INTERVERTEBRAL DISC DETECTION IN X-RAY IMAGES USING FASTER R-CNN

Ruhan Sa*, William Owens*, Raymond Wiegand^, Mark Studin^^, Donald Capoferri**, Kenneth Barooha**, Alexander Greaux**, Robert Rattray**, Adam Hutton**, John Cintineo**, Vipin Chaudhary*
*State University of New York (SUNY) at Buffalo
^Spine Metrics, Inc.
^^University of Bridgeport College of Chiropractic
**Academy of Chiropractic

INTRODUCTION

Automatic identification of specific osseous landmarks on the spinal radiograph can be used to automate calculations for diagnosing ligament instability and injury, which affect 75% of patients injured in motor vehicle accidents. In this work, we propose to use deep learning based object detection method, Faster RCNN, as the first step towards identifying landmark points in lateral lumbar X-ray images. We show that, by using only 81 lateral lumbar X-Ray training images, one can achieve much better performance compared to traditional sliding window detection method on hand-crafted features. Furthermore, we fine-tuned the network using 974 training images and tested on 108 images, which achieved average precision of 0.905 with average computation time of 3 second per image, which greatly outperformed traditional methods in terms of accuracy and efficiency. We also proposed different fine-tuning techniques and compared and discussed the performance through extensive experiments.

METHOD

Fig.1 shows the deep learning network structure we applied in our work, which is the state of the art object detection method, Faster RCNN network. In our work, we fine tuned the network parameters due to the lack of medical image data and modified the anchor size and numbers suits for our data.

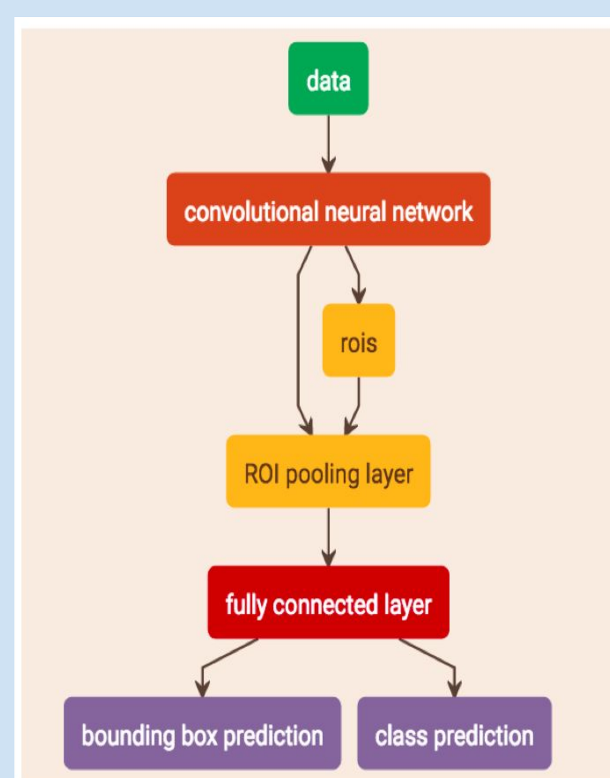

Fig. 1 Faster RCNN

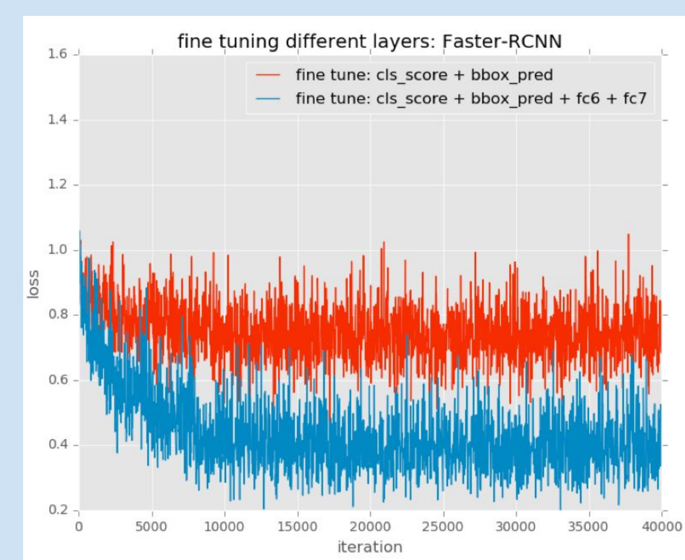

Fig. 2 Effect of shallow and deep fine-tuning on loss change for Faster-RCNN. We can see that deeper fine-tuning yields lower loss convergence.

EXPERIMENTS

We experimented on two different datasets to fully compare the performances and different tuning techniques. Base architecture is ZF net. Dataset 1 used 92 images and dataset 2 used 974 images. Fig. 2 and 3 shows the fine-tuning techniques performance.

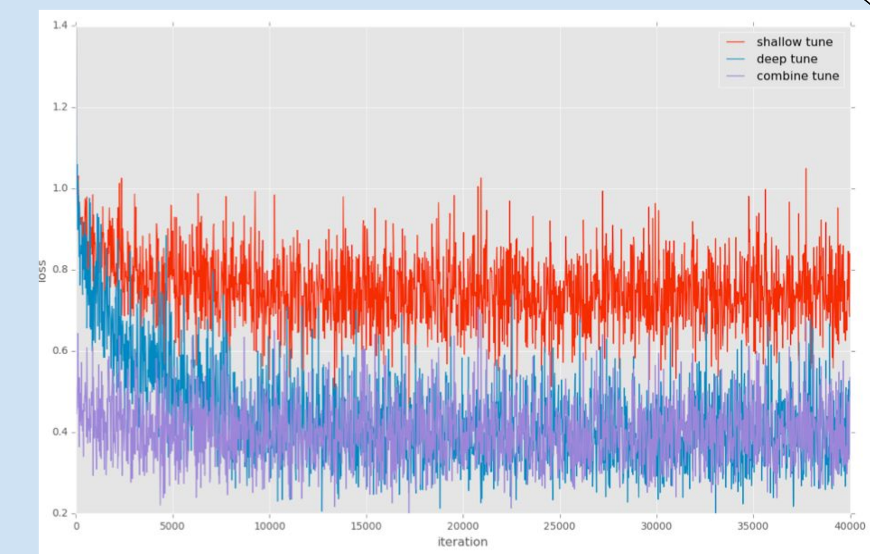

Fig. 3 Effect of shallow and deep fine-tuning VS two-stage training on loss change. We can see that combined tuning does not give lower convergence loss.

Table 1 shows the testing results on two datasets. We can see that compared to the traditional methods, deep learning based method performed better in speed and precision. Fig. 4 shows some of the qualitative results in challenging images.

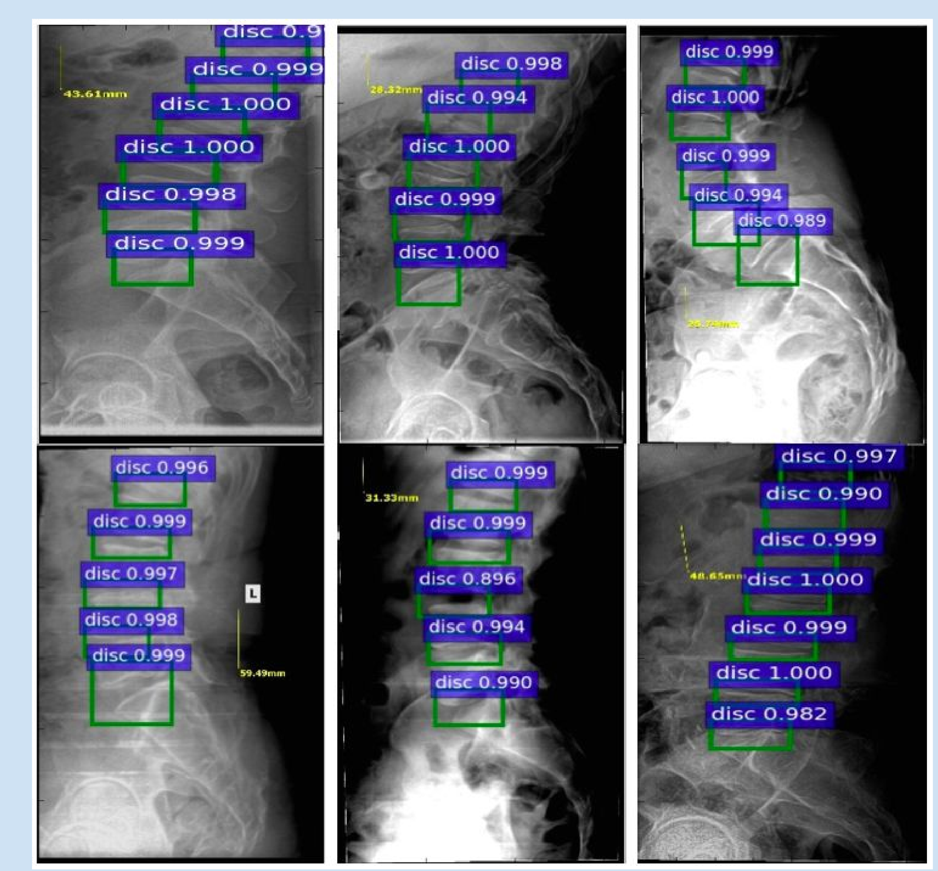

Fig. 4. Sample results

	AVG precision	AVG time (sec)
HOG+SVM (smaller dataset)	0.032	26
Faster-RCNN (smaller dataset)	0.651	10
HOG+SVM (larger dataset)	0.091	82
Faster-RCNN (larger dataset)	**0.905**	2

Table 1. Quantitative results.