Automatic Segmentation of 3D Ultrasound Spine Curvature Using Convolutional Neural Network

Sunetra Banerjee[1], Sai Ho Ling[1], Juan Lyu[2], Steven Su[1], Yong-Ping Zheng[3]

***Abstract*— Scoliosis is a 3D spinal deformation where the spine takes a lateral curvature, which generates an angle in a coronal plane. For periodic detection of scoliosis, safe and economic imaging modality is needed as continuous exposure to radiative imaging may cause cancer. 3D ultrasound imaging is a cost-effective and radiation-free imaging modality which gives volume projection image. Identification of mid-spine line using manual, semi-automatic and automatic methods have been published. Still, there are some difficulties like variations in human measurement, slow processing of data associated with them. In this paper, we propose an unsupervised ground truth generation and automatic spine curvature segmentation using U-Net. This approach of the application of Convolutional Neural Network on ultrasound spine image, to perform automatic detection of scoliosis, is a novel one.**

Keywords- Scoliosis, Segmentation, Ultrasound imaging, Convolutional Neural Network.

I. INTRODUCTION

Scoliosis is a deformation of the spinal cord, usually, S or C shaped, where a curvature, generally greater than 10°, occurs in the plane between the dorsal and ventral parts, i.e. coronal plane. In some cases, the degree of curvature is stable, while in others, it increases over time. This disease affects the adolescents primarily. Hence it is necessary to screen the school children. The current study shows, the overall occurrence of adolescent idiopathic scoliosis (AIS) is 0.47-5.2% of the total population and rampant in china (5%) and Hong Kong (3-4%) [1].

The exact cause of scoliosis is still unknown, but doctors theorize that it may happen during birth or maybe the side effect of some nerve or muscle problems such that cerebral palsy. The doctors also attribute responsibilities of genetic factors and hormonal issues for the occurrence of scoliosis. There are no such early signs for this disease.

As the disease matures, patients develop some symptoms like uneven shoulders, inflated curvature of the spine, disproportional alignment of hips or back pain.

1 S. Banerjee, S.H. Ling and S. Su are with the School of Biomedical Engineering, University of Technology Sydney, NSW, Australia Sunetra.Banerjee@student.uts.edu.au, Steve.Ling@uts.edu.au, Steven.Su@uts.edu.au

2 J. Lyu is with the College of Information and Communication Engineering, Harbin Engineering University, Harbin, China Juanlyu91@gmail.com

3 Y. P. Zheng is with the Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong yongping.zheng@polyu.edu.hk.

If not treated, progressive scoliosis may cause constant back pain, breathing problem or even physical disability for life. So, the early detection of scoliosis is critical.

The conventional method of diagnosis of scoliosis is using Cobb's measurement technique to measure the curvature angle of the spine on a standing radiograph [2]. If two lines are drawn, one perpendicular to the upper endplate of the uppermost vertebra involved and another drawn through the lower endplate of the most inferior vertebra involved, the angle between them is known as the Cobb angle.

Traditional radiography method is not preferred for periodic diagnosis of scoliosis as radiation may cause cancer. Excessive exposure to radiation can cause cancer. There are three primary in-vivo imaging modalities to detect scoliosis – The EOS imaging (invented by French scientist Georges Charpak in 1992), Magnetic resonance imaging [3], Conventional Ultrasound imaging [4].

Ultrasound imaging is a non-radiating imaging modality which is used extensively in the medical field due to its low cost. But there are visibility issues to detect bony structures from B-mode ultrasound images and thereby 2D B-mode ultrasound image is not feasible to detect spine deformation. A new invention in this field is freehand 3D ultrasound where the conventional 2D ultrasound is combined with position sensors, and the system can overcome the drawbacks of standard 2D ultrasound imaging technique. 3D ultrasound imaging system is generally experimental prototypes, and they are not ready enough to use for large clinical applications.

Usage of Convolutional Neural Network (CNN) for analysis of medical images is the recent trend [5]. CNN architecture is most used deep learning technique for analyzing medical image [6]. In biomedical image segmentation, enormous success was achieved by using U-Net architecture [7]. In U-Net, circumstantial information is transmitted to up-sampling layers, providing more number of feature channels.

The objective of our research is to segment the mid spine line from the whole image automatically and make it ready for automatic scoliosis angle measurement. In this research, we have used image generated from Scolioscan system. The Scolioscan system is used to generate 3D volume projection image (VPI) depending on conventional 3D ultrasound imaging technique. The 3D volumes were split in nine 2D images of different depths, and the best three images were selected for each patient [8]. Few sample images are shown in Fig 1.

The key contributions of this paper are:

1) An automatic unsupervised ground truth generator is introduced which generates pixel-wise annotations of the mid spine line for the input train 2D scoliosis image.

2) Automatic segmentation of mid-spine line of 2D scoliosis images is proposed for further processing of automatic scoliosis angle calculation.

The size of the image used is 1790*640. Afterwards, we calculated various loss functions to evaluate effectiveness.

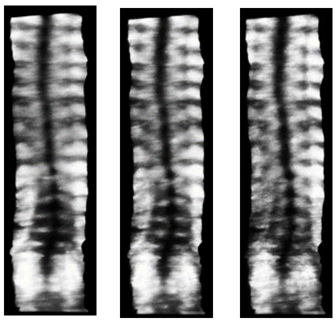


Fig 1. The top three 2D images of one patient

The outline of this paper is as follows. In section II, details of U-Net architecture along with its parameters and details about the input image, respectively, are described. Section III depicts the automatic unsupervised ground truth generation and automatic segmentation of mid-spine line along with the results of various loss function calculation. We conclude this paper in section IV with the future scope of this project

I. METHOD

Our method is comprised of three steps. In the first step, for mid-spine line segmentation of ultrasound images (Fig 1) first, we need to generate ground truth using active contour detection method [9]. The generation process of ground truth mask includes complementation of generated image, available from the previous method. In the second step, segmentation of the mid-spine line is done by U-Net architecture, using the ground truth mask. Finally, in the third step, the segmentation performance of U-Net on our image dataset is evaluated using three loss functions.

A. *Active Contour Detection*

Chan-Vese algorithm [9] is used for active contour detection of the region of interest, i.e. the mid spine line for our image set. Chan-Vese level set methodology is used for better tailoring of the boundary of the region of interest. Some objects do not have clearly defined boundaries, and the Chan-Vase segmentation algorithm is aimed to segment them. The Chan-Vase algorithm is based on level sets, and they are changed iteratively to reduce energy.

The Chan-Vese algorithm is shown in Algorithm 1. To detect the boundary of objects in an image, an active contour model evolves a curve, subject to constraints from a given image u_0. The curve C represented implicitly via a Lipschitz function *Ø*, by $C = \{(x,y)|\phi(x,y) = 0\}$. υ is a constraint on the area inside the curve, increasing the propagation speed. λ attracts the contour toward the object in the image. *H* is the Heavyside step function. *Ω* is a bounded open subset of parameterized curve. *F*(c_1,c_2 ,*C*) is the energy function defined by $F(c_1, c_2, C) = \mu.Length(C) + \upsilon.Area(inside(C)) + \lambda_1 \int_{inside(C)} |u_0(x,y) - c_1|^2\, dx\, dy + \lambda_2 \int_{outside(C)} |u_0(x,y) - c_2|^2\, dx\, dy.$

Keeping *Ø* fixed and minimizing the energy *F(c₁,c₂, Ø)* with respect to the constants c_1 and c_2 are defined as follows:

$$c_1(\emptyset) = \frac{\int_\Omega u_0(x,y) H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy} \tag{1}$$

$$c_2(\phi) = \frac{\int_\Omega u_0(x,y)(1-H(\phi(x,y)))dxdy}{\int_\Omega (1-H(\phi(x,y)))dxdy} \tag{2}$$

Parameterizing the decent direction by an artificial time $t \geq 0$, the equation in $\phi(t,x,y)$ is,

$$\frac{\partial\phi}{\partial t} = \partial_\varepsilon(\phi)\left[\mu\, div\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - v - \lambda_1(u_0 - c_1)^2 + \lambda_2(u_0 - c_2)^2\right] = 0\ in\ (0,\infty)\times\Omega \tag{3}$$

$$\phi(0,x,y) = \phi_0(x,y) in \Omega \tag{4}$$

$$\frac{\partial_\varepsilon(\phi)}{|\nabla\phi|}\frac{\partial\phi}{\partial \underset{n}{\rightarrow}} = 0\ on\ \partial\Omega \tag{5}$$

where $\underset{n}{\rightarrow}$ denotes the exterior normal to the boundary$\partial\Omega$, and $\frac{\partial\phi}{\partial\underset{n}{\rightarrow}}$ denotes the normal derivative of ϕ at the boundary. ε is a dirac delta function, μ is the length parameter, which does a scaling role.

Algorithm 1: Chan-Vese Algorithm

Data: Initialize the iteration number n=0.
Data: Construct the initial level set function ϕ_0
if n=0 **then**
Compute $c_1(\emptyset)^0$ and $c_2(\emptyset)^0$ using (1) and (2).
Set n=n+1
Compute the new level set ϕ^n by solving Partial Differential Equation using (3), (4) and (5).
Set regularized level set function ϕ_r^n as $\phi_r^n = \phi^n$.
end

while $((\phi_r^n - \phi^{n-1}) \neq 0)$ do
Compute $c_1(\emptyset)^n$ and $c_2(\emptyset)^n$ using (1) and (2).
Set n=n+1.
Compute the new level set by solving the PDE using (3), (4) and (5).
if $(|\nabla\phi^n|) \neq 0$ then
Reinitialize level set function ϕ^n to ϕ_r^n as:
$\frac{\partial\phi_r^n}{\partial t} = sign(\phi^n)(1 - |\nabla\phi^n|)$
end
else
Set as $\phi_r^n = \phi^n$
end.
end

Automatic segmentation of mid spine line is done using U-Net.

B. U-Net Architecture

U-Net architecture [7] has two major paths, i.e. a narrowing path and an expanding path. The narrowing path (left path in fig 2) consists of some recurrence of two convolution and one max-pooling layers. The kernel sizes of the convolution and max-pooling layers are 3×3 and 2×2, respectively. The expanding path (right path in fig 2.) recurrently concatenates of the features, which have been extracted from corresponding layers in narrowing path, two convolutions and one up-sampling layer. The final layer is made of one 1 × 1 convolution kernel.

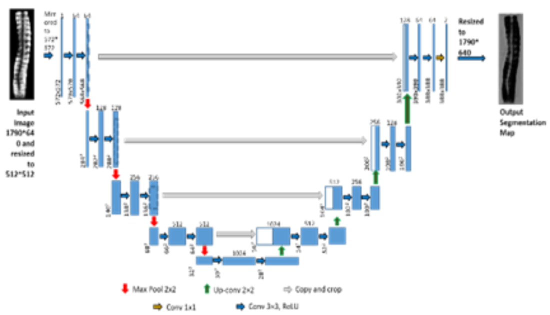


Fig 2: Basic U-Net Architecture

In convolutional layers, the activation functions for all layers are set to Rectified Linear Units (ReLU) [10], except for the last layer. In the last layer, the activation function is set to Softmax. The loss functions are set-to binary cross-entropy, Dice loss and Jaccard loss.

C. Loss Functions

We evaluated the performance of the segmentation using various loss functions. Cross-entropy loss measures the performance of a classification model whose output is a probability value between 0 and 1. In case of binary cross-entropy loss, the binary label 1is denoted as [9]. Dice loss is a measurement used to estimate the similarity of two samples. Jaccard index [11] also known as intersection over union (IOU) is a measurement used for measuring the similarity and diversity of sample sets.

II. RESULTS

A. Data Collection and experiment setup

The Scolioscan system (Model SCN801, Telefield Medical Imaging Ltd), developed in Hong Kong, is used to generate 3D volume projection image (VPI) depending on conventional 3D ultrasound imaging technique. Our data is collected from this Scolioscan system. The curve near the middle line curve of 3D VPI image is considered as a spine. This spine curve is used to measure the spinal deformity, and the deformity angle is called the Scolioscan angle, which is similar to Cobb angle of each patient, the 3D spine ultrasound image is sliced into 9, 2D vertebral anatomical images based on imaging depths. The 3D ultrasound images have different sizes according to different patients.

For future processing, we resized all of them in 1790*640 size and they all are in bitmap (.bmp) format. Then three best-selected images are selected from nine 2D images by CNN-RankNet[8]. The input images are shown in Fig 2.

The data for training contains 30 images, which we have resized to 512*512 size, and this number is not enough to train a deep neural network. So we did data augmentation to artificially increase the size of the training dataset and thereby enhance the performance of the model. ImageDataGenerator module of keras.preprocessing.image is used for data augmentation. The testing dataset also contains 30 images. We have again resized the final segmented image to 1790*640 size for comparing it with the input image.

The learning rate, optimizer, and weight initializer function are set to 0.00001, Adam [12], and He-normal [13], respectively. For the remaining parameters, we set it to the initial values as proposed in [7]. U-Net is trained through 5 epochs. The codes for implementing U-Net are scripted on python (version 3.7) using Keras with Tensorflow backend. A NVIDIA GeForce RTX 2060 OC is used for training and testing.

B. Ground Truth Generation

Ground truth mask for segmentation of mid-spine line is generated using two steps. In the first step, we applied Chan-Vese algorithm to detect the active contour, and the result is shown in Fig 3 (a)-(d). We have set the parameters like following. $\lambda_{1}= \lambda_{2}=1,\ \mu = 0.25, v = 1, h = 1, dt = 0.5$

Fig 3(b) and 3(c) show the final level set, i.e. the contours of the mid spine and the segmented mid-spine line after 200 iterations respectively. Fig 3(d) shows the number of iterations vs. energy curve, which proves that energy minimizes while we are very close in finding the active contour of the mid spine line for segmentation.

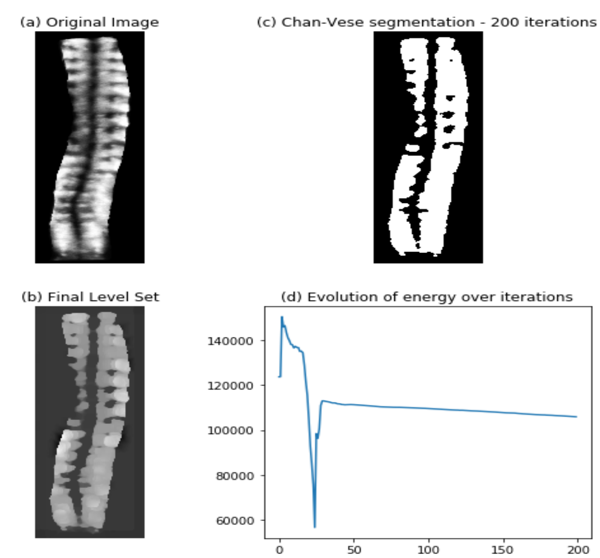


Fig 3. Generation of ground truth image using Chan-Vese method (a) Original image, (b) Final level set image, (c) Segmented image using Chan-Vese algorithm, (d) No. iterations vs. energy curve.

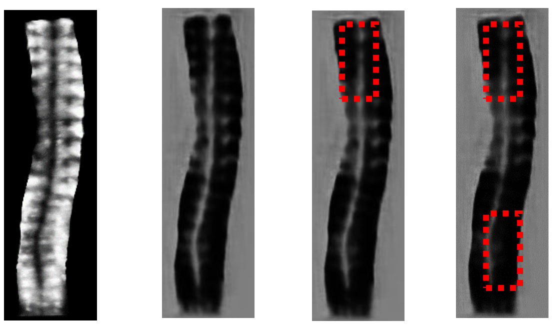
For masking we need to complement the generated image by Chan-Vese algorithm. The output image we will use for segmenting the mid-spine line.

In the second step, we complemented the generated image to produce ground truth mask for automatic segmentation. Fig 4 shows a ground truth mask.

Fig 4. Generated ground truth mask

C. Segmentation

Segmentation results using Convolutional Neural Network U-Net architecture are shown in Fig 5.



(a) (b) (c) (d)

Fig 5. Segmented mid-spine line (a) input image and segmented images using training with (b) binary cross-entropy, (c) dice loss, (d) Jaccard Loss

The mid-spine line was segmented automatically from the whole image and the outputs, of different loss functions, for same input image and model is not same. In all cases, the segmentation of mid-spine line is done successfully. But in few images, the ribs were almost indistinguishable from mid spine line. For example, Fig 5(b) is shows better result than Fig 5(c) and 5(d). In fig 5(d), a few rib lines are still visible on mid-spine line, but the segmentations are discontinuous, as shown in the red marked areas.

So, the quantitative evaluation of our system with ultrasound scoliosis dataset using three loss functions is presented here.

Loss Functions	**Accuracy (%)**
Binary Cross-entropy	63.03 ± 0.009
Dice Loss	59.12 ± 0.107
Jaccard Loss	40.04 ± 0.022

Table 1: Performance evaluation of segmentation method using various loss functions

Table 1 details the quantitative evaluation of our model using three loss functions and it can be inferred that the binary cross entropy function gives the best accuracy

III. CONCLUSION

Convolutional neural networks (CNNs) are successful in biomedical image analysis. Also, in radiology, CNN is becoming popular. To automate the scoliosis detection process, segmentation of the mid-spine line from the whole image is necessary. As no ground truth is available with the input images, the generation of ground truth is the vital part of this work. The model gives an average of 63% accuracy in segmentation of the mid-spine line using U-Net architecture and binary cross-entropy loss function.

In future, we will try to select the inflection point on the segmented mid-spine line automatically for scoliosis angle measurement.

REFERENCES

[1] Y.-P. Zheng *et al.*, "A reliability and validity study for Scolioscan: a radiation-free scoliosis assessment system using 3D ultrasound imaging," *Scoliosis and spinal disorders,* vol. 11, no. 1, p. 13, 2016.

[2] J. Cobb, "Outline for the study of scoliosis," *Instr Course Lect AAOS,* vol. 5, pp. 261-275, 1948.

[3] R. Morrissy, G. Goldsmith, E. Hall, D. Kehl, and G. Cowie, "Measurement of the Cobb angle on radiographs of patients who have," *J Bone Joint Surg Am,* vol. 72, no. 3, pp. 320-327, 1990.

[4] W. Chen, E. H. Lou, and L. H. Le, "Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 8495-8498: IEEE.

[5] 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, 2017.

[6] W. E. Thong, H. Labelle, J. Shen, S. Parent, and S. Kadoury, "Stacked auto-encoders for classification of 3d spine models in adolescent idiopathic scoliosis," in *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*: Springer, 2015, pp. 13-25.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234-241: Springer.

[8] J. Lyu *et al.*, "3D Ultrasound Spine Image Selection Using Convolution Learning-to-Rank Algorithm," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 4799-4802: IEEE.

[9] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing,* vol. 10, no. 2, pp. 266-277, 2001.

[10] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 807-814.

[11] P. Jaccard, "The distribution of the flora in the alpine zone. 1," *New phytologist,* vol. 11, no. 2, pp. 37-50, 1912.

[12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980,* 2014.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026-1034