Deep Learning for Instance-Aware Edge Detection of Vertebra on Lateral Radiographs

Author(s) names withheld
Address withheld

Editors: Under Review for MIDL 2020

Abstract

Vertebral fractures often exist in radiographs obtained for other purposes of patients without symptoms indicative of a fracture, resulting in many cases of underdiagnosis for mild and moderate fractures. The delineation of vertebral contours on medical images, particularly in X-ray images that frequently contain imaging artifacts, can provide an overview of spinal health and prevent neglect and misrecognition of potential fractures. Existing vertebra segmentation methods fall short of detecting the biconcave appearance of vertebral endplates often present in lateral radiographs due to image distortion, which may be erroneously interpreted as a fracture. We present a method for the delineation of vertebral contours that can also discriminate biconcave endplates on lateral radiographs, by adapting the Mask-RCNN framework for vertebra instance edge-detection. Our method achieved promising results when evaluated against a radiologist-generated ground truth. We discuss the potential of extending our approach to translate fracture localization and diagnosis into an end-to-end manner.

Keywords: deep learning, vertebra edge detection, convolutional neural network, x-ray

1. Introduction

When vertebral fractures (VFs) go unnoticed, it may lead to more severe fractures that can cause significant disability and morbidity. Early recognition can enable the appropriate administration of treatment, with reports demonstrating a reduction of subsequent fractures by 40-70% from osteoporosis therapies (Boonen et al., 2009; Compston, 2009). Radiography remains to be the most popular imaging technique for initial VF screening (Adams et al., 2010). However, VFs are often unsuspected clinically and present in radiographs obtained for other purposes (e.g. lung and heart examinations), prompting clinicians and radiologists to overlook the possibility of the condition (Lenchik et al., 2004). As a result, many studies observed a tendency towards underdiagnosis of VFs (Gehlbach et al., 2000, 2002; Probst et al., 2002; Delmas et al., 2005).

The precise delineation of vertebral body contours provides a clear visualization of any abnormalities in the spine, especially in the presence of imaging artifacts commonly found in lateral radiographs. Notably, the upper and lower contours of each vertebral endplate are often not perfectly superimposed due to misalignment between the diverging X-ray beam and the vertebral endplates. The apparent biconcave appearance, so-called “bean-can effect” shown in Figure 1, may be falsely recognized as a fracture (Griffith and Guglielmi, 2010). Despite the significance of this phenomenon long discussed in literature (Hurxthal, 1968), existing vertebra segmentation methods, including recently published deep learning
Deep Learning for Instance-Aware Edge Detection of Vertebra on Lateral Radiographs

Figure 1: The “bean-can effect” caused by oblique projection of the X-ray beam.

approaches (Sekuboyina et al., 2017; Janssens et al., 2018; Lessmann et al., 2019), focus on segmenting all pixels within the outermost contours and thus unable to address this issue. Methods that acknowledge the biconcave appearance of vertebral endplates are rare and do not detect the rest of the vertebral body (Gardner et al., 1998; Kamalakannan et al., 2010). To the best of our knowledge, our method is the first deep learning approach for automatic edge detection of individual vertebral bodies that also considers endplate biconcavity.

2. Methods

Data preprocessing. In this work, we used a private dataset of 120 lateral radiographs of the thoracic spine and 120 lateral radiographs of the lumbar spine with vertebral contours labelled by a certified radiologist. The images were preprocessed in the following steps: (1) downsampling, (2) contrast limited adaptive histogram equalization (CLAHE) (Zuiderveld, 1994), and (3) standardization of mean and standard deviation pixel values. These steps allow for more efficient training, given the memory constraints of the GPU.

Network architecture. The overall network architecture was inspired by the Mask R-CNN (He et al., 2017) instance segmentation framework and followed a two stage pipeline: (1) object detection based on RetinaNet (Lin et al., 2017), and (2) edge detection based on HED (Xie and Tu, 2015). The output of the object detection sub-network was a bounding box in which the edge detection sub-network predicted in. The building blocks of each sub-network were “Residual Units” (ResBlock) (He et al., 2016a) which consisted of the optimal layer combination suggested by He et al. (2016b).

Training. The network was trained by minimizing focal loss for object detection and class-balanced binary cross entropy for edge detection using the AMSGrad (Reddi et al., 2019) optimizer. The learning rate was $10^{-3}$ and weight decay was $10^{-4}$. We performed 5-fold cross-validation on the training set to minimize the bias of our model’s performance.

3. Result

Our network achieved an average Dice score on each vertebral contour of 0.73. It should be noted that unlike existing segmentation methods, we do not consider the pixels within the edges, resulting in a lower likelihood of overlap of between pixels. Figure 2 shows our results for two example cases from the thoracic spine and the lumbar spine. The network prediction showed significant qualitative similarity to the radiologist-generated ground truth.
Figure 2: Example of network performance on the thoracic spine (first row) and lumbar spine (second row). Our network detected the biconcave appearance of vertebral endplates when present. (a) Input image. (b) Radiologist-annotated ground truth (blue). (c) Network prediction (green). (d) Fusion map of ground truth and network prediction. Green represents true positive, blue represents true negative, and red represents false positive.

4. Discussion and Future Work

In this study, we showed the first application of deep learning for the detection of vertebral contours on lateral radiographs which is robust to the biconcave appearance of endplates. Our preliminary analysis showed that our framework motivated by Mask-RCNN could provide a contour map of vertebrae comparable to one produced by an expert radiologist.

For future work, we plan to include more prior knowledge, such as vertebra key points and curvature of the vertebral column, to improve the detection accuracy and generalizability of our method. As our method distinguishes individual vertebra instances, it should theoretically easily translate for vertebra localization given the appropriate training labels. Moreover, we aim to include vertebral morphometry and fracture information for an artificial intelligence-enabled tool that can support VF diagnosis.
Acknowledgments

Acknowledgments withheld.

References


