Group 4: Automatic Cobb Angle Measurements for Scoliosis Patients

Siddhi Pandare
M.S. ECE

Brett Baxley
M.S. ECE

Suman Ghosh
M.S. ECE

Zexuan Liu
M.S. BMED

Contribution Table

Member Names	Contribution
Siddhi Pandare	Model selection, Landmarks using centroids method, Results for centroid method
Brett Baxley	Image pre-processing, 4 Point Landmark Image Segmentation CNN, Results for 4 landmarks method
Suman Ghosh	Image pre-processing, Landmark to Cobb's angle, Centroids to Cobb's angle, Result validation
Zexuan Liu	Model selection, GUI implementation

Automatic Cobb Angle Measurements for Scoliosis Patients

Brett L. Baxley
Electrical and Computer Engineering
Georgia Institute of Technology Atlanta Georgia
bbaxley7@gatech.edu

Siddhi V. Pandare
Electrical and Computer Engineering
Georgia Institute of Technology Atlanta Georgia
spandare3@gatech.edu

Suman Ghosh
Electrical and Computer Engineering
Georgia Institute of Technology Atlanta Georgia
sghosh323@gatech.edu

Zexuan Liu
Biomedical Engineering
Georgia Institute of Technology Atlanta Georgia
zliu884@gatech.edu

ABSTRACT

Scoliosis is a sideways curvature of the spine that is most often diagnosed in adolescents [1]. Scoliosis severity is tracked with a measurement called the Cobb Angle. Measurements are normally done manually by clinicians, which is time-consuming and subject to human errors due to inter- and intra-observational variations. Automating this process allows objective quantification of the Cobb angles and can drastically reduce time and error from medical professionals, while allowing for changes in spinal curvature to be monitored consistently. We propose an automated Graphical User Interface system for clinicians that estimates Cobb angles for patients with scoliosis and determines disease severity upon uploading an x-ray image file. We would address this problem in the domain of Cognitive AI (Artificial Intelligence) to evaluate the performance of the decision-making algorithm in comparison to human decision making.

CCS CONCEPTS

• Deep-Learning • Cognitive AI

KEYWORDS

Scoliosis, Cobb Angle Measurement, Deep Learning, User Interfaces, Image Pre-processing

1 Problem Statement

Scoliosis is commonly diagnosed in teenagers as a sideways curvature. If left untreated, the severity of the patient's condition can worsen. Quantifying the Cobb angles on a curved spine allows clinicians to provide more accurate diagnoses and suitable treatment plans for patients with scoliosis. This measurement is taken by finding the most tilted vertebrae above and below the apex vertebrae, which is the vertebrae that is located at the deepest part of the scoliosis curve. The intersection of the slopes of the 2 found vertebrae are then used to calculate the Cobb Angle of the spine. The severity of scoliosis is related to the patient based on the following criteria. A mild diagnosis is a Cobb angle less than 20 degrees. A moderate case measures an angle between 20 and 40 degrees, and a severe case is above 40.

Traditionally, this measurement is taken manually by surgeons, which is tedious and can lead to errors and inconsistencies. Automating this entire process can drastically reduce time required and inconsistencies from the man-in-the-middle approach. We propose to create an entire user interface program that allows surgeons to upload an x-ray image and automatically view the Cobb's angle of the patient, along with severity diagnosis.

2 Background and Related Work

Although automating Cobb's angle measurements could enhance efficiency greatly and streamline the process, the computer-based image analysis required for this task is quite challenging. Processing x-ray images presents numerous difficulties, such as dealing with dynamic ranges of high exposures and varying contrasts between x-ray images, accounting for inconsistent levels of adipose tissue content that may affect image quality, and detecting implants that can appear in the images. In [2] authors proposed filtering the spine out of the image via Canny Edge Detection, which resulted in a mean absolute error of 6.6°, but still ran into issues with image noise. The noise in distinct parts of the image impacted the accuracy for different cases of spinal curvature. While [3] used K-means and curve fitting but was not resistant to noisy and low contrast images. As a result, challenges arise for accurately segmenting and processing images to obtain precise results.

3 Methodology and Approach

As shown in Figure 1, the approach implemented involves pre-processing the image, detecting spinal landmarks, measuring Cobb's angle based on the predicted landmarks, and finally validating the results. The following sections will explain in further detail the methodologies and algorithms used to predict the measurements

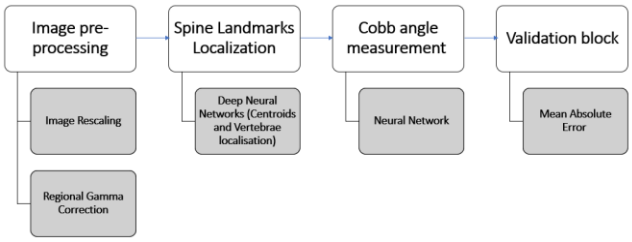

Figure 1: System Flow diagram

3.1 Dataset Description

The dataset we used is the Dataset 16 spine dataset [4] which has 609 spinal anterior-posterior x-ray images. The dataset is split into 481 training images and 128 test images. The images include both S and C scoliosis images with body implants. Most of the images have low contrast with varying sizes. The images come with 4 landmarks for each spine vertebrae annotated by professionals. The ground truth Cobb's angle is given by three angles, out of which the largest is considered as the main Cobb's angle as it defines the maximum bend in the spine. It takes at least about 10° deviation from the vertical straight axis before scoliosis is defined.

3.2 Image Preprocessing

Initial attempts to preprocess images involved iterative gaussian filtering techniques to extract the spinal features from the image. Ideally, it could be assumed that the outer edges of the X-Ray images were of lower intensity, due to the lack of detection. Blurring the images iteratively would allow for the average pixel intensities near the center of the image to stay consistent, while the edges and their neighboring pixels would be driven to lower intensities due to averaging. As shown in Figure 2, this did not prove successful due to the processes' inability to properly extract the spinal feature for every image. Some images had spinal pixel

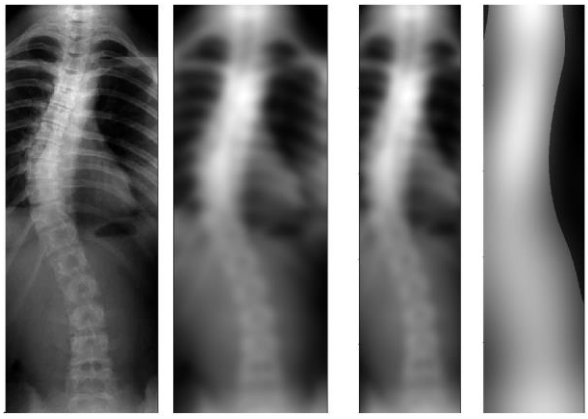

Figure 2: Pre-processing image by blurring

intensities that were identical to the intensities in other areas of the X-Ray, therefore extracting more features than necessary.

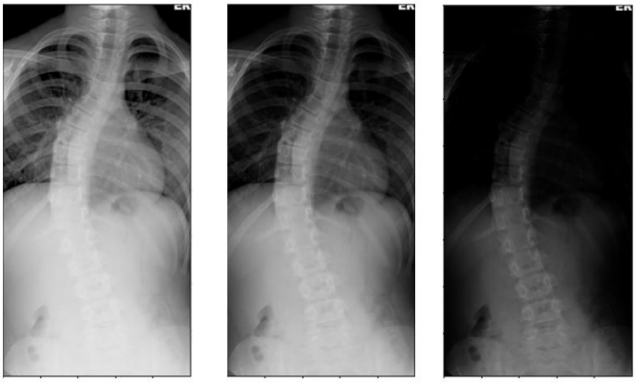

Figure 4: Different Gamma Values for Pre-processing from left to right, gamma = 0.75, 1, 3

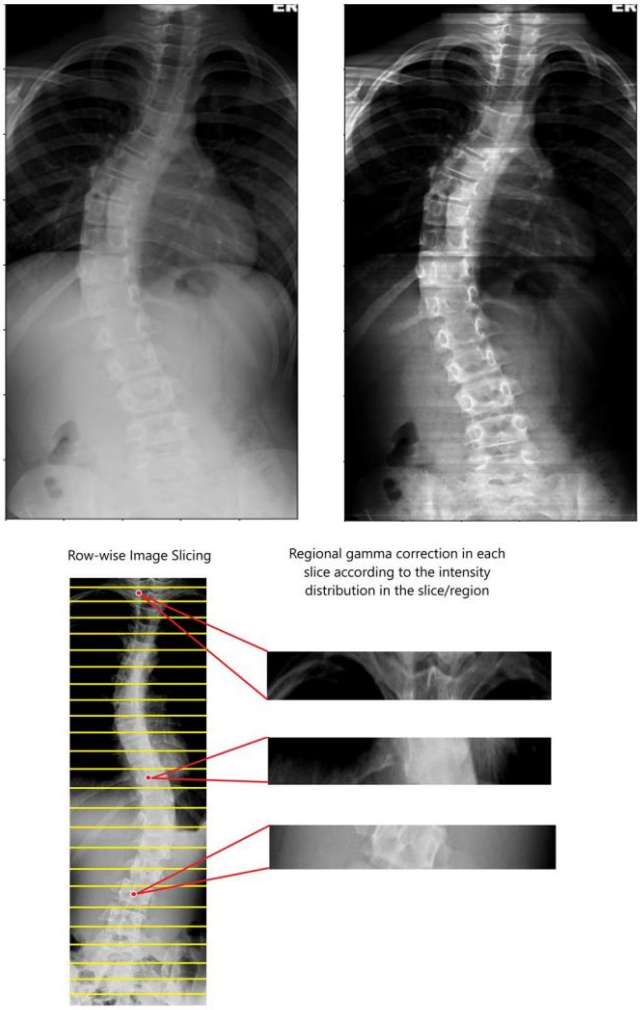

Figure 5: Regional Gamma correction applied to X-ray image

The image preprocessing technique that was selected involved Regional Gamma Correction. Images were enhanced and then resized to standardize the input into the deep learning architecture. In [5], gamma correction was applied to enhance the images for deep learning, by suppressing noise developed from soft tissues, and adjusting the contrast of the spinal structure. When applying

this method to the proposed dataset, different gamma levels enhanced certain features, and restricted others that are necessary for spinal segmentation, Figure 4. To resolve this issue across images, a regional gamma correction method was applied (Figure 5). The images are divided into multiple sections by rows, and the gamma value used for transformation is calculated based on the average pixel intensity for each section. After the image enhances spinal features, the images are resized to 512x256 pixels.

3.3 Landmark Predictions

Two approaches were attempted to detect landmarks for spinal segmentation. The first approach was referenced from [6], and involved detecting four landmarks, one for each corner for each vertebra on the spine. A deep neural network with the architecture in Table 1.

The network was trained for 100 epochs with an Adam Optimizer with a Learning Rate of 0.0005. A Mean Squared Error Loss function was used for backpropagation, and the batch sizes were 32 images. The training and testing loss plots are shown in Figure 7. The results from the following neural network were not ideal. Figure 6 demonstrates some of the pitfalls using this technique. The network generalized the spinal curve on some images but could not accurately predict landmarks. Some test cases were unable to predict the spinal curvature at all.

Layer	Input	Kernel Size	Padding	Stride	Output
Conv1	1x256x128	5	3	1	64x258x130
Maxpool1	64x258x30	2	0	2	64x129x65
Conv2	64x129x65	3	1	1	128x129x65
Maxpool2	128x129x65	2	0	2	128x64x32
Conv3	128x64x32	3	1	1	256x64x32
Maxpool3	256x64x32	2	0	2	256x32x16
Conv4	256x32x16	5	2	1	512x32x16
Maxpool4	512x16x8	2	0	1	512x16x8
FC1	65536	-	-	-	10000
FC2	10000	-	-	-	5000
FC3	5000	-	-	-	2500
FC4	2500	-	-	-	512
FC5	512	-	-	-	136

Table 1. 4-Landmarks Detection Neural Network Architecture

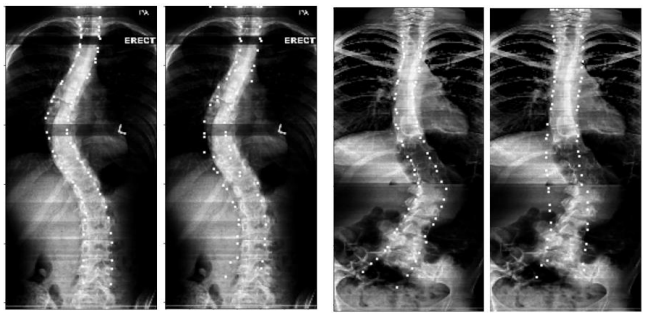

Figure 6: Ground Truth Images (Left Images) and Predicted Values (Right Images)

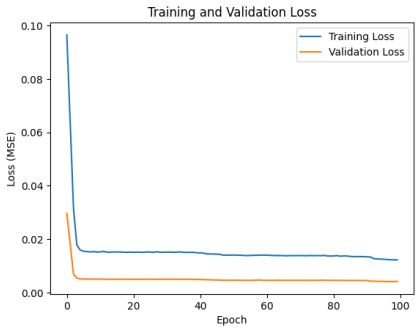

Figure 7: Training and Validation loss for 4 landmarks model

3.4 Centroids Localization

The second model outputs the centroids of the spine rather than the 4 landmarks as discussed in Section 3.2. The intuition behind modifying the architecture was to decrease the number of input and output parameters. Moreover, we modified the loss function to penalize the error in standard deviation of the abscissa values of a landmark since the abscissa values of the centroids remain in a small range of values. The loss function used is given in equation (1).

$$Loss = \lambda * Landmark_{error} + (1 - \lambda) * (Abscissa\ Displacement_{error}) \quad (eq\ 1)$$

The deep neural network architecture given in Table 2 was trained for 100 epochs with Adam Optimizer and the learning rate was set to 0.0005. The lambda in the loss function was set to 0.7. The training and testing loss after every epoch are plotted in Figure 8. From the results in Figure 9, the model predicts the spine curvature better than the one described before. The evaluations and comparisons are mentioned in the forthcoming sections.

Layer	Input	Kernel Size	Padding	Stride	Output
Conv1	1x256x128	7	3	1	32x256x128
Relu1	32x256x128	-	-	-	32x256x128
Maxpool1	32x256x128	2	0	2	32x128x64
Conv2	32x128x64	3	1	1	64x128x64
Relu2	64x128x64	-	-	-	64x128x64
Maxpool2	64x128x64	2	0	2	64x64x32
Conv3	64x64x32	3	1	1	128x64x32
Relu3	128x64x32	-	-	-	128x64x32
Maxpool3	128x64x32	2	0	2	128x32x16
Conv4	128x32x16	5	2	1	256x32x16
Relu4	256x32x16	-	-	-	256x32x16
Maxpool4	256x32x16	2	0	2	256x16x8
FC1	32768	1	0	1	512
Relu5	512	-	-	-	512
FC2	512	1	0	1	512
Relu6	512	-	-	-	512
FC3/output	512	1	0	1	34

Table 2: Centroid Landmark Detection Neural Network Architecture

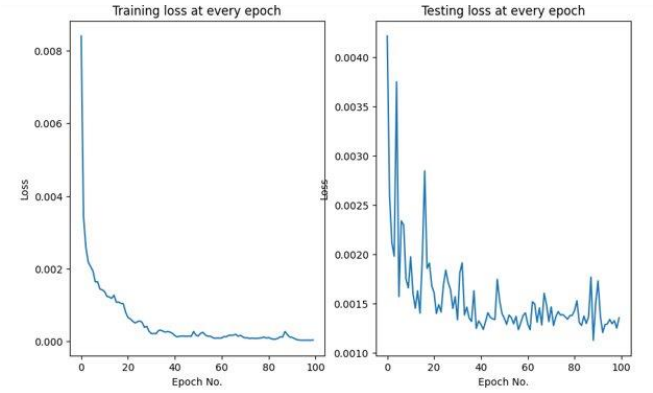

Figure 8: Training and Validation Data for the Centroid Landmark Detection Neural Network

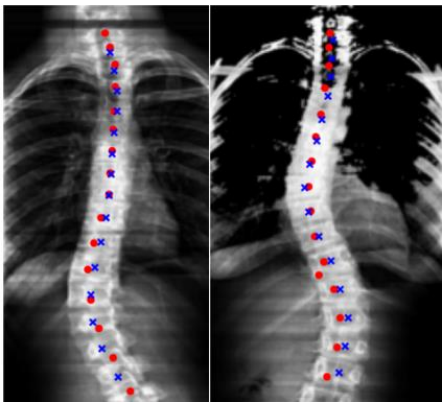

Figure 9: Ground Truth Centroid Values (Blue crosses) and Predicted Values (Red Dots) on test images

3.5 Cobb Angle Measurements

3.5.1 Methodology

Each X-ray image of the spine is linked to the three Cobb's angles. Three different Convolutional Neural Network (CNN) architectures were employed in this study. The first architecture predicts all three angles simultaneously, while the second architecture predicts one angle at a time. The third architecture predicts the first and second angles together, then the second and third angles together, and finally calculates the mean of the two predictions for the second angle.

Tuning only Architecture 1 did not give desirable results, which led to Architecture 2 predicting all angles separately from the input features, and the idea that the 2nd angle to be predicted might have some relationship with the 1st and 3rd angle, gave rise to Architecture 3. Architecture 3 came up because the 2nd angle was harder to predict than the other 2 angles, and there is a relation between the 1st and 2nd and 2nd and 3rd angles' sets. Every model of every architecture pattern was trained with noisy inputs where we manually added noise to every input data point before training the CNN architectures used for the calculation of Cobb's angle. The number of neurons in the first hidden layer for a particular model in each layer was kept at around 2/3 of the number of inputs added to the number of outputs for the model, as per the standard starting points of tuning model hyperparameters selected. For the following hidden layers, the number of neurons was half of the previous hidden layer.

All the 3 CNN architectures were considered for each of the 4-point landmark and centroid based segmentation approaches and their performances were compared. CNN Architecture 3 gave the best results overall when combined with the centroid-based segmentation approach.

In each of the models, Adam optimizer was used with a learning rate of 0.0001, the exponential decay rate for 1st moment estimates was 0.85 exponential decay rate for the 2nd moment estimates was 0.99, constant for numerical stability was 1e-10 along with mean squared error loss and mean absolute error loss tracking. A batch size of 32 and validation split of 0.3 was used during the training on the 4-point landmarks or centroids data from the test set images. The number of epochs was 1000 for Architecture 1 model and 10000 for each Architecture 2 and 3 models.

3.5.1.1 CNN Architecture 1

This standard architecture (Figure 10) consists of a single neural network model with 3 densely connected layers. It takes either 136 input values for 4-point landmarks or 34 input values for centroids. The model generates three output values that correspond to the three angles calculated by the Cobb's Angle calculation method.

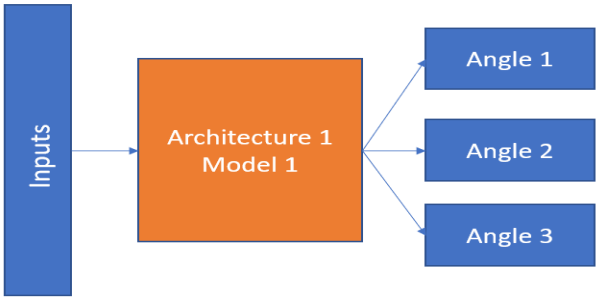

Figure 10: CNN Architecture 1

3.5.1.2 CNN Architecture 2

There are three similar models (Figure 11), each comprising 3 densely connected layers. Each model takes either 136 input values for 4-point landmarks or 34 input values for centroids. The output for each model is a single Cobb's angle of the set of 3 Cobb's angles. Therefore, there are three separate models that take the same inputs and produce a single output each, and the only difference is the specific angle being calculated.

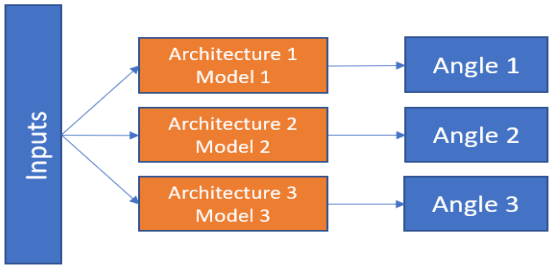

Figure 11: CNN Architecture 2

3.5.1.3 CNN Architecture 3

In this architecture we have two similar models, each consisting of 3 densely connected layers. The inputs for the first model are 136 values corresponding to 4-point landmarks, while the inputs for the second model are 34 values corresponding to centroids. Both models have an output of 2 angles. First model predicts the 1st and 2nd angle while the 2nd model predicts the 2nd and 3rd Cobb's angles. The mean value is taken for the 2nd Cobb's angle prediction as shown in Figure 12.

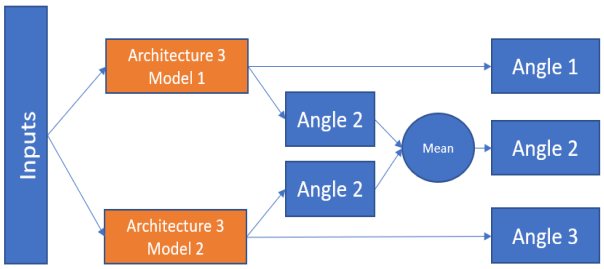

Figure 12: CNN Architecture 3

4 Results and Analysis

4.1 Hyperparameter tuning

To ensure that the values used in [6] were justified to use in the models described in sections 3.2 and 3.3, the observed training loss at the 20th epoch for a subset of the dataset is plotted in Figure 13. The acceptable combination of the optimizers with learning rates are Stochastic Gradient Descent (SGD) and Adam Optimizers with learning rate set to 0.0005 or 0.00025. However, to converge faster we chose the learning rate as 0.0005. Thus, we see both Adam and SGD can be used during the training of the models.

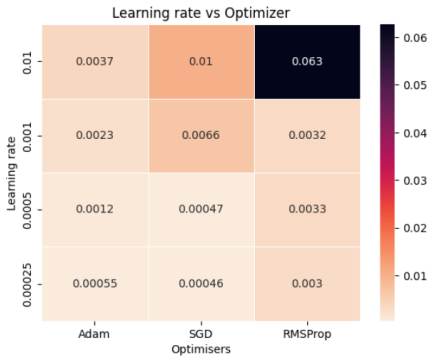

Figure 13: Heatmap of different learning rate versus Optimizers

4.2 Prediction error plots

4.2.1 Prediction error for centroids model

In Figure 14, we see more outliers (wrong prediction values) for ordinate values than abscissa values. Thus, the abscissa values are better predicted than the ordinates values. This could be because of the loss function given in equation 1 which penalizes the error of deviation of the abscissa values.

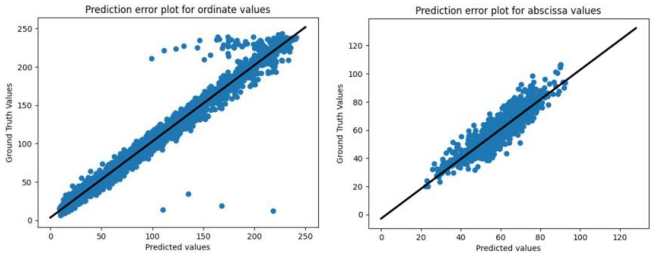

Figure 14: Prediction error plot for ordinate and abscissa values for centroids model

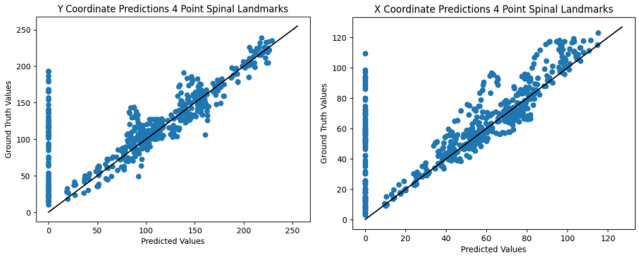

Figure 15: Prediction error plot for ordinate and abscissa values for 4 landmark model

4.2.3 Predictability Plots

The comparison of the two models' prediction for the x and y co-ordinates of the landmarks is given in Figures 15 and 16. The centroid model performs much better as both the Mean Square Error and Mean Absolute Error in predicting the x and y coordinates is lower than that of 4 landmarks detection model. Since we need accurate spine curvature detection, we thus conclude that the regional gamma correction along with the centroid detection-based approach is much suited for calculating Cobb's angle.

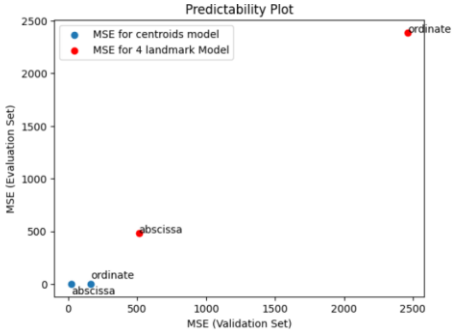

Figure 15: Comparison of Mean Square Error (MSE) of abscissa and ordinate values for centroids and 4-landmarl model

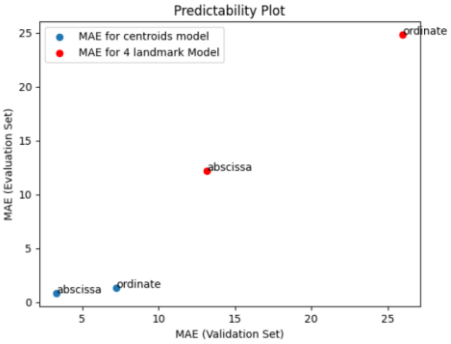

Figure 16: Comparison of Mean Absolute Error (MAE) of abscissa and ordinate values for centroids and 4-landmarl model

4.2.4 Landmarks to Cobb's angle results

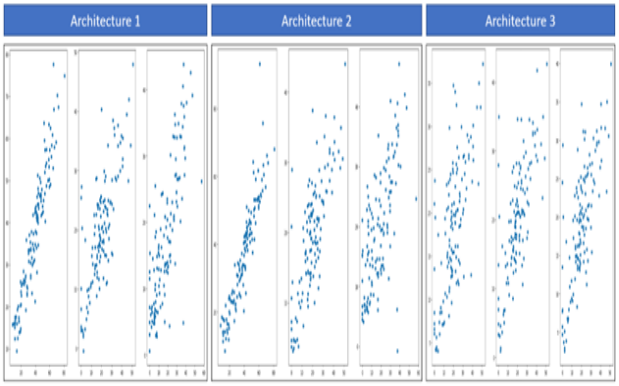

Figure 17: Predicted Vs. Truth Results for 4-point Landmark Detection Neural Network Cobb Angle Calculation

Architecture 1 and 2 showed impressive performance for the 1st angle, and satisfactory performance for 3rd angle, but the 2nd angle was the most difficult for the CNN architectures to learn. So, the 3rd architecture solved this in a way that it produced balanced performance for all the 3 angles, thus making the system robust and more capable of accurately predicting the set of 3 angles together. Performances were measured in terms of mean absolute error for every case described above. It was found that the centroids approach performed better than the 4-point landmark approach across all the 3 architectures as seen from Figure 17 and Figure 18, given the limited number of epochs the models were trained for. The lesser number of outlier values and clumping of points along a straight line which is the ideal case, can be noted for 3rd Architecture in the case of centroids in Figure 18. So, at the current state the proposed system would comprise of a centroid-dependent segmentation of regional-gamma corrected X-ray images for measuring the Cobb's angles with the CNN Architecture 3 proposed here. The predictions can thus be visualized and classified based on the known thresholds for scoliosis severity using a graphical user interface (GUI) which would be of great assistance to a medical professional.

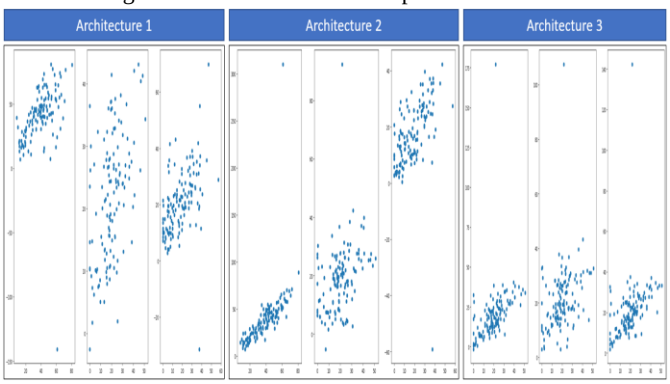

Figure 18: Predicted vs. Truth Results for Centroid Localization Detection Neural Network Cobb Angle Calculation

Mean absolute error is usually the metric used to evaluate the performance of the model in literature [2][3][4][5]. Thus, the mean absolute error while calculating the Cobb's angle for the entire process pipeline using centroids was ~ 9 degrees.

5 Graphical User Interface

A GUI was developed to aid the diagnosis of scoliosis. The system is automated and allows the clinician to upload a single jpg or jpeg file of an x-ray image of the spine. Within a second, three images, including the original x-ray image ("Uploaded Image"), the gamma-corrected image ("Preprocessed Image"), and landmarks (white circles) and centroids (red circles) of each vertebra overlaid onto the gamma-corrected image ("Predicted Image") as well as calculated Cobb angle will be displayed on the same window (Figure 19). The largest Cobb angle is displayed, followed by a message indicating the severity of scoliosis [6]. The clinician may upload one image at a time but may reupload as many times as necessary without closing the window and ending the session when done.

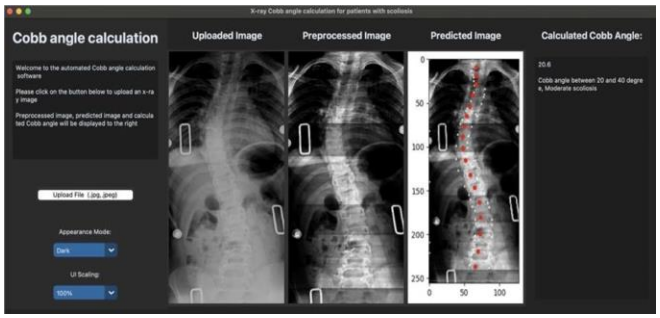

Figure 19: Graphic User Interface allowing for Surgeons to view the entire Cobb Angle Calculation Workflow.

6 Conclusions and Future Work

Current limitations involve the accuracy metrics of the proposed systems. A drawback of regional gamma correction, as observed in the second image in the GUI interface, is that values can be skewed by other parts of the image such as implants, or higher intensity pixel values in tissue. Another method to normalize that can remove discrepancies in these images can potentially lead to further accurate results in the Neural Network systems. This research demonstrated the robustness of a centroid placement algorithm instead of landmark detections. Better accuracy metrics can be attributed to the number of outputs to train is four times less than the landmark architecture. Cobb's angle prediction neural networks also support this hypothesis as well. Moreover, we get less than 10 degrees of Mean Absolute Error by training the model for just 100 epochs whereas in [6] the model is trained for 3000 epochs. The limitation we encountered is that the accuracy of the cobb's angle measurement depends on ROI (region of interest) cropping. Thus, in the future we would add a functionality that the surgeon can interact with the system to define the ROI and thereby accurately cropping the image. The GUI proposes a simple yet robust implementation to allow clinicians and surgeons to make decisions on treatment for patients. If a more robust model is discovered to provide better results, the underlying software architecture allows for a simple refactor to accept the new model for centroid and landmark algorithms and Cobb's angle predictions.

References

[1] Anon. 2022. Scoliosis. (May 2022). Retrieved April 25, 2023 from https://www.mayoclinic.org/diseases-conditions/scoliosis/symptoms-causes/syc-20350716

[2] Al-Bashir, A. K., Al-Abed, M. A., Amari, H. K., Al-Rousan, F. M., Bashmaf, O. M., Abdulhay, E. W., Al Abdi, R. M., Arunkumar, N., Bapu, B. R., & Al-Basheer, A. K. (2018). Computer-based Cobb angle measurement using deflection points in adolescence idiopathic scoliosis from radiographic images. Neural Computing and Applications, 31(5), 1547–1561. https://doi.org/10.1007/s00521-018-3614-y

[3] B. A. Kusuma, "Determination of spinal curvature from scoliosis X-ray images using K-means and curve fitting for early detection of scoliosis disease," 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 2017, pp. 159-164, doi: 10.1109/ICITISEE.2017.8285486.

[4] Wu, H., Bailey, Chris., Rasoulinejad, Parham., and Li, S., 2017.Automatic landmark estimation for adolescent idiopathic scoliosis assessment using boostnet. Medical Image Computing and Computer Assisted Intervention:127-135

[5] Linzhen Xie et al. 2022. Automatically measuring the Cobb angle and screening for scoliosis on chest radiograph with a novel artificial intelligence method. American Journal of Translational Research 14, 11 (November 2022), 7880–7888. PMID: 36505309; PMCID: PMC9730103

[6] Bo Chen, Qiuhao Xu, Liansheng Wang, Stephanie Leung, Jonathan Chung, and Shuo Li. 2019. An automated and accurate spine curve analysis system. IEEE Access 7 (2019), 124596–124605. DOI:http://dx.doi.org/10.1109/access.2019.2938402

[7] Ali Baaj, et al. Cobb angle used to measure scoliosis curves. Spine. Retrieved March 10, 2023, from https://www.spine-health.com/conditions/scoliosis/cobb-angle-used-measure-scoliosis-curves