

Deep convolutional networks for automated detection of posterior-element fractures on spine CT

Holger R. Roth^a, Yinong Wang^a, Jianhua Yao^a, Le Lu^a, Joseph E. Burns^b, and Ronald M. Summers^a

^aImaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD 20892-1182, USA.

^bDepartment of Radiological Sciences, University of California-Irvine, Orange, CA 92868, USA.

ABSTRACT

Injuries of the spine, and its posterior elements in particular, are a common occurrence in trauma patients, with potentially devastating consequences. Computer-aided detection (CADe) could assist in the detection and classification of spine fractures. Furthermore, CAD could help assess the stability and chronicity of fractures, as well as facilitate research into optimization of treatment paradigms.

In this work, we apply deep convolutional networks (ConvNets) for the automated detection of posterior element fractures of the spine. First, the vertebra bodies of the spine with its posterior elements are segmented in spine CT using multi-atlas label fusion. Then, edge maps of the posterior elements are computed. These edge maps serve as candidate regions for predicting a set of probabilities for fractures along the image edges using ConvNets in a 2.5D fashion (three orthogonal patches in axial, coronal and sagittal planes). We explore three different methods for training the ConvNet using 2.5D patches along the edge maps of ‘positive’, i.e. fractured posterior-elements and ‘negative’, i.e. non-fractured elements.

An experienced radiologist retrospectively marked the location of 55 displaced posterior-element fractures in 18 trauma patients. We randomly split the data into training and testing cases. In testing, we achieve an area-under-the-curve of 0.857. This corresponds to 71% or 81% sensitivities at 5 or 10 false-positives per patient, respectively. Analysis of our set of trauma patients demonstrates the feasibility of detecting posterior-element fractures in spine CT images using computer vision techniques such as deep convolutional networks.

Keywords: spine CT, computer-aided detection, posterior-element fractures, deep learning

1. INTRODUCTION

Injuries of the spine and its posterior elements in particular (see Figure 1) are a common occurrence in traumatic patients, with potentially devastating consequences.¹ Spine fractures are detected using volumetric imaging such as computed tomography (CT) in order to access the degree of injury. Spine injuries are a critical concern in blunt trauma, particularly in cases of motor vehicle collision and fall from significant heights. More than 140,000 vertebral fractures occur in the U.S. each year.² However, the traditional method of qualitative visual assessment of images for diagnosis could miss fractures, and is time-consuming, potentially causing delays in time-critical situations such as the treatment of spine injuries. Computer-aided detection (CADe) has the potential to expedite the assessment of trauma cases, reduce the chance of misclassification of fractures of the spine, and decrease inter-observer variability. Furthermore, CAdE could help assess the stability and chronicity of fractures, as well as facilitate research into optimization of treatment paradigms.

Further author information contact Holger Roth (holger.roth@nih.gov) or Ronald Summers (rms@nih.gov)

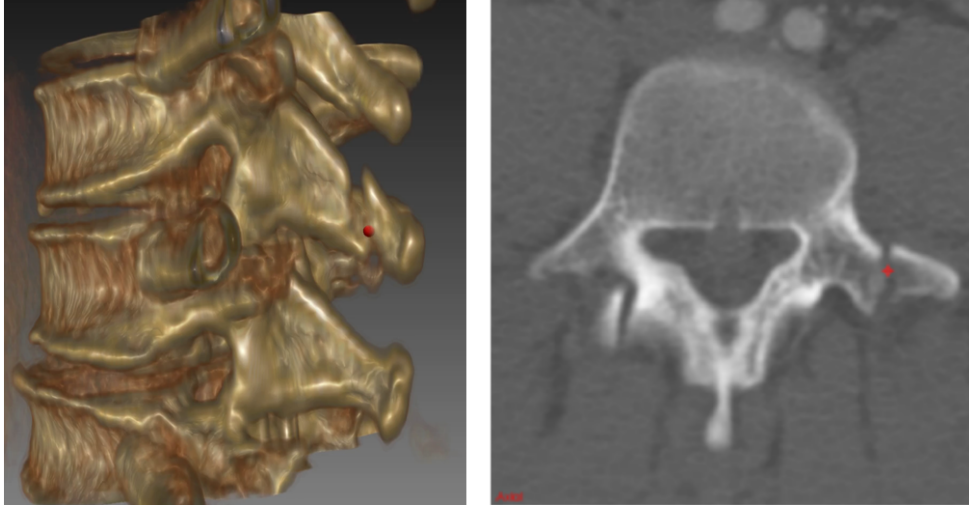


Figure 1. Example 3D rendering of a posterior element fracture of the spinous process (left) and reformatted images (right) show a posterior element fracture of the transverse process in axial and sagittal views. Manual gold standard marks by radiologist are shown in red.

2. METHODS

Recently, the availability of large annotated training sets and the accessibility of affordable parallel computing resources via GPUs has made it feasible to train deep convolutional networks (ConvNets), also popularized under the keyword deep learning, for computer vision classification tasks. Great advances in classification of natural images have been achieved.^{3,4} Studies that have tried to apply deep learning and ConvNets to medical imaging applications also showed promise.⁵⁻⁸

2.1 Convolutional networks

ConvNets are named for their convolutional filters which are used to compute image features for classification.³ In this work, we use a ConvNet similar to^{3*} with 5 convolutional layers, 3 fully connected layers, and a final softmax layer for classification. However, we use input images of size 64×64 and a stride of 1 in the first convolutional layer. All convolutional filter kernel elements are trained from the data in a fully-supervised fashion. This has major advantages over more traditional CADe approaches that use hand-crafted features, designed from human experience. ConvNets have a better chance of capturing the “essence” of the imaging data set used for training than when using hand-crafted features.^{3,4} Furthermore, we can train the similarly configured ConvNet architectures from randomly initialized or pre-trained model parameters for detecting different lesions or pathologies (with heterogeneous appearances), with no manual intervention of system and feature design. Examples of trained filters from the first convolutional layer for posterior-element fracture detection are shown in Fig. 2. Dropout⁹ is used during training as a form of regularization that avoids overfitting to the training data.

*<https://code.google.com/archive/p/cuda-convnet2/>

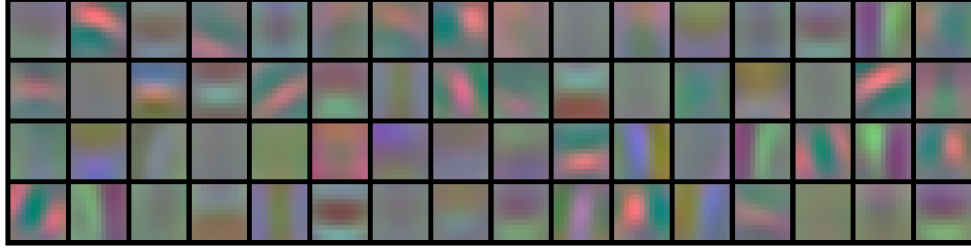


Figure 2. Convolution filters trained on examples of posterior-element fractures and non-fractures.

2.2 Application to spine CT

In this work, we apply ConvNets for the automated detection of posterior element fractures of the spine. First, the vertebra bodies of the spine with its posterior elements are segmented in spine CT using multi-atlas label fusion.^{10,11} A set of atlases of the vertebra bodies are registered to the target image with free-form deformation. Then, an edge map E of the posterior elements is computed using the Sobel operators

$$S_x = \begin{Bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{Bmatrix} \quad \text{and} \quad S_y = \begin{Bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{Bmatrix} \quad (1)$$

in order to find the horizontal ($G_x = S_x * I$) and vertical ($G_y = S_y * I$) approximations of the image derivative for each CT slice. Here, $*$ denotes a convolutional operation. Edge points are located at the maximum of the absolute gradient $|G| = \sqrt{G_x^2 + G_y^2}$. The edge maps E serve as candidate regions for predicting a set of probabilities P for fractures along an image edge using ConvNets. An example of posterior-element segmentation and edge map estimation is shown in Fig. 3.

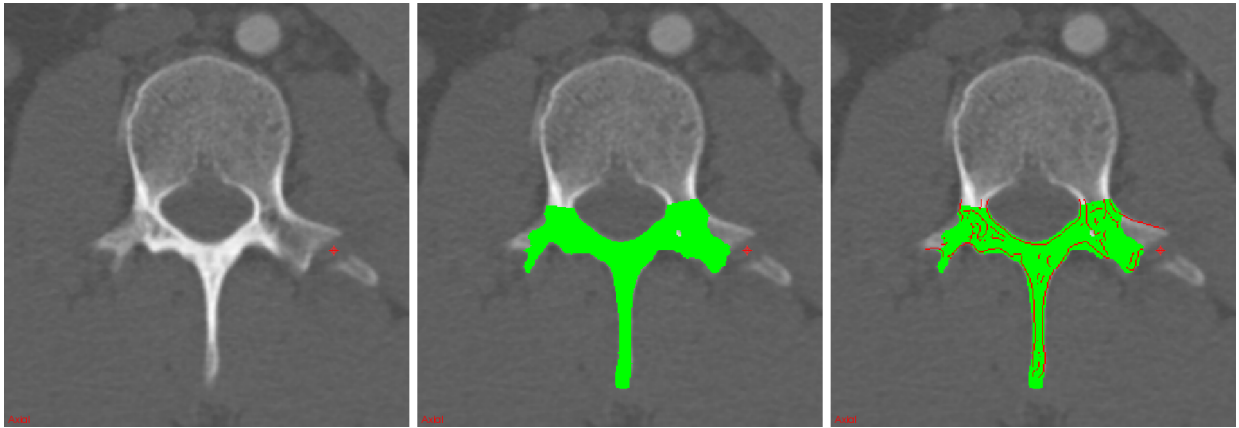


Figure 3. (a) Posterior element fracture; (b) posterior element segmentation using multi-atlas label fusion;¹⁰ (c) edge map estimation.

We explore three different methods for training the ConvNet using 2.5D patches (three orthogonal patches in axial, coronal and sagittal planes) along the edge maps of positive, i.e. fractured posterior-elements and negative, non-fractured elements:

1. P_{original} : 2.5D patches along each edge voxel $e \in E$ are aligned to original scanner coordinates in axial, coronal, and sagittal planes.
2. P_{mirrored} : 2.5D patches at voxels e near fractures are mirrored along the y -axis of the image in order to increase positive training examples.
3. P_{oriented} : 2.5D patches are oriented along the principal axis of an edge at voxel e (while positive training examples are handled as in case 2, i.e. they are also mirrored).

The alignment of a training patch along its edge orientation is illustrated in Fig. 4. This step can be also seen as a form of data augmentation that artificially increases the variation of training examples (as opposed to just having slightly shifted versions of nearby patches as in case 1 and 2). Orientating a patch along the anatomical structures of interest has also been shown to improve performance in other applications such as in detection of pulmonary embolism¹² where the image patch can be aligned along the direction of a blood vessel. The orientation of an edge is estimated as in Equation 2. In 2D, the eigenvector v_1 corresponding to the largest eigenvalue λ_1 corresponds to the major axis of a local edge. Hence, the orientation θ of the edge can be estimated as the angle between this eigenvector and the origin v_0 :

$$\theta = \tan^{-1} \left(\frac{v_1}{v_0} \right) \quad (2)$$



Figure 4. Illustration of 2D patches aligned along the candidate edge orientation with respect to the origin for training convolutional networks (ConvNets).

3. RESULTS

3.1 Data

An experienced radiologist retrospectively marked the location of 55 displaced posterior-element fractures in 18 trauma patients admitted for traumatic emergencies to the University of California-Irvine Medical Center. Image sizes range within $512 \times 512 \times [259 - 582]$ and resolutions range within $[0.29 - 0.41] \times [0.29 - 0.41] \times [1.0 - 2.0]$ mm. We use a random subset of 12 of these patients with spine fractures for training ConvNets as described in Section 2; 6 patients are reserved for testing. An additional set of 5 spine CTs of healthy patients were added to the training set in order to increase the number of non-fractured examples. A total of 800,000 2.5D patches are randomly selected from the candidate edge maps $\{E\}$ of the training set and used for learning the ConvNet parameters. After convergence, the ConvNet is applied to a testing case edge map E in order to produce a probability map P for fractures. Figure 5 shows examples of probability maps for posterior-element fracture detection along candidate edges. The radiologist’s markings are indicated by crosshairs.

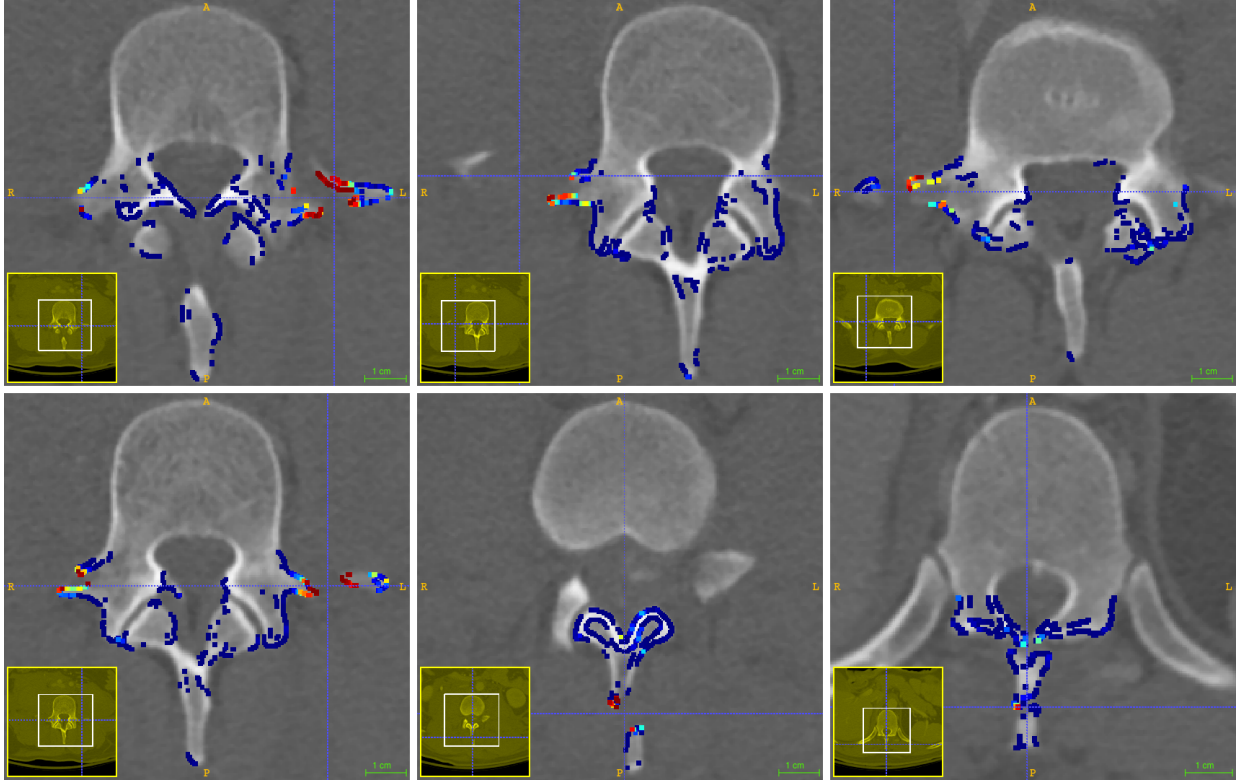


Figure 5. Examples of probability maps for posterior-element fracture detection along candidate edges. The posterior element fractures are marked by crosshairs.

3.2 Performance

The classification performance is evaluated using voxel-wise ROC and per posterior-element process (left, right and spinous process) FROC analysis. Fig. 6 shows the voxel-wise ROC and processes-wise FROC performance using the differently trained ConvNets. A clear advantage of using ‘Oriented’ patches for training can be observed in testing, compared to ‘Original’ and ‘Mirrored’ patches for training. In testing, we achieve area-under-the-curve (AUC) values of 0.761, 0.796 and 0.857 for ‘Original’, ‘Mirrored’, and ‘Oriented’ patch training respectively. This corresponds to 71% or 81% sensitivities at 5 or 10 false-positives per patient, respectively, in the case of training and testing with ‘Oriented’ patches.

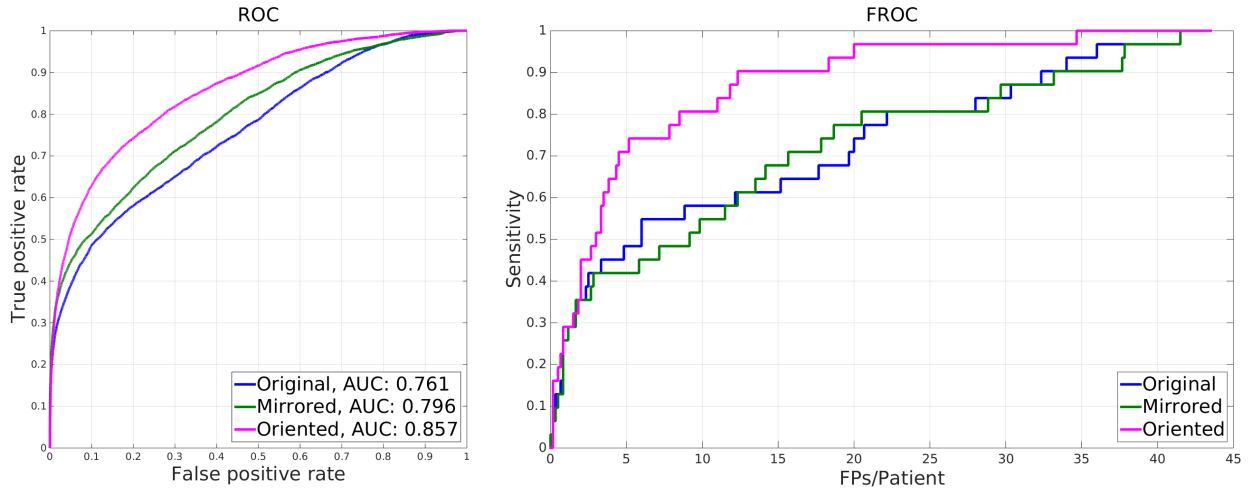


Figure 6. Voxelwise-ROC performance (left) and per-fracture FROC performance (right) evaluating the classification accuracy.

4. CONCLUSION

While there has been work on automated detection of other spinal injuries, such as vertebral compression fractures,² to the best of our knowledge, this is the first work to explore automated posterior-element fracture detection. It demonstrates that deep convolutional networks (ConvNets) can be useful for detection tasks, such as the detection of fractures in spine CT.

Results from analysis of our set of trauma patients demonstrate the feasibility of detecting posterior-element fractures in spine CT images using computer vision techniques such as ConvNets. We evaluated three methods for training a ConvNet and show the advantages of applying an oriented patch extraction method (P_{oriented}) for better classification performance.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the NIH Clinical Center.

REFERENCES

- [1] Parizel, P., van der Zijden, T., Gaudino, S., Spaepen, M., Voormolen, M., Venstermans, C., De Belder, F., van den Hauwe, L., and Van Goethem, J., “Trauma of the spine and spinal cord: imaging strategies,” *European Spine Journal* **19**(1), 8–17 (2010).
- [2] Yao, J., Burns, J. E., Muñoz, H., and Summers, R. M., “Cortical shell unwrapping for vertebral body abnormality detection on computed tomography,” *Computerized Medical Imaging and Graphics* **38**(7), 628–638 (2014).
- [3] Krizhevsky, A., Sutskever, I., and Hinton, G. E., “Imagenet classification with deep convolutional neural networks,” in [*Advances in neural information processing systems*], 1097–1105 (2012).
- [4] Zeiler, M. D. and Fergus, R., “Visualizing and understanding convolutional networks,” in [*Computer Vision—ECCV 2014*], 818–833, Springer (2014).
- [5] Yan, Z., Zhan, Y., Peng, Z., Liao, S., Shinagawa, Y., Metaxas, D. N., and Zhou, X. S., “Bodypart recognition using multi-stage deep learning,” in [*Information Processing in Medical Imaging*], 449–461, Springer (2015).
- [6] Cireşan, D. C., Giusti, A., Gambardella, L. M., and Schmidhuber, J., “Mitosis detection in breast cancer histology images with deep neural networks,” in [*Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*], 411–418, Springer (2013).
- [7] Roth, H. R., Lu, L., Seff, A., Cherry, K. M., Hoffman, J., Wang, S., Liu, J., Turkbey, E., and Summers, R. M., “A new 2.5 d representation for lymph node detection using random sets of deep convolutional neural network observations,” in [*Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014*], 520–527, Springer (2014).
- [8] Roth, H. R., Lu, L., Liu, J., Yao, J., Seff, A., Kevin, C., Kim, L., and Summers, R. M., “Improving computer-aided detection using convolutional neural networks and random view aggregation,” *arXiv preprint arXiv:1505.03046* (2015).
- [9] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., “Dropout: A simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research* **15**(1), 1929–1958 (2014).
- [10] Wang, Y., Yao, J., Roth, H. R., Burns, J. E., and Summers, R. M., “Multi-atlas segmentation with joint label fusion of osteoporotic vertebral compression fractures on ct,” *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging* (2015).
- [11] Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., Yushkevich, P., et al., “Multi-atlas segmentation with joint label fusion,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **35**(3), 611–623 (2013).
- [12] Tajbakhsh, N., Gotway, M. B., and Liang, J., “Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks,” in [*Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*], 62–69, Springer (2015).