

Prédiction automatique de la maturité osseuse à partir de
radiographies EOS chez les patients atteints de la scoliose
idiopathique de l'adolescent

par

Hilary Cintia DJUIKOUA WOUAFO

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE
LA MAITRISE EN TECHNOLOGIES DE LA SANTÉ
M. Sc. A.

MONTRÉAL, LE 20 DECEMBRE 2023

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC



Hilary Cintia Djuikoua Wouafo, 2023



Cette licence [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

PRÉSENTATION DU JURY

CE RAPPORT DE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Luc Duong, directeur de mémoire
Département de génie logiciel et des TI à l'École de technologie supérieure

Mme Sylvie Ratté, présidente du jury
Département de génie logiciel et des TI à l'École de technologie supérieure

M. Gnahoua Zoabli, membre du jury
Chef de service du génie biomédical,
Centre intégré de santé et de services sociaux des Laurentides

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 8 DÉCEMBRE 2023

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de recherche, Luc Duong, pour son accompagnement tout au long de mon parcours à la maîtrise. Ses encouragements et conseils m'ont permis de toujours chercher à me dépasser et à gagner en confiance. Dès mon stage, il a su me mettre en confiance sur mes travaux dans un domaine qui m'était totalement nouveau. Monsieur Duong a été d'une grande aide pour mon développement dans mon cheminement de second cycle.

Je remercie Dr Stefan Parent et Dr Hubert Labelle pour leur soutien clinique dans le projet. Je remercie également l'équipe de Sainte-Justine pour leur disponibilité et leur assistance technique pour les bases de données cliniques. Les travaux effectués par Julie Joncas, Marjolaine Roy Beaudry et Soraya Barchi sur les bases de données nous ont permis de valider nos approches avec des données uniques et je les remercie chaudement pour cela. Je tiens à remercier Sarita Sakoto, stagiaire en médecine, pour les annotations du Risser et du Sanders.

Je tiens également à remercier Mark-Andrew Noudel qui a été d'une grande aide pour le traitement des données de radiographie incluant la main des patients. Cette base de données a été d'une grande aide pour la prédiction de l'indice de Sanders. Je remercie mes collègues du laboratoire LIVE avec qui j'ai pu échanger des idées et surtout à Melisa Mateu, avec qui j'ai partagé son poste de travail pour rouler quelques expériences.

Ce travail de recherche a été possible grâce au financement de la Chaire de recherche en sciences du mouvement du CHU Sainte-Justine (CHU mère-enfant) et de l'Université de Montréal.

Des remerciements spéciaux vont aussi à l'égard d'Izabela et BoneXpert pour leur support pour l'utilisation de l'application et la compréhension des rapports fournis par l'application.

Je remercie finalement ma famille et mes amis qui m'ont soutenue lors de tout mon parcours académique.

Prédiction automatique de la maturité osseuse à partir de radiographies EOS chez les patients atteints de la scoliose idiopathique de l'adolescent

Hilary Cintia DJUIKOUA WOUAFO

RÉSUMÉ

La scoliose idiopathique chez les adolescents (SIA) est une pathologie caractérisée par la déformation de la colonne vertébrale. On sait aujourd'hui qu'il existe des facteurs génétiques qui influencent l'évolution de la scoliose idiopathique. L'âge osseux s'avère être un bon indicatif du potentiel de croissance et permet aux cliniciens de prescrire le traitement approprié, soit la chirurgie ou le port d'un corset. Traditionnellement, l'acquisition de l'âge osseux se faisait par l'observation d'une radiographie 2D de la main des patients en suivant les échelles de gradations telles que Tanner-Whitehouse-III (TW3) et Greulich and Pyle (GP). L'échelle de Risser est toutefois la plus utilisée pour le suivi des patients atteints de SIA puisqu'elle permet d'avoir en une radiographie, le degré de la courbure scoliothique et le niveau maturité osseuse des patients. Cette méthode d'évaluation démontre toutefois un haut niveau de variabilité de lecture entre les observateurs. De plus, les évidences de corrélation entre l'échelle de Risser et l'évolution de la courbure scoliothique sont plutôt faibles. Un protocole récemment intégré au CHU Sainte-Justine de Montréal permet d'avoir une radiographie des patients scoliothique qui inclut leur main et permet donc d'observer leur âge osseux suivant les méthodes traditionnelles et l'échelle de Risser. Des outils d'automatisation de l'acquisition de l'âge osseux autant pour les méthodes de gradation traditionnelles que pour les méthodes plus modernes comme l'échelle de Risser ou de Sanders ont récemment étaient développés et ces outils se basent sur des modèles d'apprentissage profond et machine pour la classification de l'âge osseux.

Cette étude a comme but de développer un outil qui évaluera automatiquement la maturité osseuse des patients suivis en scoliose idiopathique de l'adolescent. Pour ce faire, trois modèles ont été développés pour prédire l'âge osseux selon les échelles de Sanders et de Risser à l'aide d'un modèle de régression basé sur les machines à vecteur de support. Le premier modèle a permis de prédire de l'échelle de Risser et de Sanders grâce aux caractéristiques de l'image. Le deuxième modèle consistait en la prédiction de l'échelle de Risser et de Sanders basée sur les caractéristiques cliniques. Et finalement, le dernier modèle consistait en une prédiction de l'échelle de Risser et de Sanders basée sur une combinaison des caractéristiques de l'image et des caractéristiques cliniques.

Le premier modèle VGG16-RVS a démontré une meilleure performance pour la prédiction de l'âge osseux basé sur le Risser avec trois bases de données avec un coefficient de corrélation R^2 de 0,76, une erreur moyenne absolue de 0,17 et une erreur moyenne au carré de 0,20. Les performances du dernier modèle VGG16-BoneXpert-RVS pour la prédiction de l'indice de Sanders étaient meilleures que celle des deux autres modèles avec un coefficient de corrélation R^2 de 0,85, une erreur moyenne absolue de 0,11 et une erreur moyenne au carré de 0,15.

Mots clés: Prédiction de la maturité osseuse, SIA, Rayon-X Apprentissage profond, SVR

Automatic prediction bone maturity from EOS radiographs of adolescent idiopathic scoliosis

Hilary Cintia DJUIKOUA WOUAFO

ABSTRACT

Adolescent idiopathic scoliosis (AIS) is a pathology characterized by spinal deformity. We now know that genetic factors influence the development of idiopathic scoliosis. Bone age is a good indicator of growth potential, enabling clinicians to prescribe the appropriate treatment, either surgery or the use of a corset. Traditionally, bone age was acquired by observing a 2D radiograph of the patient's hand, using grading scales such as Tanner-Whitehouse-III (TW3) and the Sanders index. The Risser scale, however, is the most widely used for monitoring AIS patients, since it provides a radiograph, the degree of scoliotic curvature and the patient's bone age. However, this method of assessment shows a high level of variability in reading between observers. Moreover, there is little evidence of correlation between the Risser scale and the evolution of scoliotic curvature. A recently integrated protocol at Montreal's Sainte-Justine Hospital provides X-rays of scoliotic patients that include their hands, making it possible to observe their bone age using both traditional methods and the Risser scale. Tools for automating bone age acquisition using both traditional grading methods and more modern methods such as the Risser or Sanders scales have recently been developed, based on deep learning and machine learning models for bone age classification.

The aim of this study is to develop a tool that will automatically determine the bone age of patients with adolescent idiopathic scoliosis. To this end, three models were developed to predict bone age according to the Sanders and Risser scales, using a regression model based on support vector machines. The first model was a prediction of the Risser scale based on image characteristics. The second model consisted in the prediction of the Risser and Sanders scales based on clinical characteristics. And finally, the last model consisted of a prediction of the Risser and Sanders scale based on a combination of image features.

The first VGG16-SVR model demonstrated better performance for Risser-based bone age prediction with all the three databases with a maximum R2 correlation coefficient of 0.76, a mean absolute error of 0.17 and a mean squared error of 0.20. The VGG16-BoneXpert-SVR model performance for Sander index prediction was better than that of the other two models, with an R2 correlation coefficient of 0.85, a mean absolute error of 0.11 and a mean squared error of 0.15.

Key words: Bone age prediction, adolescent idiopathic scoliosis, X-ray, Deep learning, SVR

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LA LITTÉRATURE.....	5
1.1 Anatomie et définition de la SIA	5
1.2 Définition de l'âge osseux pour le suivi de la SIA	7
1.2.1 L'indice de Risser pour le suivi des patients scoliotiques	7
1.2.2 Évaluation de l'âge osseux à partir de radiographies de la main et du poignet.....	9
1.2.3 Progression de la maturité squelettique pour la SIA.....	11
1.3 L'implication de l'apprentissage profond dans la caractérisation de l'âge osseux.....	13
1.3.1 Les modèles d'apprentissage profond en imagerie médicale	13
1.3.2 Les modèles d'attention pour l'évaluation des modèles d'apprentissage profond.....	16
1.3.3 Caractérisation automatique de la maturité osseuse	17
1.3.4 Les modèles prédictifs pour la caractérisation de la croissance osseuse.....	19
1.4 Sommaire de la revue de littérature	22
CHAPITRE 2 MÉTHODOLOGIE.....	25
2.1 Description de la base de données	25
2.2 Prétraitement des données	26
2.2.1 Données d'entraînement du modèle d'apprentissage machine basée sur les caractéristiques d'image	26
2.2.2 Données d'entraînement du modèle d'apprentissage machine basée sur les caractéristiques cliniques.....	28
2.3 Les modèles d'apprentissage machine basée sur les caractéristiques d'images.....	30
2.4 Les modèles d'apprentissage machine basée sur les caractéristiques cliniques	33
2.5 Évaluation des modèles	34
CHAPITRE 3 RÉSULTATS ET DISCUSSION.....	37
3.1 Résultats.....	37
3.2 Discussion.....	43
3.2.1 Collecte et traitement de données	43
3.2.2 L'entraînement des modèles de prédiction	44
3.2.3 L'évaluation des modèles de prédiction	44
CONCLUSION	47
RECOMMANDATIONS	49

ANNEXE I Article présenté à la conférence SPIE Medical Imaging	51
ANNEXE II Liste des paramètres extraits de BoneXpert	57
ANNEXE III Rapport de l'évaluation de l'âge osseux de l'outil BoneXpert.....	60
BIBLIOGRAPHIE	62

LISTE DES TABLEAUX

	Page
Tableau 1-1	Approches significative pour la caractérisation de l'âge osseux et l'analyse d'images médicale21
Tableau 2-1	Présentation de la base de données25
Tableau 2-2	Tableau des paramètres des modèles se servant des caractéristiques d'images extraites du modèle d'apprentissage profond.....30
Tableau 2-3	Hyperparamètres pour l'optimisation des modèles de régression par machine à vecteur de support33
Tableau 3-1	Résultats de la recherche grillardée pour l'optimisation des Hyperparamètres des modèles de prédiction37
Tableau 3-2	Résultats des prédictions des différents modèles implémentés a) l'erreur quadratique moyenne RMSE b) l'erreur moyenne absolue MAE c) le coefficient de corrélation (R2)38

LISTE DES FIGURES

	Page
Figure 1.1	Colonne vertébrale en vue postero-antérieure6
Figure 1.2	Les indices de Risser pour le système français et le système américain8
Figure 1.3	a) Évaluation du stade du développement du radius selon TW3 b) Radiographie de la main d'un homme extrait de l'atlas de GP c) Indices de Sanders pour l'évaluation de l'âge osseux Tirée de Maggio (2016)10
Figure 1.4	Les méthodes d'acquisition de l'âge squelettique selon les étapes de la maturité Tirée de Dimeglio et Canavese (2013)11
Figure 1.5	Évolution de la courbure scoliotique selon le niveau de maturité squelettique Tirée de Dimeglio & Canavese (2013).....12
Figure 1.6	a) Principe de l'application de l'apprentissage par transfert, b) L'apprentissage par transfert avec la méthode du fine-tuning, c) L'apprentissage par transfert avec la méthode de l'extraction de caractéristiques Tirée de Sarkar (2018)15
Figure 2.1	Processus de détection automatique de la main de patients a) repérage du squelette b) détection des mains grâce à VotingClassifier c) résultats de la segmentation des mains28
Figure 2.2	Augmentation des données de patient pour l'entraînement du modèle d'apprentissage profond.....32
Figure 2.3	Structure du modèle pour la prédiction de l'âge osseux basée sur les échelles de Risser et l'indice de Sanders avec en orange la structure du modèle VGG16-RVS, en bleu celle du modèle BoneXpert-RVS et en rouge celle du modèle VGG16-BoneXpert-RVS35
Figure 3.1	Nuage de points des observations des indices de Risser observés par rapport aux prédictions pour l'indice de Risser du modèle VGG16-RVS de la base de données ARC 1 et 240
Figure 3.2	Nuage de points des observations des indices de Risser observés par rapport aux prédictions pour l'indice de Risser des modèles de prédiction de la base de données ARC 3 a) modèle BoneXpert-RVS b) modèle VGG16-RVS c) modèle BoneXpert-VGG16-RVS41

Figure 3.3 Nuage de points des observations des indices de Sanders observés par rapport aux prédictions pour l'indice de Sanders des modèles de prédiction de la base de données ARC 3 a) modèle BoneXpert-RVS b) modèle VGG16-RVS c) modèle BoneXpert-VGG16-RVS 42

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

SIA	Scoliose idiopathique chez les adolescents
TW3	Tanner-Whitehouse-III
GP	Greulich and Pyle
RVS	Modèle de régression par machine à vecteur de support
R2	Coefficient de corrélation
RMSE	Erreur quadratique moyenne
MAE	Erreur moyenne absolue
RAG	Réseaux antagonistes génératifs
CNN	Réseau de neurones convolutif
SVM	Support vector machine
SVR	Support Vector Regression
GLCM	Gray-Level Co-Occurrence Matrix
RSNA	Radiological Society of North America
URCO	Unité de Recherche Clinique en Orthopé
EOS	Système d'acquisition simultanée de deux images radiographiques, face et profil, commercialisé par la compagnie française EOS imaging

INTRODUCTION

La scoliose idiopathique de l'adolescent (SIA) se caractérise par une déformation en 3D de la colonne vertébrale. On sait aujourd'hui qu'il existe des facteurs génétiques qui influencent l'évolution de la scoliose idiopathique. Mais aucun de ces facteurs n'entre réellement en compte lors du suivi des patients ; le suivi dépend plutôt de l'amplitude de la courbure et du potentiel de croissance des patients pour l'identification du risque d'évolution de la courbure (Acheson, 1954) qui sert ensuite de guide pour décider du type de traitement adapté à chaque patient (Zhang et al., 2020).

Le suivi des patients se fait généralement par l'acquisition d'une radiographie par rayon X. Puisque l'observation de la colonne vertébrale est nécessaire, les radiographies des patients ciblent la région supérieure de leur corps et une acquisition frontale et latérale permet aux cliniciens de cibler les facteurs de croissance chez les patients et faire un suivi de l'évolution de leur courbure scoliothique. La radiographie conventionnelle est utilisée de routine pour la visualisation de la colonne vertébrale. Le système d'imagerie EOS (*Voir* Figure 0.1) permettant d'avoir une radiographie frontale et latérale en simultané est souvent utilisé pour l'acquisition de radiographie pour le suivi des patients scoliothiques.

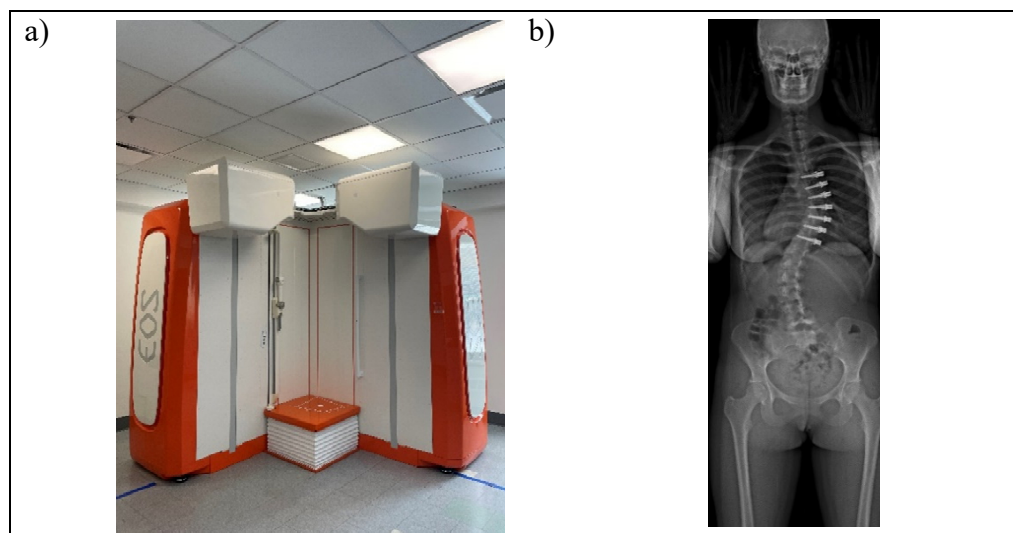


Figure 0.1 a) Équipement d'imagerie pour radiographie EOS du CHU Sainte-Justine b) Radiographie frontale de patients scoliothiques

Les méthodes d'évaluation de la scoliose reposent sur l'observation des caractéristiques de maturité osseuse des patients. Pour les patients scoliotiques, la maturité osseuse et l'angle initial de la courbure sont de bons indicateurs pour un bon suivi. Lors du suivi, des mesures cliniques telles que la taille, le poids et l'alignement du bassin sont enregistrées dans le dossier clinique afin de documenter leur croissance. L'indice de Risser est utilisé de routine en clinique pour l'évaluation de la maturité osseuse. Cet indice repose sur l'observation de l'ossification de la crête iliaque, associée à la maturité squelettique des patients scoliotiques. Cette méthode d'évaluation permettrait d'associer la maturité squelettique à l'évolution de la courbure scoliotique entre 10 et 15 ans (Hacquebord & Leopold, 2012). Le système de gradation de Risser a deux variantes : le système français et le système américain avec tous les deux 6 stades distincts de maturité.

La maturité osseuse peut être utilisée de pair avec l'observation de la courbure scoliotique pour décider du traitement approprié. Il existe particulièrement deux moments clés dans l'évolution de la maturité osseuse qui permettent de prendre une décision clinique. Pendant la poussée de croissance, la courbure scoliotique évolue de façon plus prononcée et l'identification de cette période permet donc au clinicien de déterminer la période pendant laquelle le corset serait efficace. Dès que le patient atteint sa taille maximale et selon l'angle initial de la courbure scoliotique, il est possible pour le clinicien de décider si la chirurgie sera efficace (Yin Cheung & Luk, 2017, Dimeglio & Canavese, 2013). Avant l'atteinte de la taille maximale, le patient est considéré comme ayant un Risser 0 et sa maturité squelettique est estimée en déterminant son âge osseux à l'aide des méthodes Tanner-Whitehouse-III (TW3), Greulich and Pyle (GP) ou alors Sanders. Ce n'est donc qu'après la fermeture du cartilage triradié que l'ossification de la crête iliaque commence.

L'assignation de l'âge osseux pour les patients scoliotiques basé sur l'indice de Risser comporte une importante variabilité en raison de la différence qu'on retrouve dans la qualité des images, les systèmes d'acquisition et la visibilité limitée de la crête iliaque due aux déformations frontales chez les patients scoliotiques. Ce mémoire a pour objectif d'étudier un modèle de régression qui sera en mesure de prédire l'âge osseux de patient scoliotique selon

les indices de Risser et de Sanders en se basant sur les caractéristiques cliniques des patients et les caractéristiques d'images extraites de radiographies grâce à un modèle de convolution. Les méthodes mises en place seront évaluées en fonction de l'erreur quadratique moyenne (RMSE), l'erreur moyenne absolue (MAE) et le coefficient de corrélation (R^2).

Ce mémoire est composé de trois chapitres. Le premier chapitre porte sur la revue de littérature introduisant la pathologie de la scoliose idiopathique chez les adolescents, les modèles d'apprentissage profond et machine ainsi que leur utilité pour la caractérisation de l'âge osseux. La méthodologie utilisée pour la réalisation des expériences est présentée dans le chapitre 2. Et les différents résultats d'expérimentation sont présentés dans le chapitre 3. La discussion des différents résultats obtenus est développée dans le chapitre 3. La conclusion du mémoire et les recommandations sur les futurs travaux sont présentées dans le chapitre 4 de ce mémoire.

CHAPITRE 1

REVUE DE LA LITTÉRATURE

Dans ce chapitre, une revue de la littérature de l'état de l'art est présentée. La première partie de cette revue porte sur une présentation de la scoliose idiopathique chez les adolescents. Une deuxième partie aborde le sujet de l'âge osseux pour le suivi des patients atteints de SIA. La troisième partie présente les modèles d'apprentissage profond utilisés dans la caractérisation de l'âge osseux. Finalement une présentation des différents modèles de l'état de l'art pour la prédiction de l'âge osseux est faite.

1.1 Anatomie et définition de la SIA

Les vertèbres du rachis du corps humain se divisent en quatre parties. Les sept premières vertèbres constituent le rachis cervical. Les douze vertèbres suivantes constituent le rachis dorsal. Le rachis lombaire et le sacrum se partagent également les dix dernières vertèbres. Les déformations rachidiennes surviennent lorsque l'alignement naturel de la colonne vertébrale est modifié. Les premiers signes des déformations se font généralement voir durant l'enfance et leur apparence s'aggrave pendant la période de croissance de l'adolescence. Les déformations rachidiennes portent le nom de scoliose idiopathique de l'adolescent (SIA) lorsque la cause de leur apparition n'est pas reliée à un processus pathologique. Ce type de déformation est généralement localisée entre le rachis cervical et le rachis lombaire (*Voir Figure 1.1*).

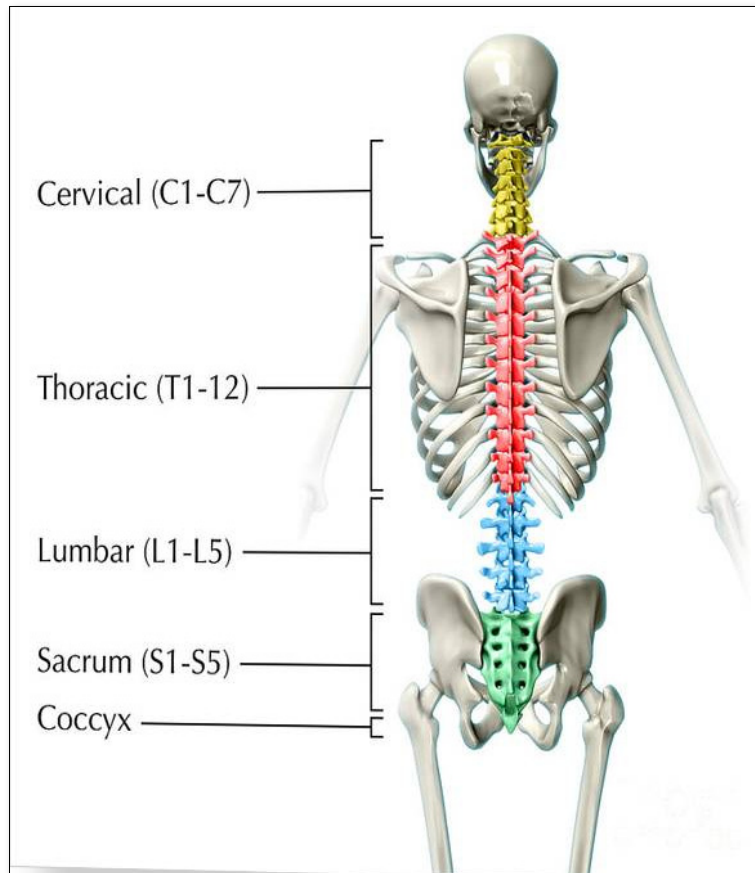


Figure 1.1 Colonne vertébrale en vue postero-antérieure

On sait aujourd'hui qu'il existe des facteurs génétiques qui influencent l'évolution de la SIA. Mais aucun de ces facteurs n'entre réellement en compte lors du suivi des patients ; le suivi dépend plutôt de l'amplitude de la courbure et du potentiel de croissance des patients pour identifier le risque d'évolution de la courbure (Acheson, 1954). L'identification du risque d'évolution de la courbure scoliootique sert ensuite de guide pour décider du type de traitement adapté à chaque patient (Zhang et al., 2020).

1.2 Définition de l'âge osseux pour le suivi de la SIA

Les méthodes d'évaluation de la scoliose reposent sur l'observation des caractéristiques de maturité osseuse des patients. Lors des suivis des patients, des mesures cliniques telles que la taille, le poids et l'alignement du bassin sont enregistrées dans leur dossier clinique afin de documenter leur croissance. L'utilisation des radiographies de la main et de la colonne vertébrale permet de suivre le progrès de la maturité des patients et de la courbure scoliotique à l'aide d'indices tels que l'indice de Risser, TW-III, GP et l'indice de Sanders.

1.2.1 L'indice de Risser pour le suivi des patients scoliotiques

L'acquisition de l'âge osseux est une interprétation de la maturité squelettique et l'âge osseux correspond donc typiquement à l'observation du niveau d'ossification de certaines régions telle que le poignet ou le genou. En 1958, Joseph C. Risser propose un indice basé sur l'observation de l'ossification de la crête iliaque, associé à la maturité squelettique des patients scoliotiques. Cette nouvelle méthode d'évaluation permettrait d'associer la maturité squelettique à l'évolution de la courbure scoliotique entre 10 et 15 ans (Hacquebord & Leopold, 2012). Le système de gradation de Risser a deux variantes : une pour le système français et une autre pour le système américain avec tous les deux 6 stades distincts de maturité. Dans le système français, au stade de Risser 0 il n'y a pas d'ossification de l'apophyse et les degrés d'ossification de la crête iliaque sont représentés par les Risser 1, 2, et 3. C'est au Risser 4 que la fusion de l'apophyse commence et elle est complète au Risser 5. Pour le système américain, le Risser 0 représente une fois de plus une absence d'ossification, les Risser 1, 2, 3 et 4 représentent les différents degrés d'ossification de la crête iliaque et la fusion de l'apophyse correspond au Risser 5 (*Voir* Figure 1.2).

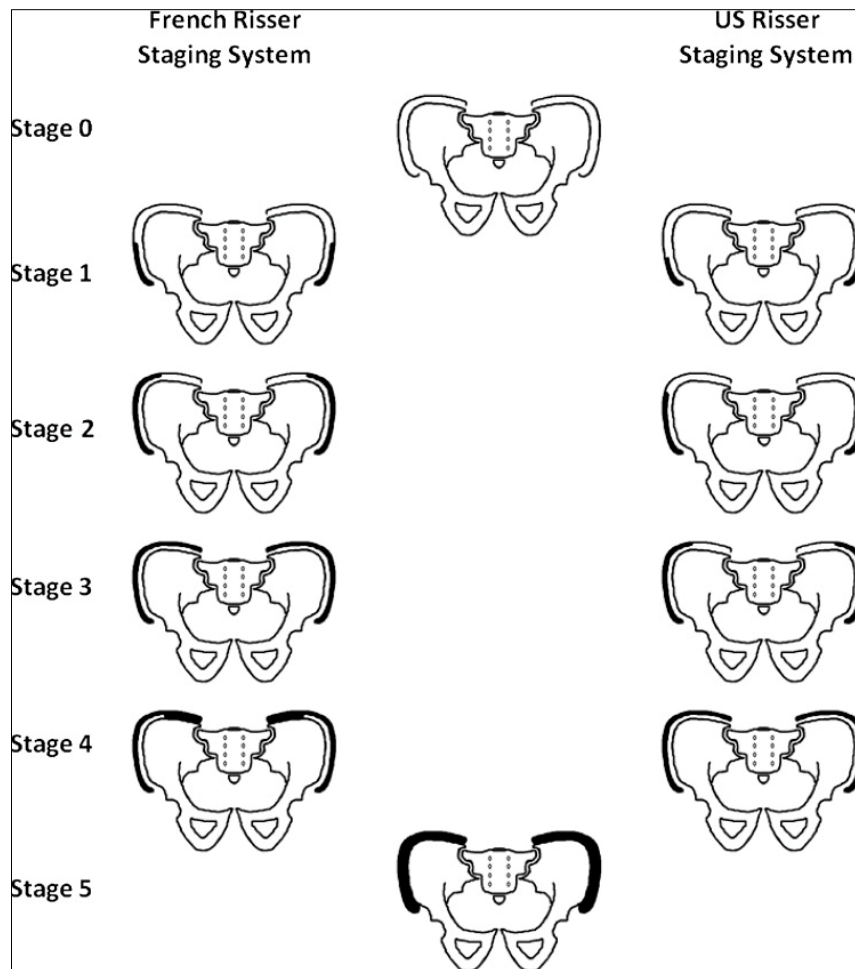


Figure 1.2 Les indices de Risser pour le système français et le système américain

L'utilisation de l'échelle de Risser pour le suivi de la maturité osseuse des patients scoliotiques a pour avantage d'éviter aux cliniciens le besoin d'avoir recours à une seconde radiographie de la main, qui nécessiterait une radiographie supplémentaire. En effet, pour l'observation de la maturité osseuse basée sur l'échelle de Risser, le clinicien n'a besoin que de la radiographie antérieure des patients qui est déjà utilisée pour la courbure scoliotique.

1.2.2 Évaluation de l'âge osseux à partir de radiographies de la main et du poignet

TW3 et GP sont les méthodes d'acquisition de l'âge osseux les plus classiques et sont souvent utilisées pour caractériser l'âge osseux entre 0 et 18 ans. Ces classifications sont obtenues en observant des radiographies de la main ou du poignet (*Voir* Figure 1.3-a et b). La classification de TW3 se fait en attribuant un score de maturité à des zones spécifiques telles que le radius et les différentes phalanges de la main. La classification de GP est une simplification de la méthode de TW3 par le fait qu'elle utilise les changements morphologiques d'os spécifiques de la main pour déterminer les différents niveaux de maturité osseuse. Cette classification fournit également un atlas de radiographies contrairement au TW3. En orthopédie, la plus récente et la plus utilisée des méthodes d'acquisition de l'âge osseux à partir de radiographies de la main sont la méthode d'évaluation selon l'échelle du Sanders.

La méthode de Sanders est aussi une simplification de la méthode TW3, qui nécessite d'avoir de longues observations qui sont moins convenables pour le milieu clinique et qui cause une variabilité trop importante de lecture entre les observateurs (*Voir* Figure 1.3-c) (Sanders et al., 2008). En effet l'acquisition de l'âge osseux par TW3 consiste en l'observation de 20 régions d'intérêts distinctes qui sont ensuite associés à un score et la somme des scores permet donc d'obtenir l'âge osseux qui diffère pour les garçons et les filles (Prokop-Piotrkowska, Marszałek-Dziuba, Moszczyńska, Szalecki, & Jurkiewicz, 2021). L'indice de Sanders comprend huit stades distincts de la croissance osseuse et donne aussi des indications sur le potentiel restant de croissance des patients. La méthode de caractérisation par le radius et l'ulna qui diffère du Sanders offre également une bonne indication quant au niveau de croissance squelettique avant l'atteinte de la maturité osseuse. Elle a l'avantage, comme l'indice de Sanders, d'être plutôt simple d'utilisation (Cheung, Cheung, & Luk, 2019).

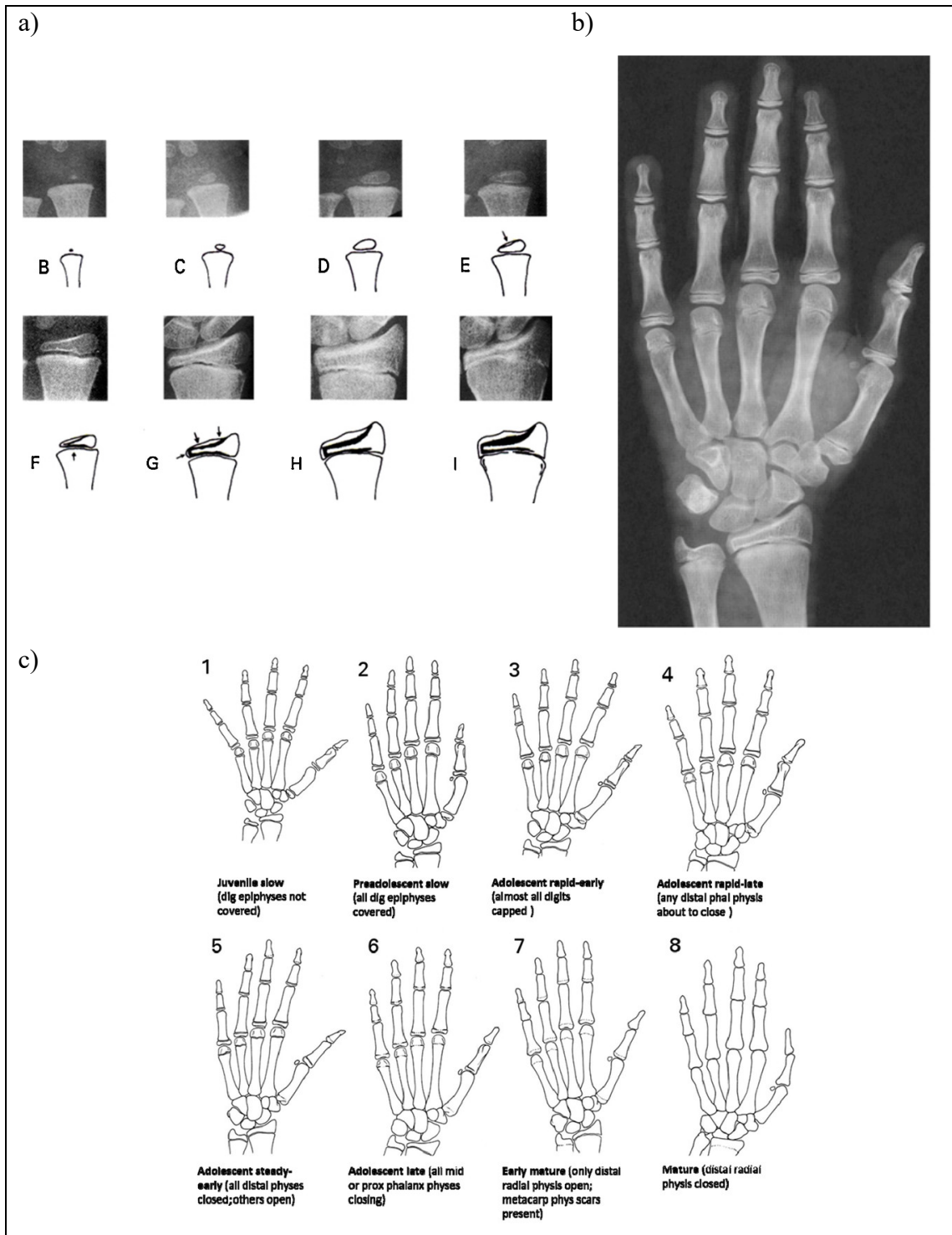


Figure 1.3 a) Évaluation du stade du développement du radius selon TW3

b) Radiographie de la main d'un homme extrait de l'atlas de GP

c) Indices de Sanders pour l'évaluation de l'âge osseux

Tirée de Maggio (2016)

1.2.3 Progression de la maturité squelettique pour la SIA

Pour les patients scoliotiques, la maturité osseuse et l'angle de Cobb initial de la courbure sont de bons indicateurs pour le suivi de la progression. La maturité osseuse peut être utilisée de pair avec l'observation de la courbure scoliotique pour décider du traitement approprié. Il existe particulièrement deux moments clés dans l'évolution de la maturité osseuse qui permettent de prendre une décision clinique. Pendant la poussée de croissance, la courbure scoliotique évolue de façon plus prononcée et l'identification de cette période permet au clinicien de déterminer la période pendant laquelle le corset serait efficace. Dès que le patient atteint sa taille maximale et selon l'angle initial de la courbure scoliotique, il est possible pour le clinicien de décider si la chirurgie sera efficace (Yin Cheung & Luk, 2017, Dimeglio & Canavese, 2013) (Voir Figure 1.5). Avant l'atteinte de la taille maximale, le patient est considéré comme ayant un Risser 0 et sa maturité squelettique est estimée en déterminant son âge osseux à l'aide des méthodes TW3, GP ou Sanders. Ce n'est donc qu'après la fermeture du cartilage triradié que l'ossification de la crête iliaque commence (Voir Figure 1.4).

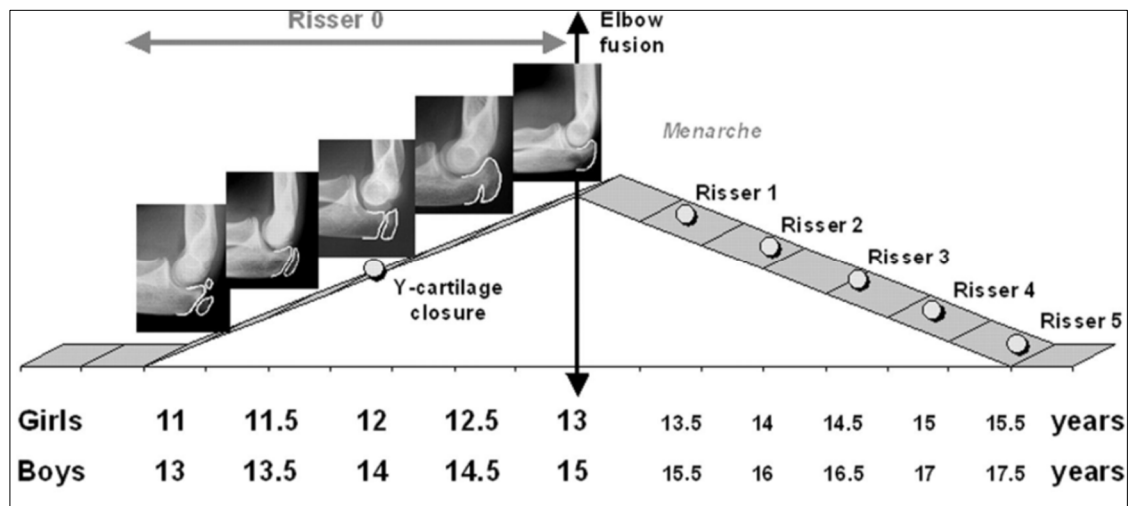


Figure 1.4 Les méthodes d'acquisition de l'âge squelettique selon les étapes de la maturité

Tirée de Dimeglio et Canavese (2013)

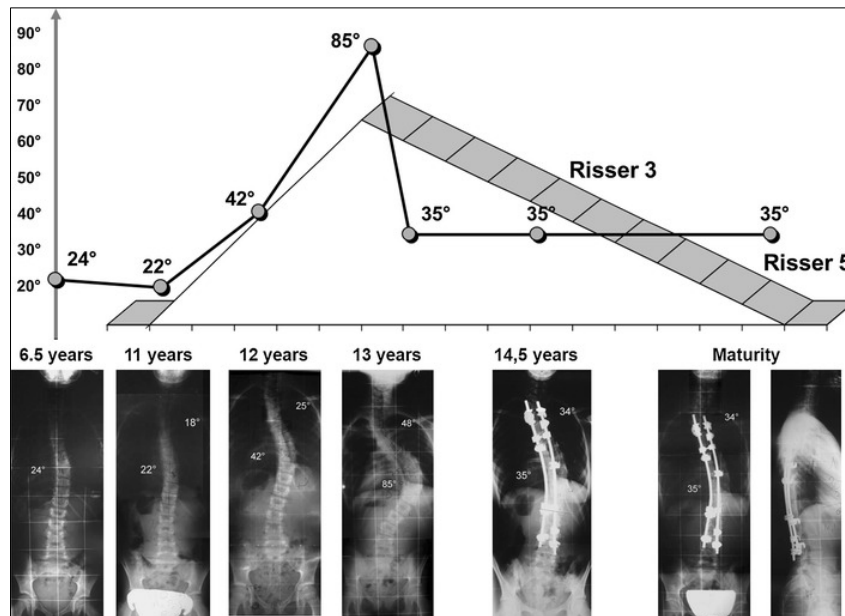


Figure 1.5 Évolution de la courbure scoliothique selon le niveau de maturité squelettique
Tirée de Dimeglio & Canavese (2013)

Selon Dimeglio & Canavese, quoique le risque de progression de la courbure scoliothique soit plus bas après l'atteinte de la taille maximale, il reste un certain risque de progression qui peut être surveillé en observant la maturité osseuse avec les méthodes telles que le Sanders et le Risser (Dimeglio & Canavese, 2013). Récemment, Minkara et al. ont mené une étude démontrant que l'indice de Sanders aurait une meilleure corrélation avec l'évolution de la courbure scoliothique que l'échelle de Risser (Minkara et al., 2020). Ils ont conclu que l'indice de Risser comme méthode d'acquisition de l'âge osseux durant le pic de croissance serait moins approprié puisqu'il présente un risque de décalage et une grande variabilité inter-observateur. L'indice de Sanders serait donc plus approprié pour le suivi des patients atteints de scoliose idiopathique adolescente.

1.3 L'implication de l'apprentissage profond dans la caractérisation de l'âge osseux

La caractérisation manuelle de l'âge osseux est une tâche répétitive, et pour des observateurs non experts, elle peut introduire une grande variabilité intra-observateur et inter-observateur. Des approches automatiques ont été proposées pour l'inférence automatique de l'âge osseux. Dans les dernières années, l'apprentissage profond a connu un essor important pour l'analyse automatique de l'image. L'apprentissage profond est une branche de l'intelligence artificielle où la tâche de décision est modélisée comme un ensemble de neurones qui apprendront de façon automatique à faire des tâches de segmentation, de classification et de recalage à partir d'indications d'apprentissage leur étant fournies par l'humain (apprentissage supervisé). En imagerie médicale, l'utilisation de l'apprentissage profond permet d'effectuer des tâches aidant au diagnostic et au traitement de nombreuses pathologies. Dans cette section, nous présenterons les différentes méthodes d'apprentissage profond utilisées pour le diagnostic de la scoliose idiopathique chez les adolescents.

1.3.1 Les modèles d'apprentissage profond en imagerie médicale

Afin d'utiliser un modèle d'apprentissage profond, il est nécessaire dans un premier temps d'entraîner le modèle à réaliser la tâche voulue et pour ce faire, il est nécessaire d'avoir accès à une base de données assez grande et variée pour permettre au modèle d'avoir assez de cas différents. Il existe plusieurs types de modèles d'apprentissage profond pré-entraînés et ceux-ci sont majoritairement entraînés avec des bases de données d'images de scènes naturelles. Ces modèles pré-entraînés sont alors adaptés à partir d'un ensemble représentatif du problème. En effet, pour que les algorithmes soient performants, il est nécessaire non seulement d'avoir un bon étiquetage des données, une base de données comprenant des cas anormaux, mais aussi d'avoir une grande base de données représentant adéquatement la population cible (Chan, Samala, Hadjiiski, & Zhou, 2020). Pour pallier le manque de données, il est commun d'utiliser des méthodes d'augmentation de données ayant pour but d'introduire des transformations rigides aléatoires aux images. Il existe également des modèles permettant d'introduire des

transformations de caractéristiques aux images d'entrée. Ces méthodes exploitent les réseaux antagonistes génératifs (RAG) pour générer des images synthétiques (Goodfellow et al., 2020).

L'exploitation de l'apprentissage profond en imagerie médicale se fait de diverses façons, mais récemment, la plupart des recherches dans le domaine se sont penchées sur l'utilisation de l'apprentissage par transfert (*Voir* Figure 1.6-a). Cette méthode consiste à utiliser des algorithmes entraînés sur des images de scènes naturelles pour réaliser des tâches de classification, de segmentation et de recalage dans un cadre médical par exemple. Les modèles pré-entraînés sont repris avec les connaissances qu'ils ont acquises dans un domaine précédent, ils reçoivent en entrée des données d'un autre domaine et les diverses couches du modèle vont extraire des informations génériques et d'autres, plus spécifiques, des images médicales pour en faire une analyse (Chan et al., 2020).

L'application de l'apprentissage par transfert peut être faite de deux façons distinctes. La première consiste à refaire un entraînement de toutes les couches du modèle en modifiant aléatoirement ses paramètres pour arriver à avoir un modèle adapté à un certain domaine (*Voir* Figure 1.6-b). La deuxième consiste à n'entraîner que les couches superficielles du modèle pour ensuite extraire les caractéristiques apprises par ces couches (*Voir* Figure 1.6-c). Ces caractéristiques sont donc ensuite données à un modèle d'apprentissage machine qui fera la classification des données en entrées selon les caractéristiques qui lui seront données. Lors de l'exécution de la seconde méthode, il est possible de choisir les caractéristiques les plus importantes en utilisant des outils d'évaluation de caractéristiques. Pan et al. ont mis en application cette technique qui a permis que le modèle d'apprentissage machine ait une meilleure précision de classification (Pan, Pouyanfar, Chen, Qin, & Chen, 2017).

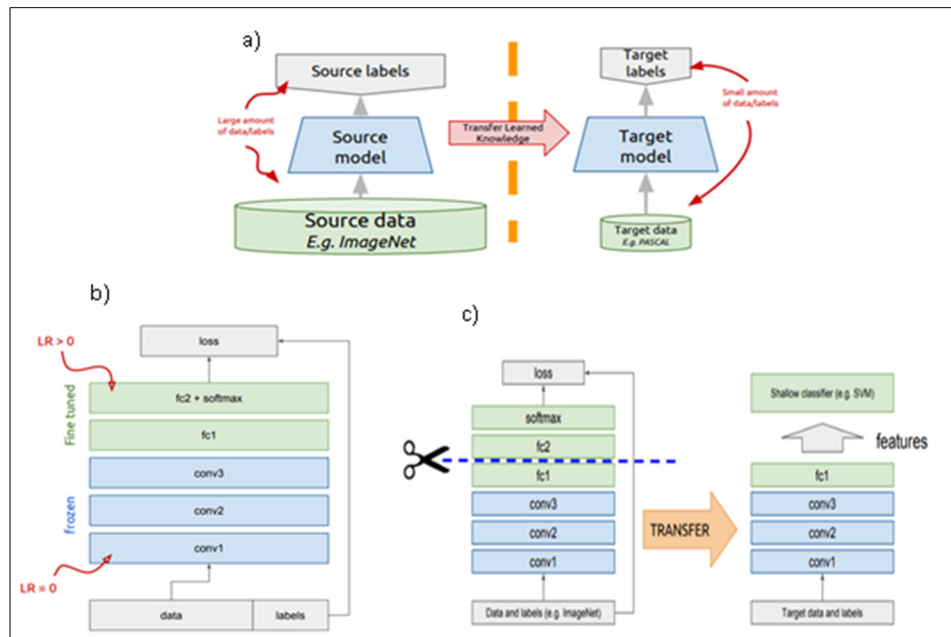


Figure 1.6 a) Principe de l'application de l'apprentissage par transfert, b) L'apprentissage par transfert avec la méthode du fine-tuning, c) L'apprentissage par transfert avec la méthode de l'extraction de caractéristiques
Tirée de Sarkar (2018)

Plusieurs études telles que Mei et al. ont récemment été faites dans le but de trouver un moyen d'entraîner des modèles d'apprentissage profond avec des bases de données d'image médicale nommé RadImageNet. Cette étude a démontré que l'utilisation de modèles entraînés sur des images médicales permettrait d'obtenir de meilleurs résultats dans les cas de classification d'images médicales. La classification à l'aide de RadImageNet a permis d'améliorer de 0,9 % à 9,4 % la courbe caractéristique de la performance en test (Mei et al., 2022). Cette étude compare un modèle entraîné sur la base de données d'images de scènes naturelles (ImageNet) à un modèle entraîné sur des images médicales (RadImageNet) pour de la classification binaire.

Les métriques d'évaluation des modèles se limitent jusqu'à présent à l'utilisation de la précision et de l'exactitude des résultats de classification, de segmentation et de recalage des algorithmes. Selon la revue faite par (Chan et al., 2020), il est important de se rappeler que

dans le domaine de l'analyse des images médicale à l'aide de modèles d'apprentissage profond, l'interprétation des résultats doit permettre d'établir un lien entre les données d'entraînement et les données cliniques.

1.3.2 Les modèles d'attention pour l'évaluation des modèles d'apprentissage profond

Il peut être difficile de relier la décision des modèles d'apprentissage profond à la décision clinique. Dans plusieurs cas, l'utilisation de métriques d'évaluation comme la précision ne suffit pas à démontrer l'efficacité du modèle et il est donc nécessaire d'avoir recours à des modèles d'attention qui montrent précisément les zones d'attention sur des données d'images. Yahara et al. ont fait l'usage de Grad-CAM pour visualiser les régions d'intérêts d'un modèle d'apprentissage profond pour une tâche de prédiction du risque de progression de la courbure scoliotique. Les régions d'intérêts ayant les signaux d'attention les plus importants ont ensuite été quantifiées pour comparer les régions d'intérêt dans les cas de bonne et de mauvaise classification (Yahara et al., 2022). Ce type d'évaluation permet donc de mieux comprendre le comportement des modèles utilisés afin de les corriger en imposant des régions d'intérêts aux modèles dans le but d'augmenter leurs résultats de prédiction.

En utilisant Grad-CAM, Magnide et al. ont également pu visualiser les régions d'intérêts d'un modèle de classification d'indices de Risser en se basant sur six régions d'intérêts distinctes (les 2 crêtes iliaques, les 2 têtes humérales et les 2 têtes fémorales) de radiographies de visites longitudinales de patients scoliotiques jusqu'à l'atteinte de leur maturité. Les résultats ont démontré qu'en plus des six régions d'intérêts ciblées, le modèle portait attention à des régions tels que la boîte crânienne et la cage thoracique lorsque les patients avançaient en niveau de maturité (Magnide et al., 2021). Ces différences d'attention du modèle n'ont toutefois pas encore été analysées en profondeur. En effet, si le modèle n'accorde pas la même importance aux caractéristiques au travers des visites, il serait donc intéressant de spécifier des régions d'intérêts selon le niveau de maturité et intégrer des éléments tels que la ménarche et la maturité pour mieux orienter la décision du modèle de classification. En plus d'aider à visualiser les éléments entraînant la décision du modèle de prédiction, les modèles d'attention peuvent aussi permettre de faire un lien avec les décisions cliniques et servir d'outils d'évaluation de

l'exactitude des modèles d'apprentissage profond. Dans Wang et al., l'auteur utilise une approche de cartographie des pixels de radiographies de la main basée sur des gradients. Dans ce travail, des modèles de régression linéaire sont bâtis pour observer l'importance de cinq régions d'intérêts sur la prise de décision d'un réseau de neurones convolutif (CNN) pour différents âges osseux. Le choix de cette approche était principalement de permettre une meilleure évaluation pour des modèles de classification par régression et elle pourrait être comparée à l'approche Grad-CAM (Wang, 2021).

1.3.3 Caractérisation automatique de la maturité osseuse

L'acquisition de la maturité squelettique basée sur l'échelle de Risser était jusqu'ici une tâche se basant sur des approximations visuelles faites par des cliniciens, ce qui pouvait occasionner des divergences entre différents observateurs. Cette différence de résultat pourrait être la raison pour laquelle l'indice de Risser obtenait jusqu'ici une plus faible corrélation avec l'évolution de la courbure scoliotique. Dans les dernières années, de nombreux outils d'intelligence artificielle ont trouvé une utilité dans le domaine de l'imagerie médicale. Ces outils ont aussi démontré leur efficacité dans le cas de l'acquisition de l'âge osseux à partir d'images radiographiques dans des études telles que (Kaddioui et al., 2020) pour l'indice de Risser et (Thodberg, Kreiborg, Juul, & Pedersen, 2009) pour l'évaluation de l'âge osseux à partir de radiographies de la main.

La caractérisation de l'âge osseux à partir des caractéristiques des radiographies de la main est une tâche complexe qui résulte souvent en des erreurs de diagnostic. L'automatisation de cette tâche est donc d'un grand apport pour le domaine clinique. BoneXpert est l'un des logiciels les plus populaires pour la caractérisation de l'âge osseux à partir de radiographies de la main (Thodberg et al., 2009). En se fiant à des images d'atlas provenant de TW3 ou GP, l'outil cherchera à segmenter 15 os spécifiques sur lesquels il prendra des mesures afin de déterminer la maturité osseuse des individus de façon automatique. L'approche de BoneXpert consiste en l'utilisation de méthodes d'apprentissage machine telle que les scores d'intensité et de textures issues d'analyse de composantes principales. Les régions de la main sont segmentées et un score leur est attribué afin de déterminer un score final qui permet au logiciel de prédire l'âge

osseux des patients selon GP et TW3. Cette approche permet d'obtenir une plus faible erreur d'âge avec GP que l'acquisition faite par les experts. Cet outil est actuellement utilisé dans plusieurs centres hospitaliers. En orthopédie et particulièrement pour le traitement de la SIA, cet outil peut être utilisé dans la mesure où les cliniciens arrivent à standardiser une méthode de radiographie. Cette méthode permettrait d'avoir les mains des patients lors de l'acquisition de la radiographie, permettant d'analyser la courbure de la colonne vertébrale.

Kaddioui et al. démontrent dans leur étude que l'accord des observateurs pour la gradation du Risser ($k=0.62$) est inférieur à celle d'un CNN de gradation ($k=0.72$) (Kaddioui et al., 2020). Magnide et al., en utilisant un modèle ResNet101-SVM ont obtenu une précision de 80 % à 84 % pour la classification de l'indice de Risser avec des images de la crête iliaque de patient scoliotique (Magnide et al., 2021). Ces études ont donc réussi à démontrer la performance de modèles d'apprentissage profond et machine pour la classification de l'âge osseux. L'outil de Kaddioui et al. a toutefois utilisé des radiographies complètes de la région du bassin sans pour autant spécifier une zone d'intérêt au modèle. La zone d'intérêt dans ce cas serait la crête iliaque des radiographies et pourrait être extraite manuellement de l'image ou alors de façon automatique en utilisant des marqueurs et un masque de segmentation automatique comme cela a été fait dans (Guo, Wang, Teng, & Chen, 2022).

Dans les cas de classification, il est commun d'avoir recours à un réseau neuronal convolutif pour extraire des caractéristiques des données et d'effectuer ensuite une classification à l'aide d'un algorithme d'apprentissage machine. Magnide et al. ont effectué une étude dans laquelle ils comparent les performances d'un modèle ResNet101-SVM à un modèle GLCM-SVM (Gray-Level Co-Occurrence Matrix – Support vector machine) pour la classification d'indices de Risser en se basant sur 6 régions d'intérêts (les crêtes iliaques, les têtes fémorales et les têtes humérales). Pour le premier modèle, ResNet101 est entraîné par apprentissage par transfert pour apprendre des caractéristiques des données et le SVM fait la classification des données en se basant sur les caractéristiques extraites par le modèle. Le modèle GLCM-SVM est plutôt basé sur l'apprentissage machine, et avec ce modèle, le GLCM permet d'extraire les caractéristiques avec une approche statistique et le SVM quant à lui permet de faire la

classification (Magnide et al., 2021). Cette étude a démontré que l'utilisation du premier modèle permettait d'avoir une meilleure précision de classification.

1.3.4 Les modèles prédictifs pour la caractérisation de la croissance osseuse

La caractérisation de l'âge osseux à partir d'étude longitudinale permet de faire des observations sur de grands groupes de patients et d'approximer des comportements de croissances des enfants et des adolescents. C'est dans ce sens que l'étude réalisée par Boeyer, Leary, Sherwood, & Duren a été faite. L'étude avait pour but d'établir une trajectoire de croissance chez les enfants et les adolescents dans un groupe de patients afin de déterminer la période de croissance rapide entre l'enfance et l'adolescence à partir de radiographies de la main récupérée depuis l'âge de 3 mois jusqu'à l'âge de 20 évalués selon la méthode Fels. Selon cette étude, 81% des patients observés avaient une croissance rapide pendant l'enfance et ou l'adolescence (Boeyer et al., 2020). Cette étude a mené à la conclusion que la croissance osseuse n'est pas linéaire contrairement à la croissance de l'âge biologique. Pour pallier le manque d'installation pour effectuer des radiographies dans certains hôpitaux, Wan et al. ont démontré statistiquement qu'il est possible et fiable d'acquérir l'âge osseux à partir d'image ultrasonique et d'y associer un nouveau score osseux comparable aux scores conventionnels tels que le TW2 et TW3 (Wan et al., 2021). Cette étude démontre qu'il est possible d'établir un score osseux par ultrason ayant une forte corrélation avec l'âge osseux acquis par radiographie. Cette étude démontre l'utilité de modèle prédictif pour la caractérisation de l'âge osseux à l'aide de nouvelle méthode d'acquisition de l'âge osseux. De plus, il existe des outils statistiques tels que ceux utilisés dans Pan et al. qui permettent d'évaluer statistiquement des caractéristiques d'un modèle d'apprentissage profond (Pan et al., 2017). (« Physis Bone Age Calculator », s.d.) fut l'application gagnante en 2017 de la compétition de la RSNA (Radiological Society of North America) pour l'inférence automatique de l'âge osseux. Ce modèle basé sur l'utilisation de l'architecture d'Inception V3 a pu obtenir une erreur moyenne absolue dans la prédiction de l'âge osseux de 4,265 mois. Dans le but d'analyser la généralisation et les biais présents dans les applications de l'apprentissage profond pour l'acquisition de l'âge osseux Beheshtian, Putman, Santomartino, Parekh et Yi ont mené une

étude en utilisant l'application 16-bit sur des données d'un hôpital externe. Cette étude arriva donc à la conclusion que l'application avait une bonne généralisation, mais que l'utilisation de l'apprentissage profond pourrait encore avoir des biais basés sur le sexe, l'âge et la maturité squelettique. Ce qui ramène aux limites auxquelles l'apprentissage par transfert fait encore face, comme la limite de cas variés de pathologies ou la limite de représentation ethnique dans les bases de données d'entraînement. Finalement, l'étude faite par (Mei et al., 2022) avait pour but d'entraîner des modèles de réseau de neurones convolutif (CNN) avec un grand ensemble de données médicales en utilisant l'apprentissage par transfert. Ces nouveaux modèles entraînés avec des images médicales ont eu de meilleures performances de détection pour des images médicales que les modèles entraînés sur des images naturelles. Cette nouvelle approche promet donc pour l'utilisation de CNN dans le domaine de l'analyse d'images médicales. Le Tableau 1-1 présente un résumé des différentes approches les plus significatives présentées dans cette section du chapitre 1.

Tableau 1-1 Approches significative pour la caractérisation de l'âge osseux et l'analyse d'images médicale

Approches	Âge osseux	Région d'intérêt	Modalités d'imagerie	Métrique d'évaluation
BoneXpert avec apprentissage machine	GP et TW	Main	X-ray	Erreur BoneXpert VS : TW- 0,8 ans¹ GP – 0.42 ans¹ Erreur des experts avec TW : 0,5 ans¹ Erreur des experts avec GP : 0,82 ans¹
16-bit avec l'apprentissage profond	Âge osseux de la main	Main	X-ray	La différence moyenne absolue par rapport à l'annotation : 4,265 mois¹
Automatic Risser sign VGG16	Gradation de Risser	Crêtes iliaques	X-ray	Coefficient Kappa automatique : 0,72 Précision automatique : 78 % Coefficient kappa expert : 0,6 Précision des experts : 74.5%
ResNet101-SVM	Gradation de Risser	Crêtes iliaques, têtes fémorales et humérales	Radiographie EOS	Précision crêtes iliaques : 84% Précision tête humérale : 78% Précision têtes fémorales : 80%
RadImageNet	Détection de pathologies	14 régions anatomiques	CT, IRM et Ultrason	Aire sous la courbe caractéristique : amélioration entre 0,9% et 9,4% par rapport au modèle entraîné sur ImageNet

¹ L'âge osseux peut ne pas correspondre exactement à l'âge physique

1.4 Sommaire de la revue de littérature

Dans cette revue de littérature, nous avons tout d'abord expliqué ce qu'est la scoliose idiopathique de l'adolescent. Cette section a permis de mieux cerner comment cette pathologie affecte les patients qui en sont atteints et comment, de nos jours, les cliniciens suivent les patients pour leur offrir un traitement convenable. Nous avons introduit dans cette section l'utilisation de méthodes classiques pour l'acquisition de l'âge osseux. Nous avons particulièrement expliqué le rôle de l'échelle de Risser pour la caractérisation de la maturité osseuse et la prédiction de la courbure scoliotique.

Nous avons ensuite introduit l'apprentissage profond et son utilité pour la caractérisation de l'âge osseux. Cette section a permis de mieux comprendre les défis qui sont présents lors de l'application de l'apprentissage profond dans le domaine de l'imagerie médicale. Bien que ces défis soient présents, nous avons aussi démontré que plusieurs études sont faites pour chercher à éliminer ces obstacles en appliquant l'apprentissage par transfert ou alors en entraînant des modèles avec des bases de données d'images médicales pour améliorer les résultats de prédiction. Nous avons également vu que l'application de l'apprentissage profond pour la caractérisation de l'âge osseux est un sujet qui a déjà fait l'objet de plusieurs études.

Nous avons ensuite abordé le sujet de l'application de l'apprentissage machine pour la caractérisation de l'âge osseux. Cette section a particulièrement permis de voir que l'apprentissage profond était jusqu'ici un meilleur outil pour la caractérisation de l'âge osseux. Nous avons finalement abordé le sujet de l'utilisation des modèles prédictifs pour la caractérisation de l'âge osseux. Ces modèles sont principalement utilisés pour démontrer l'efficacité d'une nouvelle méthode de caractérisation de l'âge osseux et comparer ce dernier à des méthodes plus conventionnelles.

Puisque la caractérisation de l'âge osseux est un élément important pour le suivi des patients scoliotiques, il est important d'avoir une méthode d'acquisition permettant également de prédire la progression de la croissance osseuse des patients puisque c'est grâce à cela qu'il est possible de prédire l'évolution de la courbure scoliotique. Bien que l'échelle de Risser soit une

méthode d'acquisition de l'âge osseux qui convienne plus aux patients scoliotiques, son utilité pour la prédiction de la courbure scoliotique et la progression de la croissance squelettique n'a pas encore été démontrée. Mais cela pourrait être dû au fait que cette méthode d'acquisition démontre une grande variabilité inter-observateurs. Puisque récemment les outils d'apprentissage profond ont démontré une efficacité pour l'automatisation de l'acquisition de l'âge osseux, il pourrait être intéressant d'utiliser les caractéristiques apprises par ces modèles pour caractériser la maturité squelettique des patients scoliotiques. Avec ces caractéristiques, il serait possible d'utiliser des modèles d'apprentissage machine pour prédire la maturité osseuse des patients scoliotiques.

CHAPITRE 2

MÉTHODOLOGIE

Dans ce chapitre, la méthodologie utilisée pour ce mémoire est présentée. Dans le but de concevoir un modèle de prédiction automatique de l'âge osseux, trois modèles ont été développés et comparés selon leur performance. Le chapitre de la méthodologie est divisé en cinq parties. La première partie consiste en une description des bases de données. La deuxième partie porte sur le prétraitement des données. La troisième et la quatrième partie présentent les deux premiers modèles utilisés et la cinquième partie présente l'évaluation des modèles et le troisième modèle développé pour la prédiction de l'âge osseux.

2.1 Description de la base de données

Trois bases de données ont été sélectionnées pour entraîner les modèles de prédiction. Ces données ont été collectées à l'Unité de Recherche Clinique en Orthopédie (URCO) du CHU Sainte-Justine et elles sont composées de radiographies EOS de patients atteints de scoliose idiopathique. On compte en tout 2380 radiographies frontales représentant la colonne vertébrale et le bassin de patients atteint de la SIA. Les patients inclus dans cette étude sont compris entre l'âge de 2 à 23 ans. La première base de données de 250 images est une base de données longitudinale avec un suivi de patient variant de 6 à 12 mois et donne l'âge osseux des patients basé sur l'indice de Risser. Elle a été choisie, car elle pourrait permettre d'éventuellement évaluer l'outil de prédiction pour sa capacité à prédire l'atteinte des stades tels que la ménarche, le pic de croissance et la maturité complète chez les patients atteints de la SIA. La deuxième base de données de 1830 images a été validée par des experts et donne l'âge osseux des patients basé sur l'indice de Risser. Cette base de données à l'avantage d'avoir été utilisé précédemment pour l'entraînement d'un modèle de prédiction automatique de l'indice de Risser. Finalement, la troisième base de données de 300 images donne l'âge osseux des patients basé sur les indices de Risser et de Sanders. La dernière base de données offre également les détails sur l'état du triradié des patients et leur taille. Elle permettra d'avoir des caractéristiques cliniques supplémentaires grâce aux radiographies de la main des patients pour

entraîner les modèles de prédiction. Le Tableau 2-1 détaille des groupes d'âge des trois bases de données ainsi que les méthodologies de description de l'âge osseux qu'elles permettent de lire.

Tableau 2-1 Présentation de la base de données

Modalité d'imagerie	EOS		
Nom	ARC-1	ARC-2	ARC-3
Type de base de données	Validée par des experts	Longitudinale (suivi entre 6 et 12 mois)	Protocole d'imagerie incluant les mains et le bassin
Population	10 à 18 ans	13,65±1,23 ans	2,46 à 13,84 ans
Nombre d'images	1830	250	300
Gradation de l'âge osseux	Échelle de Risser	Échelle de Risser	Échelle de Risser et Sanders

2.2 Prétraitement des données

2.2.1 Données d'entraînement du modèle d'apprentissage machine basée sur les caractéristiques d'image

Pour extraire les caractéristiques d'images, il a fallu faire l'entraînement d'un modèle d'apprentissage profond. Pour y arriver, une sélection de la région du bassin a été effectuée sur la base de données validée par les ARC1 et 2 (n=2080). Pour la base de données ARC 3 (n=254), une sélection automatique du bassin et des mains des patients a été faite suivant les travaux de Noudel et al. où le squelette des patients est d'abord repéré, par la suite, les mains des patients et leur bassin sont automatiquement détectés grâce à un VotingClassifier et finalement une segmentation automatique des régions est faite grâce à un traitement d'image (Voir Figure 2.1) (Noudel et al.,2023). Les radiographies des régions d'intérêts ont ensuite été inspectées visuellement pour s'assurer de leur conformité. Certaines radiographies segmentées automatiquement ont ensuite dû être sélectionnées manuellement. Pour toutes les images de la région du bassin, un filtre médian fut appliqué dans le but d'entraîner le modèle

d'apprentissage profond. Les images de la main sont toutefois restées intactes pour l'entraînement du modèle d'apprentissage profond. Toutes les images ont ensuite été redimensionnées à une taille de (224x224) pixel et transformées en tableau de pixels. Les tableaux ont ensuite été normalisés sur une échelle de 0 à 1. Finalement les données ont été divisées en deux groupes pour l'entraînement et le test du modèle d'apprentissage profond. Les caractéristiques d'images de bassins des bases de données ARC1 et 2 furent combinées pour l'entraînement du modèle d'apprentissage machine. Pour ce qui est de la base de données ARC 3, les données d'images des bassins et des mains ont été extraites séparément pour entraîner le modèle d'apprentissage machine à la prédiction des indices de Risser avec les données du bassin et l'indice de Sanders avec les données de la main. Une fois les caractéristiques extraites, elles ont toutes été normalisées entre 0 et 1 pour faciliter l'analyse des données et améliorer la performance du modèle de prédiction. Une fois les données normalisées, elles ont été séparées en deux groupes, un correspondant aux données d'entraînement (70%) et l'autre correspondant aux données de test (30%) du modèle d'apprentissage machine. L'entraînement de trois modèles d'apprentissage machine s'est fait grâce aux caractéristiques d'image du modèle d'apprentissage profond. Les bases de données ARC-1 (n=1830) et ARC-2 (n=250) furent combinées pour faire un premier entraînement du modèle d'apprentissage machine grâce aux caractéristiques extraites d'images. Les données du bassin et de la main des patients de la base de données ARC-3 (n=254) ont aussi été entraînées par les deux autres modèles d'apprentissage profond.

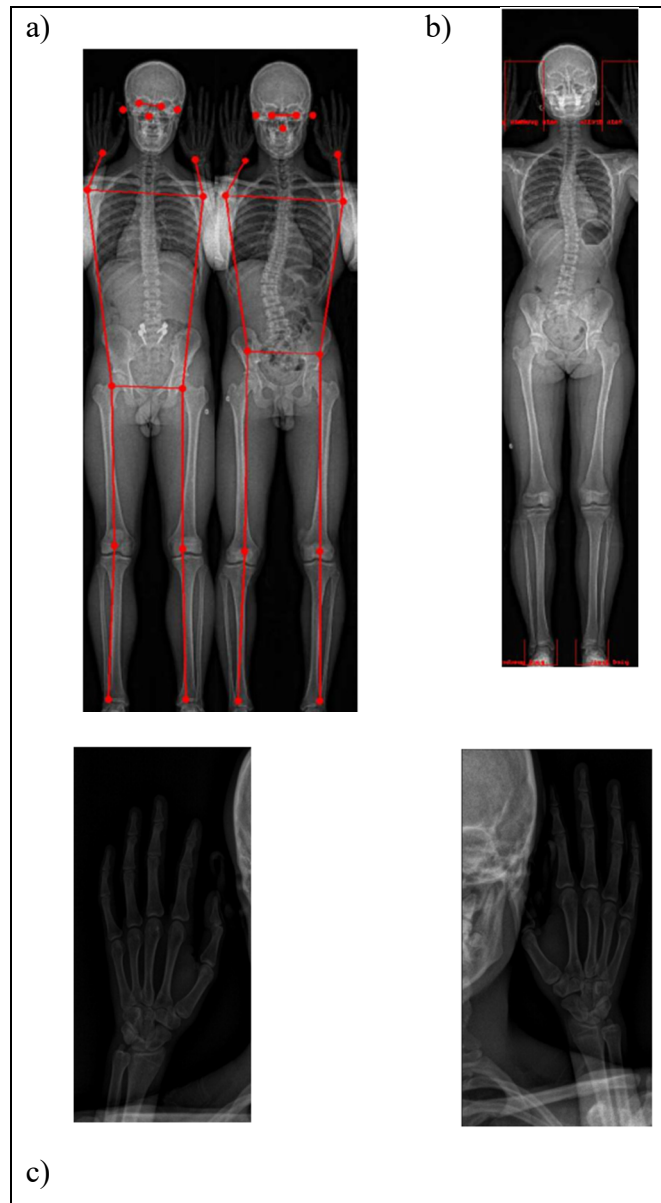


Figure 2.1 Processus de détection automatique de la main de patients a) repérage du squelette
 b) détection des mains grâce à VotingClassifier c) résultats de la segmentation des mains.
 Tirée de Noudel (2023)

2.2.2 Données d'entraînement du modèle d'apprentissage machine basée sur les caractéristiques cliniques

La base de données ARC-3 (n=254) a été utilisée pour faire un entraînement du modèle d'apprentissage machine grâce aux caractéristiques cliniques des patients. Les caractéristiques cliniques des patients furent prélevées de deux sources ; tout d'abord de leur dossier médical

avec leur âge et l'état d'ouverture de leur triradié et ensuite d'une évaluation de l'âge osseux des patients obtenue grâce à l'outil BoneXpert. Avec BoneXpert, il a été possible de déterminer l'âge osseux basé sur TW3 et d'avoir des informations supplémentaires sur les scores d'âge pour chaque os de la main. Pour l'extraction des caractéristiques à partir de BoneXpert, les radiographies des mains gauches ont été prises en compte pour la plupart des patients, puisque cette main est la plus utilisée par les orthopédistes pour la lecture de l'âge osseux (Manzoor Mughal, Hassan, & Ahmed, 2014). Dans les cas où la main gauche a été moins visible, la main droite a été prise. L'ANNEXE II présente les différentes caractéristiques extraites de BoneXpert et leur description. L'analyse faite par l'outil donne également accès à un rapport sur le patient qui inclut l'âge suivant les méthodes GP et TW3, leur indice de santé osseuse et une prédiction de leur taille adulte (*Voir ANNEXE III, p.60*). Plusieurs données extraites de BoneXpert n'ont pas pu être utilisées pour l'entraînement du modèle d'apprentissage machine. En effet, les caractéristiques ayant uniquement des valeurs de 0 ou -1, indiquant que l'outil n'a pas pu en faire une analyse pour les patients, ont été rejetées. Parmi les 81 caractéristiques de BoneXpert pour 254 patients, uniquement 30 caractéristiques ont été choisies pour l'entraînement du modèle avec une base de données comprenant 104 patients. Une sélection des caractéristiques et des patients à inclure dans la base de données a été faite en s'assurant tout d'abord que toutes les caractéristiques aient des valeurs uniques pour chaque patient et ensuite qu'aucun patient choisi n'ait des caractéristiques de Meta2, meta3, ... équivalentes à -1 puisque cette valeur correspond à un âge osseux manquant (*Voir ANNEXE II, p.57*). Toutes les caractéristiques des patients, leur sexe, leur âge au moment de la radiographie, l'état du cartilage du triradié et les données d'analyse de l'âge osseux à l'aide de l'outil BoneXpert ont été mis dans un fichier CSV. Une fois les données obtenues, elles ont été normalisées entre 0 et 1 pour faciliter leur analyse et améliorer la performance du modèle de prédiction. 70 % des données ont permis d'entraîner le modèle d'apprentissage machine et 30 % ont servi à tester le modèle.

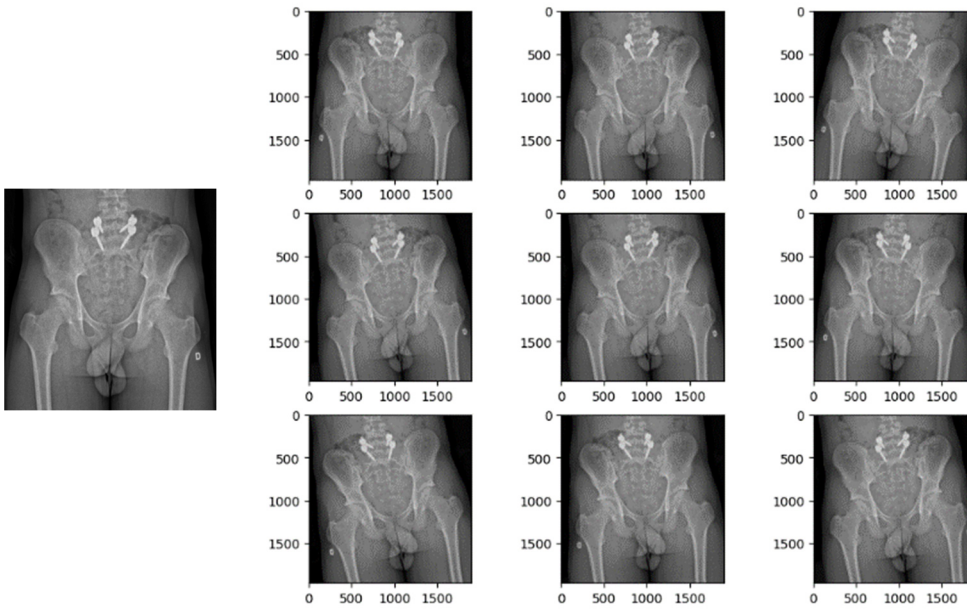
2.3 Les modèles d'apprentissage machine basée sur les caractéristiques d'images

En raison de ses performances dans les travaux précédents, le modèle VGG16 a été choisi comme modèle d'apprentissage profond. En effet, les travaux de Kaddioui et al. et de Magnide et al. ont démontré l'efficacité de ce modèle pour la classification de l'âge osseux basé sur l'échelle de Risser (Kaddioui et al., 2020) (Magnide et al., 2021). Le modèle fut implémenté grâce à l'apprentissage par transfert. L'entraînement a été fait avec les trois bases de données. L'augmentation des données a ensuite été réalisée afin d'avoir une grande variabilité dans le type d'image d'entraînement pour le modèle. Une vérification du type d'augmentation a été faite pour s'assurer de la validité des données qui seraient fournies au modèle pour son entraînement. Et après quelques essais, les paramètres d'augmentation suivante ont été sélectionnés : rotation de 15°, déplacement en longueur de 0,01% et en largeur de 0,01% respectivement, miroir horizontal et une compensation des points hors limite de type rapproché. Cette augmentation a été appliquée sur les radiographies du bassin et de la main (Voir Figure 2.1). L'entraînement des modèles s'est ensuite fait à l'aide des paramètres présentés dans le Tableau 2-2.

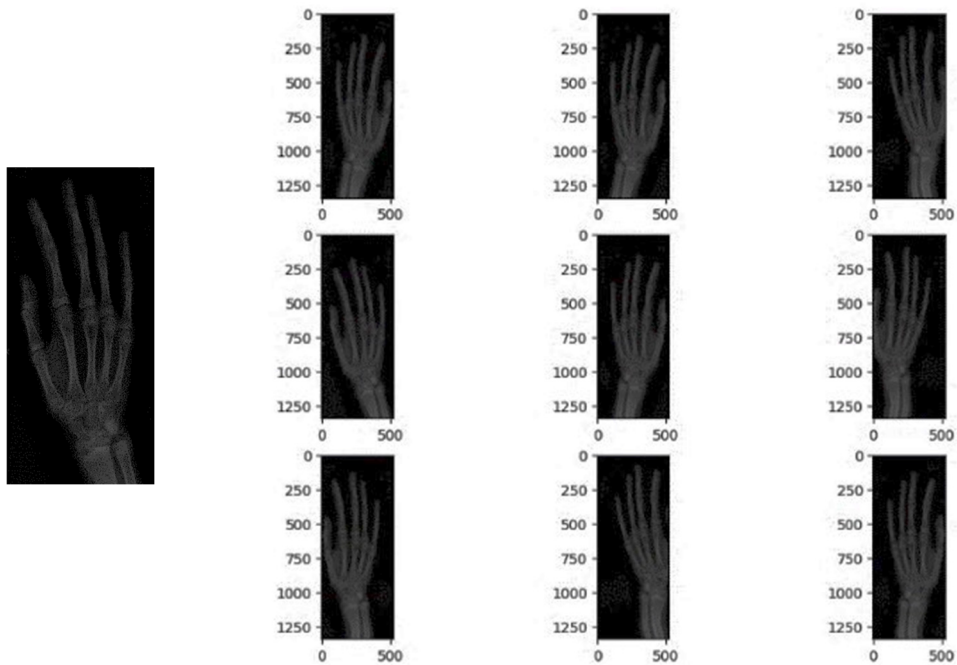
Tableau 2-2 Tableau des paramètres des modèles se servant des caractéristiques d'images extraites du modèle d'apprentissage profond

	Modèles		
	VGG16-RVS		BoneXpert-VGG16-RVS
Paramètres	ARC-1 et 2	ARC-3	ARC-3
Images par entraînements	64	8	4
Nombre de répétitions	60	60	60
Taux ou pas d'apprentissage	1 ^e -5	1 ^e -5	1 ^e -5
Couches entraînées	Superficielles	Superficielles	Superficielles
Fonction de perte	Précision	Précision	Précision
Couche d'extraction des caractéristiques	21	21	21

Après l'entraînement, les caractéristiques d'images ont été extraites pour entraîner un modèle d'apprentissage machine. En effet, en raison de la faible quantité de données dont nous disposons, nous espérons observer une meilleure performance de prédiction par l'utilisation de modèles d'apprentissage machine basée sur les caractéristiques d'images extraites du modèle d'apprentissage profond. Le choix du modèle d'apprentissage machine pour la prédiction de l'âge osseux s'est arrêté sur le modèle de régression par machine à vecteur de support (RVS). L'optimisation des paramètres du modèle a été faite grâce à une grille de recherche ou tous les paramètres listés dans le Tableau 2-3 ont été combinés et leur capacité à aider la prédiction du modèle a été évaluée. La combinaison optimale des paramètres a été obtenue en évaluant la proportion de détermination du modèle de régression. Le modèle complet proposé pour cette étude est présenté à la Figure 2.3 (en orange). On y retrouve donc le processus permettant la prédiction de l'âge osseux grâce à un modèle RVS basé sur les caractéristiques d'images extraites d'un modèle d'apprentissage profond (VGG16) selon les indices de Risser et de Sanders.



a) Image originale (à droite) et augmentée (à gauche) de la région du bassin



b) Image originale (à droite) et augmentée (à gauche) de la région de la main

Figure 2.2 Augmentation des données de patient pour l'entraînement du modèle d'apprentissage profond

2.4 Les modèles d'apprentissage machine basée sur les caractéristiques cliniques

La seconde expérimentation a été réalisée grâce à la troisième base de données comportant les données d'indices de Risser et de Sanders des patients scoliotiques. Tout d'abord, l'application BoneXpert a été utilisée pour prédire l'âge osseux des patients en se basant sur les radiographies de leur main. Par la suite les caractéristiques de BoneXpert ont été extraites pour ensuite être utilisées pour l'entraînement d'un modèle de régression par RVS. Les caractéristiques cliniques des patients telles que leur âge, indice Sanders, indice de Risser et l'état de leur triradié ont été utilisées pour l'entraînement du modèle. L'optimisation des paramètres du modèle a été faite grâce à une grille de recherche où tous les paramètres listés dans le Tableau 2-3 ont été combinés et leur capacité à aider la prédiction du modèle a été évaluée. Le modèle a été utilisé pour la prédiction de l'âge osseux basé sur l'indice de Risser et l'indice de Sanders fournit pour la troisième base de données comme présentée à la Figure 2.3 (en bleu).

Tableau 2-3 Hyperparamètres pour l'optimisation des modèles de régression par machine à vecteur de support

Paramètres	Valeurs pour le modèle VGG16-RVS	Valeurs pour le modèle BoneXpert-RVS	Valeurs pour le modèle BoneXpert-VGG16-RVS
Kernel	Linéaire et RBF	Linéaire, poly et RBF	Linéaire et RBF
Epsilon	0.00001,0.0001, 0.001, 0.01, 0.1, 0.2,0.3,0.4,0.5	0.1,0.5,0.6,0.75,1	0.00001,0.0001, 0.001, 0.01,0.1,0.2,0.3,0.4,0.5
C	0.1,0.2,0.3,0.4,0.5,0.6, 0.75,1	0.1,0.5,0.6,0.75,1	0.1,0.2,0.3,0.4,0.5,0.6, 0.75,1
Gamma	0.000001, 0.00001,0.0001, 0.001, 0.01,0.1,0.5	0.1,0.5,0.6,0.75,1	0.000001, 0.00001,0.0001, 0.001, 0.01,0.1,0.5
Degré	N/a	1,2,3,4,5,6	1,2,3,4,5,6
Coefficient	N/a	1,2,3,4	1,2,3,4

2.5 Évaluation des modèles

Les deux modèles proposés seront finalement évalués selon leur capacité à trouver une corrélation entre les caractéristiques (d'images, cliniques) et les métriques d'évaluation de l'âge osseux (Risser, Sanders). Puisque la classification est faite par un modèle RVS, l'évaluation se fera par l'erreur quadratique moyenne (RMSE), l'erreur moyenne absolue (MAE) et le coefficient de corrélation (R2) qui sont formulés dans les équations 2.1 à 2.3 où ;

\hat{y} – valeur prédite

y_i – valeur observée

\bar{y}_i – valeur moyenne des valeurs de l'ensemble de caractéristiques

N – nombre d'observations

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n}} \quad (2.1)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}|}{n} \quad (2.2)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2} \quad (2.3)$$

L'erreur quadratique moyenne permettra de mesurer la capacité du modèle à faire des prédictions se rapprochant des données d'entrées. Le modèle ayant l'erreur la plus faible sera le meilleur puisqu'il permettra d'obtenir des prédictions qui se rapprochent le plus de la réalité. L'erreur moyenne absolue indique quant à elle une quantification de l'erreur de la prédiction par rapport à la vérité absolue correspondant à la valeur observée. Finalement le coefficient de corrélation représente la proportion de variation de la variable dépendante (les caractéristiques) par la variable indépendante (l'âge osseux). Cette métrique exprime donc le niveau de corrélation qui relie les deux variables. Une valeur de R2 de 1 signifiera donc que la variance des caractéristiques explique parfaitement la variance de l'âge osseux.

Une évaluation finale d'un dernier modèle sera faite. Ce modèle évaluera la capacité du modèle RVS à prédire les âges osseux (Risser et Sanders) en se basant sur la combinaison des

caractéristiques d'images et cliniques de la troisième base de données ARC 3 (n=104). Cette implémentation présentée à la Figure 2.3 (en rouge) permettra d'avoir une comparaison de la performance d'un modèle RVS avec en entrée deux différents types de caractéristiques suivant les paramètres présentés dans le Tableau 2-3.

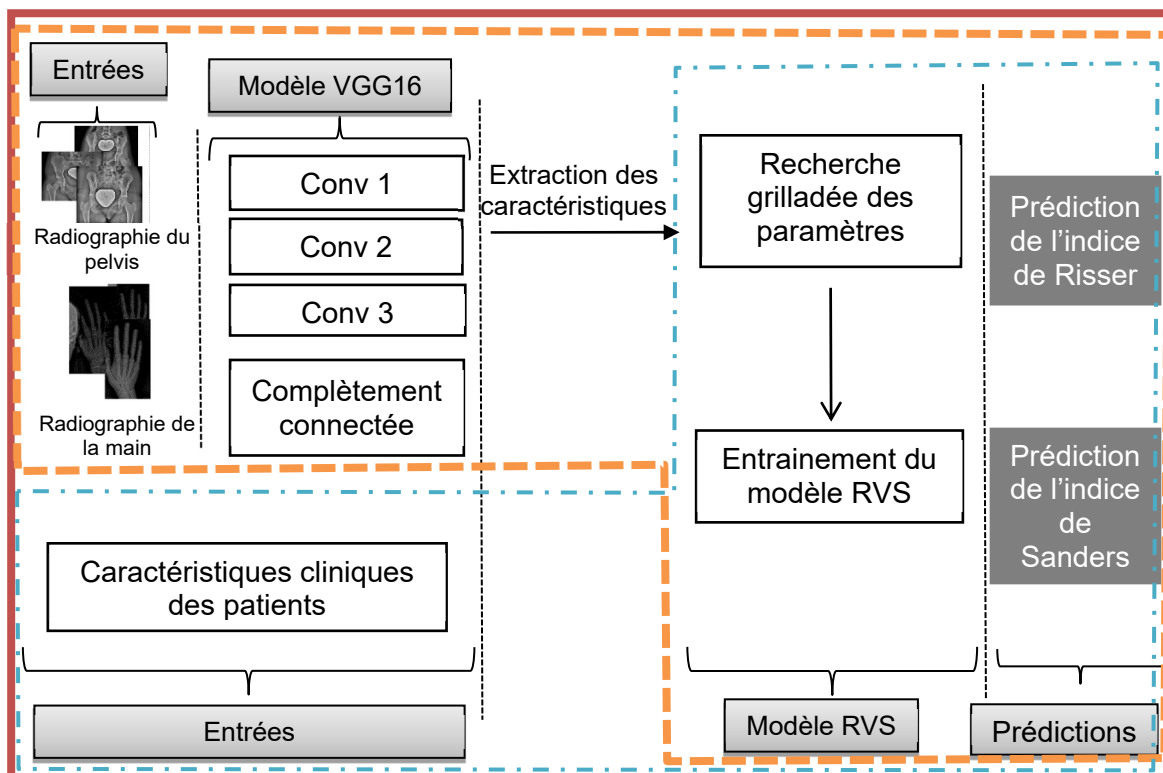


Figure 2.3 Structure du modèle pour la prédiction de l'âge osseux basée sur les échelles de Risser et l'indice de Sanders avec en orange la structure du modèle VGG16-RVS, en bleu celle du modèle BoneXpert-RVS et en rouge celle du modèle VGG16-BoneXpert-RVS

CHAPITRE 3

RÉSULTATS ET DISCUSSION

3.1 Résultats

Nous nous sommes tout d'abord intéressés à déterminer les paramètres optimaux de modélisation des modèles de prédictions. Le Tableau 3-1 présente les différents paramètres d'entraînement du modèle RVS pour la prédiction des indices de Risser et de Sanders. On remarque que pour la prédiction de l'indice de Risser les paramètres ayant permis d'avoir les classifications optimales. On remarque donc que la plupart des modélisations des fonctions de prédictions sont faites par des fonctions Kernel de type RBF à l'exception des modèles prédictifs de l'indice de Sanders pour les modèles BoneXpert-RVS et BoneXpert-VGG16-RVS qui eux ont été modélisé par des fonctions kernel de type polynomial. Les autres paramètres nous démontrent la simplicité des différents modèles de prédiction.

Tableau 3-1 Résultats de la recherche grillardée pour l'optimisation des Hyperparamètres des modèles de prédiction

Bases de données	Valeurs pour le modèle VGG16-RVS			Valeurs pour le modèle BoneXpert-RVS		Valeurs pour le modèle BoneXpert-VGG16-RVS	
	ARC 1 et 2	ARC 3		ARC 3		ARC 3	
Indice	Risser	Risser	Sanders	Risser	Sanders	Risser	Sanders
Kernel	RBF	RBF	RBF	RBF	Poly	RBF	Poly
Epsilon	0.01	-	0.1	0.00001	0.0001	0.00001	0.001
C	0.3	1	0.75	1	0.5	1	0.1
Gamma	0.1	0.01	0.01	0.00001	0.5	0.1	0.1
Degré	-	-	-	1	1	-	2
Coefficient	-	-	-	1	4	-	1

Le Tableau 3-2 présente les résultats de performance des trois modèles évalués sur les trois bases de données. Le premier modèle VGG16-RVS a démontré une meilleure performance pour la prédiction de l'âge osseux basé sur le Risser avec les bases de données ARC 1-2 et ARC 3. En effet, le modèle a obtenu des scores R2 de 0,75 et 0,76 respectivement pour les deux bases de données. Ces scores signifient que les caractéristiques d'images extraites au modèle d'apprentissage profond corrélaient à 75 % et 76 % avec l'indice de Risser des bases de données respectives. La performance du modèle pour la prédiction de l'indice de Sanders de la base de données ARC 3 s'est avérée être très faible avec un score R2 de 0,50.

Tableau 3-2 Résultats des prédictions des différents modèles implémentés a) l'erreur quadratique moyenne RMSE b) l'erreur moyenne absolue MAE c) le coefficient de corrélation (R2)

Modèles de prédiction	Base de données ARC 1 et 2			Base de données ARC 3					
	Risser			Risser			Sanders		
	RMS E	MA E	R2	RMSE	MA E	R2	RMS E	MA E	R2
VGG16 + RVS	0.18	0.14	0.75	0.20	0.17	0.76	0.26	0.21	0.50
BoneXpert + RVS	-	-	-	0.17	0.12	0.83	0.13	0.09	0.89
BoneXpert +VGG16 + RVS	-	-	-	0.30	0.24	0.44	0.15	0.11	0.85

Le second modèle BoneXpert-RVS qui a uniquement été testé sur la base de données ARC 3 a obtenu les meilleures performances pour la prédiction de l'indice de Risser et de Sander avec un score R2 de 0,83 et 0,89 respectivement pour les deux indices. L'erreur moyenne absolue (MAE) des deux prédictions a été de 0,15 et 0,14 respectivement et leur RMSE respectif a été de 0,21 et 0,17. Ce modèle démontre donc que les caractéristiques de BoneXpert ont une forte corrélation avec les indices de Sanders et de Risser.

Le dernier modèle BoneXpert-VGG16-RVS testé sur la base de données ARC 3 a obtenu des performances moins bonnes que le modèle précédent pour l'indice Risser avec un score R2 de 0,44, une MAE de 0,24, un RMSE de 0,30. Les performances du modèle pour la prédiction de l'indice de Sanders étaient meilleures que celle des deux modèles précédant avec un score R2 de 0,85, une MAE de 0,11 et un RMSE de 0,15. Ce modèle a permis de démontrer qu'une

combinaison des caractéristiques de BoneXpert et de VGG16, permettre d'améliorer la prédiction, mais uniquement pour l'indice de Sanders qui semble avoir une plus forte corrélation avec ces caractéristiques cliniques puisque le Sanders se base sur une observation des radiographies de la main comme c'est le cas avec l'outil BoneXpert plutôt que celle du bassin comme c'est le cas avec l'indice de Risser.

Nous nous sommes ensuite intéressés aux différentes erreurs obtenues par les modèles de prédictions. Nous savons que l'erreur moyenne absolue représente le MAE du modèle prédictif alors que le RMSE représente une erreur moyenne prenant plus en compte les cas de prédictions aberrantes. Tout d'abord, lorsqu'on s'intéresse au MAE, on remarque que la plus faible erreur pour les prédictions d'indices de Risser est de 0,12 avec le modèle BoneXpert-RVS. Cette erreur représente une erreur moyenne de prédiction d'un peu moins que 1 Risser, ce qui est assez bon pour mon modèle prédictif. Le modèle BoneXpert -RVS a également obtenu le RMSE le plus faible de 0,17; ce score est un peu plus élevé que le MAE en raison des cas de prédiction aberrante visible à la Figure 3.2-aFigure 3.1 . Lorsqu'on compare les cas de prédiction de l'indice de Risser des Figure 3.1 et 3.2, on remarque ce dont les prédictions des indices sont moins dispersées dans le cas du modèle BoneXpert -RVS.

Par la suite, pour la prédiction de l'indice de Sanders, nous avons remarqué que le modèle BoneXpert-RVS le plus faible MAE de 0,09 représentant une erreur moyenne d'un peu moins qu'un Sanders pour chaque prédiction faite par le modèle. Une comparaison des deux erreurs nous permet encore de remarquer que le RMSE de 0,13 est plus élevé que le MAE en raison des cas de prédictions aberrantes. La Figure 3.3-a permet donc de visualiser les cas aberrants qui ont pu influencer le RMSE du modèle prédictif. Et lorsqu'on compare le nuage de prédiction du modèle BoneXpert-RVS, on remarque que les points de prédictions sont beaucoup plus proches des observations que les points des Figure 3.3-b et c.

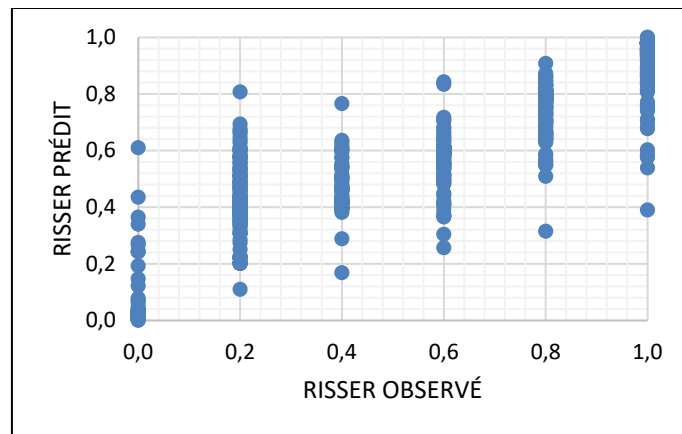


Figure 3.1 Nuage de points des observations des indices de Risser observés par rapport aux prédictions pour l'indice de Risser du modèle VGG16-RVS de la base de données ARC 1 et 2

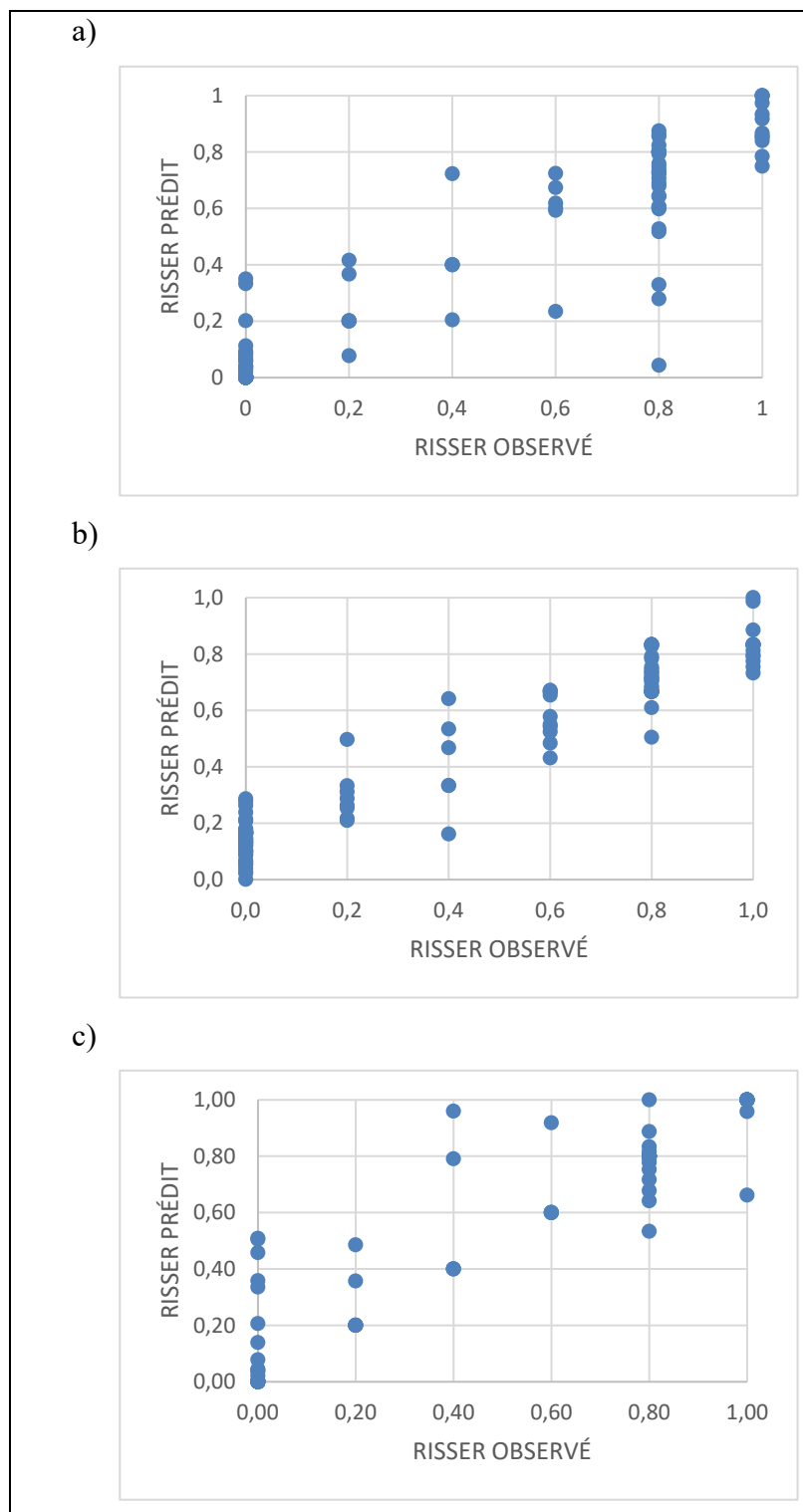


Figure 3.2 Nuage de points des observations des indices de Risser observés par rapport aux prédictions pour l'indice de Risser des modèles de prédiction de la base de données ARC 3
 a) modèle BoneXpert-RVS b) modèle VGG16-RVS c) modèle BoneXpert-VGG16-RVS

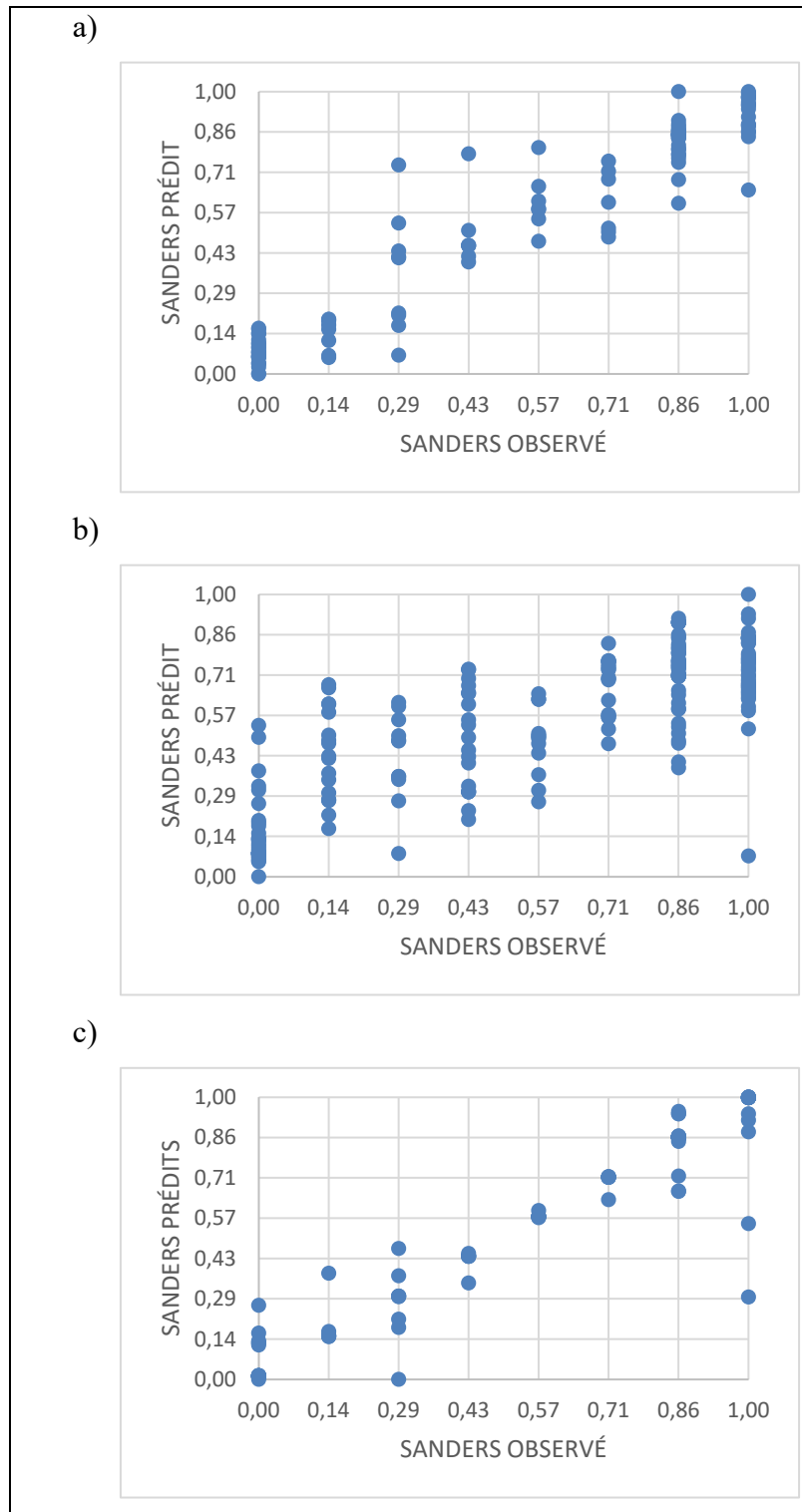


Figure 3.3 Nuage de points des observations des indices de Sanders observés par rapport aux prédictions pour l'indice de Sanders des modèles de prédiction de la base de données ARC 3
 a) modèle BoneXpert-RVS b) modèle VGG16-RVS c) modèle BoneXpert-VGG16-RVS

3.2 Discussion

L'évaluation de l'âge osseux est un aspect important dans le suivi des patients atteints de scoliose idiopathique chez les adolescents (SIA). Cette évaluation permet entre autres de faire le choix de traitement pour les patients, que ce soit pour la prescription de corset ou pour la modulation de la croissance. L'objectif de ce mémoire était de développer un outil de prédiction automatique des indices de Risser et Sanders. Pour atteindre l'objectif de ce mémoire, nous avons tout d'abord procédé à une collecte et un traitement des données pertinentes pour l'étude; pour par la suite nous avons entraîné des modèles de prédictions pour les différentes bases de données sélectionnées et nous avons trouvé pour chacun de ces modèles des hyperparamètres optimaux; finalement, nous avons pu évaluer la performance de chacun des modèles grâce à des métriques d'erreur et de corrélation.

3.2.1 Collecte et traitement de données

Tout d'abord, nous avons évalué les modèles de prédictions sur deux bases de données distinctes. La première base de données ARC1 et 2 nous a permis de valider l'importance d'avoir des radiographies EOS incluant les mains des patients. Et nous avons remarqué que l'inclusion des mains dans les radiographies des patients atteints de la SIA permet d'améliorer les performances des modèles prédictifs. La détection automatique de la main gauche faite pour les radiographies de la base de données ARC-3 était de 89,1 % grâce au modèle de détection de pose ce qui démontre que le protocole d'acquisition des radiographies est fiable et que ces images peuvent être utilisées pour une prédiction de l'indice de Sanders qui s'avère être utile pour un meilleur suivi des patients scoliotique (Noudel et al., 2023), (Neal & Shirley, 2018). Une sélection des patients et des caractéristiques incluent dans l'étude a ensuite été faite pour éviter d'inclure des patients ayant eu une analyse clinique non complète ou alors des caractéristiques absentes pour tous les patients sélectionnés. Ce travail de tri de caractéristiques a permis que n'inclure que les caractéristiques cliniques les plus significatives.

3.2.2 L'entraînement des modèles de prédiction

Le modèle prédictif choisi pour cette étude était le modèle SVR qui est un modèle qui se base sur une analyse statistique des données. Ces données sont ensuite modélisées pour définir une courbe de prédiction contrôlée par plusieurs paramètres. Ce modèle a l'avantage d'offrir une prédiction à la suite d'une analyse du degré avec lequel des données (caractéristiques) décrivent les observations (indices de maturités). Ce type de modèle a l'avantage d'être simple à comprendre et son évaluation est facilement interprétable dans le domaine clinique. Pour entraîner les modèles de prédictions, il a fallu sélectionner les paramètres optimaux. Le choix de ces paramètres a été fait par une recherche grillardée durant laquelle plusieurs combinaisons de paramètres ont été faites et la combinaison ayant permis au modèle d'obtenir la meilleure performance a été choisie. Nous avons pu démontrer ainsi la simplicité de calcul des différents modèles de prédiction.

3.2.3 L'évaluation des modèles de prédiction

Les études précédentes ont obtenu des précisions de classification de l'âge osseux basé sur l'échelle de Risser de 78 % à 85 %. Dans cette étude, nous avons plutôt développé un outil de prédiction des indices de Risser et de Sanders sur une échelle continue. Le coefficient de corrélation entre les caractéristiques des patients et leurs indices de Risser pour cet outil est de 83 % lorsqu'on se base sur les caractéristiques cliniques extraites de BoneXpert. Alors que le coefficient de corrélation entre les caractéristiques des patients et leurs indices de Sanders est de 89 % lorsqu'on se sert des caractéristiques de BoneXpert. Cette étude a pu démontrer la performance d'un modèle de régression pour la prédiction de l'âge osseux. Et le fait d'avoir une prédiction continue nous a permis de nous rapprocher de l'observation des cliniciens qui est pour la plupart du temps une observation continue de la progression osseuse. En plus de se rapporter aux caractéristiques des radiographies des patients, le modèle se sert des données cliniques qui sont souvent prises en compte pour le traitement des patients. Les données cliniques considérées dans cette étude ont particulièrement amélioré la performance du modèle pour la prédiction de l'indice de Sanders. Cela est dû au fait que ces données étaient plus en lien avec les caractéristiques d'image de la main des patients. Les données de radiographies

des patients ont quant à elles permis au modèle de prédiction d'avoir une meilleure performance pour l'indice de Risser lorsqu'on les combinait aux données cliniques.

Nous avons obtenu un modèle de prédiction de l'indice de Risser basé sur les caractéristiques cliniques de la main des patients et un second modèle de prédiction de l'indice de Sanders basé sur les caractéristiques cliniques de la main des patients. Et les résultats obtenus pour les deux modèles sont prometteurs et seront utiles dans un contexte clinique où la prédiction des points clés de la croissance des patients tels que la ménarche et le pic de croissance sont importants pour le choix de soins appropriés. Nous pensons donc qu'en plus d'être d'un support au clinicien pour l'observation des degrés de maturité osseuse, cet outil sera utile pour prédire la courbe d'évolution de la maturité osseuse basée sur les indices de Risser et de Sanders.

CONCLUSION

L'objectif de ce mémoire est de mettre sur pied un outil de prédiction de l'âge osseux de patients scoliotiques basé sur les indices de Risser et de Sanders et validé avec une annotation clinique et avec BoneXpert. Cet outil a pour avantage d'offrir aux cliniciens un moyen permettant d'observer avec une seule radiographie deux indices d'âge osseux et de faire un suivi de l'évolution la courbure des patients scoliotiques. En effet, l'utilisation de l'indice de Risser uniquement pour la prédiction de l'évolution de la courbure scoliotique a jusqu'ici démontré une faible fiabilité et l'outil proposé permet d'avoir en une seule radiographie, un point de comparaison entre la prédiction de deux indices de développement osseux. De plus, cet outil permet d'avoir une bonne précision de prédiction des indices de Risser et de Sanders et se base sur des caractéristiques extraites d'un outil qui a déjà fait ses preuves dans le domaine clinique. L'application clinique de cet outil sera d'éventuellement permettre aux cliniciens d'avoir une prédiction de l'évolution de la maturité squelettique des patients scoliotiques dans le but de prendre une décision éclairée quant aux traitements convenables.

Bien que les modèles aient obtenu de bonnes performances de prédiction, l'outil développé comporte quelques limites. En effet, les modèles ayant obtenu la meilleure performance n'ont été entraînés que pour un ensemble de données de 104 patients, ce qui est un faible échantillon pour la généralisation de l'outil. De plus, les images de main utilisées pour extraire les caractéristiques de BoneXpert n'avaient pas la forme idéale pour l'outil. Il aurait fallu avoir des radiographies où les poignets et les mains des patients sont parfaitement alignés. Il serait donc nécessaire de revoir le protocole d'acquisition des radiographies. Finalement, le modèle d'apprentissage profond n'arrivait pas à une haute précision de ces résultats, ce qui a pu influencer les résultats de modèles de régression par machine à vecteur de support se servant des caractéristiques d'images extraites des modèles d'apprentissage profond.

RECOMMANDATIONS

L'évaluation de l'âge osseux des patients atteints de SIA se faisait jusqu'ici grâce à des méthodes plus conventionnelles comme TW-3 en observant les radiographies de la main et du poignet des patients. L'automatisation de l'évaluation par les méthodes conventionnelles a été rendue possible grâce aux outils comme BoneXpert qui permettent d'avoir une évaluation détaillée de chaque phalange et une prédiction de croissance des patients. Néanmoins, des méthodes d'évaluation de l'âge osseux plus modernes comme l'indice de Risser et de Sanders ont pris en ampleur dans le milieu clinique en raison de leur simplification. Nous proposons un outil qui pourra éventuellement permettre d'avoir une évaluation simultanée de la maturité osseuse basée sur l'observation de la main et l'observation du bassin grâce aux indices de Risser et de Sanders. Pour y arriver, il faudra établir un protocole d'acquisition des radiographies permettant d'avoir avec une seule image la colonne vertébrale, le bassin, et les mains des patients parfaitement en alignement avec leur poignet. Le protocole devra permettre d'appliquer les travaux de Noudel et al. pour être en mesure de faire une segmentation automatique des régions d'intérêt, soit le bassin et la main gauche. Pour être en mesure d'avoir un outil qui est applicable à une plus grande population, la construction d'une base de données révisée par des experts du domaine sera primordiale. En effet, avec une base de données complète, il sera possible d'augmenter le nombre de caractéristiques de chaque patient sans avoir à utiliser les caractéristiques d'un outil externe comme ça a été le cas dans ce mémoire. Une étude de validation de l'outil grâce à une comparaison inter-observateur et intra-observateur permettra d'évaluer sa robustesse.

ANNEXE I

Article présenté à la conférence SPIE Medical Imaging

Automatic assessment of skeletal maturity in Adolescent Idiopathic Scoliosis patients using support vector regression on deep features

Hilary Cintia Djuikoua Wouafo^a, Julie Joncas^b, Marjolaine Roy Beaudry^b, Soraya Barchi^b, Stefan Parent^b, Hubert Labelle^b and Luc Duong^a

^aDept. of Software and IT Engineering, École de technologie supérieure, Montreal, Canada

^bResearch center CHU-Sainte-Justine, Montréal, Canada

Article publié à la conférence SPIE Medical Imaging, San Diego, février 2023

ABSTRACT

Skeletal maturity assessment is an important step in the treatment of adolescent idiopathic scoliosis (AIS). Traditional methods rely on the bone age assessment using 2D X-ray radiography. Assessment is generally performed using the Risser stages that is routinely used to assess the skeletal maturity through the observation of the level of ossification of iliac crests. This bone maturity assessment method is preferred in AIS but in practice, shows a rather high-level interobserver variability. This study aims to use an automatic Risser stage classification for a longitudinal study of follow-up visits to observe growth indicators of AIS patients. A regression model will then be used to evaluate the maturity changes of patients from a Risser stage to another. For the classification task, the pre-trained model of VGG16 was implemented with Python 3.10. The network parameters were changed since the task we were training it for contained a smaller dataset. The first experiments of this work were for the classification of the patient's Risser signs. After several tests of optimization of the SVR classifiers hyperparameter, a mean square error of 0,35, a mean absolute error of 0,29 and an R^2 of 0,33. An optimization of the network and a pre-processing of the images will be done in the next phases of this project.

Key words: Bone age assessment, adolescent idiopathic scoliosis, X-ray, Deep learning, SVR, PCA

1. DESCRIPTION OF THE PURPOSE

Adolescent idiopathic scoliosis (AIS) is a spinal deformity affecting 1–3% of the population and female patient represent a larger ratio than males. The bone age assessment for the AIS patient is crucial from the age of 10 years old. Bone age assessment generally requires an additional radiography of the hand for the AIS patient, because the Risser stage has been proven to have a smaller correlation with the remaining growth potential, but it is not always feasible and involves a second radiograph. In this work, we use an automatic Risser stage classifier and use the model's last attention layer to build a longitudinal regression model to evaluate the changes in growth indications on AIS patient's pelvis. This could provide more insights on growth indicators on the pelvis and eliminate the need for a second radiography on AIS patients.

2. METHODS

a) Transfer learning to extract features

From the complete radiographs of a large dataset ($n=304$), the pelvic regions were selected with a size of 224×224 pixels. Each region of interest was associated with a Risser sign annotated by a clinician. To get a better image quality for the classification task, a contrast enhancement is applied to enhance features on the image for the future classification task. The preprocessed data were input in the network for a training. Following the work of (Kaddioui et al., 2019), deep features obtained by transfer learning of a VGG16 will be used to automatically classify the Risser stage of the patient's radiographs ($n=608$) since for each patient the two pelvic regions were selected. This classification method has been proven to get an overall accuracy of 78%. The classification models attention layers will then be extracted. For this study, the dataset was smaller than the one used to train the original network and for that reason, an augmentation was applied to the data to improve the network performances.

The dataset was divided in two, the training dataset will be used to train the convolutional neural network (CNN) and the support vector regression (SVR) model, the validation dataset will be used to evaluate the performance of the SVR model. First, the head of the network was trained with a reduced learning rate on 50 epochs. The features learned from the training dataset were then extracted from the attention layers of the CNN model.

b) Regression Model

To observe the correlation between the model features and the growth of the patient, on the longitudinal study, a support vector regression (SVR) model was trained. To increase the performance of the SVR model, the features from the CNN model and the labels of the data were scaled using the normalization technique. The normalization was applied to the labels to reduce the error possibility during the prediction since the labels would be between a continuous range of 0 to 1 instead of a discrete range of 0 to 4. The training was done with a grid search to get all the optimal parameters for the training. The different optimized parameters are presented in the table 2-1.

Table-A I-1 Parameters of the SVR optimizer

Hyperparameters	Values	Hyperparameters	Values
kernel	“linear”, “rbf”	C	0.5,1,2,3,4,5,10,20,50
Epsilon	0.001,0.01,0.1,0.2,0.3,0.4,0.5	Gamma	0.001,0.01,0.1,0.2,0.3,0.4,0.5

3. RESULTS

The transfer learning implementation was done using Python 3.10. The dataset contained 608 radiography of AIS patients. The dataset was divided in two different set one for the training of the models and the other one to test the SVR classifier after its training. To begin the training,

the first layers of the classifier were frozen to accommodate the smaller size of the dataset we used. The other parameters of the network have also been optimized to obtain a better performance of the network. The classifier was evaluated by obtaining its mean square error, its mean absolute error and its R^2 , which were respectively 0.35, 0.29 and 0.33. The future work will focus on the optimization of the network parameter to get a better classification accuracy.

4. NEW WORK TO BE PRESENTED

This study presented a new approach to characterize maturity in patients with AIS. Through the automatic classification of Risser stage, a statistical approach has been introduced to monitor different growth indicators. The introduction of the menarche in female maturation process and the identification of the growth spurt using pelvis radiographic and Risser stage is a novel technique. This technique is particularly interesting in the case of AIS because it can improve the Risser sign staging by adding additional indicator of growth. And for AIS patients, the elimination of the need to have a second radiography of the hand after their spine radiography reduce the exposure to radiation.

CONCLUSION

This study proposed an automatic method of assessment skeletal maturity in adolescent with idiopathic scoliosis using support vector regression on deep features. The first result helped us have an automatic assessment of the Risser stage of patients from an SVR classifier. The future works will focus on the optimization of the network parameters and the link between the attentions of the network and the Risser stage of AIS patients through their follow-up with SVR. This work has not been submitted elsewhere.

REFERENCES

- Automatic bone maturity grading from EOS radiographs in Adolescent Idiopathic Scoliosis | Elsevier Enhanced Reader. (s.d.). <https://doi.org/10.1016/j.compbimed.2021.104681>
- Cheung, J. P. Y., Cheung, P. W. H., Samartzis, D., & Luk, K. D.-K. (2018). APSS-ASJ Best Clinical Research Award: Predictability of Curve Progression in Adolescent Idiopathic Scoliosis Using the Distal Radius and Ulna Classification. *Asian Spine Journal*, 12(2), 202-213. <https://doi.org/10.4184/asj.2018.12.2.202>
- KADDIOUI, H. (2019). Apprentissage profond pour l'identification automatique de l'indice de Risser.
- Manzoor Mughal, A., Hassan, N., & Ahmed, A. (2014). Bone Age Assessment Methods: A Critical Review. *Pakistan Journal of Medical Sciences*, 30(1), 211-215. <https://doi.org/10.12669/pjms.301.4295>
- Minkara, A., Bainton, N., Tanaka, M., Kung, J., DeAllie, C., Khaleel, A., ... Roye, B. (2020). High Risk of Mismatch Between Sanders and Risser Staging in Adolescent Idiopathic Scoliosis: Are We Guiding Treatment Using the Wrong Classification? *Journal of Pediatric Orthopedics*, 40(2), 60-64. <https://doi.org/10.1097/BPO.0000000000001135>

ANNEXE II

Liste des paramètres extraits de BoneXpert

Tableau-A II-1: Liste des caractéristiques extraits de l'application BoneXpert

Variable	Explanation
Age	Chronological age
avL	Average length of the 19 bones after correcting for invalid lengths. 0 means there were too few valid lengths.
avW	Average width of the 19 short bones. The width of a bone is defined as the bone area divided by the bone length.
BA	Greulich Pyle bone age
BACarpal	Bone age of the carpal region
BASDS	Standard deviation score of the GP bone age
BAstatus	Code for errors in the BA determination (see Appendix A: List of errors and warnings)
BHI	Bone Health Index
BHISDS	Standard Deviation Score of the BHI
BHIsatus	Code for errors in the BHI determination (see Appendix A: List of errors and warnings)
Birth	Birth date, date format yyyyymmdd
Eth	Ethnicity used for the analysis
Fels	Fels bone age
filename12	Filename of the analysed image
HandVer	The version of the algorithm
hsize	The hand size (avL) in percent of what is expected for that bone age and sex for Caucasians
id11	ID number


Variable	Explanation
MCI	Another measure of the amount of cortical bone in the shafts of metacarpals 2-4. It has the virtue that it is independent of any magnification in the image, and therefore, the program automatically shows MCI instead of BHI, if the resolution is unknown. One can also force the display of MCI through View->BoneHealth
MCI SDS	Computed relative to children of the same bone age, gender and ethnicity. However, if bone age is above 17 years for boys and 15 years for girls, it is computed relative to children of the same age. In the present version, it is computed only for Caucasians
Meta2, meta3, ...	GP bone ages of 28 bones. -1 means missing bone age
meta2L, meta3L, etc	Bone lengths for the 19 short bones. For bones, where the epiphysis has not yet formed, the length is given as 1.08 times the actual length for metacarpal 2-5, and as 1.05 times the actual length for the phalanges and metacarpal 1.
metaL	The average length of metacarpal 2-4
metaL	The average length of metacarpals 2-4
metaT	The average cortical thickness in metacarpal 2-4
metaW	The average bone width in metacarpal 2-4
Mode	1=standard, 2=fast
no	The number of the image, also used to name the png files
noise	Image noise level
PatientName32Long	Patient name
R	0=left hand, 1=right hand
res	Image resolution
S	Gender: 1= boy, 2=girl

Variable	Explanation
sharp	Image sharpness. If >11 it means that the bone age (BA) value is invalid (for boys with BA < 5 and girls of BA < 4, the threshold is somewhat higher)
Study	Study date, date format yyymmdd
TW2	TW2 bone age
TW3	TW3 bone age
TW3x	TW3 bone age, extended up to 19 for boys and 18 for girls
TWJap	TW Japan bone age

ANNEXE III

Rapport de l'évaluation de l'âge osseux de l'outil BoneXpert

			BONE AGE
NAME	PATIENT ID
X-RAY DATE	BIRTH DATE	10101
CLINIC NAME	Name of Hospital	REPORT DATE	13/07/2023 12:04

	GP Bone Age Report:
	- An X-ray of the Left hand was analysed.
	- Chronological age could not be calculated from the DICOM file.
	- Greulich Pyle bone age according to BoneXpert is 7.88 years (7y 11m).
	TW Bone Age Report:
	- TW2 bone age according to BoneXpert is 8.47 years (8y 6m).
	- TW3 bone age according to BoneXpert is 7.57 years (7y 7m).
	- Adult height prediction based on the observed bone age can be computed by using the below link:
	https://www.bonexpert.com/ahp/#?gdr=Female&eth=CauAm&ba=7.88&bhi=4.10&mci=0.48
	Results:

Bone age (GP)	7.88 years
Bone age SDS	14.50 (CauAm)
Bone age (TW2/TW3)	8.47 years/7.57 years
Carpal bone age (GP)	not computed
Chronological Age
Gender	Female
Bone Health Index	4.10
BHI SDS	-

This report was generated automatically by BoneXpert.

BoneXpert

BoneXpert is a CE marked medical device made by Visiana, Denmark. Visit www.bonexpert.com
Version: 3.2.2/RS / Username: ETSMontreal2 / Image Sharpness: 7.41 / Population Height: 165.0

Figure-A III-1 : Rapport de la maturité osseuse d'un patient scoliotique de l'application BoneXpert

BIBLIOGRAPHIE

- Acheson, R. M. (1954). A method of assessing skeletal maturity from radiographs. *Journal of Anatomy*, 88(Pt 4), 498-508.
- Beheshtian, E., Putman, K., Santomartino, S. M., Parekh, V. S., & Yi, P. H. (2022). Generalizability and Bias in a Deep Learning Pediatric Bone Age Prediction Model Using Hand Radiographs. *Radiology*. <https://doi.org/10.1148/radiol.220505>
- Boeyer, M. E., Leary, E. V., Sherwood, R. J., & Duren, D. L. (2020). Evidence of the non-linear nature of skeletal maturation. *Archives of Disease in Childhood*, 105(7), 631-638. <https://doi.org/10.1136/archdischild-2019-317652>
- Chan, H.-P., Samala, R. K., Hadjiiski, L. M., & Zhou, C. (2020). Deep Learning in Medical Image Analysis. *Advances in experimental medicine and biology*, 1213, 3-21. https://doi.org/10.1007/978-3-030-33128-3_1
- Cheung, J. P. Y., Cheung, P. W. H., & Luk, K. D.-K. (2019). When Should We Wean Bracing for Adolescent Idiopathic Scoliosis? *Clinical Orthopaedics and Related Research*, 477(9), 2145-2157. <https://doi.org/10.1097/CORR.0000000000000781>
- Dimeglio, A., & Canavese, F. (2013). Progression or not progression? How to deal with adolescent idiopathic scoliosis during puberty. *Journal of children's orthopaedics*, 7(1), 43-49. <https://doi.org/10.1007/s11832-012-0463-6>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. <https://doi.org/10.1145/3422622>

- Guo, L., Wang, J., Teng, J., & Chen, Y. (2022). Bone Age Assessment Based on Deep Convolutional Features and Fast Extreme Learning Machine Algorithm. *Frontiers in Energy Research*, 9. Repéré à <https://www.frontiersin.org/articles/10.3389/fenrg.2021.813650>
- Hacquebord, J. H., & Leopold, S. S. (2012). In Brief: The Risser Classification: A Classic Tool for the Clinician Treating Adolescent Idiopathic Scoliosis. *Clinical Orthopaedics and Related Research*, 470(8), 2335-2338. <https://doi.org/10.1007/s11999-012-2371-y>
- Kaddioui, H., Duong, L., Joncas, J., Bellefleur, C., Nahle, I., Chémaly, O., Labelle, H. (2020). Convolutional Neural Networks for Automatic Risser Stage Assessment. *Radiology: Artificial Intelligence*, 2(3), e180063. <https://doi.org/10.1148/ryai.2020180063>
- Magnide, E., Tchaha, G. W., Joncas, J., Bellefleur, C., Barchi, S., Roy-Beaudry, M., Duong, L. (2021). Automatic bone maturity grading from EOS radiographs in Adolescent Idiopathic Scoliosis. *Computers in Biology and Medicine*, 136, 104681 <https://doi.org/10.1016/j.compbiomed.2021.104681>
- Mei, X., Liu, Z., Robson, P. M., Marinelli, B., Huang, M., Doshi, A., Yang, Y. (2022). RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning. *Radiology: Artificial Intelligence*. <https://doi.org/10.1148/ryai.210315>
- Minkara, A., Bainton, N., Tanaka, M., Kung, J., DeAllie, C., Khaleel, A., ... Roye, B. (2020). High Risk of Mismatch Between Sanders and Risser Staging in Adolescent Idiopathic Scoliosis: Are We Guiding Treatment Using the Wrong Classification? *Journal of Pediatric Orthopedics*, 40(2), 60-64. <https://doi.org/10.1097/BPO.0000000000001135>

- Neal, K. M., & Shirley, E. (2018). Skeletal Maturity Assessment for Idiopathic Scoliosis: Comparing the Sanders Maturity Scale with Risser Staging. *Pediatrics*, 141(1_MeetingAbstract), 647. <https://doi.org/10.1542/peds.141.1MA7.647>
- Pan, L., Pouyanfar, S., Chen, H., Qin, J., & Chen, S.-C. (2017). DeepFood: Automatic Multi-Class Classification of Food Ingredients Using Deep Learning. Dans 2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC) (pp. 181-189). <https://doi.org/10.1109/CIC.2017.00033>
- Physis Bone Age Calculator. (s.d.). Repéré à <https://physis.16bit.ai/>
- Prokop-Piotrkowska, M., Marszałek-Dziuba, K., Moszczyńska, E., Szalecki, M., & Jurkiewicz, E. (2021). Traditional and New Methods of Bone Age Assessment-An Overview. *Journal of Clinical Research in Pediatric Endocrinology*, 13(3), 251-262. <https://doi.org/10.4274/jcrpe.galenos.2020.2020.0091>
- Sanders, J. O., Khoury, J. G., Kishan, S., Browne, R. H., Mooney, J. F., Arnold, K. D., ... Finegold, D. N. (2008). Predicting scoliosis progression from skeletal maturity: a simplified classification during adolescence. *The Journal of Bone and Joint Surgery. American Volume*, 90(3), 540-553. <https://doi.org/10.2106/JBJS.G.00004>
- Sarkar, D. (DJ). (2018, 17 novembre). A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning. Medium. Repéré à <https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a>
- Thodberg, H. H., Kreiborg, S., Juul, A., & Pedersen, K. D. (2009). The BoneXpert method for automated determination of skeletal maturity. *IEEE transactions on medical imaging*, 28(1), 52-66. <https://doi.org/10.1109/TMI.2008.926067>

- Wan, J., Zhao, Y., Feng, Q., Lv, P., Hong, K., & Zhang, C. (2021). Statistical Confirmation of a Method of US Determination of Bone Age. *Radiology*, 300(1), 176-183. <https://doi.org/10.1148/radiol.2021204353>
- Wang, Z. J. (2021). Probing an AI regression model for hand bone age determination using gradient-based saliency mapping. *Scientific Reports*, 11(1), 10610. <https://doi.org/10.1038/s41598-021-90157-y>
- Yahara, Y., Tamura, M., Seki, S., Kondo, Y., Makino, H., Watanabe, K., Kawaguchi, Y. (2022). A deep convolutional neural network to predict the curve progression of adolescent idiopathic scoliosis: a pilot study. *BMC Musculoskeletal Disorders*, 23(1), 610. <https://doi.org/10.1186/s12891-022-05565-6>
- Yin Cheung, J. P., & Luk, K. D.-K. (2017). Managing the Pediatric Spine: Growth Assessment. *Asian Spine Journal*, 11(5), 804-816. <https://doi.org/10.4184/asj.2017.11.5.804>
- Zhang, J., Cheuk, K., Xu, L., Wang, Y., Feng, Z., Sit, T., ... Lee, W. Y. W. (2020). A validated composite model to predict risk of curve progression in adolescent idiopathic scoliosis. *EClinicalMedicine*, 18, 100236. <https://doi.org/10.1016/j.eclinm.2019.12.006>