RSNA 2022 Cervical Spine Fracture Detection

1. Background:

CT images are widely used for radiologists to make judgements about if there's a bone fracture and evaluate the severity of the fracture. While CT has its own advantages, there is still some room for improvement. The following three reasons suggest we could try some new tech to help with the existing problem of CT.(1)As the statistical data shows, over 1.5 million spinal fractures occur annually in the US[1], which indicates a foreseeable increase in CT demands. (2)Many experienced radiologists may get retired and there's always a strain on the workforce. (3) Analyzing CT images can always be a time-consuming process, leading to a delay of treatment and many possible neurologic deterioration and paralysis after trauma. So it's reasonable for us to develop machine learning models to identify and recognize the cervical spine fracture.

2. About the topic:

In this Kaggle competition, we are provided with imaging dataset with modification and segments from radiologists. We will try to develop machine learning models that match the radiologists' performance in detecting and localizing fractures to the seven vertebrae that comprise the cervical spine[2].

3. Literature review:

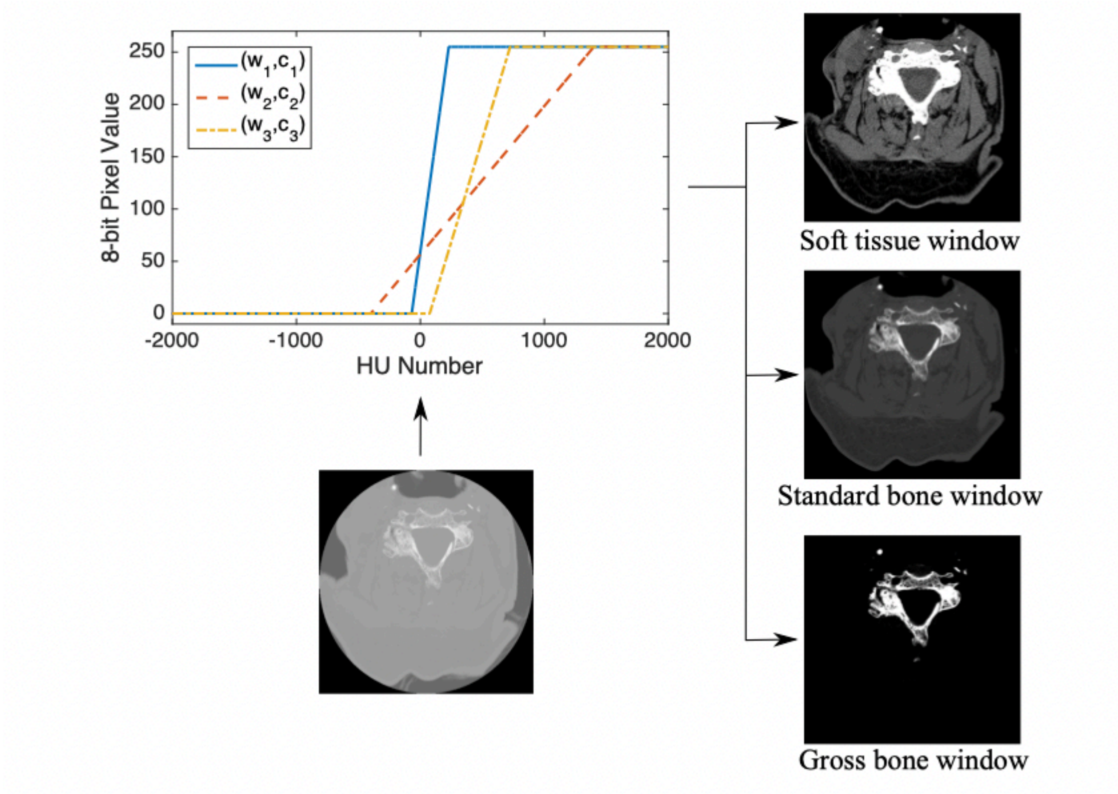
The following method and pictures mainly derive from H. Salehinejad's paper[3].

The fracture detection problem can be seen as a classification problem, which has two significant parts: preprocessing and training model. The input set x is CT images and the output is the label y (y=0 means no fracture and y=1 means the image contains at least one fracture).

a. Preprocessing part (Fig 1)

To understand how deep learning models can be used for localizing fractures, taking a glance at how CT works is super necessary. CT examinations are basically using window technique to observe normal tissues or lesions of different densities. Since various tissues or lesions have different CT values, the window width(abbreviation: w) and window level(center schemes, abbreviation: c) that are suitable for viewing a tissue or lesion should be selected to obtain the best display when displaying details of a tissue structure. Under this situation, each CT image is duplicated into three images and three different windows are chosen for each of them to better enhance the bones and make a difference between standard bones, gross bones and soft tissues. Moreover, the represented image (3) with the gross bone window x is then used to crop the images with 5% margin from each side of detected cervical spine using Otsu's method[4]. The cropped images are resized to 384*384.

Fig 1 the process of preprocessing

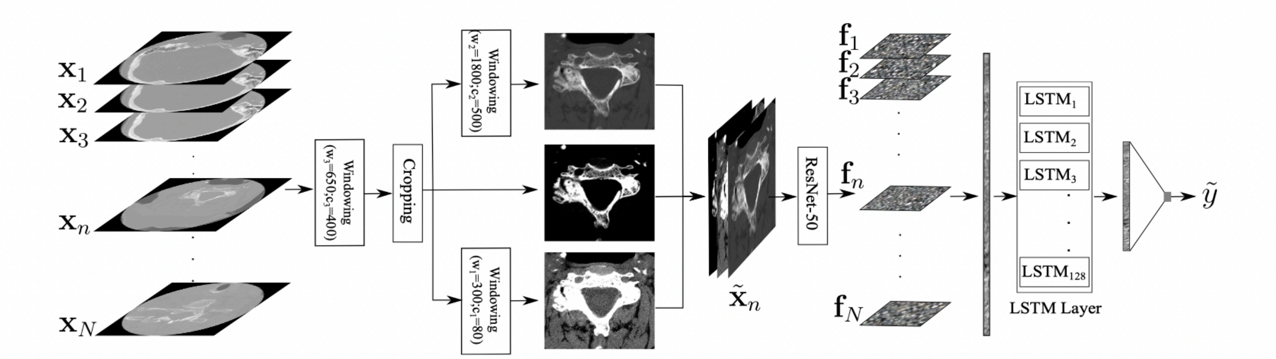


b. Training model part

ResNet-50 is trained with those preprocessed images randomly and the goal of training is to obtain the feature map(f1, f2, … fn). Then the bidirectional network of LSTM(long short term memory) units is used for learning the dependency among axial images and mapping f to label y.

In this part, the most important thing is to understand LSTM networks. LSTM is a special and improved kind of RNN network. It has three gates: input gates, forget gate, output gate. These gates are designed to help with making decisions about if the new information or the short term information is useful and should be kept in the long term memory. In other words, LSTM network is treated as a filter to enhance RNN model and improve the performance of training since it can remember information from a long time memory and find something are related.

Fig 2 The whole process of the model



4. Results analysis:

A. The performance of the ResNet-50 of fracture detection in cervical spine fracture images is presented in table1 while the performance of the combination of ResNet-50 and BLSTM model is presented in table 2. The difference of PPV(positive predictive value) between two tables suggests the significance of incorporating temporal features in training.

B. However, the NPR(negative positive rate) in table 1 is relatively high, regardless of imbalanced dataset or balanced ones. This may arise from the method to judge the value of y. To be more specific, if there's one piece of the three CT images are labeled as positive, then all three of them will be labeled as positive. That may be the reason of the large number of false positive cases and could be improved in the future.

C. The performance of the model in imbalanced dataset is much greater of the balanced dataset. This is possibily due to the BLSTM network could better capture and learn the dependency between negative cases but also may have some other leading aspects.

Table 1

Data	TPR	TNR	PPV	NPV	F1	Acc	MCC	AUC
Imblcd.	77.21 ± 4.1	80.06 ± 3.0	06.47 ± 0.6	99.50 ± 0.1	11.92 ± 1.0	80.01 ± 2.9	18.47 ± 1.0	78.63 ± 1.2
Blcd.	77.21 ± 4.1	77.62 ± 2.7	13.78 ± 1.0	98.67 ± 0.2	23.35 ± 1.3	77.61 ± 2.5	26.11 ± 1.0	77.42 ± 1.1

Table 2

Model	Data	TPR	TNR	PPV	NPV	F1	Acc	MCC	AUC
ResNet-50 + BLSTM-96	Imblcd.	64.19 ± 5.7	78.67 ± 6.6	43.62 ± 6.3	89.83 ± 1.5	51.66 ± 5.5	75.79 ± 5.2	37.84 ± 7.6	71.43 ± 3.9
ResNet-50 + BLSTM-128		62.28 ± 6.0	80.84 ± 2.9	44.83 ± 4.8	89.62 ± 1.6	52.06 ± 4.9	77.15 ± 2.9	38.54 ± 6.7	71.56 ± 3.7
ResNet-50 + BLSTM-256		59.01 ± 5.7	84.12 ± 4.9	48.54 ± 6.7	89.34 ± 1.5	52.92 ± 4.6	$\mathbf{79.18 \pm 3.8}$	40.36 ± 6.5	71.57 ± 3.1
ResNet-50 + BLSTM-96	Blcd.	64.19 ± 5.7	77.11 ± 7.3	74.14 ± 5.4	68.36 ± 3.1	68.58 ± 3.8	70.65 ± 3.5	41.90 ± 7.1	70.65 ± 3.5
ResNet-50 + BLSTM-128		62.28 ± 6.0	79.84 ± 3.1	75.55 ± 3.0	68.06 ± 3.5	68.17 ± 4.2	$\mathbf{71.06 \pm 3.1}$	42.86 ± 5.9	71.06 ± 3.1
ResNet-50 + BLSTM-256		57.75 ± 4.9	84.09 ± 5.3	78.87 ± 4.9	66.63 ± 1.8	66.44 ± 2.8	70.92 ± 1.9	43.62 ± 4.3	70.92 ± 1.9

5. References:

[1]Kalmet, P. H. S., Sanduleanu, S., Primakov, S., Wu, G., Jochems, A., Refaee, T., … Poeze, M. (2020). *Deep learning in fracture detection: a narrative review. Acta Orthopaedica, 1–6.* doi:10.1080/17453674.2019.1711

[2]https://www.kaggle.com/competitions/rsna-2022-cervical-spine-fracture-detection/overview

[3]H. Salehinejad et al., "Deep Sequential Learning For Cervical Spine Fracture Detection On Computed Tomography Imaging," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), 2021, pp. 1911-1914, doi: 10.1109/ISBI48211.2021.9434126.

[4]Sunil L Bangare, Amruta Dubal, Pallavi S Bangare, and ST Patil, “Reviewing otsu’s method for image thresholding,” International Journal of Applied Engineering Research, vol. 10, no. 9, pp. 21777–21783, 2015.