Titre: Title:	Towards a Toolbox for Quantitative and Automatic Evaluation of 3D Trunk Motion in Patients with AIS
Auteur: Author:	Ghazal Ebrahimi
Date:	2023
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Ebrahimi, G. (2023). Towards a Toolbox for Quantitative and Automatic Evaluation of 3D Trunk Motion in Patients with AIS [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/57091/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/57091/
Directeurs de recherche: Advisors:	Lama Séoud, & Carole Fortin
Programme: Program:	Génie informatique

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Towards A Toolbox for Quantitative and Automatic Evaluation of 3D Trunk Motion in Patients with AIS

GHAZAL EBRAHIMI
Département de génie informatique et génie logiciel

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

Décembre 2023

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Ce mémoire intitulé :

Towards A Toolbox for Quantitative and Automatic Evaluation of 3D Trunk Motion in Patients with AIS

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

Farida CHERIET, présidente
Lama SÉOUD, membre et directrice de recherche
Carole FORTIN, membre et codirectrice de recherche
Stefan PARENT, membre externe

DEDICATION

To my family and friends...

ACKNOWLEDGEMENTS

Before anything, I would like to acknowledge and express my appreciation toward the significant work of my colleagues: Fatima Gharsallah, Robert Maurice Etoumbe, and Léa Drolet-Roy, who helped with the data collection for this project during the summers of 2023 and 2024. I would also like to thank my supervisors, Lama Séoud and Carole Fortin, for their support and brilliant guidance during the course of this project. Finally, I want to thank my family in Iran, as well as my family here, for their support and encouragement throughout these two years.

RÉSUMÉ

La scoliose idiopathique de l'adolescent (SIA) est une déformation en trois dimensions de la colonne vertébrale et de la cage thoracique, qui affecte l'apparence générale du tronc. Si elle n'est pas traitée, la scoliose idiopathique de l'adolescent peut évoluer pendant la poussée de croissance et réduire la qualité de vie des patients. Les méthodes existantes d'évaluation de la scoliose, qu'elles soient que ce soit par radiographie ou topographie de surface, sont toutes basées sur des acquisitions statiques. Or, la colonne vertébrale est une structure articulée, qui permet la mobilité du tronc. L'évaluation du tronc sous des conditions cinématiques permet de comprendre la rigidité de la scoliose. L'exercice d'autocorrection et le test de flexion latérale sont deux exemples de mouvements utilisés par les physiothérapeutes et les orthopédistes respectivement, pour évaluer la flexibilité du tronc. Ces informations, bien que principalement qualitatives, sont prises en compte pour la planification du traitement. Ce projet est le premier à explorer la topographie de surface dynamique pour l'évaluation quantitative de la mobilité du tronc. À l'aide de modèles d'apprentissage profond 3D pré-entraînés, nous développons une boîte à outils pour la quantification automatique de la mobilité du tronc pendant des séquences de flexion latérale et d'exercices d'autocorrection. La boîte à outils que nous proposons offre des fonctionnalités telles que la détection de repères anatomiques de la surface du dos, la segmentation anatomique de la surface du dos, la quantification des paramètres de la surface du tronc au fil du temps, et la visualisation des trajectoires des points de repère au cours d'une séquence de mouvement. Malgré certaines limites, nos expérimentations montrent des résultats prometteurs quant à l'application clinique potentielle de la méthode proposée. Dans l'ensemble, ce travail constitue un premier pas vers l'automatisation de l'évaluation quantitative dynamique du mouvement du tronc et fournit une base pour les travaux futurs dans cette direction.

ABSTRACT

Adolescent idiopathic scoliosis (AIS) is a 3D deformity of the spine and the ribcage, that affects the general appearance of the trunk. If left untreated, AIS can progress during growth spurt and reduce patients' quality of life. Existing methods for assessing scoliosis, whether through radiographs or surface topography are all based on static acquisitions. However, the spine is an articulated structure allowing for the mobility of the trunk. The evaluation of the trunk under kinematic conditions provides insights on the rigidity of scoliosis. The self-correction exercise and the lateral bending test are two examples of motion used by physiotherapists and orthopedic surgeons respectively to evaluate the flexibility of the trunk. This information, although mostly qualitative, is considered for treatment planning. This project is the first to explore dynamic surface topography for the quantitative evaluation of trunk mobility. Using pre-trained 3D deep learning models, we develop a toolbox for an automatic quantification of trunk motion during sequences of lateral bending and self-correction exercises. Our proposed toolbox provides functionalities such as back surface anatomical landmark detection, back surface anatomical segmentation, quantifying trunk surface parameters over time, and visualizing landmarks' trajectories during a motion sequence. Despite some limitations, our experiments show promising results regarding the potential clinical application of the proposed toolbox. Overall, this work is the first step toward automating dynamic quantitative evaluation of the trunk motion and provides a basis for future work in this direction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Idiopathic scoliosis is a 3-dimensional deformity in the trunk with no known underlying cause and most commonly appears in adolescent population, thus referred to as Adolescent Idiopathic Scoliosis (AIS) [1]. The deformity is characterized by a spinal curve in the frontal plane and an axial rotation in the transverse plane. AIS is reported to affect 2-3% of adolescent population. Other than the aesthetic trunk deformity, AIS can result in functional problems such as mobility and breathing difficulties. If neglected and left untreated, AIS can get progressively worse and negatively affect patients' quality of life [1].

Treatment methods for AIS are categorized into invasive and non-invasive treatments. Invasive treatment plans such as surgery are advised for severe curves ($> 40°$), while non-invasive treatments such as bracing and physiotherapy are advised for milder curves ($< 40°$) [2]. Clinical assessment of AIS plays an important role in determining the treatment plan. Trunk mobility is an important parameter in decisions for treatment planning. In fact, curve reducibility and curve flexibility are assessed on lateral bending movement as part of surgical treatments' pre-operative analysis [3]. Trunk motion assessment is also an important part of physiotherapeutic scoliosis-specific exercises (PSSE). PSSE include corrective exercises such as self-correction exercise, where the patient recruits the back muscles to momentarily improve the posture by reducing the spinal curve [4]. This exercise is used by physiotherapists to assess trunk mobility and curve reducibility, and recommended for preventing curve progression and progressively improving patient's posture [4].

Over the years, there have been studies on quantitative analysis of trunk deformity in patients with AIS [2]. However, all of these studies are based on static acquisition of the trunk surface, and to our knowledge, there is no study that evaluates the dynamic trunk motion. Dynamic evaluation of the trunk motion through repeatable and accurate metrics has the potential to complement current AIS assessment by providing a quantification of trunk mobility, curve flexibility, and curve reducibility, which are determinant parameters in AIS treatment planning.

In recent years, 3D imaging technologies have matured and are becoming more affordable and available. A dynamic 3D surface topography system, available at CHU Sainte Justine (CHUSJ), allows to acquire a sequence of densely sampled 3D point clouds with a high temporal resolution. In parallel, with the advances in 3D data acquisition, 3D deep learning models are growing and improving each day and have shown strong performances on 3D point clouds for a variety of tasks.

In this project, we explore existing 3D deep learning models and the 3D dynamic imaging system at CHUSJ to develop quantitative tools for an automatic evaluation of the trunk motion in lateral bending and self-correction movements in patients with AIS. We propose an automatic anatomical landmark localization, an automatic anatomical back surface segmentation, a visual representation of the 3D motion, and an automatic computation of clinical metrics over time.

In this document, first, a literature review is presented in chapter 2. Next, the rationals and objectives are detailed in chapter 3. Chapter 4 includes the details of our methodology for the different parts of the project. Chapter 5 provides and discusses the results and findings. Finally, chapter 6 propose concluding remarks.

CHAPTER 2 LITERATURE REVIEW

This chapter aims at providing an overview of the clinical concepts and previous work in the literature, related to the research project. The chapter includes an overview of the anatomy of the back, a description of scoliosis, its assessment and treatment. More technically, it provides also an overview of 3D surface data acquisition and representation, deep learning models for 3D point clouds, and finally existing methods for automatic anatomical landmark localization and point cloud segmentation.

2.1 Anatomy of the Trunk

In this section, we take a look at the anatomy of the human's trunk. Trunk refers to the middle part of the body, and it includes the vertebral column, the ribcage, and the pelvis[5].

The vertebral column, or the spine, is a string of 33 vertebrae (small cube-like bones) extending from the base of the skull to the pelvis. The function of the vertebral column is to protect the spinal cord, stabilize the trunk, and provide mobility and flexibility [6, 5]. The vertebral column can be divided into five segments. Starting from the top, the 7 vertebrae positioned along the neck constitute the cervical segment (C1 to C7), the successive 12 vertebrae along the chest are collectively called the thoracic segment (T1 to T12), the lower back part is called the lumbar segment and consists of 5 vertebrae (L1 to L5), the sacrum, composed of 5 fused vertebrae, and the coccyx, composed of 4 fused vertebrae come next and are positioned near the pelvis [6, 5]. The vertebrae in the cervical, thoracic, and lumbar segments are separated by inter-vertebral disks that allow for the flexibility and mobility of the spine [6] (see figure 2.1).

The thoracic cage or the ribcage refers to the collection of distinct bones within the thorax region, including the 12 thoracic vertebrae, the 12 pairs of ribs, and the sternum [5]. The sternum is a long bone positioned at the front center of the ribcage. There are 24 ribs, 12 on the right and 12 on the left. All of them are connected to the 12 thoracic vertebrae (T1 to T12) at one end. The upper seven ribs are attached to the sternum in the front, forming close loops. The next three ribs, from T8 to T10, are connected to the upper ribs. And the last two ribs (T11 and T12) are not attached to anything in the front and are referred to as floating ribs. As the name suggests, the thoracic cage is structured like a cage. The cage-like structure serves to protect the vital organs that are located in the chest area (see figure 2.1).

Trunk muscles play an important role in maintaining the balance and spinal stability. Dif-

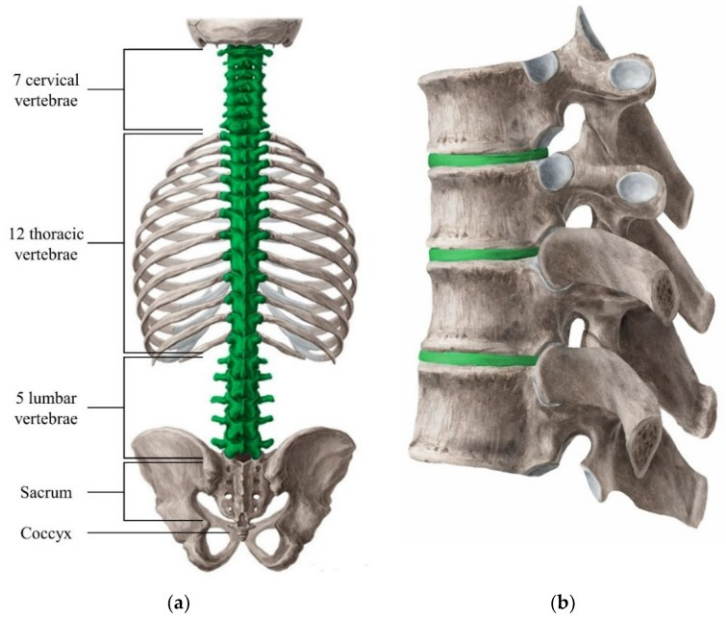


Figure 2.1 Skeleton of the spine and thoracic cage[6]

ferent muscle groups in the trunk area are responsible for motor control and adapting a stabilized posture [7]. Figure 2.2 displays a detailed illustration of the back muscles.

2.2 Adolescent Idiopathic Scoliosis

2.2.1 Definition

Scoliosis is a 3-dimensional deformity of the spine and the thorax. In the absence of a known underlying cause, the condition is referred to as idiopathic scoliosis [1]. Idiopathic scoliosis most commonly appears is adolescents, and therefor is referred to as Adolescent Idiopathic Scoliosis (AIS). The prevalence of AIS varies in different studies and populations, but it's widely reported to affect 2-3% of adolescents. It's more prevalent in girls than in boys, and in some cases, the curvature can progress as the individual grows, potentially causing health issues, decreased quality of life, cosmetic deformity and visible disability, pain and functional limitations if left untreated [1].

2.2.2 AIS Imaging Modalities

AIS is a complex deformity of the spine and 3D imaging modalities are required to comprehensively assess and quantify this 3D deformity. The most common imaging methods used in clinical practice are conventional radiography, bi-planar radiography, magnetic reso-

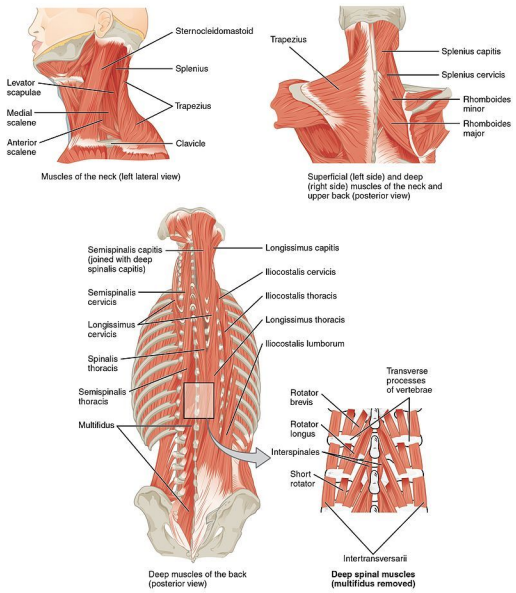


Figure 2.2 Muscles of the back[8]

nance imaging (MRI), computed tomography scans (CT), and surface topography [9]. This section aims to offer a comprehensive overview of these methods, outlining their respective advantages and disadvantages.

Radiography

Radiography is the most common and accessible modality for clinical assessment of AIS. The assessment is based on two 2D projections of the thorax: posterior–anterior (the coronal plane) and lateral (sagittal plane) views. These images are acquired for diagnosis and follow-up assessments to calculate parameters such as the curve type, the Cobb angle, and the kyphosis and lordosis angles [9], [10].

Although radiographic images provide the means to calculate the primary scoliosis parameters and indices, they consist in 2d projections of the bone structures. Thus, they do not provide a comprehensive assessment of the 3D deformity. Moreover, radiography acquisition entails a level of radiation exposure, which, if done frequently, is shown to be correlated with serious long-term health problems [9].

The existing literature on radiographic assessment for scoliosis diagnosis, evaluation, and follow-up emphasizes on minimizing X-ray exposure in children due to the potential risks. According to the SOSORT consensus, the initial evaluation of scoliosis should include X-rays in both projections and subsequent X-rays every 6–12 months are advised to limit the

exposure [1].

Surface Topography

Surface topography (ST), which refers to 3D acquisition of the back or the full trunk surface, provides a 3D representation of the trunk's deformity. The trunk appearance is a great concern of the AIS population since it affects their confidence and quality of life. Back surface topography assists the clinicians to conduct a more informed assessment of the back surface's appearance through measuring 3D back surface parameters. Surface topography's advantages such as non-invasiveness (no radiation exposure), low-cost, quick and accurate assessment of the body posture in three planes have made it an attractive evaluation method for clinicians [9].

Other Imaging Techniques

Apart from conventional radiography and surface topography some other imaging techniques for AIS assessment exist such as biplanar radiography, MRI, and CT. This section provides a brief review of these methods.

Biplanar stereo-radiography, also more commonly known as EOS imaging, is a low-dose radiographic method that uses an ultra-sensitive multi-wire proportional chamber detector [9]. In this system, anteroposterior and lateral 2D images are simultaneously acquired and used to reconstruct the spine in 3D and provide metrics in 3D to better quantify the deformity. EOS allows for an assessment in the standing position, revealing the natural shape of the spine. Despite its advantage in terms of radiation dose, the cost of EOS imaging system is high, which limits its availability in countries with lower income [9].

Another imaging method is CT which addresses the limitations of conventional radiography by providing accurate 3D reconstructions. Despite providing an accurate representation of the deformity, due to a high cost CT and radiation exposure, CT is not used for AIS assessment and is mainly used in surgical planning for adults with scoliosis [9].

MRI is another volumetric imaging method and is mostly used in cases where the patients suffer from pain and functional problems. It provides structural insights into the soft tissue of the trunk. Although it is radiation-free, it is very costly and thus not widely available [9].

In this section we covered the most common imaging techniques for AIS assessment. Overall, conventional radiography remains the main method for diagnosis and initial evaluation. However, this method fails to provide a comprehensive assessment of the 3D morphology and it involves radiation exposure. Surface topography is a radiation-free, low-cost method

that provides a precise representation of the back surface but lacks the bone structure level of information. Surface topography can be used as a complement to radiography in initial assessments and as an alternative in follow-up sessions. As for other imaging methods, they all have limitations and are not as widely used as these two methods.

2.2.3 Quantitative assessment

Scoliosis is a 3-dimensional deformity, and it is at best evaluated in three mutually perpendicular planes. The reference planes used for scoliosis evaluation are referred to as coronal, sagittal, and transverse planes [11] (see figure 2.3). Different assessment metrics have been proposed and practiced to evaluate the deformity in each plane. Assessment provides useful information for AIS classification which is used in treatment planning. AIS can be classified based on the severity of the curve (the Cobb angle), and the topography which refers to the anatomical level of the apex in the coronal plane [1]. Topographically speaking, there are four classes of scoliosis: thoracic, lumbar, thoraco-lumbar, and S-shaped (double major curve) [1].

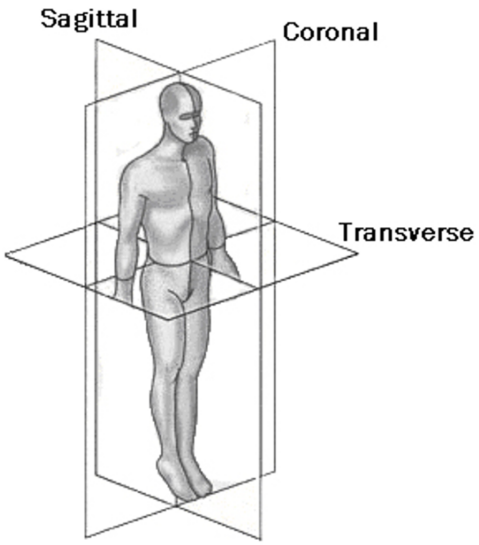


Figure 2.3 Reference anatomical planes for AIS evaluation[11]

Radiographic assessment

Radiography images of posterior–anterior (the coronal plane) and lateral (sagittal plane) views of the trunk are used to calculate parameters such as the curve type (the curve shape and the level of the apex), the Cobb angle (coronal plane), and the kyphosis and lordosis angles (sagittal plane).

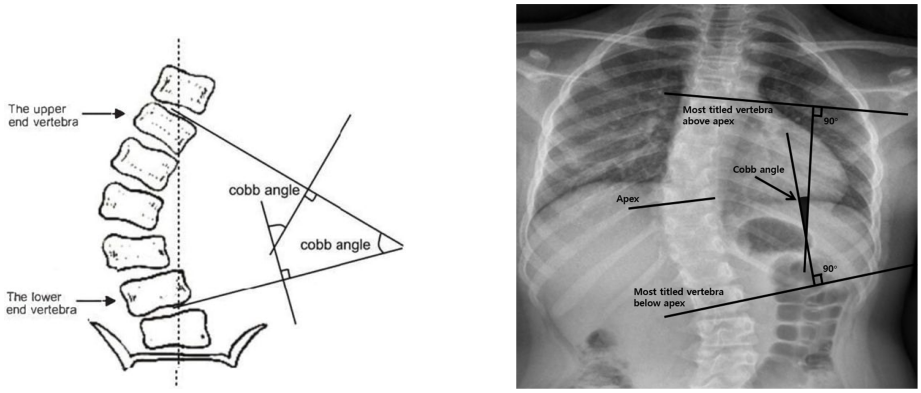
Cobb angle is the most popular AIS evaluation metric and is referred to as the gold standard. Cobb angle is the angle between the lines drawn along the upper-end plate of the highest most tilted vertebrae (the upper end of the curve) and the lower-end plate of the lowest most tilted vertebrae (the lower end of the curve). Figure 2.4 illustrates the Cobb angle calculation method. In scoliosis terminology, the most deviated vertebra is called the apex. The Cobb angle is measured using the radiographs of the coronal plane (see figure 2.3). A Cobb angle of at least 10° is required to confirm the diagnosis of AIS [1].

Lateral X-ray images are used to assess the deformity in the sagittal plane. The kyphosis and lordosis angles are used to quantify the natural spinal curvatures in the sagittal plane [10].

Topographic assessment

Assessment of the outer appearance of the trunk is possible through surface topography. Since surface topography modalities perceive the deformity through the outer 3D surface of the back, they do not provide any direct insight into the bone structure deformity [9]. This limits their usage in calculating parameters such as Cobb angle. However, studies suggest that there is a correlation between topography parameters and the radiography-based parameters [13], implying that significant changes in the cobb angle can be identified through surface topography [14]. Therefor, surface topography can be considered as an alternative to radiography in monitoring the progression of AIS and potentially decreasing the frequency of x-ray exposure. Other than an estimation of the Cobb angle, the surface parameters provided by surface topography provide a quantitative assessment of the rib hump through the axial surface rotation (ASR) [13], illustrated in Figure 2.5 (measured in the transverse plane). The rib hump is conventionally measured through clinical examination using a measuring tool called the scoliometer [1]. The ASR was proved to be correlated with vertebral axial rotation [5].

Surface topography has also been used in quantifying the trunk flexibility and reducibility. In [15], three static surface topography acquisitions of the patients in standing position and maximum right and left lateral bending were made. Using their developed methods to quantify trunk symmetry metrics and these acquisitions they were able to measure the reducibility of trunk asymmetry.



(a) [12] (b) [12]

Figure 2.4 Cobb angle calculation

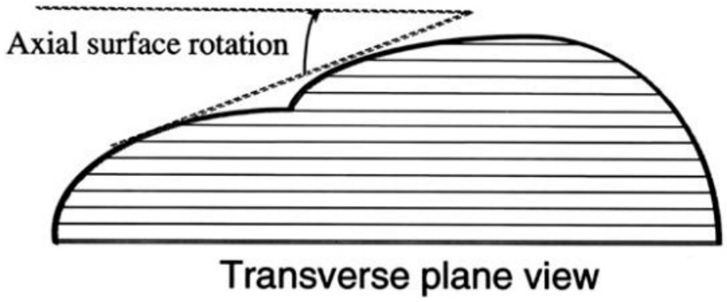


Figure 2.5 Axial rotation of the trunk surface[11]

2.2.4 Treatment

If neglected, AIS deformity can progress during growth and reduce patients' quality of life. AIS negatively affects patients' lives in different aspects including aesthetics, functional disabilities, pain and discomfort, and psychological well-being [1]. To prevent these problems, different methods have been proposed and practiced over the years to control curve progression and fix the AIS deformity.

AIS treatment methods fall into two groups: non-conservative such as surgery and conservative treatments such as bracing and physical therapy. The best treatment plan for each individual is devised according to the type and the severity of AIS. For curves with Cobb angles smaller than 25°, observation is recommended [1]. Patients will be clinically evaluated at regular interval and will have their radiographs taken to assess the progression of the curve. Observation based on surface topography is currently studied to limit patients exposure to X-rays. For more severe deformities, more active treatments are applied. Patients diagnosed

with moderate AIS (Cobb angle in the rage of 25°–40°) typically undergo non-surgical interventions, while those with severe AIS (Cobb angle >40°) resort to surgical treatment if non-surgical methods fail to limit the progression of the deformity [2]. In the following, we provide more detail regarding the physiotherapy and surgical treatment for AIS.

Physiotherapy Physiotherapy for scolioisis addresses two sets of goals: morphological and functional. From the morphological perspective, the goal is to improve the aesthetics via posture correction and to prevent the curve progression. From the functional perspective, the goal is preventing functional disabilities and pain [1].

Physiotherapeutic Scoliosis-Specific Exercises (PSSE) refer to all forms of physiotherapy exercises that are proven to be effective in achieving the AIS treatment objectives. PSSE can be used as an independent treatment or as a complementary treatment during bracing period and before or after surgical therapy for rehabilitation purposes [1]. There are multiple approaches to PSSE that are distinguished based on the exercises that they use [16]. These approaches include the Lyon approach [17], the Schroth approach [18], the Scientific Exercise Approach to Scoliosis (SEAS) [19], the Barcelona Scoliosis Physical Therapy School approach (BSPTS) [20], the Dobomed approach [21], the Side Shift approach [22], and the Functional Individual Therapy of Scoliosis approach (FITS) [23, 24].

PSSE should include:

- Corrective exercises in 3D
- Training in activities of daily living
- Stabilizing the corrected posture
- Patient education

Global postural re-education (GPR) is among PSSE treatments and aim to improve the posture and the back surface symmetry through passive and active corrective exercises [4]. Self-correction and manual correction are two respectively active and passive corrective exercises that apply a momentary posture correction in order to assess the curve reducibility and progressively improve the posture [4].

Manual correction involves a therapist applying targeted force to specific parts of the trunk to achieve a spinal correction. During the exercise, the therapist stands behind the patient and exerts force with their right hand just beneath the right thoracic apex until achieving a satisfactory spine correction. Simultaneously, their left hand is placed on the left side of the back to stabilize the pelvis and maintain the overall posture [4]. Figure 2.6 illustrates a surface

topography of manual correction. Manual correction allows the therapist to qualitatively assess the back muscle stiffness and curve reducibility [4].

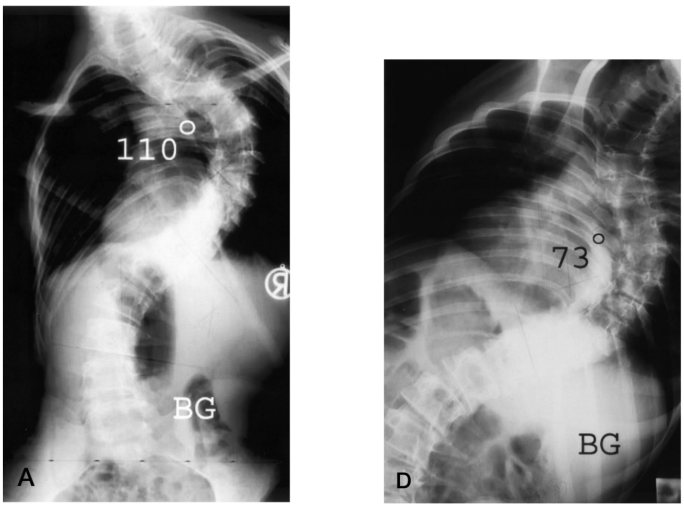
Figure 2.6 Surface topography of the manual correction (taken from [4])

Self-correction refers to active muscle recruitment by the patient to correct the posture in 3D. Self-correction allows for assessing the patients' ability to conduct the correction and progressive improvement of the posture [4]. Figure 2.8 illustrates X-ray images before and after self-correction.

Recent studies show promising results on the efficacy of PSSE in addressing AIS treatment goals and improving the scoliosis parameters such as the Cobb angle. However, further quantitative studies are required to better comprehend the biomechanics involved and to quantify and confirm their efficacy.

S. Depuis et al. in [4] conduct a clinical study with 17 AIS patients to quantitatively evaluate the PSSE corrective exercises. In this study, the curve reduction, the posture compensations, and the applied force (an indication of muscle stiffness) for manual correction and self-correction are reported. The results for self-correction show a significant reduction in the main spinal curve with slight posture compensations that varied for each patient. This observation suggests that self-correction exercise should be progressively integrated into the AIS treatment plan considering the patients' muscle control and ability to perform the correction with minimal posture compensations.

Studies combining PSSE with bracing have displayed positive outcomes, notably reducing surgery rates. Multiple studies have shown better results in terms of curve reduction after integrating PSSE in the bracing treatment comparing to bracing alone [1].



(a) (b)

Figure 2.7 Radiographs of the resting (a) and (b) maximal lateral bending position including the cobb angle at each position [3]. At the maximal bending position there is still a cobb angle of 73° which indicates a rigid curve.

Surgery

Surgical intervention is required for curves with Cobb angle larger than 40° [25]. There are different techniques for correcting the AIS deformity through surgery. Various techniques involve implanting rods on different sides of the curve and then using them to correct the curve and then fusing the realigned vertebrae into a solid segment [26, 27]. The objective is to correct the posture with minimum number of fused vertebrae to preserve the balance and mobility [27]. Locating the fusion area is very important and is determined based on curve flexibility and reducibility assessment in pre-operative analysis.

Different methods have been proposed and practiced for assessing the curve flexibility over the years. In [28], different static methods for pre-operative curve analysis such as supine lateral bending, supine traction, and fulcrum bending radiographs are compared. The results show a variability in curve flexibility analysis obtained by different methods. These methods involve using static images of movements such as lateral bending. Lateral bending on the convex side of the curve causes a curve (cobb angle) reduction, which is an indication of the curve flexibility. Higher reduction indicates a higher flexibility, whereas lower reduction indicates rigidity (see Figure 4.2) [3].

Now that clinical concepts were defined, the next sections present more technical aspects related to this work, starting from 3D data acquisition techniques and representations.

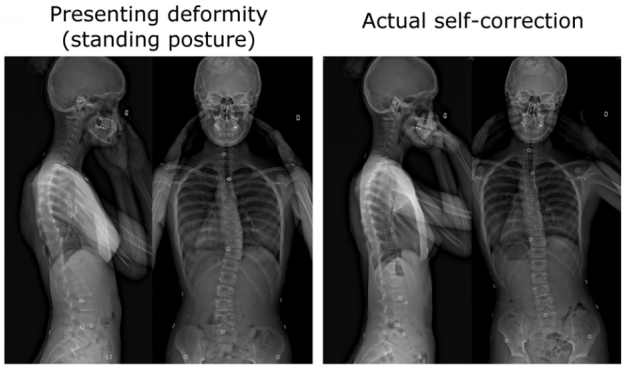


Figure 2.8 X-ray images in the resting position presenting the deformity and right after the self-correction [4]

2.3 3D Data Acquisition

In recent years, with technological advances in depth sensing technologies, 3D cameras have become more affordable and available. In this section, we provide an overview of different depth sensing technologies and 3D data representations.

2.3.1 3D Acquisition and Imaging Methods

Range imaging systems mostly consist of a color subsystem and a depth subsystem. The depth subsystem is either based on triangulation or time-of-flight method [29]. Triangulation-based approaches fall into passive or active category. In passive triangulation methods [29], referred to as stereo vision systems, two calibrated cameras with known parameters are installed in a certain distance called the baseline. These cameras are used to capture two images at the same scene. The corresponding points in the two images are then used to retrieve the actual position of the reference point through triangulating the lines passing through the image points and the corresponding camera center.

In active triangulation systems [29], one of the measurement cameras in the stereo vision systems is replaced by a projection device. In the simplest setting, a single beam of light is projected onto the scene, illuminating a point-like spot, and an image is captured. The projection line and the line between the illuminated point and its image are then used for triangulation and retrieving the depth of the point. If the intrinsic and extrinsic parameters of the camera and the projector are known, we can find the equations of the intersecting lines. Thus, the depth of the actual point can be retrieved. The projector then scans the whole scene and for each point the same process is repeated.

The problem with single point projection is that the number of required images for capturing the depth of the whole scene is equal to the number of pixels of the image, which is usually very large [29]. In order to solve this problem, other projection methods have been proposed that instead of projecting a single beam, project a light plane onto the seen, producing a vertical line of light. In this setting, the equation of the plane of light is known and is used for triangulation to retrieve the depth.

In the so-called sheet-of-light projection, the number of required images are reduced to the number of columns which can still be very large. In order to further reduce the number of images and scanning overhead, structured-light methods have been proposed. As the name suggests, these methods project light with a specific pattern (structure) and retrieve the 3D coordinates using the known pattern [29]. The challenge of these methods is figuring out the illumination source for each point. One common encoding strategy is the gray code method [29], where a series of light patterns including multiple light stripes are sequentially produced and an image is captured for each pattern. Each light stripe is specified with a binary number. For 2^N light stripes, each stripe is assigned an N-bit binary number. For example if there are 7 stripes, the 3rd stripe is assigned 011. Each projection pattern corresponds to a binary bit. In each pattern, the stripes with 1 at that bit are present. The index of the stripe at each image point is then retrieved based on the value of that stripe in each image. For example, in the case of 7 stripes there should be 3 bits and 3 projection patterns. In the first pattern, 100, 101, and 111 are illuminated because the first bit is 1. In the second pattern, 010, 011, and 111 are illuminated, and in the third pattern, 001, 011, and 111 are illuminated. Thus, if a pixel is illuminated in in the first and the third images, the retrieved index is 101 or 5 which is determines source projector of that stripe. The depth can then be inferred through triangulation.

Another category of structured-light methods are phase-shift methods [29]. These methods project light patterns with periodic intensity over the image width, producing vertical light fringes on the target surface. The light patterns are distinguished by having different phase shifts. The periodic light patterns with different phase shifts are projected sequentially and for each pattern an image is captured. Using the known equations of the light patterns (the frequency and phase shift), and the intensity of the points in the phase-shift patterns' images, a system of equations is derived and solved to find the light intensity value at that point. One of the challenges of this phase-shift pattern is finding the corresponding projector column (phase value) at this stage. This step is referred to as phase unwrapping. One solution to this problem is to use fiducial points blended in the light pattern. The Inspeck system, used for surface topography acquisitions in patients with AIS [30], is based on this phase-shift method.

Another solution for phase unwrapping is combining phase shift and Gray code together. The smaller stripes in Gray code patterns are replaced with phase shift patterns. The binary patterns are used to encode the projector information in the image pixels and facilitate the phase unwrapping.

The Dynamic Surface Topography (DST) system, recently acquired at Sainte Justine hospital, combines phase shift and binary light patterns. It is composed of a single digitization unit, consisting of a high-resolution dynamic projector, and a camera with a frame rate of 240fps. To acquire a surface in 3D, Nps phase shift patterns are projected for range measurements and Nb binary patterns are projected for precise phase unwrapping. The numbers Nps and Nb are configurable and allow for a trade-off between reconstruction rate and depth accuracy. With Nps =5 and Nb =7, we can acquire 2M 3D points on the back surface at 20fps with sub-millimeter depth uncertainty and a lateral resolution of 1mm, which, based on preliminary experiments, is sufficient for the targeted application and rate of shape deformations.

Existing range sensing systems, incorporating the above technologies, are capable of acquiring the coordinates of 3D surfaces along with texture information (color subsystem) with a high resolution and can be used in different applications that rely on 3D data. One of the challenges with 3D data is choosing the best representation. A number of most common 3D data representation methods are explained in the next section.

2.3.2 3D Data Representation

Depending on the target application, 3D data may be represented in different ways. The most common 3D representations are point clouds, depth maps, surface meshes, and voxels. In this section, overviews of these methods are provided.

A point cloud [29] is a set of acquired 3D coordinates of a scene. It is the simplest and most immediate form of 3D data. Since the data does not undergo any modification, point clouds preserve all the acquired 3D information. Point clouds are unstructured and there is no specific order to the points. There is no explicit information about the surface and the shape of the objects in the scene. Moreover, the sampling density varies in different parts of the point cloud, and the points are distributed unevenly. This unruliness of point clouds makes them challenging to work with. For example, conventional deep learning models that are designed for structured data (grids or sequences) are not compatible with point clouds.

Depth maps [29] are another form of 3D data representations and are stored as 2D arrays of depth values. The main advantage of depth maps over point clouds is their grid-like structure. Because of their 2D grid-based structure, depth maps are suitable representations

for conventional convolutional neural networks.

Another way to represent 3D data is by retrieving the surface of the objects. Surface meshes [29] are obtained by using computer graphics algorithm to reconstruct the surface of objects from the acquired 3D points. Surface representations are useful in geometrical algorithms in computer vision and graphics. Comparing to point clouds, surface meshes contain more information about the connectivity and the shape geometry of the objects. However, constructing meshes has higher computational demand.

3D data can be represented by 3D volumetric grids made of smaller units called voxels. Voxels [29] are 3D alternatives of pixels. Each voxel represents a small cubic unit in space. Like pixels, voxels keep Boolean values indicating their presence or absence. Again, similar to pixels, voxels can also hold values representing features of the points such as color. Voxels representations are useful in applications that rely on volumetric features. The problem with voxels however, is that the storage requirement is very high and grows cubically with the resolution.

Among the 3D data representations, point clouds best preserve the 3D geometrical information of the scene since they do not alter the raw acquired data. In addition to that, due to their low memory requirement and preprocessing overhead, point clouds are preferred to other representations of 3D data [31].

2.4 Deep Learning on 3D Point Clouds

Since the world around us is 3-dimensional, it is best described through 3D acquisition rather than 2D images. These days, with the recent advances in depth sensing technologies and 3D data representation, 3D data acquisition is becoming increasingly more available and affordable [31]. Consequently, 3D deep learning models are attracting more attention and are getting integrated in different applications. According to the input data format, existing 3D deep learning models can be divided into three categories: multi-view based methods, volumetric-based methods, and point-based methods [31]. In this section, we review most popular and successful 3D deep learning models, focusing on point-wise architectures.

2.4.1 PointNet and PointNet++

Due to specific properties of point clouds, such as lacking structure and non-uniform spatial distribution, conventional deep learning models designed for structured data representations are not suitable for them [31]. Therefor, a new category of models referred to as point-based MLP methods, pioneered by PointNet [32], have been proposed that are designed specifically

for point clouds. These models consider the properties of point clouds in their architecture by relying on no structure in the input data.

PointNet [32] is a deep learning model for sets, which are unordered collections of data such as point clouds. As a set function approximation, PointNet network is invariant to permutation in the input data. PointNet directly takes the point cloud as input and extracts a set of per-point features through several layers of Multi-Layer Perceptrons (MLPs). The per-point features then go through a symmetric function (max pooling), where the most informative points that best describe the input point cloud are selected to produce a set of global features (see figure 2.9). The key architectural element of PointNet, which makes it suitable for point clouds, is the use of this symmetric aggregating function. Since it's symmetric, it's invariant to the order of the input points by definition. The extracted global feature can be used for segmentation or classification. For classification, the global features are fed to a layer of MLPs to compute class labels. For segmentation, the learned global features are concatenated with each per-point features and then fed to MLPs to compute a new set of per-point features that is aware of both local geometry (per-point features) and global semantics.

Another feature of PointNet is invariance to rigid transformations. To achieve this, PointNet uses a joint alignment network (T-Net) to predict an affine transformation and applies it to the input points and the extracted per-point features at deeper layers (see figure 2.9).

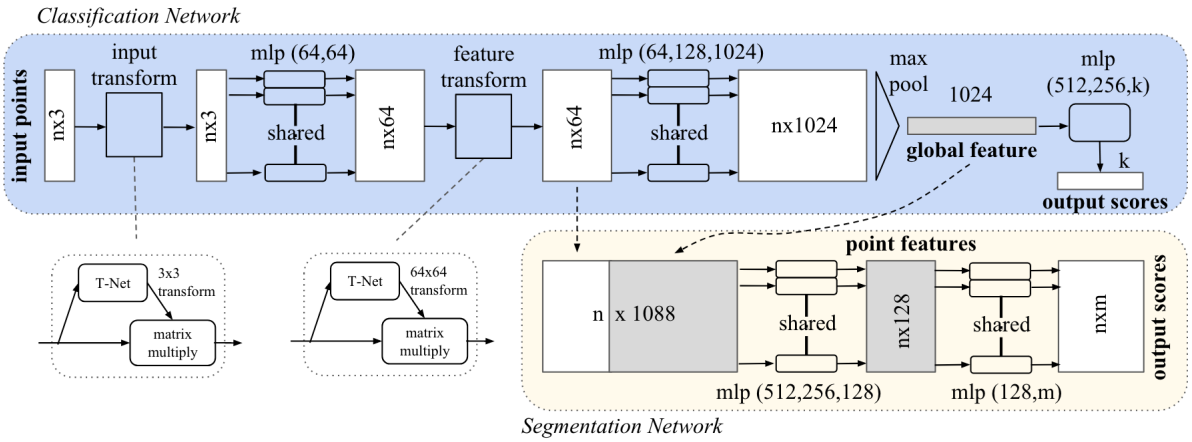


Figure 2.9 Architecture of PointNet [32]

PointNet introduced a deep learning model for set functions, incorporating brilliant design ideas, and succeeded to outperform its predecessors on various 3D benchmarks. However, PointNet network does not incorporate the local geometrical structures encoded by the relative distance of the points in a metric space. Point clouds contain meaningful local neighborhoods based on the points' geometrical proximity. By ignoring this valuable information,

PointNet fails to capture finer details and structures in the input point clouds.

Addressing this limitation, Qi et al. proposed a hierarchical feature learning approach based on PointNet [33]. The model, referred to as PointNet++, is constructed from several set-abstraction levels. Each set-abstraction level consists of a sampling layer, a grouping layer, and a PointNet layer. The sampling layer selects a sample of points using farthest point sampling (FPS) algorithms. The Grouping layer divides the points into overlapping local regions of a fixed radius around the sampled points. The PointNet module is then applied to points in each region to extract local features of the neighborhoods. Several set-abstraction layers (two-three layers in the original paper) are applied hierarchically, extracting increasingly higher-level features (see figure 2.10).

For classification, the feature vectors of the last layer are fed to another PointNet layer for extracting a final global features vector. For segmentation, the feature vectors go through a feature propagation layers, analogous to the inverse of set abstraction layers, that are based on point interpolation. At each layer, the local features are propagated to their closest points and are concatenated with features from the corresponding abstraction layer transferred by a residual connection. This process repeats until the per-point features for all the original points are obtained. Lastly, a unit PointNet (one MLP) is applied to compute the segmentation scores.

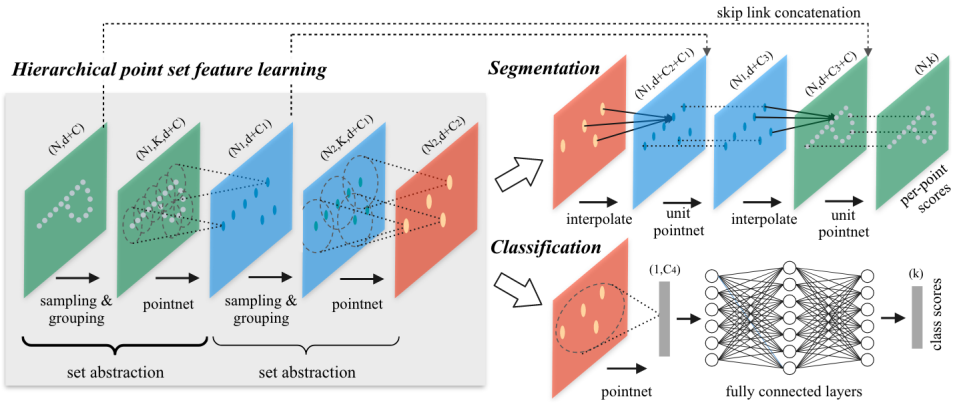


Figure 2.10 Architecture of PointNet++ [33]

By applying PointNet in a hierarchical setting and deploying the built-in structure of points in a metric space, PointNet++ outperformed PointNet on 3D benchmarks and achieved state-of-the art at the time of its release.

2.4.2 Other 3D Models and PointNeXt

Since the release of PointNet and PointNet++, many other 3D deep learning approaches have been proposed, some surpassing those models on 3D benchmarks by significant margins.

On of the recent successful models is the point transformer introduced in [34]. In this work, the authors propose using a transformer-inspired model to learn the features of a 3D point cloud. Since the self-attention mechanism does not assume any order in the input sequence, it works effectively on point cloud data.

In [35], the authors claim that the success of the modern models largely depends on modern training techniques and the size of the model rather than architectural changes. Regarding this claim, this paper sets to explore the full potential of PointNet++ by applying new training strategies and model scaling. Through comprehensive additive experiments with multiple training strategies, including data augmentation and optimization techniques, it is shown that modifying the training strategies alone can boost the performance of PointNet++ on various benchmarks. The newly trained PointNet++ improves upon PointNet++ ScanObjectNN [36] classification benchmark by 8.2% overall accuracy (OA), outperforming the state-of-the-art models. For S3DIS [37] segmentation benchmark, the mIoU (mean intersection of union) is increased by 13.6% comparing to original PointNet++, surpassing many newly-introduced models. This is while PointNet++ has significantly fewer parameters comparing to other 3D models. Other than new training strategies, a new scaled version of PointNet++ is proposed, referred to as PointNext to further boost the performance of PointNet++. This paper, once again, proves the strength of PointNet++ architecture and its superiority over other 3D models.

2.4.3 Scene Flow Estimation

Many applications in today's world rely on understanding the 3D motion of the surrounding environment [38]. The 3D motion field of points in a scene is referred to as scene flow, and the task of finding scene flow is called scene flow estimation. Like many other tasks in computer vision, deep learning models have been proposed for scene flow estimation. The next section presents one of the prominent existing scene flow estimation models.

FlowNet3D

FlowNet3D [38] is a 3D deep learning model that uses the feature extraction layers of PointNet++ [33] for estimating the scene flow between two consecutive point clouds. The proposed model is composed of three types of layers (see figure 2.11). First the two input point clouds

go through a set-conv layer which is the same as the set-conv layer in PointNet++. Just like PointNet++, the input to each set-conv layer is a point cloud with n points, including coordinates and features for each point, and the output is new feature vectors for $n^{'} < n$ sampled points. Next the learned features of input point clouds go through a flow embedding layer, where the network learns to find a set of corresponding candidates for each point based on feature and spatial similarities and makes a weighted decision to predict the flow embedding for the points in the first frame. The last layer is a set-upconv layer that propagates the learned flow embedding for the sampled points to all the points in the original point cloud by aggregating it's neighboring points flow embedding. The set upconv layer acts as a set conv layer but instead of using farthest point sampling to select the samples, the samples are the target points that we want to propagate the features to. The final output of scene flow pipeline is a set of 3D motion vectors for all the points in the input source point cloud (see figure 2.12).

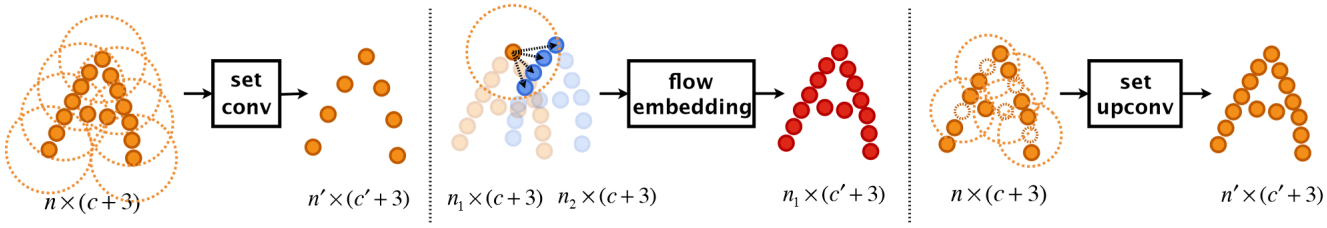


Figure 2.11 Different layers in FlowNet3D [38]

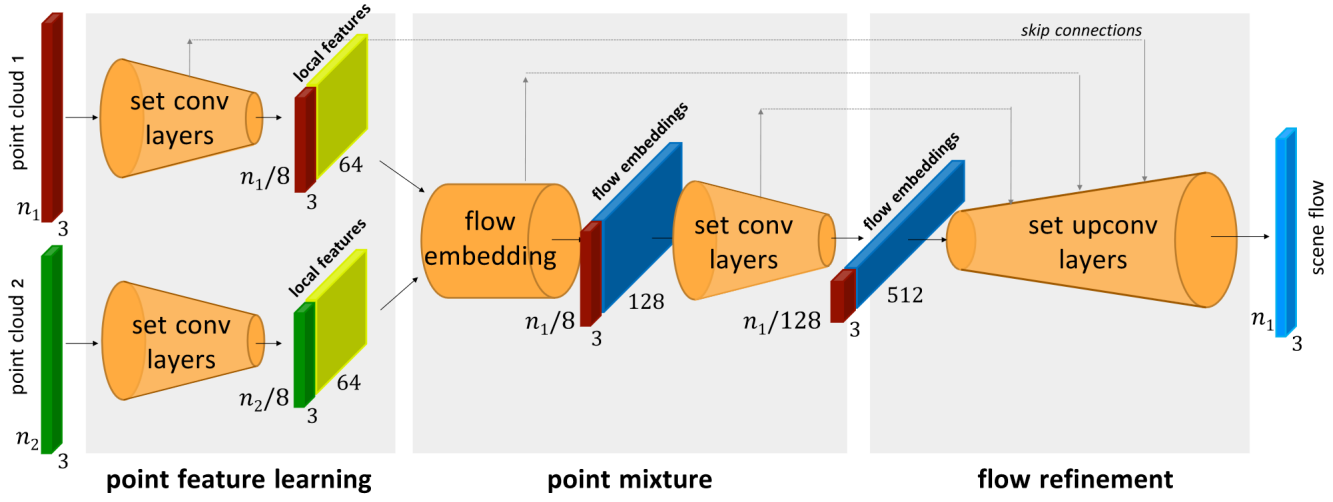


Figure 2.12 Pipeline of FlowNet3D [38]

Now that deep learning models for 3D points clouds processing have been introduced, the next two sections provide an overview of application-specific previous work related to anatomical landmark localization and point cloud segmentation.

2.5 Anatomical Landmark Localization

Over the years, many anatomical landmark detection methods have been proposed for different parts of the body. In this section, we take a look at a number of successful 3D landmark detection methods in point clouds.

In [39], the authors propose a method for knee landmark detection in knee joint point clouds based on PointNet++ feature extraction layers. They use the pretrained feature extraction part of the classification PointNet++ network and add 3 fully-connected layers that output the predicted coordinates of three anatomical landmarks on the knee. They've trained the new layers on their collected knee dataset along with fine-tuning the pretrained layers. They provide results for several experiments with different model configurations (where the number of set abstraction layers vary) and compare the results to the alternative approaches and state of the art. The provided evidence show that their method has a higher accuracy (evaluated by mean square error) comparing to other similar methods.

In [40], the authors use a regression CNN network to predict the approximate locations of anatomical landmarks and then they apply a classification CNN find the best local image patch around the predicted location. They use pretrained deep CNNs for both tasks and train them on CEASEAR dataset which consists of 3D full-body laser scans manually annotated with 73 landmarks. They select 8 landmarks located on the left side of the body visible from front. For each body scan, they project the 3D scans onto 2D planes to create a 2D gray-scale image, extract the depth map, and compute 2D projected curvature maps with the curvature value as the pixel value. They use modified VGG networks for the regression and the classification tasks, altering the last layers. They consider the four corner points of a small square around each predicted landmark as phantom landmarks. Square patches with a predetermined size centered at each of the five landmarks (the real one and the phantoms) are used to train the classification network. The patch around the real landmark labeled as positive and the phantom patches labeled as negative. A classification network is trained for each landmark. At inference, a region with predetermined area around the predicted location is searched through a sweeping window and the patch with the highest classification score is considered as the landmark patch with the landmark at its center. The provided results show that training the regression network with curvature maps, and the classification network with depth maps result in lower error and outperforms state-of-the-art on most landmarks.

In [41], the authors propose a method for back surface automatic landmark detection in point clouds acquired with a time-of-flight camera in D4S setting. The method is a combination of learning based clustering and regression and using some known auxiliary points to determine the location of the landmarks. The acquisition setting allows dynamic acquisition of movements. The preprocessing and the detection time is reported to be very low making the method applicable in real-time with 3 fps sampling frequency. The results also show a very high accuracy in landmark detection, comparable with the manual marking of experts.

These methods try different approaches to localize the specific anatomical landmarks and achieve acceptable accuracy. However, the task of 3D landmark detection remains a challenging task due to variability of landmarks in different applications and low error tolerance.

CHAPTER 3 RATIONALS AND OBJECTIVES

Adolescent idiopathic scoliosis (AIS) refers to a 3D deformity of the trunk with no known underlying cause in adolescents, which if left untreated, can progress during growth and reduce quality of life. According to the severity and the curve type, different treatment methods are advised for AIS such as surgery, wearing braces, and physical therapy.

Physiotherapeutic Scoliosis-Specific Exercises (PSSE) include corrective exercises such as self-correction to gradually correct the posture and reduce the curve. Other than that, PSSE allow the clinicians to evaluate the curve reducibility and patients' motor control over the back muscles. Due to its efficacy and minimal health risks, PSSE is increasingly drawing attention and becoming more integrated into the AIS treatment process. Existing experimental studies of PSSE show promising results but highlight the need for stronger quantitative evaluation methods. Over the years, a number of methods have been proposed and practiced for PSSE evaluation. However, all these methods are based on static acquisitions, static momentary frames of exercises. To our knowledge, there is no study in the literature that does dynamic evaluation of the exercises, despite their inherently dynamic nature.

Other than PSSE, some specific movements such as lateral bending are used to assess the curve flexibility as part of the pre-operative analysis in surgical treatments. This movement is also evaluated statically in all existing studies and in practice.

The lack of studies on dynamic evaluation of AIS can be attributed to the absence of effective methods and tools for dynamic evaluation.

These days, 3D acquisition technologies are more accessible and affordable. In fact, our affiliation with the Sainte-Justine hospital research center provides access to a dynamic surface topography system, enabling non-invasive 3D dynamic acquisitions with a high point sampling and temporal resolution.

Moreover, in recent years, 3D deep learning models have shown remarkable performance in various tasks for 3D point clouds. Recognizing the potential of 3D deep learning models, in this project, we aim to leverage these technologies and the 3D imaging system at our disposal to fill the gap in quantitative dynamic evaluations of trunk motion in AIS movements.

The research hypothesis behind this work are the following:

- H1: Trunk motion can be evaluated by tracking the 3D motion of anatomical landmarks or segments on the back surface throughout a motion sequence.

- H2: Fine-tuning PointNet++ on AIS data can provide landmark detection and back surface segmentation on markerless surface topographies with an error in the range of inter-observer variability.

To test the above hypothesis, our objectives are threefold:

- First, to collect, preprocess, and annotate a dataset of 3D dynamic acquisitions of PSSE exercises using the dynamic surface topography system at Sainte-Justine hospital.
- Second, to develop and validate deep-learning-based tools to automatically localize the anatomical landmarks and anatomical segments on the back surface.
- Third, to evaluate the clinical potential of the proposed tool to quantify trunk motion in lateral bendings and self-correction exercises.

Through these objectives, we take a step toward automating dynamic evaluation of AIS, ultimately aiming to refine the treatment of AIS and quantifying the trunk motion in 3D.

CHAPTER 4 METHODOLOGY

According to the objectives, the methodology of this project falls into three steps. The first step is collecting, preprocessing, and annotating a dataset of AIS 3D dynamic sequences. The next one is developing a tool for automatic detection of anatomical landmarks and segments on the back surface. The last part is calculating the trunk surface quantification metrics over time for a given sequence. In this chapter, each of these steps are described in separate sections.

4.1 Data Collection

As the first step of this work, we used the DST system, located at Sainte Justine hospital, for data collection during 2022 and 2023 summers. In this section, first the acquisition protocol, and then the data preparation and annotation steps are explained.

4.1.1 Acquisition Protocol

We obtained ethical approval from Sainte Justine hospital and Polytechnique Montreal for the recruitment of patients with AIS at Sainte Justine. The eligibility criteria are as follows:

- 10-18 years old
- Adolescent Idiopathic Scoliosis
- Main Thoracic curvature, Cobb angle between 15° and 45°

In each session, 5-10 markers (5mm round stickers) were placed on the participant's back surface, marking the anatomical landmarks of interest, including C_7, T_{up}, T_{apex}, T_{down}, Sc_{left}, Sc_{right}, I_{left}, I_{right}, L_{apex}, L_{down} (see figure 4.1). These specific landmarks were selected since they are used in trunk surface metrics quantification. After attaching the markers, the participants were asked to stand with their backs facing the DST system and perform a set of exercises. The acquisition protocol in summer 2023 was modified comparing to summer 2022. In summer 2023, we increased the number of landmarks to 10 (see figure 4.1). Moreover, a well-trained Physiotherapy PhD student helped us with markers placement and explaining the movements to the participants. The last modification in summer 2023 was adding static acquisitions of three versions of manual correction applied by Robert. The movements include (In the order of acquisitions):

- Free lateral bending: To the left and to the right (see figure 4.2)
- Minimum manual correction: Reducing the curve with minimal postural compensation.
- Maximum manual correction: Maximal reduction of the curve without considering the postural compensations.
- Manual correction: Manual correction of the spine.
- Self-correction: The participants were instructed to perform the self-correction exercise and trained in front of a mirror before doing the acquisition.

Each movement was performed and recorded at least two times. The manual corrections were performed by a physiotherapy PhD student. Right after the correction, the participants were asked to remain still, and a static acquisition was made.

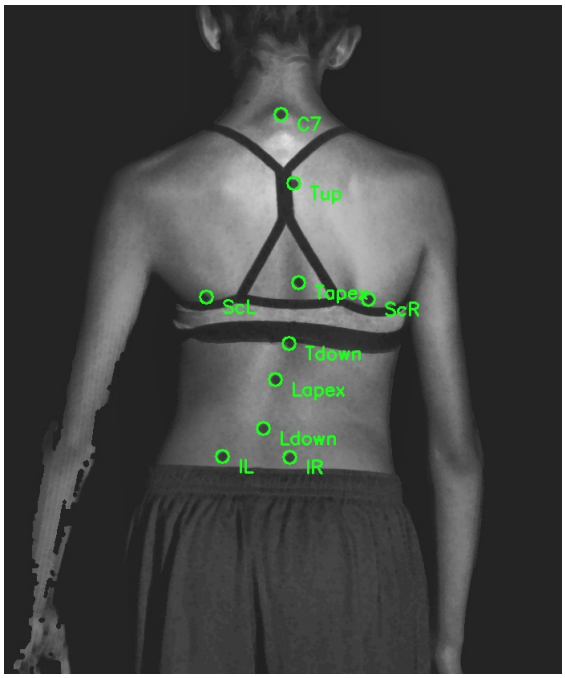
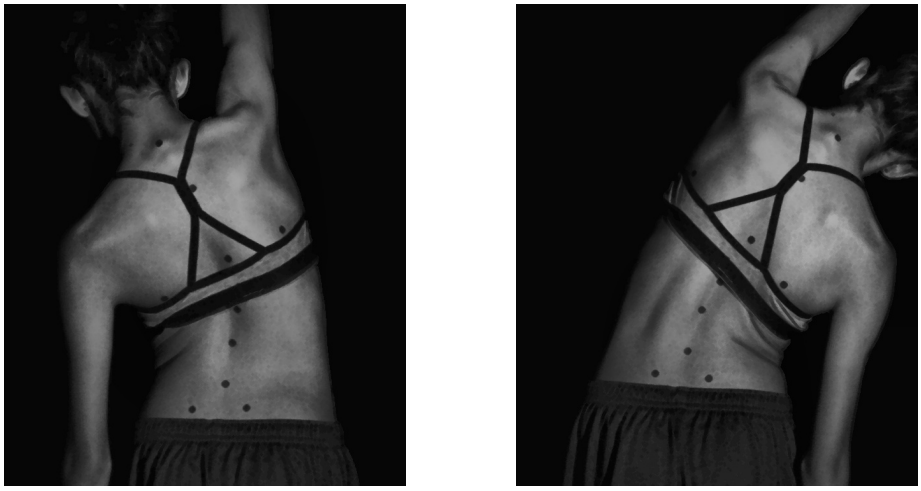


Figure 4.1 The targeted landmarks and the attached markers illustrated on participant 23, recruited in summer 2023

The recordings were limited to 5 seconds. For each frame, the 3D coordinates and gray-scale intensity values were acquired. The coordinates and the intensity values were stored separately in raw files. In the next section, the pre-processing and annotation steps are explained.



(a) Bending to the left (b) Bending to the right

Figure 4.2 Preprocessed intensity frames of the free bending movement performed by participant 23.

4.1.2 Pre-processing and Annotation

Each sequence of the dataset includes raw intensity and 3D coordinates files for all the frames. We implemented a preprocessing module to prepare the raw frames for annotation and training. The preprocessed frames are then used for annotation, finding the 3D coordinates of manually placed markers. The preprocessing module includes:

- Loading the raw intensity files, removing the header, and converting them to gray-scale images with dimensions 1176×1936.
- Loading the raw coordinates files, removing the header, and converting to an array with dimensions $1176 \times 1936 \times 3$ array
- Removing the background points using the depth information 4.3. We find the median of z values to find the estimation of the body's depth and use that value to define a threshold. All the points that are located deeper than the threshold are marked as background (assigned 0 for z coordinate).
- Automatic cropping. We use the non-background points to define boundaries and crop the intensity images based on them. The purpose of this step is to only keep the parts of the back that contain the markers to avoid unnecessary effort in the next steps.
- Lastly, we apply Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve the contrast of the intensity images, including the contrast between the markers

and the skin. To make the markers even more detectable, we use openCV preimplemented erosion and dilation function.

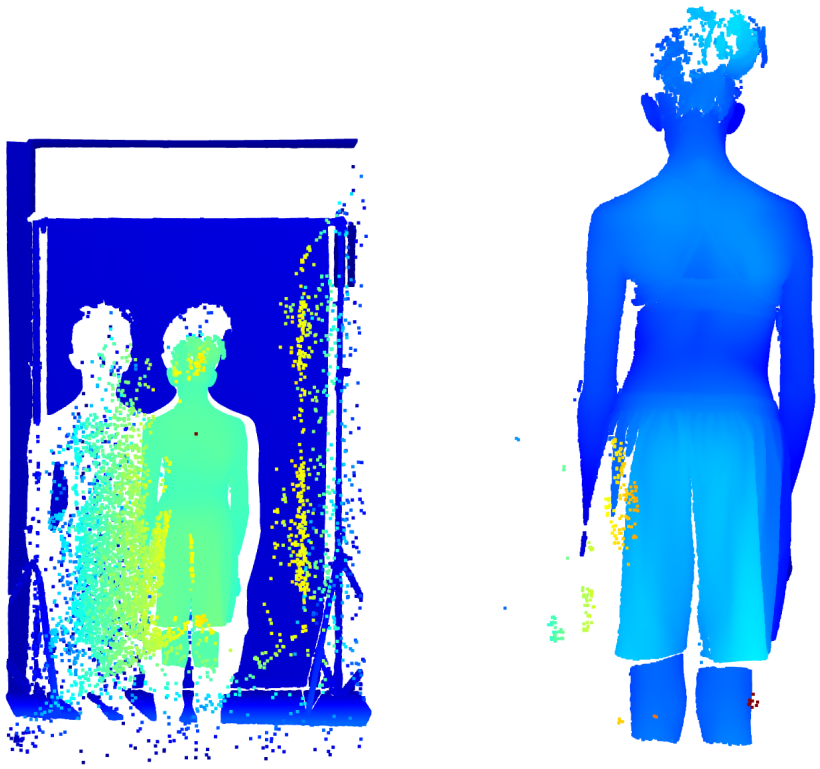
The preprocessed intensity frames are then passed to the marker detection function. The function uses the OpenCV pre-implemented blob detection method to find the 2D coordinates of the markers. Due to the texture of the skin and the presence of clothing, the marker detection algorithm is not perfectly accurate. There are some non-marker spots detected as markers while some of the markers remain undetected. With the help of Lea, our intern during the winter and the summer of 2023, we designed and implemented a graphical user interface for visualizing and manually correcting the results of the marker detection algorithm. With the help of this application, we modified and validated the 2D coordinates obtained by marker detection function. Figure 4.4 illustrates the stages of the annotation process. The corresponding depths (3rd coordinate) of the markers were then retrieved from the point clouds.

4.2 Automatic Anatomical Landmark Detection and Back Surface Segmentation

4.2.1 Back Surface Segmentation

For anatomical back surface segmentation, we used the PointNet++ part segmentation model, pretrained on ShapeNet part dataset [42], only modifying the last layer to match the number of segments which is 9 (8 landmarks + 1 outlier). For each marked landmark in the dataset, including C_7, T_{up}, T_{apex}, T_{down}, Sc_{left}, Sc_{right}, I_{left}, I_{right}, L_{apex}, L_{down} (see figure 4.1), we define a segment including its neighboring points. In order to do that, an area with a pre-determined radius (r) around each point is searched to find the closest landmark. That point is then assigned a segmentation label corresponding to that landmark. If no landmark is found close enough to a point, that point will be labeled as outlier. Collectively, we have 9 segments consisting of 8 segments around anatomical landmarks and 1 outlier segment. Using this annotation system we fine-tuned the model on the collected AIS dataset for the task of back surface segmentation. The input to the model is the point cloud along with intensity values of the back surface, and the output is the segmentation labels for all the points.

The loss function for training was negative log likelihood. We used AdamW optimization function a multi-step learning rate scheduler. Due to the requirements of PointNet++ model, we do a uniform down-sampling on the input point cloud. The model was trained for 50 epochs with initial learning rate 0.001. The segmentation results are presented in 5.



(a) Original (b) Removed background

Figure 4.3 Point clouds of participant 23. The one on the left is the initial point cloud and the one on the left after back ground removal.

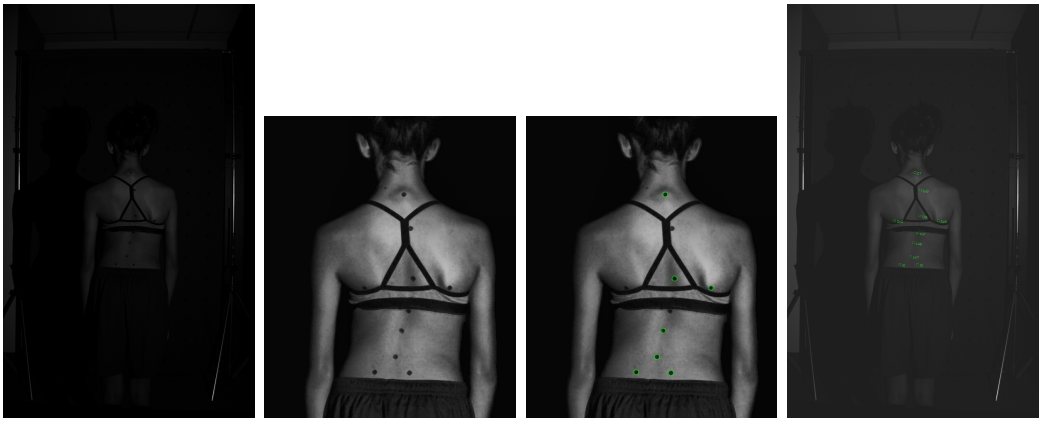


Figure 4.4 The intensity frame throughout the preprocessing and annotation pipeline. From left to right: the original image, the preprocessed image, the detected markers, manually corrected and labeled markers mapped to the original image.

4.2.2 Back Surface Landmark Detection

Regression

For the task of landmark detection, our first approach was training a regression model to estimate the 3D coordinates of the back surface landmarks. We used the pre-trained feature extraction layers of PointNet++ classification model (pre-trained on ModelNet40 [43]) and added 3 fully connected layers for regression. We trained the regression layers and fine-tuned the feature extraction layers on AIS training dataset.

The loss function used for training was the Mean Square Error (MSE). The 3D coordinates of the markers, which are provided in the annotated dataset, are used as ground truth for computing the MSE. The optimization function was AdamW with a multi-step learning rate scheduler. Due to the requirements of PointNet++ model, we do a uniform down-sampling to keep 10000 and 20000 points (we train for 10000 and 20000 separately). The input to the model is the preprocessed point cloud (removed background), also present in the annotated dataset, and the output is the predicted 3D coordinates of 8 anatomical landmarks on the back surface. We trained for 50 epochs with initial learning rate 0.001. The results are presented in 5.

Segments' Centers

The other approach that we tried for landmark detection was using the fine-tuned back surface segmentation model to find the back surface segments. After finding the segments, we calculate the the average coordinates of the points in each segment to find the center of that segment. Since the segments are defined to be centered around the landmarks, the center of each predicted segment can be considered an estimation of the anatomical landmark corresponding to that segment.

4.3 Tracking

Another approach that we tried was applying the FlowNet3D model to track the landmarks throughout the sequence. For that purpose, we used our back surface landmark detection model to find the landmarks in the first frame of the sequence. Then, used the FlowNet3D model to find the scene flow between the first and the second point cloud, and calculated the coordinates of the landmarks in the next point cloud by adding the estimated flow to the landmarks coordinates in the first sequence. This process was repeated for the second point cloud with the estimated landmarks and the third point cloud. By repeating this process for

all the consequent point clouds in a movement sequence, we tracked the estimated positions of the landmarks throughout the sequence.

4.4 Dynamic Quantification of Back Surface Metrics

Addressing the last objective of this project, we implemented functions to quantify trunk surface metrics over time for lateral bending and self-correction movements. The purpose of this part was to explore the potential of the implemented tools in dynamic quantification of the trunk surface metrics over the duration of a movement sequence. Using the developed landmark detection model, we can find the 3D coordinates of the back surface landmarks. These estimated coordinates can be used to quantify the trunk surface metrics at each frame of a sequence and monitor their changes over time.

4.4.1 Lateral Bending

For the lateral bending sequence, we calculate the Cobb angle for all the frames of the sequence. It is notable that Cobb angle is originally based on the radiography parameters. However, we use surface landmarks corresponding to the radiography parameters and calculate an estimation of the Cobb angle [1], [13]. The calculated Cobb angle is different with the reported X-ray-based cobb angle. Since we only care about the changes in the Cobb angle throughout a sequence, the translation (from radiography to surface parameters) error can be neglected as long as the relative measurements remain the same.

For calculating the cobb angle, we use the coordinates of the landmarks T_{up}, the upper-end vertebrae of the curve, T_{apex}, the most-deviated vertebrae, and T_{down}, the upper-end vertebrae of the curve. Below, you can find the formula for cobb angle calculation.

$$\text{Cobb angle} = 180^\circ - \theta$$

$$\theta = \begin{cases} 360 + atan_2(det, dot), & \text{if } atan_2(det, dot) < 0. \\ atan_2(det, dot), & \text{otherwise.} \end{cases}$$

$$dot = (T_{up,x} - T_{apex,x}).(T_{down.x} - T_{apex,x}) + (T_{up,y} - T_{apex,y}).(T_{down,y} - T_{apex,y})$$

$$det = (T_{up,x} - T_{apex,x}).(T_{down,y} - T_{apex,y}) - (T_{up,y} - T_{apex,y}).(T_{down,x} - T_{apex,x})$$

4.4.2 Self-Correction

For the self-correction sequences, in addition to Cobb angle, we calculate four other metrics. The first two metrics were the angles between the right and left scapula in the coronal and the transverse planes. For these metrics, we used the coordinates of the right and left scapula landmarks, Sc_{right} and Sc_{left}. The below formulas show the calculation steps for those metrics.

$$\text{Coronal scapulae angle} = tan^{-1}(\frac{Sc_{right,y} - Sc_{left,y}}{Sc_{right,x} - Sc_{left,x}})$$
$$\text{Transverse scapulae angle} = tan^{-1}(\frac{Sc_{right,z} - Sc_{left,z}}{Sc_{right,x} - Sc_{left,x}})$$

Another metric, referred to as trunk imbalance [44], indicates the horizontal distance between the middle of the pelvis landmarks, I_{right} and I_{left}, and the the C_7 landmark, in the frontal plane.

$$\text{Trunk imbalance} = \frac{I_{right,x} + I_{left,x}}{2} - C_{7,x}$$

The fifth metric, referred to as scapulae asymmetry (D), is the difference between the distance of right scapula (Sc_{right}) and the left scapula (Sc_{left}) from the spine. The calculation steps are shown below.

$$\Delta D = |d_{right} - d_{left}|$$
$$d_{right} = |(aSc_{right,y} + b) - Sc_{right,x}|$$
$$d_{left} = |(aSc_{left,y} + b) - Sc_{left,x}|$$
$$a = \frac{T_{down,x} - C_{7,x}}{T_{down,y} - C_{7,y}}$$
$$b = C_{7,x} - aC_{7,y}$$

Figure 4.5 illustrates the metrics on an intensity image of participant 23. Since the figure

captures the coronal plane view, the transverse angle is not included.

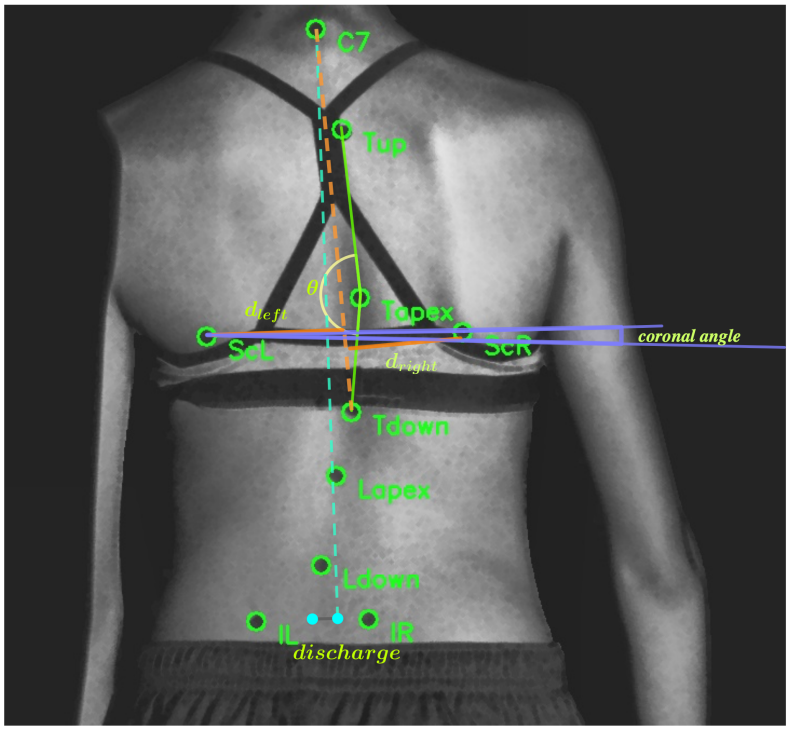


Figure 4.5 AIS metrics used for clinical evaluation, illustrated on participant 23.

4.5 Evaluation Metrics

In this section, the evaluation metrics that were used for different parts of the methodology are presented. For back surface segmentation, we calculate the accuracy (ACC) and the mean Intersection of Union (mIoU). The accuracy is computed by finding the percentage of correctly labeled landmarks over the total number of landmarks. To calculate the mIoU metric, for each segment we calculate the IoU of the ground truth points and the predicted points. IoU is the number of shared points of the ground truth and prediction (intersection) over the number of all the points in the ground truth and prediction (union). Then, we find the average of IoU over all the segments to find the mIoU.

For landmark detection, the evaluation metric is the Euclidean distance between the predicted coordinates and the ground truth coordinates (markers).

For evaluating the dynamic trunk surface metrics quantification, we calculate the Pearson correlation coefficient between the calculated metric based on automatic landmarks and calculated metrics based on markers over the duration of a sequence. The formula of Pearson correlation coefficient can be written as

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

where r is the correlation coefficient, x_is are values of the x-variable in a sample, $\bar{x}$ is the mean of the values of the x-variable, y_is are the values of the y-variable in a sample, and $\bar{y}$ is the mean of the values of the y-variable. For each metric, x and y correspond to the calculated values for that metric based on predicted landmarks and the values based on ground truth (markers). The correlation is proportional to r. Higher values of r indicate higher correlation between the prediction and the ground truth.

CHAPTER 5 RESULTS AND DISCUSSION

This section includes quantitative results, visualizations, and discussion. First, the results of automatic landmark detection and back surface segmentation are presented. Next, in the tracking section, visualizations of landmarks tracking are presented. Lastly, the results of dynamic trunk metrics quantification are presented. The last part is a discussion reviewing the results and exploring the potential clinical applications of the project.

5.1 Data Collection, Preparation, and Annotation

We recruited 26 patients during the summers of 2022 and 2023 at Sainte Justine Hospital. During acquisitions of 2 of the recruited 26 participants, 4 and 17, the lighting of the DST system was not set-up properly, resulting in problematic 3D coordinates. Those acquisitions were not used in this work. Overall, the proposed dataset includes 3D dynamic sequences of 24 patients with AIS.

The number of marked landmarks per participants varies: the first 12 participants in summer 2022 had 5-8 markers and the participants from 2023 have 10 markers. This is due to the fact that the acquisition protocol was modified for summer 2023. In some cases, some of the markers are not visible in intensity images due to clothing or lighting (especially the ones on the pelvis). Table 5.1 includes the number of acquired frames for each sequence and average number of visible landmarks for each participant. The annotated sequences in the collected dataset consists of the point clouds, the intensity files, and the 3D coordinates of anatomical landmarks on the back for 6 dynamic acquisitions of 3 different movements and (summer 2023 acquisitions) 6 static acquisitions of the final state of 3 different manual corrections for 26 participants. The duration of each dynamic acquisition is 5 seconds and the number of frames is around 80 (differs slightly for each sequence). The dataset was used to train, validate, and test our deep learning models for automatic landmark detection and segmentation. Table 5.3 shows the dataset splitting for training, validation, and testing. For training, we only use lateral bending sequences since they include a more diverse set of frames (from straight to bending), and for testing, we use self-correction sequences too. Participants 1, 7, 11, 14, and 16 do not have enough visible markers (missing the principle landmarks) and were not used for training and testing. The marked landmarks for participant 15 were not compatible with the rest of the dataset and therefor was excluded from training and testing. Participants 18 and 21 could not follow the movements properly and were not used for training or testing either. Some of the newly added sequences are not annotated yet (bending sequences for

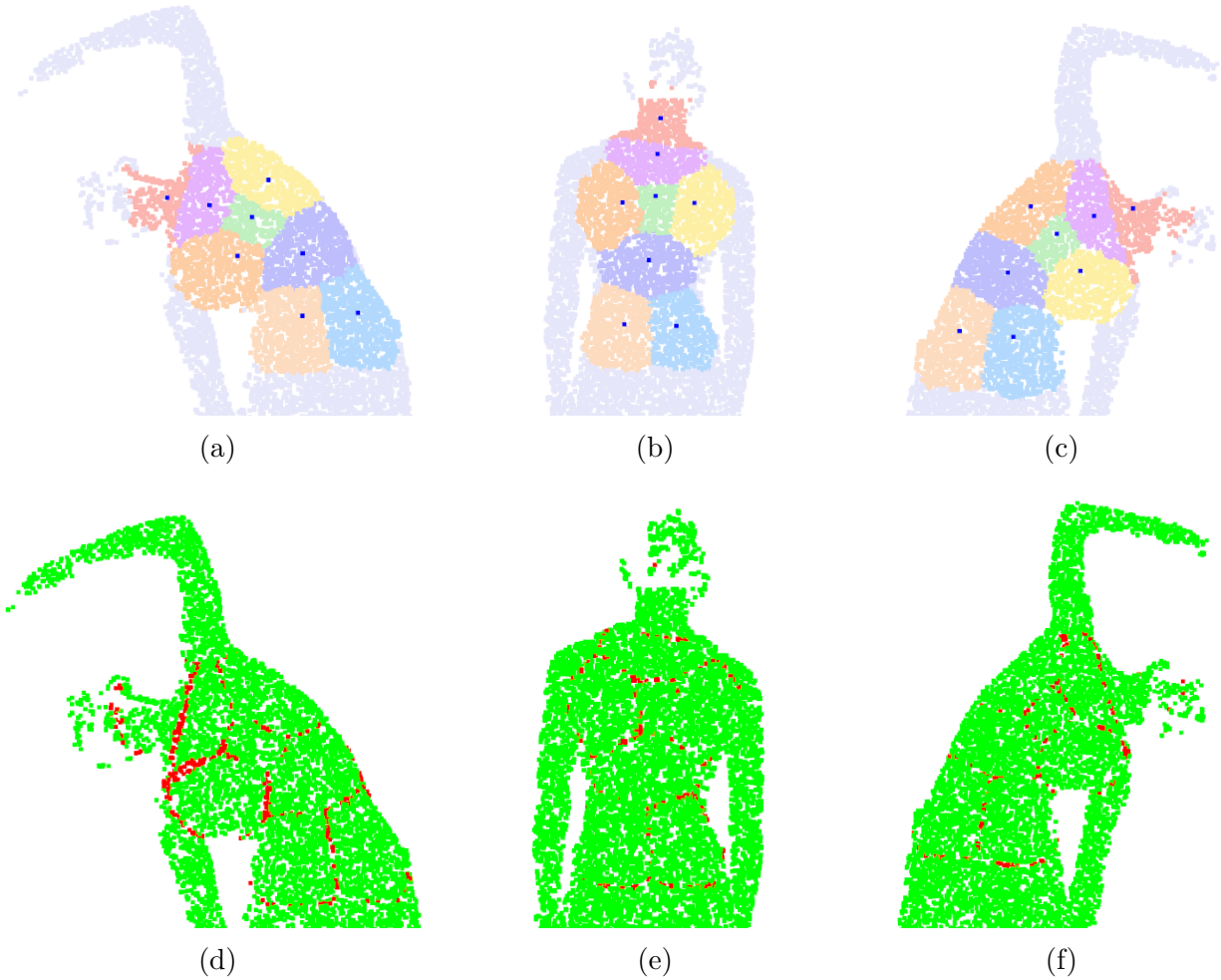
participants 9, 10, 12), which is why these participants were not used for testing or training. Table 5.2 includes the clinical information of the participants.

Table 5.1 Number of frames and markers for each sequence. NA stands for non-acquired. The blue, yellow, and green cells indicate sequences used for training, validation, and testing respectively. The gray rows indicate participants that were not used.

	self-correction				bending left		bending right		
Pt	take 1	take 2	take 3	take 4	take 1	take 2	take 1	take 2	Number of Markers
1	82	79	NA	NA	84	82	78	84	5
2	89	91	NA	NA	92	89	84	91	9
3	81	81	79	NA	81	77	82	77	8
5	81	78	NA	NA	77	83	84	81	8
6	73	79	NA	NA	82	82	71	83	8
7	NA	84	NA	NA	83	74	79	80	7
8	81	69	NA	NA	82	74	68	63	8
9	72	70	NA	NA	83	74	83	77	6
10	76	71	NA	NA	79	72	80	70	6
11	81	74	NA	NA	73	77	77	80	0
12	82	78	NA	NA	79	82	81	63	5
13	81	78	NA	NA	39	76	77	80	8
14	68	70	NA	NA	82	71	82	75	4
15	78	71	NA	NA	76	75	72	74	9
16	NA	NA	NA	NA	68	69	67	73	4
18	69	79	NA	NA	78	75	74	74	9
19	63	73	NA	NA	76	65	65	69	10
20	73	75	NA	NA	74	72	74	75	10
21	70	94	72	NA	97	79	103	103	10
22	76	70	70	NA	13	71	76	75	10
23	70	71	74	NA	76	72	75	72	9
24	73	75	69	69	34	65	69	72	9
25	78	72	NA	NA	73	71	74	55	10
26	77	73	NA	NA	70	74	76	74	9

5.2 Back Surface Segmentation

In this section, the results for back surface segmentation are presented. Figure 5.1 illustrates the point cloud representation of the predicted segments and the mislabeled points for participant 8. The evaluation metric for segmentation is overall accuracy and mIoU. The mean required time for segmentation on one sequence is also reported. Table 5.4 presents quantitative results of back surface segmentation.



(a) (b) (c)

(d) (e) (f)

Figure 5.1 Back surface segmentation results on straight, bent to the right, and bent to the left postures. The first row illustrates the predicted segments, and the second row illustrates the correct (green) and wrong (red) segmentation labels.

Table 5.2 Patients information

Year	Number	Age	Height (cm)	Weight (kg)	Scoliosis type	Cobb angle (degree)
2022	1	12	147	33.5	thoracic/thoraco-lumbar	8
	2	13	158.9	41.1	thoracic	14
	3	15	160	71	thoracic/lumbar	11
	4	13	160	43	thoraco-lumbar	21
	5	14	163.5	58	thoracic/thoraco-lumbar	19
	6	14	152	50	thoracic/thoraco-lumbar	31
	7	16	171	50	thoracic/thoraco-lumbar	21
	8	14	-	43	thoracic	22
	9	13	166.7	49.9	thoracic/thoraco-lumbar	50
	10	15	-	-	thoracic/thoraco-lumbar	46
	11	17	169	64.9	thoracic	55
	12	15	170	53.5	thoracic/thoraco-lumbar	21
	13	16	165	53.6	thoraco-lumbar	30
	14	12	-	44.9	thoraco-lumbar	40
2023	15	13	153	98	thoracic/lumbar	35
	16	14	161	51	thoracic/thoraco-lumbar	22
	17	14	152	42	thoracic	33
	18	10	129	23	thoracic	50
	19	15	163	52	thoracic/thoraco-lumbar	36
	20	11	153	43	thoracic	30
	21	18	149	51.2	thoracic/thoraco-lumbar	40
	22	12	153	36	thoraco-lumbar	22
	23	12	152	35	thoracic/lumbar	28
	24	14	-	-	thoracic/lumbar	44
	25	11	-	-	thoracic/thoraco-lumbar	44
	26	12	154	-	thoracic/lumbar	23

Table 5.3 Dataset splitting for training, validation, and testing the deep learning models.

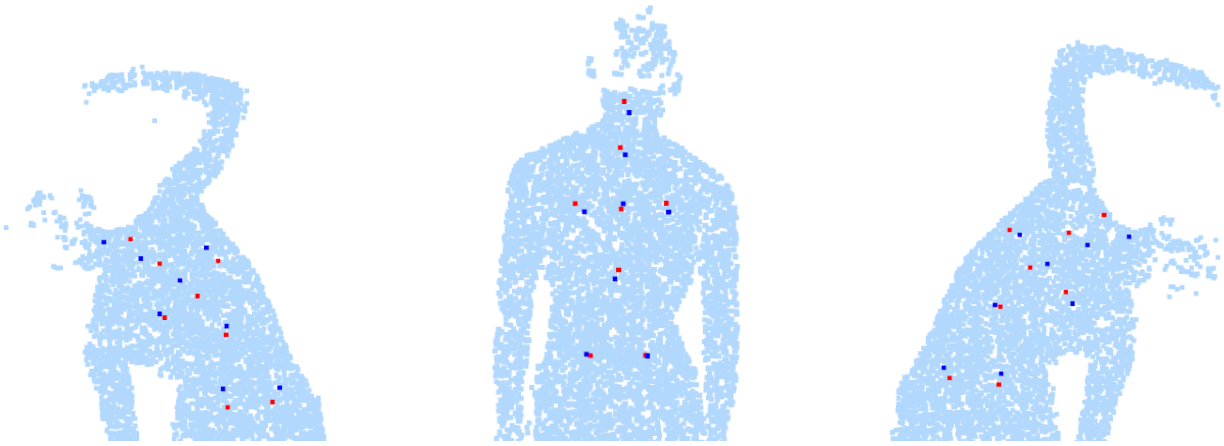
Split	Number of sequences	Number of frames
Training	31	3451
Validation	4	325
Testing	21: 4 bending + 17 self-correction	1539

5.3 Back Surface Landmark Detection

In this section, the results of our different approaches for landmark detection are presented in separate sections.

5.3.1 Regression

The first approach that we tried for landmark detection is training a regression model based on PointNet++. Figure 5.2 illustrates the landmarks predicted by the regression model for participant 8. Table 5.5 presents the quantitative results on the test dataset. Here, we present the results for 10000 and 20000 down-sampled points. The coordinates of the markers are used as the ground truth and the reported error is the Euclidean distance between the predicted coordinates and the markers' coordinates in mm.



(a) (b) (c)

Figure 5.2 The predicted landmarks for lateral bending of participant 8. The blue points are the actual markers and the red points are the predicted landmarks.

5.3.2 Segments' Centers

Our other approach for landmark detection was finding the centers of the predicted segments using the segmentation model. Figure 5.3 illustrates the segments' centers and the markers. Table 5.6 presents quantitative evaluation of the landmark detection through segmentation on the test dataset. We present the results for two different values for the segments' radius, and two down-sampling rates. According to the reported errors, reducing the radius significantly improves the accuracy of landmarks. This is explainable since by considering a smaller

Table 5.4 Back surface segmentation results on test dataset

Points	ACC	mIoU	Segmentation time per frame (s)
2500	94.81	84.32	0.15
10000	95.09	84.74	0.43

Table 5.5 Evaluation of landmark detection regression model on test dataset

	Error (Euclidean distance in mm)									
Points	C7	Sc-left	Sc-right	I-left	I-right	T-up	T-apex	T-down	Mean	Regression time per frame (s)
10000	39.67	30.10	30.68	23.79	24.26	33.93	34.75	18.84	29.50	0.1
20000	43.67	29.85	30.67	21.72	22.78	36.99	30.17	18.96	29.35	0.31

radius around each landmark when labeling the points, we will get more circular segments. Therefor, the predicted segments would be more circular, and consequently, their centers would be closer to the actual landmarks. Figure 5.3 demonstrates this point more clearly.

Table 5.6 Evaluation of landmark detection through segmentation on test dataset

	Error (Euclidean distance in mm)								
Radius (mm)	C7	Sc-left	Sc-right	I-left	I-right	T-up	T-apex	T-down	Mean
100	27.68	41.77	38.1	31.39	34.59	28.22	22.15	23.2	30.89
30	16.99	28.01	23.47	7.29	10.56	26.66	24.99	30.34	25.0

In table 5.7 the quantitative results achieved by segmentation with radius 0.05 and 10000 down-sampled points in the input point clouds are presented per participant (test dataset) and landmarks. The last three participants (9, 10, 12) do not have the markers on the pelvis landmarks. The error is considerably higher for participants 9 and 10 comparing to other participants. This can be attributed to the fact that these participants have a larger Cobb angle. Since most of our recruited participants had small curves (according to the acquisition protocol), our training dataset does not contain many participants with more severe curves. As a result, the performance for participants 9 and 10 is not as good as participants with smaller Cobb angles (6, 8, 23). This is a limitation of our method that can be addressed by collecting more data sampled from a more diverse AIS population (larger Cobb angles, different types of AIS). Figure 5.4 depicts the predicted landmarks for three participants. Here again, the difference between the performance for participant 10 with a larger curve and participants 8 and 23 with smaller curves is noticeable.

5.3.3 Comparisons

According to the reported errors, the segmentation model trained with radius 0.05 (around 30mm) and 10000 down-sampling rate achieves the most accurate estimation of the landmarks, comparing to other segmentation experiments and regression model.

Table 5.8 presents a comparison between the results of our method and the results reported

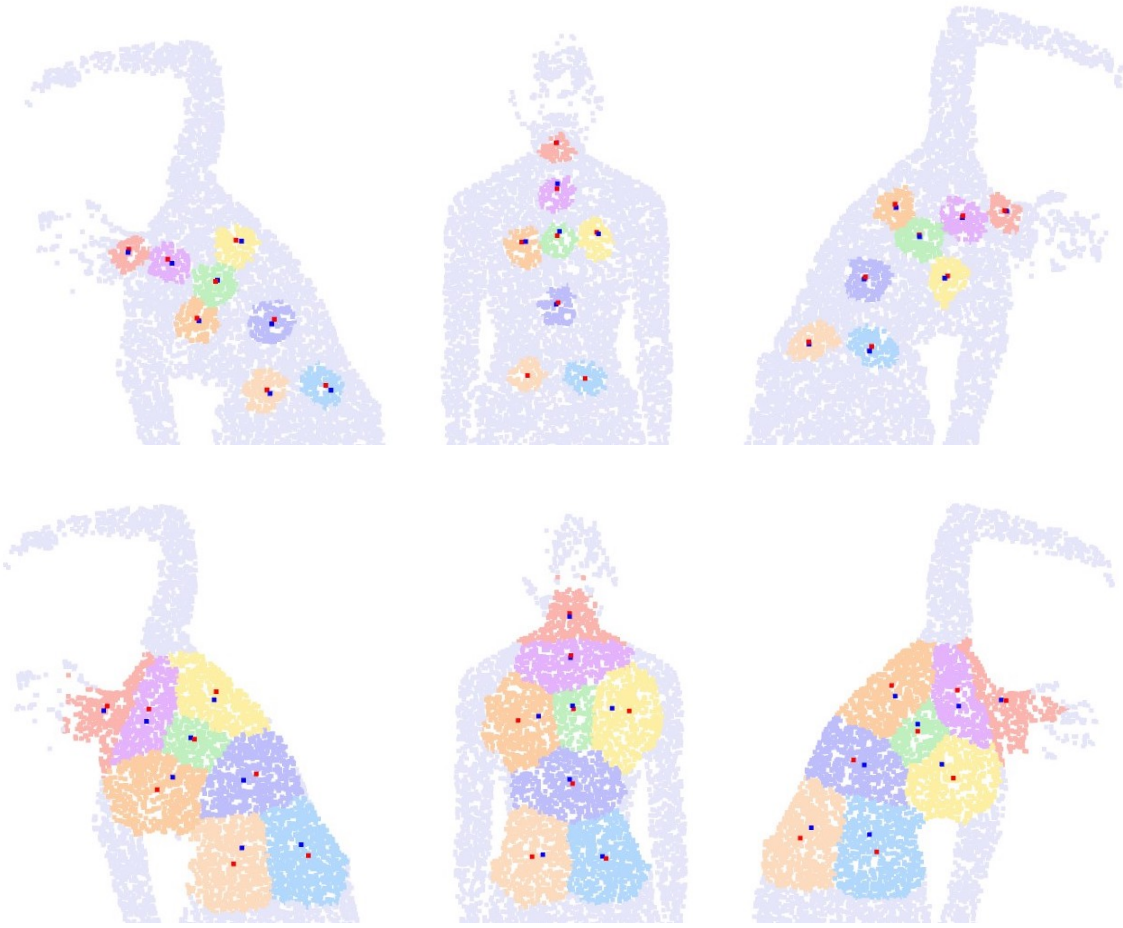


Figure 5.3 Predicted segments and their centers (red) and the markers (blue) for different postures of participant 8 with 10000 down-sampled points. The first row shows the results for search radius 0.05 (around 30mm), and the second one for radius 0.15 (around 100mm).

Table 5.7 Evaluation of landmark detection through segmentation on test dataset. The symbol - indicates that the marker is missing for the associated participant.

	Error (Euclidean distance in mm)								
Participant	C7	Sc-left	Sc-right	I-left	I-right	T-up	T-apex	T-down	Mean
8	4.41	7.04	6.62	6.42	9.48	5.32	4.86	5.61	6.22
6	15.54	11.02	9.91	8.63	21.59	43.4	25.8	10.5	18.3
23	9.2	11.76	15.88	7.51	7.97	8.07	6.33	10.32	9.63
9	21.38	89.32	38.76	-	-	16.47	55.51	36.86	43.05
10	26.24	34.83	44.99	-	-	48.53	41.24	50.27	41.02
12	25.19	13.72	24.65	-	-	38.17	15.81	34.23	25.29
mean	16.99	28.01	23.47	-	-	26.66	24.99	30.34	25.08

in [41] and the inter-observer variability, also reported in [41], based on mean error measured by Euclidean distance in mm. Since the other method is designed for different back surface landmarks, only the reported mean errors are compared rather than per-landmark errors. The method proposed in [41] is the closest to ours since it specifically focuses on back surface landmarks, sharing a few common landmarks with ours, such as C7. The mean landmark error is not reported, and we computed it based on the per-landmark errors. In the same work, landmarks localized by three experts are used as the reference. The error between landmarks localized by each pair of experts' landmarks is also reported since there is variability among them. Here, we took the average of each pair's errors over the landmarks and then took an average of the three errors over the pairs to compute the inter-observer variability.

According to the comparisons, our method with back segmentation has a lower landmark localization error comparing to the method in [41], and it is comparable to the inter-observer variability of experts' landmark localization also reported in [41].

5.4 Landmarks Tracking

In this section, we present visualizations of the landmarks' motion in 3D based on the predicted landmarks throughout a lateral bending and a self-correction sequence of participant 8. Since landmark detection through segmentation has achieved better results, we use this model for finding the trajectories. Figure 5.5 illustrates the 3D trajectories of landmarks in a bending sequence. Figures 5.6 and 5.7 illustrate the 2D trajectories projected on the three anatomical planes (Figure 2.3).

The other approach that we tried for landmark tracking was scene flow estimation by FlowNet3D model. However, since the scene flow errors accumulate throughout the sequence, the results were not satisfying and are not included here.

(a) Pt 8 (b) Pt 10 (c) Pt 23

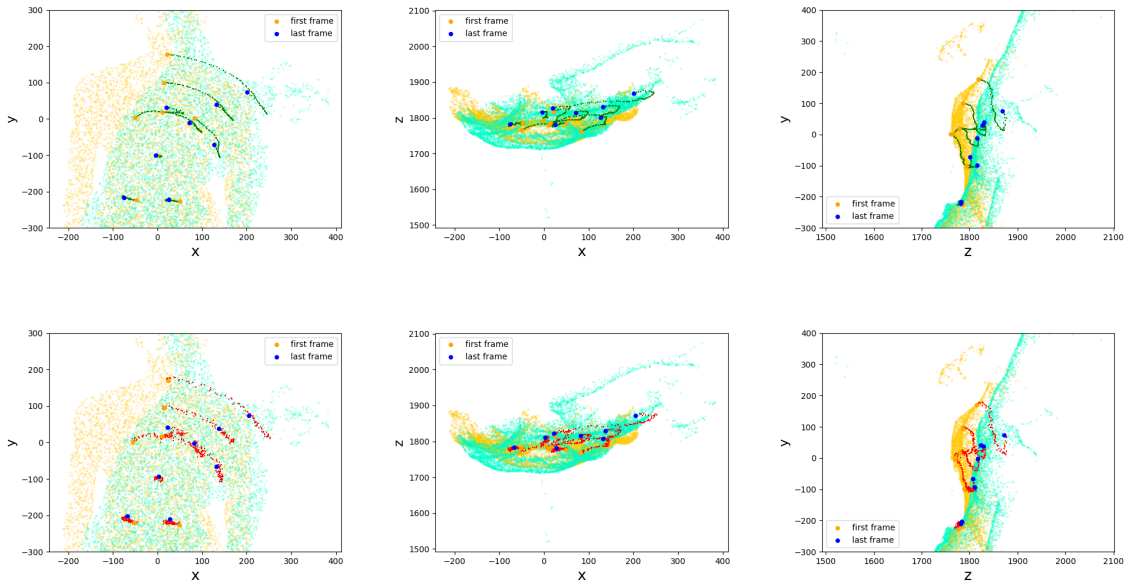
Figure 5.4 Predicted landmarks for participant 8, 10, and participant 23. The red points are the predicted landmarks, and the blue points are the manually placed markers.

Table 5.8 Comparison between our results and the results reported in [41], based on mean Euclidean distance (mm).

Method	Error (mm)
Inter-observer variability [41]	23.86
Proposed method [41]	27.68
Regression 20000 pts (ours)	29.35
Segmentation 10000 pts (ours)	**25.08**

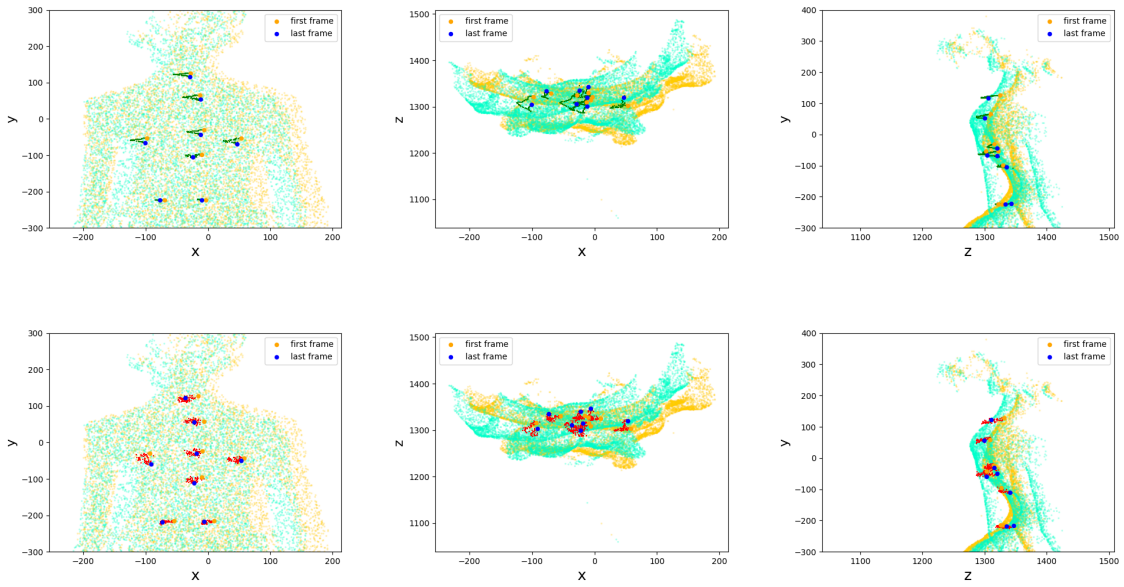
(a) (b) (c)

Figure 5.5 3D representations of tracking the landmarks throughout a lateral bending to the right. Green trajectories are based on actual markers and red trajectories are based on the predicted landmarks.



(a) Coronal (b) Transverse (c) Sagittal

Figure 5.6 2D projections of the trajectories on the Coronal, Sagittal, and Transverse planes for free lateral bending, participant 8.



(a) Coronal (b) Transverse (c) Sagittal

Figure 5.7 2D projections of the trajectories on the Coronal, Sagittal, and Transverse planes for self-correction, participant 23.

5.5 Dynamic Quantification of Trunk Surface Metrics

To test the performance of our method in dynamic trunk metrics quantification, we compare the calculated metrics over time based on automatically detected landmarks with the calculated metrics based on the markers' coordinates for self-correction and lateral bending sequences. Our evaluation metric is Pearson correlation coefficient. We also provide the graphs of the metrics plotted over time for visual evaluation. The correlation is calculated and presented for participants 8, 23, and 24. We could not include more participants for self-correction, since the other test participants (6, 9, 10, and 12) did not have all the 8 markers required for calculating the metrics. Moreover, the lateral bending sequences of these participants (6, 9, 10, and 12) are not annotated yet and therefor, are not included here. The visualizations are only presented for participant 23 and 24, since they have larger Cobb angles and the metrics change in a wider range, resulting in more representative graphs.

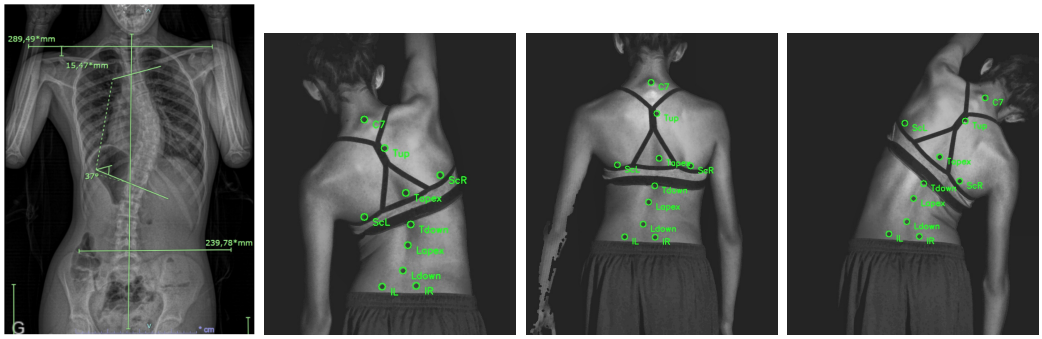
5.6 Lateral Bending

The values of the Pearson correlation coefficient computed for Cobb angle values based on markers' coordinates and the Cobb angle values based on predicted landmarks for lateral

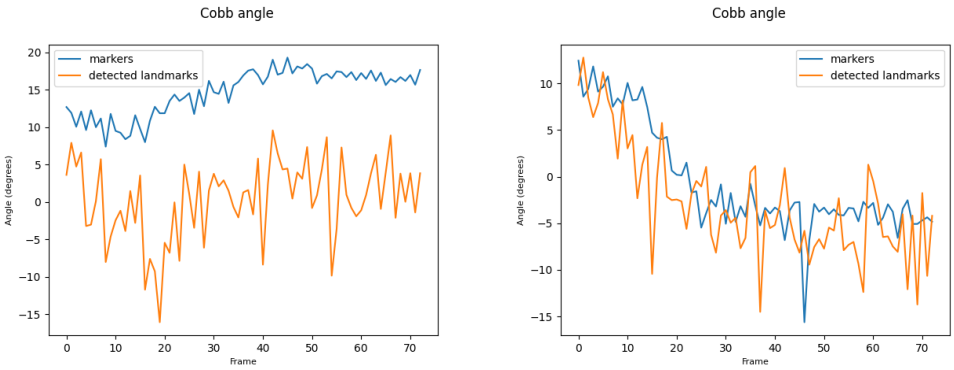
bendings of participant 8, 23, and 24 are presented in table 5.9. The included lateral bendings are on the convex(right) side of the curve. According to the reported values, we can see there is a high correlation for participants 23 and 24. However, the result for participant 8 shows a very low correlation. This can be attributed to the fact that participant 8 has a very small Cobb angle (22°) and therefor, the Cobb angle changes during a lateral bending are in a narrow range, which results is higher sensitivity to error and a low correlation. Figures 5.8 and 5.9 depict Cobb angle analysis of free lateral bending sequences for participants 23 and 24. As discussed before, the analysis of Cobb angle in lateral bendings on the convex side of the curve is used in assessing the AIS curve flexibility before surgery.

Table 5.9 Correlation between clinical metrics based on automatically detected landmarks and markers coordinates. Metrics 1, 2, 3, 4, and, 5 refer to Cobb angle, transverse scapulae angle, coronal scapulae angle, trunk imbalance, and scapulae asymmetry respectively. The landmark detection error is the mean Euclidean distance in mm between the predicted landmarks and markers' coordinates for each participant. The correlation is the Pearson correlation coefficient computed for the values of each metric based on markers and based on predicted landmarks. The range of the values for correlation is between -1 and 1 and absolute values closer to 1 indicate a higher correlation. Statistically significant correlations at p-value < 0.05 level are indicated with *.

Participant	Landmarks error (mm)	Sequence	Metric	Correlation (<= 1 and >= -1)
8	6.33	Lateral bending	1	-0.11
	6.22	Self-correction	1	0.12
			2	0.93*
			3	0.51*
			4	0.78*
			5	0.62*
23	6.825	Lateral bending	1	0.60*
	9.63	Self-correction	1	-0.24*
			2	0.99*
			3	-0.21
			4	0.59*
			5	0.67*
24	6.45	Lateral bending	1	0.78*
	8.88	Self-correction	1	-0.012
			2	0.92*
			3	0.41*
			4	0.75*
			5	-0.69*



(a) X-ray (b) Left (c) Resting (d) Right



(e) Left bending (f) Right bending

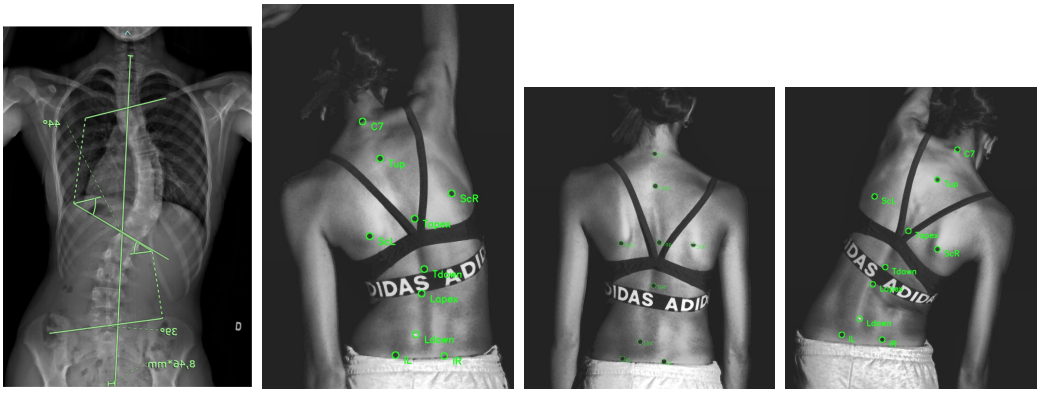
Figure 5.8 Cobb angle analysis of free lateral bending performed by participant 23. The blue graphs represent the actual markers and the orange ones represent the predicted landmarks (segmentation method).

5.7 Self-Correction

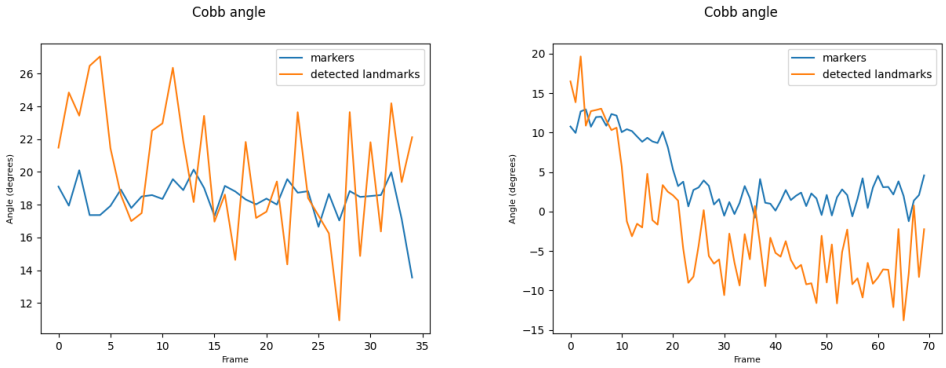
Table 5.9 includes the Pearson correlation coefficient values for 5 different metrics during self-corrections of participants 8, 23, and 24. According to the reported values, the metrics for trunk imbalance and transverse scapulae angle (metrics 4 and 2) are highly correlated for all three participants. This can be attributed to a wider range of motion in the transverse plane and the pelvis, which makes these metric more tolerant to landmark detection error. For Cobb angle and coronal angle, the results show a low correlation. Figures 5.10 and 5.11 illustrate the graphs of the metrics over time for self-correction of participant 23 and 24.

5.8 Discussion

According to the presented experiments, our results for back surface segmentation are highly accurate. In addition, our landmark detection method based on segmentation achieves an



(a) X-ray (b) Left (c) Resting (d) Right



(e) Left bending (f) Right bending

Figure 5.9 Cobb angle analysis of free lateral bending performed by participant 24. The blue graphs represent the actual markers and the orange ones represent the predicted landmarks (segmentation method).

accuracy comparable to inter-observer variability (see table 5.4) and better than the method proposed in [41]. It is notable that our results are based on experiments on both bent and straight postures of subjects with AIS, while the other method and the inter-observer error are solely based on straight positions of healthy subjects.

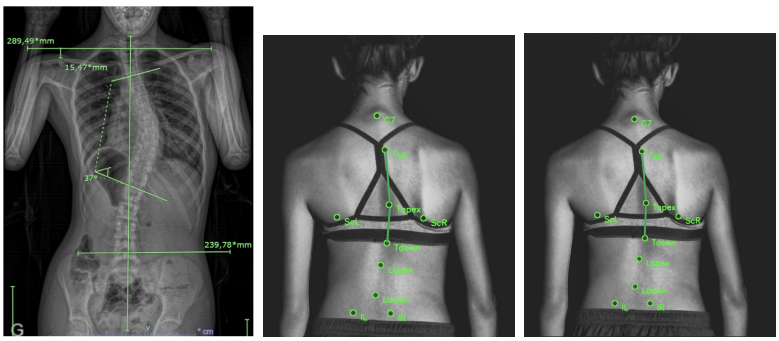
Using the automatically detected landmarks, the trajectories of the 3D landmarks throughout a self-correction or lateral bending sequence can be estimated (see figures 5.5, 5.7, and 5.6). These trajectories provide a better understanding of the dynamic trunk motion. In that regard, it is notable that according to both trajectories and Cobb angle graphs (figures 5.8 and 5.9), the maximum bending position is achieved before the last frame. Evaluating the movement dynamically allows us to monitor all the frames and find the frame where maximum bending is achieved, which is not possible through static evaluation. In the current clinical practice, a static radiograph is acquired at the end of the movement and is used to assess

curve flexibility [28]. As observed in our experiments, an static acquisition at the end of the lateral bending does not necessarily correspond to the maximum bending capacity. Therefor, dynamic acquisitions of lateral bending provide a better estimation of curve flexibility.

As presented in the results of quantification of trunk metrics over time, for lateral bending sequences, the Cobb angle is computed for all the frames in a sequence and is plotted over the duration of an acquisition. Lateral bending on the convex side of the curves causes a visible reduction in the curve and the Cobb angle. By measuring the curve reduction, the clinicians are able to assess the curve flexibility. Both graphs for bending sequences (figures 5.8 and 5.9) on the convex side of the curve show a reduction in the Cobb angle. The markers and predicted landmarks graphs seem to be highly correlated implying that our method provides accurate clinical analysis for Cobb angle in lateral bending. More experiments with more participants are needed to further test the performance of our method, but the results of the current experiments are promising.

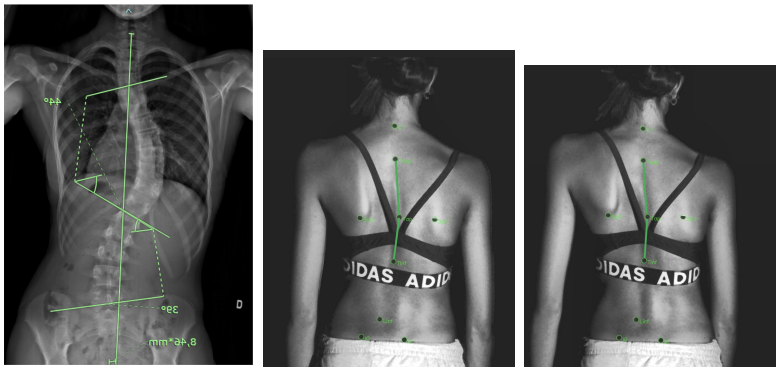
For self-correction sequences, we computed five clinical metrics and plotted the graphs over the duration of the sequence. Self-correction is a very subtle movement and the changes in the metrics during this movement are small. Due to the small scale of the movement, even slight errors in landmark detection can result in a relatively high error in calculating the metrics. For example, for participant 23 (figure 5.10), the range of changes in Cobb angle during the self-correction is around 10°. Therefor, as illustrated in the plots, the landmark detection error changes the graphs significantly and results in a low correlation. To improve these results , the landmark detection algorithm needs to be further improved.

Overall, based on all the results, our proposed tool can obtain accurate landmark detection, segmentation, and landmarks trajectories. For trunk metrics quantification over time, our experiments show promising results but still not accurate enough to be applied in clinical practice. Despite all its limitations, this project is an original work which takes a determinate step toward automatic dynamic evaluation of AIS. Our method can be used as a baseline for future work. The ultimate objective of this project is a real-time automatic dynamic trunk motion quantification tool that can be used by clinicians in assessing AIS movements. In the next section, the future steps that can improve upon the proposed tools are discussed.



(a) X-ray (b) Resting (c) Corrected

Figure 5.10 Clinical analysis of self-correction performed by participant 23. The blue graphs represent actual markers and the orange ones represent the automatically detected landmarks (segmentation method).



(a) X-ray (b) Resting (c) Corrected

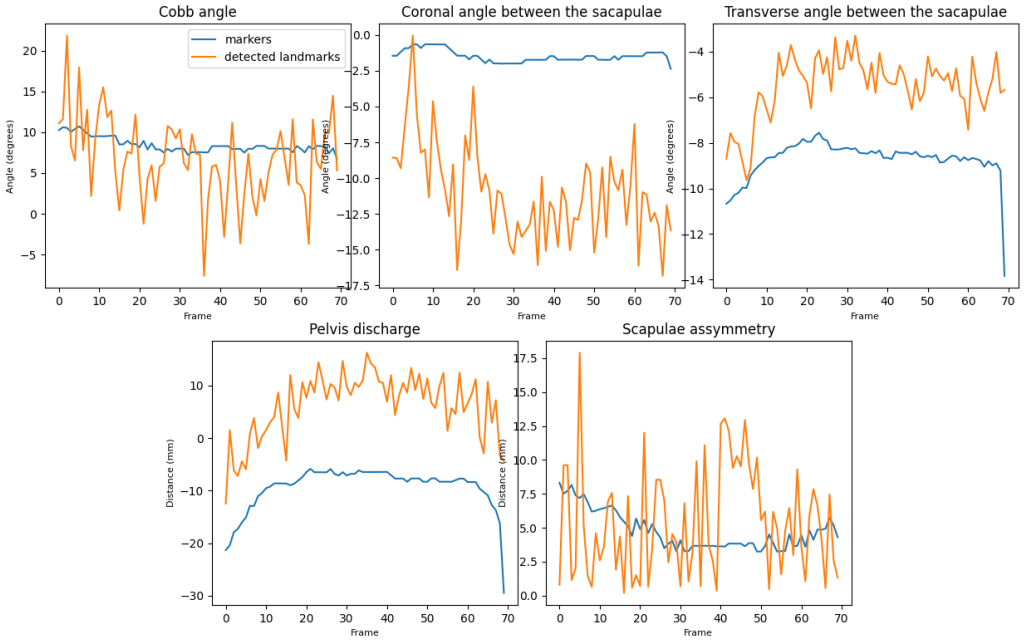


Figure 5.11 Clinical analysis of self-correction performed by participant 24. The blue graphs represent actual markers and the orange ones represent the automatically detected landmarks (segmentation method).

CHAPTER 6 CONCLUSION

6.1 Summary of Works

In this work, we collected, preprocessed, and annotated a 3D dynamic dataset of 24 patients with AIS, which can be used for future studies, and implemented a dynamic AIS evaluation method based on deep learning models. To our knowledge, this is the first work in the literature to bring time into AIS evaluation.

The functionalities of our method include back surface segmentation, back surface landmark detection, visualizing the 3D trajectories of the landmarks throughout a motion sequence, and quantifying trunk metrics over time for lateral bending and self-correction sequences. According to the presented results, our segmentation and landmark localization methods show strong performance on the test dataset. The performance of our method in quantifying the back surface metrics is promising but still requires higher accuracy to be applicable in clinical analysis.

This work takes the first step toward automating the dynamic 3D trunk motion quantification. Considering the originality of this work, the future holds many potential improvements and modifications that can boost the performance of our method and ultimately, integrate it with the clinical analysis of AIS. In the following sections, the limitations, recommendations, and guidelines for future work that can improve upon this project are discussed.

6.2 Limitations

Our proposed tools for dynamic quantification of trunk metrics, despite achieving promising results, have certain limitations.

First, the landmark detection model's performance varies for different types of curves with different severity. Our results show that the landmark detection error increases for participants with higher Cobb angles. This is because the majority of the participants in our training dataset had small Cobb angles and the model is not trained properly for higher Cobb angles.

The results of clinical analysis of self-correction movement are not accurate for most participants. The reason is that the self-correction exercise has a very small range of motion, therefor even small landmark detection error results in a relatively high error in metrics analysis for this exercise. The next section provide some recommendation and guidelines for

improving upon the proposed tools.

6.3 Recommendations and Future Work

This project is, to our knowledge, the first step toward quantifying dynamic evaluation of AIS. Hopefully, this will act as the baseline for future work and will be the first step of the many steps to come. In this section, we provide recommendations and guidelines for future work in this direction.

A certain solution for improving the accuracy of landmark detection and thus trunk motion quantification is acquiring more data from a more diverse AIS population. Our proposed dataset mostly includes mild AIS curves that are located in the upper half of the trunk. Including patients with more severe AIS deformities and curves located in the middle or lower parts of the trunk and using them for training would definitely strengthen the performance of the models.

Another future direction that can be explored is experimenting with other deep learning models. Although our experiment with FlowNet3D did not achieve good results, fine-tuning this model on AIS data might change the performance of this model. FlowNet3D is just one of the possible models to explore, but lots of other deep learning architectures exist and new models are emerging every day. We recommend experimenting with different models and comparing their results.

Another future modification would be using more markers to mark more landmarks for future data acquisitions. Having more landmarks, especially on the spine, would help with better quantifying the deformity and a more complete and accurate trunk motion quantification.

REFERENCES

[1] Stefano Negrini et al. “2016 SOSORT guidelines: orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth”. In: *Scoliosis and Spinal Disorders* 13.1 (2016), p. 3.

[2] Bhavna Mehta et al. “Non-Invasive Assessment of Back Surface Topography: Technologies, Techniques and Clinical Utility”. In: *Sensors* 23.20 (2023).

[3] Azmi Hamzaoglu et al. “Assessment of Curve Flexibility in Adolescent Idiopathic Scoliosis”. In: *Spine* 30(14) (2005), pp. 1637–42.

[4] Sarah Dupuis et al. “Global Postural Re-Education in Pediatric Idiopathic Scoliosis: A Biomechanical Modeling and Analysis of Curve Reduction during Active and Assisted Self-Correction”. In: *BMC Musculoskeletal Disorders 19* (2018).

[5] Lama Seoud. “Analyse de la relation entre les déformations scoliotiques du tronc et celles des structures osseuses sous-jacentes”. URL: `https://publications.polymtl.ca/801/`.

[6] Brody A. Frost, Sandra Camarero-Espinosa, and E. Johan Foster. “Materials for the Spine: Anatomy, Problems, and Solutions”. In: *Materials* 12.253 (2019).

[7] G. R. Ebenbichler et al. “Sensory-motor control of the lower back: implications for rehabilitation.” In: *Medicine and science in sports and exercise* 33(11) (2001), pp. 1889–98.

[8] J. Gordon Betts et al. *Anatomy and Physiology*. OpenStax, 2013.

[9] Ana San Román Gaitero et al. “Imaging Methods to Quantify the Chest and Trunk Deformation in Adolescent Idiopathic Scoliosis: A Literature Review”. In: *Healthcare* 11.10 (2023).

[10] Alessandra B Porto and Victor H A Okazaki. “Thoracic Kyphosis and Lumbar Lordosis Assessment by Radiography and Photogrammetry: A Review of Normative Values and Reliability”. In: *Journal of Manipulative and Physiological Therapeutics* 41.8 (2018), pp. 712–723.

[11] Petros Patias et al. “A review of the trunk surface metrics used as Scoliosis and other deformities evaluation indices”. In: *Scoliosis* 5.1 (2010), p. 12.

[12] Zhifeng Zhou, Jia Zhu, and Chengxian Yao. “Vertebral Center Points Locating and Cobb Angle Measurement Based on Deep Learning”. In: *Applied Sciences* 13.6 (2023).

[13] Pazos V et al. “Reliability of trunk shape measurements based on 3-D surface reconstructions”. In: *Eur Spine J.* 16(11) (2007), pp. 1882–91.

[14] C J Goldberg et al. “Surface topography and the several components of scoliotic deformity”. In: *Studies in health technology and informatics* 88 (2002), pp. 67–9.

[15] P. Debanné et al. “Evaluation of reducibility of trunk asymmetry in lateral bending”. In: *Studies in health technology and informatics* 158 (2010), pp. 72–78.

[16] H. et al Berdishevsky. “Physiotherapy scoliosis-specific exercises – a comprehensive review of seven major schools”. In: *Scoliosis* 11.20 (2016).

[17] De Mauroy JC et al. “The new lyon ART brace versus the historical Lyon brace: a prospective case series of 148 consecutive scoliosis with short time results after 1 year compared with a historical retrospective case series of 100 consecutive scoliosis; SOSORT award 2015 winner.” In: *Scoliosis* 10.26 (2015).

[18] Weiss HR. “The Schroth scoliosis-specific back school—initial results of a prospective follow-up study.” In: *Z Orthop Ihre Grenzgeb* 133.2 (1995), pp. 114–7.

[19] Romano M et al. “SEAS (Scientific Exercises Approach to Scoliosis): a modern and effective evidence based approach to physiotherapic specific scoliosis exercises.” In: *Z Orthop Ihre Grenzgeb* 10.3 (2015).

[20] Jelačić M et al. “Barcelona Scoliosis Physical Therapy School – BSPTS – based on classical Schroth principles: short term effects on back asymmetry in idiopathic scoliosis.” In: *Scoliosis* 7 Suppl 1:057 (2012).

[21] Piotrowski J. Durmala J Kotwicki T. “Stabilization of progressive thoracic adolescent idiopathic scoliosis using brace treatment and DoboMed physiotherapy.” In: *Scoliosis* 4 Suppl 2:O29 (2009).

[22] Betts T. “The development of a classification system for the treatment of scoliosis by the side shift.” In: *Scoliosis* 9 Suppl 1:O66 (2014).

[23] M’hango A. Białek M. “FITS concept - functional individual therapy of scoliosis.” In: *studies in health technology and informatics. the conservative scoliosis treatment. 1st SOSORT instructional course lectures book* 9 Suppl 1:O66 (2008).

[24] Białek M. “Conservative treatment of idiopathic scoliosis according to FITS concept: presentation of the method and preliminary, short term radiological and clinical results based on SOSORT and SRS criteria”. In: *Scoliosis* 6(1).25 (2011).

[25] R. Compagnon et al. “Side bending radiographs and lowest instrumented vertebra in adolescent idiopathic scoliosis: A French quality-of-care study”. In: *Orthopaedics and Traumatology: Surgery and Research* (2022).

[26] Philippe Violas et al. “Correction of thoracic adolescent idiopathic scoliosis via a direct convex rod manoeuvre”. In: *Orthopaedics and Traumatology: Surgery and Research* (2019).

[27] C. Hirsch et al. “Flexibility analysis in adolescent idiopathic scoliosis on side-bending images using the EOS imaging system”. In: *Orthopaedics & Traumatology: Surgery & Research* (2016).

[28] Rodrigues LM et al. “Comparison between different radiographic methods for evaluating the flexibility of scoliosis curves”. In: *Acta Ortop Bras* 22(2) (2014), pp. 78–81.

[29] Yonghuai Liu et al. *3D Imaging, Analysis and Applications*. Springer Cham, 2020.

[30] Harmouche R et al. “Multimodal image registration of the scoliotic torso for surgical planning”. In: *BMC Med Imaging* 13, 1 (2013).

[31] Yulan Guo et al. “Deep Learning for 3D Point Clouds: A Survey”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.12 (2021), pp. 4338–4364.

[32] Charles R. Qi et al. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 2017.

[33] Charles Ruizhongtai Qi, Li Yi, and Leonidas J. Su Haoand Guibas. “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space”. In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc., 2017.

[34] Hengshuang Zhao et al. “Point Transformer”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2021, pp. 16259–16268.

[35] Guocheng Qian et al. “PointNeXt: Revisiting PointNet++ with Improved Training and Scaling Strategies”. In: *Advances in Neural Information Processing Systems*. Vol. 35. Curran Associates, Inc., 2022, pp. 23192–23204.

[36] Mikaela Angelina Uy et al. “Revisiting Point Cloud Classification: A New Benchmark Dataset and Classification Model on Real-World Data”. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019, pp. 1588–1597.

[37] Iro Armeni et al. “3D Semantic Parsing of Large-Scale Indoor Spaces”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 1534–1543.

[38] Xingyu Liu, Charles Ruizhongtai Qi, and Leonidas J. Guibas. “Learning Scene Flow in 3D Point Clouds”. In: *CoRR* abs/1806.01411 (2018).

[39] Weiya Wang et al. “An automatic extraction method on medical feature points based on PointNet++ for robot-assisted knee arthroplasty”. In: *The International Journal of Medical Robotics and Computer Assisted Surgery* 19.1 (2023), e2464.

[40] Pengcheng Xi, Chang Shu, and Rafik Goubran. “Localizing 3-D Anatomical Landmarks Using Deep Convolutional Neural Networks”. In: *2017 14th Conference on Computer and Robot Vision (CRV)*. 2017, pp. 197–204.

[41] Daniel Ledwoń et al. “Real-Time Back Surface Landmark Determination Using a Time-of-Flight Camera”. In: *Sensors* 21.19 (2021).

[42] Li Yi et al. “A Scalable Active Framework for Region Annotation in 3D Shape Collections”. In: *SIGGRAPH Asia* (2016).

[43] Zhirong Wu et al. “3D ShapeNets: A Deep Representation for Volumetric Shapes”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2015.

[44] Carole Fortin et al. “Trunk imbalance in adolescent idiopathic scoliosis”. In: *The spine journal* (2016).