



Master of Science conjoint HES-SO - UNIL en Sciences de la santé Av. de Provence 6 CH-1007 Lausanne

Master of Science conjoint HES-SO – UNIL en Sciences de la santé Orientation Technique en Radiologie Médicale

L'intelligence artificielle au service de la détection, de la position et du mouvement :

une assistance pour la technique en radiologie médicale ?

Fait par

Butt Meryam

Sous la direction de Prof. Jérôme Schmid, PhD, Haute école de santé Genève (HEdS), HES-SO (Président du Jury)

Experts Mme Caecilia Charbonnier, PhD, présidente de la fondation Artanim à Genève Mr Thomas Ramoussin, MSc, chef TRM au centre d'imagerie d'Onex

Lausanne, HES-SO Master, 2020

Remerciements

En premier lieu, je tiens à remercier très sincèrement mon directeur de mémoire, Prof. Jérôme Schmid pour la qualité de son encadrement, ses précieux conseils, son suivi tout au long de ce master, ses remarques pertinentes, ainsi que les heures passées à la relecture du mémoire.

Je remercie tous les membres du projet AIRx avec lesquels j'ai été amenée à collaborer de près ou de loin, avec une mention particulière à Amy Scheurer et Christophe Chênes de la Haute École de Santé de Genève, et Rémy Gardier de l'EPFL pour leur implication dans la réalisation de la partie expérimentale de ce travail. Je remercie les participants à l'étude sans qui je n'aurais pas pu réaliser ce travail.

Merci également à Madame Caecilia Charbonnier et Monsieur Thomas Ramoussin pour leur temps consacré à la lecture de ce mémoire et pour avoir accepté de faire partie de mon jury.

Je tiens à remercier et à exprimer ma profonde reconnaissance à mes chers collègues du Master pour l'ambiance qu'ils ont créé, leur motivation, leur bienveillance, leur réactivité, et encouragements, je ne sais pas ce que j'aurais fait sans vous les amis !

J'adresse enfin une pensée à tous mes collègues de la clinique Convert, pour leur bon sens, et pour m'avoir encouragée à poursuivre mes études.

Mes sincères remerciements à ma famille pour leur patience et leur soutien durant ce master, tout spécialement à mon père sans qui je n'aurais pas pu arriver jusqu'ici. Je remercie enfin tous ceux qui, de près ou de loin, ont contribué directement ou indirectement à ce travail et qui se reconnaîtront.

Table des matières

RE	REMERCIEMENTS					
LIS	STES DES	TABLEAUX	7			
LIS	STES DES	FIGURES	8			
LIS	LISTES DES ABRÉVIATIONS1					
RE						
ABSTRACT						
1	INTR		15			
1. 2	нун. Стат		10			
Ζ.	EIAI	DE LA LITTERATORE	10			
	2.1.	La physique des revens X	16			
	2.1.1	La positionnament des insidences et les critères de réussites	10 17			
	2.1.2	La fluorosconie	/۱ ۱۶			
	2.1.5	l'apprentissage de la radiographie	10 10			
	2.1.4		1J 21			
	2.2.	Architecture et classification	21			
	2.2.1	Canture de mouvement en médecine	23 24			
	2.2.2	IA annliquée à la canture de mouvement et ses métriques d'évaluation	27			
	2.2.4	Synthèse des différentes méthodes d'IA de détection de corps humain avec les principales	s limites			
		30				
3.	PROF		32			
	31		37			
	3.2		32 33			
_	0.2.		00			
4.	METI	HODE ET MATERIEL	34			
	4.1.	Preambule	34			
	4.2.	Methodes d'IA	34			
	4.3.	ÉTUDE PRELIMINAIRE	37			
	4.3.1	. Participants	37			
	4.3.2	Protocole	37			
	4.3.3	. Résultats et conclusion	37			
	4.4.	ÉTUDE PRINCIPALE	39			
	4.4.1	. Participants	39			
	4.4.2	. Caméras et calibration	39			
	4.4.3	. Acquisitions	41			
	4.5.	TRAITEMENTS DES DONNEES	43			
	4.5.1	. Annotation des points de repères	43			
	4.5.2	. Mesures statistiques	45			
	4.6.	GENERATION DES RADIOGRAPHIES VIRTUELLES	46			
5.	RESU	ILTATS	47			
	5.1.	DETECTION	47			
	5.1.1	. OpenPose versus Densepose	47			
	5.1.2	. Impact de la lumière	49			
	5.1.3	. Impact du type d'incidence	5 50			
	5.1.4	. Impact de la détection des incidences en fonction du genre du suiet	50			
	5.1.5	. Impact de la détection sur des bons et des mauvais positionnements radioaranhiaues	50			
	5.2.	RADIOGRAPHIES VIRTUELLES	50 51			
~						
6.	DISC	USSIUN	53			
	61	PERFORMANCES ET LIMITES DES METHODES DE DETECTION IA	53			

6.2.	UTILISATION DE CES METHODES POUR L'ENSEIGNEMENT	55
6.3.	UTILISATION DE CES METHODES POUR LA PRATIQUE CLINIQUE	56 57
6.4.		
6.5.	Aborder l'IA en radiologie	58
7. CONCLUSION		59
8. REF	. REFERENCES	
9. ANI		
9.1.	LISTES NON-EXHAUSTIVES DES METHODES IA DE DETECTION DE POSE HUMAINE	65
9.2.	PRINCIPALES BASES DE DONNEES POUR L'ESTIMATION DE POSE HUMAINE	69
9.3.	Formulaire de consentement	71
9.4.	DETAILS DES INCIDENCES RADIOLOGIQUES UTILISEES POUR L'ETUDE PRINCIPALE	72
9.5.	RADIOGRAPHIES VIRTUELLES	75

Listes des tableaux

Tableau 1 : Liste non-exhaustive des méthodes d'IA d'estimation de pose. Les références des bases de données se trouvent en Annexe 9.2. 30 Tableau 2 : résumé des éléments de notre base de données initiale 43 Tableau 3 : Détails de la réalisation des mauvais positionnements _____ 43 Tableau 4 : moyenne des métriques OKS et PED [m] pour les points d'OpenPose _____ 48 Tableau 5 : moyenne d'OKS et de PED [m] pour les points de Densepose. Le point de repère « RArmPit » n'as pas de valeur PED car il n'est jamais simultanément détecté dans deux vues en correspondance. _____ 49 Tableau 6 : moyennes des PED [m] en fonctions des incidences 50

 Tableau 7 : tableau des moyennes d'OKS et PED [m] pour les bons et mauvais positionnements
 51

 Tableau 8 : moyennes d'OKS et de PED en fonctions des points d'OpenPose et de Densepose 53

Tableau 9 : Le tableau présente les résultats en distance euclidienne (MPJPE) disponibles dans l'état de l'art. 54

 Tableau 10 : Comparaison de notre mAP exprimé en % par rapport à quelques méthodes d'IA publiées_____5

Tableau 11 : Bases de données publiques disponibles pour l'estimation de pose humaine (liste non-exhaustive)

69

Listes des figures

Figure 1 : schéma d'un tube à rayons X, consulté le 14/01/2020 http://www.cardio-paramed.com/fr/rayonsx.html 17 Figure 2 : fantôme anatomique d'enseignement consulté le 14.06.2020 https://www.medicalexpo.fr/prod/erlerzimmer/product-68400-885865.html 20 Figure 3 : exemple de simulation semi-virtuelle du projet VirtX [5] combinant une salle de simulation (gauche) et un logiciel de contrôle sur l'ordinateur (droite) ____ 21 Figure 4 : schéma montrant le positionnement des notions d'IA, d'apprentissage automatique, et 22 d'apprentissage profond imbriquées les unes aux autres Figure 5 : Schéma d'une architecture de réseaux neuronal profond, consulté le 9/06/20 à https://ch.mathworks.com/content/dam/mathworks/ebook/gated/80879v00_Deep_Learning_ebook.pdf ___ 23 *Figure 6 : Système de capture mécanique Gypsy, consulté le 07/06/2020* http://pubs.sciepub.com/ajme/1/7/12/figure/1 24 Figure 7 : Un danseur portant une combinaison avec des marqueurs passifs réfléchissants, lors d'une capture optique de mouvement. Consulté le 07/06/2020 https://fr.wikipedia.org/wiki/Capture_de_mouvement_____25 Figure 8 : Exemple de modélisation de squelette d'OpenPose consulté le 24 juillet 2020 https://github.com/CMU-Perceptual-Computing-Lab/openpose/blob/master/doc/output.md 27 Figure 9 : valeur du facteur ki pour le calcul de la métrique OKS_ 29 Figure 10 : quelques exemples d'estimation de pose humaine par le système DeepCut [51] 34 Figure 11 : Estimation de la posture humaine avec la méthode Vnect [40] et modélisation de ce squelette 35 Figure 12 : modélisation du sujet avec la méthode HMR [31] en utilisant un modèle 3D SMPL [36] 35 *Figure 13 : détection d'une personne la tête en bas avec la méthode OpenPose* [7] 36 Figure 14 : détection de plusieurs personnes sur diverses images avec la méthode OpenPose [7] 36 Figure 15 : Exemples de détection de plusieurs personnes sur une même image par la méthode Densepose [65] 37 Figure 16 : résultats des acquisitions de l'étude préliminaire avec les méthodes Vnect (A et B), HMR (C), DeepCut (D et E) 38 Figure 17 : résultats de la méthode OpenPose lors de l'étude préliminaire 38 Figure 18 : principe général de la calibration d'une caméra, tiré de http://wcours.gel.ulaval.ca/2017/a/GIF4100/default/5notes/A2016etalonnageCameraPageWeb.pdf 40 Figure 19 : Mire de calibration de type « ChArUco » acquise par les deux caméras (ici pour la calibration des paramètres extrinsèques). 41 Figure 20 : Graphique issu du travail de Bachelor de Laura Hirschi qui montre l'utilisation de la fluoroscopie pour les incidences du rachis 42 Figure 21 : placement des points de repères de notre modèle et tableau comparatif des points d'OpenPose versus nos points de repères. Les points de repères ajoutés pour notre étude sont en rouge et les points de repères barrés sont ceux supprimés de l'annotation car ils ne représentaient pas d'utilité. 44 Figure 22 : Capture d'écran de l'interface VIA lors de l'annotation manuelle d'une radiographie de thorax de face. Les points localisés en haut à gauche sont des points ne pouvant pas être localisés de façon fiable sur la 45 personne par l'opérateur. Figure 23 : exemple du maillage de Densepose pour les incidences du bassin de face (à gauche) et épaule interne (à droite). On peut observer des erreurs, typiquement au niveau du pied droit de la volontaire, ou au niveau du cou du volontaire. 47 Figure 24 : Exemple de notre base de données montrant la détection OpenPose avec lumière (A) et sans lumière (B) pour la même position radiologique. Pas de différences significatives de détection sont observées. ____ 49 Figure 25 : Exemple de squelette reconstruit pour une radiographie du bassin de face. Rangée supérieure : images du sujet acquises par les caméras. Rangée inférieure : squelette reconstruit (rouge) avec le cône de projection (jaune) et le détecteur (vert). Un modèle 3D estimé sur la base du squelette est aussi représenté. 51 Figure 26 : Exemple de radiographies virtuelles du bassin de face (80 kV, 50 mAs, 120cm, détecteur 43 x 43 cm). A gauche : pas de rayonnement diffusé simulé ; à droite : rayonnement diffusé simulé en présence d'une grille anti-diffusante ratio 12:1, focalisée à 120cm. 52 Figure 27 : Autres exemples de radiographies du bassin. A) 80 kV, 50 mAs, 120cm, détecteur 43 x 43 cm, pas de rayonnement diffusé ; B) 80 kV, 50 mAs, 120cm, détecteur 43 x 43 cm, avec du rayonnement diffusé ; C) 80 kV, 20 mAs, 120cm, détecteur 43 x 43 cm, avec du rayonnement diffusé ; C) 80 kV, 20 mAs, 120cm, détecteur 43 x

43 cm, avec du rayonnement diffusé ; D) 120 kV, 50 mAs, 120cm, détecteur 43 x 43 cm, avec de rayonnement diffusé. ______ 75

Figure 28 : Autre exemple de radiographie du bassin avec un CT en présence d'une prothèse de hanche droite. A noter que des hyperintensités sont présentes avec le fenêtrage choisi, notamment au niveau de la tête fémorale de l'implant. Ceux-ci sont en partie dus au fait que dans le CT des artéfacts métalliques sont présents et qu'ils influencent donc la radiographie virtuelle générée. 75

 Figure 29 : exemple de la détection et la génération d'une radiographie virtuelle du bassin de face sur un sujet

 masculin______76

 Figure 30 : exemple d'un sujet bien positionné pour une incidence de thorax de face (A) et mal positionné (B) à

 savoir le centrage est trop haut et le sujet n'est pas collé au détecteur

 77

Figure 31 : squelettes 3D reconstruits pour des radiographies du thorax de face. A) et C) représentent des vues pour un bon positionnement du sujet ; B) et D) représentent des vues pour un mauvais positionnement, dans lequel le centrage est trop haut et le sujet n'est pas assez collé au détecteur._____77

Figure 32 : deux exemples de radiographies virtuelles générées en fonction d'un bon positionnement (à gauche) et d'un mauvais positionnement (à droite). A noter que la zone coupée au-dessus de la radiographie virtuelle est due aux limites du champ d'acquisition du CT.______78

Listes des abréviations

ASTR	American Society of Radiology Technologists
ASTRM	Association Suisse des Techniciens en Radiologie Médicale
CAD	Computer Aided Diagnosis
CAR	Canadian Association of Radiology
CNN	Convolutional Neural Network
DAO	Diagnostic Assisté par Ordinateur
DL	Deep Learning ou Apprentissage Profond en français
GPS	Geodesic Point Similarity
HEdS	Haute École de Sante à Genève
IA	Intelligence Artificielle
kV	Kilo voltage
mA	Milliampères
mAP	Mean Average Precision
ML	Machine Learning ou Apprentissage Machine en français
МоСар	Motion Capture ou Capture de mouvement en français
MPJPE	Mean Per Joint Position Error
MSCSA	Master ès Sciences en Science de la Santé
OKS	Object Keypoint Similarity
PED	Point Euclidean Distance
РСК	Percentage Correct Keypoint
РСР	Percentage of Correct Parts
PDJ	Percentage of Detected Joints
SFR	Société Française de Radiologie
ТР	Travaux Pratiques
TRM	Technicien(ne) ou Technique en Radiologie Médicale
RCP	Ratio of Correct Point
VIA	VGG Image Annotator
VGG	Visual Geometry Group

Résumé

But de l'étude :

Le but de ce travail est d'étudier la performance de systèmes d'intelligence artificielle (IA) appliqués à l'estimation de pose humaine pour le positionnement des patients lors des radiographies. Dans ce contexte, ces systèmes peuvent réduire l'utilisation parfois abusive de la fluoroscopie, mais aussi offrir aux étudiants techniciens en radiologie médicale (TRM) un outil d'apprentissage réaliste et sécurisé afin d'évaluer leur pratique par le biais de la simulation et de la production de radiographies virtuelles.

Méthodes et matériels :

Après avoir identifié différents systèmes d'IA et rigoureusement sélectionné une série d'incidences, nous avons procédé avec cinq volontaires (trois femmes et deux hommes) à l'acquisition de 488 photos dans les différentes positions radiologiques.

L'évaluation de la performance de ces systèmes d'IA s'est basée sur différents indicateurs tels que la détection des points de repères en 2D et 3D par ces systèmes versus des annotations manuelles d'une TRM de formation, la sensibilité de ces systèmes à un champ de vision partiel, avec ou sans la présence de la lumière du collimateur, etc.

<u>Résultats :</u>

Parmi les systèmes testés, les systèmes OpenPose et Densepose ont eu des résultats encourageants. Avec une détection moyenne précise à 93% OpenPose semble être viable pour une utilisation sur des images dans un contexte d'apprentissage pour certaines incidences. En revanche nous avons mesuré une différence spatiale d'environ 3 cm avec l'estimation d'une TRM. En comparaison, les autres systèmes d'IA testés ont généralement échoué à détecter les sujets dans les conditions d'acquisitions radiographiques.

Conclusion :

Bien que l'intelligence artificielle puisse être utile aux TRMs pour le positionnement des patients, les systèmes d'IA actuel pour l'estimation de pause humaine ne sont pas assez fiables et avancés pour être utilisés dans un cadre clinique, mais ils offrent un grand potentiel pour l'enseignement.

Cependant, la technologie de l'IA progresse rapidement et les TRM pourraient contribuer à l'amélioration des systèmes de détection de la pose humaine à usage médical.

<u>Mots clés</u>: intelligence artificielle, positionnement, détection, radiographies virtuelles, enseignement

Abstract

Purpose:

This study investigates the use of artificial intelligence (AI) systems originally designed to detect human pose in everyday color photographs, in the context of radiography for patient positioning. Such systems have the potential to reduce the reliance on fluoroscopy – too often used for some radiographs. Similarly, AI could support innovative teaching tools where radiographer students could safely assess the quality of subject positioning with e.g. the generation of virtual radiographs.

Methods and Materials:

We selected a series of radiographic exams based on their clinical relevance and conducted a study assessing the major limits of different publicly available AI systems for pose detection. 488 photographs of 5 volunteers (2 males and 3 females) were taken in a radiographic setup and were provided to the selected AI systems. We assessed their performance with various indicators such as 2D and 3D joint detection, robustness to partial field of view, presence or absence of collimation light, etc.

Results:

Among the tested AI systems only OpenPose and Densepose consistently produced satisfactory results. With an average precision of 93% OpenPose was notably robust although the detection lacked some accuracy (3 cm of difference between the system and the radiographer tagging). The remaining systems usually failed in detecting human pose in conditions of radiographic acquisitions.

Conclusion:

Although AI could be helpful for radiographers in patient positioning, current AI systems for human detection are not mature enough to be used in a clinical setup but they offer a great potential for teaching. However, AI technology is quickly progressing, and radiographers could contribute to the improvement of human pose detection systems for medical use.

Key words: artificial intelligence, positioning, detection, virtual radiographs, teaching tools

Dans le présent document, le genre masculin est utilisé au sens neutre et désigne les femmes autant que les hommes.

1. Introduction

Depuis son début en 1895 par le physicien allemand Wilhelm Röntgen, la radiologie ne cesse d'améliorer sa technique, ses appareillages et ses performances. Totalement informatisée et numérisée en Suisse, et domaine de médecine hautement technologique, la radiologie s'est très tôt intéressée à l'intelligence artificielle (IA) qui lui offre des possibilités remarquables.

Le terme « intelligence artificielle » est officialisé comme étant un domaine scientifique en 1956 lors de la conférence de Dartmouth¹ (Dartmouth Summer Research Project on Artificial Intelligence) aux États-Unis, mais les débuts de l'IA datent des années cinquante grâce au mathématicien Alan Turing et son article sur la conscience d'une machine [66].

L'intelligence artificielle a connu des succès et beaucoup d'échecs mais depuis les années 2000, elle représente un enjeu pour la société dans de nombreux domaines comme la cinématographie, les jeux-vidéo, la médecine, et notamment dans le diagnostic radiologique à travers par exemple la reconnaissance des lésions pathologiques, le dépistage du cancer du sein [10], des poumons, ou de la prostate.

Cette recherche est réalisée dans le cadre de la formation Master ès Sciences en Sciences de la Santé (MScSa) avec orientation Technique en Radiologie Médicale (TRM) et porte sur l'intelligence artificielle en imagerie médicale. Plus précisément, nous allons étudier l'IA au service de la détection et du positionnement en radiologie conventionnelle, comme assistance pour les TRMs en formation lors de l'apprentissage et la réalisation de radiographies.

La première partie de ce travail va introduire quelques concepts entourant la radiologie ainsi que l'intelligence artificielle, s'en suivra une description de la partie exploratoire pour aboutir aux résultats de cette recherche. Pour finir nous terminerons par vous montrer une application concrète de ce travail de master pour le projet AIRx, un projet qui a pour mission d'apporter aux étudiants TRMs un outil de simulation à des fin d'apprentissage, en générant des radiographies virtuelle grâce à ces systèmes d'estimation de pose humaine.

¹ <u>https://www.dartmouth.edu/~ai50/homepage.html</u>

2. État de la littérature

2.1. La radiologie

Avec plus de 3,6 milliards d'examens de radiodiagnostic dans le monde [48], la radiographie conventionnelle est un outil diagnostique important dans la prise en charge en médecine de premier recours. Le développement des appareils de radiologies qui sont de plus en plus performants ont permis de réduire les doses de radiations ionisantes et d'améliorer la qualité des images. Néanmoins, pour obtenir des images de bonne qualité tout tant ayant une faible irradiation du patient, il faut respecter certaines conditions :

• une bonne formation des TRMs dans la réalisation des radiographies,

• un appareillage d'imagerie sensible avec notamment un bon diaphragme,

• des contrôles qualité réguliers pour assurer la sécurité des patients ainsi que celle des professionnels de santé.

Pour le professionnel, la réalisation de radiographie nécessite des compétences spécifiques en radioprotection, de bonnes connaissances en anatomie humaine et un savoir-faire en matière de positionnement radiologique.

2.1.1. La physique des rayons X

Une radiographie repose sur l'utilisation de rayons X et est réalisée grâce à deux appareils, un tube qui génère ces rayons (Figure 1) et un détecteur. Le tube à rayons X libère des électrons à partir de la cathode, qui sont accélérés vers l'anode. La collision entre l'émission d'électrons chargés négativement et la cible de tungstène de l'anode chargée positivement produit des photons de rayons X selon deux phénomènes :

- le freinage des électrons par les atomes de la cible crée un rayonnement continu (rayonnement de freinage ou Bremsstrahlung) dont une partie de rayons X
- l'énergie des électrons accélérés suffit à exciter les atomes de la cible et perturbe leurs couches électronique internes. Ces atomes excités diffusent des rayons X en retournant à leur état fondamental.

Ils sont ensuite dirigés vers le détecteur grâce à la collimation du tube. Un cliché radiologique est formé lorsque ces photons sont arrêtés par le détecteur après avoir traversé le patient.

Les deux paramètres importants d'un tube à rayons X et que l'on peut modifier en tant que TRM sont :

- La différence de potentiel entre cathode et anode réglé par le choix du kilovoltage (kV)
- Le nombre d'électrons traversant le tube à chaque seconde, traduisant l'intensité du courant mesurée en milliampères (mA) et qui est directement proportionnelle à la quantité de photons X produits.

La quantité de photons X absorbés dépend de la composition des tissus traversés. Par exemple, plus les tissus sont denses (ex. les os) plus ils arrêtent les photons [21]. Sans risque à faible dose, les rayonnements ionisants peuvent s'avérer dangereux à haute dose car l'énergie peut ioniser les atomes en arrachant leurs électrons et, ainsi, endommager les tissus vivants.



Figure 1 : schéma d'un tube à rayons X, consulté le 14/01/2020 http://www.cardio-paramed.com/fr/rayons-x.html

2.1.2. Le positionnement des incidences et les critères de réussites

Pour que le diagnostic soit satisfaisant, les TRMs doivent assurer une reproductibilité du positionnement pour que le médecin radiologue puisse lire correctement les clichés radiographiques et les comparer de façon efficace lors du suivi d'un patient. En général, pour chaque partie du corps, trois incidences sont essentielles, l'incidence de face, l'incidence de profil et celle de ³/₄. Cependant en fonction de la région du corps radiographiée, le positionnement est différent, de ce fait, chaque incidence est protocolée et varie en fonction

des repères anatomiques du patient, des repères spatiaux par rapport au détecteur, et du centrage du tube. Le point de centrage ainsi que la collimation sont très importants dans la réalisation d'un cliché radiologique. Ils permettent de limiter l'irradiation sur la partie du corps étudié.

Les repères anatomiques sont des repères palpables ou visuels du corps humain, pour permettre au TRM de positionner le patient. Certaines incidences nécessitent également des angles patient-détecteur dans le but d'orienter des structures osseuses, ou des rotations des membres comme celle des pieds vers l'intérieur lors d'une radiographie du bassin de face afin de corriger l'antéversion des cols fémoraux [4].

Le positionnement de la radiographie est considéré comme une réussite lorsque tous ces critères sont respectés.

Dans un second temps pour juger de la qualité de la radiographie, des critères de réussites sont définis. Pour le positionnement, en général, ce sont des structures anatomiques qui doivent être visualisées sur le cliché final, la superposition de deux détails anatomiques ou la forme et l'emplacement de certains os du corps par rapport à d'autres structures. Par exemple pour les cervicales de profil, on doit visualiser toutes les cervicales, des conduits auditifs externes jusqu'à C7, les disques vertébraux doivent être enfilés, et les articulations zygapophysaires superposées [4]. Pour la qualité de l'image, les critères de réussites sont liés aux paramètres techniques qui pour simplifier doivent refléter un bon équilibre entre la résolution spatiale (la capacité à distinguer des structures fines dans l'image) et le contraste, le rapport signal sur bruit.

2.1.3. La fluoroscopie

Même s'il s'agit d'un examen médical fréquemment pratiqué, une radiographie n'est pas totalement sans risque, puisqu'elle repose sur l'utilisation des rayons X. Les TRMs effectuent les radiographies en manipulant le patient, en palpant les repères anatomiques et le centrent à l'aide de la lumière de la collimation. Cependant ils peuvent avoir recours à la fluoroscopie pour contrôler le positionnement avant de prendre le cliché. Le terme fluoroscopie désigne l'utilisation des techniques d'imagerie par rayons X pour visualiser l'anatomie interne de façon dynamique et en temps réel [32].

Au départ, inventée essentiellement pour une utilisation en radiologie interventionnelle [38], la fluoroscopie est maintenant de plus en plus courante en radiologie conventionnelle. Elle est

employée pour les examens tels que les arthrographies ou les transits œsophagiens et lavements barytés. Cependant en pratique, on constate qu'elle est aussi utilisée pour positionner le patient depuis le poste de commande, essentiellement pour les radiographies où les critères de réussites dépendent beaucoup de la morphologie du patient [55], pour contrôler le positionnement avant de prendre le cliché et parfois même pour effectuer le centrage du patient.

Sans surprises, la dose à la peau augmente en réalisant un cliché avec l'aide de la fluoroscopie comparé à un cliché avec un positionnement à l'aveugle [20]. Malgré cela, une étude a montré que l'utilisation de la fluoroscopie pouvait être justifiée pour certaines incidences complexes, par exemple l'incidence de profil L5-S1 [55]. D'après Statkiewicz-Sherer et al. [57], sans fluoroscopie, environ 8% des radiographies sont répétées, ce qui veut dire que si l'utilisation de la fluoroscopie serait recommandée pour vérifier le positionnement des incidences avant de prendre le cliché final, environ 92% des patients seraient inutilement surexposés.

C'est pour cela qu'il existe des règles et recommandations de radioprotection misent en place par les différents organismes internationaux de radioprotection, comme la Société Française de Radiologie (SFR), l'American Society of Radiology Technologists (ASRT), ou la Canadian Association of Radiology (CAR) qui découragent fortement l'utilisation de la fluoroscopie mais ne l'interdisent pas pour autant [22].

2.1.4. L'apprentissage de la radiographie

Dès la première année de formation pour devenir TRM, en complément des apports théoriques, les élèves réalisent des travaux pratiques (TP) pour l'apprentissage des positionnement radiologiques. De la simulation à l'aide des camarades de classe dans des salles de radiologies fictives, à la pratique sur le terrain en situation réelle, l'enseignement se présente sous différentes formes selon l'organisation et les moyens de l'établissement de formation.

En ce qui concerne les stages sur le terrain, cette pratique est très formatrice pour les étudiants qui se rendent mieux compte des enjeux nécessaire à la réalisation des radiographies. En effet, avec une grande diversité de patients, des contraintes de temps, et l'avantage d'avoir de vrais clichés radiologiques fidèles aux patients, la réalité du terrain leur permet de mieux se préparer à leur futur professionnel. Malgré ces nombreux points positifs, la pratique en établissement de santé peut être une source d'angoisse pour l'étudiant s'il n'a pas eu assez d'entrainement préliminaire, d'où la nécessiter de réaliser des TP à l'école, avec un formateur qualifié et pédagogue. L'étudiant apprend et intègre les informations à son rythme et non au rythme du terrain qui peut parfois être intense. Les simulations sont exercées sur des mannequins, des fantômes anatomiques (Figure 2), ou des camarades de classes, mais pour des raisons d'éthique et de sécurité, cela ne suffit pas à reproduire des situations réelles.



Figure 2 : fantôme anatomique d'enseignement consulté le 14.06.2020 https://www.medicalexpo.fr/prod/erler-zimmer/product-68400-885865.html

Pour combler le manque entre l'apprentissage à l'école et l'apprentissage sur le terrain, de nombreux logiciels de simulation virtuelle et semi-virtuelle ont fait leur apparition dans l'enseignement radiologique [5]. Avec la possibilité de varier les scénarios, l'étudiant choisit la position du patient, ainsi que les paramètres techniques de l'appareil d'imagerie pour avoir une radiographie plus proche de la réalité.

Les logiciels de simulation virtuelle répliquent un scénario de positionnement radiologique de façon 100% virtuelle [59][39], les étudiants interagissant avec le patient et l'appareillage à travers des interfaces traditionnelles telles que le clavier et la souris. En comparaison, les logiciels de simulation semi-virtuelle combinent une partie pratique, tel que le positionnement du patient grâce à un mannequin, et une partie sur ordinateur, par exemple le choix des constantes et l'évaluation de la radiographie virtuelle (Figure 3) Dans le cas de virtX, le logiciel de simulation de Bott, O. J. et al. [5] l'étudiant va d'abord positionner le mannequin dans la

salle de radiologie fictive, puis reviendra à la console pour régler les paramètres techniques comme il le ferait sur le terrain. Le mannequin, la table de radio ainsi que l'appareil d'imagerie sont équipés de capteurs de mouvement reliés à un système de suivi qui permet à l'ordinateur de faire correspondre les objets de la salle de radiologie aux objets de l'espace virtuel.

L'avantage d'avoir ces deux parties pratiques et virtuelle font de ces logiciels de simulation semi-virtuelle des méthodes d'apprentissage efficaces car elles sont plus proches de la réalité et n'utilisent pas de rayons X, respectant ainsi les besoins et la sécurité des étudiants. Toutefois ces équipements étant très onéreux, tous les centres de formation TRM ne peuvent pas forcément se permettre d'investir dans ces systèmes.



Figure 3 : exemple de simulation semi-virtuelle du projet VirtX [5] combinant une salle de simulation (gauche) et un logiciel de contrôle sur l'ordinateur (droite)

2.2. L'intelligence artificielle et sa terminologie

Nécessitant des ressources matérielles élevées pour faire tourner les algorithmes, l'IA est longtemps restée en sommeil limitée par les ressources de l'époque. Mais depuis les années 2010, la combinaison des composants informatiques de plus en plus puissants et moins chers ainsi que l'accès à un grand volume de données ont permis aux projets de recherche et développement de voir le jour plus facilement, et l'IA est aujourd'hui en pleine essor. Souvent utilisée dans l'automatisation des processus et dans diverses applications informatiques, l'IA est notamment portée par l'apprentissage profond et l'apprentissage machine. Ces termes peuvent parfois porter à confusion, il nous semble nécessaire d'expliquer ces termes généraux avant de continuer la lecture et rentrer dans les détails (Figure 4).

L'intelligence artificielle est un terme général qui désigne la branche de l'informatique consacrée au développement d'algorithmes pour accomplir des taches associées à l'intelligence humaine, telles que la capacité à apprendre, résoudre des problèmes, la reconnaissance faciale ou d'objet.

L'apprentissage automatique/machine ou Machine Learning en anglais (ML) est un sous ensemble de l'IA et est une forme d'apprentissage qui tente de fournir des connaissances aux systèmes sans être explicitement programmée et d'améliorer continuellement ses performances à résoudre des tâches en fonction des données traitées. En général, les deux principaux types d'algorithme ML utilisés aujourd'hui sont:

- l'apprentissage supervisé, avec ce modèle une intervention humaine guide et enseigne à l'algorithme les conclusions qu'il doit tirer avec une base de données étiquetées.
- l'apprentissage non supervisé, l'algorithme apprend à identifier lui-même la tâche demandée ou de façon très faiblement guidée.

L'apprentissage profond ou Deep Learning (DL) est l'une des principales approches de l'apprentissage automatique qui s'affranchit de l'exigence de l'apprentissage machine traditionnel consistant à ce que des « caractéristiques » (*features* en anglais) soient définies et extraites à priori. La définition de ces caractéristiques est souvent faite de façon manuelle et par des experts. Les méthodes de DL déduisent de façon autonome quelles caractéristiques sont pertinentes dans les données afin de mieux répondre à la tâche demandée. Le DL est aujourd'hui basé sur les réseaux de neurones organisé en couches et modélisé par des milliers voire des millions de paramètres.



Figure 4 : schéma montrant le positionnement des notions d'IA, d'apprentissage automatique, et d'apprentissage profond imbriquées les unes aux autres

Étant donné que l'IA moderne est en grande partie soutenue par l'apprentissage profond, nous allons nous concentrer sur cette branche d'IA pour la suite de ce travail.

2.2.1. Architecture et classification

Au cœur des évolutions technologiques, l'intelligence artificielle basée sur l'apprentissage profond s'est révélée dans plusieurs applications bien plus performante que l'être humain, surtout pour classifier ou segmenter des structures. Les algorithmes d'apprentissage profond comme par exemple les réseaux neuronaux convolutifs (convolutional neural network en anglais (CNN)) ont franchi une étape importante dans la détection des images. Grâce à une base de données contenant des milliers d'images, ce système apprend à reconnaitre une structure, ses contours, sa forme.

Trois éléments composent un CNN, l'entrée, les couches et la sortie (Figure 5). Les éléments d'entrée sont les données fournies au réseau. Les éléments de sortie représentent le type de réponse demandé au réseau. Les couches intermédiaires du CNN sont plus complexes. Généralement, on distingue quatre types de couches : la couche de convolution, la couche de regroupement (pooling), la couche de rectification d'unité linéaire (ReLU), et la couche entièrement connectée. Chaque type de couche a une fonction précise et chaque couche du CNN traite une information en utilisant des différents types d'opérations mathématiques et de traitement d'image (filtre). Chaque entrée est connectée à plusieurs neurones de la couche suivante par des poids.

L'entrainement du réseau se fait essentiellement par itérations successives en cherchant à minimiser une fonction dite « d'erreur » ce qui traduit la capacité du réseau à apprendre quels poids sont les plus adaptés pour répondre à la tache demandée.



Figure 5 : Schéma d'une architecture de réseaux neuronal profond, consulté le 9/06/20 à https://ch.mathworks.com/content/dam/mathworks/ebook/gated/80879v00_Deep_Learning_ebook.pdf

Plusieurs types de tâches peuvent être demandées au réseau, par exemple:

- La segmentation comme pour l'aide au diagnostic assisté (Computer aided diagnosis en anglais (CAD)) avec la segmentation automatisée des structures pulmonaires sur les tomodensitométries thoracique [34][10]
- La classification comme pour distinguer les masses mammaires malignes des masses bénignes grâce au CAD en mammographie [70]
- La détection de points comme pour les systèmes d'estimation de pose avec la détection des articulations dans l'image [7][65]

2.2.2. Capture de mouvement en médecine

La capture de mouvement (ou motion capture ou MoCap) est une technique qui consiste à déterminer les mouvements d'un objet ou analyser les déplacements d'un corps humain dans l'espace, et de les restituer sur un ordinateur. Plus connus dans le monde des jeux vidéo et de l'industrie cinématographique, les applications de la capture de mouvement sont cependant très variées.

Il existe plusieurs techniques de captures de mouvement, dont les plus populaires :

La capture mécanique. Cette technique se fait à l'aide d'un exosquelette mécanique attaché sur la personne qui permet d'enregistrer ses mouvements grâce aux capteurs intégrés à la combinaison (Figure 6). Ce système étant précis il a tout de même l'inconvénient d'être très encombrant ce qui limite son utilisation.



Figure 6 : Système de capture mécanique Gypsy, consulté le 07/06/2020 http://pubs.sciepub.com/ajme/1/7/12/figure/1

La capture magnétique. Des capteurs magnétiques sont placés sur les différentes articulations et parties du corps de la personne à enregistrer. L'acquisition et les données sont traitées en temps réels et contrairement aux systèmes mécaniques, il n'y a aucune gêne pour la personne en mouvement qui porte la combinaison, excepté le câblage qui relie tous les capteurs. Il faut tout de même faire attention aux perturbations électromagnétiques qui pourraient fausser les calculs et donc fausser les résultats.

La capture optique avec marqueurs. Ce système utilise des marqueurs réfléchissants et des caméras infrarouges disposées tout autour de la zone de capture (Figure 7). Lorsque le participant entre dans cette zone, les caméras détectent les marqueurs de la combinaison et renvoient l'information à un ordinateur qui retranscrit leurs positions dans un espace virtuel. Cette technologie est généralement précise, temps réel et bien maitrisée mais peut nécessiter un équipement et des logiciels onéreux, ainsi que le placement de marqueurs sur le corps du sujet ou le port d'une combinaison.



Figure 7 : Un danseur portant une combinaison avec des marqueurs passifs réfléchissants, lors d'une capture optique de mouvement. Consulté le 07/06/2020 https://fr.wikipedia.org/wiki/Capture_de_mouvement

La caméra Kinect. Une technologie très populaire dans le domaine ludique ouvert au public. Avec un principe de fonctionnement différent des précédentes techniques, elle nécessite qu'une seule caméra et n'a pas besoin de marqueurs. Cette technologie optique utilise des faisceaux lumineux infrarouges (IR) qui sont émis dans l'environnement de la scène observé. Ce système émet un motif IR spécifique et selon la déformation sur les lignes du motif on peut estimer les mouvements du corps.

Répandue également dans le milieu médical, la capture de mouvement est utilisée dans le but d'améliorer le diagnostic et d'assurer un meilleur traitement aux patients, grâce notamment à l'étude de la physiologie articulaire à différents niveaux, par exemple en chirurgie pour les prothèses articulaires [56], en physiothérapie elle vise à aider les professionnels de santé à parfaire leurs exercices en évaluant l'état et le progrès des patients en rééducation [53]

A notre connaissance, il n'y a, à ce jour, aucune application utilisant la capture de mouvement liée à la radiologie conventionnelle et au positionnement du patient.

2.2.3. IA appliquée à la capture de mouvement et ses métriques d'évaluation

L'estimation de la position humaine vise à estimer la position des articulations d'une personne sur une photo ou une vidéo puis de les relier pour créer un squelette graphique. Nous pouvons classer les systèmes de capture de mouvement en fonction de différents critères :

Les modalités d'entrées. Sous trois principales formes, les images fournies aux systèmes peuvent être des images RVB [31], des images de profondeurs, ou des images infrarouges. Ces modalités d'entrée peuvent être statiques [40] ou dynamique (vidéo).

La modélisation du corps à la sortie. Le squelette de la capture de mouvement [36] représente l'orientation d'une personne sous un format graphique simple avec N points de repère ou articulations [7], qui s'accorde à l'algorithme du système d'IA (par ex. Figure 8). Il s'agit d'un ensemble de coordonnées qui peuvent être reliées pour décrire la posture de la personne. Chaque coordonnée représente un point de repère, une articulation, ou une partie anatomique (par ex. coude, œil, cheville), et il est possible de relier ces points de repère pour former un membre (par ex. le membre supérieur avec l'ensemble poignet-coude-épaule).



Figure 8 : Exemple de modélisation de squelette d'OpenPose consulté le 24 juillet 2020 https://github.com/CMU-Perceptual-Computing-Lab/openpose/blob/master/doc/output.md

La base de données. Une des raisons principales du succès de ces systèmes d'estimation de pose sont les bases de données [86][83] utilisées pour entraîner et valider les approches développées (Annexe 9.2). En effet, grâce à la disponibilité de grandes quantités de données annotées pour l'entrainement, les systèmes d'apprentissage profond se sont développés. Ces bases de données comprennent un large panel d'images de diverses activités comme le sport, les loisirs, la mode, ou des photos de groupes et de rassemblement.

L'estimation de la pose d'une ou plusieurs personnes. L'approche la plus facile de la détection de corps humain est lorsqu'une seule personne est présente dans l'image.

Les méthodes de détection de pose étaient initialement axées sur la détection d'une seule personne [40][72] mais avec la disponibilité d'énormes bases de données multi-personnes [80][86], le problème de l'estimation de plusieurs corps à récemment fait l'objet d'attention accrue [7][51]. Ce scénario est plus proche de la réalité car le plus souvent il y a plusieurs personnes sur une image. En revanche en radiologie, seulement une personne est radiographiée, à l'exception de la pédiatrie où l'aide d'un parent est souvent sollicitée pour réaliser la radiographie de l'enfant.

Il y a deux manières d'estimer la position de plusieurs personnes :

- La méthode « Top-down » [45] : qui consiste en premier temps à créer des boites de délimitation pour chaque corps présent sur l'image. Ensuite chaque boite est traitée pour estimer la posture humaine.
- La méthode « Bottom-Up » [51] : qui consiste à repérer toutes les articulations de tous les corps présents sur l'image puis à regrouper et associer ces articulations à des personnes distinctes.

2D versus 3D. L'estimation de la position du corps humain en 2D [7] consiste à prédire l'emplacement des points de repère qui correspondent aux articulations du corps dans l'image. D'autre part, l'estimation de la pose en 3D [18] correspond à la configuration spatiale tridimensionnelle de tous les points de repères. La plupart des systèmes 3D prédisent d'abord une pose 2D puis essaient de la faire correspondre à une position 3D dans l'espace dans un repère de référence défini de façon arbitraire (par ex. l'espace 3D des données avant la projection perspective de la caméra), d'autres estiment directement la positon 3D des points de repère.

Évaluation de la prédiction des points de repères. Il existe différentes métriques pour évaluer ces systèmes d'IA appliqués à la capture de mouvement, ce sont des méthodes de calcul qui permettent de mesurer la performance de ces algorithmes. La position d'une articulation est considérée comme correcte si la distance entre le point de repère prédit par l'IA et le point de repère réel est inférieure à un certain seuil (dépendant de la métrique utilisée). Au vu des nombreuses IA, il existe une multitude de métriques pour l'évaluation de ces points de repères, telles que les métriques « Object Keypoint Similarity » (OKS) [7], Point Euclidean Distance (PED) et mAP (mean Average Precision).

La métrique **Object Keypoint Similarity** (OKS) comprise entre 0 et 1 mesure la similitude entre le point réel cliqué et le point détecté par le système d'IA en fonction de la taille du personnage sur l'image. Plus la détection du point de repère est précise plus l'OKS tend vers 1. Le facteur k_i dépend de l'articulation et de la précision à déterminer manuellement la position du centre de l'articulation. Plus le point de repère sera facile à déterminer plus le facteur k_i sera petit.

$$\mathsf{OKS} = \exp(-\frac{d_i^2}{2s^2k_i^2})$$

Avec :

- *d_i* distance Euclidienne entre le point de repère réel et le point de repère prédit
- s^2 surface de la boîte englobant les points détectés (*bounding box*)
- k_i facteur du point de repère (Figure 9)



Figure 9 : valeur du facteur k_i pour le calcul de la métrique OKS

Le facteur k_i a été calculé en mesurant l'écart type pour chaque point de repères (joints articulaires) en se basant sur 5000 images annotées aléatoirement par plusieurs personnes². De manière général, les points de repères tels que le genou, la hanche et l'épaule ont tendance à avoir un écart type plus grand que les yeux ou le nez qui sont plus faciles à annoter sur une personne.

La métrique **mAP (Mean Average Precision)**, proposée par Shotton et al. [58], permet de quantifier la précision de l'estimation. L'Average Precision (AP) représente le ratio des points de repères prédits avec la métrique OKS au-dessus d'un certain seuil. En variant ce seuil entre 0.5 et 0.95 tous les 0.05, on peut calculer le mAP comme la moyenne des AP calculés pour tous les seuils et tous les points détectés. Plus la valeur mAP se rapproche de 1, plus l'estimation est précise.

Point Euclidean Distance (PED) est la distance Euclidienne usuelle entre deux points de l'espace 3D. Plus cette distance est petite, plus la détection est performante. Dans notre travail, l'unité est le mètre.

Pour information, la métrique PED est équivalente à la métrique MPJPE (Mean Per Joint Position Error) que l'on retrouve dans de nombreux articles et qui nous servira pour comparer les résultats de notre travail avec ceux des articles du Tableau 9.

² <u>https://cocodataset.org/#keypoints-eval</u>

2.2.4. Synthèse des différentes méthodes d'IA de détection de corps humain avec les principales limites

Dans cette section, nous présentons quelques méthodes d'IA (Tableau 1) car il en existe beaucoup d'autres qu'elles soient 2D, 3D, ou avec détection de plusieurs personnes ou non. Une liste plus exhaustive des systèmes existants est proposée en Annexe 9.1

Les méthodes Vnect [40], HMR [31], OpenPose [7], DeepCut [51] et Densepose [18] seront présentées dans la section 4.2.

Tableau 1 : Liste non-exhaustive des méthodes d'IA d'estimation de pose. Les références des bases de données se trouvent en Annexe 9.2.

	Méthode	Base de données	
		2D	
	Iskakov K. el al. [28]	Human3.6m	
	Deeply Learned Compositional	FLIC, LSP, MPII	
	Models [62]		
	Convolutional Pose Machines [72]	MPII, LSP, FLIC	
Détection de personne	Iterative Error Feedback [8]	MPII , LSP	
seule	3D		
	Vnect [40]	MPII, LSP, Human3.6m, MPI-INF-3DHP	
	Human Mesh Recovery [31]	MS COCO, LPS, LPSET, MPII, Human3.6m,	
		MPI-INF-3DHP	
	Pavlakos G. et al. [49]	Human3.6m, KTH Multiview Football II	
	Tome D. et al. [64]	Human3.6m	
	Kostrikov I. et al. [35]	Human3.6m	
	2D		
	Regional Multi-person Pose	MPII, MS COCO,	
	Estimation [13]		
	OpenPose [7]	MS COCO, MPII	
Détection de plusieurs	DeepCut [51]	LSP, LSPET, MPII	
personnes	Generative Partition Networks [43]	WAF	
	Single-Stage Multi-Person Pose	MPII, MS COCO	
	Machines [44]		
	3D		
	Densepose [18]	MS COCO, COCO Densepose	

Pour qu'un système d'estimation de pose soit performant, il doit être robuste aux nombreuses sources de variations qui peuvent affecter une image [74] :

Les vêtements : les personnes portent une large variété d'habits avec des matières, des motifs et des couleurs différentes.

La luminosité : nous pouvons imaginer que les conditions d'éclairage des salles de radiologie peuvent perturber ces systèmes d'IA, en raison de la lumière de la collimation, du laser du tube ou de la position des sources de lumières.

L'occlusion : dans le cas de l'estimation de pose humaine, l'auto-occlusion se produit souvent.

L'arrière-plan : Il est probable que l'arrière-plan de l'image soit encombré par ses objets irréguliers ou complexes. Ces objets ont tendance à détourner l'attention des systèmes d'IA. Dans le contexte de la radiologie, cela pourrait être des motifs sur le mur ou sur le détecteur, des accessoires de positionnement accrochés au mur dans le champ de vision tel que les mousses ou coussins de positionnement.

Les variations radiométriques : qui correspondent aux changements de texture tels que la couleur de peau, des cheveux, de la présence de barbe ou non, etc.

Les déformations géométriques : liées à la morphologie (membre amputé, arthrose sévère, etc.)

L'espace dimensionnel : le corps humain peut prendre une grande variété de poses corporelles. Chaque articulation a plusieurs degrés de liberté, l'espace peut être extrêmement complexe lors des mouvements ou de certaines incidences radiologiques.

Néanmoins certaines faiblesses citées dans la littérature ne sont pas à prendre en compte dans le cadre de la radiologie comme le flou de mouvement qui est minime ou inexistant lors de la réalisation d'incidence car le TRM est formé pour éviter ce phénomène.

3. Problématique

3.1. Question de recherche et objectifs de l'étude

Notre question de recherche est la suivante : « est-ce qu'un système d'intelligence artificielle peut être utilisé comme aide au positionnement pour les étudiants TRM encore en formation et pour les professionnels de terrain ? ».

Pour répondre à cette question, nous avons défini les objectifs suivants :

- Évaluer de façon qualitative différents systèmes d'IA de détection disponibles dans le domaine public sous un angle TRM dans un contexte de positionnement radiographique, notamment en évaluant les systèmes d'IA à :
 - Indiquer la latéralité
 - Détecter les articulations sous une blouse patient
 - Détecter toutes les parties du corps visibles sur une image avec une vue partielle ou avec une vue globale de la personne
- 2. Comparer de manière quantitative la précision de détectabilité et localisation des points de repères anatomiques (joints articulaires et points surfaciques) de l'IA versus celle de l'annotation du TRM, à l'aide de métriques d'évaluation 2D et 3D. Pour ce faire nous allons évaluer la performance de ces systèmes :
 - En comparant la précision de détection des points surfaciques (Densepose) et les points liés aux articulations (OpenPose)
 - A détecter les repères anatomiques avec ou sans la présence de la lumière du collimateur
 - A détecter les repères anatomiques 2D et 3D en fonction du type d'incidences radiologiques
 - En comparant la détection des sujets féminins et des sujets masculins
 - En comparant la détection des images avec un bon et un mauvais positionnement radiographique

3.2. Intérêt de la recherche en pédagogie et en clinique

L'intérêt de ce travail peut être divisé en trois points :

- Nous avons vu que les systèmes d'IA existants n'ont pas été développés initialement pour une utilisation en radiologie, ce travail permettrait donc de mieux analyser ces méthodes, faire le tri des différentes approches d'IA et mettre en lumière les critères de ces systèmes qui sont nécessaires pour que ces approches soient viables dans le domaine de la radiologie et du positionnement radiographique.
- Pour la formation de base, cela apporterait aux TRM étudiants un outil d'apprentissage du positionnement radiographique, qui leur permettrait d'évaluer la qualité de leur pratique en toute sécurité (sans générer des rayons X) tout en étant au plus proche de la réalité du terrain en produisant des radiographies virtuelles.
- Et pour finir en pratique bien que la fluoroscopie soit fortement déconseillée, nous avons observé sur le terrain son utilisation abusive. En effet, les TRMs disposant d'un appareil de radiographie couplé à la fluoroscopie ont tendance à l'utiliser même lorsqu'il n'est pas nécessaire et ceci de façon automatique. Cet outil de diagnostic est petit à petit devenu un outil d'aide au positionnement. Si l'étude aboutit à un résultat concluant, l'intérêt de cette recherche sera d'implémenter ces systèmes d'IA dans la pratique clinique en vue de limiter l'utilisation de la fluoroscopie tout en gardant un outil d'aide au positionnement. Cela permettrait en finalité de diminuer les doses reçues par le patient lors de son examen.

4. Méthode et matériel

4.1. Préambule

Les travaux présentés dans ce mémoire ont contribué à la réalisation du projet AIRx³ mené par le Prof. Jérôme Schmid à la HEdS-Genève dans la filière TRM et soutenu par la fondation Hasler. Le but principal de ce projet est de fournir un simulateur de radiographie utilisant l'IA pour détecter le positionnement et générer des radiographies virtuelles mais réalistes.

4.2. Méthodes d'IA

Toutes les méthodes d'IA étudiées dans ce travail sont accessibles gratuitement au grand publique. Comme présenté dans la section 2.2.4, il existe une multitude d'approches et nous en avons sélectionné une série, présentant des caractéristiques différentes.

DeepCut [51]

Cette méthode utilise plusieurs types de CNN pour détecter la position 2D des joints du squelette d'une ou plusieurs personnes (Figure 10). Contrairement à plusieurs méthodes, elle combine la détection des personnes (Top-Down) présentent sur l'image et l'estimation de pose (Bottom-Up) en une seule étape (voir section 2.2.3). Sur la base de données MPII Multi-Person [75] DeepCut a un mAP de 54.7%.



Figure 10 : quelques exemples d'estimation de pose humaine par le système DeepCut [51]

³ <u>https://www.hesge.ch/heds/recherche-developpement/projets-recherche/en-cours/airx</u>

VNect [40]

Méthode 3D calculant directement la position 3D des joints d'un squelette (Figure 11) et qui est optimisée pour le suivi dynamique et temps réel de la pose (environ 33 millisecondes/image pour la détection et l'estimation de pose en temps réel). Bien que la méthode soit 3D, nous pouvons obtenir un squelette 2D dans chaque image. Avec un score MPJPE de 8.5 cm, cette méthode permet la détection avec des images de faible qualité (prises avec des téléphones portables) mais n'estime la pose que d'une seule personne.



Figure 11 : Estimation de la posture humaine avec la méthode Vnect [40] et modélisation de ce squelette

HMR [31]

Méthode visant à estimer la forme et la position d'un sujet présent dans l'image en utilisant le modèle 3D SMPL [36] (Figure 12). Un squelette 2D est calculé. Cette méthode permet la modélisation d'un squelette 3D sans passer par l'emplacement des joints en 2D. Avec un MPJPE à 8.7cm, cette méthode permet l'estimation en temps réel avec environ 40ms/image.



Figure 12 : modélisation du sujet avec la méthode HMR [31] en utilisant un modèle 3D SMPL [36]

OpenPose [7]

Méthode 2D capable de détecter plusieurs sujets (Figure 14) et pouvant aussi calculer des positions 3D avec des caméras calibrées. Le squelette est assez riche avec la possibilité de détecter aussi plusieurs points du visage et les doigts des mains. Avec un temps de calcul de 5 ms/image cette méthode a un score mAP de 75.6. Elle montre néanmoins une faiblesse de détection quand les personnes sur les images ont la tête en bas (Figure 13), rencontrées dans des positions de yoga ou de gymnastique par exemple.



Figure 13 : détection d'une personne la tête en bas avec la méthode OpenPose [7]



Figure 14 : détection de plusieurs personnes sur diverses images avec la méthode OpenPose [7]

Densepose [65]

Supportant la détection de plusieurs personnes, cette méthode calcule une correspondance entre la surface d'une personne détectée et celle du modèle générique SMPL (Figure 15). Il est ainsi possible de définir des marqueurs *surfaciques*, contrairement à la plupart des méthodes qui détectent des articulations situées à l'intérieur du corps. En réalité, la méthode Densepose détecte en interne un squelette mais celui-ci n'est pas accessible par défaut aux utilisateurs. Selon la documentation, l'accès à cette information nécessiterait un réentraînement complet du réseau pour être disponible ce qui s'avèrerait trop prohibitif en termes de temps de calcul. Cette méthode permet d'estimer en temps réel (25 images par seconde) la pose d'une ou de plusieurs personnes avec des vues complètes et partielles des corps.


Figure 15 : Exemples de détection de plusieurs personnes sur une même image par la méthode Densepose [65]

4.3. Étude préliminaire

L'objectif de cette première étude préliminaire était de tester les différentes méthodes IA à disposition du grand public, pour mieux se rendre compte des limites et des critères nécessaires pour établir un protocole plus rigoureux pour la suite ce de travail de recherche.

4.3.1. Participants

Deux volontaires de 21 et 22 ans ont participé à cette première partie de l'étude. Les participantes ont été informées du déroulement de l'étude et ont donné leur consentement éclairé par écrit.

4.3.2. Protocole

Les acquisitions se sont déroulées sur une après-midi mi-juillet 2019, dans la salle de radiologie de l'HEdS à Genève, à l'aide d'un appareil photo de la marque Canon. Des vues partielles et des vues complètes de plusieurs incidences radiologiques ont été prises, certaines avec blouse, d'autres sous des angles différents par rapport au détecteur.

4.3.3. Résultats et conclusion

193 images ont été prises et fournies en entrée aux différents systèmes d'IA. Suite à l'analyse des détections des différentes méthodes, nous avons pu identifier de façon qualitative différentes faiblesses (voir le premier objectif de l'étude, section 3.1) :

 VNect se trompe parfois sur la latéralité (Figure 16 A) et B)), point très important en radiologie.

- HMR ne détecte pas certains points comme les trois points de repères du bras gauche sur la Figure 16C), alors qu'il n'y a aucune occlusion sur l'image.
- DeepCut renseigne tous les points de repère même si la partie du corps ou l'articulation correspondante n'est pas sur l'image (Figure 16 D) et E))



Figure 16 : résultats des acquisitions de l'étude préliminaire avec les méthodes Vnect (A et B), HMR (C), DeepCut (D et E)

Cette analyse qualitative (latéralité, détection avec blouse et vue partielle) nous montre que dans l'ensemble ces trois méthodes ne sont pas précises dans l'estimation de la pose humaine, il nous semble donc judicieux de les retirer de l'étude. En revanche, la méthode OpenPose est encourageante (Figure 17) pour la suite, elle fournit les informations sur la latéralité, semble être plus précise pour la détection des articulations par rapport aux autres méthodes même sous une blouse, et détecte plusieurs personnes sur la même image ce qui peut être utile dans les cas de radiographie pédiatrique où l'aide du parent est sollicité. Densepose est une méthode qui ne détecte pas de squelette mais dessine une surface, elle nous semble utile pour appréhender la détection de marqueurs surfaciques.



Figure 17 : résultats de la méthode OpenPose lors de l'étude préliminaire

Au vu de ces premiers résultats, il nous a semblé pertinent d'approfondir notre recherche avec les méthodes OpenPose et Densepose.

4.4. Étude principale

Grâce à notre étude préliminaire nous avons pu effectuer un tri de façon qualitative dans les systèmes d'IA à garder, et élaborer un protocole plus rigoureux que nous allons détailler par la suite.

4.4.1. Participants

De nouveaux participants ont été recrutés pour ce deuxième protocole expérimental, trois femmes âgées de 24 à 26 ans et deux hommes âgés de 23 ans et 31 ans. Comme pour l'étude préliminaire, il leur a été demandé de se présenter le jour de l'acquisition en tenu près du corps (type cycliste, legging) pour que le cadre soit conforme aux travaux pratiques réalisés à l'école de radiologie, ainsi que sans bijoux et sans objets métalliques (jeans, ceinture) afin de reproduire les exigences du terrain.

Les participants à l'étude avaient les capacités physiques et mentales à suivre les instructions et rester debout le temps des acquisitions. Ils ont tous signé un formulaire de consentement donnant l'accord pour l'enregistrement photo et/ou vidéo (Annexe 9.3).

4.4.2. Caméras et calibration

La calibration de la caméra est le processus par lequel les valeurs des paramètres (intrinsèques et extrinsèques) d'une caméra sont estimées. C'est un processus essentiel pour corriger tout artefact de distorsion optique, estimer la distance d'un objet par rapport à une caméra, mesurer la taille des objets dans une image, et construire des vues 3D pour les systèmes d'estimation de pose humaine par exemple. Les paramètres intrinsèques comprennent la distance focale de la caméra, le centre optique et les coefficients de distorsion. Les paramètres extrinsèques décrivent la position et l'orientation de la caméra dans des coordonnées du monde en 3D. Dans notre projet, nous utilisons deux caméras afin de pouvoir reconstruire en 3D la position de marqueurs 2D détectés dans chaque caméra, par le principe de triangulation.

Les paramètres intrinsèques sont invariants au déplacement des caméras, tant que les paramètres des lentilles restent inchangés (mise au point, zoom, etc.). Ces paramètres sont estimés par une procédure de calibration [73] durant laquelle une mire (ou cible de calibrage) est photographiée sous différent angles et positions (Figure 18). La mire contient des points 3D qui sont connus et définis dans un repère attaché à la mire. A travers les multiples images

acquises, une détection automatique estime de façon très précise la position de ces points 3D dans les images. On obtient ainsi les correspondances des points de l'image avec les points du monde réel, permettant d'estimer les paramètres intrinsèques.



Figure 18 : principe général de la calibration d'une caméra, tiré de <u>http://wcours.gel.ulaval.ca/2017/a/GIF4100/default/5notes/A2016etalonnageCameraPageWeb.pdf</u>

Les paramètres extrinsèques doivent être recalculés à chaque fois que les caméras sont déplacées mais contrairement à l'estimation des paramètres intrinsèques nécessitant au moins une centaine d'images pour être précise, une seule image par caméra est suffisante pour la calibration des paramètres extrinsèques car les paramètres intrinsèques sont connus.

Dans notre projet la mire présente en outre la caractéristique de pouvoir être imagée de façon partielle grâce à une mire de type « ChArUco » (Figure 19) disponible dans la librairie de vision par ordinateur *OpenCV* [6] et combinant la précision des mires de type damier avec la robustesse des marqueurs de type « QR code ». Ceci facilite donc la calibration des paramètres extrinsèques par des utilisateurs non expérimentés. La mire doit être juste placée sur le détecteur chaque fois que celui-ci est déplacé. Les caméras utilisées sont des caméras Ethernet du constructeur FLIR⁴.

⁴ Modèle RoHS 1.3MP Color Blackfly S GigE PoE, lentille RoHS Fujinon avec distance focale fixe à 6mm.



Figure 19 : Mire de calibration de type « ChArUco » acquise par les deux caméras (ici pour la calibration des paramètres extrinsèques).

4.4.3. Acquisitions

Pour notre étude principale nous avons collecté 763 images qui incluent les images de mire pour la calibration et les positionnements radiologiques, lesquels ont été effectués avec les incidences suivantes (voir le détail des incidences en Annexe 9.4) :

- Thorax de face (postéro-antérieur)
- Thorax de profil
- Cervicale de face
- Cervicale de profil
- Épaule neutre
- Épaule interne
- Profil de Neer
- Lombaire de face
- Lombaire de profil
- Bassin de face
- Faux profil de Lequesne
- Genou de profil

Ces incidences ont été sélectionnées avec des critères précis. Tout d'abord, toutes ces incidences se feront au potter mural pour simplifier la mise en place du matériel (caméra sur trépied). Ensuite en s'appuyant sur le travail de Bachelor de Hirschi Laura [22] nous avons sélectionné ces incidences d'après leurs difficultés à les réaliser sans l'aide de la fluoroscopie. En effet en Suisse romande, 75% des TRM utilisent la fluoroscopie pour positionner le genou de profil, elle est également utilisée pour le profil de Neer avec 65% des TRM de l'enquête. En raison du point de palpation (symphyse pubienne) embarrassant de l'incidence du bassin de face, cette radiographie a été incluse dans la liste. Comme le montre la Figure 20, les incidences du rachis qui sont les plus assistées par la fluoroscopie sont les cervicales de face (45%) et de profil (40%) et les lombaires de face et de profil (40%).



Figure 20 : Graphique issu du travail de Bachelor de Laura Hirschi qui montre l'utilisation de la fluoroscopie pour les incidences du rachis

Pour terminer et dans le but de pouvoir évaluer ces systèmes sur tout le corps humain nous voulions au moins une incidence par région anatomique. Ne pouvant pas être réalisées sur le potter mural les extrémités ont été exclues, tout comme le crâne car la plupart des systèmes ne distinguent que très peu les points de repères de la face.

Résumé dans le Tableau 2, pour chacune de ces douze incidences et pour les cinq participants de l'étude nous avons réalisé deux photos (une par caméra) avec un positionnement du patient correct (avec et sans la lumière de la collimation) et deux photos avec un mauvais positionnement (Tableau 3) (avec et sans lumière). Chaque incidence a été photographiée avec les deux caméras sous des angles différents, le positionnement des caméras étant resté fixe tout au long de l'acquisition. L'acquisition de notre jeu de données s'est déroulée sur une journée complète fin janvier 2020 à l'HEdS de Genève.

	Sujet 1	Sujet 2	Sujet 3	Sujet 4	Sujet 5	Total
	féminin	féminin	féminin	masculin	masculin	
Bon positionnement	48	46	43	45	46	228
Mauvais positionnement	51	51	52	56	50	260
Total	99	97	95	101	96	488

Tableau 2 : résumé des éléments de notre base de données initiale

Tableau 3 : Détails de la réalisation des mauvais positionnements

Incidences	Détails du mauvais positionnement
Thorax de face (postéro-antérieur)	Mauvais centrage en hauteur
Thorax de profil	Rotation du buste pour ne pas être de profil strict
Cervicale de face	Menton du sujet baissé
Cervicale de profil	Sujet de 3/4, tête posée sur le potter
Épaule neutre	Mauvais centrage en latéral
Épaule interne	Sujet dos contre la plaque, pas de rotation à 45°
Profil de Neer	Sujet ventre contre la plaque
Lombaire de face	Centrage trop bas, pas de collimation
Lombaire de profil	Sujet en extension
Bassin de face	Pied en rétroversion, rotation du bassin de 3/4
Faux profil de Lequesne	Bassin décollé du potter
Genou de profil	Genou décollé du potter, genou de ¾

4.5. Traitements des données

Pour pouvoir analyser nos données, toutes les images de mires furent utilisées pour calibrer les paramètres extrinsèques (section 4.4.2). Les images mal appariées (suite à une incorrecte synchronisation des images entre les caméras ou à une mauvaise manipulation de notre part) ou redondantes furent retirées car la 3D ne peut se calculer sur la base d'une seule image. Au final, notre base de données comprend un total de 488 images à annoter.

4.5.1. Annotation des points de repères

Nous allons maintenant décrire comment nous avons annoté notre collection d'images. Afin d'adapter le système aux incidences radiologiques nous avons ajouté aux points de repères du squelette d'OpenPose des points de repères qui nous semblait pertinent pour le positionnement.

Plusieurs points de repères ont été supprimés du squelette d'OpenPose et huit nouveaux points de repères ont été ajoutés (voir Figure 21) : des points au niveau des aisselles pour estimer si le patient n'est pas penché en avant ou en arrière lors de la radiographie du thorax de profil, deux points sur la partie interne du coude pour apprécier la rotation du bras lors des radiographies des épaules, deux points à la hauteur des crêtes iliaques pour mesurer l'équidistance par rapport au détecteur lors de la radiographie du bassin de face et les deux derniers points au niveau du condyle médial des genoux pour déterminer si l'on est bien de profil lors du positionnement de l'incidence du genou de profil. Tous ces points sont surfaciques et sont détectés par la méthode Densepose. Pour cela, ces points sont définis une fois pour toute sur le modèle générique SMPL afin de faire la correspondance avec la surface générée par Densepose.



Figure 21 : placement des points de repères de notre modèle et tableau comparatif des points d'OpenPose versus nos points de repères. Les points de repères ajoutés pour notre étude sont en rouge et les points de repères barrés sont ceux supprimés de l'annotation car ils ne représentaient pas d'utilité.

Image source : https://fr.wikipedia.org/wiki/Fichier:Silhouette_humain_asexue_anterieur_posterieur.

Vu le nombre conséquent d'images de notre jeu de données, il était essentiel d'utiliser une interface simple mais efficace. La totalité de la base de données produite lors de cette étude a été annotées sur le logiciel VGG Image Annotator (VIA)⁵. Ce logiciel d'annotation simple peut servir à annoter des images (Figure 22), des vidéos et de l'audio. VIA fonctionne dans un

⁵ <u>http://www.robots.ox.ac.uk/~vgg/software/via/</u>

navigateur web et ne nécessite aucune installation ou configuration. Utile à la fois pour des projets universitaires et des applications commerciales, VIA est une interface open source développée au sein du Visual Geometry Group (VGG).



Figure 22 : Capture d'écran de l'interface VIA lors de l'annotation manuelle d'une radiographie de thorax de face. Les points localisés en haut à gauche sont des points ne pouvant pas être localisés de façon fiable sur la personne par l'opérateur.

Pour ce travail, l'annotation a été réalisé par un unique opérateur de formation TRM (M. Butt), et chaque image a été annoté une seule fois. Par manque de temps nous n'avons pas fait annoter nos images par plusieurs TRMs et nous n'avons pas pu ainsi estimer la reproductibilité des annotations par le même opérateur.

4.5.2. Mesures statistiques

A l'aide des métriques 2D et 3D détaillées dans la section 2.2.3, nous avons effectué différents tests statistiques pour déterminer des différences significatives entre :

- Les prédictions des systèmes d'IA et l'annotation d'une TRM
- L'estimation des points de repères par OpenPose versus Densepose
- La détection avec lumière du collimateur versus sans lumière
- L'impact du genre du sujet pour une même incidence
- L'impact de la détection avec des bons et des mauvais positionnements radiographiques

Les tests statistiques ont été réalisés à l'aide du logiciel Stata⁶ (version 15.1).

⁶ https://www.stata.com

4.6. Génération des radiographies virtuelles

Nous n'avons pas participé au développement du système de génération de radiographies virtuelles du projet AIRx, nous présentons néanmoins ici le fonctionnement de la génération de radiographies virtuelles car celui va de pair avec notre travail. En effet, notre analyse des systèmes d'IA notamment sous l'angle TRM est essentielle pour cette génération car la détection du positionnement du patient influence la création des radiographies synthétiques. Dans notre chapitre Résultats, nous présenterons des exemples de radiographies virtuelles tout en discutant des forces et des limites du système actuel.

La génération de radiographies virtuelles suit les étapes suivantes :

- Détection de la position du sujet par la reconstruction 3D de son squelette. Suivant la même méthodologie que pour notre travail, chaque acquisition est réalisée sous l'angle de vue des deux caméras préalablement calibrées comme le décrit la section 4.4.2. La reconstruction 3D du squelette se fait grâce au principe de triangulation sur les deux images 2D.
- 2. Sélection, adaptation et positionnement d'un volume CT. La génération de radiographies virtuelles suit le principe des « digitally reconstructed radiographs » (DRR) [67] qui utilise une image tomodensitométrique. Dans l'idéal, ce volume CT devrait être une acquisition du sujet dans la même position afin de représenter sa morphologie, et pour qu'il soit correctement localisé par rapport à un tube et un détecteur virtuels.
- 3. Création de la radiographie virtuelle à partir du CT et de paramètres techniques fournis par l'utilisateur. Une DRR s'obtient en simulant la physique de rayons X traversant le corps. Le CT permet de déduire les coefficients linéaires d'atténuation, et il est donc possible de générer une image ressemblant à une radiographie grâce à la connaissance de la géométrie du système radiographique tube-détecteur.

Dans le projet AIRx, une radiographie plus réaliste est obtenue en modifiant l'approche de Unberath M. et al. [67], basée sur des méthodes de Deep Learning afin de :

- Tenir compte de la nature polychromatique du spectre du rayonnement X en utilisant différentes segmentations de l'image CT (os, tissu mou, air, etc.),
- Générer du bruit en le couplant notamment à la charge du tube,
- Générer du rayonnement diffusé en présence ou non de grilles anti-diffusantes.

5. Résultats

5.1. Détection

5.1.1. OpenPose versus Densepose

Dans cette section nous mesurons la différence entre les métriques des points de type « OpenPose » (i.e. détectés par la méthode OpenPose) et les points « Densepose » (section 4.5.1).

A noter que ces deux méthodes ne soient pas strictement comparables puisqu'elles ne détectent pas la même chose (squelette vs. surface) et que les points détectés diffèrent (points articulaires pour OpenPose, points surfaciques pour Densepose). Cependant, nous voulons déterminer laquelle de ces deux méthodes semblent être plus sensible à la détection.

D'après les écrits [7][18] et de manière subjective, la détection Densepose semble moins précise que la méthode Openpose, lié au fait que Densepose crée un maillage sur le corps qui prend en compte les vêtements et qui n'épouse pas toujours parfaitement le contour (Figure 23).



Figure 23 : exemple du maillage de Densepose pour les incidences du bassin de face (à gauche) et épaule interne (à droite). On peut observer des erreurs, typiquement au niveau du pied droit de la volontaire, ou au niveau du cou du volontaire.

Le Tableau 4 et le Tableau 5 résument respectivement les moyennes des métriques OKS et PED pour les points de repères relatifs à la méthode OpenPose et à la méthode Densepose. Nous pouvons observer que la moyenne des OKS des points de repère OpenPose est supérieure à celle de Densepose (0.927 vs. 0.9), soulignant les meilleures performances de OpenPose. Cette observation est confirmée avec la distance Euclidienne en 3D, qui est inférieure avec OpenPose (2.6 cm vs. 4.5 cm). A noter que le point de repère « RArmPit » n'a pas de métrique 3D (Tableau 5) car il n'est jamais simultanément détecté dans deux vues en correspondance. En effet, Densepose n'extrapole pas la surface détectée dans les parties non visibles.

Points de repère OpenPose	Moyenne OKS	Moyenne PED [m]
1 : Nose	0.880	0.012
2 : REye	0.889	0.011
3: LEye	0.889	0.011
4: Neck	0.976	0.064
5: RShoulder	0.856	0.018
7: RElbow	0.928	0.026
9 : RWrist	0.929	0.020
10 : LShoulder	0.986	0.019
12 : LElbow	0.915	0.027
14 : LWrist	0.927	0.020
16 : RHip	0.945	0.036
17 : RKnee	0.944	0.030
19 : RAnkle	0.982	0.017
20 : RBigToe	0.946	0.031
21 : RHeel	0.942	0.039
23 : LHip	0.948	0.035
24 : LKnee	0.927	0.032
26 : LAnkle	0.976	0.017
27 : LBigToe	0.942	0.035
28 : LHeel	0.948	0.036
Moyenne des points OpenPose	0.927	0.026

Tableau 4 : moyenne des métriques OKS et PED [m] pour les points d'OpenPose

Points de repère Densepose	Moyenne OKS	Moyenne PED [m]
6: RArmPit	0.856	-
8: RIntELbow	0.898	0.046
11 : LArmPit	0.820	0.057
13 : LIntELbow	0.886	0.045
15 : RMidWaisT	0.899	0.045
18 : RMedialCondyle	0.923	0.043
22 : LMidWaisT	0.913	0.035
25 : LMedialCondyle	0.967	0.045
Moyenne des points Densepose	0.900	0.045

 Tableau 5 : moyenne d'OKS et de PED [m] pour les points de Densepose. Le point de repère « RArmPit » n'as pas de valeur

 PED car il n'est jamais simultanément détecté dans deux vues en correspondance.

5.1.2. Impact de la lumière

D'après les écrits de Zhu A. [74], la lumière et les jeux d'ombres peuvent avoir une influence sur la détection par les systèmes IA, un test de Wilcoxon [25] appliqué sur les mesures d'OKS a été réalisé pour déterminer s'il y a une différence significative entre les images prises avec la lumière du collimateur versus les images prises sans cette lumière, avec une p-valeur > 0.05 (0.575) nous pouvons dire qu'il n'y a probablement pas de différence significative entre la présence de lumière du collimateur sur les images et l'absence de cette lumière (Figure 24).



Figure 24 : Exemple de notre base de données montrant la détection OpenPose avec lumière (A) et sans lumière (B) pour la même position radiologique. Pas de différences significatives de détection sont observées.

5.1.3. Impact du type d'incidence

Pour évaluer la performance d'une façon globale de notre système hybride combinant les deux systèmes par rapport au type d'incidence, nous avons observé la différence 3D d'estimation entre tous les points annotés par une TRM et ces mêmes points de repères détectés par l'IA grâce à la métrique PED [m]. En moyenne, on mesure une différence de 2.9 cm pour toutes les incidences (Tableau 6). Nous pouvons observer que l'incidence la mieux détectée est l'épaule neutre (2.3 cm), alors que celle qui présente l'erreur la plus élevée est l'incidence des lombaires de profil (3.8 cm).

INCIDENCES	PED [m]
thorax_face_PA	0.030
thorax_profil	0.031
epaule_neutre	0.023
epaule_interne	0.025
profil_neer	0.035
genou_profil	0.035
cervicales_profil	0.028
cervicales_face	0.026
bassin_face	0.025
lombaires_face	0.024
lombaires_profil	0.038
hanche_profil	0.032
Moyenne	0.029

Tableau 6 : moyennes des PED [m] en fonctions des incidences

5.1.4. Impact de la détection des incidences en fonction du genre du sujet

Nous avons effectué un t-test sur les mesures de la métrique OKS pour rechercher une différence significative entre la détection des hommes et des femmes par les systèmes d'IA. Avec une p-valeur > 0.05 (p-valeur=0.31) nous pouvons dire qu'il n'y a probablement pas de différence entre la détection des hommes versus celle des femmes par les systèmes d'IA.

5.1.5. Impact de la détection sur des bons et des mauvais positionnements radiographiques

Comme nous le montre les résultats (Tableau 7), il n'y a pas en apparence d'écart important entre l'estimation de pose d'un bon positionnement versus d'un mauvais positionnement. Un t-test appliqué aux mesures PED et OKS des deux méthodes d'IA (respectivement p-valeur=0.37 et 0.68) confirme qu'il n'y a pas de différence significative entre la détection de bonnes et de mauvaises radiographies.

Incidences	PED [m]	OKS
Bon positionnement	0.029	0.93
Mauvais positionnement	0.030	0.92

Tableau 7 : tableau des moyennes d'OKS et PED [m] pour les bons et mauvais positionnements

5.2. Radiographies virtuelles

Nous avons testé la génération de radiographies virtuelles sur certaines de nos acquisitions. Les incidences actuellement supportées sont celles du bassin, des hanches et du thorax. Actuellement il n'est pas possible de déformer le CT afin de reproduire la morphologie ou la position des membres articulés du sujet présent dans les images. Ceci empêche donc de réaliser des radiographies virtuelles d'incidences où la position du patient dans le CT diffère trop de celle du sujet détecté.

La Figure 25 illustre un exemple de pose de bassin de face AP reconstruite sur la base des images acquises d'un sujet.



Figure 25 : Exemple de squelette reconstruit pour une radiographie du bassin de face. Rangée supérieure : images du sujet acquises par les caméras. Rangée inférieure : squelette reconstruit (rouge) avec le cône de projection (jaune) et le détecteur (vert). Un modèle 3D estimé sur la base du squelette est aussi représenté.

La Figure 26 illustre les radiographies virtuelles résultantes avec les constantes 80 kV et 50 mAs, une distance foyer-détecteur à 120 cm, en absence et en présence de rayonnement diffusé. Lorsque le diffusé est présent une grille anti-diffusante est simulée. Un détecteur de taille 43 x 43 cm est utilisé.



Figure 26 : Exemple de radiographies virtuelles du bassin de face (80 kV, 50 mAs, 120cm, détecteur 43 x 43 cm). A gauche : pas de rayonnement diffusé simulé ; à droite : rayonnement diffusé simulé en présence d'une grille anti-diffusante ratio 12:1,focalisée à 120cm.

D'autres exemples de génération de radiographies virtuelles sont présentés en Annexe 9.5.

6. Discussion

Dans cette partie nous allons discuter des résultats 2D et 3D calculés sur notre base de données. Bien qu'ils puissent être améliorés, les résultats sont globalement positifs.

6.1. Performances et limites des méthodes de détection IA

En synthétisant les résultats des métriques (Tableau 8), nous pouvons voir que les points d'OpenPose sont mieux détectés que les points Densepose avec une moyenne d'erreur de 2.6 cm pour les points d'OpenPose et 4.5 cm pour les points Densepose, ainsi qu'une moyenne d'OKS supérieure pour OpenPose. Nous rappelons que ce sont deux méthodes différentes avec des points surfaciques pour Densepose et des joints articulaires pour OpenPose, la comparaison de ces deux méthodes permet seulement d'apprécier la précision des méthodes pour un futur travail qui utiliserait une seule méthode par exemple.

Ces résultats peuvent s'expliquer d'une part par le petit échantillon de données testé pour cette étude : 488 images pour notre étude versus 105698 images pour Densepose, et d'autre part par le fait que nos images ont été capturées toutes dans un même contexte avec des conditions similaires, rappelons que les images issues de la base de données COCO [81] sont très diverses, avec des postures et des descriptions très différentes (activité sportives, rassemblement, etc.)

	Moyenne OKS	Moyenne PED [m]
Points OpenPose	0.927	0.026
Points Densepose	0.900	0.045

Néanmoins les performances de ces systèmes d'IA dans un contexte d'imagerie semblent prometteuses. D'un point de vue TRM, il est tout à fait envisageable de positionner certaines incidences comme le bassin de face ou le thorax de face grâce aux points d'OpenPose et une erreur de 2.6 cm, en effet les points de centrage de ces radiographies (voir Annexe 9.4) sont respectivement 3 à 5 cm au-dessus de la symphyse pubienne et à hauteur de la pointe des scapulas. Comme expliqué au début de ce travail, le point de repère de la symphyse pubienne est délicat à palper, ce n'est plus une pratique courante en clinique, de plus le centrage de cette incidence varie entre 3 à 5 cm ce qui me permet de dire qu'une erreur de 2.6 cm est acceptable pour la réalisation de cette incidence. Si nous suivions le même raisonnement pour l'incidence du thorax de face, en fonction de comment le patient a rentré les coudes en avant

au moment de son positionnement, les scapulas sont plus ou moins bien dégagées des poumons et donc la hauteur des scapulas varient de quelques centimètres.

En revanche, pour une incidence d'épaule de face par exemple, une erreur de 2.6 cm nous semble pas raisonnable. Centré sur la fossette sous acromiale, la collimation joue un rôle très important dans la qualité de cette radiographie, cette dernière doit être aux limites de la peau, il est même conseillé d'utiliser un cône localisateur⁷ et pour cela un décalage de plus de 2 cm risque d'augmenter la dose au patient et d'entraver la qualité de la radiographie réalisée.

Le Tableau 9 présente les résultats en centimètres de quelques méthodes issues de la littérature. Comme expliqué à la section 2.2.3, ses résultats sont en MPJPE, une métrique équivalente au PED de notre travail. Nous pouvons voir que notre erreur est inférieure à celles de toutes les méthodes citées dans ce tableau, sauf [28] qui a une erreur de 2.08 cm.

Tableau 9 : Le tableau présente les résultats en distance euclidienne (MPJPE) disponibles dans l'état de l'art.

Méthode	Moyenne distance euclidienne [cm]
Pavlakos G. et al. [49]	5.69
<i>Tome D. et al. [64]</i>	5.28
Iskakov K. el al. [28]	2.08
Kostrikov I. et al. [35]	4.03
Notre modèle [en PED]	2.99

En ce qui concerne la présence de lumière, les résultats prouvent qu'il n'y a pas de différence de détection avec ou sans la présence de la lumière du collimateur. C'est un point positif pour ces systèmes d'IA puisque ceci permet au TRM de ne pas attendre la disparition de la lumière du collimateur qui pour certains appareils de radiologie se fait automatiquement après une vingtaine de secondes.

Puisque le TRM est la personne décidant de tamiser sa salle d'examen, il serait tout de même intéressant de tester ces systèmes dans des salles de radiologies avec des éclairages différents, pour voir s'ils sont sensibles aux changements de luminosité.

Un autre point apprécié de ces systèmes est qu'ils ne font pas de différence entre les hommes et les femmes, le contraire aurait été problématique puisqu'il aurait fallu limiter son utilisation a un seul genre. Ceci n'est concevable ni en pratique, ni dans l'enseignement.

⁷ Le cône localisateur permet une réduction de la dose patient grâce à la collimation forcée du faisceau et l'élimination du rayonnement secondaire en sortie de tube.

Le Tableau 10 présente les résultats de la métrique mAP (section 4.5.2) pour plusieurs systèmes d'IA. Nous pouvons voir que notre méthode a un mAP bien supérieur aux autres références citées. Comme évoqué précédemment, cela s'explique notamment par le fait que les bases de données testées ne sont pas identiques. Néanmoins, étant donné le caractère exploratoire de cette étude, nous sommes très satisfaits de nos résultats qui sont prometteurs et encourageants pour continuer la recherche sur ce même sujet.

Méthode	mAP [%]
Regional Multi-person Pose Estimation [13]	76.7
OpenPose [7]	75.6
Generative Partition Networks [43]	84.8
DeeperCut [27]	71.2
Deep High [61]	75.3
Single-Stage Multi-Person Pose Machines [44]	78.5
DensePose [18]	55.8
Notre approche	91.6

Tableau 10 : Comparaison de notre mAP exprimé en % par rapport à quelques méthodes d'IA publiées

6.2. Utilisation de ces méthodes pour l'enseignement

Les méthodes d'enseignement des radiographies n'étant pas optimales, un des intérêts de conduire cette recherche était de trouver une solution pour combler les lacunes entre les travaux pratiques effectués à l'école et les stages cliniques sur le terrain pour l'apprentissage des radiographies. Les résultats de notre étude (PED à seulement 3 cm en moyenne) montrent que les systèmes d'IA d'estimation de pose sont suffisamment précis pour pallier aux manques des étudiants TRM. Ils auraient la possibilité d'apprendre et de s'entrainer sans avoir recours aux rayons X avec la possibilité de recréer les conditions du terrain (temps limité pour chaque radiographie, etc.) en fonction de leurs attentes.

La génération des radiographies virtuelles est un outil précieux pour les étudiants lors de l'apprentissage des incidences. En effet cette simulation complète l'apprentissage du positionnement en offrant aux étudiants la possibilité de visualiser de façon concrète leurs pratiques et de l'évaluer. Grâce aux radiographies virtuelles ils peuvent modifier les paramètres d'acquisitions (kV, mAs, présence de diffusé, distance foyer-détecteur) et juger de la qualité de la radiographie produite. Les radiographies virtuelles étant générées par des

examens CT, elles offrent la possibilité d'avoir des morphologies variées ainsi que la présence de certaines pathologies ou de corps étrangers (Figure 28). En revanche il n'est pas certain de pouvoir faire concorder la radiographie virtuelle à la morphologie de chaque étudiant. Pour ce faire il faudrait avoir des examens CT de tous genres, âges, tailles et poids différents ou de mettre en place une technique de déformation de volumes CT qui peut s'avérer complexe.

Ces systèmes sous la forme d'aide aux positionnements pourraient aussi être utilisés comme une démonstration lors des portes ouvertes pour la présentation du métier de TRM, une activité ludique que tous futurs étudiants ou personnes souhaitant s'informer sur le métier pourraient essayer.

Pour rendre ce travail plus riche, nous pourrons y ajouter une partie qualitative, en effet d'après les écrits et notre expérience personnelle, nous sommes partis du principe qu'il y a des lacunes dans l'apprentissage de la radiographie. Il nous semble donc important de questionner les étudiants sur leurs ressentis et si un système d'IA tel que présenté dans ce travail pour l'enseignement leur serait bénéfique.

6.3. Utilisation de ces méthodes pour la pratique clinique

Un des objectifs de ce travail de recherche était d'étudier si l'utilisation de la fluoroscopie en tant qu'aide au positionnement pourrait être diminuée grâce à ces systèmes d'IA d'estimation de pose. Utiliser ces systèmes tels quels ne nous semble pas pertinent, la position des patients en clinique se doit d'être précise pour réaliser un bon diagnostic et pour pouvoir reproduire les radiographies à l'identique en cas de suivi. Une différence de 3 cm peut complétement changer le point de centrage d'une incidence et ainsi on risque d'irradier le patient de façon abusive. En revanche si nous pouvions entrainer le système avec des images de positionnement radiographiques et avec un squelette présentant des points de repères pertinents pour les incidences radiologiques, nous pensons que les résultats pourraient être améliorés. Néanmoins, une des limites de cette approche serait d'annoter avec une très grande précision des points de repères visuels sur les images qui sont à la base des points de palpation lors de positionnement des patients (par ex. crêtes iliaques)

De plus comme décrit dans la section 5.1.5, les systèmes ne font pas la différence entre une bon positionnement et un mauvais positionnement du fait qu'ils ne sont pas entrainés pour cette tâche spécifique. C'est un objectif que nous avions souhaité atteindre en début de projet mais par manque de temps et de ressources nous n'avons pas pu aller jusqu'au bout de cet objectif. En effet, nous pouvons imaginer entrainer ces systèmes d'IA en leur fournissant une base de données étiquetées avec des bonnes et des mauvaises radiographies pour qu'ils apprennent à faire la différence. Cela nécessiterait beaucoup de travail et l'acquisition d'une grande variété de données, mais supposons qu'après avoir spécifié à ces systèmes tous les critères de réussites d'une bonne radiographie, la position des points de repères et le degré des angles des parties anatomiques par rapport au détecteur (ce sont là tous les critères d'apprentissage pour le positionnement d'une radiographie), nous pouvons concevoir un système d'IA qui puisse classifier les bonnes radiographies des mauvaises.

6.4. Perspectives

Pour des raisons de temps, de ressources et limitations des systèmes d'IA utilisés, nous avons fait le choix de cibler certaines incidences pertinentes au vu de nos objectifs. Mais pour une étude a plus large échelle, il serait important de tester ces systèmes d'IA sur toutes les incidences radiologiques (avec un plus grand nombre de données), réalisables au potter mural et à la table d'examen. Par exemple, les nouveaux systèmes d'IA sont désormais plus sensibles dans la détection des mains [41], il serait donc intéressant de tester ces systèmes sur les incidences des extrémités.

De plus, pour ce travail l'annotation des images a été réalisée par une seule TRM, pour calculer la variabilité intra et inter-opérateur et avoir des résultats plus significatifs il faudrait faire annoter les images de la base de données par plusieurs TRMs, et annoter les mêmes images plusieurs fois.

Pour la suite de ce travail, nous pouvons imaginer créer une nouvelle méthode d'IA spécialement conçue pour une utilisation en radiologie. Grâce à cette étude nous pouvons donc d'ores et déjà cibler certains paramètres de ces systèmes à mettre en avant comme la précision de la détection et d'autres à ne pas privilégier comme par exemple la possibilité de détecter la pose en temps réel ou d'estimer la position de plusieurs personnes sur une même radiographie qui ne serait pas utiles dans l'enseignement comme en pratique clinique.

Bien que la génération de radiographies virtuelles soit satisfaisante pour les incidences du bassin de face et du thorax (section 4.6 et annexe 9.5), nous avons vu que ce procédé rencontrait des limites. En effet en se basant sur les acquisitions du CT, les radiographies résultantes ne peuvent pas répondre aux exigences des positionnements plus complexes notamment pour les incidences avec des ouvertures d'angles (hanche de profil, genou de profil etc.) Pour approfondir le travail sur la génération des radiographies virtuelles, nous pouvons imaginer déformer les images CT en se basant sur le maillage du système Densepose pour se conformer aux exigences radiologiques.

6.5. Aborder l'IA en radiologie

Enfin pour cette dernière partie, je ferais part de mon ressenti quant à ce travail sur l'intelligence artificielle. J'ai rencontré quelques difficultés à appréhender ces systèmes d'IA et d'imaginer la suite de ce travail, d'où l'importance d'avoir réalisé une première étude préliminaire, puis après avoir saisi les concepts clés, j'ai pu mettre en application mes compétences de TRM et grâce au Prof. Jérôme Schmid et son équipe, réaliser ce travail.

En raison de l'importance de l'intelligence artificielle dans le domaine de la radiologie, il serait avantageux pour les TRMs d'avoir des cours d'introduction à l'IA pour comprendre le langage informatique et mathématique parfois délicat. Je suis persuadée que le TRM peut contribuer à l'amélioration et au développement de l'IA en imagerie médicale.

Dans la même optique, la conception d'une commission⁸ portant sur la digitalisation et l'intelligence artificielle a récemment été planifiée par l'ASTRM (Association Suisse des Technicien en Radiologie Médicale), ouverte à tous les TRMs qui souhaitent y collaborer, cette décision traduit bien l'importance et la nécessité d'impliquer les TRMs dans ce domaine d'IA.

⁸ <u>https://www.astrm.ch/verband/news/detail/aufloesung-fachstelle-diagnostik-fachstelle-digitalisierung-kuenstliche-intelligenz-in-planung</u>

7. Conclusion

Dans ce travail nous avons évalué la performance de différents systèmes d'IA d'estimation de pose humaine pour une utilisation lors des positionnements en radiodiagnostic. Les deux méthodes d'IA retenues et analysées plus finement dans cette étude sont OpenPose et Densepose. Nous avons démontré qu'OpenPose était plus précise dans notre contexte que Densepose, cependant comparés à l'état de l'art ils sont tous deux performants.

Ce travail nous a permis de s'ouvrir sur un certain nombre de perspectives, telles que l'utilisation de ces systèmes comme aide au positionnement lors de l'apprentissage des incidences radiographiques à l'école, ou la mise en évidence des critères à privilégier pour un usage de ces systèmes en imagerie médicale, comme la précision de détection et le choix des points de repères. Ce travail nous a également permis de comprendre les besoins de ces systèmes d'IA à savoir l'entrainement sur une large base de données qui inclut notamment des cas issus de la pratique visée.

Bien que cette étude soit exploratoire, nous nous sommes rendus compte que ces systèmes d'IA d'estimation de pose n'étaient aujourd'hui pas encore suffisamment fiables pour une utilisation en clinique mais ils présentent un fort potentiel et au vu de l'avancé rapide et exponentiel de ces méthodes, il ne serait pas étonnant d'en voir en radiologie dans un futur proche.

Il n'y a ce jour pas d'articles combinant ces méthodes d'IA et le positionnement en radiologie mais il semble tout à fait pertinent au vu de leurs compétences que les techniciens en radiologie médicale participent au développement de la recherche. En outre, ce travail prouve que les TRM ont une place dans le monde de l'intelligence artificielle, ils peuvent se positionner en tant qu'experts et avec l'aide des autres professionnels du domaine médical, technique et scientifique mener à bien ces projets.

8. Références

[1] Aït si selmi T. Examens complémentaires devant un genou douloureux de l'adulte. In: 18 Mises au Point en Chirurgie du Genou [Internet]. Elsevier; 2013; p. 3-18. Disponible sur: https://linkinghub.elsevier.com/retrieve/pii/B9782294740534000017

[2] Ayache N. L'imagerie médicale à l'heure de l'intelligence artificielle. In Villani Cédric et Nordlinger Bernard (dir.), Santé et intelligence artificielle. CNRS Editions, Paris ; 2018 ; p. 151-154

[3] Boet S, Granry J-C, Savoldelli G. La simulation en santé: De la théorie à la pratique. Springer Science & Business Media; 2013; p. 439

[4] Bontrager KL, Lampignano JP, Bontrager KL. Bontrager's handbook of radiographic positioning and techniques. 7th ed. St. Louis, Mo: Mosby/Elsevier; 2010; p. 323

[5] Bott OJ, Dresing K, Wagner M, Raab B-W, Teistler M. Informatics in Radiology: Use of a C-Arm Fluoroscopy Simulator to Support Training in Intraoperative Radiography. RadioGraphics. mai 2011;31(3):E65-75.

[6] Bradski, G. The OpenCV Library. Dr. Dobb's Journal of Software Tools. 2000. p. 120; 122-125. Disponible sur: <u>https://www.drdobbs.com/open-source/the-opencv-library/184404319</u>

[7] Cao Z, Hidalgo G, Simon T, Wei S-E, Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. 30 mai 2019; Disponible sur: <u>http://arxiv.org/abs/1812.08008</u>

[8] Carreira J, Agrawal P, Fragkiadaki K, Malik J. Human Pose Estimation with Iterative Error Feedback. 12 juin 2016; Disponible sur: <u>http://arxiv.org/abs/1507.06550</u>

[9] Charbonnier C. Extreme Hip Movements Based on Optical Motion Capture. Université de Genève. 2010 ;197.

[10] Christe A, Leidolt L, Huber A, Steiger P, Szucs-Farkas Z, Roos JE, et al. Lung cancer screening with CT: Evaluation of radiologists and different computer assisted detection software (CAD) as first and second readers for lung nodule detection at different dose levels. European Journal of Radiology. déc 2013;82(12):e873-8.

[11] Deprez T, Schrijvers L, Pauwels H, Bosmans H, Marchal G. Can patient positioning using an ultrashort fluoroscopic pulse be justified. IAEA International Conference on Radiological Protection of patients in diagnostic and interventional radiology, nuclear medicine & radiotherapy. Mars 2001; p. 89-92

[12] Dogan E. Estimation de pose humaine et reconnaissance d'action par un système multi-robots. Modélisation et simulation. Université de Lyon. 2017.

[13] Fang H-S, Xie S, Tai Y-W, Lu C. RMPE: Regional Multi-person Pose Estimation. 3 févr 2018; Disponible sur: <u>http://arxiv.org/abs/1612.00137</u>

- [14] Filliettaz L, Durand I. L'activité de conception des tuteurs comme modalisation de l'expérience du travail : le cas de la formation des techniciens en radiologie médicale. Les Sciences de l'éducation Pour l'Ère nouvelle. 2016; p. 83.
- [15] Germonpré S, Vermaut G, Trybou J. The acceptance of radiographers to use Non-fluoroscopy Guided Positioning and the influence of leadership: a survey research. Radioprotection. oct 2019; p. 295-301.
- [16] Germonpré S, Gemmel P, Beeckman D, Trybou J. Determinants of using (non-)fluoroscopyguided positioning for radiographs: a case study research. Journal of Radiological Protection. sept 2016;36(3):667-79.

- [17] Golodetz S, Cavallari T, Lord NA, Prisacariu VA, Murray DW, Torr PHS. Collaborative Large-Scale Dense 3D Reconstruction with Online Inter-Agent Pose Optimisation. IEEE Transactions on Visualization and Computer Graphics. nov 2018;24(11):2895-905.
- [18] Güler RA, Neverova N, Kokkinos I. DensePose: Dense Human Pose Estimation In The Wild. 1 févr 2018; Disponible sur: <u>http://arxiv.org/abs/1802.00434</u>
- [19] Habba MR. Radiological Technology and Positioning For X ray Technician Part I. Suez Canal University. 2019
- [20] Haynes, K., & Curtis, T. Fluoroscopic vs blind positioning : comparing entrance skin exposure. Radiologic technology. oct 2009 ;81:15- 23
- [21] Hendee WR, Ritenour ER. Medical Imaging Physics [Internet]. John Wiley & Sons; 2003. Disponible sur: <u>https://books.google.fr/books?id=55lh1B82SLsC</u>
- [22] Hirschi L. La fluoroscopie de positionnement : raisons de son utilisation par les Trm. Travail de Bachelor. Haute Ecole de Santé Vaud. 2016; 71 p.
- [23] Holm T, Palmer PES, Lehtinen E. Manual of radiographic technique. World Health Organization. 1986; 254 p.
- [24] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nature Reviews Cancer. août 2018;18(8):500-10.
- [25] Howell, D., Bestgen, Y., Yzerbyt, V., & Rogier, M. Méthodes statistiques en sciences humaines. 2008. Disponible sur : <u>https://dial.uclouvain.be/pr/boreal/object/boreal:76106</u>
- [26] Hricak H, Brenner DJ, Adelstein SJ, Frush DP, Hall EJ, Howell RW, et al. Managing Radiation Use in Medical Imaging: A Multifaceted Challenge. Radiology. mars 2011;258(3):889-905.
- [27] Insafutdinov E, Pishchulin L, Andres B, Andriluka M, Schiele B. DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model. 30 nov 2016; Disponible sur: <u>http://arxiv.org/abs/1605.03170</u>
- [28] Iskakov K, Burkov E, Lempitsky V, Malkov Y. Learnable Triangulation of Human Pose. 14 mai 2019; Disponible sur: <u>http://arxiv.org/abs/1905.05754</u>
- [29] Iqbal U, Gall J. Multi-Person Pose Estimation with Local Joint-to-Person Associations. 31 août 2016; Disponible sur: <u>http://arxiv.org/abs/1608.08526</u>
- [30] Jain A, Tompson J, LeCun Y, Bregler C. MoDeep: A Deep Learning Framework Using Motion Features for Human Pose Estimation. 28 sept 2014 ; Disponible sur: <u>http://arxiv.org/abs/1409.7963</u>
- [31] Kanazawa A, Black MJ, Jacobs DW, Malik J. End-to-End Recovery of Human Shape and Pose. In: Conference on Computer Vision and Pattern Recognition [Internet]. Salt Lake City UT; IEEE; 2018. p. 7122-31. Disponible sur: <u>https://ieeexplore.ieee.org/document/8578842/</u>
- [32] Karellas A, Vedantham S, Suryanarayanan S. System and method for x-rays fluoroscopic imaging. 2005;109.
- [33] Knossow D. Paramétrage et Capture Multicaméras du Mouvement Humain. Interface hommemachine. Institut National Polytechnique de Grenoble - INPG, 2007. :275.

- [34] Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. Medical Image Analysis. janv 2017;35:303-12.
- [35] Kostrikov I, Gall J. Depth Sweep Regression Forests for Estimating 3D Human Pose from Images. In: Proceedings of the British Machine Vision Conference 2014 [Internet]. Nottingham: British Machine Vision Association; 2014. p. 80.1-80.13. Disponible sur: http://www.bmva.org/bmvc/2014/papers/paper060/index.html
- [36] Loper M, Mahmood N, Romero J, Pons-Moll G, Black MJ. SMPL: a skinned multi-person linear model. ACM Transactions on Graphics. 4 nov 2015;34(6):1-16.
- [37] Luo Y, Ren J, Wang Z, Sun W, Pan J, Liu J, et al. LSTM Pose Machines. 9 mars 2018; Disponible sur: <u>http://arxiv.org/abs/1712.06316</u>
- [38] Mahesh M. Fluoroscopy: Patient Radiation Exposure Issues. RadioGraphics. juill 2001;21(4):1033-45.
- [39] Maruyama T, Yamamoto H. Positioning simulator of radiography. In: International Conference on Imaging Systems and Techniques [Internet]. Thessaloniki, Greece: IEEE; 2010. p. 161-5. Disponible sur: <u>http://ieeexplore.ieee.org/document/5548521/</u>
- [40] Mehta D, Sridhar S, Sotnychenko O, Rhodin H, Shafiei M, Seidel H-P, et al. VNect: real-time 3D human pose estimation with a single RGB camera. ACM Transactions on Graphics. 20 juill 2017;36(4):1-14.
- [41] Moon G, Chang JY, Lee KM. V2V-PoseNet: Voxel-to-Voxel Prediction Network for Accurate 3D Hand and Human Pose Estimation from a Single Depth Map. 16 août 2018; Disponible sur: http://arxiv.org/abs/1711.07399
- [42] Nakano N, Sakura T, Ueda K, Omura L, Kimura A, Iino Y, et al. Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras [Internet]. Neuroscience; 2019 nov. Disponible sur: <u>http://biorxiv.org/lookup/doi/10.1101/842492</u>
- [43] Nie X, Feng J, Xing J, Yan S. Generative Partition Networks for Multi-Person Pose Estimation. 28 nov 2017; Disponible sur: <u>http://arxiv.org/abs/1705.07422</u>
- [44] Nie X, Zhang J, Yan S, Feng J. Single-Stage Multi-Person Pose Machines. 24 août 2019; Disponible sur: <u>http://arxiv.org/abs/1908.09220</u>
- [45] Ning G, Liu P, Fan X, Zhang C. A Top-down Approach to Articulated Human Pose Estimation and Tracking. 22 janv 2019; Disponible sur: <u>http://arxiv.org/abs/1901.07680</u>
- [46] Nunes RB, Amaral DT, Oliveira VS de. Propedêutica radiológica do impacto femoroacetabular em tempos de tomografia computadorizada e ressonância magnética: o que o radiologista precisa saber. Radiologia Brasileira. août 2011;44(4):249-55.
- [47] Omran M, Lassner C, Pons-Moll G, Gehler P, Schiele B. Neural Body Fitting: Unifying Deep Learning and Model Based Human Pose and Shape Estimation. In: International Conference on 3D Vision (3DV) [Internet]. Verona: IEEE; 2018. p. 484-94. Disponible sur: https://ieeexplore.ieee.org/document/8491000/
- [48] Organisation Mondiale de la Santé site internet : <u>https://www.who.int/news-room/feature-stories/detail/to-x-ray-or-not-to-x-ray-</u>
- [49] Pavlakos G, Zhou X, Derpanis KG, Daniilidis K. Harvesting Multiple Views for Marker-less 3D Human Pose Annotations. 16 avr 2017 ; Disponible sur: http://arxiv.org/abs/1704.04793

- [50] Phommahavong S, Haas D, Yu J, Krüger-Ziolek S, Möller K, Kretschmer J. Evaluating the microsoft kinect skeleton joint tracking as a tool for home-based physiotherapy. Current Directions in Biomedical Engineering. 1 sept 2015;1(1):184-7.
- [51] Pishchulin L, Insafutdinov E, Tang S, Andres B, Andriluka M, Gehler P, et al. DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation. 26 avr 2016; Disponible sur: http://arxiv.org/abs/1511.06645
- [52] Realini X, Amez-Droz M, Gremion I. Mandat "Profil professionnel du Technicien en Radiologie Médicale" rapport final. Lausanne. 18 juillet 2008 :81.
- [53] Regazzoni D, Vitali A, Rizzi C, Colombo G. A Method to Analyse Generic Human Motion With Low-Cost Mocap Technologies. In: Volume 1B: 38th Computers and Information in Engineering Conference [Internet]. Quebec City, Quebec, Canada: American Society of Mechanical Engineers; 2018. Disponible sur: <u>https://asmedigitalcollection.asme.org/IDETC-CIE/proceedings/IDETC-CIE2018/51739/Quebec%20City,%20Quebec,%20Canada/273607</u>
- [54] Rodríguez-Ruiz A, Krupinski E, Mordang J-J, Schilling K, Heywang-Köbrunner SH, Sechopoulos I, et al. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology. févr 2019;290(2):305-14.
- [55] Saunders M, Budden A, MacIver F, Teunis M, Warren-Forward H. Dose implications of fluoroscopy-guided positioning (FGP) for lumbar spine examinations prior to acquiring plain film radiographs. The British Journal of Radiology. févr 2005;78(926):130-4
- [56] Schmid J, Chênes C, Chagué S, Hoffmeyer P, Christofilopoulos P, Bernardoni M, et al. MyHip: supporting planning and surgical guidance for a better total hip arthroplasty: A pilot study. International Journal of Computer Assisted Radiology and Surgery. oct 2015;10(10):1547-56.
- [57] Sherer MAS, Paula J. Visconti PDD, Ritenour ER, Kelli Haynes MRT. Radiation Protection in Medical Radiography [Internet]. Elsevier Health Sciences; 2013. Disponible sur: https://books.google.ch/books?id=IUBYAQAAQBAJ
- [58] Shotton J, Fitzgibbon A, Cook M, Sharp T, Finocchio M, Moore R, et al. Real-Time Human Pose Recognition in Parts from Single Depth Images. 2011 :8
- [59] Soderstrom T, Hall L, Nilsson T, Ahlqvist J. How Does Collaborative 3D Screen-Based Computer Simulation Training Influence Diagnostic Skills of Radiographic Images and Peer Communication? Contemporary Educational Technology [Internet]. 1 déc 2012 ;3(4). Disponible sur: https://www.cedtech.net/article/how-does-collaborative-3d-screen-based-computer-simulation-training-influence-diagnostic-skills-of-6085
- [60] Stahl CM, Meisinger QC, Andre MP, Kinney TB, Newton IG. Radiation Risk to the Fluoroscopy Operator and Staff. American Journal of Roentgenology. oct 2016;207(4):737-44.
- [61] Sun K, Xiao B, Liu D, Wang J. Deep High-Resolution Representation Learning for Human Pose Estimation. 25 févr 2019 ; Disponible sur: <u>http://arxiv.org/abs/1902.09212</u>
- [62] Tang W, Yu P, Wu Y. Deeply Learned Compositional Models for Human Pose Estimation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision – ECCV 2018 [Internet]. Cham: Springer International Publishing; 2018. p. 197 214. Disponible sur: <u>http://link.springer.com/10.1007/978-3-030-01219-9_12</u>
- [63] Thakur Y, Bjarnason TA, Hammerstrom K, Marchinkow L, Koch T, Aldrich JE. Assessment of Patient Doses in CR Examinations Throughout a Large Health Region. Journal of Digital Imaging. févr 2012;25(1):189-95.

- [64] Tome D, Toso M, Agapito L, Russell C. Rethinking Pose in 3D: Multi-stage Refinement and Recovery for Markerless Motion Capture. 4 août 2018; Disponible sur: http://arxiv.org/abs/1808.01525
- [65] Toshev A, Szegedy C. DeepPose: Human Pose Estimation via Deep Neural Networks. IEEE Conference on Computer Vision and Pattern Recognition. juin 2014;1653-60.
- [66] Turing AM. Computing Machinery and Intelligence. In: Epstein R, Roberts G, Beber G, éditeurs. Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer [Internet]. Dordrecht: Springer Netherlands; 2009. p. 23-65. Disponible sur: https://doi.org/10.1007/978-1-4020-6710-5_3
- [67] Unberath M, Zaech J-N, Lee SC, Bier B, Fotouhi J, Armand M, et al. DeepDRR A Catalyst for Machine Learning in Fluoroscopy-Guided Procedures. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, éditeurs. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018 [Internet]. Cham: Springer International Publishing; 2018. p. 98-106. Disponible sur: <u>http://link.springer.com/10.1007/978-3-030-00937-3_12</u>
- [68] Van Rikxoort EM, van Ginneken B. Automated segmentation of pulmonary structures in thoracic computed tomography scans: a review. Physics in Medicine and Biology. 7 sept 2013;58(17):R187-220.
- [69] Varol G, Ceylan D, Russell B, Yang J, Yumer E, Laptev I, et al. BodyNet: Volumetric Inference of 3D Human Body Shapes. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision [Internet]. Cham: Springer International Publishing; 2018. p. 20-38. Disponible sur: <u>http://link.springer.com/10.1007/978-3-030-01234-2_2</u>
- [70] Wang Y, Aghaei F, Zarafshani A, Qiu Y, Qian W, Zheng B. Computer-aided classification of mammographic masses using visually sensitive image features. Journal of X-Ray Science and Technology. 28 janv 2017;25(1):171-86.
- [71] Waymel Q, Badr S, Demondion X, Cotten A, Jacques T. Impact of the rise of artificial intelligence in radiology: What do radiologists think? Diagnostic and Interventional Imaging [Internet]. mai 2019 ; Disponible sur: <u>https://linkinghub.elsevier.com/retrieve/pii/S2211568419300907</u>
- [72] Wei S-E, Ramakrishna V, Kanade T, Sheikh Y. Convolutional Pose Machines. 11 avr 2016; Disponible sur: <u>http://arxiv.org/abs/1602.00134</u>
- [73] Zhang Z. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence. nov 2000;22(11):1330-4.
- [74] Zhu A. Articulated human pose estimation in images and video. Université de Technologie de Troyes. 2016:138.

9. Annexes

9.1. Listes non-exhaustives des méthodes IA de détection de pose

humaine

Systèmes d'IA 2D :

- Chen Y, Shen C, Wei X-S, Liu L, Yang J. Adversarial PoseNet: A Structure-aware Convolutional Network for Human Pose Estimation. 1 mai 2017; Disponible sur: <u>http://arxiv.org/abs/1705.00389</u>
- Wei S-E, Ramakrishna V, Kanade T, Sheikh Y. Convolutional Pose Machines. 11 avr 2016; Disponible sur: <u>http://arxiv.org/abs/1602.00134</u>
- Sun K, Xiao B, Liu D, Wang J. Deep High-Resolution Representation Learning for Human Pose Estimation. 25 févr 2019; Disponible sur: <u>http://arxiv.org/abs/1902.09212</u>
- Pishchulin L, Insafutdinov E, Tang S, Andres B, Andriluka M, Gehler P, et al. DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation. 26 avr 2016; Disponible sur: <u>http://arxiv.org/abs/1511.06645</u>
- Insafutdinov E, Pishchulin L, Andres B, Andriluka M, Schiele B. DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model. 30 nov 2016; Disponible sur: <u>http://arxiv.org/abs/1605.03170</u>
- Tang W, Yu P, Wu Y. Deeply Learned Compositional Models for Human Pose Estimation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision [Internet]. Cham: Springer International Publishing; 2018. p. 197-214. Disponible sur: http://link.springer.com/10.1007/978-3-030-01219-9 12
- Toshev A, Szegedy C. DeepPose: Human Pose Estimation via Deep Neural Networks. Conference on Computer Vision and Pattern Recognition. juin 2014;1653-60.
- Tompson J, Goroshin R, Jain A, LeCun Y, Bregler C. Efficient object localization using Convolutional Networks. In: Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Boston, MA, USA: 2015. p. 648-56. Disponible sur: http://ieeexplore.ieee.org/document/7298664/
- Pfister T, Charles J, Zisserman A. Flowing ConvNets for Human Pose Estimation in Videos. 8 nov 2015; Disponible sur: <u>http://arxiv.org/abs/1506.02897</u>
- Sun K, Lan C, Xing J, Zeng W, Liu D, Wang J. Human Pose Estimation using Global and Local Normalization. 21 sept 2017; Disponible sur: <u>http://arxiv.org/abs/1709.07220</u>
- Carreira J, Agrawal P, Fragkiadaki K, Malik J. Human Pose Estimation with Iterative Error Feedback. 12 juin 2016; Disponible sur: <u>http://arxiv.org/abs/1507.06550</u>
- Nie X, Feng J, Zuo Y, Yan S. Human Pose Estimation with Parsing Induced Learner. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition [Internet]. Salt Lake City, UT, USA; 2018. p. 2100-8. Disponible sur: <u>https://ieeexplore.ieee.org/document/8578322/</u>
- Tompson J, Jain A, LeCun Y, Bregler C. Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation. 17 sept 2014; Disponible sur: http://arxiv.org/abs/1406.2984
- Ning G, Zhang Z, He Z. Knowledge-Guided Deep Fractal Neural Networks for Human Pose Estimation. 8 août 2017; Disponible sur: <u>http://arxiv.org/abs/1705.02407</u>

- Yang W, Li S, Ouyang W, Li H, Wang X. Learning Feature Pyramids for Human Pose Estimation. 3 août 2017; Disponible sur: <u>http://arxiv.org/abs/1708.01101</u>
- Jain A, Tompson J, Andriluka M, Taylor GW, Bregler C. Learning Human Pose Estimation Features with Convolutional Networks. 23 avr 2014; Disponible sur: <u>http://arxiv.org/abs/1312.7302</u>
- Luo Y, Ren J, Wang Z, Sun W, Pan J, Liu J, et al. LSTM Pose Machines. 9 mars 2018; Disponible sur: <u>http://arxiv.org/abs/1712.06316</u>
- Jain A, Tompson J, LeCun Y, Bregler C. MoDeep: A Deep Learning Framework Using Motion Features for Human Pose Estimation. 28 sept 2014; Disponible sur: <u>http://arxiv.org/abs/1409.7963</u>
- Chu X, Yang W, Ouyang W, Ma C, Yuille AL, Wang X. Multi-Context Attention for Human Pose Estimation. 23 févr 2017; Disponible sur: <u>http://arxiv.org/abs/1702.07432</u>
- Ke L, Chang M-C, Qi H, Lyu S. Multi-Scale Structure-Aware Network for Human Pose Estimation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision ;Cham: Springer International Publishing; 2018. p. 731-46. Disponible sur: http://link.springer.com/10.1007/978-3-030-01216-8 44
- Osokin D. Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose. 29 nov 2018; Disponible sur: <u>http://arxiv.org/abs/1811.12004</u>
- Cao Z, Simon T, Wei S-E, Sheikh Y. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. 13 avr 2017; Disponible sur: <u>http://arxiv.org/abs/1611.08050</u>
- Belagiannis V, Zisserman A. Recurrent Human Pose Estimation. 5 août 2017; Disponible sur: <u>http://arxiv.org/abs/1605.02914</u>
- Chou C-J, Chien J-T, Chen H-T. Self Adversarial Training for Human Pose Estimation. 15 août 2017; Disponible sur: <u>http://arxiv.org/abs/1707.02439</u>
- Xiao B, Wu H, Wei Y. Simple Baselines for Human Pose Estimation and Tracking. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision. Cham: Springer International Publishing; 2018. p. 472-87. Disponible sur: <u>http://link.springer.com/10.1007/978-3-030-01231-</u> <u>1_29</u>
- Newell A, Yang K, Deng J. Stacked Hourglass Networks for Human Pose Estimation. 26 juill 2016; Disponible sur: <u>http://arxiv.org/abs/1603.06937</u>
- Papandreou G, Zhu T, Kanazawa N, Toshev A, Tompson J, Bregler C, et al. Towards Accurate Multi-person Pose Estimation in the Wild. 14 avr 2017; Disponible sur: <u>http://arxiv.org/abs/1701.01779</u>

Systèmes d'IA 3D :

- Yang W, Ouyang W, Wang X, Ren J, Li H, Wang X. 3D Human Pose Estimation in the Wild by Adversarial Learning. 16 avr 2018; Disponible sur: <u>http://arxiv.org/abs/1803.09722</u>
- Pavllo D, Feichtenhofer C, Grangier D, Auli M. 3D human pose estimation in video with temporal convolutions and semi-supervised training. 29 mars 2019; Disponible sur: <u>http://arxiv.org/abs/1811.11742</u>

- Martinez J, Hossain R, Romero J, Little JJ. A simple yet effective baseline for 3d human pose estimation. 4 août 2017 ; Disponible sur: <u>http://arxiv.org/abs/1705.03098</u>
- Varol G, Ceylan D, Russell B, Yang J, Yumer E, Laptev I, et al. BodyNet: Volumetric Inference of 3D Human Body Shapes. 18 août 2018 ; Disponible sur: <u>http://arxiv.org/abs/1804.04875</u>
- Pavlakos G, Zhou X, Derpanis KG, Daniilidis K. Coarse-to-Fine Volumetric Prediction for Single-Image 3D Human Pose. 26 juill 2017 ; Disponible sur: <u>http://arxiv.org/abs/1611.07828</u>
- Sun X, Shang J, Liang S, Wei Y. Compositional Human Pose Regression. 1 août 2017; Disponible sur: <u>http://arxiv.org/abs/1704.00159</u>
- Neverova N, Guler RA, Kokkinos I. Dense Pose Transfer. 6 sept 2018; Disponible sur: <u>http://arxiv.org/abs/1809.01995</u>
- Güler RA, Neverova N, Kokkinos I. DensePose: Dense Human Pose Estimation In The Wild. 1 févr 2018; Disponible sur: <u>http://arxiv.org/abs/1802.00434</u>
- Wang M, Chen X, Liu W, Qian C, Lin L, Ma L. DRPose3D: Depth Ranking in 3D Human Pose Estimation. 24 mai 2018; Disponible sur: <u>http://arxiv.org/abs/1805.08973</u>
- Kanazawa A, Black MJ, Jacobs DW, Malik J. End-to-end Recovery of Human Shape and Pose. 23 juin 2018; Disponible sur: <u>http://arxiv.org/abs/1712.06584</u>
- Sun X, Xiao B, Wei F, Liang S, Wei Y. Integral Human Pose Regression. 18 sept 2018; Disponible sur: <u>http://arxiv.org/abs/1711.08229</u>
- Dabral R, Mundhada A, Kusupati U, Afaque S, Sharma A, Jain A. Learning 3D Human Pose from Structure and Motion. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision. Cham: Springer International Publishing; 2018. p. 679-96. Disponible sur: <u>http://link.springer.com/10.1007/978-3-030-01240-3_41</u>
- Pavlakos G, Zhu L, Zhou X, Daniilidis K. Learning to Estimate 3D Human Pose and Shape from a Single Color Image. In: Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT: IEEE; 2018. p. 459-68. Disponible sur: <u>https://ieeexplore.ieee.org/document/8578153/</u>
- Tome D, Russell C, Agapito L. Lifting from the Deep: Convolutional 3D Pose Estimation from a Single Image. Conference on Computer Vision and Pattern Recognition (CVPR). juill 2017;5689-98.
- Mehta D, Rhodin H, Casas D, Fua P, Sotnychenko O, Xu W, et al. Monocular 3D Human Pose Estimation In The Wild Using Improved CNN Supervision. 4 oct 2017; Disponible sur: <u>http://arxiv.org/abs/1611.09813</u>
- Omran M, Lassner C, Pons-Moll G, Gehler PV, Schiele B. Neural Body Fitting: Unifying Deep Learning and Model-Based Human Pose and Shape Estimation. 17 août 2018; Disponible sur: <u>http://arxiv.org/abs/1808.05942</u>
- Lin M, Lin L, Liang X, Wang K, Cheng H. Recurrent 3D Pose Sequence Machines. 30 juill 2017 ; Disponible sur: <u>http://arxiv.org/abs/1707.09695</u>
- Mehta D, Sotnychenko O, Mueller F, Xu W, Sridhar S, Pons-Moll G, et al. Single-Shot Multi-Person 3D Pose Estimation From Monocular RGB. 28 août 2018; Disponible sur: <u>http://arxiv.org/abs/1712.03453</u>
- Tekin B, Katircioglu I, Salzmann M, Lepetit V, Fua P. Structured Prediction of 3D Human Pose with Deep Neural Networks. 17 mai 2016; Disponible sur: <u>http://arxiv.org/abs/1605.05180</u>

- Zhou X, Huang Q, Sun X, Xue X, Wei Y. Towards 3D Human Pose Estimation in the Wild: a Weakly-supervised Approach. 30 juill 2017; Disponible sur: <u>http://arxiv.org/abs/1704.02447</u>
- Rhodin H, Salzmann M, Fua P. Unsupervised Geometry-Aware Representation for 3D Human Pose Estimation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, éditeurs. Computer Vision . Cham: Springer International Publishing; 2018. p. 765-82. Disponible sur: <u>http://link.springer.com/10.1007/978-3-030-01249-6_46</u>
- Mehta D, Sridhar S, Sotnychenko O, Rhodin H, Shafiei M, Seidel H-P, et al. VNect: real-time 3D human pose estimation with a single RGB camera. ACM Transactions on Graphics. 20 juill 2017;36(4):1-14.

9.2. Principales bases de données pour l'estimation de pose humaine

Base de donnée	Nombres de poses annotées	Multi personnes	Vidéos	Type de données
LSP [76]	2000			Sports
LSP Extended [80]	10000			Sports
MPII Single Person [77]	26429			Divers
FLIC [82]	5003			Films
FashionPose [79]	7305			Blog de mode
We Are Family [87]	3131	Х		Photos de groupe
MPII Multi- Person [75]	14993	Х		Divers
MS COCO KeyPoints [81]	105698	Х		Divers
Penn Action [78]	159633		Х	Sports
JHMDB [86]	31838		Х	Divers
Youtube Pose [84]	5000		Х	Divers
Video Pose 2.0 [83]	1286		Х	Séries télé
Multi-Person PoseTrack [85]	16219	Х	Х	Divers

Tableau 11 : Bases de données publiques disponibles pour l'estimation de pose humaine (liste non-exhaustive)

- [75] Andriluka M, Pishchulin L, Gehler P, Schiele B. 2D Human Pose Estimation: New Benchmark and State of the Art Analysis. In: Conference on Computer Vision and Pattern Recognition [Internet]. Columbus, OH, USA: IEEE; 2014 p. 3686-93. Disponible sur: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6909866
- [76] Johnson S, Everingham M. Clustered Pose and Nonlinear Appearance Models for Human Pose Estimation. In: Proceedings of the British Machine Vision Conference 2010 [Internet]. Aberystwyth: British Machine Vision Association; 2010. p. 12.1-12.11. Disponible sur: http://www.bmva.org/bmvc/2010/conference/paper12/index.html
- [77] Pishchulin L, Andriluka M, Schiele B. Fine-grained Activity Recognition with Holistic and Pose based Features. 28 juill 2014; Disponible sur: http://arxiv.org/abs/1406.1881
- [78] Zhang W, Zhu M, Derpanis KG. From Actemes to Action: A Strongly-Supervised Representation for Detailed Action Understanding. In: International Conference on Computer Vision [Internet]. Sydney, Australia: IEEE; 2013. p. 2248-55. Disponible sur: http://ieeexplore.ieee.org/document/6751390/
- [79] Dantone M, Gall J, Leistner C, Van Gool L. Human Pose Estimation Using Body Parts Dependent Joint Regressors. In: Conference on Computer Vision and Pattern Recognition [Internet]. Portland, OR, USA: IEEE; 2013. p. 3041-8. Disponible sur: http://ieeexplore.ieee.org/document/6619235/

- [80] Johnson S, Everingham M. Learning effective human pose estimation from inaccurate annotation. In: Colorado Springs, CO, USA: IEEE; 2011. p. 1465-72. Disponible sur: http://ieeexplore.ieee.org/document/5995318/
- [81] Lin T-Y, Maire M, Belongie S, Bourdev L, Girshick R, Hays J, et al. Microsoft COCO: Common Objects in Context. 20 févr 2015; Disponible sur: http://arxiv.org/abs/1405.0312
- [82] Sapp B, Taskar B. MODEC: Multimodal Decomposable Models for Human Pose Estimation. In: Conference on Computer Vision and Pattern Recognition [Internet]. Portland, OR, USA: IEEE; 2013. p. 3674-81. Disponible sur: http://ieeexplore.ieee.org/document/6619315/
- [83] Sapp B, Weiss D, Taskar B. Parsing human motion with stretchable models. In: Colorado Springs, CO, USA: IEEE; 2011. p. 1281-8. Disponible sur: http://ieeexplore.ieee.org/document/5995607/
- [84] Charles J, Pfister T, Magee D, Hogg D, Zisserman A. Personalizing Human Video Pose Estimation. In: Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Las Vegas, NV, USA: IEEE; 2016. p. 3063-72. Disponible sur: http://ieeexplore.ieee.org/document/7780703/
- [85] Iqbal U, Milan A, Gall J. PoseTrack: Joint Multi-person Pose Estimation and Tracking. In: Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Honolulu, HI: IEEE; 2017. p. 4654-63. Disponible sur: http://ieeexplore.ieee.org/document/8099978/
- [86] Jhuang H, Gall J, Zuffi S, Schmid C, Black MJ. Towards Understanding Action Recognition. In: 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE; 2013. p. 3192-9. Disponible sur: http://ieeexplore.ieee.org/document/6751508/
- [87] Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, Mitchell JC, et al. We Are Family: Joint Pose Estimation of Multiple Persons. In: Daniilidis K, Maragos P, Paragios N, éditeurs. Computer Vision. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010. p. 228-42. Disponible sur: http://link.springer.com/10.1007/978-3-642-15549-9_17

9.3. Formulaire de consentement





Master of Science conjoint HES-SO - UNIL en Sciences de la santé Av. de Provence 6 CH-1007 Lausanne

Formulaire de consentement

Autorisation pour l'enregistrement photo/vidéo et l'exploitation des données enregistrées

Présentation :

Ces enregistrements sont réalisés dans le cadre d'un mémoire de recherche en imagerie médicale.

Ce travail de master porte sur la détection du positionnement des patients en radiologie diagnostique à l'aide d'un système d'intelligence artificielle.

Cette recherche est menée par Meryam Butt et encadrée par Prof. Jérôme Schmid de la Haute école de santé de Genève. Elle n'est néanmoins possible que grâce au consentement des personnes qui acceptent d'être enregistrées, à qui nous demandons par conséquent une autorisation écrite.

Autorisation :

Je soussigné(e) _____

autorise par la présente à utiliser à titre gratuit ma prestation photo et/ou vidéo

- pour les besoins de ce travail de master, uniquement au sein de l'établissement HES-SO Master.
- pour une utilisation plus large dans un cadre institutionnel, à l'extérieur de l'établissement, à des fins de recherches scientifiques (thèses, conférences, congrès, séminaires, colloques, articles scientifiques)

Pour toutes ces utilisations, les photos et vidéos enregistrées seront anonymisées, ceci signifie que les visages seront floutés, et aucune information pouvant porter à l'identification des participants ne sera dévoilée.

Fait à _____, le _____, le _____,

Signature

9.4. Détails des incidences radiologiques utilisées pour l'étude principale

Thorax de face PA :

Position du patient : Patient debout thorax en appui contre la plaque dos contre le potter (PA), menton relevé, mains sur les hanches, épaules en avant pour bien dégager les scapulas Centrage : rayon directeur perpendiculaire au potter, centrage sur la ligne médiale à hauteur de la pointe des scapulas Collimation : latéralement limite de la peau

Thorax de profil :

Position du patient : Patient debout, côté gauche contre la plaque, bras relevés et croisés audessus de la tête, menton légèrement relevé

Centrage : rayon directeur perpendiculaire au potter, 10cm en dessous de l'aisselle au milieu de la distance antéro-postérieur ou à hauteur de la pointe de la scapula

Collimation : latéralement limite de la peau, limite supérieure au niveau de l'aisselle ou ras du cou

Cervicale de face :

Position du patient : Patient debout, dos contre la plaque, bras le long du corps, menton légèrement relevé

Centrage : commencer avec le rayon droit centré au niveau du manubrium sternal puis incliné en caudo-crânial jusqu'à la proéminence laryngée. Rayon incliné d'environ 15°

Collimation : latéralement limite de la peau, limite supérieure au niveau du menton, limite inférieur 5cm en dessous du manubrium sternal.

Cervicale de profil :

Position du patient : patient debout, côté droit ou gauche contre la plaque, épaule relâchées et abaissées, menton légèrement relevé, tête bien droite.

Centrage : rayon directeur perpendiculaire au niveau de la proéminence laryngée Collimation : de part et d'autre des limites de la peau en latéral. Bord supérieur au conduit auditif externe, bord inférieur au niveau de l'épaule.

Épaule de face AP, rotation neutre, interne :

Position du patient : Patient debout ou assis, dos contre la plaque, angle de 45° du coté à examiner par rapport au détecteur

Rotation neutre : main le long du corps
Rotation interne : main sur le ventre (pronation)

Centrage : rayon directeur incliné de 20° à 25° caudalement, centré sur la fossette sous acromiale

Collimation : diaphragmer strictement à la zone d'intérêt

Profil de Neer :

Position du patient : Patient debout, ventre contre la plaque, angle de 45° du coté à examiner et l'épaule du côté à examiner contre la plaque, bras fléchi a 90°, paume de la main sur le ventre

Centrage : rayon directeur cranio-caudal de 10° à 15°, bord médial de la scapula, 2 travers de doigt sous l'épine de la scapula

Collimation : diaphragmer jusqu'à la pointe de la scapula

Genou de profil en charge :

Position du patient : Patient debout, genou à radiographier contre la plaque, flexion de 20° et talon éloigné du détecteur de quelques degrés

Centrage : milieu du genou à la hauteur du bord sup de la patella, rayon directeur ascendant de 5 à 10 ° pour superposer la face inférieure des deux condyles fémoraux

Collimation : latéralement aux limites de la peau, du tiers inférieur du fémur jusqu'au tiers supérieur de la jambe

Bassin de face AP :

Position du patient : Patient debout, PSM aligné, dos contre la plaque, les mains sur la poitrine, rotation interne des pieds de 15° (écarté les talons et touché la pointe des pieds entre eux) Centrage : sur le PSM, 3 à 5cm au-dessus du bord supérieur de la symphyse pubienne Collimation : des crêtes iliaques aux 1/3 supérieur des fémurs et en latéral selon corpulences du patient, visible au moins 1cm sur les 4 côtés du potter

Faux profil de Lequesne :

Position du patient : debout patient en oblique postérieur 25° (le dos du patient fait un angle de 65° avec la table)

Centrage : entre les têtes fémorales, rayon droit

Collimation : diaphragmer jusqu'aux limites de la peau

Lombaire de face :

Position du patient : patient debout, dos contre la plaque, membres supérieurs écartés de la zone d'intérêt

Centrage : rayon directeur perpendiculaire au potter, 2 travers de doigt au-dessus de la ligne bi-crêtes

Collimation : toute la longueur du rachis lombaire

Lombaire de profil :

Position du patient : patient debout, pieds légèrement écartés, membres supérieurs croisés et surélevés

Centrage : rayon directeur perpendiculaire au potter, 2 travers de doigt au-dessus du milieu de la crête iliaque et 4 travers de doigt en avant du plan postérieur

Collimation : toute la longueur du rachis lombaire

Toutes ces informations sont tirées du livre :

Bontrager, K. L., Lampignano, J. P., & Bontrager, K. L. Bontrager's handbook of radiographic positioning and techniques (7th ed). St. Louis, Mo: Mosby/Elsevier. 2010.

9.5. Radiographies virtuelles

La Figure 27 illustre d'autres exemples de génération de radiographies virtuelles du bassin de face avec des changements au niveau des paramètres d'acquisitions.



Figure 27 : Autres exemples de radiographies du bassin. A) 80 kV, 50 mAs, 120cm, détecteur 43 x 43 cm, pas de rayonnement diffusé ; B) 80 kV, 50 mAs, 120cm, détecteur 43 x 43 cm, avec du rayonnement diffusé ; C) 80 kV, 20 mAs, 120cm, détecteur 43 x 43 cm, avec du rayonnement diffusé ; C) 80 kV, 20 mAs, 120cm, détecteur 43 x 43 cm, avec du rayonnement diffusé ; D) 120 kV, 50 mAs, 120cm, détecteur 43 x 43 cm, avec de rayonnement diffusé .



La Figure 28 présente le cas où l'image CT présente une prothèse totale de hanche à droite. Cet exemple illustre l'avantage du scénario pédagogique qui permet aussi de choisir le CT en fonction de la présence de corps étrangers ou de différentes pathologies que les étudiants seront amenés à rencontrer lors de leurs pratiques professionnelles.

Figure 28 : Autre exemple de radiographie du bassin avec un CT en présence d'une prothèse de hanche droite. A noter que des hyperintensités sont présentes avec le fenêtrage choisi, notamment au niveau de la tête fémorale de l'implant. Ceux-ci sont en partie dus au fait que dans le CT des artéfacts métalliques sont présents et qu'ils influencent donc la radiographie virtuelle générée.

La Figure 29 illustre un exemple d'une radiographie de bassin de face sur un sujet masculin. En fonction de la diversité des CT récoltés nous pouvons nous rapprocher de la réalité des sujets.



Figure 29 : exemple de la détection et la génération d'une radiographie virtuelle du bassin de face sur un sujet masculin

Un dernier exemple illustrant des radiographies de thorax de face (Figure 32) avec un bon et un mauvais positionnement du sujet lors des acquisitions (Figure 30) avec le squelette OpenPose estimé grâce à la reconstruction 3D (Figure 31).



Figure 30 : exemple d'un sujet bien positionné pour une incidence de thorax de face (A) et mal positionné (B) à savoir le centrage est trop haut et le sujet n'est pas collé au détecteur



Figure 31 : squelettes 3D reconstruits pour des radiographies du thorax de face. A) et C) représentent des vues pour un bon positionnement du sujet ; B) et D) représentent des vues pour un mauvais positionnement, dans lequel le centrage est trop haut et le sujet n'est pas assez collé au détecteur.



Figure 32 : deux exemples de radiographies virtuelles générées en fonction d'un bon positionnement (à gauche) et d'un mauvais positionnement (à droite). A noter que la zone coupée au-dessus de la radiographie virtuelle est due aux limites du champ d'acquisition du CT.