

Automated Rod Length Measurements on Radiographs and Sonograms in Children with Early Onset Scoliosis

by

Mohammad Humayun Kabir

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Software Engineering and Intelligent Systems

Department of Electrical and Computer Engineering
University of Alberta

© Mohammad Humayun Kabir, 2024

Abstract

Early Onset Scoliosis (EOS) is a medical condition that is defined as a lateral curvature of the spine with vertebral rotation in children under age 10. Approximately 2-3% of children worldwide have scoliosis. Surgical intervention is the most effective management to treat these children who have severe scoliosis. Currently, the magnetically controlled growing rods (MCGR) surgery which can gradually extend externally to straighten the curve is the most cost-effective method to treat EOS. Radiographs are taken at every visit twice, once before the rod length adjustment and once after, to measure manually the rod length change.

The existing manual measurement techniques face various obstacles, including inconsistencies between different raters, poor quality of radiographs, patients' motion, and their different postures during radiography taken, and disparities in image resolution. In addition, taking radiography exposes children to harmful ionizing radiation. Therefore, the ultrasound (US) imaging method has also been reported in the literature to image the MCGR for measuring the change of the growing rods.

The objective of this work was to develop machine learning (ML) algorithms to automatically measure the rod length adjustment on both radiographs and US images at each visit. Clinical data were collected, and studies were conducted to validate both algorithms for radiographs and sonograms.

Driven by the imperative for accuracy, this study utilized ML algorithms to construct an autonomous system. Both radiography and US automated systems were developed using Mask Regional Convolution Network (Mask RCNN) techniques, which were widely recognized and

commonly used in the field of computer vision for object detection and instance segmentation.

The application of Mask RCNN was implemented utilizing the Detectron2 framework.

Instead of using absolute measurements, a calibration technique was applied. Three ML models: rod model, 58mm fixed length model, and head-piece model were developed to extract the rod length from radiographs. Three-hundred and eighty-seven radiographs were used for model development, and 60 radiographs with 118 rods were used for testing. The radiography automated system eventually demonstrated an acceptable inter-method correlation coefficient of 0.90 and a mean absolute difference (MAD) of 0.98 ± 0.88 mm when compared to manual adjustment measurements.

In the US imaging system, similar ML algorithms were developed. Two ML models: the Boundary model and the Rod model, were developed. A study that included 90 US images acquired from 23 EOS patients was conducted. Among the 90 images, 70 images were used for model development, and 20 images were used for testing. The MAD between the Artificial Intelligence (AI) measurements versus the manual measurements was 1.2 ± 1.46 mm and the reliability of the inter-method correlation coefficient was 0.96.

In summary, this thesis reported a new and automated approach to measuring rod adjustments in children who have MCGR installation. The automated approaches were accurate to measure the adjustments in a faster way, saving time of clinicians, while reducing the chance of error made by the traditional manual approach.

Preface

This thesis is an original work by Mohammad Humayun Kabir, under the supervision of Drs. Edmond Lou and Marek Reformat. This study received ethics approval (Pro00109499) from the University of Alberta Health Research Ethics Board. The base research of this thesis was a collaboration work of the Department of Electrical and Computer Engineering and the Department of Surgery at the University of Alberta. Ms. Sarah Southon Hryniuk, a nurse practitioner, is responsible for participants' recruitment and adjustment of the MCGR. I was responsible for designing, planning, and implementing the overall system associated with Machine Learning.

The materials in chapter 3 were exported from the manuscript which was submitted to a prestigious journal.

M. H. Kabir, M. Reformat, S. S. Hryniuk, K. Stampe, and E. Lou, "Validity of Machine Learning Algorithms for Automatically Extract Growing Rod Length on Radiographs in Children with Early Onset Scoliosis," *Medical & Biological Engineering & Computing*.

And materials in chapter 4 was exported from another submitted manuscript.

M. H. Kabir, M. Reformat, S. S. Hryniuk, K. Stampe, and E. Lou, "Automated Method for Growing Rod length Measurement on Ultrasound Images in Children with Early Onset Scoliosis", *Ultrasound in Medicine and Biology*.

Throughout both chapters, I played a crucial part in the entirety of the research process, encompassing tasks such as data preparation, model construction, analysis, and paper writing. More precisely, I led the efforts in preparing the data, creating the models, and conducting the analysis, guaranteeing the strength and reliability of the research results. In addition, I meticulously composed the manuscripts, integrating the research findings into coherent narratives that are fit for publishing.

Dr. Edmond Lou played a crucial role in the study project by providing essential assistance in data gathering and making excellent contributions to the editing of the publication. His knowledge and insights greatly improved the quality and rigour of the research outputs. In addition, both Drs. Edmond Lou and Marek Reformat acted as supervisory authors, providing

direction and expertise in developing the research, improving model designs, and completing the manuscript.

The research team, consisting of myself, Drs. Edmond Lou, and Marek Reformat, highlight the interdisciplinary nature of the research and the synergistic combination of knowledge applied to the project. Collectively, we successfully managed the complexities of the research process, resulting in the submission of chapters that are well-positioned to make a significant impact on the scientific conversation within our areas of expertise.

Acknowledgments

I would like to express my sincere appreciation to my thesis supervisor and co-supervisor, Dr. Edmond Lou and Dr. Marek Reformat, for their invaluable guidance and unwavering assistance, which have been instrumental in facilitating this research project. Their contributions of encouragement, helpful suggestions, and invaluable ideas have significantly accelerated my academic progress.

I would like to express my profound gratitude to my parents, Abdur Rahim and Ferdous Begum, for their steadfast faith in my abilities and consistent encouragement. The encouragement and support they have provided have acted as a catalyst for advancements in this academic pursuit.

I wish to express my heartfelt gratitude to my beloved wife, Ummay Habiba Wani, for her motivational remarks, patience, and understanding during the challenging phases of this research undertaking. She has consistently relied on me as a pillar of strength.

In addition, my siblings, who have each made a unique contribution to my academic journey, merit recognition. I would like to extend special recognition to Tamzid Kabir, my youngest sibling, for his exceptional encouragement and support.

I am writing to express my gratitude and recognition for the collaborative atmosphere that is present in the academic community at the University of Alberta. My sincere appreciation goes out to my colleagues and fellow researchers. Due to the members' companionship and sagacity, an intellectually stimulating and fruitful environment has been created for research.

The achievement would not have been feasible without the collaborative motivation and assistance provided by these individuals. Every individual contribution, irrespective of its scale, has had a substantial influence on this scholarly endeavor, and for that, I am extremely appreciative.

Table of Contents

Chapter 1 Introduction.....	1
1.1 Background	1
1.2 Motivation	3
1.3 Objectives	4
1.4 Contribution	4
1.5 Thesis Outline	5
Chapter 2 Background and Literature Review	6
2.1 Overview	6
2.2 Background of EOS	6
2.3 TGR Complications for EOS	7
2.4 Magnetically Controlled Growing Rod (MCGR)	11
2.5 Cost Analysis and Complication Rate Analysis Between TGRs and MCGRs	11
2.6 Measurement of the MCGR Rod Adjustment	14
2.7 Drawbacks of Radiography and Emerging US for EOS Applications	14
2.8 Literature of Applying AI in Medical Imaging	16
2.8.1 CNN in Medical Images	16
2.8.2 Detectron2	18
2.9 Summary	19
Chapter 3 Automated Rod Lengthening Measuring Technique on Radiographs.....	21
3.1 Overview	21
3.2 Introduction	21
3.3 Methods	24
3.3.1 Dataset	24
3.3.2 Overview of the Automated Rod Length Measurement	24
3.3.3 Input Radiograph Preprocessing	26
3.3.3.1 Filters.....	26
3.3.3.2 Annotation and Augmentation of Training Dataset	27
3.3.4 Detectron2 with Mask RCNN Models	28
3.3.4.1 Backbone of the Models	28
3.3.4.2 Model Optimization.....	29
3.3.4.3 Evaluation Metrics.....	31

3.4 Result	33
3.5 Discussion	37
Chapter 4 Automated Method for Growing Rod Length Measurement on Ultrasound.....	39
4.1 Overview	39
4.2 Introduction	39
4.3 Method	41
4.3.1 Participants Selection	41
4.3.2 Data Acquisition	42
4.3.3 Overview of Rod length Detection	43
4.3.4 Data Preprocessing Using Filters	43
4.3.5 Data Annotation and Augmentation.....	44
4.3.6 Mask RCNN Models.....	45
4.3.7 Model Evaluation and Statistical Analysis.....	47
4.4 Result and discussion	48
4.4.1 Results	48
4.4.2 Limitation.....	53
4.4.3 Advantages of using US.....	53
4.4.4 Disadvantages of using US	54
4.4.5 Effect of Image Quality.....	54
4.5 Conclusion and Contribution	56
Chapter 5 Conclusion and Future Recommendations.....	57
5.1 Summary	57
5.2 Future Recommendations	58
Reference	60
Appendix	70

List of Figures

Figure 2.1 Radiograph of an EOS patient, Adapted from [30].....	7
Figure 2.2 (a) A radiograph of VEPTR growing rod for EOS treatment [15], (b) A traditional growing rod [35] , and (c) A MCGR for EOS treatment [36]	9
Figure 3.1 (a) shows an actual titanium rod in practice, (b) represents how a radiograph of an EOS patient looks like, and (c) is showing how the clinicians hold the remote controller at the back of the patient while doing the process	23
Figure 3.2 Overview of the process.....	25
Figure 3.3 Visual representation of different parts of the rod in radiograph. (a) shows the whole radiograph, (b) is the ‘58mm’ and ‘head-piece’ segment of the rod after detecting the rod, (c) detected ‘head-piece’ and (d) detected ‘58mm’ portion	26
Figure 3.4 (a) A raw radiograph and (b) The radiograph after passing through the combination of filter	27
Figure 3.5 Mask RCNN structure with RestNet-50 and FPN	29
Figure 3.6 Training loss and validation loss plots of a ‘rod’ model	33
Figure 3.7 Sequential stages from raw image to detecting and measuring the desired length	34
Figure 3.8 Data points from the 56 adjustments.....	35
Figure 3.9 Correlation of AI and Manual measurement.....	36
Figure 3.10 Bland Altman analysis showing agreement between AI and Human measurement for the adjustments	37
Figure 4.1 a) An actual MCGR rod used in this study, b) the portion A-B which gets elongated through the process, and c) the A-B portion in a sonogram	42
Figure 4.2 Overview of the process.....	43
Figure 4.3 a) A raw unprocessed US image, and b) a filtered and processed US image	44
Figure 4.4 a) Exported US image from clinic, b) the expected output of the boundary model, c) cropped version of the first model’s output, and d) expected output rod length from the rod model ..	45
Figure 4.5 Backbone Architecture of the Mask RCNN models	46
Figure 4.6 Loss plots of the rod model.....	48
Figure 4.7 Sequential steps of the overall system to measure rod length.....	49
Figure 4.8 Datapoints of measurements from 20 US test dataset.....	50
Figure 4.9 Bland Altman Analysis between AI and Manual measurements	52
Figure 4.10 Adjusted Bland Altman Analysis between AI and Manual measurements.....	55

List of Tables

Table 3.1 Hyperparameters used in the mask RCNN models with radiographs	30
Table 3.2 AP for the mask RCNN models with radiographs	34
Table 3.3 Statistical comparison of rod measurements between AI and manual approach in radiographs	35
Table 4.1 Hyperparameters used in the Boundary and Rod models and the AP values of both models	49
Table 4.2 Comparison of rod length measurements between AI and R1 (with 20 test points)	51
Table 4.3 Comparison of rod length measurements between AI and manual measurements from both raters	52
Table 4.4 Comparison of rod length measurements between AI and manual methods after the fuzzy images were removed	55

List of Abbreviation

AP	Average Precision
AI	Artificial Intelligence
BA Analysis	Bland Altman Analysis
CNN	Convolutional Neural Network
COCO	Common Objects in Context
CLAHE	Contrast Limited Adaptive Histogram Equalization
DL	Deep Learning
EOS	Early Onset Scoliosis
FPN	Fully Connected Network
ICC	Interclass Correlation Coefficient
IoU	Intersection over Union
LOA	Limit of Agreement
MAD	Mean Absolute Difference
Mask RCNN	Mask Region-Based Convolutional Neural Network
MCGR	Magnetically Controlled Growing Rod
ML	Machine Learning
RCNN	Region-Based Convolutional Neural Network

RPN	Region Proposal Network
SD	Standard Deviation
TGR	Traditional Growing Rod
US Image	Ultrasound Image
VEPTR	Vertical Expandable Prosthetic Titanium Ribs

Chapter 1 Introduction

1.1 Background

Scoliosis is a medical condition characterized by an abnormal lateral curvature of the spinal column with vertebral rotation. The condition exhibits a range of severity levels and has a global prevalence of approximately 2-3% [1]. If the degree of lateral spinal curvature, known as the Cobb angle, in the patient exceeds 10 degrees, it is indicative of scoliosis [2], [3]. In general, a curve is considered severe when it is over 45 degrees. Early onset scoliosis (EOS) refers to the condition of a child exhibiting a spinal curvature before the age of ten [4]. There are three types of management: a) observation, b) conservative treatments include both exercises and bracing, and c) surgical treatment. Treatment decision depends on the age and the severity of the spinal curvature.

Conservative medical therapies such as casting and bracing are commonly employed in the management of the condition. The utilization of casting or bracing techniques can contribute to the enhancement of flexibility, muscular strength, and posture. The aforementioned procedures have the potential to yield positive outcomes for the spine and enhance overall spinal well-being [5]. When the degree of spine curvature becomes larger, Cobb angle $>45^\circ$, conservative treatment become ineffective and surgical interventions are required. Currently, surgical treatments encompass the Traditional Growing Rod technique (TGR) and the Magnetically Controlled Growing Rod (MCGR) technique [4], [5]. In the absence of surgical intervention, the patient's spinal deformity may progress, leading to further physiological consequences, including cosmetic disfigurement [6].

The discovery of the TGR approach occurred throughout the late 1990s and early 2000s, serving as the initial way for surgical therapy in children with EOS. To accommodate the child's growth, rods along the spine are surgically placed and then gradually extended over time. At each time, operation is performed. The rod extension operation halts or controls the deterioration rate of the curve progression while allowing for continuing growth of the child

[7]. Various types of implants can be utilized, including single rods, dual rods, and Vertical Expandable Prosthetic Titanium Ribs (VEPTR) [7], [8]. Given the higher pace of growth observed in children compared to adolescents, the rod extension procedure is continued at regular intervals until the end of the treatment regimen. Numerous surgical procedures might result in notable physiological repercussions due to their inherent involvement in the creation of wounds [9], [10]. The technique is characterized by its high cost and inherent difficulties.

The efficacy of the VEPTR approach in the treatment of children with EOS who do not exhibit rib abnormalities has been demonstrated [11]. To achieve stabilization, growth, and improvement of lung development in pediatric patients diagnosed with thoracic insufficiency syndrome associated with EOS, the utilization of VEPTR devices has been implemented. These devices consist of expandable rods that are surgically attached to either the ribs or the spine [12]. A study found that people with congenital scoliosis, rib fusion, and expansion thoracostomy with VEPTR implantation exhibited a decrease in hemoglobin levels [13]. The utilization of VEPTR in pediatric patients with EOS has been found to be accompanied by some complications [12], [14], [15], [16].

A cost-effective and non-invasive growing rod technique has been developed as an alternative to conventional growing rods. The new approach uses the MCGR technique [17], [18]. This technique can be controlled externally without requiring surgery. Due to its non-invasive nature, this approach is more cost-effective and presents less challenges for patients. Multiple studies have demonstrated the therapeutic effectiveness [17], [19], and economic feasibility [20] of this intervention. The studies [16], [19], [21] demonstrated that MCGR was better than both TGR and VEPTR. Despite the presence of certain complexities associated with the MCGR approach [22], [23], [24], [25], it still demonstrated to be better in terms of number of complications and the associated cost of treatment when compared to the previously employed invasive growing rod technique [26], [27]. One significant benefit is the reduction of surgeries required. After the installation of the rod, the MCGR technique allows for a higher frequency with safer extension compared to typical growing rod techniques.

Although the rod can be lengthened externally, an immediate assessment of the rod length is needed. A precise estimation of the 'rod-length' of the implanted rod is crucial because the rods

used in the process have a limited length. Once the rod reaches its extension limit, rod replacement surgery may be needed if the child still has growth left.

In each visit, a radiograph is taken before the rod lengthening process, and one is taken right after. Difference between the two rod lengths from these pre and post images provides the required adjustment length to the clinicians. That length helps the clinicians to keep a record of the length of the rod and its extension limit. Until now, the length from the radiograph images has been measured manually, which may lead to human errors in measurement and eventually misjudged the treatment planning. Hence, there is a need to detect the adjustment length accurately to eradicate the existing issues.

Recently, another imaging technique, known as ultrasound (US) has emerged as an alternative in different medical diagnosis. One notable advantage of this approach is the absence of radiation, rendering it particularly advantageous in medical contexts involving women and children. Moreover, the measurement of US image reliability is notably high when utilizing the radiograph rod adjustment measurement technique [28], [29]. Similarly, human measurement errors also exist because the rod lengths are measured manually. The manual approach for conducting these measurements, however, has encountered inherent inaccuracies, resulting in significant clinical complexities.

1.2 Motivation

Clinicians have been actively seeking a precise and automated method to measure rod length modifications accurately from both x-ray and US images in children who have MCGR. The requirement is influenced by several factors.

Firstly, the clinicians experience affects the measurement accuracy. Secondly, the quality of image exhibits significant variation. In certain scenarios, the edge of the rod is difficult to identify. The accurate rod length measurement is difficult to achieve. Thirdly, the unstable body posture during image acquisition may affect the scale or focus of the images. This poses

challenges in maintaining measurement consistency. Finally, the scale of the radiographs, specifically the number of pixels per centimeters may affect the measured results.

In addition, manually measuring the rod length on radiographs and sonograms also takes up clinicians' time. Hence, an automated method is needed. Furthermore, the use of US imaging to eliminate the radiation exposure is a novel image approach. Automatically measuring the rod length on US images is another motivation in this thesis study.

1.3 Objectives

The objectives of this study were:

- a) To develop machine learning (ML) algorithms to automatically extract features from radiographs to assist rod length measurements.
- b) To evaluate the accuracy, reliability, and speed of the developed algorithms on in-vivo radiograph data.
- c) To develop ML algorithms to automatically extract features from sonograms to assist rod length measurements.
- d) To evaluate the accuracy, reliability, and speed of the developed algorithms on in-vivo US data.

1.4 Contribution

This dissertation presents a significant contribution to the field of EOS surgical treatment. New and novel ML algorithms were developed for both radiographs and sonograms to automatically

extract the rod length of MCGR. With providing precise rod length measurements, clinicians can make surgical planning more precisely and optimally.

1.5 Thesis Outline

This thesis is divided into 5 chapters.

Chapter 1 provides a succinct overview of the problem at hand in its introductory section. Subsequently, the rationale for conducting the study and the objectives of the investigation are provided. This chapter also includes the thesis outline.

Chapter 2 provides an overview of the historical context and background of EOS, various treatment modalities and traditional rod length measurement method. Subsequently, different studies in applying artificial intelligence (AI) to assist measurements on radiographs and sonograms for scoliosis and other medical conditions are reported.

In chapter 3, the development of an AI system to automatically extract the rod length of MCGR on radiographs are presented. An in-vivo study has been conducted to determine the accuracy, reliability and the speed of the developed algorithms.

In chapter 4, the development of an AI system to automatically extract the rod length of MCGR on sonograms are provided. The validation was performed by comparing the AI measurements with manual measurements which were conducted by an experienced rater.

In chapter 5, a summary of the thesis is reported. Future work and recommendation are also provided.

Chapter 2 Background and Literature Review

2.1 Overview

In this chapter, sections 2.2 to 2.7 provide a summary on the cost, risks, obstacles, and complications experienced for children with EOS using different surgical treatment methods. Different image modalities which can be used to monitor the post surgical treatment outcomes are reported. Section 2.8 reports some studies that incorporates ML and AI in the field of medical imaging. Especially, this section discusses some of the studies that laid the foundation to choose the ML model and the framework of this study.

2.2 Background of EOS

EOS is a three-dimensional spinal disorder which occurs mostly in children under 10 years old. The lateral curvature of spine which is defined as the Cobb angle is the most common method to diagnose the severity of scoliosis. The Cobb angle is made by drawing lines along the top and bottom of the vertebrae that exhibit the most tilt in the curve. The angle created by the point of intersection of these two lines is the Cobb angle which is usually measured on posteroanterior radiographs. This angle aids clinicians in evaluating the extent of the spinal curvature, deciding treatment, and tracking any progression over time.

EOS can result from a variety of causes. Congenital abnormalities can lead to malformations that affect the growth of the spine. Neuromuscular conditions like cerebral palsy and muscular dystrophy can contribute to the development of EOS. Idiopathic scoliosis refers to cases where the actual cause is often unknown. It can result from a combination of genetic, environmental, and developmental factors. Figure 2.1 depicts the physical appearance of a severe conditioned EOS patient in radiographs. The Cobb angle of the curve is around 81° which is significantly over the surgical treatment threshold of 45° .



Figure 2.1 Radiograph of an EOS patient, Adapted from [30]

Early diagnosis is very important to manage EOS. This way it is possible to work on preventing progression of the Cobb angle and minimize the impact of EOS. A study [31] reported a multidisciplinary approach of management of EOS. The management team includes orthopedic surgeons, pediatricians, pulmonologists, physical therapists, and other specialists. Children who have mild EOS with Cobb angle $\leq 25^\circ$ are usually put under observation to monitor regularly until the curvature changes to moderate. When the spinal curvature becomes moderate in between 25° to 45° , conservative treatment like casting and bracing are used to halt the further curvature progression. Unfortunately, if the conservative treatment is ineffective, the spinal curvature exceeds 45° , surgical treatment is the final option. Traditional growing rod surgery to treat EOS was introduced in 1990.

2.3 TGR Complications for EOS

When the TGR was introduced, it came as a blessing for children with EOS were happy since the body growth could be compensated by adjusting the rod length. The basic growing rods comprise of a rod connected to the spine and anchors (hooks, screws, or wires) to the vertebrae. The rod might be composed of stainless steel, titanium alloy, or alternative materials. Due to

the rapid growth in young patients, it is necessary for the rods to adjust and expand along with the child's spine. Therefore, this method requires regular spinal surgeries. During each surgery, the back of the spine is opened. The rod is manually adjusted using a wrench or equivalent equipment based on specific requirements and factors such as age, growth pattern, and curve progression. Due to the inclusion of operations at each visit, the technique exhibited a significantly elevated incidence of post-surgical complications. Numerous studies have been conducted on this topic[8], [32], [33].

Initially, it is important to acknowledge that the use of growing rod approaches presents certain challenges that necessitate solution. Despite the observed efficacy of rod lengthening in achieving desired outcomes, empirical investigations have revealed an associated escalation in the likelihood of postoperative problems following the implementation of this technique in growing rod surgery. Given the frequency of the procedure and the likelihood of post-operative problems, the maintenance of this technique poses challenges for both children and their parents. The patients endure significant pain and stress due to these wounds and complications. Managing the cost and stress during this process is often challenging for the family.

In a study conducted by Akbarnia et al., it was suggested that the rod length adjustment surgery should be operated on a biannual basis [8]. According to a study conducted by Thompson et al., it was reported that dual rods exhibited enhanced strength and stability when compared to a single rod operation. This characteristic lead to superior initial correction and maintenance of correction [34].

Although utilization of growing rods helps to minimize the spinal curvature of the patient and provide stability to the spine, complications may still occur. A study conducted by Sankar et al. (2010) addressed the challenges encountered by EOS patients who underwent implantation of several types and styles of growing rods [32]. The patients in this study were divided into three groups: Standard Growing Rods, Hybrid Growing Rods, and Vertical Expandable Prosthetic Titanium Rib (VEPTR). Standard growing rods are the basic growing rods mentioned above. Hybrid growing rods were constructed using traditional spinal implants with hooks functioning as rib anchors. The VEPTR was used mostly on the EOS patients who had

thoracic insufficiency. VEPTR is not directly attached to the spine, unlike conventional growing rods. It extends over the chest wall and ribs, offering support and enabling regulated expansion of the thoracic chamber. Figure 2.2 presents radiographs of EOS patients undergoing VEPTR, TGR, and MCGR treatment approach.

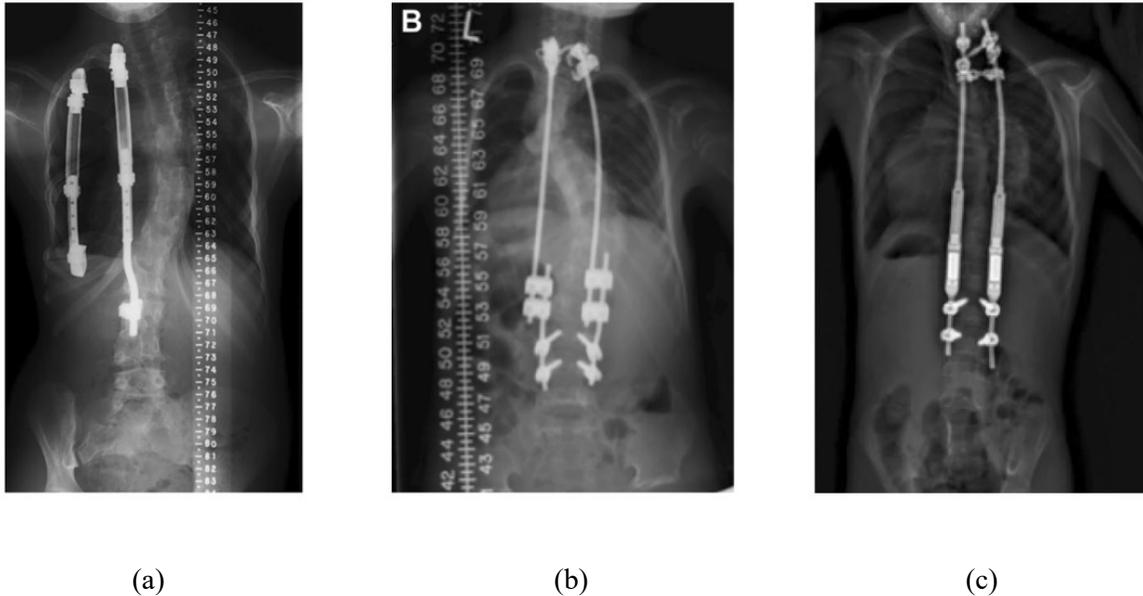


Figure 2.2 (a) A radiograph of VEPTR growing rod for EOS treatment [15], (b) A traditional growing rod [35] , and (c) A MCGR for EOS treatment [36]

The mean age at the time of the initial implantation was 4.8 years, whereas the mean duration of follow-up was 51 months, ranging from 24 to 117 months. The reported study comprised of a sample size of 36 participants, but 26 participants (constituting 72% of the patient cohort) experienced the need for at least one unforeseen surgical procedure or encountered a notable complication over the entire duration of the follow-up period. A total of 72 unplanned surgeries were required. Out of the total sample size, 18 cases needed revisions due to rod fracture, 31 cases required operations to address loose implants, and 18 cases necessitated irrigation and debridement for the treatment of infections. Two young individuals required to undergo the extraction of an implanted rod because of the neurological complications they experienced. The patients who received standard growing rods exhibited an average of 2.3 problems per

patient. In contrast, the patients who received hybrid growing rods experienced a lower complication rate of 0.86 per patient, while VEPTR patients had a little higher complication rate of 2.37 per patient. More studies have been reported that using standard growing rods for the treatment of EOS [10], [14], [15] showed a large number of complications.

The risk factors associated with problems in children with EOS were published by Watanabe et al. in 2013, specifically focusing on the use of TGRs [33]. This study involved a total of 88 participants, with ages ranging from 1.5 to 9.8 years old. The overall follow-up time spanned from 2 to 12 years, with a mean duration of 3.9 ± 2.6 years. The intervals between the rod-lengthening treatments had an average duration of 6.6 ± 2.1 months, ranging from 5 to 12.5 months. Multiple factors associated with complication were considered throughout the planning phase of the investigation. Variables such as age, gender, the number of lengthening rods utilized, and the Cobb angles of the thoracic and lumbar curves were among the factors considered. Approximately 57% of the patients experienced problems. Also, each patient might have multiple surgeries which ended up with 538 surgeries. Among a total of 538 surgeries, it was observed that 119 of them experienced complications accounting for approximately 22% of the total. Among the cases examined, a significant proportion of 72% were found to have experienced implant failures. This was followed by 19 instances of infections, accounting for 16% of the total cases. Additionally, there were 3 reported cases of neurological diseases and 11 miscellaneous abnormalities. The reported study also employed Kaplan-Meier analysis and demonstrated that an increase in complication rate was associated with the number of rod lengthening surgeries. Based on the findings of this study, three significant risk factors associated with problems. These factors include undergoing six or more surgeries for rod-lengthening, experiencing an increase of 20° Cobb angle when comparing with the preoperative thoracic curve, and encountering a 20° increment in preoperative thoracic kyphosis. The above studies provided the foundation for exploring an improved surgical method to treat children with EOS.

2.4 Magnetically Controlled Growing Rod (MCGR)

To resolve the several existing issues of TGR, researchers and surgeons aimed to develop newer and better approaches. An important achievement in this area was the development of MCGR, designed to reduce the necessity for frequent surgeries and minimize the related difficulties. This non-invasive approach for spinal development modulation was developed in the late 2000s. The first instance of its use was documented in 2009 on a patient with scoliosis. The TGR approach includes surgical procedures during the initial implantation phase and any subsequent replacements of the developing rod, if necessary.

MCGR systems are comprised of a magnetic mechanism connecting two telescopic rods. An external controller is positioned close to the patient's spine during every visit, ensuring that the controller is in alignment with the internal magnet inside the rod. An operator is then set up the desired length adjustment that would like to achieve. Once the controller provides the electromagnetic field to the internal magnet, the rods will extend according to the setup. By extending the headpiece in this manner, progressive spinal growth is facilitated. The precise regulation of the magnetic force by the external control unit guarantees gradual and controlled elongation, which is of utmost importance to minimize patient distress and optimize spinal alignment.

MCGR technology makes it possible to external lengthening modifications without invasive surgery. This novel method enables non-invasive lengthening treatments to be performed in an outpatient clinic, providing increased convenience for patients and caregivers.

2.5 Cost Analysis and Complication Rate Analysis Between TGRs and MCGRs

While the MCGR approach addresses several significant shortcomings of the TGR approach, it is not a complication-free approach either. Numerous investigations and study endeavors were conducted to ascertain the extent of challenges associated with these two methods [16],

[17], [21], [23], [37]. The following studies reported a comparison of the costs and complications associated with the MCGR and TGR techniques.

In 2012, a pioneering study was undertaken by Cheung et al., wherein five patients underwent the installation of MCGR [17]. These individuals were diagnosed with EOS and participated in monthly interventions for a consecutive period of two years. This study provided evidence supporting the non-invasive nature and increased frequency of the rod adjustments unaffacting patients' life. The average Cobb angle of the participants was 67° prior to the commencement of the trial, and after a duration of two years, the average of Cobb angle was 29°. The study revealed that patients did not experience any pain, and no complications were noted. Subsequently, other investigations were conducted to examine several facets of MCGR in the context of EOS treatment.

Peiro-Garcia et al. conducted a study which revealed that the MCGR method exhibited a significantly lower incidence of early complications compared to VEPTR [16]. There were 35 participants in the study. 15 of them were undergoing MCGR approach, and 20 of them were under VEPTR approach. After a 2-year research, it was found that VEPTR had significantly higher complication rates (approximately 5 times) and reoperation rates (around 4 times) compared to MCGR.

A comparative analysis of the surgical procedures required for TGRs and MCGRs based on their research findings was conducted [23]. The duration of the trial spanned a period of two years. The selection of participants for this study was conducted with careful consideration of multiple factors. Besides the patients should be under 10 years old, the inclusion criteria included: a) the Cobb angle had to equal to or exceed 30 degrees, b) the body height between T1-T12 (Thoracic 1-12) had to less than 22 cm, and c) the patient had no history of prior spine surgery. Out of the total sample size of 17 potential participants, 12 participants met the inclusion criteria of the study. In which, five participants were male, while the remaining seven were female. The growth patterns of both groups were found to be similar. Furthermore, the primary adjustment made in both approaches were same. During that period, the MCGR patients had 57 fewer surgeries when compared to the TGRs. The total number of open

procedures conducted for TGRs was 73, with a specific subset of 56 surgeries dedicated solely to the purpose of lengthening. In contrast, a mere 16 surgeries were necessitated for MCGRs, whereas a total of 137 non-invasive lengthening procedures were conducted. Based on the analysis of the available evidence, the authors reported that MCGR was superior to TGR due to its reduced complexity, decreased occurrence of unforeseen procedures, and more frequent utilization of non-invasive lengthening techniques.

Apart from the number of complications and surgeries, the consideration of cost has significance in the selection of the way of treatment. Several studies have been conducted to compare the cost of TGR with MCGR. The cost neutrality analysis of a particular study in 2016 revealed a treatment duration of six years [21]. On this investigation, Polly et al. provided further evidence supporting the notion that MCGRs exhibit a higher initial cost compared to TGRs during the initial implantation phase. The direct medical expenses were compared within the context of the integrated healthcare delivery system in the United States. A study including a sample size of 1,000 patients over a span of 6 years demonstrated that the utilization of MCGR resulted in a statistically significant reduction of 270 surgical infections and 197 revisions compared to TGR. At the conclusion of the six-year research period, the TGR and the MCGR incurred direct expenditures amounting to \$149,234 and \$149,295, respectively. The numerical data presented in this study substantiates the assertion of cost neutrality. However, a lingering inquiry remains: what would be the outcome if a patient were to require a treatment duration exceeding six years?

To answer this question, Wong et al. conducted a cost-analysis study and indicated that the MCGR method exhibited higher precision and cost-effectiveness when compared to the TGR for the treatment of EOS [37]. This study aimed to compare various growth rod treatments employed for the treatment of EOS, yielding commendable results. A hypothetical case analysis was used. Assuming a 5-year-old patient diagnosed with EOS. During the experiment, two distinct types of MCGR and TGR were utilized. There were two kinds in each, comprising four different types of growing rods in total. In general, the initial cost of implementing MCGRs was higher when compared to TGRs. However, as the MCGR treatment was more advance and considering a 4-year cycle was needed and completed, the overall cost of the

MCGR was comparable to that of the TGR. After that, the cost of MCGR went down to than that of TGR. The findings of this study clearly demonstrated that the use of MCGR has a superior cost-effectiveness when compared to TGR.

2.6 Measurement of the MCGR Rod Adjustment

As described in the section 1.2, clinicians currently measure the rod length manually at the scoliosis clinics. The measurements are performed on radiographs. However, due to the manual measurement, there are some factors that affect the accuracy of the rod length measurement. Diverse level of competence among the raters, poor radiograph quality, excessive patient movement, and resolution discrepancies within the images are the reasons behind the inaccuracies. This inability to measure the rod length accurately eventually hampers proper treatment planning, unplanned surgeries, complications and so on. Hence, clinicians have been looking for a universal and accurate approach to perform this measurement.

2.7 Drawbacks of Radiography and Emerging US for EOS Applications

To diagnose and monitor the progression of scoliosis, radiographs are taken at every clinic visit. However, taking a radiograph means exposing a patient to ionizing radiation. Children with EOS are diagnosed under the age of 10. Several studies have demonstrated that repeated exposure of ionizing radiation to children can significantly facilitate the proliferation of cancer cells in them [38]. Various types of cancer, such as leukaemia, thyroid cancer, breast cancer, brain cancer, skin cancer, and other malignancies, have the potential to develop later in their lives. A study which was conducted in 1962 was designed to investigate the relationship between prenatal exposure to X-rays and the occurrence of cancer in children [39]. A total of 734,243 children born during the period from 1947 to 1954 which were recruited from 37 hospitals were included in the analysis. Among all the children that participated, 556 cancer-

related deaths observed, but 85 cases (15.3%) were associated with exposure to X-rays during intrauterine development. The study's conclusion indicated that children who received X-ray examinations had a 40% higher cancer death rate compared to those who did not have X-ray examinations. Findings indicate that cancer death rates were estimated to be around 40% higher among individuals who underwent X-ray examinations. Hence, due to the radiation exposure MCGR treatment approach had some limitations and drawbacks for which the researcher community was looking for alternatives.

The US imaging approach is an innovative imaging modality that utilizes high-frequency sound waves to generate dynamic visual representations of the inside anatomical structures of the human body. The primary advantage of this approach lies in its complete absence of radiation emission. As previously mentioned, the present utilization of this intervention extends to the management of pregnancy, with the aim of preventing life-threatening complications and treating fusions. Moreover, it demonstrates efficacy in seeing and analyzing soft tissues. Typically, practitioners specializing in the field of Scoliosis choose to utilize US techniques as an alternative to radiography, hence eliminating the necessity for patients to undergo X-ray procedures. In 2014, a study was conducted to examine if US could be used to image the MCGR in children with EOS [28]. The study recruited six participants (4F, 2M), aged between 8 to 16 years old, who were undergoing MCGR treatment for their scoliosis. From that study, researchers were able to measure the rod length manually. A high level of reliability and accuracy were reported, which lead the authors to report US as an alternative of radiography.

However, one issue faced by the US images is the clarity of the image. Usually, a curve object which scans using US technique may have fuzziness at the curve area. Hence, manually measuring the length of the rod in an unclear image is highly dependent on the experience of the rater. Due to the fuzziness and the need of automatic measurement, AI approach is suggested.

2.8 Literature of Applying AI in Medical Imaging

In recent years, AI has emerged as a significant facilitator in enhancing the quality of human existence by streamlining and augmenting several aspects of daily life. ML falls within the domain of AI. Numerous algorithms have been created within the subject of ML research to effectively address various objectives [40], [41]. Due to the varying requirements and advantages associated with different types of algorithms, their utilization is limited to specific applications. Deep Learning (DL) has significantly transformed the field of image analysis in a wide range of domains [42], [43], [44], [45], [46]. It has demonstrated impressive effectiveness in deciphering complex patterns within medical images, enabling accurate diagnosis and treatment planning. Moreover, DL, along with ML techniques, has been successfully applied to address various image-related challenges in diverse applications beyond the medical domain [47], [48]. Especially, a Convolutional Neural Network (CNN), an ML technique, and its potential applications in the field of medical imaging were reviewed.

2.8.1 CNN in Medical Images

As mentioned in the section 2.8, different ML algorithms include diverse attributes that make them suitable for specific applications. CNNs are commonly employed in the field of computer vision for the purposes of object detection and segmentation [49], [50], [51], [52]. Multiple CNN architectures have been identified for the purpose of computer vision tasks. Several examples of CNNs commonly used in computer vision tasks include the Regional Convolutional Neural Network (RCNN), Fast RCNN, Faster RCNN, Mask RCNN, YOLO (You Only Look Once), and SSD (Single Shot Multibox Detector). Some of the breakthrough studies [53], [54], [55], [56] that have laid the foundation for the use of CNNs have been reviewed.

In the year 2022, Zhang et al. introduced a novel approach for the placement of a Spine model using an RCNN-based methodology [53]. The authors of the study used radiographs as the primary data source. Following a comprehensive discussion on the various models including a) RCNN, b) Fast RCNN, c) Faster RCNN, and d) Mask RCNN, the decision was made to employ Mask RCNN for the purpose of conducting detection and segmentation operations. The use of a feature pyramid network (FPN) in Mask RCNN serves as an enhancement to Faster RCNN, enabling the mitigation of region transformation limitations, enhancement of algorithm performance, and effective utilization of pixel-level position information from training data image labels. Furthermore, the proposed approach replaces Region of Interest pooling (RoIPool) with RoI alignment, hence mitigating the loss of data during the training of the model. The utilization of a mask surrounding the outcome of the Region of Interest (RoI) alignment enables its application in object prediction and segmentation, specifically in conjunction with a bounding box. The training process utilized two graphics processing units (GPUs), with the ResNet-50 architecture serving as the foundational basis for the model. The models exhibited outstanding accuracy, with an average accuracy of 97.4% for detection boxes and 96.8% for segmentation.

Mask RCNN had also been employed in the context of patients with scoliosis [54]. In the year 2021, Cui et al. conducted a study to employ CNN to quantify the Cobb angle in patients with scoliosis [55]. The U-net framework was utilized in this study for the purpose of object detection, yielding significant advantages. Mask RCNN was employed for the purpose of segmenting and labeling the different components that required detection. In this study, 609 images were original obtained; with applying augmentation technique to increase the size of training set, a total number of 5772 images were used. Given that the validation and training loss graphs exhibited a decline in performance after the 4000 data point. The decision was made to limit the number of epochs or iterations to 4000. The researchers reported that the maximum error on the Cobb angle measurement was 4.7° , while the average error was 3.0° . The analysis yielded a significant finding in the form of the Symmetric Mean Absolute Percentage Error (SMAPE), which was calculated to be 21.7%.

Another study in applying Deep Neural Network (DNN) for spinal applications was reported in [56]. The investigation analyzed a total of 40 radiographs in spine phantom models and 65 radiographs (14M, 51F) in in-vivo data. The average age of the participants was 12.5 ± 3.6 years old. The results reported that the ML method to automatically measure the Cobb angle had a high accuracy measurement (mean absolute difference $\leq 5^\circ$) and excellent reliability (both intra and inter-observer ICC > 0.9). These studies established the groundwork for the utilization of Mask RCNN in the context of object detection for medical images.

2.8.2 Detectron2

In the year 2018, the Facebook AI Research (FAIR) community introduced a framework named Detectron, that has the capability to perform object detection and instance segmentation tasks. In due course, it garnered significant acclaim among the computer vision community, emerging as a widely adopted paradigm. In the latter part of 2019, FAIR unveiled Detectron2 as the successor to its previous iteration. The revised version of the framework acknowledged and sought to overcome the constraints identified in the previous version, while also providing a more user-friendly and streamlined approach. Detectron2 possesses the notable capability of accommodating unique datasets and models, in addition to its ability to incorporate many widely used pre-trained models for computer vision tasks. The approach exhibits superior efficiency compared to traditional algorithms, enabling rapid processing. Moreover, it offers a robust codebase that is conducive to conducting object detection research with a high level of accuracy and precision.

A study conducted in 2023 reported the utilization of Detectron2 for the purposes of identifying objects and image manipulation. The COCO dataset served as the fundamental dataset for the development of models. In this study [57], a pre-existing model with a default architecture and Mask RCNN was utilized. Prior to being inputted into the model, the images underwent preprocessing techniques such as Gaussian blurring, Median blurring, sharpening with a 2D convolution kernel. Overall, they achieved notable success in developing exemplary models for the tasks of instance segmentation and object detection.

There is existing research that assesses the effectiveness and precision of object detection algorithms by utilizing a dataset containing two distinct weed species [58]. The two species are Convolvulus and Phalaris Paradoxa. Four separate models were employed to evaluate the detection, then subject to comparison according to a predetermined set of criteria. The models encompassed in this study comprised EfficientDet, Faster R-CNN, YOLOv5, and Detectron2. In all cases, the COCO dataset was utilized, employing a range of Average Precision (AP) values across various Intersection over Union (IoU) thresholds. The study utilized several augmentation techniques. Upon careful examination of the results, it was shown that Detectron2 exhibited the highest AP among all the applied methodologies, achieving an impressive 97%. Hence, Detectron2 became familiar for its utility in the construction of object detection algorithms that achieve enhanced levels of accuracy.

Detectron2 has also been employed for the purpose of object detection and segmentation in the domain of medical imaging. A recent study published in 2021 employed the detectron2 architecture in conjunction with Mask RCNN to effectively detect and separate breast cancer lesions from medical images [59]. The development of the model involved the utilization of the public dataset INbreast. The segmentation of the lesion was performed utilizing the Mask RCNN architecture, which furthermore yielded the bounding box. Mammogram images were employed in this study. A total of 115 individuals underwent 410 full digital mammograms. 90% of them were allocated for the purpose of model training and validation. The remaining 10% were specifically reserved for model testing. To leverage the superior performance of RCNN bounding boxes and masks, the researchers choose to utilize pre-trained models, specifically R101-FPN and X101-FPN. The suggested strategy exhibited a high level of accuracy, with a value of 95.87%. Additionally, the F1 score, a metric used to evaluate the model's performance, reached 81.05.

2.9 Summary

In summary, MCGR is the most used surgical tools for the treatment of EOS. Due to its cost-effectiveness, minimal complications, and non-invasive nature, MCGR has become an

extensively utilized method for treating EOS. To measure the rod length changes on both radiographs and sonograms, clinicians have suggested to use automatic methods to eliminate human measurement errors, save time, and increase the reliability. Among many ML methods, Mask RCNN models and Detectron2 may have a potential to develop a model to automatically measure the rod length.

Chapter 3 Automated Rod Lengthening Measuring Technique on Radiographs

3.1 Overview

This chapter provides a detailed report regarding the development of the automated system for measuring the rod length of MCGR on radiographs. An in-vivo data study was conducted to determine the accuracy, reliability, and speed of the developed ML algorithms.

The materials in this chapter are mainly exported from the manuscript which has been submitted to the Journal of Medical & Biological Engineering and Computing in Feb 2024.

M. H. Kabir, M. Reformat, S. S. Hryniuk, K. Stampe, and E. Lou, "Validity of Machine Learning Algorithms for Automatically Extract Growing Rod Length on Radiographs in Children with Early Onset Scoliosis", Medical & Biological Engineering & Computing.

3.2 Introduction

EOS refers to a condition characterized by the presence of a lateral curvature of the spine exceeding 10° in children who are under 10 years old. Surgical intervention is recommended when the degree of curvature exceeds 60° [60] or shows a significant progression after 45° . Surgical intervention is the most effective way to stop and reduce spinal curvature; however, if families decline to perform the spinal surgeries, the deformity may subsequently lead to other physiological consequences, including the manifestation of cosmetic disfigurement, as well as respiratory inefficiencies [6], [12], [61]. The TGR technique was developed to address the gradual correction needed for spinal curvature in children with EOS [8], [62]. According to Akbarnia et al. [62], it was recommended to do the rod adjustment process should be performed biannually to account for the growth rate disparity between children and adolescents. However,

each surgical procedure leads to the formation of wounds and scars, and a substantial number of surgeries are associated with notable immediate and post-surgical complications [10], [34], [63], [64]. In addition, spinal surgery is a expensive and rigorous procedure [20]; hence an excessive number of procedures is unfavorable. Therefore, a new non-invasive method known as the MCGR was developed and aimed to minimize the number of surgeries [18], [24], [27]. Rod length adjustment of the MCGR can be performed externally. To increase the length of the rod, clinicians can set a desired adjustment on an external remote controller. It has been reported that the MCGR method was an efficient and cost-effective approach to treat children with severe EOS [21], [22], [23], [32], [34].

To measure the rod length, clinicians perform the measurements on the pre-rod and post-rod adjustment radiographs. The manual measurement process can introduce human measurement errors and is time-consuming. The imprecise measurement affects future surgical planning as the rod has maximum distraction capacity. Figure 3.1 shows a titanium growing rod, a radiograph with MCGR installed in an EOS child, and an external controller machine is held at the back of an EOS child.

ML, a subset of AI, is defined to use data and algorithms to replicate the learning processes observed in humans. Among many ML algorithms, the DL technique is the most widely adopted for image-related applications [42], [43], [44], [45], [46]. CNN under the DL are extensively employed in the field of medical imaging for object detection and image segmentation [50], [51], [55], [56]. From the literature, the Regional Convolutional Neural Network (RCNN) had been utilized in spinal applications [53], [54]. Liu et al. employed a Faster RCNN model to perform automated measurement of spinal curvature on US spine images through the detection of vertebral laminae [65]. Pham et al. also employed two CNN models, namely inception ResNet v2 and mobileNet v2, to measure the severity of hip displacement automatically in children with cerebral palsy [52]. The Mask RCNN represents an enhanced version of the RCNN, specifically designed to detect and segment objects present in an image, and hence generating masks [59], [66], [67].

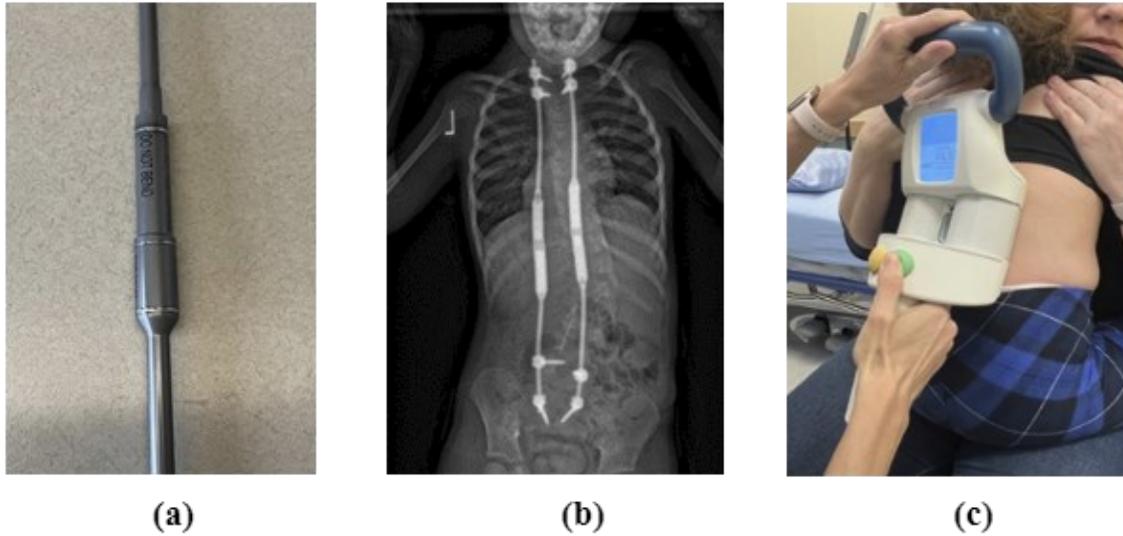


Figure 3.1 (a) shows an actual titanium rod in practice, (b) represents how a radiograph of an EOS patient looks like, and (c) is showing how the clinicians hold the remote controller at the back of the patient while doing the process

Detectron2 is also a DL method developed by Facebook AI Research (FAIR) that encompasses many object identification algorithms, including the Mask RCNN. The utilization of Detectron2 was extensively applied in several studies for object detection and segmentation [57], [58], [68] which provided excellent results. Ammar et al. introduced a computerized approach to detect the existence of COVID-19 on chest radiographs. Their method also employed Detectron2 and Faster RCNN [69].

The objective of this study was to report the development and validation of a Detectron2 Mask RCNN algorithms to automatically detect the MCGR rod length before and after the rod adjustment on radiographs.

3.3 Methods

3.3.1 Dataset

This study received the local health research ethics board approval (Pro00109499). Prior the participation, parental or guardian consents were duly signed. A cohort of 24 children diagnosed with EOS underwent spinal radiography, resulting in a total of 449 radiographs. The inclusion criteria were children with EOS who a) had MCGR rod instrumented, and b) required rod length adjustment at the clinic. Each participant would have two radiographs (pre-rod and post-rod adjustments), but 1 subject was accidentally taking 3 radiographs, at the clinic visit. Among the 24 participants, only 1 participant had a single rod and the rest had 2 rods installed.

3.3.2 Overview of the Automated Rod Length Measurement

Figure 3.2 depicts a flowchart illustrating the sequential steps commencing from radiograph input until reporting the automated measurement of rod length.

Three ML models were developed to detect and segment features. Before inputting radiographs into the model, the original radiograph underwent preprocessing to enhance the image quality. The first model was to detect the implanted rods. The subsequent model identified a designated segment of the MCGR rod, measuring 58mm in length, that served the calibration purpose. The head-piece of the rod within the MCGR rod was discovered by the third model. Figure 3.3 shows all the different portions of a MCGR rod that need to be detected in different ML models in the system.

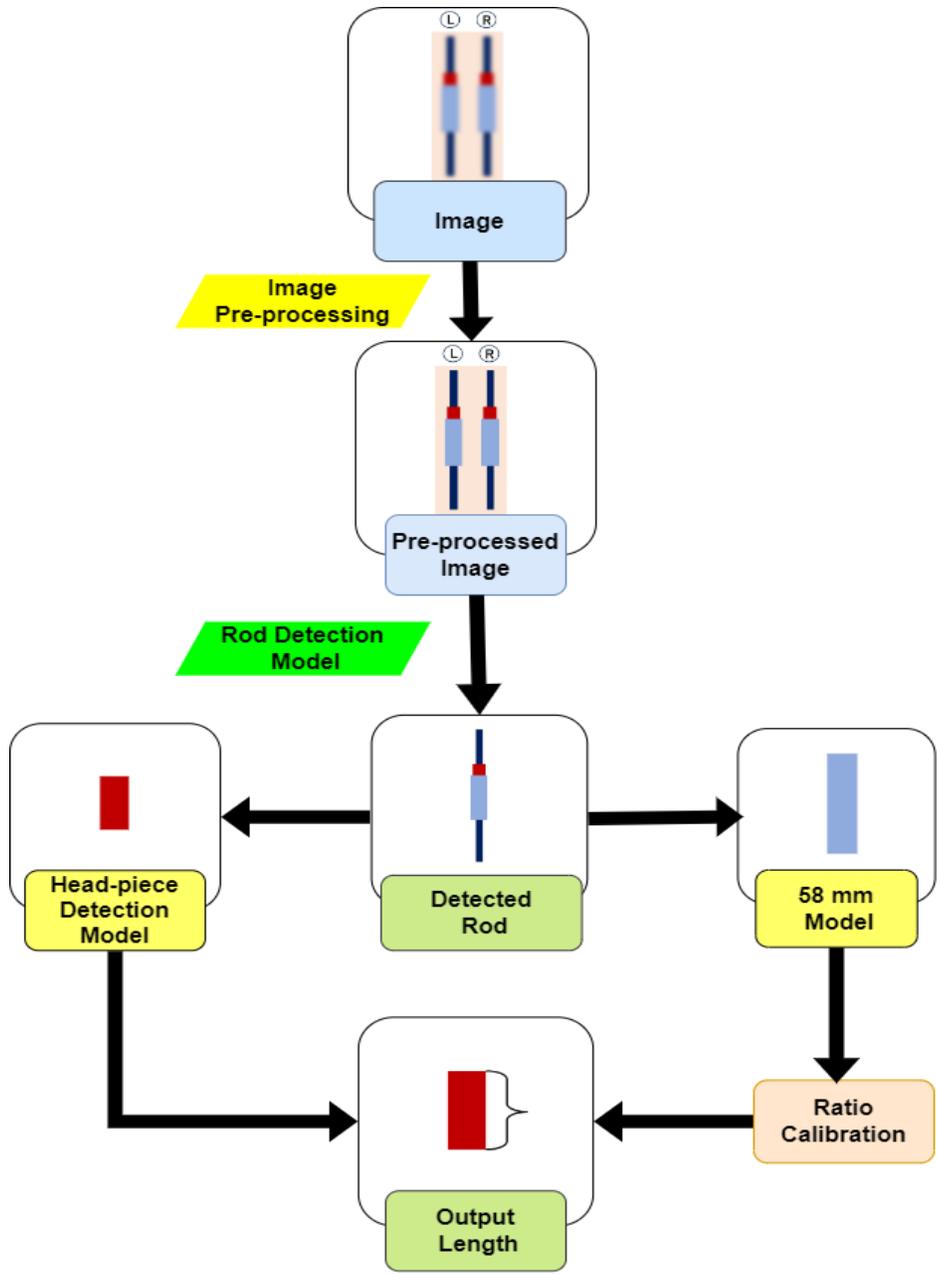


Figure 3.2 Overview of the process

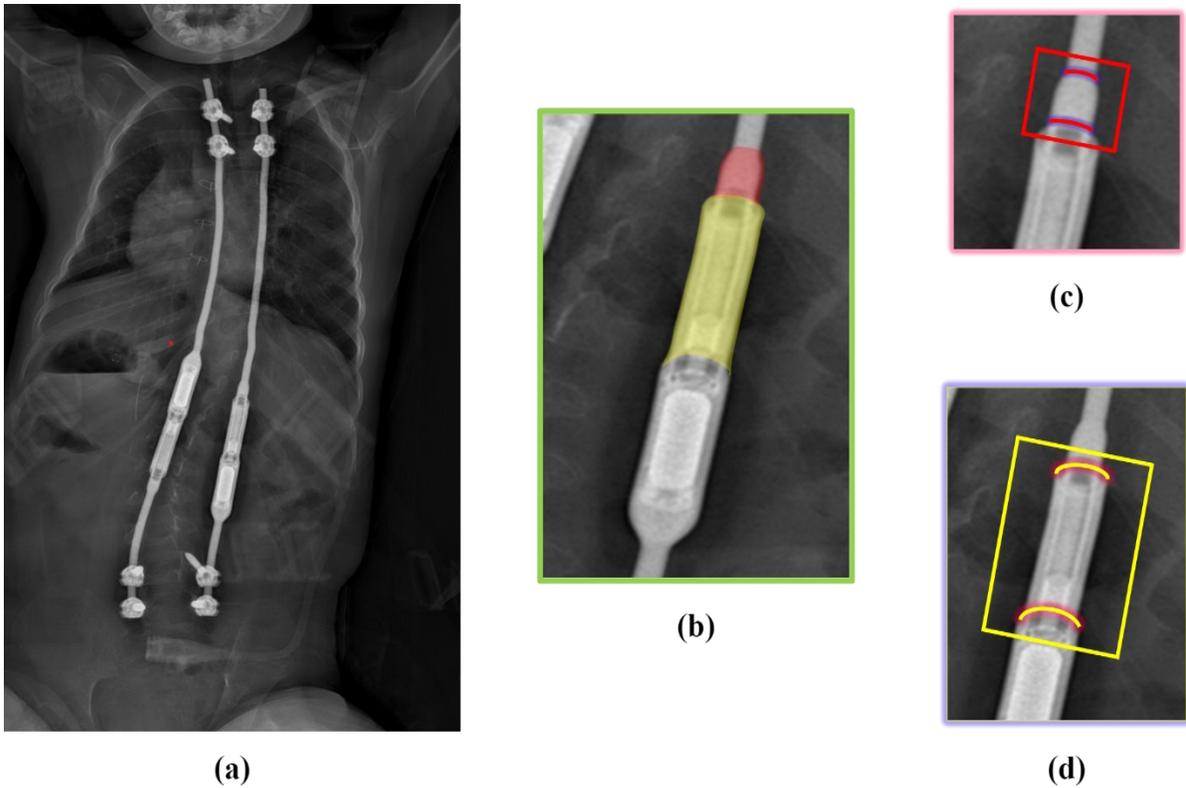


Figure 3.3 Visual representation of different parts of the rod in radiograph. (a) shows the whole radiograph, (b) is the ‘58mm’ and ‘head-piece’ segment of the rod after detecting the rod, (c) detected ‘head-piece’ and (d) detected ‘58mm’ portion

3.3.3 Input Radiograph Preprocessing

3.3.3.1 Filters

After a spinal radiograph was input into the program, two filters were applied in the preprocessing stage. The first filter was called the unsharp mask and the second filter was Contrast Limited Adaptive Histogram Equalization (CLAHE). The unsharp mask filter was to enhance the edges and details in the images by emphasizing high-frequency components [70], [71], [72] and the CLAHE filter was used to improve the images quality by enhancing their contrast [73], [74]. To increase the sharpness along the edges of images, two parameters, “radius” and “amount”, could be controlled. The radius defined the size of the local region

around each pixel to apply the mask. The amount was to define the degree of sharpening impact. Regarding the contrast, the adjustable parameters in the CLAHE filter include 'cliplimit' and 'tilegridsize'. The cliplimit is the contrast limit for localized changes in contrast. The tilegridsize is to determine the size of the contextual zone for histogram equalization. Figure 3.4 displays the (a) original and (b) filtered radiograph of the same participant.

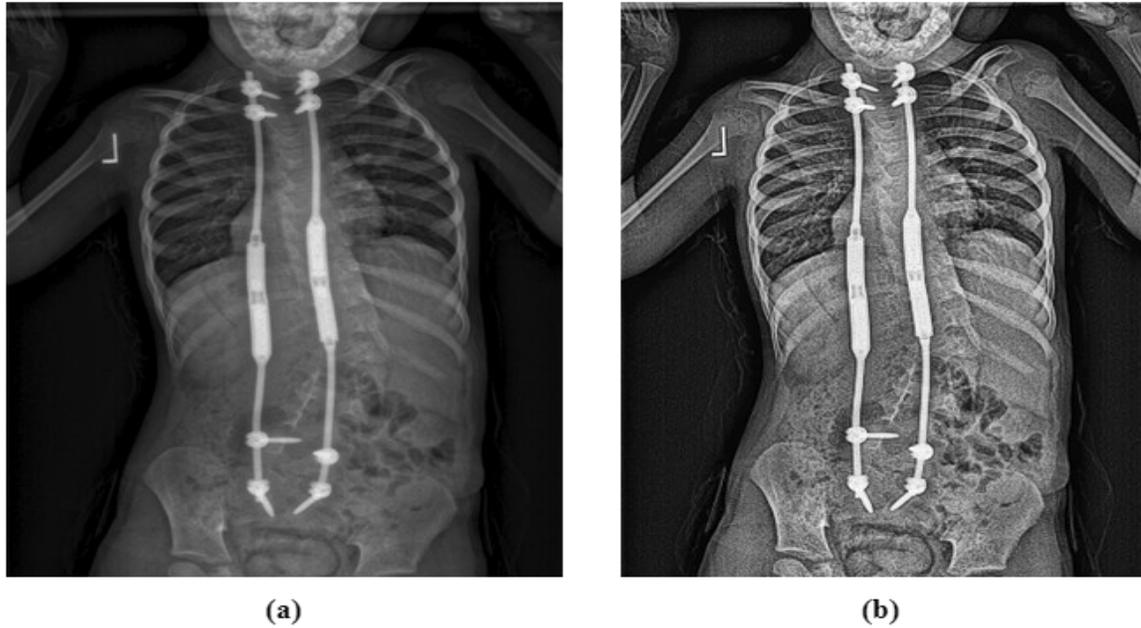


Figure 3.4 (a) A raw radiograph and (b) The radiograph after passing through the combination of filter

3.3.3.2 Annotation and Augmentation of Training Dataset

Among 449 radiographs, only 387 radiographs were used for models' development and validation. For the rod model, the 387 radiographs were divided into 327 for training, 50 for validation and 10 for testing. After the preprocessing of the input radiographs, the training dataset underwent annotation and feature labeling. To label the rod, the outline of each rod was identified by putting 15-20 points along the edge of each individual rod. This labeled area was thereafter designated as the "rod" region. The annotation task was facilitated by employing a software application called LabelMe [75]. Subsequently, the second and third models were

annotated using the outputs obtained from the first model, employing the similar technique, annotating the 58 mm rod length region and the head-piece region. However, on each radiograph, only 1 rod was annotated. Hence, 387 rods were used for the rod model development. For the remaining 2 models, since 9 of the 387 radiographs were failed to be used, the last 2 models only had 378 segmented datasets. Hence, 313 was used for the models' training, 50 for models' validation, and 15 for models' testing.

Historically, DL models have conventionally necessitated a considerable volume of training data. The research conducted by Shorten et al. [76] examined the effects of picture data augmentation on model performance and overfitting, revealing its beneficial effects. The utilization of data augmentation has proven to be advantageous in the domain of DL. Hence, the 180-degree flipping as well as mirroring strategies were applied separately to increase the training datasets for all 3 models. Using the proposed strategies, the number of training datasets were increased by three times.

3.3.4 Detectron2 with Mask RCNN Models

3.3.4.1 Backbone of the Models

The framework employed for each model utilized the Detectron2 with the mask RCNN technique. Each algorithm generated an item mask together with the corresponding class, labels (rod, 58mm, and head-piece), and bounding boxes (outline of the features' region). Figure 3.5 shows the sequential steps involved in the Mask RCNN models.

The first step of the mask RCNN was to process the input images through a fusion of the Residual Network (ResNet-50) and the Feature Pyramid Network (FPN). This combination of networks was utilized to extract feature maps from radiographs at different scales, sizes, and placements. The second step of the model was to apply the Region Proposal Network (RPN), which was tasked with creating a proposed region of interest (ROI). Following selecting the ROI, the process of pooling features from the FPN was carried out at specified locations using the technique known as ROI Align. The data acquired using ROI Align was consolidated and

subjected to processing utilizing a Fully Connected Network (FCN) to generate final predictions about the class label and enhance the precision of the bounding box coordinates.

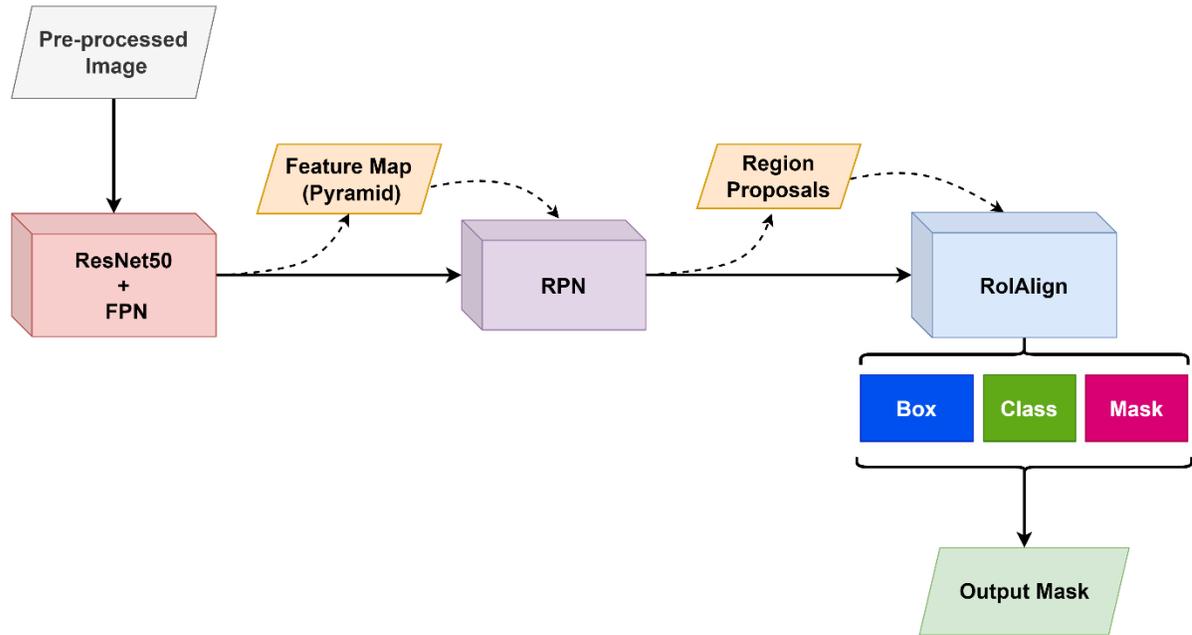


Figure 3.5 Mask RCNN structure with RestNet-50 and FPN

3.3.4.2 Model Optimization

Prior to finishing the models, a critical step involves the fine-tuning of the hyperparameters. The parameters for the model were selected by employing a trial-and-error methodology and analyzing the loss plots generated by the training and validation datasets. The factors encompassed in this study consist of the learning rate, batch size per image, the number of data-loader workers, the number of epochs or iterations, and the number of images per batch.

In this study, the hyperparameters for the first ML model were selected using the following approach: The learning rate employed in the experiment was set at 0.00025. The number of epochs or iterations utilized during the training process was 7800. A batch size per image 128

was chosen, and the data loader number of worker was 2. Once the training of the model was completed, a confidence score threshold of 0.45 was chosen for this model.

The 58 mm model exhibited a disparity in the quantity of radiographs available for analysis. In this study, a total of 378 images were utilized, as certain images were found to be undetectable in the rod model. Subsequently, to augment the training dataset, the same augmentation technique was applied. The hyperparameters were modified through a process of trial and error, resulting in the creation of many models. The model that was chosen for this study was configured with a set of specific hyperparameters. The learning rate utilized in the experiment was set to 0.00050. The number of epochs or iterations performed throughout the training process amounted to 2800. Additionally, a batch size per image of 256 was chosen. The number of dataloader workers in this model was also 2. Once the training of the model was completed, a confidence score threshold of 0.70 was chosen for this model. Table 3.1 summarizes the hyperparameters used for finalizing all three different models.

Table 3.1 Hyperparameters used in the mask RCNN models with radiographs

Model Name	Learning rate	Number of epoch	Batch size per image	Data-loader num workers	Confidence threshold
Rod	0.00025	7800	128	2	45%
58mm	0.00050	2800	256	2	70%
Rod-head	0.001	3400	512	2	80%

The third model, referred to as the head-piece detection model, had a comparable distribution of images for both training and validation, mirroring the 58mm model. The hyperparameters that were chosen are as follows: The learning rate used in the experiment was set to 0.001. The number of epochs or iterations performed during the training process was 3400. A batch size per image of 512 was used. The number of workers for the dataloader in this model was also 2. In this model, a confidence score threshold of 0.80 was chosen.

3.3.4.3 Evaluation Metrics

The entire structure and the AI system were implemented using Google Colab®. To assess the efficacy of the model before finalization, several metrics were considered, including Intersection over Union (IoU), Average Precision (AP), and the training loss-validation loss plots. The Intersection over Union (IoU) metric is utilized to measure the degree of overlap between the bounding boxes generated by the ground truth and the prediction, thereby providing a quantitative assessment of their proximity. The ratio between the area of junction and the aggregate area of the boxes is being referred to IoU.

$$\text{IoU} = \text{Area of Overlap} / \text{Area of Union} \quad \dots\dots\dots (1)$$

On the other hand, the AP is utilized as a quantitative metric within the Mask RCNN framework to evaluate the effectiveness of the model. The computation of the integral of the precision-recall curve is used to determine this. The metric of precision (P) is computed by dividing the count of true positive (TP) instances by the overall count of positive occurrences (TP+ FP). On the other hand, recall (R) is found by dividing the count of true positive (TP) instances by the whole count of positive instances (TP+FN) in the ground truth. The total count of positive cases in the ground truth can be determined by adding the number of real positive instances and the number of false negative instances. The computation of the AP involved the evaluation of different intersection over union (IoU) thresholds, which spanned from 0.50 to 0.95, incrementing by 0.05.

$$P = \frac{TP}{TP+FP} \cdot 100\% \quad \dots\dots\dots (2)$$

$$R = \frac{TP}{TP+FN} \cdot 100\% \quad \dots\dots\dots (3)$$

$$AP = \sum_n (R_n - R_{n-1})P_n \quad \dots\dots\dots (4)$$

Where P_n and R_n are the precision and recall at the nth threshold.

The loss plots from training loss and validation loss were another performance model finalizing metric for evaluating the performance of the models. The loss functions for the Mask RCNN could be defined as

$$L = L_{\text{class}} + L_{\text{box}} + L_{\text{mask}} \dots\dots\dots (5)$$

The plots depicting training loss and validation loss exhibit a relation with equation 5. In each iteration of the training process, the model computes the overall loss (L) by aggregating the various losses (L_{class}, L_{box}, L_{mask}). The trend of the metric first exhibits a decline as the number of epochs grows, but then demonstrates an upward trajectory when the model is subjected to over-training. To enhance the performance of the algorithms, hyperparameters which consisted of the learning rate, batch size per image, the number of data-loader workers, the number of epochs or iterations, and the number of images per batch were iterated based on analyzing the plots of the losses and AP. The number of epochs or the criterion for halting the epochs were determined using the minimum between the training and validation loss, and simultaneously achieving around the maximum AP.

When considering DL algorithms, the hyperparameters of the model were optimized by analyzing the plots of the losses and AP, leading to the finalization of the models. The number of epochs or the criterion for terminating epochs were determined using these two criteria as well. The objective of the epoch was to conclude at the specific place where the discrepancy in loss is minimal and the AP attains its greatest or near-maximum value. The presence of overfitting or underfitting in the model was also observed through the examination of the loss plots.

Although 62 radiographs were left for testing, 2 radiographs had poor image quality. Therefore, only 60 radiographs which consisted of 118 rods (58 x 2 rods and 2 x single rod) were finally used. An individual rod length extracted from each radiograph and the overall rod length adjustment were measured using the AI and manual methods. An experienced rater having 25 years of scoliosis experience did the manual measurements twice with a minimum of 1 week apart to reduce memory bias. An intra-rater reliability ICC_[2,1] and mean absolute

difference(MAD) \pm standard deviation (SD) between the 2 sessions' measurements were analyzed. To determine the reliability, the inter-method interclass correlation coefficient (ICC_[2,1]) with 95% confidence interval was used. To determine the agreement and bias between the manual and automatic measurements, a Bland-Altman analysis was employed. In all comparison, the length adjustment between the post-rod and pre-rod length was used.

3.4 Result

Figure 3.6 illustrates the training loss and validation loss plots of the 'head-piece' model from 0 to 8000 epochs. After the 8000 epochs, the validation loss started rising, meaning the model started overfitting. Hence, the final training model was stopped at 7800 epoch point. Table 3.2 presents the APs associated with each model.

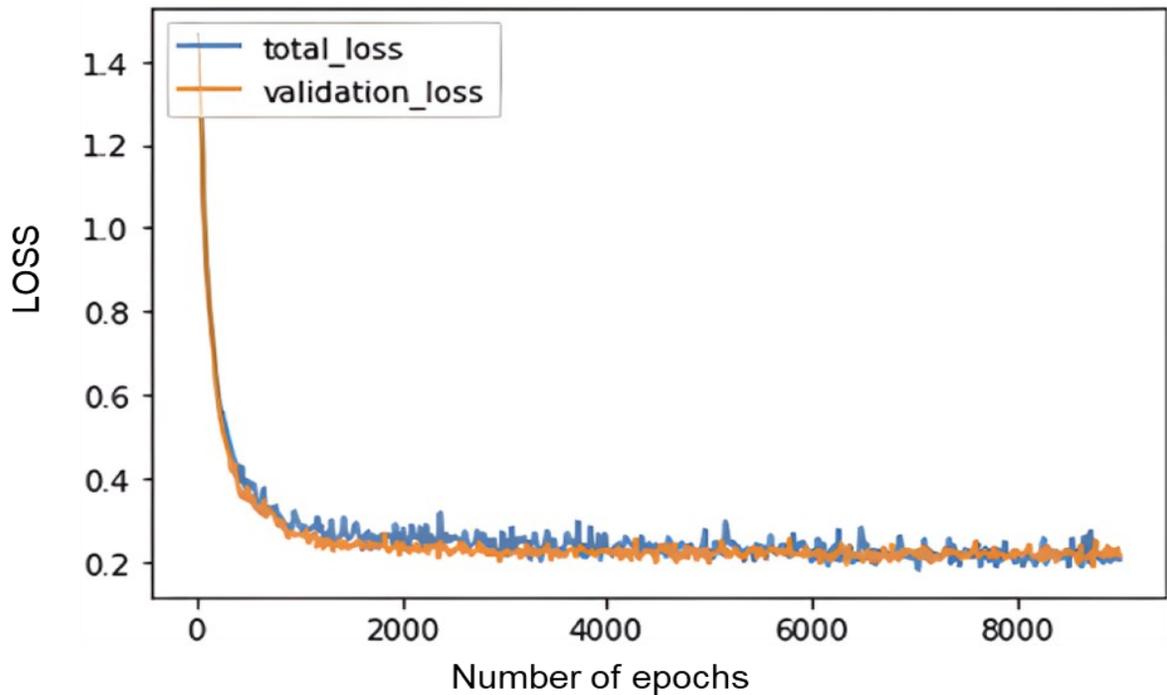


Figure 3.6 Training loss and validation loss plots of a 'rod' model

Table 3.2 AP for the mask RCNN models with radiographs

Model	AP
Rod model	67.7%
58mm model	94.8%
Head-piece model	86.3%

Figure 3.7 illustrates the sequential stages of the process starting from the original radiograph until both the 58 mm and the head-piece outputs.

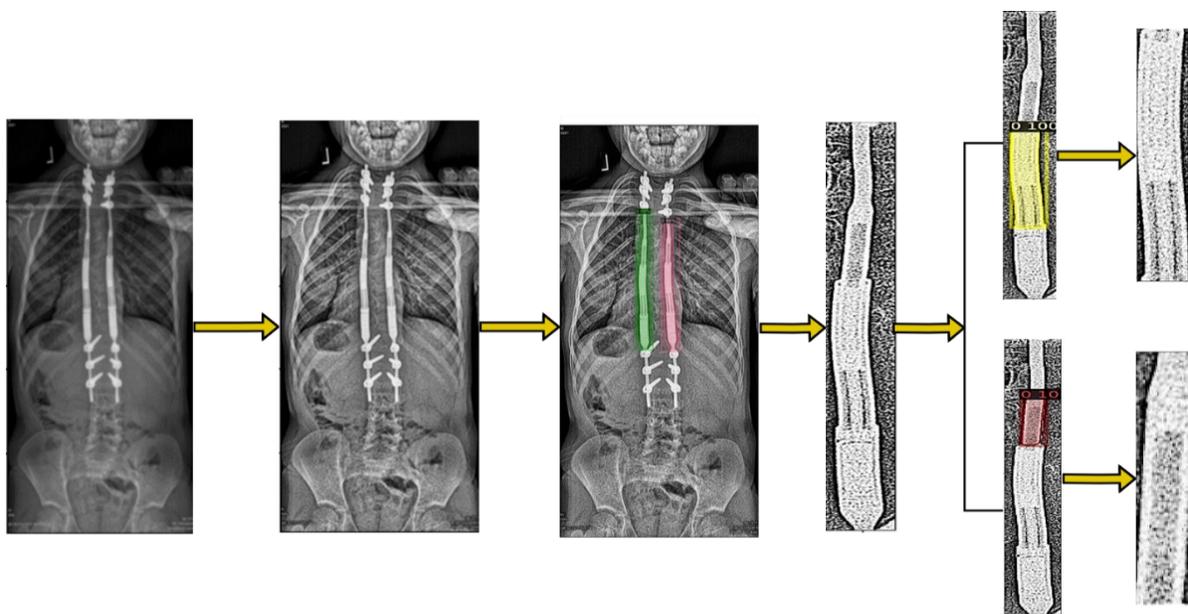


Figure 3.7 Sequential stages from raw image to detecting and measuring the desired length

For the 60 testing radiographs, the AI method exhibited a great rod model detection accuracy of 98.3% (116/118). Among the 116 rods (58 pre-rod and 58 post-rod adjustment), there were 2 post rods which were not able to report the measurements from the rod-head model. Hence, 56 rod length adjustments were used for final evaluation.

Table 3.3 shows a summary of a comparison between the AI and manual rod length adjustment measurement.

Table 3.3 Statistical comparison of rod measurements between AI and manual approach in radiographs

Method of measurement	Average length of adjustments (mm)	Inter-method correlation coefficient (ICC [2,1])	MAD (mm)	Bland Altman analysis agreement	Adjusted MAD (mm)
AI	3.20±2.94	0.902	0.98 ± 0.88	>90%	0.93 ±0.80
Manual	3.17 ± 2.73				

methods were 3.20 ± 2.94 mm and 3.17 ± 2.73 mm, respectively. The $MAD \pm SD$ between the AI and manual of the rod length adjustment was 0.98 ± 0.88 mm, which was within the clinically acceptance error ± 1.5 mm. Moreover, a significant proportion of the detections, precisely 87.5% (49 out of 56), fell within this specified range, indicating a high degree of precision exhibited by the AI model while comparing to the manual method. Figure 3.8 gives an illustration of the datapoints of all the adjustment difference between the automatic and manual approach.

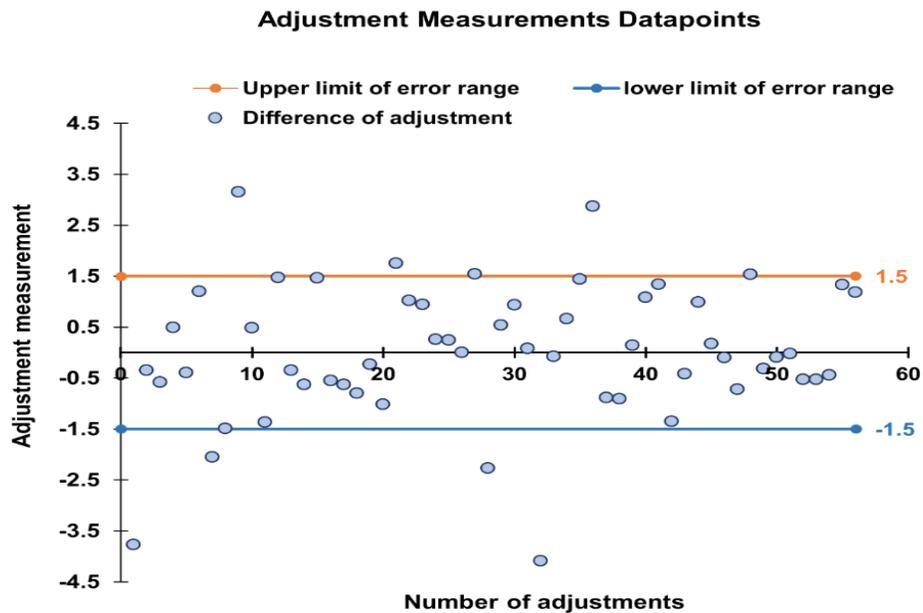


Figure 3.8 Data points from the 56 adjustments

Inter-method correlation coefficient (ICC 2,1) between the automatic adjustment measurement and the manual adjustment measurement was determined to be 0.902, demonstrating great reliability between the two techniques of measurement. The Pearson Correlation Coefficient between two approaches, was determined to be 0.89, which also indicates extremely excellent reliability. From Figure 3.9, we can see the correlation trendline to be positive and linear.

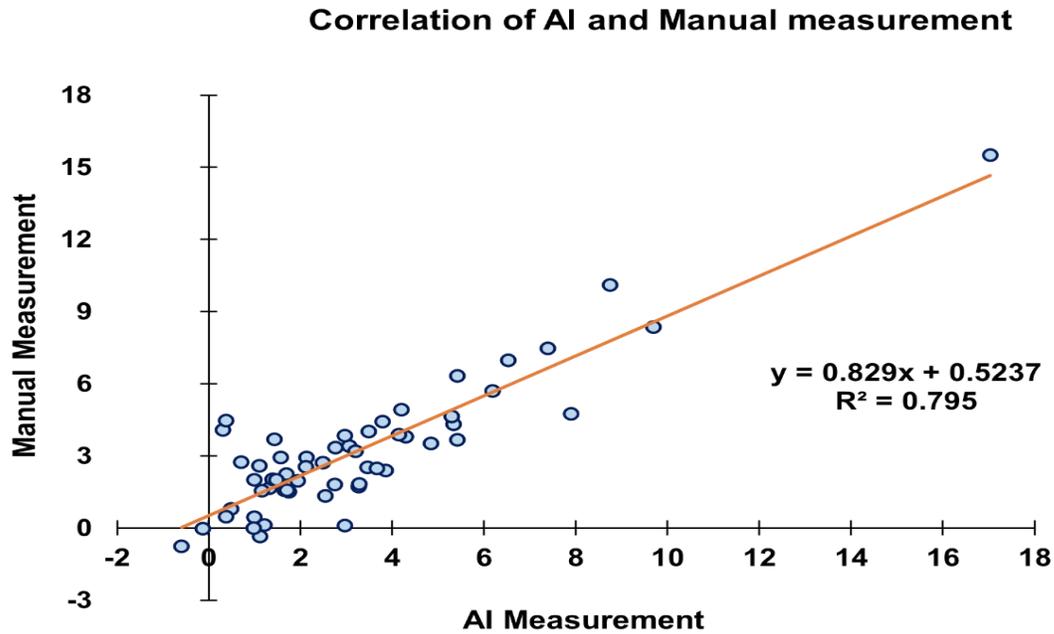


Figure 3.9 Correlation of AI and Manual measurement

In-term of the rater measurement accuracy and intra-rater reliability, the $MAD \pm SD$ and the $ICC_{[2,1]}$ were 0.37 ± 0.30 mm, 0.98, respectively. The rater's measurements also showed excellent accuracy and reliability. Figure 3.10 shows a Bland-Altman plot between the AI and human measurements. From the plot, there is no bias (mean = 0.02 mm) and only 4/56 points are outside the 95% confidence interval.

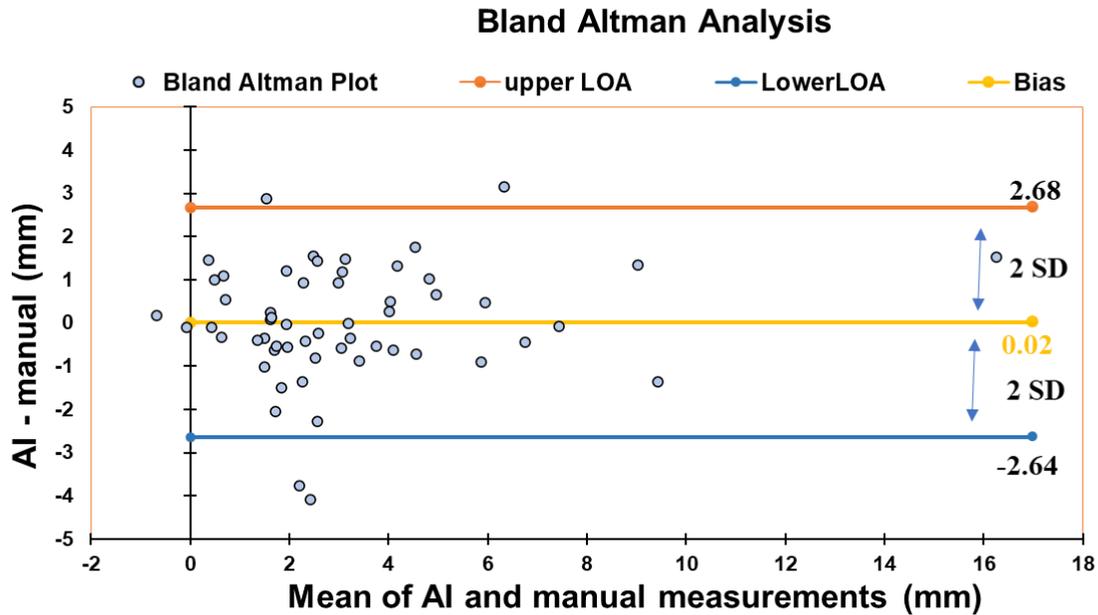


Figure 3.10 Bland Altman analysis showing agreement between AI and Human measurement for the adjustments

3.5 Discussion

Currently, the MCGR is the most commonly used surgical technique for children who have severe EOS. Setting external remote controller to adjust the rod length changes does not accurately correspond to the actual rod length extension. A study from 11 EOS patients validated the statement [77]. On the other hand, manual measurement of the rod length changes on radiographs also introduces errors. In addition, during the X-ray taken, the posture of the EOS children may vary between pre-rod and post-rod adjustment radiography. This means the sagittal tilt angle relative to the coronal plane is different. The projection of the rod onto the coronal plane affects the rod length measurements. This extra error is different from the human measurement error. In this study, there are 4 cases, which showed negative manual rod length changes with values for -0.35mm, -0.01mm, -0.76mm, -0.03mm. In this study, the developed AI method is based on calibration approach. A 58 mm rod length is used as a calibration feature.

The AI measurements from these 4 rods are 1.11mm, 0.98mm, -0.59mm, and -0.13mm. Another factor which may affect measurements is the image quality of the original radiograph. The low-dose or micro-dose configuration of the EOS X-ray system (EOS Imaging Inc., Paris, France) impacts the clarity of the image. A rod length adjustment of less than 1 mm may not be able to detect accurately. Further investigation will be conducted in the next study.

In this study, the rod model exhibits a comparatively lower AP of 67.6% when comparing with the 58 mm and rod-head models (94.8% and 86.3%). One possible explanation is that the whole rod covers the entire length of the radiograph, it has a higher chance to contain superfluous pixels and areas. Hence outlining a perfect rod region is much more difficult than the other 2 features.

Furthermore, the developed method has shown a significant improvement in the time needed to measure the length of the rod. With the manual approach, it usually takes up to 100 seconds to measure both rods. Using the AI system, it only takes 6 seconds per rod, which is approximately 15 times faster than the manual approach. This offers a great relief to clinicians especially when they are at a busy clinic.

Regarding the limitation, there are only a small number of participants, 24. Although these 24 participants have been followed for a few years and generated over 400 radiographs, there is still limited variation on these radiographs. Using augmentation method is able to increase the size of the training dataset, but it is not as good as more variation cases.

This study represents the first use of AI in this specific activity, which makes it difficult to compare with other AI models because there is no previous research available. Nevertheless, utilizing precise human measures from skilled raters facilitated significant comparisons with the newly constructed AI system in this research. In the future, I expect to see additional comparisons based on the recommendations. The new AI system has a notable speed advantage, being approximately 15 times faster than manual measuring procedures, providing substantial time-saving advantages in a clinical environment.

Chapter 4 Automated Method for Growing Rod Length Measurement on Ultrasound

4.1 Overview

This chapter provides a detailed report regarding the development of the automated system for measuring the rod length of MCGR on sonograms. An in-vivo data study was conducted to determine the accuracy, reliability, and speed of the developed ML algorithms.

The materials in this chapter are mainly exported from the manuscript is submitted to the Journal of Ultrasound in Medicine and Biology.

M. H. Kabir, M. Reformat, S. S. Hryniuk, K. Stampe, and E. Lou, " Automated Method for Growing Rod length Measurement on Ultrasound Images in Children with Early Onset Scoliosis", " Ultrasound in Medicine and Biology.

4.2 Introduction

Any spinal abnormality in children under the age of 10 with lateral curvature greater than 10 degrees is defined as EOS. In case the severity of scoliosis is significant, surgery is highly recommended [3], [4]. If spinal surgeries are not performed immediately, the spinal curvature is likely to worsen, which may lead to further physiological effects such as cosmetic deformity, and pulmonary and respiratory inefficiencies [1], [2], [5]. A surgical growing rod technique for treating EOS patients came to light in the late 1990s [3], [7]. This process is expensive [15], and is associated with post-surgical wounds and complications [11], [12], [13], [14]. To compensate for the growth of the children with EOS after surgery, the instrumented growing rod is adjusted every 6 to 12 months. Another spine surgery is required for every rod length adjustment. In the late 2000s, a novel non-invasive technique, which used a MCGR, was

introduced [16], [18], [24], [27]. This method only necessitates surgery during the rod installation. Safer rod lengthening is made externally by a machine. During every visit, an average of 3-6 months, a small rod length adjustment is made to reduce pain while getting the children's reaction in perfect harmony. If it hurts the patient too much during the process, it can be stopped. This growing rod modification aids in regulating spine abnormalities and significantly reduces the curvature of the spine.

Currently, radiography is the most common image modality to capture the internal spinal structure with the growing rods. The change of rod length is measured on the pre-operative and post-operative rod length adjustment radiographs. Understanding the change in rod length can assist surgeons in planning for the next surgery when the rod length reaches its limit. In the last decade, the US imaging technique has been promoted by taking advantage of no ionizing radiation exposure and being portable. In 2015, Yoon et al. [29] conducted a study comparing the rod length measured on US images versus on radiographs. They discovered that the interclass correlation coefficient (ICC) of the inter-method between radiography and US was 0.992 [29]. Their pilot study also showed the US method had a high inter- and intra-rater reliability with $ICC > 0.98$. In another study [78] which was conducted by Karlen et al., it was discovered that the US reduced average patients' radiation exposure by 83% when compared to both pre-rod and post-rod length adjustments radiographs. The patient wait time was also decreased by 64%, while the time required to complete the rod lengthening procedure was lowered by 50%. However, all X-ray and US measurements were performed manually, and the measurement time on US images was longer. Human measurement errors always exist. Furthermore, the experience and spinal knowledge of the rater affect the accuracy of measurements. The inaccuracy may affect the surgical planning which may influence patients' quality of life. Hence, monitoring the rod length adjustments precisely may optimize the treatment planning.

To reduce human measurement errors, ML algorithms, which is a subset of AI have been introduced in many applications [45], [46], [49], [50], [51]. Especially, Convolution Neural Network (CNN), a subset of ML, is commonly used for object recognition and segmentation in medical images. Recently, RCNNs have been applied in medical US images for spinal object

recognition, segmentation, and other studies [65], [66], [79]. Liu et al. [65] employed the faster RCNN on US images to automatically quantify spine curvature. They discovered that the model's AP was 76.1%, and the MAD between the AI technique and the manual method was within the clinical acceptance error. Tamas et al. [80] also conducted a study in using CNN (U-net architecture) for scoliosis visualization and spinal transverse process angle estimation on US images.

To ensure efficient model training and quick inference, ML algorithms have been implemented into Detectron2 framework for few studies [57], [59], [68]. Kumar et al. proposed in using a mask RCNN model with detectron2 for breast cancer cell identification on US images, they were able to identify cancer cells accurately with an average accuracy of 97.9% [81].

The objective of this study was to develop an ML system to identify and quantify the rod lengthening of the MCGR for children who have EOS surgeries. The reliability, accuracy, and speed of measurements were reported.

4.3 Method

4.3.1 Participants Selection

The ethics approval of this study was granted by the local health research ethics board (Pro00109499). Parental consents and child assents were signed before participation. The inclusion criteria were children who a) were diagnosed with EOS, b) had spinal surgeries with MCGR rods, and c) required rod length adjustment at the local clinic. A total of 23 participants (10 M, 13 F) with an average of 6.2 ± 1.4 years old were recruited from the local clinic. Among the 23 participants, 22 had dual rods installed, and one had a single rod.

4.3.2 Data Acquisition

To acquire the US images, a linear phase array wireless Claris US scanner L15HD3 (BC, Canada) equipped with a 5-15 MHz high-frequency transducer and a maximum depth of penetration of 7 cm was chosen. A nurse practitioner, who was responsible for the rod length adjustment, used an external magnet to identify the internal magnet inside the rod. Then she marked the approximate area on the skin. A US operator then applied warmed ultrasound gel on the surface of skin to provide a coupling medium. The location of rod head was estimated based on the magnet region. Once the rod head image was displayed on the screen, the operator captured the image. The left and right rods in the dual rod case were captured separately into 2 images. The practitioner then adjusted the rod length to an expected extension. The same US procedures were performed after the rod adjustment. Figure 4.1 shows a US image and an MCGR highlighted at the same area.

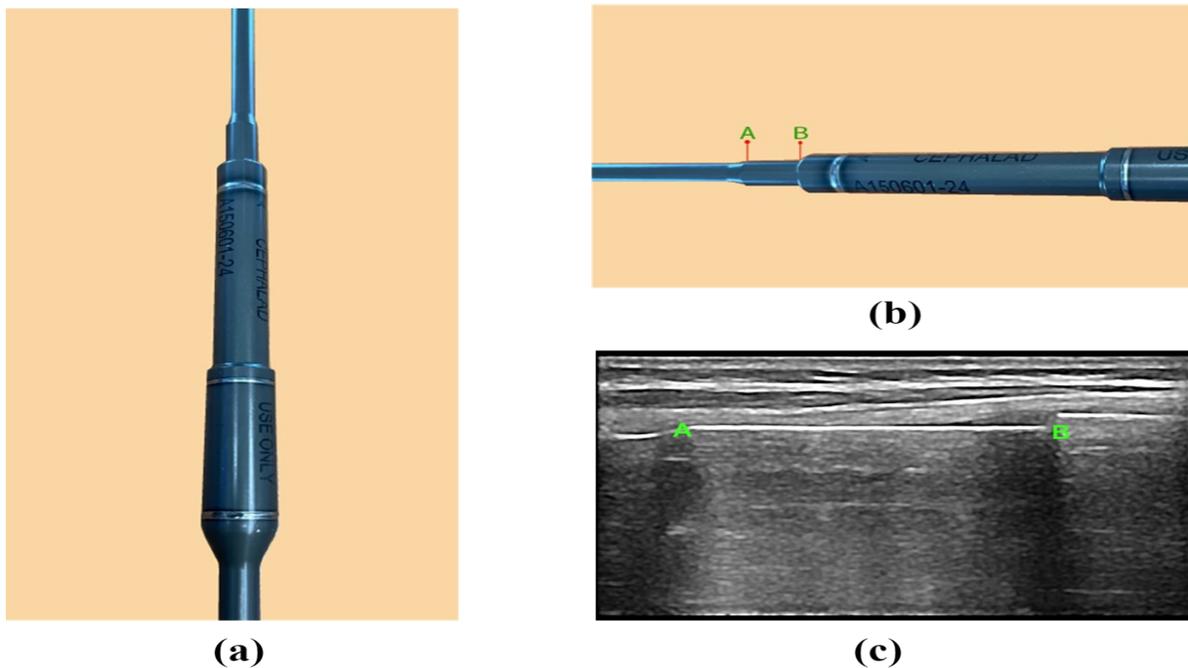


Figure 4.1 a) An actual MCGR rod used in this study, b) the portion A-B which gets elongated through the process, and c) the A-B portion in a sonogram

4.3.3 Overview of Rod length Detection

Figure 4.2 presents an overview of the developed system. The process commences with an initial unprocessed US image and culminates in the determination of the length of a specific segment of the rod. Filter was then applied to enhance the edge of the rod. After that, the ‘Boundary model’ was employed to delineate a region of interest around the rod head area. The second ML model, ‘Rod model’, was then applied to identify the rod portion precisely. Finally, the rod length was calculated.

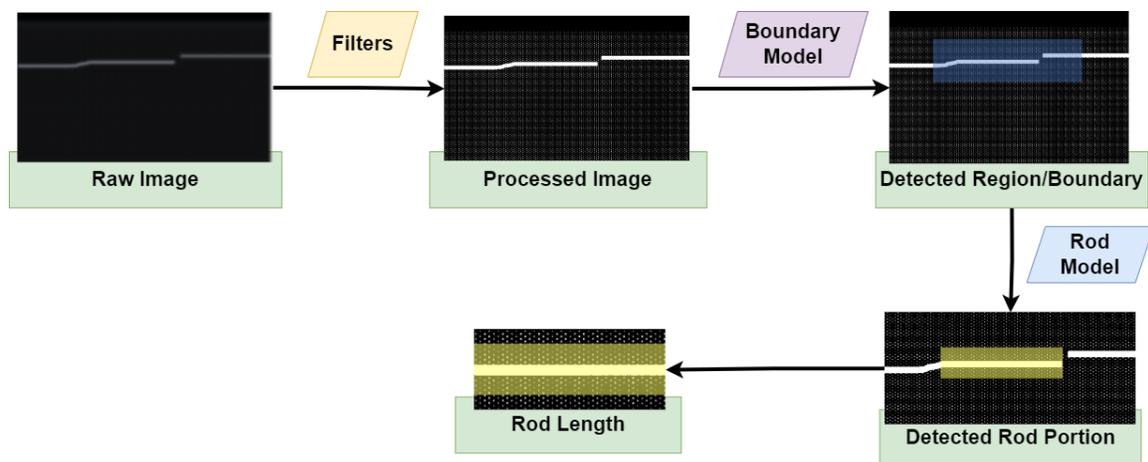


Figure 4.2 Overview of the process

4.3.4 Data Preprocessing Using Filters

US images are usually fuzzy at the edge of objects. To enhance the efficacy of model training, it was necessary to ensure the input images provide more clarity and sharper edges. A 3x3 kernel was applied to the raw image, which underwent erosion, dilation, and bilateral filtering. Subsequently, the image underwent median filtering, followed using an unsharp mask filter to enhance edge sharpness and clarity. Figure 4.3 illustrates the process by which a raw image undergoes a series of filtering techniques to enhance its quality.

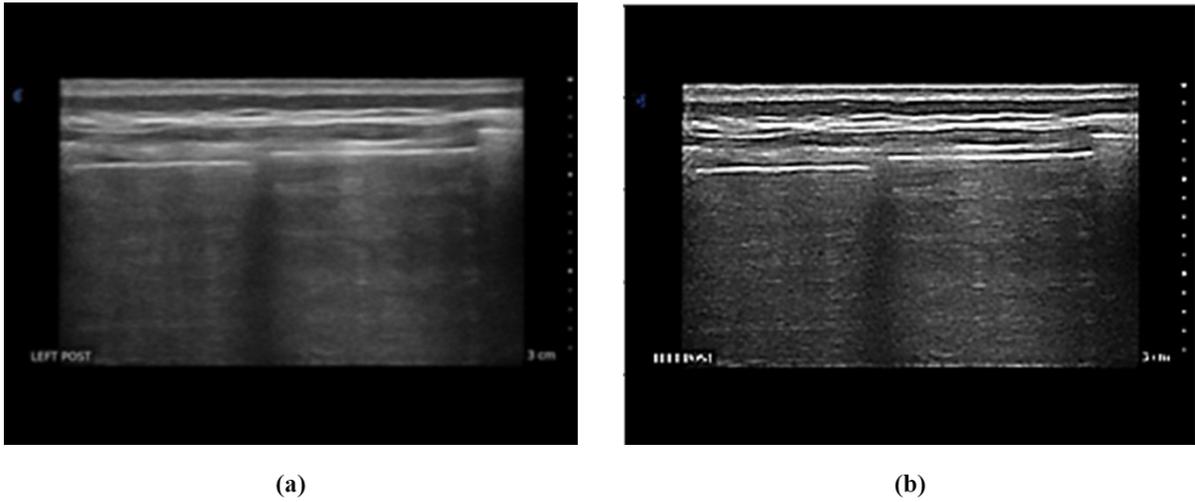


Figure 4.3 a) A raw unprocessed US image, and b) a filtered and processed US image

4.3.5 Data Annotation and Augmentation

The 90 images were distributed into three portions. Under the model development, 55 was used for training and 15 was used for validation. The last 20 images were used for testing. Prior to develop the model through training, it was crucial to annotate the training and validation datasets. To streamline the annotation process, 'LabelMe' [75] software was used to annotate the feature for both models. For the boundary model, a rater had been trained to identify the rod length region (A to B) from the US image (Figure 4.4a), then the region of interest (ROI) of the outermost points of the rod was selected using a rectangular box (Figure 4.4b). The boundary model provided delineation or demarcation for the rod. After the ROI was cropped (Figure 4.4c), a straight line was detected for the rod model (Figure 4.4d). The end goal of the automation was to measure the length of A-B. After annotation, each file was assigned a category and labeled accordingly. The file was saved with the file extension “. json”.

In the context of ML models that utilize images, datasets containing a substantial quantity of data are typically favored. The process of augmentation significantly contributes to the

expansion of data quantities in a remarkable manner [76], especially in image classification and segmentation [82], [83] applications. In this study, 180-degree horizontal flipping was applied to double the size for models training.

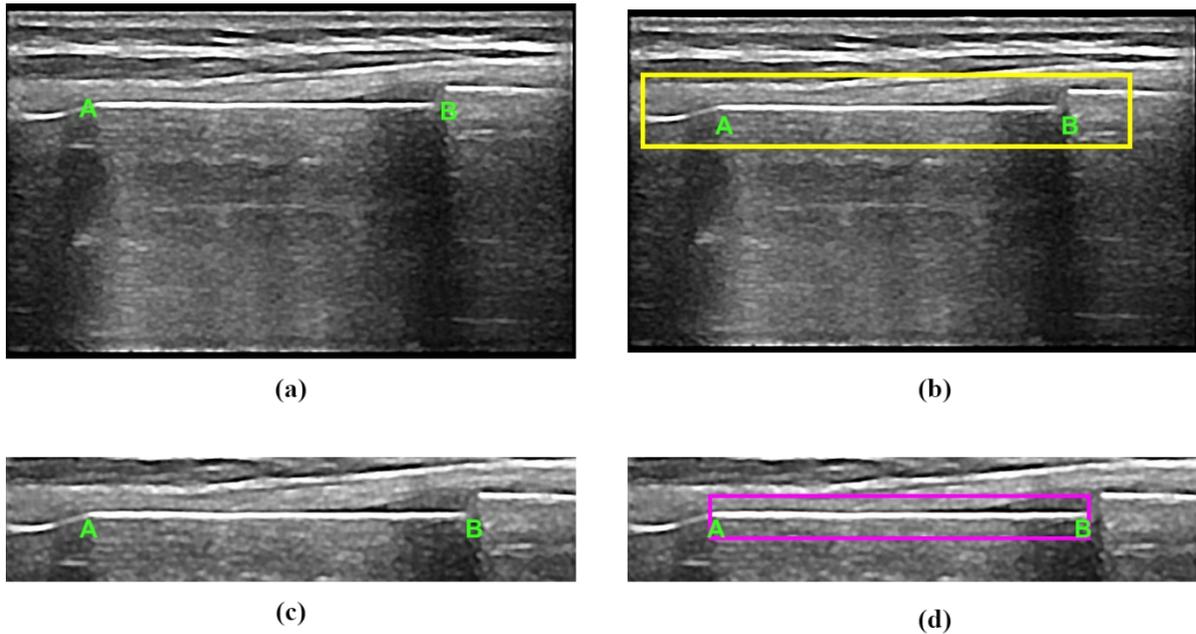


Figure 4.4 a) Exported US image from clinic, b) the expected output of the boundary model, c) cropped version of the first model's output, and d) expected output rod length from the rod model

4.3.6 Mask RCNN Models

Figure 4.5 presents the backbone architecture of the models. The two ML models which were developed used the Mask RCNN. The network was built with a Residual Network (ResNet-50), a Feature Pyramid Network (FPN) and the Regional Proposal Network (RPN). After filters were applied to the original US images, the rendered images were sent to the ResNet-50 to extract feature maps at different scales. The FPN was followed to further improve feature maps by generating a pyramid of feature maps. The maps were then input into a RPN to generate region proposals. After developing the region proposals, ROI Align was used to gather the

features from the FPN at the indicated locations. Using the collected information from the ROI Align, a fully connected network (FCN) was employed to predict the class label and refine the bounding box coordinates for each region proposal. Detectron2 was used as the framework and google Colab® was used for overall study involving ML algorithms.

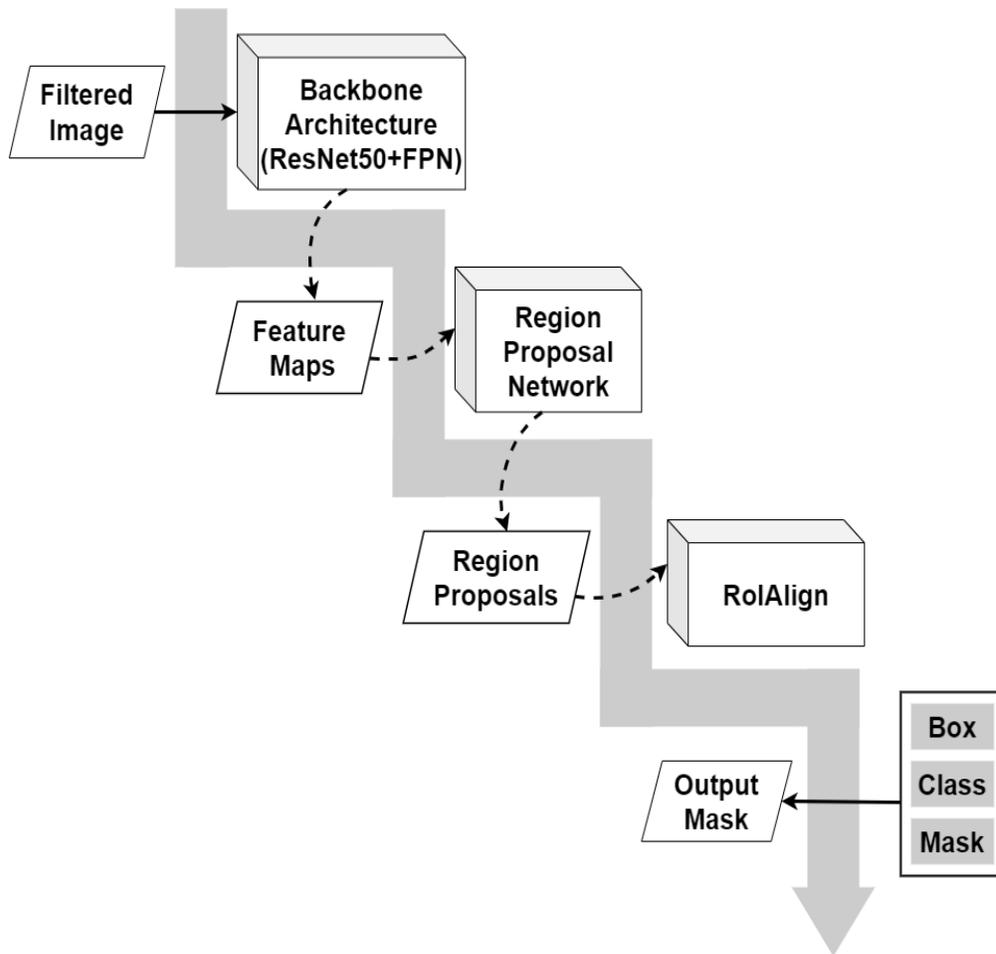


Figure 4.5 Backbone Architecture of the Mask RCNN models

4.3.7 Model Evaluation and Statistical Analysis

There were several metrics applied to finalize the models. Intersection over Union (IoU) was one of them. IoU helped to measure the degree of overlap between the bounding boxes created by the ground truth and the predicted data. The ratio was calculated by dividing the area of intersection by the total area of the combined boxes. Then a metric called AP was used to evaluate the accuracy of object detection and segmentation. The calculation of the AP required assessing several IoU thresholds, ranging from 0.50 to 0.95, with an increment of 0.05. The loss plot from training loss and validation loss with the change in epochs was another metric. The epoch was used to identify the stop point at which the difference between the training loss and validation loss was minimum, and at the same time the AP value was close to maximum. The hyperparameters of the models were optimized using trial and error approach to find the best possible combination.

This study involved two raters to measure the US images manually. An experienced rater 1 (R1) with 13 years of experience on sonograms measured the rod length of MCGR which was considered as the ground truth. To evaluate the variation of the US measurement, the primary author of this manuscript worked as the rater 2 (R2) who has no US imaging experience also measured the length manually.

The inter-method reliability of the interclass correlation coefficient ($ICC_{[2,1]}$), and $MAD \pm SD$ between the R1 and the AI system were compared. A Bland-Altman plot was used to determine the bias and agreement between the R1 and AI measurement. Furthermore, the inter-rater reliability ($ICC_{[2,1]}$) and $MAD \pm SD$ between R1 and R2 measurements were analyzed.

4.4 Result and discussion

4.4.1 Results

Figure 4.6 shows the sequential steps of an original US image passing through the stages of the system, and eventually providing the measured rod length.

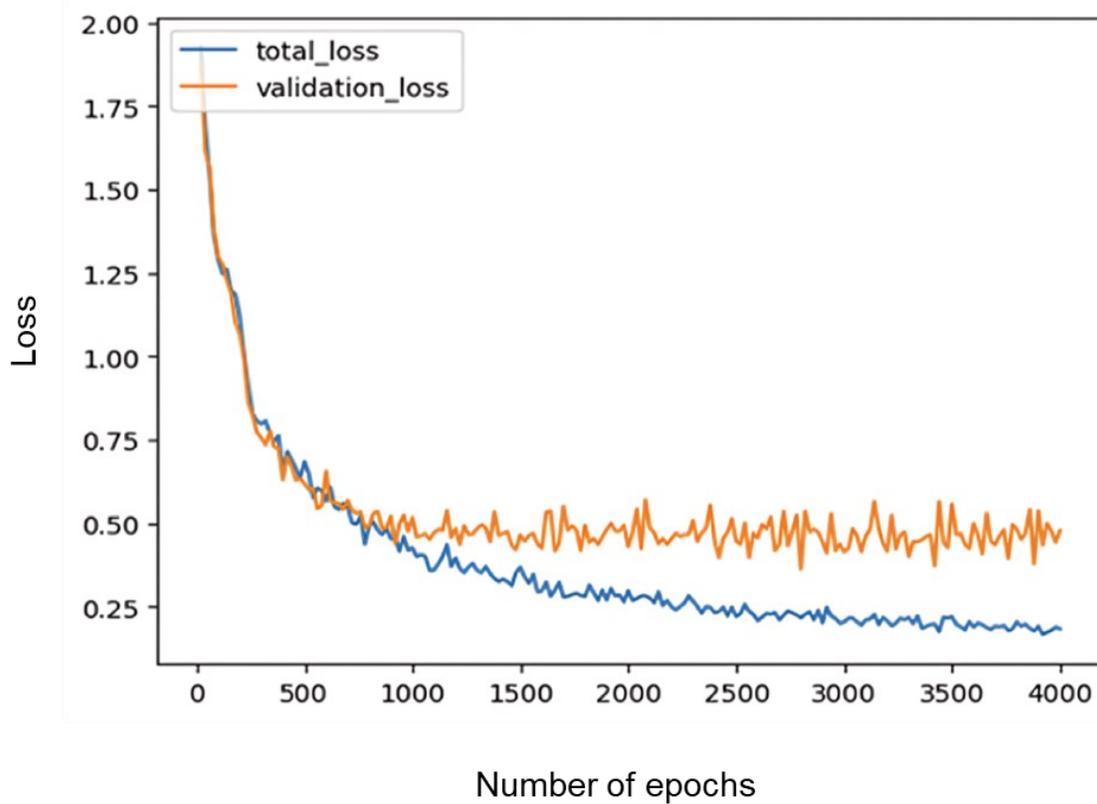


Figure 4.6 Loss plots of the rod model

Table 4.1 provides a concise overview of the hyperparameters included in both models, and the AP values achieved by the models. The AP value of the 'Boundary model' was better than the 'Rod model'.

Table 4.1 Hyperparameters used in the Boundary and Rod models and the AP values of both models

Model Name	Learning Rate	Number of epochs	Batch size per image	Data-loader num workers	Confidence threshold	AP
Boundary	0.00025	4200	128	2	80%	88.5%
Rod	0.001	1250	512	2	75%	60.2%

Figure 4.7 displays the loss plots of the rod model throughout 0 to 4000 epochs. At the 1200 epoch point, the model achieved the desired conditions of minimal loss difference and maximal AP.

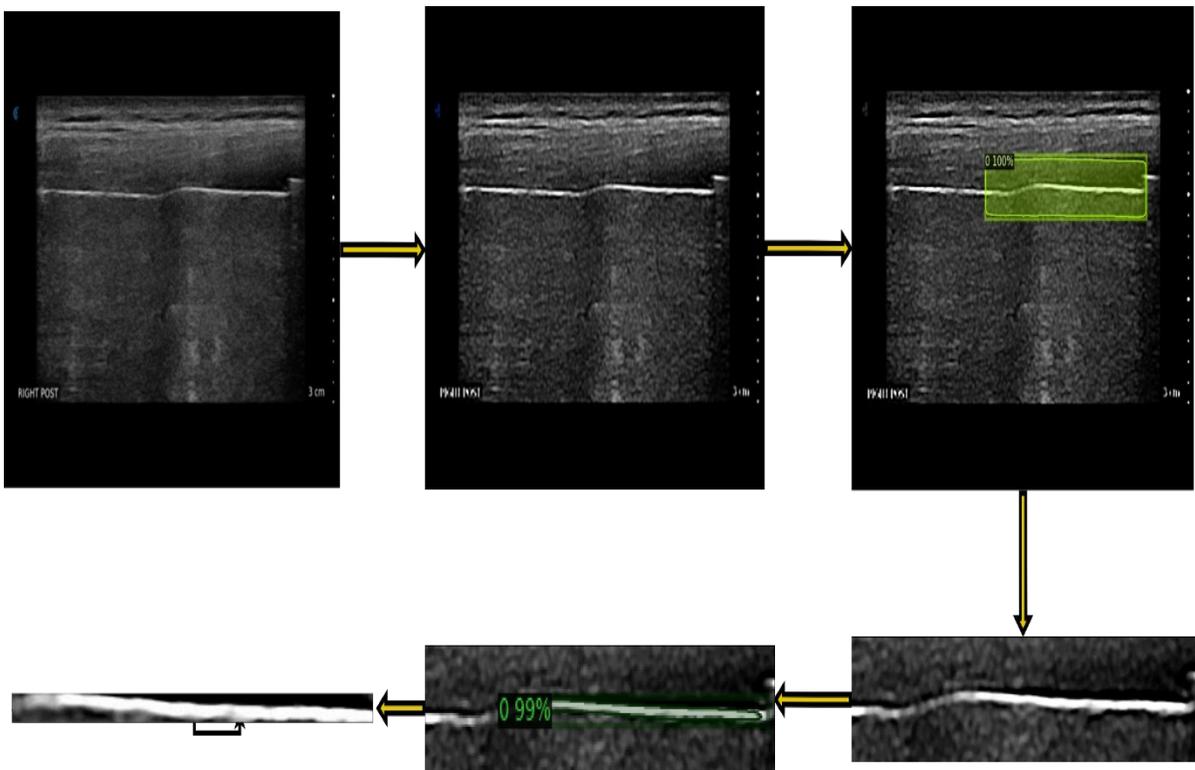


Figure 4.7 Sequential steps of the overall system to measure rod length

The average rod length measured from the AI system was 20.35 ± 7.33 mm, while the average length from the R1 was 20.33 ± 6.88 mm. The $MAD \pm SD$ between them was 1.2 ± 1.46 mm. Among the 20 test images, 16 (80%) were within the range of clinically accepted error of ± 1.5 mm. Figure 4.8 displays the data points representing the measurements obtained from both AI and manual measurements. This observation suggests that the AI model has a considerable level of accuracy when compared to the manual approach.

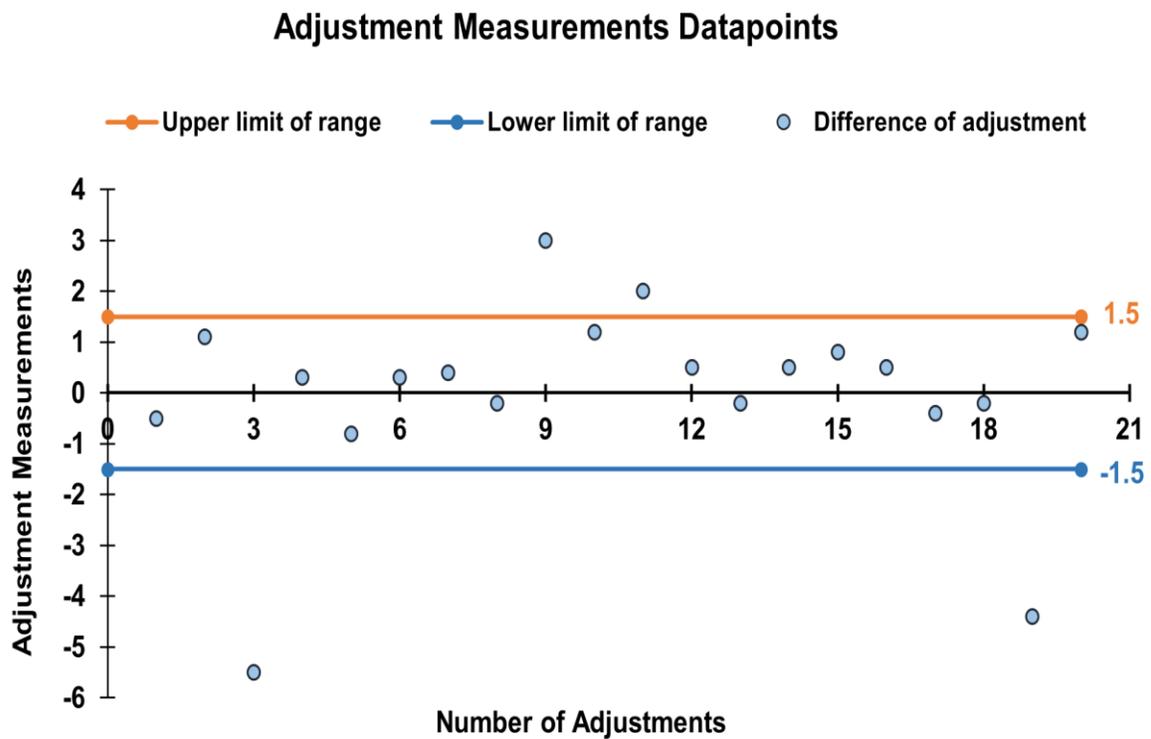


Figure 4.8 Datapoints of measurements from 20 US test dataset

The inter-method reliability of the interclass correlation coefficient ($ICC_{[2,1]}$) was 0.96, demonstrating great reliability between the two techniques of measurement. Table 4.2 provides a clear summary of the test data analysis from the manual (R1) and AI measurements.

Table 4.2 Comparison of rod length measurements between AI and R1 (with 20 test points)

Method of measurement (n = 20)	Average length (mm)	Inter-method correlation coefficient ($ICC [2,1]$)	MAD (mm)
AI	20.35 ± 7.33	0.96	1.2 ± 1.46
Manual	20.33 ± 6.88		

From the Bland-Altman plot as shown in Fig 4.9, the bias between the AI and the manual measurements (R1) was -0.02 mm which was almost close to zero. 18 out of 20 datapoints were within the 95% confidence interval of $\pm 2SD$.

For the inter-rater analysis, the inter-rater reliability $ICC_{[2,1]}$ was 0.99 which meant the US measurements were reliable and independent of rater experience. Additionally, the $MAD \pm SD$ between the two raters was 0.68 ± 0.47 mm. With 20 test data, the average length for the R1 was 20.33 ± 6.88 mm, while the average length for R2 was 20.51 ± 7.03 mm.

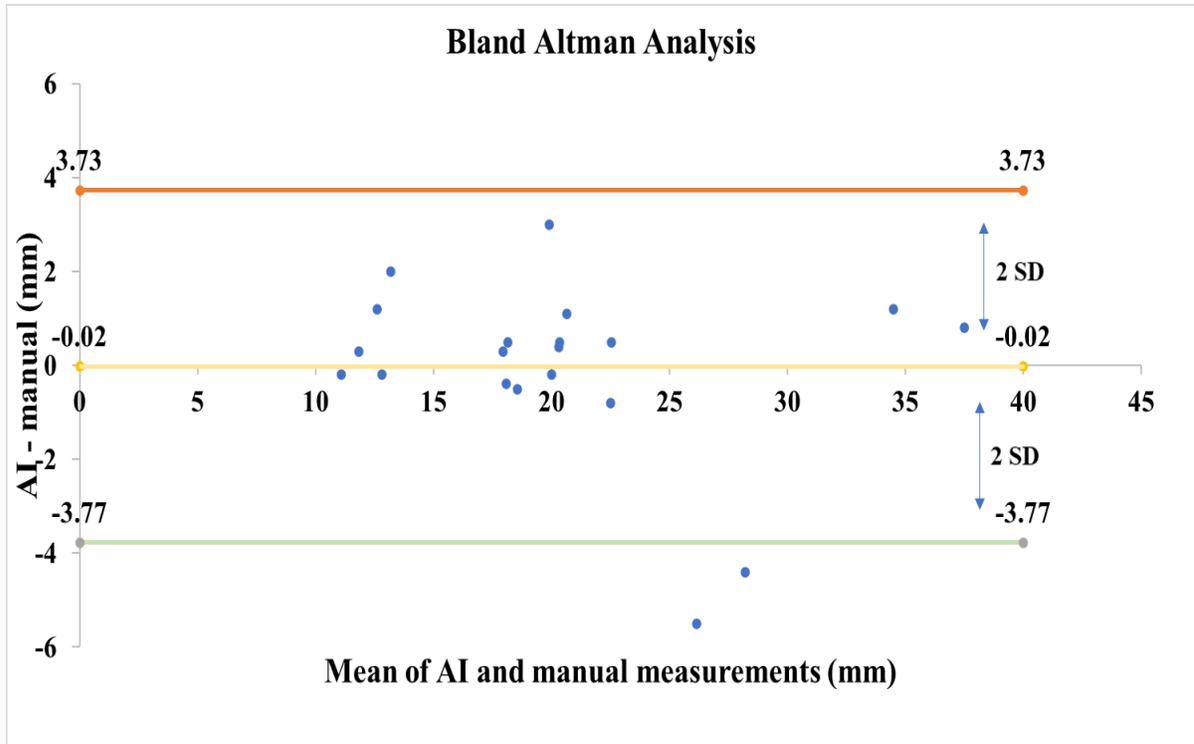


Figure 4.9 Bland Altman Analysis between AI and Manual measurements

Table 4.3 presents the $ICC_{[2,1]}$ and the $MAD \pm SD$ among R1 (main rater), R2 (secondary rater), and the AI measurement. Overall, the system displayed the results in 4.6 seconds after the US images were inputted.

Table 4.3 Comparison of rod length measurements between AI and manual measurements from both raters

	MAD \pm SD (mm)	Reliability ($ICC_{[2,1]}$)
R1 vs AI (n = 20)	1.2 \pm 1.46	0.96
R2 vs AI (n = 20)	1.64 \pm 1.69	0.95
R1 vs R2 (n = 20)	0.68 \pm 0.48	0.99

Among the 3 comparisons, manual US measurements between raters had less discrepancy and highest reliability. This meant measuring the rod length on US images was quite reliable.

Regarding the speed, the AI method displayed the results in an average of 4.6 seconds after the US images were inputted, the manual measurement requires about 60 seconds to calculate the result. It makes great difference for clinicians in a busy clinic.

4.4.2 Limitation

In this study, the major limitation is the data size. The training dataset only consists of 55 US images, and the rod lengths are mainly within 4 to 12 mm as shown in Figure 4.9. The number of training images is significantly lower than conventional DL models. ML models extract valuable information from the different characteristics and classifications found in the training images. The reduced quantity of images in the training dataset is recognized as a contributing element to the decreased AP, especially in the context of the rod model.

4.4.3 Advantages of using US

The US, being a non-ionizing imaging technology, has clear benefits compared to radiographic modalities, as it eliminates radiation exposure to patients. The lack of radiation hazard enables the acquisition of multiple images without jeopardizing patient well-being. This flexibility is helpful if clinicians decide to adjust the rod length more within the clinic to compensate for the rapid growing period. Furthermore, similar to the study [78], it saves a lot of clinic time. Using a portable US machine, it is easy to capture the image within the clinical room without requiring the patients to walk to the X-ray acquisition room. In a clinical setting, it is a 10-minute walk between the examination area and the diagnostic imaging department. Furthermore, the MCGR occasionally encounters complications such as the slippage of the distraction mechanism, resulting in unsuccessful rod length extension attempts. The capability

of US to take repeated images combined with the precision of AI-driven measurements can provide immediate feedback within the clinic so that the overall clinical efficiency can be improved.

4.4.4 Disadvantages of using US

However, the US acquisition takes longer time to acquire an image when comparing with radiography. It usually takes 10-15 seconds for a technician to identify the location of the rod magnet and save the image. The body motion of patients affects the US image quality. Sometimes it may need multiple trials to get a high-quality image.

4.4.5 Effect of Image Quality

Upon further examination of the large discrepancy points (beyond the clinical acceptance range) between the AI and manual measurements, it has been discovered that 2 of those 4 images exhibit a lack of clarity, which presents difficulties in identifying points A and B accurately. Since US images have the flexibility of taking multiple shots, the poor quality and unclear images could be avoided from analysis. There were 2 images of this type which was completely impossible to comprehend the A-B even with the naked eye. If the 2 unclear images were eliminated from this study, the revised average rod length from the AI system was 19.98 ± 7.44 mm, while the average length for the R1 became 20.10 ± 7.23 mm. The MAD \pm SD between the AI model and R1 dropped to 0.86 ± 1.0 mm. In addition, the inter-method reliability ($ICC_{[2,1]}$) improved from 0.96 to 0.98. After the elimination, there was just one point outside the 95% interval. Table 4.4 provides a summary of the revised analysis after the 2 fuzzy images were removed. Therefore, more training sets including some fuzzy images are required to improve accuracy and robustness.

Table 4.4 Comparison of rod length measurements between AI and manual methods after the fuzzy images were removed

Method of measurement (n = 18)	Average length (mm)	Inter-method correlation coefficient (ICC _[2,1])	MAD (mm)
AI	19.98 ± 7.44	0.98	0.86 ± 1.0
Manual	20.10 ± 7.23		

With the adjusted test set on 18 images, the revised Bland-Altman plot is shown in figure 4.10. The bias between the AI and the R1 was 0.11 mm which was still very small. There were only 1 datapoint outside the 95% confidence interval of ±2SD, means 94.4% datapoints were there within the confidence interval.

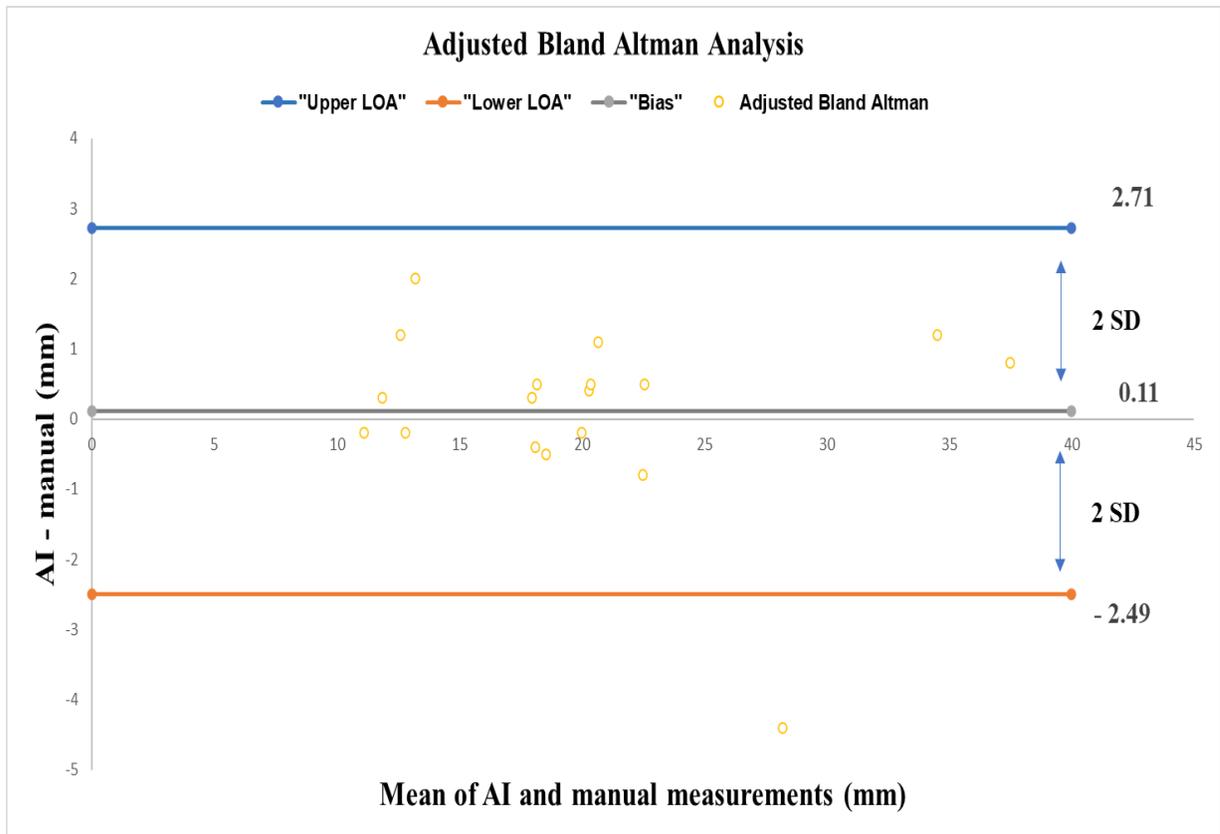


Figure 4.10 Adjusted Bland Altman Analysis between AI and Manual measurements

4.5 Conclusion and Contribution

Using US to monitor the rod length of MCGR has a significant benefit. Adding an AI system into a MCGR rod length adjustment clinic can further enhance its efficiency and cost saving. The AI-system developed in this study was the first reported automatic rod length measurement of MCGR on US images for children who have EOS surgeries. The reliability, accuracy and speed were good for clinical use.

Moreover, this groundbreaking study represents the inaugural application of AI within EOS research with US images for detecting and measuring the length of the head-piece in MCGR treatment. Despite being a pilot study, it already demonstrates accurate detection and measurement capabilities, albeit with acknowledged limitations. With anticipated improvements addressing these limitations and implementing future recommendations, the study holds promise to significantly streamline clinicians' workflow while enhancing research accuracy in this domain.

Chapter 5 Conclusion and Future Recommendations

5.1 Summary

EOS refers to a condition that affects children who are under the age of ten. When it comes to treating severe EOS with high Cobb angles, the most widely acknowledged surgical method is known as MCGR. The occurrence of manual measuring errors during the MCGR rod lengthening process results in challenges for the child's treatment and consumes a significant amount of time of the clinicians in a busy facility. Utilizing radiographs for the MCGR rod lengthening procedure entails the risk of radiation exposure. This thesis introduces an innovative approach to automate the measurement of rod length and adjustment in children undergoing MCGR therapy for EOS, using both radiographs and sonograms.

For radiographs, the system was developed using 449 radiographs. Initially, the training images were processed by using filters, then got annotated and augmented. Mask RCNN was used to build all the three different models and Detectron2 was used as the framework. The objective of the study was achieved by developing an automated ML system that could detect and measure the rod length of the MCGR from the radiographs. The system successfully detected 98.3% of the total rods. While comparing the adjustment difference with the manual measurements, the AI measurements demonstrated a strong inter-method correlation coefficient ($ICC_{[2,1]}$) of 0.902. The MAD between adjustment differences of the measurement methods was 0.98 ± 0.88 mm, with 87.5% falling within the clinically recognized error range of ± 1.5 mm. The study was proved to be accurate and 10 times faster than the manual approach.

To reduce the radiation effect associated with the radiographs, there has been an introduction of US in the field of EOS lately. But the possibility of human measurement errors exists here as well. Hence, another study was conducted in this thesis that used ML to develop an automated measuring method for MCGR treatment in EOS using the sonograms. 90 sonograms were used in this study. The ML system provided inter-method correlation coefficient ($ICC_{[2,1]}$)

of 0.96 when compared to rater 1 and 0.95 with rater 2. The inter-rater correlation coefficient ($ICC_{[2,1]}$) was 0.99 between the raters. The system displayed measurement output in 4.6 seconds only.

5.2 Future Recommendations

Although the detection algorithms have shown effectiveness, there is still significant room for improvement. Both radiograph system and US system exhibit potential for improving accuracy and AP.

Here are some of the key future recommendations that I think would help to get a more accurate and even faster automated system:

- The enhancement of system performance and effectiveness can be achieved by including a more extensive dataset. The number of images in the datasets was a limitation of this study, specially in case of sonograms. Increasing the number of overall images in the dataset would definitely help towards achieving a more a robust model.
- Another way of improvement in this study is the annotation. In some images, it was really difficult to understand the ‘58mm’ portion in the radiograph, and the A-B portion in the US. With better annotation on a larger dataset would help to achieve a more precise model.
- Augmentation is a way of increasing the image number in the datasets. Adding different types of augmentation would definitely be another way of increasing the size of the dataset.
- ML algorithms are evolving with time. Even though Mask RCNN has been one of the fastest models for object detection and instance segmentation, applying more ML algorithms and experimenting with the hyperparameters would probably help to find better and faster models.

- The inability to get adjustment measurements in US images has hindered the ability to undertake a comparative analysis of AI adjustment data between the two imaging modalities. With a more accurate and faster model, the comparative analysis of the US length adjustment and the radiograph length adjustment would be possible.

This work signifies a huge achievement in the realm of academic research, and its practical consequences are remarkable. While the software system is not yet fully developed, its application to newly captured images from a local clinic serves as a successful demonstration of its potential practical use for scoliosis patients who are having MCGR treatment. The utilization of this automated approach exhibits potential in facilitating accurate treatment planning, diminishing the probability of unforeseen surgical interventions, and ensuring the safety and reassurance of patients undergoing MCGR treatment for EOS patients.

Reference

- [1] S. L. Weinstein, L. A. Dolan, J. C. Cheng, A. Danielsson, and J. A. Morcuende, “Adolescent idiopathic scoliosis,” *The Lancet*, vol. 371, no. 9623, pp. 1527–1537, May 2008, doi: 10.1016/S0140-6736(08)60658-3.
- [2] H. Kim *et al.*, “Scoliosis Imaging: What Radiologists Should Know,” *RadioGraphics*, vol. 30, no. 7, pp. 1823–1842, Nov. 2010, doi: 10.1148/rg.307105061.
- [3] L. A. Karol, “The Natural History of Early-onset Scoliosis,” *Journal of Pediatric Orthopaedics*, vol. 39, p. S38, Jul. 2019, doi: 10.1097/BPO.0000000000001351.
- [4] J. E. Tis *et al.*, “Early Onset Scoliosis: Modern Treatment and Results,” *Journal of Pediatric Orthopaedics*, vol. 32, no. 7, p. 647, Nov. 2012, doi: 10.1097/BPO.0b013e3182694f18.
- [5] J. E. Lonstein and J. M. Carlson, “The prediction of curve progression in untreated idiopathic scoliosis during growth,” *JBJS*, vol. 66, no. 7, p. 1061, Sep. 1984, Accessed: Nov. 22, 2023. [Online]. Available: https://journals.lww.com/jbjsjournal/abstract/1984/66070/the_prediction_of_curve_progression_in_untreated.13.aspx
- [6] C. J. Goldberg *et al.*, “Respiratory Function and Cosmesis at Maturity in Infantile-onset Scoliosis,” *Spine*, vol. 28, no. 20, p. 2397, Oct. 2003, doi: 10.1097/01.BRS.0000085367.24266.CA.
- [7] G. H. Thompson, B. A. Akbarnia, and R. M. J. Campbell, “Growing Rod Techniques in Early-Onset Scoliosis,” *Journal of Pediatric Orthopaedics*, vol. 27, no. 3, p. 354, May 2007, doi: 10.1097/BPO.0b013e3180333eea.
- [8] B. A. Akbarnia, D. S. Marks, O. Boachie-Adjei, A. G. Thompson, and M. A. Asher, “Dual Growing Rod Technique for the Treatment of Progressive Early-Onset Scoliosis: A Multicenter Study,” *Spine*, vol. 30, no. 17S, p. S46, Sep. 2005, doi: 10.1097/01.brs.0000175190.08134.73.
- [9] I. J. Helenius *et al.*, “Outcomes of growing rod surgery for severe compared with moderate early-onset scoliosis: a matched comparative study,” *The Bone & Joint*

Journal, vol. 100-B, no. 6, pp. 772–779, Jun. 2018, doi: 10.1302/0301-620X.100B6.BJJ-2017-1490.R1.

- [10] S. Bess *et al.*, “Complications of Growing-Rod Treatment for Early-Onset Scoliosis: Analysis of One Hundred and Forty Patients,” *JBJS*, vol. 92, no. 15, p. 2533, Nov. 2010, doi: 10.2106/JBJS.I.01471.
- [11] R. El-Hawary, M. Kadhim, M. Vitale, J. Smith, A. Samdani, and J. M. Flynn, “VEPTR Implantation to Treat Children With Early-Onset Scoliosis Without Rib Abnormalities: Early Results From a Prospective Multicenter Study,” *Journal of Pediatric Orthopaedics*, vol. 37, no. 8, pp. e599–e605, Dec. 2017, doi: 10.1097/BPO.0000000000000943.
- [12] R. M. J. Campbell *et al.*, “The Effect of Opening Wedge Thoracostomy on Thoracic Insufficiency Syndrome Associated with Fused Ribs and Congenital Scoliosis,” *JBJS*, vol. 86, no. 8, p. 1659, Aug. 2004, Accessed: Nov. 22, 2023. [Online]. Available: https://journals.lww.com/jbjsjournal/fulltext/2004/08000/the_effect_of_opening_wedge_thoracostomy_on.9.aspx?casa_token=lh4qMpVZRD4AAAAA:29R8OWaMyyNrgIzeUpQo7asfysu-iEu2hh9XbRFBGMyyq00tjwhHQDCFFjNlgJDRMgrrakvwhO6aQny_RLxU6jJIqtDDr7W8
- [13] J.-F. Caubet *et al.*, “Increased Hemoglobin Levels in Patients With Early Onset Scoliosis: Prevalence and Effect of a Treatment With Vertical Expandable Prosthetic Titanium Rib (VEPTR),” *Spine*, vol. 34, no. 23, p. 2534, Nov. 2009, doi: 10.1097/BRS.0b013e3181bd0fc4.
- [14] J. H. T. Waldhausen, G. Redding, K. White, and K. Song, “Complications in using the vertical expandable prosthetic titanium rib (VEPTR) in children,” *Journal of Pediatric Surgery*, vol. 51, no. 11, pp. 1747–1750, Nov. 2016, doi: 10.1016/j.jpedsurg.2016.06.014.
- [15] B. A. Akbarnia and J. B. Emans, “Complications of Growth-Sparing Surgery in Early Onset Scoliosis,” *Spine*, vol. 35, no. 25, p. 2193, Dec. 2010, doi: 10.1097/BRS.0b013e3181f070b5.

- [16] A. Peiro-Garcia, J. Bourget-Murray, I. Suarez-Lorenzo, F. Ferri-De-Barros, and D. Parsons, “Early Complications in Vertical Expandable Prosthetic Titanium Rib and Magnetically Controlled Growing Rods to Manage Early Onset Scoliosis,” *International Journal of Spine Surgery*, vol. 15, no. 2, pp. 368–375, Apr. 2021, doi: 10.14444/8048.
- [17] K. M.-C. Cheung *et al.*, “Magnetically controlled growing rods for severe spinal curvature in young children: a prospective case series,” *The Lancet*, vol. 379, no. 9830, pp. 1967–1974, May 2012, doi: 10.1016/S0140-6736(12)60112-3.
- [18] J. M. Wick and J. Konze, “A Magnetic Approach to Treating Progressive Early-Onset Scoliosis,” *AORN Journal*, vol. 96, no. 2, pp. 163–173, Aug. 2012, doi: 10.1016/j.aorn.2012.05.008.
- [19] B. A. Akbarnia *et al.*, “Next Generation of Growth-Sparing Techniques: Preliminary Clinical Results of a Magnetically Controlled Growing Rod in 14 Patients With Early-Onset Scoliosis,” *Spine*, vol. 38, no. 8, p. 665, Apr. 2013, doi: 10.1097/BRS.0b013e3182773560.
- [20] D. Rolton, J. Richards, and C. Nnadi, “Magnetic controlled growth rods versus conventional growing rod systems in the treatment of early onset scoliosis: a cost comparison,” *Eur Spine J*, vol. 24, no. 7, pp. 1457–1461, Jul. 2015, doi: 10.1007/s00586-014-3699-7.
- [21] S. Ackerman, D. Polly, B. Akbarnia, J. Pawelek, and K. Schneider, “Cost analysis of magnetically controlled growing rods compared with traditional growing rods for early-onset scoliosis in the US: an integrated health care delivery system perspective,” *CEOR*, vol. Volume 8, pp. 457–465, Sep. 2016, doi: 10.2147/CEOR.S113633.
- [22] K. H. Teoh *et al.*, “Do magnetic growing rods have lower complication rates compared with conventional growing rods?,” *The Spine Journal*, vol. 16, no. 4, Supplement, pp. S40–S44, Apr. 2016, doi: 10.1016/j.spinee.2015.12.099.
- [23] B. A. Akbarnia *et al.*, “Traditional Growing Rods Versus Magnetically Controlled Growing Rods for the Surgical Treatment of Early-Onset Scoliosis: A Case-Matched 2-Year Study,” *Spine Deformity*, vol. 2, no. 6, pp. 493–497, Nov. 2014, doi: 10.1016/j.jspd.2014.09.050.

- [24] J. Lebon *et al.*, “Magnetically controlled growing rod in early onset scoliosis: a 30-case multicenter study,” *Eur Spine J*, vol. 26, no. 6, pp. 1567–1576, Jun. 2017, doi: 10.1007/s00586-016-4929-y.
- [25] E. Choi *et al.*, “Implant Complications After Magnetically Controlled Growing Rods for Early Onset Scoliosis: A Multicenter Retrospective Review,” *Journal of Pediatric Orthopaedics*, vol. 37, no. 8, pp. e588–e592, Dec. 2017, doi: 10.1097/BPO.0000000000000803.
- [26] C. Thakar, D. C. Kieser, M. Mardare, S. Haleem, J. Fairbank, and C. Nnadi, “Systematic review of the complications associated with magnetically controlled growing rods for the treatment of early onset scoliosis,” *Eur Spine J*, vol. 27, no. 9, pp. 2062–2071, Sep. 2018, doi: 10.1007/s00586-018-5590-4.
- [27] P. Obid, K. Yiu, K. Cheung, K. Kwan, M. Ruf, and J. P. Y. Cheung, “Magnetically controlled growing rods in early onset scoliosis: radiological results, outcome, and complications in a series of 22 patients,” *Arch Orthop Trauma Surg*, vol. 141, no. 7, pp. 1163–1174, Jul. 2021, doi: 10.1007/s00402-020-03518-z.
- [28] O. M. Stokes, E. J. O’Donovan, D. Samartzis, C. H. Bow, K. D. K. Luk, and K. M. C. Cheung, “Reducing radiation exposure in early-onset scoliosis surgery patients: novel use of ultrasonography to measure lengthening in magnetically-controlled growing rods,” *The Spine Journal*, vol. 14, no. 10, pp. 2397–2404, Oct. 2014, doi: 10.1016/j.spinee.2014.01.039.
- [29] W. W. Yoon, A. C. Chang, P. Tyler, S. Butt, S. Raniga, and H. Noordeen, “The use of ultrasound in comparison to radiography in magnetically controlled growth rod lengthening measurement: a prospective study,” *Eur Spine J*, vol. 24, no. 7, pp. 1422–1426, Jul. 2015, doi: 10.1007/s00586-014-3589-z.
- [30] R. F. Murphy and J. F. Mooney, “The First Generation of Early Onset Scoliosis Care,” *Journal of the Pediatric Orthopaedic Society of North America*, vol. 3, no. 2, p. 281, May 2021, doi: 10.55275/JPOSNA-2021-281.
- [31] B. A. Akbarnia, “Management Themes in Early Onset Scoliosis,” *JBJS*, vol. 89, no. suppl_1, p. 42, Feb. 2007, doi: 10.2106/JBJS.F.01256.

- [32] W. N. Sankar, D. C. Acevedo, and D. L. Skaggs, “Comparison of Complications Among Growing Spinal Implants,” *Spine*, vol. 35, no. 23, p. 2091, Nov. 2010, doi: 10.1097/BRS.0b013e3181c6edd7.
- [33] K. Watanabe *et al.*, “Risk Factors for Complications Associated With Growing-Rod Surgery for Early-Onset Scoliosis,” *Spine*, vol. 38, no. 8, p. E464, Apr. 2013, doi: 10.1097/BRS.0b013e318288671a.
- [34] G. H. Thompson *et al.*, “Comparison of Single and Dual Growing Rod Techniques Followed Through Definitive Surgery: A Preliminary Study,” *Spine*, vol. 30, no. 18, p. 2039, Sep. 2005, doi: 10.1097/01.brs.0000179082.92712.89.
- [35] P. Hosseini *et al.*, “Construct Levels to Anchored Levels Ratio and Rod Diameter Are Associated With Implant-Related Complications in Traditional Growing Rods,” *Spine Deform*, vol. 6, no. 3, pp. 320–326, 2018, doi: 10.1016/j.jspd.2017.11.004.
- [36] B. A. Akbarnia and G. M. Mundis, “Magnetically controlled growing rods in early onset scoliosis : Indications, timing and treatment,” *Orthopade*, vol. 48, no. 6, pp. 477–485, Jun. 2019, doi: 10.1007/s00132-019-03755-0.
- [37] C. K. H. Wong, J. P. Y. Cheung, P. W. H. Cheung, C. L. K. Lam, and K. M. C. Cheung, “Traditional growing rod versus magnetically controlled growing rod for treatment of early onset scoliosis: Cost analysis from implantation till skeletal maturity,” *Journal of Orthopaedic Surgery*, vol. 25, no. 2, p. 230949901770502, May 2017, doi: 10.1177/2309499017705022.
- [38] R. A. Kleinerman, “Cancer risks following diagnostic and therapeutic radiation exposure in children,” *Pediatr Radiol*, vol. 36, no. 2, pp. 121–125, Sep. 2006, doi: 10.1007/s00247-006-0191-5.
- [39] B. MacMahon, “Prenatal X-Ray Exposure and Childhood Cancer²³,” *JNCI: Journal of the National Cancer Institute*, vol. 28, no. 5, pp. 1173–1191, May 1962, doi: 10.1093/jnci/28.5.1173.
- [40] S. Ray, “A Quick Review of Machine Learning Algorithms,” in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Feb. 2019, pp. 35–39. doi: 10.1109/COMITCon.2019.8862451.

- [41] M. Mohammed, M. B. Khan, and E. B. M. Bashier, *Machine Learning: Algorithms and Applications*. CRC Press, 2016.
- [42] A. Shrestha and A. Mahmood, “Review of Deep Learning Algorithms and Architectures,” *IEEE Access*, vol. 7, pp. 53040–53065, 2019, doi: 10.1109/ACCESS.2019.2912200.
- [43] J. Ker, L. Wang, J. Rao, and T. Lim, “Deep Learning Applications in Medical Image Analysis,” *IEEE Access*, vol. 6, pp. 9375–9389, 2018, doi: 10.1109/ACCESS.2017.2788044.
- [44] S. Wang, D. M. Yang, R. Rong, X. Zhan, and G. Xiao, “Pathology Image Analysis Using Segmentation Deep Learning Algorithms,” *The American Journal of Pathology*, vol. 189, no. 9, pp. 1686–1698, Sep. 2019, doi: 10.1016/j.ajpath.2019.05.007.
- [45] M. I. Razzak, S. Naz, and A. Zaib, “Deep Learning for Medical Image Processing: Overview, Challenges and the Future,” in *Classification in BioApps: Automation of Decision Making*, N. Dey, A. S. Ashour, and S. Borra, Eds., in Lecture Notes in Computational Vision and Biomechanics. , Cham: Springer International Publishing, 2018, pp. 323–350. doi: 10.1007/978-3-319-65981-7_12.
- [46] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [47] S. Pouyanfar *et al.*, “A Survey on Deep Learning: Algorithms, Techniques, and Applications,” *ACM Comput. Surv.*, vol. 51, no. 5, p. 92:1-92:36, Sep. 2018, doi: 10.1145/3234150.
- [48] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, “Deep learning for visual understanding: A review,” *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016, doi: 10.1016/j.neucom.2015.09.116.
- [49] R. Yang and Y. Yu, “Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis,” *Frontiers in Oncology*, vol. 11, 2021, Accessed: Nov. 25, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fonc.2021.638182>

- [50] Z. Li, M. Dong, S. Wen, X. Hu, P. Zhou, and Z. Zeng, "CLU-CNNs: Object detection for medical images," *Neurocomputing*, vol. 350, pp. 53–59, Jul. 2019, doi: 10.1016/j.neucom.2019.04.028.
- [51] F. Yuan, Z. Zhang, and Z. Fang, "An effective CNN and Transformer complementary network for medical image segmentation," *Pattern Recognition*, vol. 136, p. 109228, Apr. 2023, doi: 10.1016/j.patcog.2022.109228.
- [52] T.-T. Pham, M.-B. Le, L. H. Le, J. Andersen, and E. Lou, "Assessment of hip displacement in children with cerebral palsy using machine learning approach," *Med Biol Eng Comput*, vol. 59, no. 9, pp. 1877–1887, Sep. 2021, doi: 10.1007/s11517-021-02416-9.
- [53] L. Zhang, J. Zhang, and S. Gao, "Region-Based Convolutional Neural Network-Based Spine Model Positioning of X-Ray Images," *BioMed Research International*, vol. 2022, p. e7512445, Jun. 2022, doi: 10.1155/2022/7512445.
- [54] L. Xie *et al.*, "Automatically measuring the Cobb angle and screening for scoliosis on chest radiograph with a novel artificial intelligence method," *Am J Transl Res*, vol. 14, no. 11, pp. 7880–7888, Nov. 2022, Accessed: Nov. 25, 2023. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9730103/>
- [55] J.-L. Cui, D.-D. Gao, S.-J. Shen, L.-Z. Wang, and Y. Zhao, "Cobb Angle Measurement Method of Scoliosis Based on U-net Network," In Review, preprint, Jul. 2021. doi: 10.21203/rs.3.rs-604546/v1.
- [56] J. Zhang, H. Li, L. Lv, and Y. Zhang, "Computer-Aided Cobb Measurement Based on Automatic Detection of Vertebral Slopes Using Deep Neural Network," *International Journal of Biomedical Imaging*, vol. 2017, p. e9083916, Oct. 2017, doi: 10.1155/2017/9083916.
- [57] A. V. S. Abhishek and S. Kotni, "Detectron2 Object Detection & Manipulating Images using Cartoonization," *International Journal of Engineering Research*, vol. 10, no. 08.
- [58] B. Jabir, N. Falih, and K. Rahmani, "Accuracy and Efficiency Comparison of Object Detection Open-Source Models," *Int. J. Onl. Eng.*, vol. 17, no. 05, p. 165, May 2021, doi: 10.3991/ijoe.v17i05.21833.

- [59] H. Soltani, M. Amroune, I. Bendib, and M. Y. Haouam, “Breast Cancer Lesion Detection and Segmentation Based On Mask R-CNN,” in *2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI)*, Sep. 2021, pp. 1–6. doi: 10.1109/ICRAMI52622.2021.9585913.
- [60] J. S. Yang *et al.*, “Growing Rods for Spinal Deformity: Characterizing Consensus and Variation in Current Use,” *Journal of Pediatric Orthopaedics*, vol. 30, no. 3, pp. 264–270, Apr. 2010, doi: 10.1097/BPO.0b013e3181d40f94.
- [61] G. J. Redding and O. H. Mayer, “Structure-Respiration Function Relationships Before and After Surgical Treatment of Early-onset Scoliosis,” *Clin Orthop Relat Res*, vol. 469, no. 5, pp. 1330–1334, May 2011, doi: 10.1007/s11999-010-1621-0.
- [62] B. A. Akbarnia *et al.*, “Dual Growing Rod Technique Followed for Three to Eleven Years Until Final Fusion: The Effect of Frequency of Lengthening,” *Spine*, vol. 33, no. 9, p. 984, Apr. 2008, doi: 10.1097/BRS.0b013e31816c8b4e.
- [63] H. B. Elsebai *et al.*, “Safety and Efficacy of Growing Rod Technique for Pediatric Congenital Spinal Deformities,” *Journal of Pediatric Orthopaedics*, vol. 31, no. 1, p. 1, Feb. 2011, doi: 10.1097/BPO.0b013e318202c1f0.
- [64] P. D. Sponseller *et al.*, “Growing Rods for Infantile Scoliosis in Marfan Syndrome,” *Spine*, vol. 34, no. 16, p. 1711, Jul. 2009, doi: 10.1097/BRS.0b013e3181a9ece5.
- [65] Z. Liu *et al.*, “Automatic spinal curvature measurement on ultrasound spine images using Faster R-CNN,” in *2021 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2021, pp. 1–4. doi: 10.1109/IUS52206.2021.9593343.
- [66] J. Liu and P. Li, “A Mask R-CNN Model with Improved Region Proposal Network for Medical Ultrasound Image,” in *Intelligent Computing Theories and Application*, D.-S. Huang, K.-H. Jo, and X.-L. Zhang, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, pp. 26–33. doi: 10.1007/978-3-319-95933-7_4.
- [67] R. Wang, J. H. Yi Voon, D. Ma, S. Dabiri, K. Popuri, and M. F. Beg, “Vertebra Segmentation for Clinical CT Images Using Mask R-CNN,” in *8th European Medical and Biological Engineering Conference*, T. Jarm, A. Cvetkoska, S. Mahnič-Kalamiza, and D. Miklavcic, Eds., in IFMBE Proceedings. Cham: Springer International Publishing, 2021, pp. 1156–1165. doi: 10.1007/978-3-030-64610-3_130.

- [68] F. J. Yagüe, J. F. Diez-Pastor, P. Latorre-Carmona, and C. I. G. Osorio, “Defect detection and segmentation in X-Ray images of magnesium alloy castings using the Detectron2 framework.” arXiv, Feb. 28, 2022. doi: 10.48550/arXiv.2202.13945.
- [69] A. A. Ali, R. Katta, R. Jasek, B. Chramco, and S. Krayem, “COVID-19 Detection from Chest X-Ray Images Using Detectron2 and Faster R-CNN,” in *Data Science and Algorithms in Systems*, R. Silhavy, P. Silhavy, and Z. Prokopova, Eds., in Lecture Notes in Networks and Systems. Cham: Springer International Publishing, 2023, pp. 37–53. doi: 10.1007/978-3-031-21438-7_3.
- [70] S. Kansal, S. Purwar, and R. K. Tripathi, “Image contrast enhancement using unsharp masking and histogram equalization,” *Multimed Tools Appl*, vol. 77, no. 20, pp. 26919–26938, Oct. 2018, doi: 10.1007/s11042-018-5894-8.
- [71] F. Y. M. Lure, P. W. Jones, and R. S. Gaborski, “Multiresolution unsharp masking technique for mammogram image enhancement,” in *Medical Imaging 1996: Image Processing*, SPIE, Apr. 1996, pp. 830–839. doi: 10.1117/12.237989.
- [72] K. Panetta, Y. Zhou, S. Agaian, and H. Jia, “Nonlinear Unsharp Masking for Mammogram Enhancement,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 6, pp. 918–928, Nov. 2011, doi: 10.1109/TITB.2011.2164259.
- [73] K. Koonsanit, S. Thongvigitmanee, N. Pongnapang, and P. Thajchayapong, “Image enhancement on digital x-ray images using N-CLAHE,” in *2017 10th Biomedical Engineering International Conference (BMEiCON)*, Aug. 2017, pp. 1–4. doi: 10.1109/BMEiCON.2017.8229130.
- [74] H. Wen, W. Qi, and L. Shuang, “Medical X-Ray Image Enhancement Based on Wavelet Domain Homomorphic Filtering and CLAHE,” in *2016 International Conference on Robots & Intelligent System (ICRIS)*, ZhangJiaJie, China: IEEE, Aug. 2016, pp. 249–254. doi: 10.1109/ICRIS.2016.50.
- [75] K. Wada *et al.*, “wkentaro/labelme: v4.6.0.” Zenodo, Nov. 18, 2021. doi: 10.5281/ZENODO.5711226.
- [76] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

- [77] K. H. Teoh *et al.*, “Does the external remote controller’s reading correspond to the actual lengthening in magnetic-controlled growing rods?,” *Eur Spine J*, vol. 29, no. 4, pp. 779–785, Apr. 2020, doi: 10.1007/s00586-020-06335-5.
- [78] J. Karlen and M. Riemann, “Optimization of a MCGR US-Guided Lengthening Clinic,” *Spine Deformity*, vol. 6, no. 6, p. 807, Nov. 2018, doi: 10.1016/j.jspd.2018.09.037.
- [79] B. Benjdira, K. Ouni, M. M. Al Rahhal, A. Albakr, A. Al-Habib, and E. Mahrous, “Spinal Cord Segmentation in Ultrasound Medical Imagery,” *Applied Sciences*, vol. 10, no. 4, Art. no. 4, Jan. 2020, doi: 10.3390/app10041370.
- [80] T. Ungi *et al.*, “Automatic Spine Ultrasound Segmentation for Scoliosis Visualization and Measurement,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3234–3241, Nov. 2020, doi: 10.1109/TBME.2020.2980540.
- [81] P. Kumar, K. V. Sumandar, G. V. S. Das, G. K. Reddy, S. Srivastava, and A. Kuamar, “Ultrasound based breast cancer recognition using Mask RCNN,” pp. 526–531, Jan. 2023, doi: 10.1049/icp.2023.1543.
- [82] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth, “A review of medical image data augmentation techniques for deep learning applications,” *Journal of Medical Imaging and Radiation Oncology*, vol. 65, no. 5, pp. 545–563, 2021, doi: 10.1111/1754-9485.13261.
- [83] A. Mikołajczyk and M. Grochowski, “Data augmentation for improving deep learning in image classification problem,” in *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, May 2018, pp. 117–122. doi: 10.1109/IIPHDW.2018.8388338.

Appendix

This part of the thesis contains the source code used for the studies. Appendix A presents the source code for the rod detection model of using radiographs. However, only one complete source code is presented here, as they contain the similar things apart from the changes in hyperparameters. Appendix B contains the combination of three ML models which was used to determine the AI measurements with the help of test images. All the source codes are taken from google colab(.py) format.

Appendix A

1. Mounting colab with google drive

```
from google.colab import drive
drive.mount('/content/drive')
```

2. Installing dependencies for detectron2 and installing detectron2

```
!pip install -U torch==1.5 torchvision==0.6 -
fhttps://download.pytorch.org/whl/cu101/torch\_stable.html

!pip install cython pyyaml==5.1

!pip install -U 'git+https://github.com/cocodataset/cocoapi.git#subdirectory=PythonAPI'

import torch

TORCH_VERSION = ".".join(torch.__version__.split(".")[2:])

CUDA_VERSION = torch.__version__.split("+")[-1]

print("torch: ", TORCH_VERSION, "; cuda: ", CUDA_VERSION)

!pip install detectron2 -f
https://dl.fbaipublicfiles.com/detectron2/wheels/\$CUDA\_VERSION/torch\$TORCH\_VERSION/index.html
```

```
import detectron2

from detectron2.utils.logger import setup_logger

setup_logger()
```

3. Importing common libraries

```
import some common libraries

import numpy as np

import os, json, cv2, random

from google.colab.patches import cv2_imshow

from detectron2 import model_zoo

from detectron2.engine import DefaultPredictor

from detectron2.config import get_cfg

from detectron2.utils.visualizer import Visualizer

from detectron2.data import MetadataCatalog, DatasetCatalog

import matplotlib.pyplot as plt

from detectron2.structures import BoxMode

from detectron2 import model_zoo

from detectron2.engine import DefaultTrainer, DefaultPredictor

from detectron2.config import get_cfg

from detectron2.utils.visualizer import ColorMode, Visualizer
```

4. Getting information on datasets (.json files)

```
def get_data_dicts(directory, classes):

    dataset_dicts = []

    for filename in [file for file in os.listdir(directory) if file.endswith('.json')]:

        json_file = os.path.join(directory, filename)

        with open(json_file) as f:
```

```

img_anns = json.load(f)

record = {}
filename = os.path.join(directory, img_anns["imagePath"])
record["file_name"] = filename

record["height"] = img_anns["imageHeight"]
record["width"] = img_anns["imageWidth"]
record["image_id"] = json_file
annos = img_anns["shapes"]
objs = []
for anno in annos:
    px = [a[0] for a in anno['points']] # x coord
    py = [a[1] for a in anno['points']] # y-coord
    poly = [(x, y) for x, y in zip(px, py)] # poly for segmentation
    poly = [p for x in poly for p in x]

    obj = {
        "bbox": [np.min(px), np.min(py), np.max(px), np.max(py)],
        "bbox_mode": BoxMode.XYXY_ABS,
        "segmentation": [poly],
        "category_id": classes.index(anno['label']),
        "iscrowd": 0
    }
    objs.append(obj)
record["annotations"] = objs
dataset_dicts.append(record)

```

```
return dataset_dicts
```

5. Registering dataset and creating metadata

```
classes = ['rod']
data_path = 'dataset directory'
for d in ["train", "validation"]:
    DatasetCatalog.register(
        "category_" + d,
        lambda d=d: get_data_dicts(data_path+d, classes)
    )
    MetadataCatalog.get("category_" + d).set(thing_classes=classes)
microcontroller_metadata = MetadataCatalog.get("category_train")
```

6. Hyperparameters setting.

```
cfg = get_cfg()
cfg.merge_from_file(model_zoo.get_config_file("COCO-
InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml"))
cfg.DATASETS.TRAIN = ("category_train",)
cfg.DATASETS.TEST = ("category_validation",)
#cfg.TEST.EVAL_PERIOD = 100
cfg.DATALOADER.NUM_WORKERS = 2
#cfg.SOLVER.CHECKPOINT_PERIOD = 600
cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-
InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml")
cfg.SOLVER.IMS_PER_BATCH = 2
#cfg.SOLVER.BASE_LR = 0.00025
#cfg.SOLVER.CHECKPOINT_PERIOD = 10
cfg.SOLVER.BASE_LR = 0.00025
```

```

cfg.SOLVER.MAX_ITER = 7800
cfg.MODEL.ROI_HEADS.NUM_CLASSES = 1
cfg.MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE = 128
cfg.OUTPUT_DIR = "/directory/"

```

7. Setting up losses and training the model

```

from detectron2.engine import HookBase
from detectron2.data import build_detection_train_loader
import detectron2.utils.comm as comm
class ValidationLoss(HookBase):
    def __init__(self, cfg):
        super().__init__()
        self.cfg = cfg.clone()
        self.cfg.DATASETS.TRAIN = cfg.DATASETS.TEST
        self._loader = iter(build_detection_train_loader(self.cfg))
    def after_step(self):
        data = next(self._loader)
        with torch.no_grad():
            loss_dict = self.trainer.model(data)
            losses = sum(loss_dict.values())
            assert torch.isfinite(losses).all(), loss_dict
            loss_dict_reduced = {"category_validation" + k: v.item() for k, v in
                                comm.reduce_dict(loss_dict).items()}
            losses_reduced = sum(loss for loss in loss_dict_reduced.values())
            if comm.is_main_process():
                self.trainer.storage.put_scalars(total_val_loss=losses_reduced,
                                                **loss_dict_reduced)

```

```

os.makedirs(cfg.OUTPUT_DIR, exist_ok=True)

trainer = DefaultTrainer(cfg)

val_loss = ValidationLoss(cfg)
trainer.register_hooks([val_loss])
# swap the order of PeriodicWriter and ValidationLoss
trainer._hooks = trainer._hooks[:-2] + trainer._hooks[-2][::-1]
trainer.resume_or_load(resume=True)
trainer.train()

```

8. Saving the model

```

import pickle

cfg.MODEL.WEIGHTS = 'model path directory'
cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST = 0.45
with open('directory for saving the model', 'wb') as f:
    pickle.dump(cfg, f)
f.close()

```

9. Using evaluator to get AP

```

predictor = DefaultPredictor(cfg)

from detectron2.evaluation import COCOEvaluator, inference_on_dataset

from detectron2.data import build_detection_test_loader

#evaluator = COCOEvaluator("category_validation",
distributed=True, tasks=None, output_dir='/content/output/')

```

```

#evaluator = COCOEvaluator('category_validation', ["bbox"],
distributed=False,output_dir='/content/output/')

evaluator = COCOEvaluator('category_validation', cfg,False,output_dir='/Output directory/')

val_loader = build_detection_test_loader(cfg, "category_validation")

inference_on_dataset(predictor.model, val_loader, evaluator)

```

After these steps are done, next step is to inference and saving the cropped rod version. The stages are same for the other models too, using radiographs/sonograms.

Appendix B

On the appendix B section, the source code of the combined AI system to measure length from the radiographs are given. The initial steps from 1-3 are just the same since in these steps, the drive is mounted with the google colab, and the required dependencies are imported and detectron2 is installed. The following steps are given here for measuring the length of the MCGR rod.

4. Importing libraries for image processing

```

import numpy as np

import os, json, random

import cv2

from google.colab.patches import cv2_imshow

from skimage.filters import unsharp_mask

from skimage import img_as_ubyte

from skimage.filters import sobel,scharr,prewitt,Roberts

```

5. Image processing

```
a= cv2.imread("/directory/",0)

plt.imshow(a)

unsharped_image =unsharp_mask(a,radius = 3, amount =5)

img_8bit = img_as_ubyte(unsharped_image)

clahe= cv2.createCLAHE (clipLimit = 2.0, tileGridSize=(8,8))

cl_img= clahe.apply(img_8bit)

b = img_as_ubyte (cl_img)

img = cv2.cvtColor(b,cv2.COLOR_GRAY2RGB)

plt.imshow(img)
```

6. Loading the rod model, cropping the rod, and saving

with open('directory of the saved rod model', 'rb') as f:

```
testConfig = pickle.load(f)

testConfig

testPredictor = DefaultPredictor(testConfig)

o = testPredictor(img)

v = Visualizer(img[:, :, :-1],

                #metadata=microcontroller_metadata,

                scale=0.5,

                instance_mode=ColorMode.IMAGE_BW # removes the colors of unsegmented

pixels

)
```

```

v = v.draw_instance_predictions(o["instances"].to("cpu"))

plt.figure(figsize = (12,8))

plt.imshow(cv2.cvtColor(v.get_image()[ :, :, :-1], cv2.COLOR_BGR2RGB))

plt.show()

from google.colab.patches import cv2_imshow

import PIL

from PIL import Image

masks = np.asarray(o['instances'].pred_masks.to("cpu"))

item_mask = masks[0]

# Get the true bounding box of the mask (not the same as the bbox prediction)

segmentation = np.where(item_mask == True)

x_min = int(np.min(segmentation[1]))

x_max = int(np.max(segmentation[1]))

y_min = int(np.min(segmentation[0]))

y_max = int(np.max(segmentation[0]))

print(x_min, x_max, y_min, y_max)

# Create a cropped image from just the portion of the image we want

cropped = Image.fromarray(img[y_min:y_max, x_min:x_max, :], mode='RGB')

plt.figure(figsize = (12,8))

plt.imshow(cropped)

plt.show()

```

```
cropped.save("/content/rod.jpg")
```

7. Loading the 58 model, cropping the 58mm part, and saving the ratio

with open('/saved 58mm model directory/', 'rb') as f:

```
testConfig58 = pickle.load(f)
```

```
testConfig58
```

```
testPredictor58 = DefaultPredictor(testConfig58)
```

```
img1 = cv2.imread("/content/rod.jpg")
```

```
o1 = testPredictor58(img1)
```

```
v1 = Visualizer(img1[:, :, ::-1],
```

```
                #metadata=microcontroller_metadata,
```

```
                scale=0.5,
```

```
                instance_mode=ColorMode.IMAGE_BW # removes the colors of unsegmented  
pixels
```

```
            )
```

```
v1 = v1.draw_instance_predictions(o1["instances"].to("cpu"))
```

```
plt.figure(figsize = (12,8))
```

```
plt.imshow(cv2.cvtColor(v1.get_image()[:, :, ::-1], cv2.COLOR_BGR2RGB))
```

```
plt.show()
```

```
masks = np.asarray(o1['instances'].pred_masks.to("cpu"))
```

```
item_mask = masks[0]
```

```
# Get the true bounding box of the mask (not the same as the bbox prediction)
```

```

segmentation = np.where(item_mask == True)

x_min1 = int(np.min(segmentation[1]))

x_max1 = int(np.max(segmentation[1]))

y_min1 = int(np.min(segmentation[0]))

y_max1 = int(np.max(segmentation[0]))

print(x_min1, x_max1, y_min1, y_max1)

# Create a cropped image from just the portion of the image we want

cropped1 = Image.fromarray(img1[y_min1:y_max1, x_min1:x_max1, :], mode='RGB')

plt.figure(figsize = (12,8))

plt.imshow(cropped1)

plt.show()

print(y_max1-y_min1)

ratio=58/(y_max1-y_min1)

print(ratio)

cropped.save("/content/58mm.jpg")

```

8. Loading the magnet model, cropping the magnet, and measuring the length using ratio from 7.

```

with open('saved magnet model directory', 'rb') as f:

testConfigmagnet = pickle.load(f)

```

```

testConfigmagnet

testPredictormagnet = DefaultPredictor(testConfigmagnet)

img2 = cv2.imread("/content/58mm.jpg")

o2 = testPredictormagnet(img2)

v2 = Visualizer(img2[:, :, ::-1],

                #metadata=microcontroller_metadata,

                scale=0.5,

                instance_mode=ColorMode.IMAGE_BW # removes the colors of unsegmented
pixels

                )

v2 = v2.draw_instance_predictions(o2["instances"].to("cpu"))

plt.figure(figsize = (24,16))

plt.imshow(cv2.cvtColor(v2.get_image()[:, :, ::-1], cv2.COLOR_BGR2RGB))

plt.show()

masks = np.asarray(o2['instances'].pred_masks.to("cpu"))

item_mask = masks[0]

# Get the true bounding box of the mask (not the same as the bbox prediction)

segmentation = np.where(item_mask == True)

x_min2 = int(np.min(segmentation[1]))

x_max2 = int(np.max(segmentation[1]))

y_min2 = int(np.min(segmentation[0]))

y_max2 = int(np.max(segmentation[0]))

```

```
print(x_min, x_max, y_min, y_max)

# Create a cropped image from just the portion of the image we want
cropped2 = Image.fromarray(img2[y_min2:y_max2, x_min2:x_max2, :], mode='RGB')

plt.figure(figsize = (12,8))

plt.imshow(cropped2)

plt.show()

y_max2-y_min2

length=ratio*(y_max2-y_min2)

print(length)
```

Using the same approach, the saved models for US system could be opened and loaded. And the length measurement could be done too.