend deux sous-couches principales : l'attention multi-tête et un réseau de neurones feedforward positionnel. Les entrées de ces sous-couches sont normalisées par la technique de normalisation de couche et passées par un mécanisme de connexion résiduelle pour stabiliser l'entraînement.

$$\text{Output}_{\text{Block}} = \text{LayerNorm}(x + \text{Sublayer}_1(x)) \quad \text{et} \quad \text{Output}_{\text{Block}} = \text{LayerNorm}(x + \text{Sublayer}_2(x))$$

L'entraînement des transformers utilise principalement la rétropropagation, optimisant les poids avec des méthodes comme Adam et en utilisant des fonctions de perte adaptées à chaque tâche (par exemple, l'entropie croisée pour les tâches de classification).

La Perte pour la classification est définie comme suit :

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(\hat{y}_{ic})$$

où y_{ic} est le label réel et $\hat{y}_{ic}$ la probabilité prédite pour la classe c .

Lors de l'entraînement des modèles tels que GPT, des techniques de masquage sont utilisées pour empêcher le modèle de voir des informations futures dans la séquence (c'est-à-dire les tokens suivants).

Les transformers ont trouvé des applications dans divers domaines :

- **Traitement du langage naturel (NLP)** : Traduction automatique, génération de texte, résumé automatique, classification de texte, etc. [Vaswani et al., 2017].
- **Vision par ordinateur** : Les transformers ont été adaptés pour des tâches de vision par ordinateur, avec des architectures telles que le Vision Transformer (ViT) [Dosovitskiy et al., 2020].
- **Modélisation multimodale** : Les transformers sont également utilisés dans des modèles multimodaux, comme CLIP, qui combinent images et texte pour des tâches telles que la recherche d'images par texte.

4.2.4.A Vision Transformer (ViT)

Le Vision Transformer (ViT) est une architecture inspirée des transformers, adaptée pour des tâches de vision par ordinateur. Contrairement aux CNNs traditionnels, ViT découpe l'image en patches non chevauchants, qui sont ensuite traités comme des séquences de tokens dans un modèle transformer.

Dans ViT, l'image est d'abord découpée en patches, chaque patch étant aplati en un vecteur et traité par un transformer. L'architecture ViT est fondée sur la même idée que les transformers traditionnels, mais appliquée aux données visuelles.

$$X_{\text{patch}} = \text{Flatten}(\text{Image})$$

Chaque patch est ensuite traité de manière similaire aux tokens dans les modèles NLP, et l'output est utilisé pour des tâches comme la classification d'images.

Avancées récentes dans les Transformers et ViT :

- **DeiT (Data-efficient Image Transformer)** : Proposé pour rendre ViT plus efficace avec moins de données, l’approche DeiT améliore l’efficacité d’entraînement en ajoutant un mécanisme de distillation des connaissances [Touvron et al., 2021].
- **Swin Transformer** : Ce modèle hiérarchique améliore l’efficacité des transformers en utilisant des fenêtres glissantes pour limiter la portée de l’attention, permettant de traiter des images de plus grande taille [Liu et al., 2021].
- **Vision Transformer avec Hybrid CNN (Hybrid ViT)** : Une approche hybride qui combine les avantages des CNNs pour extraire des caractéristiques locales et des transformers pour modéliser les relations globales dans les images [Chu et al., 2021].

4.2.5 Graph Neural Network (GNN)

GNNs sont une classe de modèles d’apprentissage profond conçus pour traiter des données sous forme de graphes. Contrairement aux données structurées en grille, comme les images ou les textes, les données de graphe sont souvent non structurées et ont des relations complexes entre les nœuds (ou sommets) et les arêtes (ou liens). Les GNNs ont été introduits pour la première fois par Scarselli et al. en 2009 [Scarselli et al., 2009], et ont depuis été largement utilisés pour des tâches telles que la classification de nœuds, la prédiction de liens et la classification de graphes, en exploitant la topologie et les informations des voisins des nœuds.

Les GNNs sont basés sur l’idée d’apprendre des représentations des nœuds en tenant compte de leurs voisins dans le graphe. La figure 4.6 illustre l’architecture typique d’un GNN qui consiste en plusieurs couches qui mettent à jour les représentations des nœuds en fonction de leurs voisins, ce qui permet au modèle de capturer des dépendances locales et globales dans le graphe.

La propagation de messages est le mécanisme fondamental des GNNs. À chaque couche, un nœud met à jour sa représentation en combinant sa propre représentation avec celles de ses voisins. Ce processus est généralement formulé comme suit :

$$h_v^{(k)} = \sigma \left(W^{(k)} \cdot \left(h_v^{(k-1)} \oplus \sum_{u \in \mathcal{N}(v)} h_u^{(k-1)} \right) \right)$$

où $h_v^{(k)}$ est la représentation du nœud v à la couche k , $\mathcal{N}(v)$ représente les voisins de v , et $\oplus$ est une opération de combinaison (comme la somme ou la concaténation). $W^{(k)}$ est une matrice de poids et σ est une fonction d’activation, comme ReLU.

Il existe plusieurs variantes des GNNs, y compris les Graph Convolutional Network (GCN), Graph Attention Network (GAT), et Message Passing Neural Network (MPNN), qui diffèrent par la manière dont ils combinent les informations des voisins.

- **GCN** : Utilise des convolutions sur le graphe pour propager l’information entre les nœuds voisins.

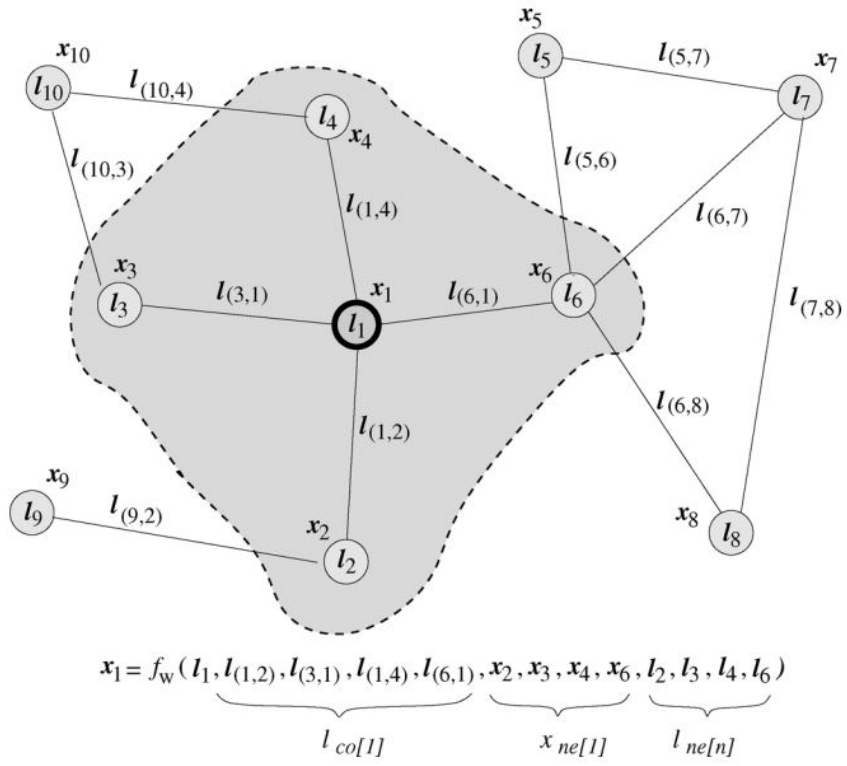


Figure 4.6 – Graphe et le voisinage d'un nœud. L'état x_1 du nœud 1 dépend de l'information contenue dans son voisinage. [Scarselli et al., 2009]

- **GAT** : Utilise un mécanisme d’attention pour pondérer l’importance des voisins lors de la mise à jour des représentations des nœuds.
- **MPNN** : Utilise un cadre général où l’information est propagée sous forme de messages entre les nœuds du graphe.

L’entraînement des GNNs repose sur la minimisation d’une fonction de perte, généralement en utilisant des méthodes de rétropropagation et d’optimisation par descente de gradient. Comme pour d’autres réseaux de neurones, des techniques de régularisation et d’optimisation peuvent être appliquées pour éviter le surapprentissage et améliorer la convergence.

Pour la classification de nœuds, une fonction de perte courante est la perte d’entropie croisée :

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(\hat{y}_{ic})$$

où y_{ic} est le label réel du nœud i pour la classe c , et $\hat{y}_{ic}$ la probabilité prédite.

L’optimisation des paramètres des GNNs se fait généralement à l’aide de méthodes comme Adam, tandis que des techniques de régularisation telles que le dropout ou la régularisation de poids peuvent être utilisées pour éviter le surapprentissage.

Les GNNs ont été appliqués dans divers domaines, en particulier ceux impliquant des données structurées sous forme de graphes :

- **Classification de nœuds** : Attribution de labels aux nœuds d’un graphe, comme dans le cas des réseaux sociaux pour prédire des caractéristiques des utilisateurs [Kipf and Welling, 2017].
- **Prédiction de liens** : Prédire les liens manquants entre les nœuds d’un graphe, une tâche courante dans les systèmes de recommandation ou les graphes de connaissances [He et al., 2020].
- **Classification de graphes** : Attribution de labels à un graphe entier, comme pour la classification des molécules ou la prédiction de propriétés de matériaux à partir de leurs graphes moléculaires [Ying et al., 2018].

Les GNNs ont connu plusieurs améliorations récentes pour améliorer leur efficacité et leur capacité à traiter des graphes à grande échelle :

- **GraphSAGE (Graph Sample and Aggregation)** : Une approche permettant d’échantillonner un sous-ensemble de voisins pour chaque nœud, ce qui réduit la complexité des calculs pour de grands graphes [Hamilton et al., 2017].
- **GAT (Graph Attention Networks)** : Utilise des mécanismes d’attention pour permettre à chaque nœud de pondérer l’importance de ses voisins, offrant ainsi plus de flexibilité dans l’agrégation des informations [Veličković et al., 2018].

- **GIN (Graph Isomorphism Networks)** : Propose un modèle qui améliore la capacité de discrimination des graphes, ce qui est utile pour des tâches telles que la classification de graphes [Xu et al., 2019].
- **GCN pour les grands graphes** : Des techniques comme le pré-traitement ou l'échantillonnage de graphes sont appliquées pour rendre les GCNs plus efficaces lorsqu'il s'agit de traiter des graphes de très grande taille [Chiang et al., 2019].

4.3 Conclusion

Ce chapitre a présenté une vue d'ensemble des architectures de réseaux de neurones les plus influentes dans le domaine de l'apprentissage profond : **AE**, **CNNs**, **RNNs**, **Transformers**, **Vision Transformers (ViTs)** et **GNNs**. Chaque type de réseau est conçu pour traiter des données spécifiques et excelle dans des domaines particuliers. Les **CNNs** sont largement utilisés pour la vision par ordinateur, grâce à leur capacité à extraire des caractéristiques locales à partir des images. Les **RNNs**, quant à eux, sont adaptés aux données séquentielles, comme celles rencontrées dans le traitement du langage naturel et les séries temporelles. Les **Transformers** ont révolutionné le traitement du langage avec leur mécanisme d'attention, surpassant les RNNs dans de nombreuses tâches. Les **ViTs**, une extension des Transformers, ont également ouvert de nouvelles possibilités pour le traitement d'images en les traitant comme des séquences de patches. Les **GNNs** se distinguent par leur capacité à traiter des données structurées sous forme de graphes, ce qui les rend particulièrement puissants pour des applications comme la modélisation de réseaux sociaux ou la prédiction de liens dans des systèmes complexes. Les **AE**, de leur côté, sont des modèles non supervisés utilisés pour la réduction de dimensionnalité, la détection d'anomalies, ainsi que l'apprentissage de représentations compactes des données.

Malgré leurs performances exceptionnelles, ces architectures rencontrent des défis tels que la complexité computationnelle, la nécessité de grandes quantités de données étiquetées et l'interprétabilité des modèles. Cependant, elles ont ouvert la voie à des progrès significatifs dans des domaines variés, allant de l'imagerie médicale à la compréhension du langage naturel et à l'analyse de réseaux complexes. Les recherches actuelles continuent de se concentrer sur l'amélioration de l'efficacité, de l'accessibilité et de l'interprétabilité de ces modèles, promettant des avancées majeures dans l'optimisation de l'apprentissage profond.

5

Impact de la localisation des Région d'Intérêts (ROI) sur la détection précoce de la gonarthrose

Sommaire

5.1	Introduction et Motivation	67
5.2	Segmentation des radiographies et extraction des ROIs	68
5.3	Données et paramètres expérimentaux	73
5.4	Résultats de la classification et discussion	74
5.5	Forces et limites de l'étude	76
5.6	Conclusion	76

5.1 Introduction et Motivation

Les techniques de DL ont montrés un potentiel prometteur dans la classification de la gonarthrose (Cf. 3).

Malgré ces avancées, détecter précocement l’arthrose du genou sur des radiographies classiques reste difficile. La similitude entre les images d’arthrose et celles sans arthrose au stade initial, combinée à l’architecture CNN qui ne prend pas en compte les détails de texture, complique cette tâche. Différentes études ont montré que les CNN traditionnels ont du mal à distinguer les caractéristiques subtiles de texture, essentielles pour une détection précoce de l’arthrose du genou [Cimpoi et al., 2015, Andrearczyk and Whelan, 2016, Zeiler and Fergus, 2014, Springenberg et al., 2014].

Dans cette étude, nous proposons une nouvelle approche basée sur des techniques d’apprentissage profond pour la détection précoce de la gonarthrose en se concentrant sur l’impact des ROI de l’os tibial. La caractérisation de la microarchitecture de l’os tibial peut fournir des informations précieuses sur la gravité et la progression de la gonarthrose [Arden and Nevitt, 2006]. Notre méthode proposée implique une série d’étapes séquentielles, exploitant des techniques avancées de traitement d’images et des algorithmes d’apprentissage automatique pour extraire des caractéristiques informatives de l’os tibial. Tout d’abord, en utilisant des radiographies standard, l’articulation du genou est segmentée [Stachowiak et al., 2016, Wang et al., 2021]. En localisant précisément l’articulation du genou, nous assurons l’extraction ultérieure des ROI en utilisant l’outil BoneFinder [Lindner et al., 2013]. Nous utilisons ensuite un algorithme de détection de sous-contours pour extraire différentes ROIs de l’os tibial. Ces sous-régions fournissent un niveau d’analyse plus fin et permettent de capturer des changements localisés dans la microarchitecture de l’os tibial, dus à la gonarthrose. Ensuite, nous sélectionnons sept ROIs sur chaque sous-contour, pour étudier leur pertinence pour la détection de la gonarthrose.

Le Stacked Sparse AutoEncoder (SSAE) est reconnu pour sa capacité à capturer des motifs et des structures complexes dans les images médicales [Shin et al., 2012]. C’est dans cette perspective que nous l’avons employé afin d’extraire les caractéristiques multidimensionnelles des ROIs sélectionnées [Ng et al., 2011]. Les caractéristiques extraites sont ensuite utilisées pour entraîner un Support Vector Machine with Radial Basis Function kernel (SVM-RBF) [Brahim et al., 2019].

En tirant parti des caractéristiques extraites, le classificateur SVM-RBF peut distinguer efficacement entre les différents stades de la gonarthrose, fournissant des informations précieuses pour le diagnostic précoce et la planification thérapeutique. Pour garantir une prise de décision robuste, nous utilisons la technique de vote maximum pour obtenir une décision unique pour chaque sous-contour. Le vote maximum combine les décisions de multiples prédictions des différentes ROIs de l’ensemble de données, favorisant des résultats de classification plus fiables et cohérents.

La principale contribution de notre étude est d’analyser l’influence de la sélection des ROI sur la classification précise et robuste de la gonarthrose, en utilisant des méthodes d’apprentissage profond.

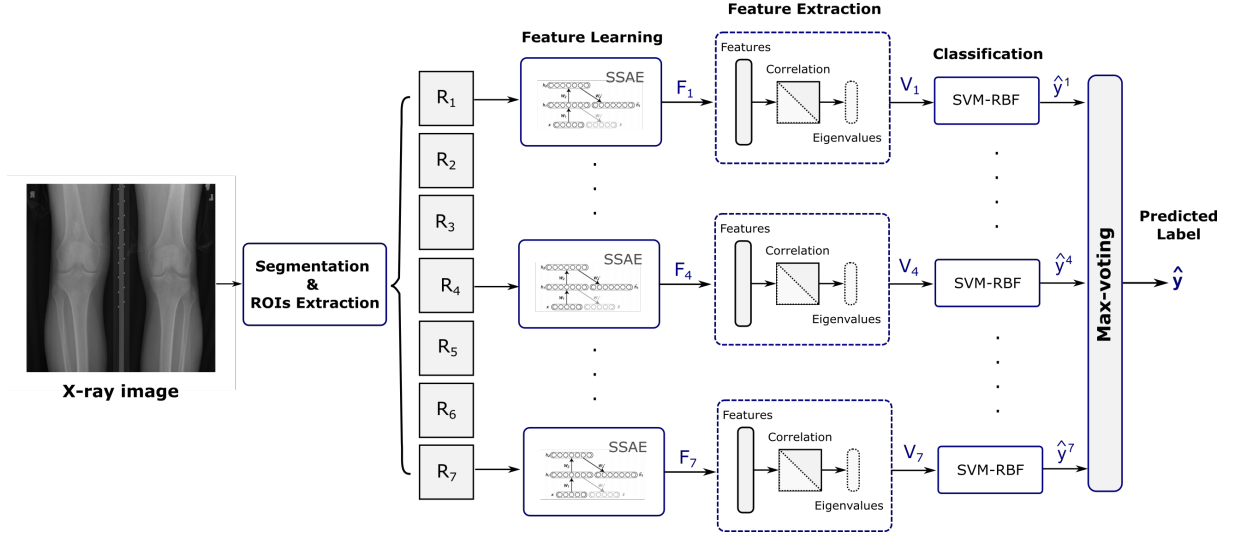


Figure 5.1 – Pipeline de l'approche proposée.

5.2 Segmentation des radiographies et extraction des ROIs

Pour classifier une image radiographique à partir des données brutes fournies par OAI, notre méthode proposée effectue une série d'étapes successives résumées dans la Figure 5.1. Pour localiser automatiquement la région de l'articulation du genou à partir d'une image radiographique standard, nous avons tout d'abord utilisé l'outil BoneFinder [Lindner et al., 2013]. BoneFinder est un outil entièrement automatique qui combine l'apprentissage automatique et l'analyse de forme pour identifier grossièrement le fémur et le tibia. Plus précisément, nous avons employé BoneFinder pour délimiter les contours des os fémoraux et tibiaux (lignes en pointillés jaunes sur la Figure 5.3.B, permettant ainsi la localisation de sous-contours à l'intérieur de l'os tibial pour extraire les ROIs nécessaires à nos expériences, comme illustré sur la Figure 5.4.

5.2.1 Bonefinder pour la segmentation automatique du plateau tibio-fémoral.

BoneFinder® représente un logiciel entièrement automatisé dédié à la délimitation et à la segmentation des structures squelettiques à partir de radiographies 2D, en plaçant un ensemble de points le long des contours osseux ou aux positions clés des repères anatomiques. Initialement développé pour la segmentation de l'articulation de la hanche, cet outil est désormais également utilisé dans d'autres contextes, tels que les articulations de la main ou l'articulation du genou. Disponible sur <https://bone-finder.com>. BoneFinder utilise une approche de type Hough Forest pour détecter la structure d'intérêt dans l'image, puis Random Forest Regression Voting dans le cadre du modèle local contraint pour affiner localement toutes les positions de points. [Lindner et al., 2013].

L'extraction des ROIs à partir des images segmentées se compose de :

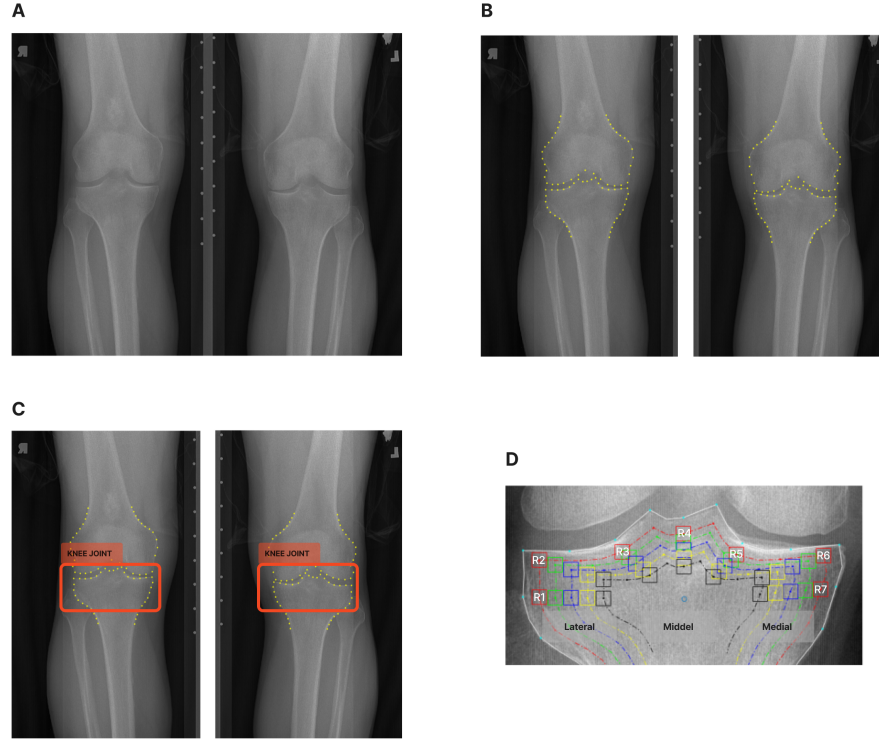


Figure 5.2 – Image radiographique des deux genoux (A), Délimitation des contours des os fémoral et tibial à l’aide de l’outil BoneFinder (B), localisation de l’articulation du genou en rouge (C), détection de cinq contours mis en évidence respectivement en rouge, vert, bleu, jaune et noir, et extraction de 7 ROIs sur chaque contour (D).

— Détection des contours : Cinq sous-contours ont été extraits en utilisant une homothétie avec différents rapports. En se référant au contour de l’os tibial comme étant le plus précis, un ensemble de points appelés $PTbf$ a été sélectionné. Ensuite, les coordonnées de ces points ont été calculées comme suit :

$$X_b = \frac{(\sum |absc| \times \sum Int)}{\sum Int}; Y_b = \frac{(\sum |ord| \times \sum Int)}{\sum Int} \quad (5.1)$$

où $absc$ et ord représentent respectivement les coordonnées x et y. Int désigne l’intensité des pixels. Pour calculer les coordonnées de chaque contour, nous avons utilisé différents rapports, $Rh = 0,8$ à $0,4$ avec un pas de $-0,1$. Par conséquent, 5 sous-contours ont été identifiés sur chaque image du genou. Les nouvelles coordonnées, $absc'$ et ord' du contour suivant ont été calculées comme suit :

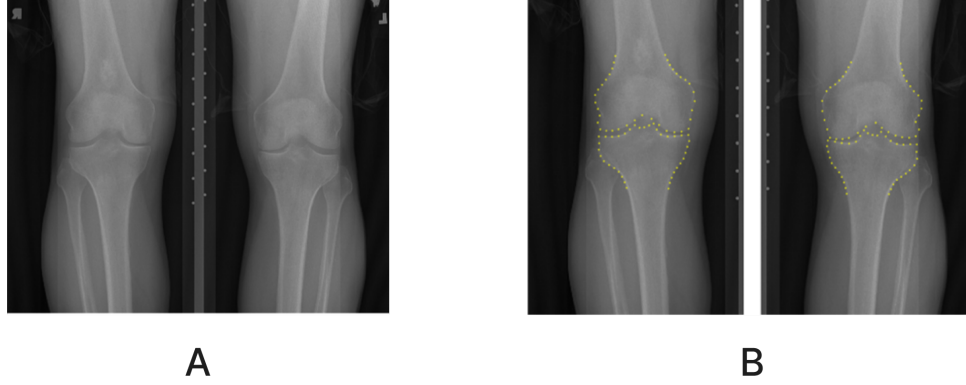


Figure 5.3 – Image radiographique des deux genoux (A), délimitation des contours des os fémoral et tibial à l’aide de l’outil BoneFinder (B).

$$\begin{aligned}
 absc' &= X_b \times (1 - Rh - px) + PTbf1_x \\
 ord' &= Y_b \times (1 - Rh) + PTbf1_y
 \end{aligned}
 \tag{5.2}$$

où,

$$\begin{aligned}
 PTbf1_x &= (Rh + px) \times PTbf_x \\
 PTbf1_y &= Rh \times PTbf_y
 \end{aligned}
 \tag{5.3}$$

Rh est le rapport d’homothétie, px est une constante égale à 0.1. $PTbf_x$ et $PTbf_y$ représentent respectivement les valeurs des coordonnées x et y des points du contour précédent.

— *Sélection des ROIs* : Parmi les 37 points de chaque contour obtenus à l’aide de BoneFinder, nous avons sélectionné 7 points spécifiques (9, 11, 15, 19, 23, 27 et 29) comme centre de chaque ROI extraite. En utilisant les 7 points sélectionnés comme centres, nous avons extrait 7 ROIs de chaque sous-contour. Un sous-contour fait référence à un segment ou une partie spécifique du contour principal. Avec 7 ROIs extraites de chacun des 5 sous-contours, nous disposons d’un total de 35 ROIs par image. Ce procédé garantit une méthode standardisée et reproductible pour analyser des régions spécifiques au sein des structures anatomiques à travers différentes images. Chaque ROI fournit des informations précieuses et différentes qui contribuent à l’analyse globale de la structure osseuse.

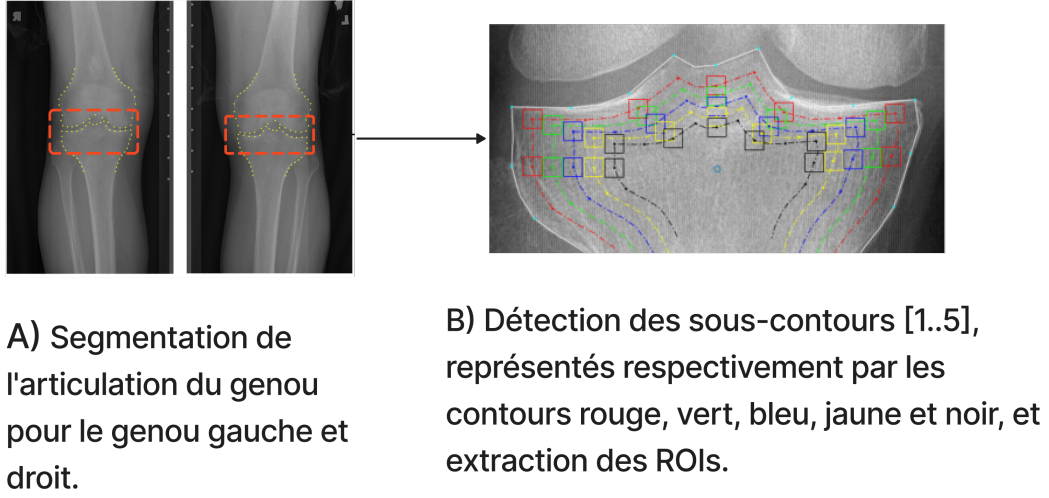


Figure 5.4 – Image radiographique des deux genoux (A), délimitation des contours des os fémoral et tibial à l'aide de l'outil BoneFinder (B).

5.2.2 Apprentissage automatique des caractéristiques à l'aide du modèle SSAE

Pour l'extraction des caractéristiques, nous avons utilisé le modèle SSAE, qui comprend plusieurs couches empilées d'auto-encodeurs épars Sparse AutoEncoder (SAE), où chaque couche apprend de manière progressive des représentations de plus en plus abstraites des données d'entrée. Chaque SAE se compose d'un réseau encodeur qui mappe l'entrée vers un espace latent de dimension inférieure et d'un réseau décodeur qui reconstruit l'entrée à partir de la représentation latente. Plus précisément, le modèle SSAE implique un apprentissage hiérarchique de caractéristiques de plus en plus complexes et significatives à partir des données d'entrée. Cela est réalisé en imposant des contraintes de parcimonie lors de l'entraînement, ce qui encourage l'activation d'une petite fraction de neurones dans l'encodeur. Le réseau décodeur vise ensuite à reconstruire l'entrée d'origine à partir de ces représentations éparses. La fonction de coût, $\mathcal{L}_{SAE}$, à optimiser lors de l'entraînement de chaque SAE est définie comme une combinaison de la perte de reconstruction, de la régularisation de parcimonie et d'un terme de régularisation des poids :

$$\mathcal{L}_{SAE} = \mathcal{L}_{MSE} + \beta \Omega_{sparsity} + \lambda \Omega_{weights} \quad (5.4)$$

où, β and λ représentent respectivement les poids des termes de régularisation de parcimonie L^2 . Le premier terme, $\mathcal{L}_{MSE}$ représente la perte de reconstruction. Il mesure la disparité entre l'entrée originale

$x_k \in X$ et la sortie reconstruite $\hat{x}_k$ en utilisant la perte de MSE. $\mathcal{L}_{MSE}$ est défini comme suit :

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K (x_{kn} - \hat{x}_{kn})^2 \quad (5.5)$$

où N est le nombre d'échantillons d'entraînement et K le nombre d'unités cachées, $\Omega_{sparsity}$ est un terme de régularisation qui encourage la parcimonie dans l'activation des neurones de l'encodeur en utilisant la divergence de Kullback-Leibler [Nasser et al., 2020]. Le troisième terme, $\Omega_{weights} = \|W\|_2^2$, correspond à la régularisation L^2 qui maintient la matrice des poids W petite, afin d'éviter le sur-apprentissage.

5.2.3 Extraction des caractéristiques

Pour tenir compte des informations spatiales liées aux caractéristiques de texture des ROIs extraites de chaque image, nous nous sommes appuyés sur les corrélations des caractéristiques. Plus précisément, si x_i représente le vecteur d'entrée d'intensité d'un ROI, R_i , nous avons tout d'abord calculé le vecteur de caractéristiques apprises correspondant, $F_i \in \mathbb{R}^K$ en utilisant notre modèle SSAE entraîné. Ensuite, la matrice de corrélation, $C \in \mathbb{R}^{K \times K}$ a été calculée comme suit :

$$C = F \times F^T \quad (5.6)$$

où K représente la taille de la dernière couche cachée dans le SSAE.

Puisque la matrice de corrélation est symétrique, nous avons sélectionné uniquement la partie triangulaire inférieure C_l de la matrice C . Ensuite, afin de réduire la dimensionnalité du vecteur de caractéristiques final, nous avons calculé les valeurs propres de la matrice C_l . Ce dernier vecteur, noté V_i , a ensuite été utilisé comme entrée pour le modèle de classification.

5.2.4 Classification de la sévérité de la gonarthrose

Pour la classification de la gonarthrose, nous avons utilisé les vecteurs de caractéristiques, V_i extraits précédemment. Ces vecteurs, qui capturent des informations spatiales et de texture des ROIs, ont été utilisés comme entrée pour notre modèle de classification. Le modèle de classification a été formé pour distinguer entre les images de genoux atteints d'arthrose et celles des genoux sains, en se basant sur les caractéristiques apprises et réduites en dimensionnalité.

Pour la tâche de classification, nous avons utilisé les séparateurs à vastes marges (SVM : Support Vector Machine) avec noyau de fonction de base radiale SVM-RBF. Pour ce faire, nous avons utilisé les vecteurs de caractéristiques extraits $\{V_j | j = 1, \dots, 7\}$ comme entrées des modèles SVM-RBF, qui cherchent à trouver un hyperplan optimal dans l'espace des caractéristiques séparant les différentes classes.

ROI Contour	R1	R2	R3	R4	R5	R6	R7
1	81.79 ± 0.20	80.71 ± 0.84	83.68 ± 1.48	84.81 ± 1.67	84.00 ± 1.38	82.86 ± 0.24	83.25 ± 1.85
2	83.91 ± 1.23	83.64 ± 0.64	83.91 ± 1.23	80.98 ± 0.21	80.92 ± 0.40	80.92 ± 0.40	83.64 ± 0.64
3	82.59 ± 1.10	80.02 ± 0.40	85.68 ± 1.33	82.59 ± 1.10	83.91 ± 0.72	80.92 ± 0.40	80.92 ± 0.40
4	82.53 ± 1.28	82.14 ± 0.63	81.76 ± 0.59	81.34 ± 0.50	84.39 ± 0.20	81.34 ± 0.45	83.58 ± 0.38
5	81.49 ± 0.92	83.11 ± 2.29	83.98 ± 2.39	83.44 ± 2.20	80.32 ± 0.55	83.34 ± 1.98	80.59 ± 0.92

Table 5.1 – Précisions de classification obtenues (%) avec chaque ROI à partir des différents contours.

Max-voting Contour	3 best ROIs	5 best ROIs	All ROIs
1	85.35 ± 1.85	84.42 ± 1.00	84.99 ± 1.10
2	83.91 ± 1.23	83.91 ± 1.23	86.20 ± 1.41
3	83.40 ± 0.45	83.34 ± 1.57	80.83 ± 0.68
4	85.29 ± 1.10	83.85 ± 0.91	83.37 ± 0.54
5	83.98 ± 2.93	83.44 ± 2.20	83.19 ± 2.04

Table 5.2 – Précisions obtenues (%) en utilisant la technique Max-voting.

Pour combiner les prédictions de toutes les ROIs de chaque sous-contour, nous avons employé la technique de vote majoritaire [Wang et al., 2024] (Max-Voting). Ce processus d’agrégation nous a permis d’obtenir un résultat unique et consolidé comme décision collective. Le vote majoritaire pour les 7 ROIs a été appliqué comme suit :

$$\hat{y}_i = \text{mode}\{\hat{y}_i^{(1)}, \hat{y}_i^{(2)}, \hat{y}_i^{(3)}, \hat{y}_i^{(4)}, \hat{y}_i^{(5)}, \hat{y}_i^{(6)}, \hat{y}_i^{(7)}\} \quad (5.7)$$

où $\hat{y}_i^{(j)}$ est le label prédit du ROI R_j extrait de la $i^{\text{ème}}$ image, mode représente la valeur la plus fréquente. Le score de vote majoritaire a été calculé en utilisant les étiquettes obtenues $\hat{y}_i$ et les étiquettes réelles y_i . Nous avons également sélectionné les trois régions qui ont donné les meilleurs résultats de la Table 5.2, et nous avons appliqué le vote majoritaire sur ces trois régions. Ensuite, nous avons répété le processus pour 5 régions.

5.3 Données et paramètres expérimentaux

Dans cette section, nous présentons les données utilisées dans cette étude et détaillons les paramètres de configuration de nos expériences.

5.3.1 Ensemble de données

La base de données OAI inclut des radiographies du genou dans le cadre de son processus complet de collecte de données. Les radiographies du genou sont couramment utilisées dans la recherche sur la gonarthrose pour évaluer les changements structuraux et la sévérité de la dégénérescence de l'articulation du genou. Nous avons utilisé un ensemble de données équilibré, correspondant aux données de base et comprenant 3347 images radiographiques, dont 1687 images de sujets en bonne santé (KL-0) et 1687 images de patients atteints de gonarthrose (KL-2). Comme nous nous intéressons à la détection précoce de la gonarthrose, seules les catégories KL-0 et KL-2 ont été conservées. Les catégories KL-1 (douteuse) et KL-3 et KL-4 (arthrose sévère) n'ont pas été prises en compte. De plus, seules les images radiographiques obtenues avec la modalité de radiographie calculée spécifique CR ont été conservées. La modalité CR fait référence à des images numériques capturées à l'aide de plaques sensibles aux rayons X et traitées ultérieurement par un dispositif dédié. La modalité CR remplace l'approche traditionnelle basée sur des films, car les images sont directement obtenues sous forme numérique. L'utilisation d'images CR offre plusieurs avantages, notamment l'accessibilité immédiate, le stockage pratique et la capacité à améliorer et manipuler les données numériques à des fins d'analyse et d'interprétation avancées.

5.3.2 Application de l'approche

Pour évaluer notre approche, nous avons réalisé une cross-validation à 10 plis sur les 3347 images. Cela signifie que nous avons divisé les données de manière aléatoire en 10 sous-ensembles égaux ou "plis" et utilisé l'un d'eux comme ensemble de validation pour tester les performances du modèle, tandis que les neuf autres ont été utilisés pour l'entraînement. De cette manière, nous nous sommes assurés que chaque instance était validée une fois.

Nous avons réglé les paramètres du SSAE en variant le nombre de couches cachées, le nombre de neurones dans chaque couche, et les valeurs des poids des termes de régularisation de parcimonie L^2 . Nous avons obtenu les paramètres optimaux après plusieurs tests approfondis.

Nous avons utilisé un modèle SSAE avec trois couches cachées contenant respectivement 256, 128 et 32 unités cachées. La proportion de parcimonie et les valeurs des poids de la parcimonie β et des termes de régularisation L^2 , λ ont été fixées à 0.01, 2 et $1e - 4$, respectivement.

5.4 Résultats de la classification et discussion

Dans cette section, nous évaluons les performances du modèle proposé pour la prédiction précoce de l'arthrose du genou (KL-0 vs KL-2). L'objectif principal est d'étudier les effets de l'emplacement des ROIs extraites sur les résultats de classification. Le tableau 5.1 présente la moyenne et l'écart-type des scores de précision obtenus par le modèle proposé pour les différentes ROIs extraites des différents sous-contours.

Les scores de précision ont été estimés en comparant les grades KL prédits avec les grades KL réels de chaque image du genou. Le tableau 5.1 montre que la méthode proposée a obtenu le meilleur score de précision de 85.68% pour la ROI R3 du sous-contour 3, qui correspond à la partie médiane du tibia. Le score le plus bas de 80.02% a été obtenu pour la ROI R2 du sous-contour 5.

En analysant chaque contour séparément, la première ligne du tableau 5.1 montre les résultats obtenus pour le sous-contour 1, qui est le plus proche du plateau tibial. Les scores de précision pour le sous-contour 1 varient de 80.71% à 84.81%, avec le score le plus élevé pour la ROI R4 et le plus bas pour la ROI R2. Cela indique que la partie médiane est la région la plus distinctive de ce sous-contour.

La deuxième ligne montre les résultats pour le sous-contour 2, qui est légèrement plus éloigné du plateau tibial. Le score de précision le plus élevé est obtenu avec les ROIs R1, R2, R3 et R7, qui sont situées respectivement dans les parties latérales et médiales du tibia. Les scores les plus bas d'environ 81% sont obtenus avec les ROIs R4, R5 et R6, qui correspondent respectivement aux parties médianes et médianes-médiales du tibia. Cela indique que la partie latérale du tibia est plus informative sur l'arthrose que les parties médiane et médianes-médiales du genou.

La troisième ligne montre les résultats pour le sous-contour 3, qui est encore plus éloigné du plateau tibial. Les scores de précision les plus élevés sont obtenus par les ROIs R3 et R5, qui correspondent respectivement aux parties médianes-médiales et médianes-latérales. Les scores les plus bas sont obtenus par les ROIs R2, R6 et R7, qui correspondent aux parties tibiales médiale et latérale, respectivement. Cela indique que la région entre la région médiane et les deux extrémités du genou (médiale et latérale) est plus informative sur l'arthrose que les autres parties de ce sous-contour.

La quatrième ligne montre les résultats pour le sous-contour 4. Le meilleur score de précision de 84.39% est atteint par la ROI R5, tandis que le score le plus bas de 81.34% est atteint par les ROIs R4 et R6.

La cinquième ligne montre les résultats pour le sous-contour 5, qui est le plus éloigné du plateau tibial. Le meilleur score de précision de 83.98% est atteint par la ROI R3, qui correspond à la partie médiane. Les scores les plus bas sont obtenus par les ROIs R1, R5 et R7.

En résumé, le tableau 5.1 démontre que les différentes ROIs extraites ont des impacts différents sur la précision de classification de l'arthrose du genou. Il montre également que certaines régions sont plus informatives que d'autres en fonction de leur emplacement et de leur distance par rapport au plateau tibial.

Afin d'obtenir une décision unique par image, nous avons appliqué la technique de Max-voting (Eq. 5.7). Selon le tableau 5.2, sauf pour le contour 3, on peut observer que les résultats ont légèrement été améliorés grâce à l'application de la technique de Max-voting. L'amélioration la plus significative est pour les 7 ROIs du sous-contour 2, qui ont atteint une précision de 86.20%.

5.5 Forces et limites de l'étude

L'utilisation de BoneFinder pour la segmentation des images radiographiques permet de délimiter précisément n'importe quelle partie du genou. Cette méthode permet également de séparer automatiquement et de manière fiable le genou gauche du genou droit, assurant une segmentation de bout en bout sans intervention manuelle. En utilisant SSAE, nous avons pu extraire des représentations caractéristiques significatives des ROIs. Cette méthode permet de capturer des informations texturales complexes et abstraites, améliorant ainsi la capacité du modèle à différencier les images de genoux sains (KL-0) et arthrosiques (KL-2).

En analysant l'impact de l'emplacement des ROIs sur les résultats de classification, nous avons identifié les régions les plus informatives. Par exemple, les résultats montrent que certaines zones, comme la partie médiane du tibia, fournissent des informations plus distinctives pour la classification de l'arthrose.

La technique de validation croisée à 10 plis assure que notre modèle est bien généralisé et évite le surapprentissage. Chaque instance est validée une fois, garantissant des résultats non biaisés par une division particulière des données. De plus, les paramètres du SSAE ont été optimisés après des tests rigoureux, assurant des performances maximales.

En utilisant la technique max-voting pour combiner les prédictions des différents ROIs, nous avons obtenu une décision consolidée par image. Cette méthode a permis d'améliorer légèrement les résultats de classification, notamment pour les sous-contours les plus informatifs, en atteignant une précision maximale de 86.20% pour le sous-contour 2.

Malgré ses avantages, notre méthode présente également certaines limites, la précision de la segmentation et de la classification dépend fortement de la qualité des images radiographiques. Les images de mauvaise qualité, avec des artefacts ou un faible contraste, peuvent entraîner des erreurs de segmentation et des prédictions incorrectes. La taille des ROIs sélectionnées peut affecter les performances du modèle. Une exploration plus approfondie de l'impact de différentes tailles de ROIs pourrait améliorer davantage les résultats. L'intégration de caractéristiques multi-échelles pourrait également capturer plus d'informations pertinentes et améliorer la robustesse du modèle. Notre étude se concentre uniquement sur les images radiographiques obtenues par la modalité CR et sur les grades KL-0 et KL-2. L'évaluation de notre méthode sur des images obtenues par d'autres modalités (comme RG, la tomographie par ordinateur ou l'imagerie par résonance magnétique) et sur d'autres grades de l'arthrose (KL-1, KL-3, KL-4) n'a pas été réalisée, ce qui limite la généralisation des résultats.

5.6 Conclusion

Cette étude présente une approche automatique pour extraire plusieurs régions d'intérêt à partir d'images radiographiques du genou de l'os tibial. En utilisant le réseau d'autoencodeurs empilés parci-

monieux, nous avons pu étudier l'impact de l'emplacement des ROIs sur les résultats de classification de l'arthrose précoce du genou (KL-0 vs KL-2). Nos résultats contribuent à optimiser la sélection de la taille et de l'emplacement des patchs pour une classification précise et fiable de l'arthrose du genou.

Bien que notre méthode montre des résultats prometteurs pour la détection précoce de l'arthrose du genou, il est essentiel de continuer à explorer et à améliorer les aspects liés à la qualité des images, à la taille des ROIs, et à la généralisation sur d'autres modalités et bases de données pour renforcer encore sa robustesse et son applicabilité clinique.

Les travaux futurs pourraient impliquer l'exploration de facteurs supplémentaires, tels que l'influence de la taille des ROIs, l'intégration de caractéristiques multi-échelles, et l'évaluation des capacités de généralisation de la méthode proposée pour d'autres bases de données. De telles avancées peuvent améliorer davantage la détection précoce et le suivi de l'arthrose du genou.

6

Combinaison de caractéristiques de forme et de texture pour le diagnostic de la gonarthrose

Sommaire

6.1	Introduction	81
6.2	Approche Proposée	81
6.3	Données et Paramètres Expérimentaux	87
6.4	Résultats expérimentaux et discussion	88
6.5	Forces et limites de l'étude	95
6.6	Conclusion	96

6.1 Introduction

Dans ce chapitre, nous introduisons une approche novatrice basée sur l'apprentissage profond qui combine des caractéristiques de forme et de texture pour prédire l'arthrose du genou à partir d'images radiographiques. Pendant de nombreuses années, l'arthrose était considérée comme une maladie du cartilage, des études récentes ont démontré que le remodelage osseux peut également contribuer à sa progression. Parmi les divers facteurs de qualité osseuse, la microarchitecture osseuse est devenue un sujet d'intérêt majeur pour comprendre la douleur associée à l'arthrose.

Plus récemment, certaines approches ont combiné des caractéristiques de forme et de texture extraites de l'ensemble de l'articulation du genou, conduisant à des prédictions plus précises de l'arthrose du genou. Motivés par ces avancées, nous proposons une nouvelle stratégie pour la détection précoce de l'arthrose du genou, qui vise à améliorer la classification en combinant des informations de texture extraites de l'os tibial avec des informations de forme de la zone de l'articulation du genou.

À cette fin, nous utilisons une architecture double, basée d'une part sur un CNN [Tiulpin et al., 2017] pour extraire des informations de forme et d'autre part sur un AE pour capturer les détails de texture [Nasser et al., 2020]. Ces deux ensembles de caractéristiques sont ensuite fusionnés à travers le modèle Compact Bilinear Pooling (CBP) [Gao et al., 2016] pour obtenir une représentation globale des images radiographiques. La classification finale est réalisée grâce à une couche entièrement connectée.

Les principales contributions de cette étude sont les suivantes :

- Proposition d'une nouvelle approche d'apprentissage profond, combinant des caractéristiques de forme et de texture pour évaluer la sévérité de l'arthrose du genou.
- Réalisation d'une étude ablative pour mettre en évidence les effets des différents composants du modèle proposé.
- Conduite d'expériences claires et complètes sur un ensemble de données publiquement disponible, la base de données OAI.

6.2 Approche Proposée

Cette section présente un aperçu des deux modèles d'apprentissage utilisés pour extraire les caractéristiques de forme et de texture associées à la gonarthrose. Le diagramme de flux de l'architecture proposée est illustré sur la Fig. 6.1.

6.2.1 Extraction des caractéristiques de forme

Pour extraire les caractéristiques liées à la forme de la zone distale du genou, nous avons affiné le modèle Siamese profond proposé dans [Tiulpin et al., 2018]. La principale nouveauté de ce modèle est un

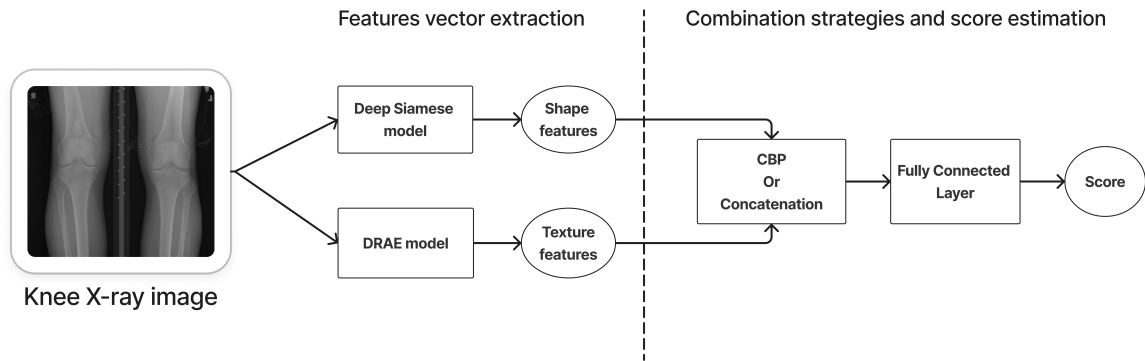


Figure 6.1 – Organigramme de l’approche proposée combinant les modèles DRAE et siamois avec le CBP pour la classification des premiers stades de la gonarthrose.

nouveau CNN efficace qui exploite la symétrie des caractéristiques visuelles au sein du genou. Nous avons sélectionné ce modèle en raison de ses critères de classification très proches des critères radiologiques utilisés pour le diagnostic de l’arthrose du genou (accord Kappa de 0.83).

Le réseau Siamese est un réseau de neurones artificiels composé de deux sous-réseaux ayant une architecture identique, chacun traitant simultanément deux échantillons différents et comparant leur similarité en calculant la distance entre leurs vecteurs de caractéristiques en utilisant soit la distance cosinus, soit la distance euclidienne. Les deux sous-réseaux partagent les mêmes paramètres, réduisant ainsi la complexité du modèle. Dans [Tiulpin et al., 2018], les auteurs ont d’abord localisé la zone de l’articulation du genou dans chaque image et l’ont découpé à une taille fixe en utilisant la méthode précédemment proposée par Antony et al. [Tiulpin et al., 2017]. Ensuite, l’image découpée est divisée en deux patches : un de la partie latérale et un de la partie médiale de l’articulation. Le patch médial est retourné horizontalement pour exploiter la symétrie de l’articulation du genou. La méthode alimente ensuite les deux patches à un réseau de neurones convolutif Siamese profond, qui se compose de deux branches identiques partageant les mêmes poids. Chaque branche extrait des caractéristiques d’un patch et produit un vecteur de caractéristiques. Les vecteurs de caractéristiques des deux branches sont concaténés et utilisés comme entrée d’une couche entièrement connectée qui prédit la distribution de probabilité de la note KL (KL-0 ou KL-2) pour l’image d’entrée, comme illustré dans la Fig. 6.2.

6.2.2 Extraction des caractéristiques de texture

Pour l’extraction des caractéristiques de texture, nous nous sommes appuyés sur le modèle d’apprentissage DRAE introduit dans [Nasser et al., 2017]. Avant de décrire le modèle DRAE, introduisons brièvement l’AE classique. L’AE est un réseau de neurones non supervisé composé de deux parties : un encodeur et un décodeur. L’encodeur est utilisé pour encoder les données d’entrée, tandis que le décodeur est utilisé pour reconstruire les données encodées. Pendant l’entraînement, l’auto-encodeur cherche à mi-

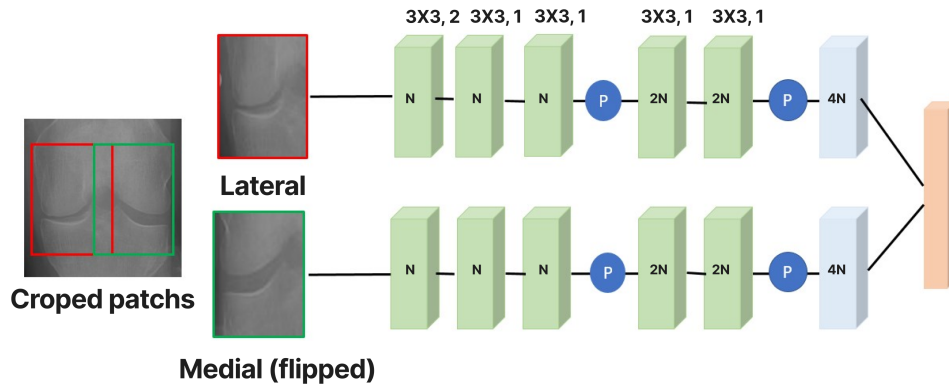


Figure 6.2 – L’architecture du modèle Deep Siamese. Le max-pooling (2×2) est représenté en bleu, les couches de convolution, la normalisation par lot (batch normalization) et le bloc ReLU en bleu clair. La couche softmax pour concaténer les sorties des deux réseaux est en orange. Les chiffres à l’intérieur des couches de convolution, de normalisation par lot et du bloc ReLU indiquent le nombre de filtres de convolution ; les chiffres au-dessus des couches de convolution correspondent à leurs paramètres ($K \times K, S$), où K représente la taille du filtre et S le pas de convolution.

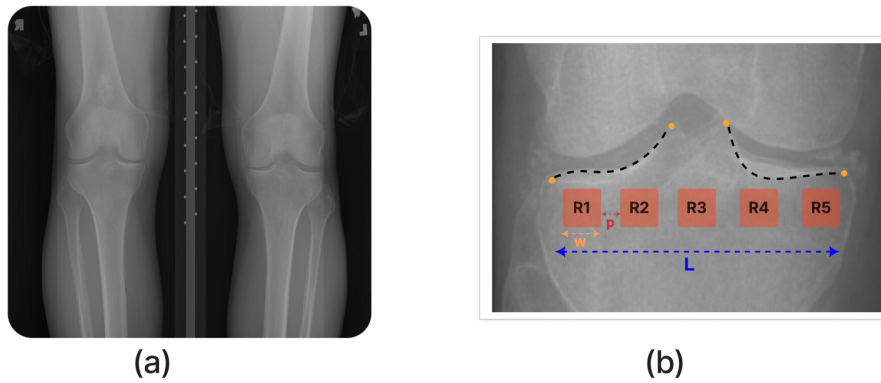


Figure 6.3 – Sélection des ROIs pour l’extraction des caractéristiques de texture. (a) Une image radiographique typique du genou issue de la base OAI. (b) Zone distale extraite avec les cinq ROIs récupérées.

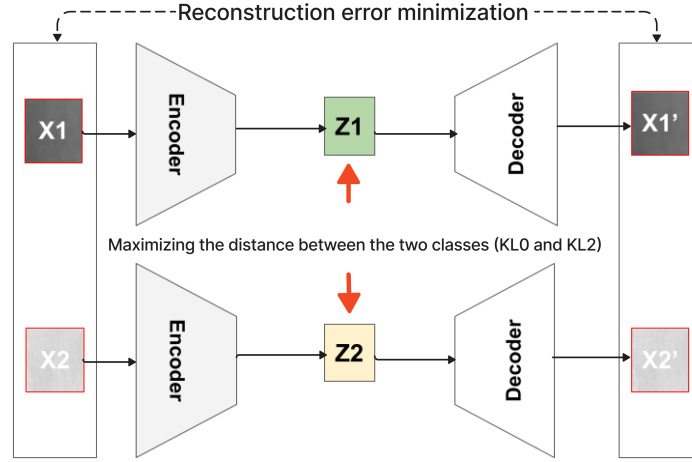


Figure 6.4 – Architecture du modèle d’apprentissage DRAE. X_1 et X_2 sont les ensembles d’entrées du réseau pour les deux classes (sujet normal et OAI minimale), X'_1 et X'_2 sont les entrées reconstruites, et Z_1 et Z_2 sont les vecteurs de caractéristiques appris en minimisant l’erreur de reconstruction et en maximisant la distance entre les caractéristiques des deux classes.

minimiser la différence entre les données d’entrée et la sortie reconstruite, en utilisant typiquement une fonction de perte telle que l’erreur quadratique moyenne définie comme suit :

$$E_{AE}(\theta) = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^I (x_{i,n} - \hat{x}_{i,n})^2 \quad (6.1)$$

où N représente le nombre d’échantillons d’entraînement, θ est le vecteur de tous les paramètres du réseau, x est le vecteur d’intensité de l’image d’entrée, et $\hat{x}$ correspond à l’entrée reconstruite.

En minimisant E_{AE} , l’autoencodeur apprend à extraire les caractéristiques les plus importantes des données d’entrée et à les encoder dans le code latent tout en éliminant les détails moins importants [Nasser et al., 2017].

Nasser *et al.* [Nasser et al., 2020] ont introduit un modèle appelé DRAE pour la détection précoce de la gonarthrose. L’objectif était d’extraire des caractéristiques discriminantes et pertinentes pour améliorer les performances de classification de la gonarthrose. L’architecture du modèle DRAE est illustrée dans la Fig. 6.4.

Le modèle DRAE utilise cinq ROIs composées essentiellement de texture (R1, R2, R3, R4, R5) de taille 32×32 extraites de la radiographie sous le plateau tibial, représentant les parties latérales, centrales et médiales du genou (voir la Fig. 6.3).

Les ROIs ont été obtenues en utilisant la méthode de segmentation décrite dans [Nasser et al., 2020].

Tout d'abord, les épines tibiales ainsi que les extrémités médiale et latérale du tibia ont été identifiées manuellement (points orange dans la Fig. 4.b). L'os sous-chondral a ensuite été automatiquement détecté comme étant le chemin le plus lumineux passant par ces marqueurs anatomiques (lignes pointillées noires dans la Fig. 6.3). Enfin, les cinq ROIs (R1 à R5) ont été localisées et extraites sous la bordure inférieure des plaques corticales en tant que parties médiale (R1, R2), centrale (R3) et latérale (R4, R5) du genou. La distance p entre les ROIs extraites a été calculée en utilisant la formule suivante :

$$p = \frac{L - 5 \cdot W}{6}, \quad (6.2)$$

où L est la distance entre les extrémités médiale et latérale, et W est la largeur de la ROI. Le modèle DRAE vise à forcer le réseau à capturer des propriétés discriminantes qui maximisent la distance entre les caractéristiques des différentes classes de la gonarthrose. À cette fin, les auteurs ont ajouté une fonction de perte discriminante à la fonction de perte standard définie dans l'Équation (6.3) :

$$E_{DRAE} = E_{AE} + \lambda \Omega_{disc} + \beta \Omega_{weights} \quad (6.3)$$

où E_{AE} est MSE correspondant à la fonction de coût de l'auto-encodeur traditionnel (Éq. 6.1), Ω_{disc} représente la pénalité discriminante, λ est le poids du terme discriminatif, et β est le terme de régularisation L^2 qui encourage le modèle à trouver minimiser les valeurs de poids, afin d'empêcher le surapprentissage.

Ω_{disc} vise à faciliter l'apprentissage de la séparabilité en maximisant la distance entre les moyennes μ_1 et μ_2 de chaque ensemble de caractéristiques encodées (Z_1 et Z_2) correspondant à chaque classe (KL-0 et KL-2) et en minimisant leurs variances, σ_1^2 et σ_2^2 . Ω_{disc} est défini comme suit :

$$\Omega_{disc} = \frac{\sigma_1^2 + \sigma_2^2}{|\mu_1 - \mu_2|^2} \quad (6.4)$$

où

$$\sigma_i^2 = \frac{1}{m^{(i)}} \sum_{j=1}^{m^{(i)}} (z_j^{(i)} - \mu_i)^2 \quad ; \quad \mu_i = \frac{1}{m^{(i)}} \sum_{j=1}^{m^{(i)}} z_j^{(i)} \quad (6.5)$$

6.2.3 Compact Bilinear Pooling (CBP)

Le CBP est un modèle utilisé en vision par ordinateur pour combiner les sorties de deux CNNs en un seul vecteur de caractéristiques [Gao et al., 2016]. Le CBP s'est révélé efficace dans diverses tâches de vision par ordinateur, notamment la classification d'images, la détection d'objets et le questionnement visuel [Liu et al., 2019].

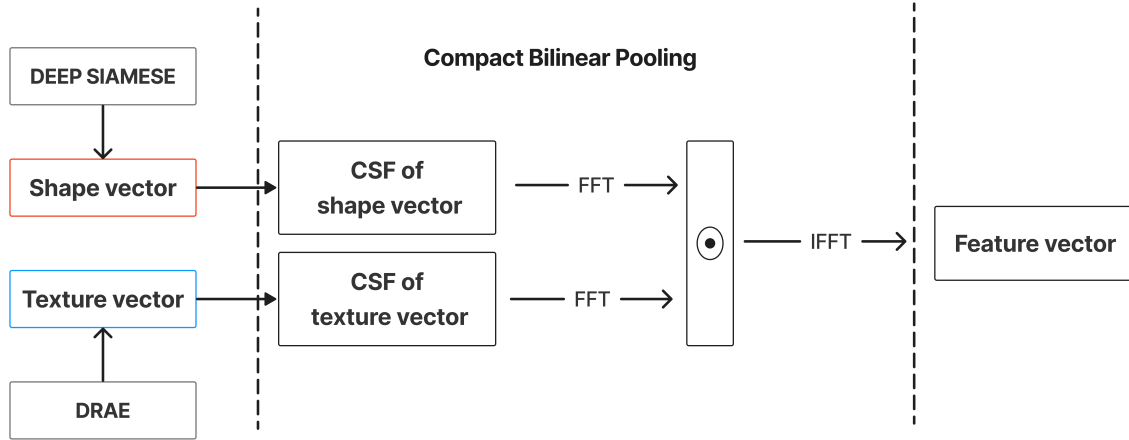


Figure 6.5 – Structure du modèle d'apprentissage CBP.

L'idée principale derrière le CBP est d'approximer le produit extérieur des deux cartes de caractéristiques des CNN avec une matrice de rang faible. Cette approximation est réalisée en appliquant une projection aléatoire à chaque carte de caractéristiques, puis en calculant leur produit élément par élément. La matrice résultante est ensuite transformée en vecteur par une opération d'aplatissement. Comparée à un produit extérieur traditionnel, le CBP permet une réduction significative de la dimensionnalité du vecteur de caractéristiques, ce qui est crucial dans des contextes où les ressources mémoire ou de calcul sont limitées. Dans cette étude, nous avons utilisé le CBP pour calculer le produit entre le vecteur de caractéristiques de texture, $V1$, et le vecteur de caractéristiques de forme, $V2$, résultant respectivement du DRAE et du réseau Siamois selon l'algorithme décrit en 1. Lorsque ces vecteurs sont trop grands, cela conduit à un grand nombre de paramètres. Pour déterminer lequel des deux vecteurs entraîne un temps de calcul long et une occupation mémoire importante et pour réduire le nombre de ses paramètres, nous avons utilisé la fonction Count Sketch Function (CSF) [Charikar et al., 2002], qui est un algorithme aléatoire utilisé pour la réduction de dimensionnalité et l'esquisse de données. CSF consiste en un tableau de d compteurs, où d est un paramètre choisi en fonction de la Accuracy (Acc) souhaitée et des contraintes de mémoire. Pour chaque élément dans le flux de données, la fonction CSF applique une fonction de hachage pour déterminer l'index du compteur à mettre à jour et une autre fonction de hachage pour déterminer s'il faut ajouter ou soustraire le compte de l'élément. L'idée clé de la CSF est de mettre à jour les compteurs de manière à préserver le produit scalaire entre le vecteur de fréquence du flux et un vecteur de requête. Cela permet d'estimer la fréquence de tout élément dans le flux en calculant son produit scalaire avec le vecteur de requête.

Algorithm 1 Compact Bi-linear Pooling (CBP)

```
1: Input :  $v_1 \in \mathbb{R}^{n_1}, v_2 \in \mathbb{R}^{n_2}$ 
2: Output :  $\Phi(v_1, v_2) \in \mathbb{R}^d$ 
3: procedure CBP( $v_1, v_2, n_1, n_2, d$ )
4:   for  $k \leftarrow 1 \rightarrow 2$  do
5:     if  $h_k$  et  $s_k$  ne sont pas initialisés then
6:       for  $i \leftarrow 1 \rightarrow n_k$  do
7:         Sample  $h_k$  from  $\{1, \dots, d\}$ 
8:         Sample  $s_k$  from  $\{1, -1\}$ 
9:       end for
10:       $v' = \Psi(v_k, h_k, s_k, n_k)$ 
11:    end if
12:  end for
13:   $\Phi = \text{FFT}^{-1}(\text{FFT}(v'_1) \odot \text{FFT}(v'_2))$ 
14:  return  $\Phi$ 
15: end procedure
16: procedure  $\Psi(v, h, s, n)$ 
17:    $y = [0, \dots, 0]$ 
18:   for  $i \leftarrow 1 \rightarrow n$  do
19:      $y[h[i]] = y[h[i]] + v[i] \cdot s[i]$ 
20:   end for
21:   return  $y$ 
22: end procedure
```

6.3 Données et Paramètres Expérimentaux

Cette section présente les données et détaille les paramètres de nos expériences.

6.3.1 Dataset

Dans cette étude, nous avons utilisé la base de données publique OAI. Plus précisément, les données baseline correspondent à la visite 0. L'une des principales idées de l'étude étant de considérer les informations de texture de l'os tibial et d'éviter les artefacts dus au processus de numérisation, seules les images obtenues par la modalité CR ont été prises en compte. De plus, pour éviter tout biais statistique associé à des données déséquilibrées, 3900 images radiographiques du genou (1300 de chaque grade, KL-0, KL-1 et KL-2) du jeu de données de base de l'OAI ont été sélectionnées.

Pour éliminer les incertitudes potentielles causées par la classe douteuse KL-1, notre approche a d'abord été évaluée en utilisant uniquement les données KL-0 et KL-2. La validation croisée en 4 plis a été employée pour valider notre approche.

6.3.2 Extraction de Caractéristiques

Pour la classification des images, l'extraction de caractéristiques implique l'identification des motifs ou attributs les plus importants susceptibles d'être utiles pour distinguer les images les unes des autres. Cela

se fait généralement en transformant l'image en un ensemble de caractéristiques, telles que des valeurs numériques représentant les couleurs, les textures et les formes présentes dans l'image.

Dans cette étude, nous nous intéressons à l'extraction et à la combinaison des caractéristiques de forme et de texture. Nous avons ajusté le modèle Siamois profond pour extraire les caractéristiques de forme, comme présenté dans la section 6.2.1. Ce modèle produit un vecteur de forme V_s de taille 512. Pour extraire les caractéristiques de texture des cinq régions d'intérêt (R1, R2, ..., R5), nous avons utilisé le modèle DRAE présenté dans la section 6.2.2. La taille du vecteur de caractéristiques de texture résultant, V_t est de 256 pour chaque région d'intérêt.

6.3.3 Combinaison de Caractéristiques

Pour combiner les caractéristiques de forme et de texture extraites, nous devons aligner la taille du vecteur de caractéristiques de forme, V_s (512) avec celles du vecteur de caractéristiques de texture, V_t (256) pour chaque région (R1 à R5). Chaque vecteur de caractéristiques de texture a été combiné avec le vecteur de caractéristiques de forme correspondant en utilisant la procédure de remplissage par des zéros. La combinaison des caractéristiques de texture garantit que les informations provenant des côtés latéral, central et médial du genou soient incorporées. Pour ce faire, nous avons commencé par combiner deux vecteurs de texture parmi les cinq vecteurs existants en utilisant une simple concaténation (par exemple, (R1, R2), (R1, R3), ..., (R2, R3), (R2, R4), ...). Les combinaisons possibles entre deux vecteurs ont été calculées, résultant en 10 vecteurs de caractéristiques de texture. Chacun de ces vecteurs avait une taille de 512, ce qui correspond à la taille du vecteur de caractéristiques de forme. Pour combiner trois et quatre vecteurs de caractéristiques de texture (par exemple, (R1, R2, R3) ou (R1, R2, R3, R4), etc.), la concaténation ne pouvait pas être réalisée car elle ne conduirait pas à un vecteur de taille correspondant à celle du vecteur de caractéristiques de forme. Par conséquent, nous avons utilisé le modèle CBP, qui permet de surmonter la disparité de taille et garantit la compatibilité entre le vecteur de caractéristiques de texture et le vecteur de caractéristiques de forme. Toutes les combinaisons possibles ont été réalisées, résultant en un total de 31 combinaisons. La procédure de remplissage avec des zéros a également été utilisée pour combler les éléments manquants des vecteurs de caractéristiques de texture. La combinaison des vecteurs de forme et de texture conduit à un vecteur de caractéristiques global de taille 512.

6.4 Résultats expérimentaux et discussion

Dans cette section, nous évaluons l'approche proposée en utilisant les données de la base OAI. Une comparaison avec l'état de l'art et une étude ablative prouvant la contribution de chaque composant de l'approche proposée sont également fournies.

Features	Models	Acc	Pr	Re	F1
Shape	Deep Siamese	82.68	84.76	82.12	83.86
	DenseNet121	80.76	83.65	81.34	82.64
	EfficientNetB0	81.84	85.63	80.43	83.12
	ResNet50	80.12	81.56	79.98	81.84
Texture	DRAE	82.53	88.23	79.22	83.28
	AE	73.12	81.29	71.88	75.21

Table 6.1 – Comparaison des modèles d’apprentissage sélectionnés basés sur la forme et la texture pour la classification (KL-0 vs. KL-2).

6.4.1 Sélection des modèles appropriés pour l’approche proposée

Notre approche proposée nécessite deux modèles, le premier pour analyser la forme et le second pour caractériser la texture. Pour sélectionner les modèles appropriés pour chaque tâche, nous avons comparé quatre modèles d’analyse de forme (Deep Siamese [Tiulpin et al., 2018], DenseNet121 [Ledingham et al., 1993], EfficientNet [Tan and Le, 2019], et ResNet50 [He et al., 2016]) et deux modèles d’analyse de texture (DRAE [Nasser et al., 2020] et AE [Bourlard and Kamp, 1988]).

Avec le jeu de données OAI, les modèles d’apprentissage basés sur la forme ont été comparés en utilisant les régions de l’articulation du genou extraites en utilisant l’approche proposée dans [Tiulpin et al., 2018]. Pour comparer les modèles d’apprentissage basés sur la texture, nous avons utilisé les régions extraites en utilisant l’approche décrite dans [Nasser et al., 2020]. Les performances sont évaluées en termes d’Accuracy (Acc), Precision (Pr), Recall (Re) et F1-score (F1). Le Tableau 6.1 présente les performances de chaque modèle.

Comme on peut le voir, pour les modèles d’apprentissage basés sur la forme, à l’exception de la précision, les scores les plus élevés ont été atteints en utilisant le modèle Deep Siamese (Acc = 82.68%, Re = 82.12% et F1 = 83.86%). La Pr la plus élevée (85.63%) a été atteinte par le modèle EfficientNetB0. En ce qui concerne les modèles d’apprentissage basés sur la texture, comme nous avons cinq régions par image, la technique de vote majoritaire a été employée pour obtenir une seule décision par image, comme suit :

$$\hat{y}_i = \text{mode}\{\hat{y}_i^{(1)}, \hat{y}_i^{(2)}, \hat{y}_i^{(3)}, \hat{y}_i^{(4)}, \hat{y}_i^{(5)}\} \quad (6.6)$$

où $\hat{y}_i^{(j)}$ est le label prédit de la ROI R_j extrait de la $i^{\text{ème}}$ image. *mode* représente la valeur la plus fréquente. Le score de vote majoritaire a été calculé en utilisant les labels obtenus $\hat{y}_i$ et les labels réels y_i .

Le Tableau 6.1 montre que les scores les plus élevés sont atteints en utilisant le modèle DRAE (Acc = 82.53%, Pr = 88.23%, Re = 79.22% et F1 = 83.28%).

Pour conclure cette comparaison, nous pouvons affirmer que les modèles Deep Siamese [Tiulpin et al., 2018] et DRAE [Nasser et al., 2020] sont les plus adaptés, respectivement pour les caractéristiques de

Models	ROIs	DRAE			
		Acc	Pr	Re	F
Deep Siamese	1	77.12	82.63	74.98	79.14
	2	85.33	87.44	78.09	85.88
	3	75.33	75.33	70.12	79.12
	4	78.26	82.22	79.83	80.11
	5	83.17	87.27	80.12	84.63
Max-voting		86.69	90.03	81.12	86.38
DenseNet121	1	70.14	80.75	73.92	72.45
	2	73.12	81.29	71.88	75.21
	3	71.68	82.33	74.09	75.71
	4	71.76	81.94	70.76	71.85
	5	72.83	81.42	73.51	74.39
Max-voting		74.90	84.16	75.62	77.23
EfficientNet	1	79.67	83.12	78.19	83.76
	2	81.23	85.56	78.84	84.92
	3	78.89	80.78	80.55	81.99
	4	80.01	81.47	79.21	82.10
	5	80.45	82.34	78.97	81.35
Max-voting		83.29	87.19	81.62	86.26
ResNet50	1	78.99	82.70	78.43	83.95
	2	83.15	85.35	77.56	81.04
	3	79.78	80.12	77.89	83.78
	4	81.22	81.88	77.21	81.67
	5	82.79	84.92	76.45	82.33
Max-voting		84.43	86.94	79.83	84.90

Table 6.2 – Effets du DRAE combiné à la stratégie CBP pour améliorer les modèles DL de State Of The Art (SOTA) sélectionnés.

forme et de texture pour la suite de notre étude.

6.4.2 Effets des stratégies de combinaison sur les modèles DL de l'état de l'art

Dans cette section, en utilisant les différentes régions extraites (R1, R2, R3, R4 et R5) comme entrées du modèle DRAE, nous étudions les effets du CBP pour améliorer les modèles d'apprentissage basés sur la forme sélectionnés précédemment (Deep Siamese [Tiulpin et al., 2018], DenseNet121 [Ledingham et al., 1993], EfficientNet [Tan and Le, 2019] et ResNet [He et al., 2016]), qui se concentrent sur la capture des caractéristiques géométriques ou structurelles spécifiques des objets ou des motifs dans les données. Pour obtenir une décision unique et comparer les résultats, le vote majoritaire a également été employé. Le Tableau 6.2 résume les résultats obtenus, qui devraient être comparés à ceux du Tableau 6.1, où le modèle DRAE n'est pas pris en compte. Comme on peut le voir, l'incorporation des informations de texture extraites à l'aide du modèle DRAE améliore les performances de tous les modèles de l'état de

l'art sélectionnés. Grâce aux informations de texture combinées à l'aide du modèle CBP, les scores les plus élevés ont été obtenus avec le réseau Siamese, améliorant ses performances (Acc +4.01%, Pr +5.27%, Re -1.00% et F1 +2.52%). De plus, à l'exception du modèle DenseNet121, les performances de tous les autres modèles sélectionnés se sont améliorées : Précision (EfficientNetB0 +1.45%, ResNet50 +4.31%), pour la précision (EfficientNetB0 +1.56%, ResNet50 +5.38%), pour le rappel (EfficientNetB0 +1.19%, ResNet50 -0.15%) et pour le F1 (EfficientNetB0 +3.14%, ResNet50 +3.06%). Pour les réseaux Siamese et ResNet50, on peut dire que la baisse du rappel est compensée par l'augmentation de la précision. Ces résultats suggèrent que la stratégie CBP proposée combinant les informations de texture et de forme est potentiellement prometteuse pour améliorer la détection précoce de la gonarthrose.

6.4.3 Étude ablative et évaluation des performances des stratégies proposées

Dans cette section, nous évaluons la performance des différentes stratégies de combinaison (CBP) pour la prédiction précoce de la gonarthrose (KL-0 vs. KL-2). Le Tableau 6.3 résume les résultats obtenus en utilisant à la fois la zone distale et les régions de texture (R1, R2, R3, R4 et R5) comme entrées de notre modèle proposé. Tous les vecteurs de caractéristiques obtenus à partir des régions de texture et le vecteur de forme ont été combinés pour entraîner notre modèle.

Selon le Tableau 6.3, le meilleur résultat est obtenu en utilisant le modèle CBP, avec une Acc de 91.67% et une Pr de 93.22%. Ces scores sont obtenus en combinant les caractéristiques de forme du genou de la zone distale et les caractéristiques de texture des régions R2 et R5. En utilisant les mêmes entrées, la méthode de concaténation n'atteint qu'une Acc de 84.85% et une Pr de 85.45%. En termes de F1 et de Re, les scores les plus élevés sont systématiquement obtenus en utilisant le modèle CBP avec 91.10% et 89.78%, respectivement, qui sont atteints en combinant la zone distale avec les régions de texture R2, R3 et R5. Le rappel atteint en utilisant la méthode de concaténation est inférieur (87.11%), tandis que le F1 est plus élevé (90.72%).

Les scores obtenus montrent que la combinaison des informations de la zone distale avec les caractéristiques de texture en utilisant notre modèle CBP proposé donne de meilleurs résultats par rapport à une simple concaténation des caractéristiques. Une analyse plus approfondie des résultats montre que l'incorporation de la texture de l'os tibial, englobant les informations des régions centrale, médiale et latérale du genou, améliore la détection précoce de la gonarthrose. Les deux régions qui fournissent les informations les plus utiles étant les régions R2 (extrait de la partie médiale du genou) et R5 (extrait de la partie latérale du genou). R3, extrait de la partie centrale du genou, fournit également des informations précieuses, surtout lorsqu'il est combiné avec les régions R2 et R5. Re = 89.78% et F1 = 91.10% avec le modèle CBP et un F1 = 90.72% lors de l'utilisation de l'approche de concaténation.

D'après les résultats ci-dessus, on peut affirmer que, selon notre approche proposée, les symptômes initiaux de la gonarthrose sont principalement observables dans les parties latérale et médiale de l'os

tibial, ce qui est conforme à d'autres études [Woloszynski et al., 2010, Nasser et al., 2020, Charikar et al., 2002]. Cela souligne également l'importance de sélectionner des régions d'intérêt pour classifier les premiers stades de la maladie, car cela est crucial pour l'identification et le diagnostic précoces.

6.4.4 Classification du grade douteux KL-1

Comme mentionné dans l'introduction, KL-1 est considéré comme douteux et pourrait donc entraîner des difficultés dans l'entraînement des modèles d'apprentissage. Cependant, d'un point de vue clinique, KL-1 reste un atout en tant que candidat principal pour avancer vers KL-2. Par conséquent, nous avons également effectué une classification binaire pour (KL-0 vs. KL-1) et (KL-1 vs. KL-2). Comme indiqué dans le Tableau 6.4, comparé à l'état de l'art, notre modèle obtient les meilleurs résultats de classification. Pour (KL-0 vs. KL-1), une Acc de 74.19%, une Pr de 76.33%, un Re de 71.20% et un F1 de 77.17% ont été atteints. En considérant KL-1 vs. KL-2, de meilleures performances ont été obtenues avec une Acc de 79.02%, une Pr de 81.17%, un Re de 75.12, et un F1 de 79.98%. Globalement, ces résultats montrent que les classes KL-0 et KL-1 sont très similaires, rendant la tâche de classification plus difficile.

6.4.5 Synthèse des résultats et discussion

Dans cette section, nous discutons les résultats obtenus dans notre étude. Nous avons tout d'abord évalué différents modèles pour la caractérisation de la forme (Deep Siamese, DenseNet121, EfficientNet et ResNet) et la caractérisation de la texture (DRAE et AE), identifiant les meilleurs modèles pour chaque tâche. Le modèle Deep Siamese a montré de bonnes performances pour l'extraction des caractéristiques de forme, atteignant une Acc de 82,68%, un Re de 82,12% et un F1 de 83,86%. En revanche, le modèle DRAE s'est démarqué dans l'analyse de la texture avec une Acc de 82,53%, une Pr de 88,23%, un Re de 79,22% et un F1 de 83,28%.

Ensuite, nous avons évalué l'impact de l'utilisation de CBP pour combiner les caractéristiques de texture extraites par le modèle DRAE, améliorant ainsi les modèles d'apprentissage basés sur la forme. Les résultats ont montré que l'intégration des informations de texture améliore significativement les performances de tous les modèles sélectionnés. Le modèle Deep Siamese a atteint les meilleures améliorations, avec une augmentation de la Acc de 4,01%, de la Pr de 5,27% et du F1 de 2,52%, bien qu'il y ait eu une légère baisse du Re de 1,00%. De plus, à l'exception de DenseNet121, tous les autres modèles ont montré des améliorations : EfficientNetB0 (exactitude +1,45%, précision +1,56%, rappel +1,19%, F1 +3,14%) et ResNet50 (Acc +4,31%, Pr +5,38%, Re -0,15%, F1 +3,06%). Ces résultats soulignent le potentiel de CBP pour améliorer les modèles basés sur les caractéristiques de forme afin de détecter précocement l'arthrose du genou.

Ensuite, nous avons évalué différentes stratégies de combinaison (concaténation et CBP) pour intégrer

ROIs					CBP				Concatenation			
R1	R2	R3	R4	R5	Acc (%)	Pr (%)	Re (%)	F (%)	Acc (%)	Pr (%)	Re (%)	F (%)
✓	✗	✗	✗	✗	77.12	82.63	74.98	79.14	79.42	83.12	77.57	80.09
✗	✓	✗	✗	✗	85.33	87.44	78.09	85.88	82.42	88.46	79.99	83.33
✗	✗	✓	✗	✗	75.33	75.33	70.12	79.12	80.12	83.50	75.00	79.18
✗	✗	✗	✓	✗	78.26	82.22	79.83	80.11	70.22	72.27	66.76	71.10
✗	✗	✗	✗	✓	83.17	87.27	80.12	84.63	80.33	84.17	80.12	80.60
✓	✓	✗	✗	✗	78.13	79.26	78.66	77.12	80.21	82.13	79.11	81.27
✓	✗	✓	✗	✗	83.33	85.28	79.11	82.09	83.38	83.97	82.40	84.09
✓	✗	✗	✓	✗	75.17	78.90	72.13	77.02	79.12	78.99	78.54	80.39
✓	✗	✗	✗	✓	75.00	77.89	74.33	75.46	74.13	77.13	75.00	76.12
✗	✓	✓	✗	✗	84.34	86.60	80.03	84.14	82.76	80.17	80.18	81.18
✗	✓	✗	✓	✗	79.17	79.83	75.88	80.00	76.17	77.32	75.87	76.39
✗	✓	✗	✗	✓	91.67	93.22	89.12	90.98	84.85	85.45	80.12	84.97
✗	✗	✓	✓	✗	82.63	87.08	81.43	82.67	74.93	76.20	75.39	75.12
✗	✗	✓	✗	✓	78.10	79.19	77.17	78.30	83.97	85.30	82.24	83.13
✗	✗	✗	✓	✓	83.23	85.13	82.64	82.12	80.10	81.37	80.03	81.11
✓	✓	✓	✗	✗	79.32	80.65	80.54	79.98	79.06	82.48	78.64	84.49
✓	✓	✗	✓	✗	78.23	79.62	79.43	80.09	74.70	75.89	73.55	80.57
✓	✓	✗	✗	✓	87.56	89.34	88.99	89.65	81.07	82.11	80.68	83.48
✓	✗	✓	✓	✗	77.90	78.09	78.00	79.02	74.39	76.89	73.89	78.34
✓	✗	✓	✗	✓	83.90	85.09	84.87	84.26	81.07	82.11	80.68	84.48
✓	✗	✗	✓	✓	84.12	84.90	85.29	84.55	87.07	88.0	85.59	89.93
✗	✓	✓	✓	✗	83.48	84.32	82.84	85.12	76.55	77.35	74.57	79.29
✗	✓	✓	✗	✓	90.38	91.13	89.78	91.10	88.31	89.68	87.11	90.72
✗	✓	✗	✓	✓	86.55	85.72	86.03	85.98	81.61	83.20	77.19	84.56
✗	✗	✓	✓	✓	84.43	85.01	85.23	84.12	79.40	80.26	77.67	82.48
✓	✓	✓	✓	✗	83.86	83.99	84.61	85.13	83.61	84.96	78.38	85.08
✓	✓	✓	✗	✓	86.08	85.91	86.02	86.23	88.19	87.6	87.5	88.44
✓	✓	✗	✓	✓	87.22	88.12	86.98	88.67	88.68	87.92	86.0	89.05
✓	✗	✓	✓	✓	84.29	84.98	83.17	86.00	82.25	82.90	81.300	83.73
✗	✓	✓	✓	✓	85.09	85.91	84.48	85.90	84.84	81.39	83.39	85.82
✓	✓	✓	✓	✓	83.18	84.17	83.28	84.01	79.58	81.42	80.23	82.76

Table 6.3 – Étude ablative et évaluation des performances obtenues en utilisant les caractéristiques de forme et de texture du genou provenant de (R1, R2, ... et R5) combinées à l’aide du CBP et et d’une simple concaténation.

Model	ROIs	KL-0 vs. KL-1				KL-1 vs. KL-2			
		Acc	Pr	Re	F	Acc	Pr	Re	F
DRAE + Deep Siamese	1	69.45	71.98	68.45	70.34	74.67	78.32	71.54	74.22
	2	72.77	74.23	69.23	74.12	77.14	79.68	74.21	76.45
	3	68.12	72.34	67.12	71.56	73.89	77.79	72.89	74.10
	4	67.34	70.47	66.89	69.78	73.45	76.47	70.36	73.39
	5	71.77	73.11	69.78	73.21	76.11	79.14	73.12	75.23
Max-voting		74.19	76.33	71.20	77.17	79.02	81.17	75.12	79.78
DRAE + DenseNet121	1	64.89	70.30	63.13	68.44	69.18	74.13	68.01	71.54
	2	66.56	72.13	65.86	69.17	70.49	75.48	70.00	74.40
	3	63.74	68.12	62.32	64.29	66.00	69.09	65.59	68.32
	4	59.98	62.18	59.02	60.04	65.15	67.00	64.03	66.65
	5	65.70	72.76	70.03	66.18	68.50	75.17	66.17	74.37
Max-voting		68.83	74.17	69.33	70.76	71.32	75.48	71.17	75.09
DRAE + EfficientNet	1	68.23	70.19	65.02	70.23	71.37	76.18	69.13	72.29
	2	70.12	74.09	67.12	73.17	73.29	78.01	70.42	75.03
	3	65.04	69.23	64.73	68.26	69.18	70.00	66.52	70.61
	4	63.17	66.23	60.84	62.29	66.17	69.44	64.58	67.64
	5	69.79	73.53	68.39	73.03	72.84	77.33	69.99	73.33
Max-voting		71.23	75.49	69.83	75.09	75.33	79.86	73.11	76.82
DRAE + Resnet50	1	68.36	72.65	65.84	71.54	73.12	74.21	68.42	73.11
	2	69.23	74.00	68.00	74.03	76.00	78.29	73.17	75.34
	3	67.87	69.57	66.57	70.17	70.57	71.33	67.56	70.67
	4	65.79	67.76	65.12	67.79	68.23	69.83	65.91	67.78
	5	66.45	73.12	67.23	72.31	74.23	77.59	72.15	74.89
Max-voting		72.39	75.84	70.19	75.44	77.26	80.23	75.18	77.10

Table 6.4 – Performances de classification obtenues pour (KL-0 vs. KL-1) et (KL-1 vs. KL-2) en combinant les caractéristiques du DRAE et du modèle Deep Siamese avec la stratégie CBP.

les caractéristiques de texture et de forme. Le modèle CBP a obtenu les meilleures performances avec une Acc de 91,67% et une Pr de 93,22%, en utilisant les régions de texture R2 et R5. En revanche, la concaténation a atteint une Acc (84,85%) et une Pr (85,45%) inférieures. De plus, le modèle CBP a systématiquement obtenu un F1 plus élevé (91,10%) et un Re (89,78%) supérieur en intégrant les régions de texture R2, R3 et R5, par rapport à la concaténation (F1 90,72%, rappel 87,11%).

Ces résultats montrent que l'intégration des caractéristiques de texture provenant de régions spécifiques (notamment R2 et R5) de l'os tibial améliore la détection précoce de l'arthrose du genou. La région R3 (partie centrale du genou) a également fourni des informations précieuses, surtout lorsqu'elle était combinée avec R2 et R5, suggérant que les premiers signes de l'arthrose du genou sont principalement observables dans les régions latérales et médiales de l'os tibial. Cela souligne l'importance de sélectionner des ROIs appropriées pour une identification et un diagnostic précis des stades précoces de l'arthrose du genou.

En outre, nous avons mené des expériences de classification pour les stades (KL-0 vs KL-1) et (KL-1 vs KL-2). La classe KL-1, étant incertaine, pose des défis pour l'entraînement des modèles, mais sur le plan clinique, il est essentiel de considérer KL-1 comme marquant la transition critique vers KL-2. Notre modèle proposé a surpassé les modèles de l'état de l'art dans les deux expérimentations. Pour KL-0 vs KL-1, la Pr était de 74,19%, l'Acc de 76,33%, le Re de 71,20% et F1 de 77,17%. Pour (KL-1 vs KL-2), le modèle a atteint une Pr de 79,02%, une Acc de 81,17%, un Re de 75,12% et un F1 de 79,98%. Ces résultats soulignent la difficulté de distinguer entre les classes KL-0 et KL-1 en raison de leur grande similitude.

En conclusion, notre étude démontre l'efficacité d'utiliser le CBP pour combiner des caractéristiques de forme et de texture de régions spécifiques du genou pour améliorer la détection précoce de l'arthrose du genou. Les résultats mettent également en évidence la pertinence clinique de classer avec précision KL-1 pour faciliter une intervention rapide dans l'ostéoarthrite du genou.

6.5 Forces et limites de l'étude

Notre approche propose une méthode novatrice pour la détection précoce de l'arthrose du genou en utilisant une combinaison de caractéristiques de texture et de forme extraites des images radiographiques. En intégrant des modèles avancés d'apprentissage profond comme Deep Siamese et DRAE, nous avons pu améliorer significativement la représentation des images, capturant ainsi les caractéristiques locales et globales liées aux changements dus à l'arthrose. Cette combinaison, facilitée par le modèle CBP, a permis d'augmenter l'Acc, la Pr et F1, démontrant une nette amélioration des performances de détection.

L'analyse de différentes régions du tibia a révélé que les premiers signes de la maladie sont principalement visibles dans les régions latérales et médiales du genou. Cette découverte souligne l'importance de la sélection des ROIs pour une classification précise des stades précoces de l'arthrose. L'approche basée

sur le CBP offre plusieurs avantages, notamment une meilleure représentation des caractéristiques, un pouvoir discriminant accru et une réduction de la dimensionnalité des données.

Cependant, notre méthode présente également certaines limites. La complexité calculatoire liée à l'utilisation du CBP nécessite des ressources importantes, ce qui peut poser des défis pour une application à grande échelle ou en temps réel. De plus, la méthode peut être sensible aux variations dans les données d'entrée, telles que la qualité des images radiographiques et les différences inter-patients, ce qui pourrait affecter la généralisation de la méthode.

Un autre défi majeur réside dans la classification des images de grade KL-1, souvent confondues avec celles des grades KL-0 ou KL-2, en raison de leur similitude. Cette incertitude peut nuire à la précision globale de la méthode pour les cas précoces d'arthrose. De plus, la sélection des régions d'intérêt peut introduire un biais si certaines régions pertinentes sont négligées ou mal définies, affectant ainsi l'exactitude des résultats de détection.

En conclusion, bien que notre approche montre des performances prometteuses pour la détection précoce de l'arthrose du genou grâce à la combinaison des caractéristiques de texture et de forme, elle présente également des défis en termes de complexité computationnelle, variabilité des données et précision pour les classes douteuses. Une attention particulière doit être portée à la sélection des régions d'intérêt et à la gestion des biais potentiels pour maximiser l'efficacité de la méthode dans des applications cliniques.

6.6 Conclusion

Dans cette étude, nous avons présenté une nouvelle approche utilisant les informations de texture et de forme pour la détection précoce de l'ostéoarthrite du genou à partir d'images radiographiques de la base de données OAI. En utilisant le modèle CBP, notre stratégie proposée combine les informations extraites des images radiographiques par deux modèles d'apprentissage : l'un basé sur l'extraction des caractéristiques de texture et l'autre sur celles de forme. Grâce à CBP, nous avons pu efficacement combiner et intégrer les caractéristiques de texture et de forme, capturant les caractéristiques locales et globales liées aux changements dus à l'arthrose.

En utilisant différentes régions extraites de l'os tibial, nos résultats montrent que les premiers symptômes de l'ostéoarthrite du genou sont principalement visibles dans les régions latérales et médiales du genou. Cela souligne l'importance de sélectionner des régions d'intérêt appropriées pour classifier avec précision l'ostéoarthrite du genou à un stade précoce. CBP présente plusieurs avantages, notamment une meilleure représentation des caractéristiques, une meilleure capacité discriminative et une réduction de la dimensionnalité.

L'approche proposée promet d'améliorer la prise de décision clinique, de faciliter l'intervention précoce, et enfin d'améliorer la gestion de l'arthrose du genou en pratique clinique.

7

Conclusion générale

Sommaire

7.1	Conclusion	99
7.2	Résumé des Contributions	100
7.3	Perspectives	101

7.1 Conclusion

Dans cette thèse, nous avons exploré et développé plusieurs approches innovantes pour améliorer la détection précoce de l'arthrose du genou, en particulier la gonarthrose, en utilisant des techniques avancées d'apprentissage profond appliquées aux images radiographiques.

Tout d'abord, nous avons mis en évidence l'importance des réseaux de neurones et des architectures comme les autoencodeurs, les réseaux convolutifs (CNN), les réseaux de neurones récurrents (RNN), ainsi que des modèles plus récents comme les transformers et les GNN, qui sont devenus des outils incontournables dans le domaine de la vision par ordinateur et du traitement des données. Ces architectures ont permis de traiter efficacement les images de radiographies du genou, d'extraire des caractéristiques locales et globales pertinentes et d'améliorer la précision des diagnostics.

Nous avons présenté une approche basée sur les autoencodeurs empilés parcimonieux (SSAE) pour obtenir des représentations significatives des (ROI)s extraites des images radiographiques du genou, en étudiant l'impact de leur position sur les performances de classification.

L'introduction d'une approche combinant des informations de texture et de forme, notamment à travers le modèle CBP, a permis de mieux différencier les stades précoces de l'arthrose (KL-0 vs KL-2), avec des performances notables. Nos résultats ont démontré que l'intégration des caractéristiques de texture et de forme dans les modèles d'apprentissage profond améliore considérablement la détection précoce de l'arthrose, avec une précision maximale de 91,67 %. Cette approche est prometteuse pour la pratique clinique, permettant une prise de décision plus précise et une intervention précoce pour les patients.

Les expériences menées pour la classification des grades (KL-0 vs KL-1, KL-1 vs KL-2) ont mis en lumière les défis associés à la détection du stade intermédiaire (KL-1), qui est difficile à distinguer en raison du caractère douteux de la labellisation. Cependant, grâce à l'apprentissage profond, notre modèle a surpassé les approches existantes (SOTA) pour la classification du grade KL-1, avec des performances particulièrement impressionnantes pour les stades KL-1 et KL-2.

Enfin, bien que cette étude ait montré des résultats prometteurs, plusieurs défis demeurent. La qualité des images, la taille et le placement des ROIs, ainsi que la généralisation des modèles à d'autres bases de données et modalités d'imagerie, sont des axes d'amélioration pour renforcer la robustesse et l'applicabilité clinique de la méthode. Les futures recherches pourraient explorer l'intégration de caractéristiques multi-échelles, ainsi que l'évaluation des modèles sur d'autres grades de l'arthrose et d'autres modalités d'imagerie, pour améliorer encore la détection précoce et le suivi de la progression de la maladie.

7.2 Résumé des Contributions

Cette thèse comporte deux contributions majeures, qui ont été développées et évaluées au cours des chapitres 4 et 5.

Contribution 1 : Approche basée sur des autoencodeurs empilés parcimonieux pour l'extraction de ROI.

Dans le chapitre 4, nous avons exploré une approche pour la détection précoce de l'arthrose du genou en utilisant des autoencodeurs empilés parcimonieux (SSAE) pour extraire des caractéristiques à partir des (ROIs). La sélection des ROIs a été réalisée par une méthode basée sur la segmentation des contours de l'os tibial. Nous avons utilisé un outil de segmentation automatisée, BoneFinder, pour délimiter précisément les zones d'intérêt, assurant une segmentation fiable des images radiographiques. Cette segmentation a permis de déterminer les ROIs qui représentent les zones les plus pertinentes pour l'analyse de l'arthrose.

Les ROIs sélectionnées étaient de taille 32 X 32. Les régions du tibia médian et latéral ont été considérées comme les plus importantes, car elles sont les premières à être affectées par les changements dégénératifs de l'arthrose. Après avoir segmenté et extrait les ROIs de ces zones, nous avons utilisé le modèle SSAE pour extraire des caractéristiques qui ont ensuite été utilisées pour classifier les images radiographiques des stades KL-0 et KL-2.

Les résultats ont montré que la sélection des contours et l'extraction des ROIs étaient des étapes cruciales pour améliorer les performances du modèle. La taille des ROIs a également joué un rôle important, car des zones plus petites ont permis une détection plus précise des changements locaux. Les performances du modèle ont atteint une précision maximale de 86,20% pour le sous-contour 2, ce qui représente une amélioration par rapport aux méthodes traditionnelles de classification par les radiologues, et ce, grâce à l'automatisation de l'extraction et de la classification des ROIs.

Contribution 2 : Détection précoce de l'arthrose par la combinaison de caractéristiques de texture et de forme.

Dans le chapitre 5, nous avons développé une méthode basée sur la combinaison des caractéristiques de texture et de forme pour la détection précoce de l'arthrose.

Dans cette étude, nous avons évalué plusieurs modèles pour la caractérisation de la forme et de la texture dans la détection précoce de l'arthrose du genou. Le modèle **Deep Siamese** a montré des performances prometteuses dans l'analyse de la forme, atteignant une *accuracy* de 82,68%, un *recall* de 82,12% et un *F1 score* de 83,86%. En parallèle, pour la texture, le modèle **DRAE** a montré une *accuracy* de 82,53%, une *précision* de 88,23%, un *recall* de 79,22% et un *F1 score* de 83,28%.

L'utilisation du modèle **CBP** pour combiner les caractéristiques extraites par les deux modèles **Deep Siamese** et **DRAE** a permis une amélioration significative des performances. Par exemple, le modèle **Deep Siamese** a vu une augmentation de l'*accuracy* de 4,01%, de la *précision* de 5,27% et du *F1 score*

de 2,52%, avec une légère baisse de 1% pour le *recall*. Les autres modèles (EfficientNetB0 et ResNet50) ont aussi montré des améliorations, notamment une augmentation notable du *F1 score* et de la *précision*.

L'évaluation des différentes stratégies de combinaison des caractéristiques a montré que le modèle CBP était plus performant que la concaténation classique des caractéristiques. Avec une *accuracy* de 91,67% et une *précision* de 93,22%, **CBP** a montré une amélioration par rapport à la concaténation (*accuracy* de 84,85% et *précision* de 85,45%).

En ce qui concerne la classification des stades KL-0, KL-1 et KL-2, les résultats ont montré que la classification des classes KL-0 et KL-1 reste un défi. Pour la classification de KL-0 vs KL-1, le modèle a atteint une *précision* de 74,19% et un *F1 score* de 77,17%. Pour KL-1 vs KL-2, les résultats ont montré une *précision* de 79,02% et un *F1 score* de 79,98%.

Ces résultats soulignent l'efficacité de l'approche combinée **CBP** et l'importance de la sélection des ROIs, notamment dans les zones latérales et médiales de l'os tibial, pour une détection précoce et précise de l'arthrose du genou. Cependant, la classification du grade KL-1 reste un défi important en raison du caractère douteux du label et de la similitude entre les stades précoces et intermédiaires de la maladie.

7.3 Perspectives

Les contributions développées dans ce manuscrit ont permis de proposer des approches performantes pour la détection précoce de l'arthrose du genou. Cependant, plusieurs limitations subsistent, et des pistes d'amélioration peuvent être explorées pour renforcer l'impact et la portée de ces travaux. Ces perspectives incluent des axes suivants :

- **Généralisation et robustesse à travers d'autres bases de données :** Nos expérimentations ont été réalisées principalement sur la base de données OAI, qui constitue une ressource précieuse pour l'étude de l'arthrose. Afin d'évaluer la robustesse et la généralisation de nos modèles, il est essentiel d'intégrer des bases de données complémentaires, telles que la base MOST. Une telle démarche permettrait non seulement de vérifier la stabilité des performances de nos approches sur des populations variées, mais également d'adapter les modèles à différentes conditions d'acquisition, améliorant ainsi leur applicabilité en conditions cliniques réelles.
- **Extension vers une classification multi-classes :** Notre travail s'est concentré sur la détection précoce de l'arthrose, particulièrement sur des grades spécifiques (KL-0 vs KL-2). Une extension naturelle serait d'envisager une classification multi-classes pour prédire tous les grades de l'arthrose (KL-0 à KL-4). Cette approche fournirait une évaluation plus complète et précise de la gravité de la maladie, aidant ainsi à une meilleure prise en charge des patients selon le stade de progression de l'arthrose.
- **Incorporation de données cliniques textuelles :** Les données d'imagerie offrent une richesse

d'informations visuelles sur les articulations, mais elles ne capturent pas toujours l'ensemble des facteurs contextuels influençant la progression de l'arthrose. Les données cliniques, telles que les antécédents médicaux, les scores fonctionnels, ou les symptômes rapportés par les patients, pourraient être intégrées sous forme de texte dans une architecture multimodale. Cela permettrait de tirer parti de ces informations complémentaires pour enrichir les modèles, conduisant à des prédictions plus précises et contextualisées.

- **Exploration de nouvelles modalités d'imagerie :** Actuellement, nos approches exploitent principalement des radiographies XR. L'ajout de données issues d'autres modalités d'imagerie, comme l'IRM, pourrait fournir des informations supplémentaires sur les structures osseuses, les tissus mous et le cartilage. En combinant ces différentes modalités, il serait possible de concevoir des modèles capables de capturer une vue multidimensionnelle plus enrichie des pathologies articulaires.
- **Développement d'architectures multidimensionnelles :** L'intégration des données radiographiques 2D avec des données volumétriques 3D issues de l'IRM nécessiterait le développement d'architectures multidimensionnelles capables de traiter conjointement des informations 2D et 3D. Ces architectures pourraient mieux modéliser les relations spatiales et structurelles complexes dans les données médicales, ouvrant la voie à des diagnostics encore plus précis.
- **Adoption de modèles de deep learning émergents :** Les Vision Transformers (ViT) et leurs dérivés ont démontré leur efficacité dans diverses applications de vision par ordinateur grâce à leur mécanisme d'attention globale. Ces modèles pourraient être explorés pour traiter à la fois les radiographies et les IRM, permettant de capturer des relations globales dans les images, souvent essentielles pour détecter les signes subtils de l'arthrose. De plus, ces modèles pourraient être adaptés à un cadre multimodal pour combiner efficacement les données d'imagerie et les données cliniques.
- **Optimisation des architectures et des hyperparamètres :** L'utilisation de techniques avancées d'AutoML (Automated Machine Learning) pourrait permettre d'automatiser la recherche des architectures et des hyperparamètres optimaux. Cela réduirait la complexité du développement des modèles tout en garantissant des performances optimales pour différentes tâches de diagnostic.

Publications Associées

Les travaux présentés dans ce manuscrit ont donné lieu aux publications suivantes :

Article de revue

- T. Mehdi, Y. Nasser, S. Aloui, R. Jennane. *Deep learning based approach combining shape and texture features for knee osteoarthritis prediction from X-ray images*. **Biomedical Signal Processing and Control**, 2025, vol. 101, p. 107172. [[Mehdi et al., 2025](#)]

Papiers conférences

- T. Mehdi, Y. Nasser, A. Almhdie, S. Aloui, R. Jennane. *Effects of region of interest location on osteoarthritis detection using deep feature learning*. In : **2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA)**. IEEE, 2023, p. 16. [[Tinhinane et al., 2023](#)]

Autres

- T. Mehdi, S. Aloui, E. Lespessailles, R. Jennane. *Contribution aux techniques d'apprentissage automatique pour le diagnostic de la gonarthrose*, **33^{ème} Colloque BIOTECHNOCENTRE**, Center Parcs de Sologne, 7-8 Octobre 2021.

Bibliographie

- [Almhdie-Imjabbar et al., 2021] Almhdie-Imjabbar, A., Podsiadlo, P., Ljuhar, R., Jennane, R., Nguyen, K.-L., Toumi, H., Saarakkala, S., and Lespessailles, E. (2021). Trabecular bone texture analysis of conventional radiographs in the assessment of knee osteoarthritis : review and viewpoint. *Arthritis Research & Therapy*, 23 :1–13.
- [Altman and Gold, 2007] Altman, R. D. and Gold, G. (2007). Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis and cartilage*, 15 :A1–A56.
- [Andrearczyk and Whelan, 2016] Andrearczyk, V. and Whelan, P. F. (2016). Using filter banks in convolutional neural networks for texture classification. *Pattern Recognition Letters*, 84 :63–69.
- [Anifah et al., 2011] Anifah, L., Purnama, I. K. E., Hariadi, M., and Purnomo, M. H. (2011). Automatic segmentation of impaired joint space area for osteoarthritis knee on x-ray image using gabor filter based morphology process. *IPTEK The Journal for Technology and Science*, 22(3).
- [Antony et al., 2017] Antony, J., McGuinness, K., Moran, K., and O’Connor, N. E. (2017). Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks. In *International conference on machine learning and data mining in pattern recognition*, pages 376–390. Springer.
- [Arden and Nevitt, 2006] Arden, N. and Nevitt, M. C. (2006). Osteoarthritis : epidemiology. *Best practice & research Clinical rheumatology*, 20(1) :3–25.
- [Bahdanau et al., 2015] Bahdanau, D., Cho, K., and Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. In *International Conference on Learning Representations (ICLR)*.
- [Bayramoglu et al., 2020] Bayramoglu, N., Nieminen, M. T., and Saarakkala, S. (2020). A lightweight cnn and joint shape-joint space () descriptor for radiological osteoarthritis detection. In *Annual Conference on Medical Image Understanding and Analysis*, pages 331–345. Springer.
- [Bourlard and Kamp, 1988] Bourlard, H. and Kamp, Y. (1988). Auto-association by multilayer perceptrons and singular value decomposition. *Biological cybernetics*, 59(4) :291–294.

- [Brahim et al., 2019] Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H., and Lesspessailles, E. (2019). A decision support tool for early detection of knee osteoarthritis using x-ray imaging and machine learning : Data from the osteoarthritis initiative. *Computerized Medical Imaging and Graphics*, 73 :11–18.
- [Chan et al., 2023] Chan, K. Y., Abu-Salih, B., Qaddoura, R., Ala’M, A.-Z., Palade, V., Pham, D.-S., Del Ser, J., and Muhammad, K. (2023). Deep neural networks in the cloud : Review, applications, challenges and research directions. *Neurocomputing*, 545 :126327.
- [Charikar et al., 2002] Charikar, M., Chen, K., and Farach-Colton, M. (2002). Finding frequent items in data streams. In *International Colloquium on Automata, Languages, and Programming*, pages 693–703. Springer.
- [Chen et al., 2019] Chen, P., Gao, L., Shi, X., Allen, K., and Yang, L. (2019). Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. *Computerized Medical Imaging and Graphics*, 75 :84–92.
- [Chiang et al., 2019] Chiang, W.-L., WANG, H., YAN, L., and JIANG, S. (2019). Cluster-gen : An efficient algorithm for training deep and large graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*.
- [Cho et al., 2014] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734.
- [Chopra et al., 2005] Chopra, S., Hadsell, R., and LeCun, Y. (2005). Learning a similarity metric discriminatively, with application to face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 539–546. IEEE.
- [Chu et al., 2021] Chu, X., Wei, Z., and Liu, Y. (2021). Hybrid vit : Vision transformer with hybrid cnn backbone for visual recognition. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [Cibrián Uhalte et al., 2017] Cibrián Uhalte, E., Wilkinson, J. M., Southam, L., and Zeggini, E. (2017). Pathways to understanding the genomic aetiology of osteoarthritis. *Human molecular genetics*, 26(R2) :R193–R201.
- [Cimpoi et al., 2015] Cimpoi, M., Maji, S., and Vedaldi, A. (2015). Deep filter banks for texture recognition and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3828–3836.
- [Cross et al., 2014] Cross, M., Smith, E., Hoy, D., Nolte, S., Ackerman, I., Fransen, M., Bridgett, L., Williams, S., Guillemin, F., Hill, C. L., et al. (2014). The global burden of hip and knee osteoarthritis :

- estimates from the global burden of disease 2010 study. *Annals of the rheumatic diseases*, 73(7) :1323–1330.
- [Dacre et al., 1989] Dacre, J., Coppock, J., Herbert, K., Perrett, D., and Huskisson, E. (1989). Development of a new radiographic scoring system using digital image analysis. *Annals of the rheumatic diseases*, 48(3) :194–200.
- [Dosovitskiy et al., 2020] Dosovitskiy, A., Springenberg, J. T., Riedmiller, M., and Schultz, T. (2020). Discriminative unsupervised feature learning with exemplar convolutional neural networks. volume 38, pages 1734–1747. IEEE.
- [Duryea et al., 2000] Duryea, J., Li, J., Peterfy, C., Gordon, C., and Genant, H. (2000). Trainable rule-based algorithm for the measurement of joint space width in digital radiographic images of the knee. *Medical physics*, 27(3) :580–591.
- [Feehan et al., 1982] Feehan, J., Macfarlane, C., and Vaughan, B. (1982). Conservative management of a traumatic meniscal injury. *Sports Med*, 10(2) :90–5.
- [Finnilä et al., 2017] Finnilä, M. A., Thevenot, J., Aho, O.-M., Tiitu, V., Rautiainen, J., Kauppinen, S., Nieminen, M. T., Pritzker, K., Valkealahti, M., Lehenkari, P., et al. (2017). Association between subchondral bone structure and osteoarthritis histopathological grade. *Journal of Orthopaedic Research*, 35(4) :785–792.
- [Gan and Sayuti, 2017] Gan, H. S. and Sayuti, K. A. (2017). Comparison of improved semi-automated segmentation technique with manual segmentation : data from the osteoarthritis initiative.
- [Gao et al., 2016] Gao, Y., Beijbom, O., Zhang, N., and Darrell, T. (2016). Compact bilinear pooling. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 317–326.
- [Gornale et al., 2016] Gornale, S. S., Patravali, P. U., and Manza, R. R. (2016). Detection of osteoarthritis using knee x-ray image analyses : a machine vision based approach. *Int. J. Comput. Appl*, 145(1) :20–26.
- [Gu et al., 2022] Gu, H., Li, K., Colglazier, R. J., Yang, J., Lebhar, M., O’Donnell, J., Jiranek, W. A., Mather, R. C., French, R. J., Said, N., et al. (2022). Knee arthritis severity measurement using deep learning : a publicly available algorithm with a multi-institutional validation showing radiologist-level performance. *arXiv preprint arXiv :2203.08914*.
- [Gu et al., 2019] Gu, H., Wang, Y., Hong, S., and Gui, G. (2019). Blind channel identification aided generalized automatic modulation recognition based on deep learning. *IEEE Access*, 7 :110722–110729.
- [Gubern-Mérida et al., 2012] Gubern-Mérida, A., Kallenberg, M., Martí, R., and Karssemeijer, N. (2012). Segmentation of the pectoral muscle in breast mri using atlas-based approaches. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012 : 15th International Conference, Nice, France, October 1-5, 2012, Proceedings, Part II 15*, pages 371–378. Springer.

- [Guillemin et al., 2011] Guillemin, F., Rat, A., Mazieres, B., Pouchot, J., Fautrel, B., Euler-Ziegler, L., Fardellone, P., Morvan, J., Roux, C., Verrouil, E., et al. (2011). Prevalence of symptomatic hip and knee osteoarthritis : a two-phase population-based survey. *Osteoarthritis and cartilage*, 19(11) :1314–1322.
- [Hamilton et al., 2017] Hamilton, W. L., Ying, Z., and Leskovec, J. (2017). Inductive representation learning on large graphs. In *Neural Information Processing Systems (NeurIPS)*.
- [Haverkamp et al., 2011] Haverkamp, D. J., Schiphof, D., Bierma-Zeinstra, S. M., Weinans, H., and Waarsing, J. H. (2011). Variation in joint shape of osteoarthritic knees. *Arthritis & Rheumatism*, 63(11) :3401–3407.
- [Hayashi et al., 2016] Hayashi, D., Guermazi, A., and Roemer, F. W. (2016). Radiography and computed tomography imaging of osteoarthritis. *Oxford Textbook of Osteoarthritis and Crystal Arthropathy*, page 157.
- [He et al., 2016] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- [He et al., 2020] He, R., Huang, J., Lin, X., Zhao, L., and Wu, L. (2020). Lightgcn : Simplifying and powering graph convolution network. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*.
- [Hinton and Salakhutdinov, 2006] Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786) :504–507.
- [Hirvasniemi et al., 2019] Hirvasniemi, J., Niinimäki, J., Thevenot, J., and Saarakkala, S. (2019). Bone density and texture from minimally post-processed knee radiographs in subjects with knee osteoarthritis. *Annals of biomedical engineering*, 47 :1181–1190.
- [Hirvasniemi et al., 2014] Hirvasniemi, J., Thevenot, J., Immonen, V., Liikavainio, T., Pulkkinen, P., Jämsä, T., Arokoski, J., and Saarakkala, S. (2014). Quantification of differences in bone texture from plain radiographs in knees with and without osteoarthritis. *Osteoarthritis and cartilage*, 22(10) :1724–1731.
- [Hladůvka et al., 2017] Hladůvka, J., Phuong, B. T. M., Ljuhar, R., Ljuhar, D., Rodrigues, A. M., Branco, J. C., and Canhão, H. (2017). Femoral rois and entropy for texture-based detection of osteoarthritis from high-resolution knee radiographs. *arXiv preprint arXiv :1703.09296*.
- [Hochreiter and Schmidhuber, 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. volume 9, pages 1735–1780. MIT Press.
- [Huang et al., 2017] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.

- [Hunter and Felson,] Hunter, D. and Felson, D. Osteoarthritis bmj, 332 (7542)(2006). *Crossref View in Scopus*, pages 639–642.
- [Jahan et al., 2024] Jahan, M. M. N. Z., Hasan, M. Z., Samia, I. J., Fatema, K., Rony, M. A. H., Arefin, M. S., and Moustafa, A. A. (2024). Koa-cctnet : An enhanced knee osteoarthritis grade assessment framework using modified compact convolutional transformer model. *IEEE Access*, 12 :107719–107741.
- [Janvier et al., 2017] Janvier, T., Jennane, R., Toumi, H., and Lespessailles, E. (2017). Subchondral tibial bone texture predicts the incidence of radiographic knee osteoarthritis : data from the osteoarthritis initiative. *Osteoarthritis and cartilage*, 25(12) :2047–2054.
- [Kellgren and Lawrence, 1957] Kellgren, J. H. and Lawrence, J. (1957). Radiological assessment of osteoarthritis. *Annals of the rheumatic diseases*, 16(4) :494.
- [Kingma, 2013] Kingma, D. P. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv :1312.6114*.
- [Kipf and Welling, 2017] Kipf, T. and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*.
- [Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105.
- [LeCun et al., 1998] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324.
- [Ledingham et al., 1993] Ledingham, J., Regan, M., Jones, A., and Doherty, M. (1993). Radiographic patterns and associations of osteoarthritis of the knee in patients referred to hospital. *Annals of the rheumatic diseases*, 52(7) :520–526.
- [Lee et al., 2021] Lee, L. S., Chan, P. K., Fung, W. C., Chan, V. W. K., Yan, C. H., and Chiu, K. Y. (2021). Imaging of knee osteoarthritis : A review of current evidence and clinical guidelines. *Musculoskeletal care*, 19(3) :363–374.
- [Lim and Lau, 2011] Lim, K. and Lau, C. S. (2011). Perception is everything : Oa is exciting. *International journal of rheumatic diseases*, 14(2) :111–112.
- [Lindner et al., 2013] Lindner, C., Thiagarajah, S., Wilkinson, J. M., Wallis, G. A., Cootes, T. F., arcO-GEN Consortium, et al. (2013). Fully automatic segmentation of the proximal femur using random forest regression voting. *IEEE transactions on medical imaging*, 32(8) :1462–1472.
- [Litjens et al., 2017] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., and Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42 :60–88.

- [Liu et al., 2019] Liu, L., Chen, J., Fieguth, P., Zhao, G., Chellappa, R., and Pietikäinen, M. (2019). From bow to cnn : Two decades of texture representation for texture classification. *International Journal of Computer Vision*, 127(1) :74–109.
- [Liu et al., 2021] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B. (2021). Swin transformer : Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022.
- [Ljuhar et al., 2018] Ljuhar, R., Bertalan, Z., Nehrer, S., Norman, B., Dimai, H. P., Fahrleitner-Pammer, A., and Ljuhar, D. (2018). Fri0521 combining fractal- and entropy-based bone texture analysis for the prediction of osteoarthritis : data from the multicenter osteoarthritis study (most). volume 77, pages 787–787. BMJ Publishing Group Ltd.
- [Lynch et al., 1991] Lynch, J., Hawkes, D., and Buckland-Wright, J. (1991). Analysis of texture in macro-radiographs of osteoarthritic knees, using the fractal signature. *Physics in Medicine & Biology*, 36(6) :709.
- [Marijnissen et al., 2008] Marijnissen, A. C., Vincken, K. L., Vos, P. A., Saris, D., Viergever, M., Bijlsma, J., Bartels, L., and Lafeber, F. (2008). Knee images digital analysis (kida) : a novel method to quantify individual radiographic features of knee osteoarthritis in detail. *Osteoarthritis and cartilage*, 16(2) :234–243.
- [Mehdi et al., 2025] Mehdi, T., Nasser, Y., Aloui, S., and Jennane, R. (2025). Deep learning based approach combining shape and texture features for knee osteoarthritis prediction from x-ray images. *Biomedical Signal Processing and Control*, 101 :107172.
- [Mengko et al., 2005] Mengko, T. L., Wachjudi, R. G., Suksmono, A. B., and Danudirdjo, D. (2005). Automated detection of unimpaired joint space for knee osteoarthritis assessment. In *Proceedings of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry, 2005. HEALTHCOM 2005.*, pages 400–403. IEEE.
- [Minciullo et al., 2017] Minciullo, L., Thomson, J., and Cootes, T. F. (2017). Combination of lateral and pa view radiographs to study development of knee oa and associated pain. In *Medical Imaging 2017 : Computer-Aided Diagnosis*, volume 10134, pages 255–261. SPIE.
- [Nasser et al., 2017] Nasser, Y., El Hassouni, M., Brahim, A., Toumi, H., Lespessailles, E., and Jennane, R. (2017). Diagnosis of osteoporosis disease from bone x-ray images with stacked sparse autoencoder and svm classifier. In *2017 international conference on advanced technologies for signal and image processing (ATSIP)*, pages 1–5. IEEE.
- [Nasser et al., 2024] Nasser, Y., El Hassouni, M., Chetouani, A., and Jennane, R. (2024). Swe-net : Sliced-wasserstein efficientnet for learning discriminative texture features in early diagnosis of knee

- osteoarthritis. In *2024 IEEE Thirteenth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 01–06. IEEE.
- [Nasser et al., 2022] Nasser, Y., Hassouni, M. E., and Jennane, R. (2022). Discriminative deep neural network for predicting knee osteoarthritis in early stage. In *International Workshop on PRedictive Intelligence In MEDicine*, pages 126–136. Springer.
- [Nasser et al., 2020] Nasser, Y., Jennane, R., Chetouani, A., Lespessailles, E., and El Hassouni, M. (2020). Discriminative regularized auto-encoder for early detection of knee osteoarthritis : data from the osteoarthritis initiative. *IEEE transactions on medical imaging*, 39(9) :2976–2984.
- [Ng et al., 2011] Ng, A. et al. (2011). Sparse autoencoder. *CS294A Lecture notes*, 72(2011) :1–19.
- [Ng et al., 2006] Ng, H., Ong, S., Foong, K., Goh, P.-S., and Nowinski, W. (2006). Medical image segmentation using k-means clustering and improved watershed algorithm. In *2006 IEEE southwest symposium on image analysis and interpretation*, pages 61–65. IEEE.
- [Nguyen et al., 2020] Nguyen, H. H., Saarakkala, S., Blaschko, M. B., and Tiulpin, A. (2020). Semixup : in-and out-of-manifold regularization for deep semi-supervised knee osteoarthritis severity grading from plain radiographs. *IEEE Transactions on Medical Imaging*, 39(12) :4346–4356.
- [Norman et al., 2019] Norman, B., Pedoia, V., Noworolski, A., Link, T. M., and Majumdar, S. (2019). Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs. *Journal of digital imaging*, 32(3) :471–477.
- [Oka et al., 2010] Oka, H., Muraki, S., Akune, T., Nakamura, K., Kawaguchi, H., and Yoshimura, N. (2010). Normal and threshold values of radiographic parameters for knee osteoarthritis using a computer-assisted measuring system (koacad) : the road study. *Journal of Orthopaedic Science*, 15 :781–789.
- [Pandey et al., 2015] Pandey, M. S., Rajitha, B., and Agarwal, S. (2015). Computer assisted automated detection of knee osteoarthritis using x-ray images. *Sci. Technol*, 1(2) :74–79.
- [Peterfy et al., 2008] Peterfy, C. G., Schneider, E., and Nevitt, M. (2008). The osteoarthritis initiative : report on the design rationale for the magnetic resonance imaging protocol for the knee. *Osteoarthritis and cartilage*, 16(12) :1433–1441.
- [Pierri, 2023] Pierri, M. (2023). 6. musculoskeletal imaging : Radiology techniques for bones and joints.
- [Qaddoura et al., 2021] Qaddoura, R., M. Al-Zoubi, A., Faris, H., and Almomani, I. (2021). A multi-layer classification approach for intrusion detection in iot networks based on deep learning. *Sensors*, 21(9) :2987.
- [Rayegan et al., 2023] Rayegan, H., Nguyen, H., Weinans, H., Gielis, W., Ahmadi Brooghani, S., Custers, R., van Egmond, N., Lindner, C., and Arbabi, V. (2023). Automated radiographic measurements of knee osteoarthritis. *Cartilage*, 14(4) :413–423.

- [Redmon et al., 2016] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once : Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.
- [Reginster, 1980] Reginster, J. (1980). Long-term effects of glucosamine sulphate on osteo. *Fortschr Med*, 98 :801–6.
- [Ren et al., 2015] Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn : Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*, pages 91–99.
- [Riad et al., 2018] Riad, R., Jennane, R., Brahim, A., Janvier, T., Toumi, H., and Lespessailles, E. (2018). Texture analysis using complex wavelet decomposition for knee osteoarthritis detection : Data from the osteoarthritis initiative. *Computers & Electrical Engineering*, 68 :181–191.
- [Rifai et al., 2011] Rifai, S., Vincent, P., Muller, X., Glorot, X., and Bengio, Y. (2011). Contractive auto-encoders : Explicit invariance during feature extraction. In *Proceedings of the 28th international conference on international conference on machine learning*, pages 833–840.
- [Rumelhart et al., 1986] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by backpropagating errors. In *Nature*, volume 323, pages 533–536. Nature Publishing Group.
- [Sakellariou et al., 2017] Sakellariou, G., Conaghan, P. G., Zhang, W., Bijlsma, J. W., Boyesen, P., D’Agostino, M. A., Doherty, M., Fodor, D., Kloppenburg, M., Miese, F., et al. (2017). Euler recommendations for the use of imaging in the clinical management of peripheral joint osteoarthritis. *Annals of the rheumatic diseases*, 76(9) :1484–1494.
- [Scarselli et al., 2009] Scarselli, F., Gori, M., Tsoi, A. C., Hagen, F., and Monfardini, G. (2009). The graph neural network model. In *IEEE Transactions on Neural Networks*, volume 20, pages 61–80.
- [Schuster and Paliwal, 1997] Schuster, M. and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11) :2673–2681.
- [Segal et al., 2013] Segal, N. A., Nevitt, M. C., Gross, K. D., Hietpas, J., Glass, N. A., Lewis, C. E., and Torner, J. C. (2013). The multicenter osteoarthritis study : opportunities for rehabilitation research. *PM&R*, 5(8) :647–654.
- [Shamir et al., 2010] Shamir, L., Felson, D. T., Ferrucci, L., and Goldberg, I. G. (2010). Assessment of osteoarthritis initiative–kellgren and lawrence scoring projects quality using computer analysis. *Journal of Musculoskeletal Research*, 13(04) :197–201.
- [Shamir et al., 2008] Shamir, L., Ling, S. M., Scott, W. W., Bos, A., Orlov, N., Macura, T. J., Eckley, D. M., Ferrucci, L., and Goldberg, I. G. (2008). Knee x-ray image analysis method for automated detection of osteoarthritis. *IEEE Transactions on Biomedical Engineering*, 56(2) :407–415.

- [Shamir et al., 2009] Shamir, L., Ling, S. M., Scott, W. W., Hochberg, M. C., Ferrucci, L., and Goldberg, I. G. (2009). Early detection of radiographic knee osteoarthritis using computer-aided analysis. *Osteoarthritis and Cartilage*, 17(10) :1307–1312.
- [Shin et al., 2012] Shin, H.-C., Orton, M. R., Collins, D. J., Doran, S. J., and Leach, M. O. (2012). Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4d patient data. *IEEE transactions on pattern analysis and machine intelligence*, 35(8) :1930–1943.
- [Singh et al., 2024] Singh, R., Sharma, N., Upadhyay, D., Devliyal, S., and Singh, A. (2024). A deep learning approach for knee osteoarthritis classification using convolutional neural networks. pages 1–5.
- [Sinha and Veemaraj, 2024] Sinha, V. and Veemaraj, E. (2024). Knee osteoarthritis detection using x-rays and dnn. In *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pages 467–472. IEEE.
- [Springenberg et al., 2014] Springenberg, J. T., Dosovitskiy, A., Brox, T., and Riedmiller, M. (2014). Striving for simplicity : The all convolutional net. *arXiv preprint arXiv :1412.6806*.
- [Stachowiak et al., 2016] Stachowiak, G. W., Wolski, M., Woloszynski, T., and Podsiadlo, P. (2016). Detection and prediction of osteoarthritis in knee and hand joints based on the x-ray image analysis. *Biosurface and Biotribology*, 2(4) :162–172.
- [Sutskever et al., 2014] Sutskever, I., Vinyals, O., and Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 3104–3112.
- [Tan and Le, 2019] Tan, M. and Le, Q. (2019). Efficientnet : Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR.
- [Thomson et al., 2015] Thomson, J., O’Neill, T., Felson, D., and Cootes, T. (2015). Automated shape and texture analysis for detection of osteoarthritis from radiographs of the knee. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 127–134. Springer.
- [Tinhinane et al., 2023] Tinhinane, M., Yassine, N., Ahmad, A., Soraya, A., and Rachid, J. (2023). Effects of region of interest location on osteoarthritis detection using deep feature learning. In *2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6. IEEE.
- [Tiulpin et al., 2018] Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P., and Saarakkala, S. (2018). Automatic knee osteoarthritis diagnosis from plain radiographs : a deep learning-based approach. *Scientific reports*, 8(1) :1–10.
- [Tiulpin et al., 2017] Tiulpin, A., Thevenot, J., Rahtu, E., and Saarakkala, S. (2017). A novel method for automatic localization of joint area on knee plain radiographs. In *Scandinavian Conference on Image Analysis*, pages 290–301. Springer.
- [TO and ALEXA, 1982] TO, C. A. and ALEXA, D. I. D. (1982). Habilitation thesis.

- [Touvron et al., 2021] Touvron, H., Cord, M., Berthouze, D., Chambon, A., Toth, D. G. F., and Lemoine, P. (2021). Training data-efficient image transformers distillation through attention. In *International Conference on Machine Learning (ICML)*.
- [Vaswani et al., 2017] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 30.
- [Veličković et al., 2018] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y. (2018). Graph attention networks. In *International Conference on Learning Representations (ICLR)*.
- [Vincent et al., 2010] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A., and Bottou, L. (2010). Stacked denoising autoencoders : Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(12).
- [Vittitoe et al., 1999] Vittitoe, N. F., Vargas-Voracek, R., and Floyd Jr, C. E. (1999). Markov random field modeling in posteroanterior chest radiograph segmentation. *Medical Physics*, 26(8) :1670–1677.
- [Wang et al., 2022] Wang, C., Liu, H., Sun, Y., Wei, Y., Wang, K., and Wang, B. (2022). Dimension reduction technique based on supervised autoencoder for intrusion detection of industrial control systems. *Security and Communication Networks*, 2022(1) :5713074.
- [Wang et al., 2021] Wang, Y., Wang, X., Gao, T., Du, L., and Liu, W. (2021). An automatic knee osteoarthritis diagnosis method based on deep learning : data from the osteoarthritis initiative. *Journal of Healthcare Engineering*, 2021(1) :5586529.
- [Wang et al., 2024] Wang, Z., Chetouani, A., Jarraya, M., Hans, D., and Jennane, R. (2024). Transformer with selective shuffled position embedding and key-patch exchange strategy for early detection of knee osteoarthritis. *Expert Systems with Applications*, 255 :124614.
- [Woloszynski et al., 2010] Woloszynski, T., Podsiadlo, P., Stachowiak, G., and Kurzynski, M. (2010). A signature dissimilarity measure for trabecular bone texture in knee radiographs. *Medical physics*, 37(5) :2030–2042.
- [Xu et al., 2019] Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2019). How powerful are graph neural networks? In *International Conference on Learning Representations (ICLR)*.
- [Yazid et al., 2012] Yazid, H., Arof, H., and Yazid, H. (2012). 8. automated thresholding in radiographic image for welded joints. *Nondestructive Testing and Evaluation*.
- [Yeoh et al., 2023] Yeoh, P. S. Q., Lai, K. W., Goh, S., Hasikin, K., Wu, X., and Li, P. (2023). Transfer learning-assisted 3d deep learning models for knee osteoarthritis detection : Data from the osteoarthritis initiative. *Frontiers in Bioengineering and Biotechnology*, 11.
- [Ying et al., 2018] Ying, Z., Jin, R., and Leskovec, J. (2018). Hierarchical graph representation learning with differentiable pooling. In *Neural Information Processing Systems (NeurIPS)*.

- [Zeiler and Fergus, 2014] Zeiler, M. D. and Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer.
- [Zhang and Chen, 2004] Zhang, D.-Q. and Chen, S.-C. (2004). A novel kernelized fuzzy c-means algorithm with application in medical image segmentation. *Artificial intelligence in medicine*, 32(1) :37–50.
- [Zhou et al., 2018] Zhou, Z., Zhao, G., Kijowski, R., and Liu, F. (2018). Deep convolutional neural network for segmentation of knee joint anatomy. *Magnetic resonance in medicine*, 80(6) :2759–2770.
- [Çağlar Cengizler and Kabakci, 2023] Çağlar Cengizler and Kabakci, A. G. (2023). 1. a k-nearest neighbors-based classification approach for automated detection of knee osteoarthritis. *Cukurova Medical Journal*.