Cervical Spine Fracture Detection and Localization

Naman Raghuvanshi, Siddharth Rayabharam† and Sumant Suryawanshi†

Contributing authors: nvr5386@psu.edu; nqr5356@psu.edu; szs7220.psu.edu;
†These authors contributed equally to this work.

Abstract

Purpose: To develop a deep learning model for the detection and localization of cervical spine fractures in the axial CT scans. **Methods:** We use the dataset consisting of cervical spine CT scans provided by the Radiological Society of North America (RSNA). The dataset consists of 3000 studies of individual patients for the training and testing combined. Out of these 3000 patient studies, 83 of these studies also contains segmentation data. Additionally, for x studies the dataset contains bounding box coordinates data. The dataset population is split into 90% for training and 10% for validation. We use EfficientNetV2[1] to learn the segmentation, X model for detecting the fractures and Y model for drawing bounding boxes around the fracture area. **Results:** Results need to be updated **Conclusion:** The ability of a X model to detect and localize cervical spine fractures on axial CT radiographs with high sensitivity and specificity was demonstrated.

Keywords: Cervical Spine, Fracture Detection, Deep Learning

1 Introduction

Cervical spine injury is very common injury with more than 3 million cases per year that are being evaluated for cervical spine injury in North America[2]. In United States, more than 1 million patients with blunt force injury are suspected to suffer cervical spine injury[3]. Since cervical spine injury is associated with high morbidity and mortality, quick diagnosis of the injury is crucial. Any delay in diagnosis may result in devastating consequences for the patient. So, any additional aid to the radiologists can reduce the morbidity or mortality of the patient.

In recent years, a machine deep learning technique known as deep convolutional neural network (DCNN) has been applied to image recognition tasks. DCNN's are well suited for images. So, they have been used extensively in the field of medicine to classify medical images.

In past few years, there have been many studies that have tried to use DCNN[4][5][6] on medical radiographs. In these studies, the reference standard for the training and testing images was based on the assessment of human readers determining which were visible, only within a radiograph. Many radiologist fail to detect "occult fracture" because of the difficulty in detecting such fracture in a radiograph. These extraction methodologies could adversely influence the classification accuracy and occult fracture being assessed as a "non-fracture case". A proficient algorithm may help identify and triage studies for the radiologist to review more urgently, helping to ensure faster diagnoses.

The purpose of our study is to develop an automated deep learning system for detecting cervical spine fractures using CT a gold standard annotated by radiologists, and to evaluate the diagnostic performance inclusive of the experienced readers in detecting cervical spine fractures on radiographs.

2 Methodology

2.1 Dataset and Study Population

We use the dataset consisting of cervical spine CT scans provided by the Radiological Society of North America (RSNA). The dataset we are using is made up of roughly 3000 CT studies, from twelve locations and across six continents. Spine radiology specialists have provided annotations to indicate the presence, vertebral level and location of any cervical spine fractures. Each radiology study consists of multiple dcm files. A dcm file follows the Digital Imaging and Communications in Medicine (DICOM) format. It is the standard format used for storing medical images and related metadata. It dates back to 1983, although it has been revised many times. Out of these 3000 patient studies, 83 of these studies also contains segmentation data. Additionally, for x studies the dataset contains bounding box coordinates data. The dataset population is split into 90% for training and 10% for validation. The distribution of fractured and non-fractured dataset in the dataset is shown in the below two graphs.

Bar graph and pie graph for the dataset.

2.1.1 Segmentation

Segmentation involves classification of CT radiographs into C1 to C7 labels. Using the 83 studies which have segmentation data. We train Random Forest classifier[7] and EfficientNetV2[1]. Correct vertebrae labels are essential when training the fracture prediction model. First we extract vertebrae targets from 83 segmentation files.

Vetebrae	Random Forest Classifier	EfficientNetV2
C1	0.8891	0.9621
C2	0.8898	0.9409
C3	0.8962	0.9559
C4	0.8952	0.9605
C5	0.8893	0.9482
C6	0.8713	0.9450
C7	0.8576	0.9465
Overall	0.8841	0.9513

Table 1 Accuracies of the models for the segmentation

From Table 1, we can see that EfficientNet performs well compared to the Random Forest Classifier. EfficientNetV2 is a new family of convolutional networks that have faster training speed and better parameter efficiency than previous models. To develop these models, a combination of training-aware neural architecture search and scaling was used to jointly optimize training speed and parameter efficiency. The models were searched from the search space enriched with new ops such as Fused-MBConv[8]. Experiments show that EfficientNetV2 models train much faster than state-of-the-art models while being up to 6.8x smaller. With progressive learning, EfficientNetV2 significantly outperforms previous models on ImageNet and CIFAR/Cars/Flowers datasets. By pretraining on the same ImageNet21k, EfficientNetV2 achieves 87.3% top-1 accuracy on ImageNet ILSVRC2012, outperforming the recent ViT by 2.0% accuracy while training 5x-11x faster using the same computing resources

2.1.2 Fracture Detection

2.1.3 Localization

3 Results

4 Discussion

Discussions should be brief and focused. In some disciplines use of Discussion or 'Conclusion' is interchangeable. It is not mandatory to use both. Some journals prefer a section 'Results and Discussion' followed by a section 'Conclusion'. Please refer to Journal-level guidance for any specific requirements.

5 Conclusion

Conclusions may be used to restate your hypothesis or research question, restate your major findings, explain the relevance and the added value of your work, highlight any limitations of your study, describe future directions for research and recommendations.

In some disciplines use of Discussion or 'Conclusion' is interchangeable. It is not mandatory to use both. Please refer to Journal-level guidance for any specific requirements.

Appendix A Section title of first appendix

An appendix contains supplementary information that is not an essential part of the text itself but which may be helpful in providing a more comprehensive understanding of the research problem or it is information that is too cumbersome to be included in the body of the paper.

References

[1] Tan, M., Le, Q.V.: Efficientnetv2: Smaller models and faster training. In: Meila, M., Zhang, T. (eds.) Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event. Proceedings of Machine Learning Research, vol. 139, pp. 10096–10106. PMLR, ??? (2021). http://proceedings.mlr.press/v139/tan21a.html

[2] Milby, A.H., Halpern, C.H., Guo, W., Stein, S.C.: Prevalence of cervical spinal injury in trauma. Neurosurg Focus **25**(5), 10 (2008). https://doi.org/10.3171/FOC.2008.25.11.E10

[3] Minja, F.J., Mehta, K.Y., Mian, A.Y.: Current challenges in the use of computed tomography and mr imaging in suspected cervical spine trauma. Neuroimaging Clin N Am **28**(3), 483–493 (2018). https://doi.org/10.1016/j.nic.2018.03.009

[4] Olczak, J., Fahlberg, N., Maki, A., Razavian, A.S., Jilert, A., Stark, A., Sköldenberg, O., Gordon, M.: Artificial intelligence for analyzing orthopedic trauma radiographs. Acta Orthop **88**(6), 581–586 (2017). https://doi.org/10.1080/17453674.2017.1344459

[5] Kim, D.H., MacKinnon, T.: Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clin Radiol **73**(5), 439–445 (2018). https://doi.org/10.1016/j.crad.2017.11.015

[6] Chung, S.W., Han, S.S., Lee, J.W., Oh, K.-S., Kim, N.R., Yoon, J.P., Kim, J.Y., Moon, S.H., Kwon, J., Lee, H.-J., Noh, Y.-M., Kim, Y.: Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta Orthop **89**(4), 468–473 (2018). https://doi.org/10.1080/17453674.2018.1453714

[7] Buitinck, L., Louppe, G.: Scikit-learn: Machine Learning in Python, Random Forest Classifier. https://scikit-learn.org/stable/modules/

generated/sklearn.ensemble.RandomForestClassifier.html

[8] Gupta, S., Tan, M.: EfficientNet-EdgeTPU: Creating Accelerator-Optimized Neural Networks with AutoML. https://ai.googleblog.com/2019/08/efficientnet-edgetpu-creating.html

[9] Campbell, S.L., Gear, C.W.: The index of general nonlinear DAES. Numer. Math. **72**(2), 173–196 (1995)

[10] Slifka, M.K., Whitton, J.L.: Clinical implications of dysregulated cytokine production. J. Mol. Med. **78**, 74–80 (2000). https://doi.org/10.1007/s001090000086

[11] Hamburger, C.: Quasimonotonicity, regularity and duality for nonlinear systems of partial differential equations. Ann. Mat. Pura. Appl. **169**(2), 321–354 (1995)

[12] Geddes, K.O., Czapor, S.R., Labahn, G.: Algorithms for Computer Algebra. Kluwer, Boston (1992)

[13] Broy, M.: Software engineering—from auxiliary to key technologies. In: Broy, M., Denert, E. (eds.) Software Pioneers, pp. 10–13. Springer, New York (1992)

[14] Seymour, R.S. (ed.): Conductive Polymers. Plenum, New York (1981)

[15] Smith, S.E.: Neuromuscular blocking drugs in man. In: Zaimis, E. (ed.) Neuromuscular Junction. Handbook of Experimental Pharmacology, vol. 42, pp. 593–660. Springer, Heidelberg (1976)

[16] Chung, S.T., Morris, R.L.: Isolation and characterization of plasmid deoxyribonucleic acid from Streptomyces fradiae. Paper presented at the 3rd international symposium on the genetics of industrial microorganisms, University of Wisconsin, Madison, 4–9 June 1978 (1978)

[17] Hao, Z., AghaKouchak, A., Nakhjiri, N., Farahmand, A.: Global integrated drought monitoring and prediction system (GIDMaPS) data sets. figshare https://doi.org/10.6084/m9.figshare.853801 (2014)

[18] Babichev, S.A., Ries, J., Lvovsky, A.I.: Quantum scissors: teleportation of single-mode optical states by means of a nonlocal single photon. Preprint at https://arxiv.org/abs/quant-ph/0208066v1 (2002)

[19] Beneke, M., Buchalla, G., Dunietz, I.: Mixing induced CP asymmetries in inclusive B decays. Phys. Lett. **B393**, 132–142 (1997) arXiv:0707.3168 [gr-gc]

[20] Stahl, B.: DeepSIP: Deep Learning of Supernova Ia Parameters, 0.42,

Astrophysics Source Code Library (2020), ascl:2006.023