Convolutional Neural Network for Automatic Cobb Angle Detection

A Technical Report submitted to the Department of Biomedical Engineering

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Michael O'Hanlon

Spring, 2020.

Technical Project Team Members

Michael O'Hanlon

Worthley Burke

Vicki Spina

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Timothy Allen, Department of Biomedical Engineering

Abstract

Adolescent Idiopathic Scoliosis (AIS) affects up to 3% of children in the U.S., and it is the most common form of spinal deformation.[2] This disease develops in late childhood and results in abnormal lateral curvature of the spine. The current method of diagnosis for scoliosis involves a manual measurement of a patient's Cobb angle, which denotes degree of spinal curvature. Scoliosis is defined for spines with at least a 10° Cobb angle.[2] The Cobb angle is measured from an X-ray of the patient's spine. This process is currently prone to human error as the doctor uses a ruler, pencil, and protractor to make the measurement by hand. In this paper, a deep learning algorithm is used to automate this outdated method of AIS diagnosis. The algorithm first performs image preprocessing and spinal trace identification on a data set of de-identified patient spinal X-rays images. These images are then inputted into a convolutional neural network, along with their corresponding Cobb angle measures, to iteratively identify qualitative features of the images and ultimately predict the angle of spinal curvature. By automating Cobb angle measurement, this research aims to provide a more effective and efficient method to diagnose AIS. The machine learning algorithm ultimately achieved 32.03% accuracy in determining Cobb angles from spinal X-rays. Ultimately the automated convolutional neural network will increase accuracy of Cobb angle detection and save valuable time for spinal doctors by eliminating an outdated manual diagnostic method.

Introduction

Scoliosis develops most commonly in infancy or early childhood, although the primary age of onset for scoliosis is 10-15 years old. Scoliosis patients are required to make frequent visits to private physicians and an estimated 30,000 children are fitted with a brace while 38,000 undergo a dramatic and life-hindering spinal fusion surgery.[3] Scoliosis treatment is guided by the specific scoliosis type, the amount of growth a child has remaining, the degree of spinal deformity and its anticipated progression. A primary goal of spinal orthopedics is to diagnose and treat scoliosis early, in order to prevent the risk of curve progression and other related medical complications.[2] However, spinal fusion surgery is typically performed at the end of skeletal maturity, following the conservative methods of bracing, and presents its own risk of complications.[1] These current treatments are extremely painful and disabling to the patient's range of motion. As such, it is crucial for doctors to accurately measure the spine to determine the appropriate treatment.

The Cobb angle is a measure of lateral curvature of the spine.[6] It is currently used as the standard measurement to quantify and track the progression of scoliosis.[6] This measurement assists the doctor in determining what type of treatment is necessary. A Cobb angle of 10 degrees is regarded as a minimum angulation to define scoliosis.[4] Currently, measuring the Cobb angle is conducted as follows (shown visually in Figure 1)[3]:

1. Locate the most tilted vertebra at the top of the curve and draw a parallel line to the superior vertebral end plate.

2. Locate the most tilted vertebra at the bottom of the curve and draw a parallel line to the inferior vertebral end plate.

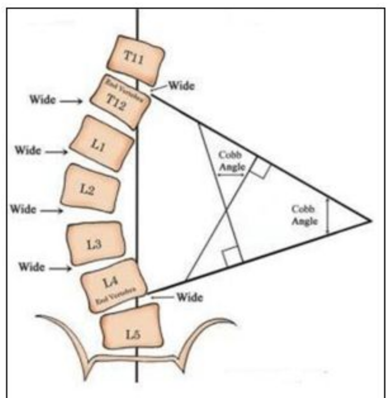


Figure 1: Manual Method for Measuring Cobb angle. This diagram shows how the Cobb angle is currently calculated. The method involves choosing the most deviated vertebrae in the upper and lower back, and then drawing the tangent lines to these two vertebrae. The angle of intersection of these lines is the Cobb angle (Safari et al., 2019).

3. Erect intersecting perpendicular lines from the two parallel lines
4. The angle formed between the two parallel lines is the Cobb angle.

From the steps described above, it is clear where potential discrepancies lie among the determination of the Cobb angle across various physicians, or observers. This method subjects the patient to a rather large degree of human error. For example, this human error could result from a discontinuity among physician skills, a skill-based error, or from a simple mistake or misread by the observer. Some specific examples of the described errors are a wrong definition of end/top vertebra, an incorrect drawing of lines through the endplates or through the pedicles, an incorrect drawing of the perpendiculars or the measurement of the angle itself.[4] Additionally, this method evidently takes a noteworthy amount of time. In one study physicians took 18.96 seconds to calculate the Cobb angle.[4] Additionally three million new cases of scoliosis are diagnosed in the United States every year.[5] This means that physicians across the United States spend 15,800 hours per year just measuring the Cobb Angle. This number is artificially low as it does not take into account remeasurement of the Cobb angle during patient checkups, and measurement of patients that have a curved back, but do not have scoliosis. It appears to be a rather simple measurement, but the multiple moving parts and involved calculations present as tedious and time-demanding in a physician's realm.

The proposed solution, a machine learning algorithm to detect and measure the Cobb angle, will not only provide a more accurate computational method to measure the Cobb angle, but will increase patient throughput by allowing doctors to spend less time focusing on a rather trivial task. Doctor's time is very valuable, and this algorithm would allow doctors to see more patients. Patients will also benefit from such technology as they will be more accurately diagnosed and therefore be provided with more accurate treatment and better care. The increased diagnostic accuracy will assist in preventing false positives that would lead to unnecessary and unwanted as well as transition the spinal deformity and scoliosis treatment platform towards becoming more minimally invasive.

To discuss related techniques that have been explored, in 2019 a group from Shiraz, Iran developed a new algorithm for determining Cobb angle of the spine.[11] This algorithm first requires the operator to identify every vertebrae by placing a point somewhere on the vertebrae. These points are then fitted to a 5th order polynomial curve. Two perpendicular lines are drawn normal to the inflection points. The Cobb angle between these two lines is measured. Even with this advanced algorithm the team was only able to achieve 81% accuracy,[11] leaving room for error.

The utilized technique described in this paper uses a deep learning algorithm to analyze the X-ray images and detect their Cobb angles. Specifically a convolutional neural network will be implemented as it excels at drawing data accurately from images.[5] The network will be trained by showing it X-rays and their corresponding Cobb angles. The algorithm will learn from these images and will iteratively become more accurate at analyzing the X-rays for the Cobb angle. Hopefully, the algorithm will be given enough test images to become more accurate than any of the prior art. Additionally, by using machine learning the algorithm will not rely on any operator action, eliminating human error and saving the doctors time.

Results

Image preprocessing results

After performing the stated preprocessing techniques the resulting images shown in Figure 2 demonstrated enhanced contrast, increased sharpness and decreased noise reduction around various pixel regions. The preprocessing steps were able to assist the convolutional neural network in feature detection

For example, MATLAB's second order statistic filter, or *ordfilt2* implements what is known as grayscale morphological erosion operation and this removes certain pixels on object boundaries, enlarging the lighter objects in image. Histeq assigns the intensity values of the pixels in the input image so that the output image contains a uniform distribution of intensities that spreads out the most frequent intensity values. This allows for areas of lower local contrast to gain a higher contrast, which more easily allows the algorithm to better detect these regions of the image. The preprocessing steps were able to assist the convolutional neural network in feature detection. The preprocessing techniques concluded to be successful, as demonstrated by the increased accuracy within the iterations in which preprocessing was performed. The preprocessing specific parameters were adjusted via trial-and-error attempts, choosing the parameters that provided the highest accuracy results upon subsequent iterations.

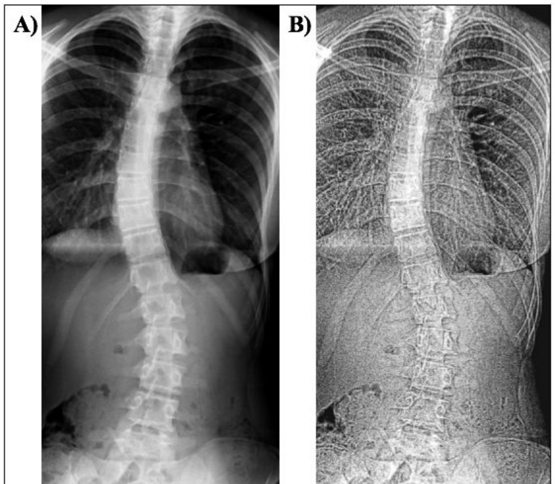


Figure 2: Image Preprocessing. The demonstrated image preprocessing techniques enhanced image contrast and increased sharpness and noise reduction. Figure 2a shows the raw X-ray image taken from our dataset. Figure 2b shows the same image after processing using several MATLAB filters and functions.

Spinal Trace Identification

While image preprocessing was able to significantly improve the accuracy of our algorithm, the images still contained a large amount of background noise. This noise consisted of other bones neighboring the spinal vertebrae, prominently the rib cage on the lateral edges, and the pelvis at the basal side of the image. While initial tests involved cropping the lateral sides of the images, this was found to cut off certain spinal vertebrae for patients with prominent spinal curvature. As a result, we eliminated the cropping step, replacing it with the spinal tracing method, shown visually in Figure 3.

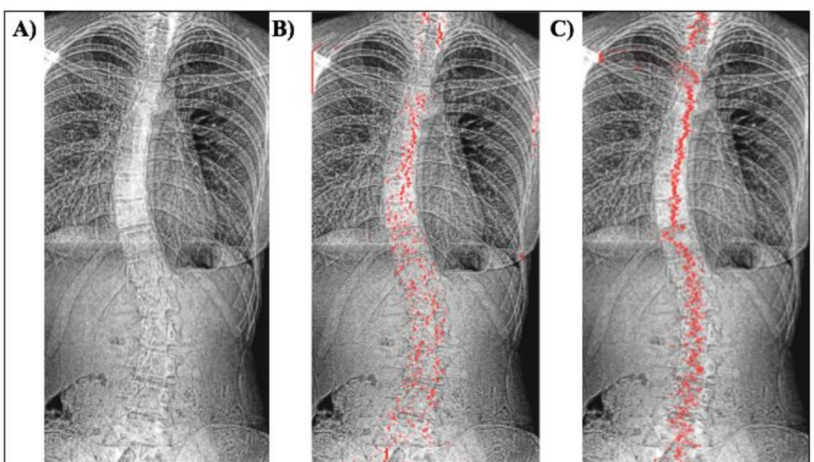


Figure 3: Spinal Trace Identification. After image pre-processing, the center of the spine is identified and marked. Figure 3a shows a spinal X-ray after image preprocessing, while Figure 3b shows the same image overlaid with spinal tracing markers denoting indicating the brightest pixel in each row. Figure 3c incorporates an innovated tracing method where the median of the 200 brightest pixel indices in each row are marked. Tracing the center of the spine allows the the convolutional neural network to more easily identify the spinal center when determining Cobb angle.

Spinal tracing involves identifying and marking the brightest pixel in each row of the image matrix as shown in Figure 3B. This brightest pixel was assumed to represent the center of the spine. In order to form a more continuous trace, this method was modified. For each row, the algorithm identifies the brightest 200 pixels and their indices within the row. The median of the spatial index is then marked, as shown in Figure 3C.

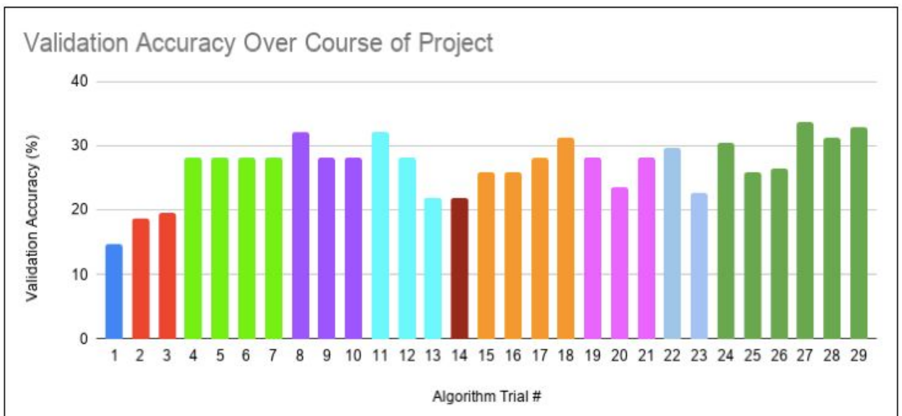
Finally, each of the images are scaled so that its dimensions are equivalent to the average width and height of each picture in the data set. After all images are made the same size, they can be inputted into the convolutional neural network, as described below. Overall our spinal tracing algorithm allowed for the convolutional neural network to more easily identify features specific to the spine.

CNN Results

We successfully developed an algorithm that classifies chest X-rays based off of the spine's Cobb angle. It classifies these images in bins of 10°. The first bin classifies images with Cobb angles greater than or equal to 5° and less than 15°. There are 6 more bins of increasing degree sizes that contain 10 degrees each. The last bin includes all Cobb angles larger than 75°.

To test our algorithm, we iteratively ran our algorithm with different combinations of layering as well as input images to achieve a higher validation accuracy. Our testing is shown in Figure 4 below. The validation accuracy of a neural network denotes how accurately it can classify images that it has not seen before. The training accuracy of a neural network denotes how accurately it can classify images that it has been trained on. On our first run, we were only able to achieve 14% validation accuracy. By using image preprocessing we were able to raise the accuracy to 33%.

We noticed that our network suffered from overfitting: when the validation accuracy is lower than the training accuracy. To rectify this we tried both dropout layers and weighting regularization. Dropout layers randomly remove some nodes from the neural network. Weighting regularization limits how much weight a certain node can gain after each training iteration. These two techniques help ensure that the neural network will not assign too much importance to non-defining characteristics. While these techniques did improve the validation accuracy of our neural network, they drastically reduced the training accuracy. They did however, solve the problem of overfitting.



Legend

Label	CNN Modification	Image Processing
(blue)	Basic Version	Force fit (no image preprocessing)
(red)	Basic Version	Different sized spinal traces
(green)	Dropout Layer	Spinal Trace 185x705
(purple)	Basic Version	Different sized spinal traces
(cyan)	Weighting Regularization	Spinal Trace 62 x 235
(dark red)	Weighting Regularization	Spinal Trace 185 x 705
(orange)	Dropout Layer + Weighting Regulatization	Spinal Trace 62 & 185
(pink)	Dropout Layer + Weighting Regulatization	New Sprinal Trace Method 388x145
(light blue)	Weighting Regularization	New Sprinal Trace Method 388x145
(dark green)	Basic Version	Image Processing Modification

Figure 4: Validation Accuracy for Every Trial. This graph shows the validation accuracy achieved for every trial during our research. For each trial one parameter was changed. The color coding shows which groups of parameters were altered. Specific modifications for each group of trials are explained in the legend.

We did not focus on the loss of our neural networks as they all functioned similarly. Loss shows the error of our neural network in classifying images. For all of our trials our loss sharply decreased within the first epoch of training and then remained low.

After completing 29 iterations, it was decided that iteration number 11 provided the best result. Trial number 11 combined a neural network that used weighting regularization and version one Spinal Trace images sized 62 pixels wide by 235 pixels high. The graph showing the validation accuracy from Trial 11 is shown below. This neural network was chosen for three reasons. Firstly, it had one of the highest validation accuracies at 32.03%. Secondly, the validation accuracy steadily trends upwards through each iteration. Theoretically, given a larger database, the validation accuracy would continue to rise. Lastly, the training accuracy for this algorithm is still relatively high. Given a more accurate dataset, the validation accuracy should more closely follow the training accuracy

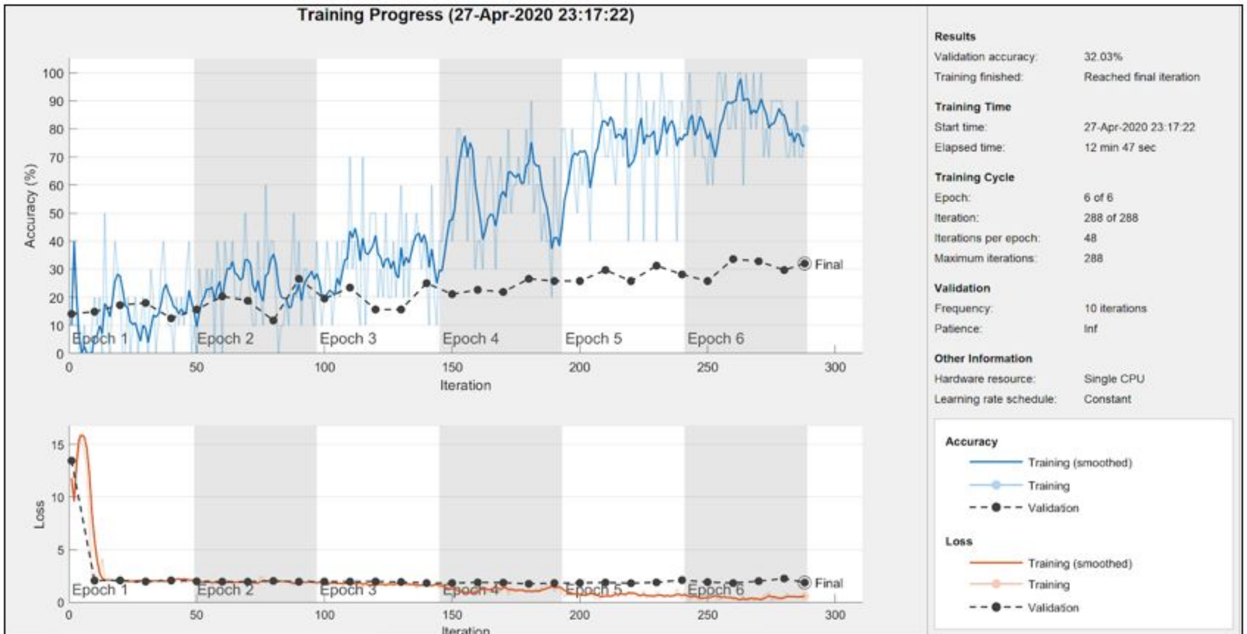


Figure 5: Convolutional Neural Network Trial Number 11. This graph shows the change in accuracy of our CNN throughout its training iterations.

curve.

Discussion

We chose to use the MATLAB Deep Network Designer to simplify the creation of the neural network. MATLAB offers a variety of toolboxes that simplify the development of a convolutional neural network. We used our specific dataset upon MIST's request. Upon conducting research to discover the best image preprocessing techniques for image feature detection on X-ray images, the previously listed MATLAB functions were discovered and therefore integrated in this research. As the bone is highlighted as white in an X-ray image, it was decided that it would be best to trace the spine for identification by locating and marking the brightest pixels along the length of the spinal curve.

Minimally Invasive Spinal Technology (MIST), a Charlottesville, Virginia based medical device startup, has constructed a minimally invasive spinal implant to treat ADS, adult degenerative scoliosis. Although the device represents an alternative treatment method for scoliosis, the Cobb angle measurement algorithm will allow them to more precisely decide which patients need treatment. Additionally, this technology would allow for device adjustment, such as device height alterations, based on the specific measurement degree of the Cobb angle.

We had several limitations during our research that stemmed from our data set. The data set we used only had 608 images. This is a small dataset to train machine learning algorithms, and resulted in a lower validation accuracy. With a larger dataset, our algorithm would be more accurate. Due to the lack of a comprehensive range of various Cobb angles within this dataset, we were not able to have smaller classifications. In an optimal environment, we would have liked to have a classification for every degree from 0 to 90. Additionally, the labeling of our dataset images was inaccurate. This inaccuracy is inherent in cobb angle detection as the angles were determined by a manual method. In the dataset we found images of the same spine, that were labeled with cobb angles varying as much as 10°. By providing our algorithm with an inaccurate data set, it was impossible to achieve a high accuracy. Our dataset was also biased by not containing any X-ray images of straight spines. If given an X-ray image of a straight spine, our dataset would likely inaccurately classify it, as it was not trained with straight spine images.

As this research project addresses scoliosis and X-ray imaging, biomedical ethical issues arise within the fields of orthopedics and radiology. Within the scoliosis industry, the decision to have surgery is not typically a matter of life or death. As a result, the patient’s deformity may not always be corrected, as in the event of minimal funding or lack of insurance coverage. Families with lower income may not be able to afford treatment, causing an unethical divide between the rich and the poor in terms of accessibility to spinal care.Within the radiology field, patients are concerned about the security of their data. For many different conditions, patients will receive some sort of imaging as part of their diagnosis. When images are taken, they are stored in the patient’s electronic medical record. These files should not be shared without the patient’s permission. Data privacy is an important ethical issue, especially in research studies that require large patient datasets in order to accurately draw conclusions. The Health Insurance Portability and Accountability Act (HIPAA), passed by congress in 1996, mandated the opportunity for people to privatize and protect their personal healthcare data. People’s data should be protected. If it were not safe, patients would not trust doctors with their data. Such a lack of trust could undermine doctor-patient confidentiality and could be detrimental to the U.S healthcare system as a whole.

The main ethical issue that will result from our project is the privacy of data. Our machine learning algorithm will be improved by additional X-ray images of the spine. An increase in learning images will allow the algorithm to be more accurate and provide patient care; however, we cannot just directly use the images of our patients. We will have to ask for their permission to incorporate their deidentified images into our algorithm. We will ask doctors to explain the process and ask them if they would like to sign a release form during their checkups. Our product simply analyzes images. The main biomedical issue that could arise would be over-exposing our patients to radiation from X-ray images. The current method used to measure the Cobb Angle already uses X-ray images, so this would be no worse than the prior art. To minimize radiation levels we would restrict taking X-ray images as much as possible. When we would take the X-ray images we would use lead coverings to protect the rest of the body. We would also prioritize using X-ray images that already exist, to avoid the need to take an abundant amount of X-rays and risk radiation exposure.

Another ethical issue that could arise surrounds the actual implementation of our technology into hospitals. It will likely be costly to expand our software beyond the prototype phase. In order to be approved by hospitals, our algorithm will have to be heavily refined, tested, and trained so that it can consistently measure the angle of the spine with utmost accuracy. This process will raise the price of our product. These costs may not be affordable for smaller hospitals with less funding, causing a potential unethical divide in access to the best scoliosis diagnosis regimens .However, to address the apprehensiveness to adopt a new procedural method, the future plan would be to provide brief training sessions for physicians. The training sessions will familiarize them with the technology and give them the confidence they need for their medical performances.

For future studies, it is important to obtain a larger sum of data. Utilizing more data within similar research, has the potential to produce higher accuracy results, therefore contributing to the novel impact of this technology. Convolutional neural networks require large sets of data since there are a vast number of parameters that need to be adjusted during the training phase. At a bare minimum, for machine translation, high dimensional data generation, or anything requiring deep learning, it is suggested to obtain about 100,000-1,000,000 samples.[12] Working closely with MIST, it is hoped that in the future we can combine our algorithm with MIST's spinal implant. Specifically, we would correlate the measured Cobb angle to angular adjustments of the implant. For example, the height of the implant could be adjusted with a larger/smaller cobb angle. This approach would allow for patient-specific customizable scoliosis treatment opportunities. Further, we hope to eventually reach the opportunity to integrate this algorithm into hospitals. Upon doing so, there would be the requirement to create a user interface in which the physician can easily interact with. Additionally, we would need to schedule consultations with physicians and develop training sessions to ensure a smooth transition from the current manual method to this innovative automated computational approach.

Materials and Methods

Database

The database of spinal X-ray JPEG images and corresponding comma-separated values (CSV) labels were obtained from MIST. The images are from real patients who have been de-identified. The database consisted of a training set and testing set of images in order to both train and test the algorithm on a wide range of images. The training set contained 480 images and the testing set contained 128 images, summating to a total of 608 images. The images were labeled using a manual computerized method of Cobb angle detection. We were not given more information on this database.

MATLAB

MATLAB was used as the primary software platform for our algorithm. MATLAB's image processing functions were helpful in preprocessing our X-ray images, as described below. MATLAB's Computer Vision toolbox was used to help mark the spine during spinal trace identification, and the Deep Network Designer was used to design the architecture of the convolutional neural network.

Preprocessing

Before inputting the images into the algorithm, the images underwent preprocessing using MATLAB functions. Image preprocessing was performed in order to assist the convolutional neural network in image feature detection. The MATLAB functions performed on each image are summarized in Table 1.

MATLAB Function	Description
imadjust	Imadjust saturates the bottom 1% and the top 1% of all pixel values. This operation increases the contrast of the output image.
histeq / adapthisteq	Adapthisteq enhances contrast of image by transforming image intensity values using contrast-limited adaptive histogram equalization (CLAHE). Histeq works on the entire image while adapthisteq works on small regions in the image. Adapthisteq combines these small regions using bilinear interpolation to eliminate artificially induced boundaries. More simply put, it evenly distributes image intensity values.
ordfilt2	Ordfilt2 computes order-statistic filters, or rank filters. These are nonlinear spatial filters whose response is based on ranking the pixels within an image region and then replacing the value in the center pixel in the region with the value determined by the ranking result. Basically, it performs grayscale morphological erosion operation that removes pixels on object boundaries, enlarging light objects in the image.
fspecial / imfilter	Fspecial is used in conjunction with imfilter and creates an unsharp contrast enhancement filter by using the negative of the Laplacian filter. This unsharp filter is an image sharpening operator

Table 1: Image Preprocessing Techniques. The functions above were used to improve the convolutional neural network's ability to identify spinal features within the X-ray images.

The result of these functions were enhanced contrast of the image, increased sharpness and decreased noise around pixel regions.

Spinal Trace

MATLAB was further used to conduct spinal tracing identification on the images after they were subject to image preprocessing. Spinal tracing was performed iteratively for every image in the dataset. For each row of the image our algorithm creates a matrix of the 200 brightest pixels alongside their corresponding spatial index within the row. The algorithm then takes the median of these 200 indices. This point is determined to be the center of the spine. MATLAB's insertMarker function is used to mark these points with a red circle, which was determined through trial and error to be the best shape and color combination in order to maximize validation accuracy. Finally, the algorithm computes the average pixel width and height of every image. Each image is then resized so that its dimensions match that of the average width and height. This is done so that every image is the same size, a requirement to be inputted into the convolutional neural network.

Convolutional Neural Network

We built our neural network with MATLAB's Deep Network Designer add-on. We cannot discuss the layering of the neural network to protect MIST's intellectual property.

References

1. Why the Cobb Angle Is Used to Diagnose Scoliosis [Internet]. Verywell Health. [cited 2020 Apr 30]. Available from: https://www.verywellhealth.com/cobbs-angle-and-scoliosis-296583
2. Scoliosis – Symptoms, Diagnosis and Treatment [Internet]. [cited 2020 Apr 30]. Available from: https://www.aans.org/
3. Michael R. Orthopedic Clinical Examination. Human Kinetics; 2015. 1152 p.
4. Measurement of scoliosis Cobb angle by end vertebra tilt angle method | Journal of Orthopaedic Surgery and Research | Full Text [Internet]. [cited 2020 Apr 30]. Available from: https://josr-online.biomedcentral.com/articles/10.1186/s13018-018-0928-5
5. Facts about Scoliosis Every Parent Should Know | Johns Hopkins Medicine [Internet]. [cited 2020 Apr 30]. Available from: https://www.hopkinsmedicine.org/health/conditions-and-diseases/scoliosis/5-facts-about-scoliosis-every-parent-should-know
6. Core Concepts Pte Ltd. (2019, March 30). Understanding Cobb Angles and what it means for Scoliosis. Retrieved from https://www.coreconcepts.com.sg/article/cobb-angle-and-scoliosis/.
7. Playbook, H. S. S. (2015, December 18). The Importance of Early Detection of Scoliosis in Children. Retrieved from https://www.hss.edu/playbook/the-importance-of-early-detection-of-scoliosis-in-childre
8. Scoliosis. (n.d.). Retrieved from https://www.aans.org/patients/neurosurgical-conditions-and-treatments/scoliosis.
9. Gstoettner, M., Sekyra, K., Walochnik, N., Winter, P., Wachter, R., & Bach, C. M. (2007). Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools. *European Spine Journal*, *16*(10), 1587–1592. doi: 10.1007/s00586-007-0401-3
10. Zhang, Q. (2018). Convolutional Neural Networks. *3rd International Conference on Electromechanical Control Technology and Transportation*. doi: 10.5220/0006972204340439
11. SHARMA S. Epoch vs Batch Size vs Iterations [Internet]. Medium. 2019 [cited 2019 Sep 22]. Available from: https://towardsdatascience.com/epoch-vs-iterations-vs-batch-size-4dfb9c7ce9c9
12. Rusom, P. (2018, September 14). Data Requirements for Machine Learning. Retrieved from https://tdwi.org/articles/2018/09/14/adv-all-data-requirements-for-machine-learning.aspx