Project Plan
Early Diagnosis of Scoliosis in Children from RGB-D Images Using Deep Learning

Huang Siyi, 3035335349, u3533534@connect.hku.hk
Li Gengyu, 3035331886, dawnlgy@connect.hku.hk

September 11, 2019

Contents

1 Introduction

Scoliosis, which can be usually found in teenagers or children, is a troublesome disease. Once after onset, scoliosis can result in severe damage to our health [4]. Therefore, great importance is now laid onto early diagnosis of scoliosis in children. Scoliosis in children is traditionally diagonalized by manually measuring the Cobb angle on X-ray which was invented in 1948 [1, 4, 5]. With the development of deep learning in computer vision field, neural network based method has now been adopted to diagnosis of scoliosis, which has yielded acceptable results [1, 5]. However, as the process to take X-ray images is radioactive and harmful to children, those X-ray based methods are not always the best options. Moreover, traditional method is both time-consuming and unreliable [4, 5]. To address these problems, we propose to use RGB-D images instead of X-ray images to diagnose scoliosis. We will design a deep learning model (CNN-based) to do landmark estimation on RGB-D images. Then we will use generative adversarial networks (GAN) [6] to synthesize X-ray images. With this procedure, children can be diagonalized correctly without taking the X-ray.

This project plan is written to propose our final year project named "Early Diagnosis of Scoliosis in Children from RGB-D Images Using Deep Learning".

2 Background

2.1 Scoliosis in Children

Scoliosis is defined as a morbid medical condition where the spine of a patient curves sideways. Scoliosis is widely seen in teenagers and children. Early diagnosis of scoliosis plays an essential role in treating it. Traditional methods including manual diagnosis using Cobb angel or even advanced method like landmark estimation on X-ray [1, 5] or segmentation [5] have yielded relatively acceptable results. However, none of the methods mentioned above can escape using X-ray, which is taken by a radioactive machine. As is the common sense, radioactivity is harmful to humans, especially children.

2.2 RGB-D Image

Traditional computer RGB image defines red, green and blue color components for each individual pixel. While a depth image for an RGB image is an image channel in which each pixel relates to a distance between the image plane and the corresponding object in the RGB image. An RGB-D

image is basically a combination of an RGB image and its corresponding depth image.

In this project, we will take RGB-D images for those children with scoliosis.

2.3 Microsoft Azure Kinect DK

Microsoft Azure Kinect DK is an advanced developer tool involving lots of computer vision technologies [3]. It has a depth sensor which can tell the distance between each pixel of the image and the corresponding point on the actual object. Meanwhile, this tool can work as a normal camera that records normal RGB images as well. Therefore, this tool is capable to take RGB-D images of the spine of a patient without the help of any radioactive machine.

In this project, we will use Microsoft Azure Kinect DK to build the training, testing and validating datasets for our model.

2.4 Landmark Estimation

Landmark estimation, or landmark detection, is one of the most fundamental tasks in computer vision. It requires a model to predict and plot some key points that identify the features of the object in the background. It is widely adopted in facial recognition and body movement tracking area [8]. Medical researchers have already been thinking of doing landmark detection on X-ray images or moire images [1, 4, 5], which have achieved a good performance.

2.5 Generative Adversarial Nets

Generative Adversarial Nets, or GAN, was introduced by Ian J. Goodfellow et. al in 2014 [6]. Since its invention, GAN has been adopted in lots of popular research fields in machine learning such as image reconstruction and image synthesis. Its basic principle is the contest between two networks in GAN namely generator and discriminator. Generator tries to generate fake data (usually images) to "fool" discriminator while discriminator tries to judge the fake data generated by generator. Ideally, in equilibrium, generator can generate the data which is "real" enough so that discriminator cannot tell whether the data generated is fake or not.

3 Objective

We plan to divide our objective into two parts, which correspond to two stages of our project process: landmark estimation and X-ray synthesis.

The first objective is to do landmark estimation by adopting latest deep learning model on RGB-D images which will be taken from the voluntary scoliosis patients.

The second objective is to synthesize X-ray by adopting GAN using original RGB-D images and the predicted landmark in the first part.

Several experiments will be carried out in both parts in order to get a good performance. Moreover, we will test our solution in real cases such as photo lab in the hospital.

4 Methodology

For data collection, RGB-D images will be collected in the hospital with the help of voluntary patients with scoliosis. Microsoft Azure Kinect DK will be the instrument to record the RGB-D images.

For model architecture, we are now doing a lot of research in related areas in order to find the best architecture for our deep learning model. Our deep learning models will be implemented using PyTorch which is a popular framework for deep learning in Python. Models will be trained on the GPU farm owned by Computer Science Department in HKU.

4.1 Part I - Landmark Estimation

We will use CNN based deep learning model for landmark estimation. Possible candidates includes BoostNet [1], MVC-Net [5] or the method applied on moire images [4]. As those methods are not for landmark estimation on RGB-D images, some modifications to the original model are necessary.

Input of the model will be RGB-D images of the back of patients with scoliosis. And output will be the landmarks which will be plotted onto corresponding RGB-D image.

4.2 Part II - Synthesis of X-ray

Generative adversarial nets (GAN) [6] will be used to synthesize X-ray. Again

some modifications are necessary because our task for GAN is different from that in the original paper. Method raised by B. Teixeira et. al [2] will also be our important reference.

Input of the model will be RGB-D images collected in the hospital and landmarks predicted in the first stage. Output will be synthetic X-ray images.

5 Schedule and Milestones

The tentative schedule and milestones are shown in the table below and they are subject to change with different situations in the future:

Date	TODOs	Milestone	Deliverable
9/1/2019 ~ 9/29/2019	1. Help to develop an App in The Duchess of Kent Children's Hospital at Sandy Bay to help them take RGB-D images and label the RGB-D images. 2. Discuss with supervisor and stakeholders to figure out the scope and make a detailed project plan. 3. Make project website. 4. Read some related papers.	1. Finish and test the App we developed for hospital. 2. Finish the detailed project plan. 3. Finish the project website.	1. A detailed project plan. 2. A project website.
9/30/2019 ~ 10/29/2019	1. Read all the related papers 2. Work on the implementation.	1. Finish all the related materials such as papers, code, books and tutorials. 2. Start the implementation for landmark detection.	N/A.
10/30/2019 ~ 11/30/2019	1. Complete a runnable implementation for landmark estimation.	1. The code can run smoothly without any bugs on CS GPU farm.	N/A.

12/1/2019 ~ 1/13/2020	1. Prepare for the first presentation. 2. Adjust hyperparameters, run several experiments on CS GPU farm and record the results.	1. Presentation can go smoothly. 2. Landmark estimation can give acceptable metrics.	1. A presentation powerpoint 2. A presentation demo.
1/13/2020 ~ 2/2/2020	1. Complete a runnable implementation for synthesis of X-ray. 2. Make interim report with the experiments results.	1. The code can run smoothly without any bugs on CS GPU farm. 2. Finish interim report.	1. A repository of implementation code. 2. An interim report.
2/3/2020 ~ 4/19/2020	1. Adjust hyperparameters, run several experiments on CS GPU farm and record the results. 2. Make final report with all the experiments results.	1. Model achieves acceptable metrics and code appears graceful with full features. 2. Finish final report.	1. Final version of implementation code. 2. A final report.
4/20/2019 ~ 5/5/2019	1. Prepare for final presentation.	1. Presentation can go smoothly with all the materials prepared.	1. A presentation poster. 2. A presentation powerpoint. 3. A presentation demo.

6 Resources and References

[1] H. Wu, C. Bailey, P. Rasoulinejad, and S. Li. Automatic Landmark Estimation for Adolescent Idiopathic Scoliosis Assessment Using BoostNet. In MICCAI, 2017.

[2] B. Teixeira, V. Singh, T. Chen, K. Ma, B. Tamersoy, Y. Wu, E. Balashova, and D. Comaniciu. Generating Synthetic X-ray Images of a Person from the Surface Geometry. In CVPR, 2018.

[3] Microsoft Azure Kinect DK: Azure Kinect DK Documentation. https://docs.microsoft.com/en-us/azure/kinect-dk.

[4] R. Choi, K. Watanabe, H. Jinguji, N. Fujita, Y. Ogura, S. Demura, T. Kotani, K. Wada, M. Miyazaki, H. Shigematsu, and Y. Aoki. CNN-based Spine and Cobb Angle Estimator Using Moire Images. In IIEEJ, 2017.

[5] H. Wu, C. Bailey, P. Rasoulinejad, and S. Li. Automatic Comprehensive Adolescent Idiopathic Scoliosis Assessment Using MVC-Net. In Medical Image Analysis, 2018.

[6] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. In arXiv: 1406.2661, 2014.

[7] M. Lootus, T. Kadir, and A. Zisserman. Vertebrae Detection and Labeling in Lumbar MR Images. In Computational Methods and Clinical Applications for Spine Imaging, 2014.

[8] A. Bulat, and G. Tzimiropoulos. Binarized Convolutional Landmark Localizers for Human Pose Estimation and Face Alignment with Limited Resources. In ICCV, 2017.