Cervical Spine Fracture Detection through Two-stage Approach of Mask Segmentation and Windowing based on Convolutional Neural Network

Doyeon Kim[1], Xujia Ning[2], Kaicheng Liang[3], Yi Ni[4], Duan Wang[5], Mingyuan Li[6], Yichuan Wang[7], Erick Purwanto[8], Ka Lok Man[9]

CONTENTS

01
02
03
04
05
06

Background

Cervical spine fracture:
break / dislocation

Most occurrence:

- 15 – 24 years
- Over 55 years

Severe consequence:

- disorder
- paralysis

Malunion / delayed union

Fractured bone debris

01
02
03
04
05
06

Background

Radiologists' detection on fracture:
=> Limitation of human capability

95%

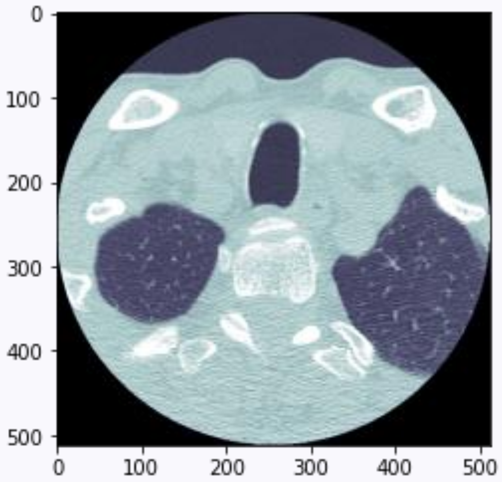
Acute fracture

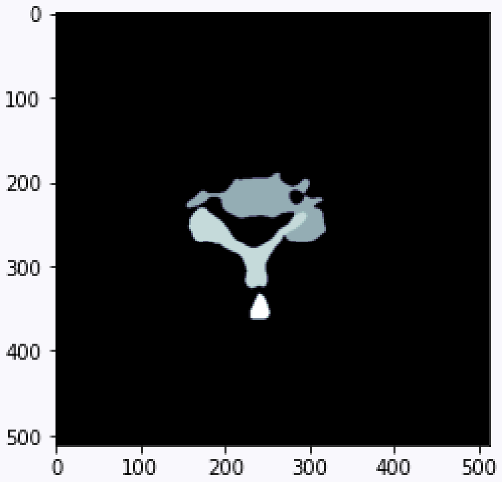
80%

Blunt trauma / no symptom

Introduction

- Kaggle RSNA Competition: Cervical Spine Fracture Detection
- Train_images: 2019 CT scans
  segmentations: 87 scans

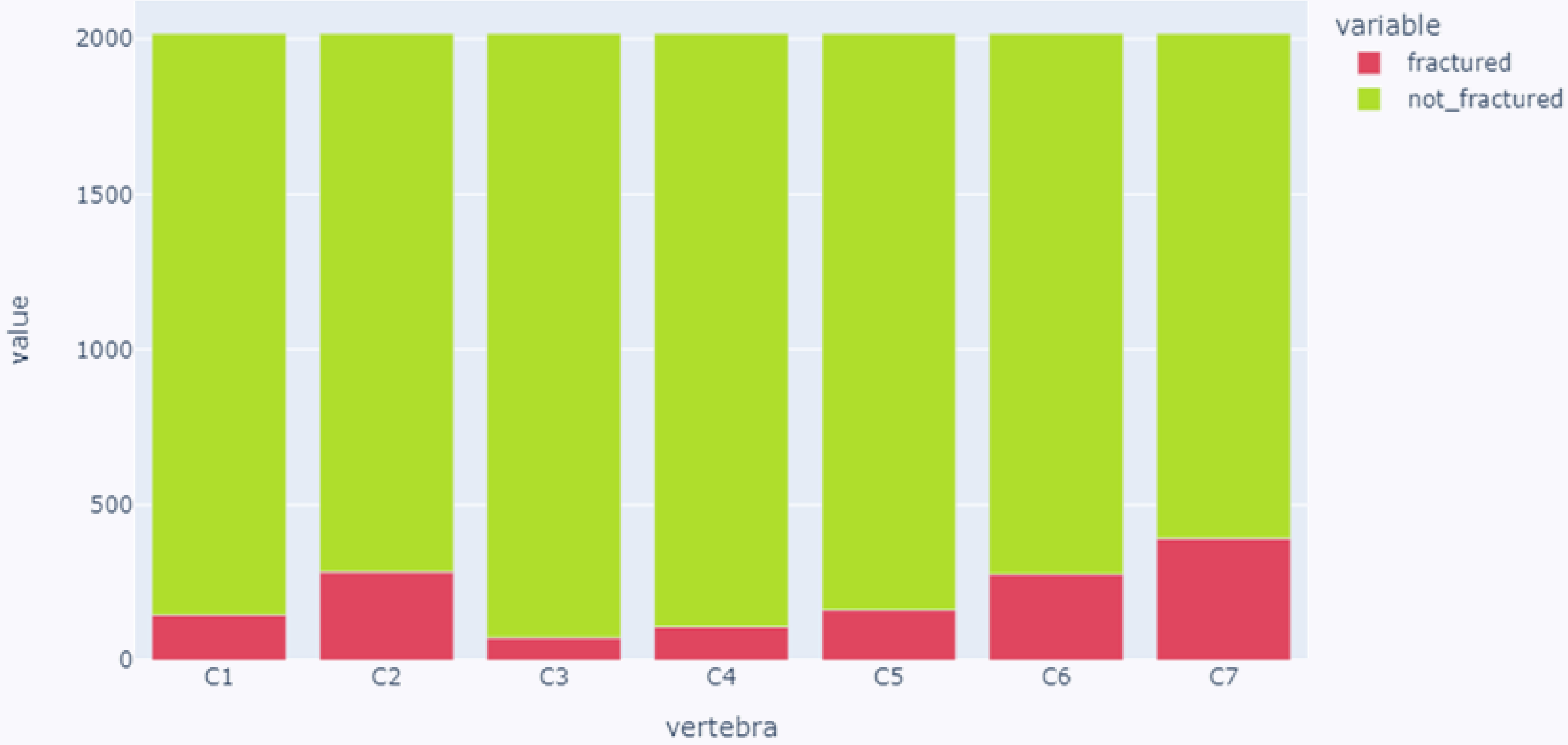




array([0., 3., 4.])

Challenges

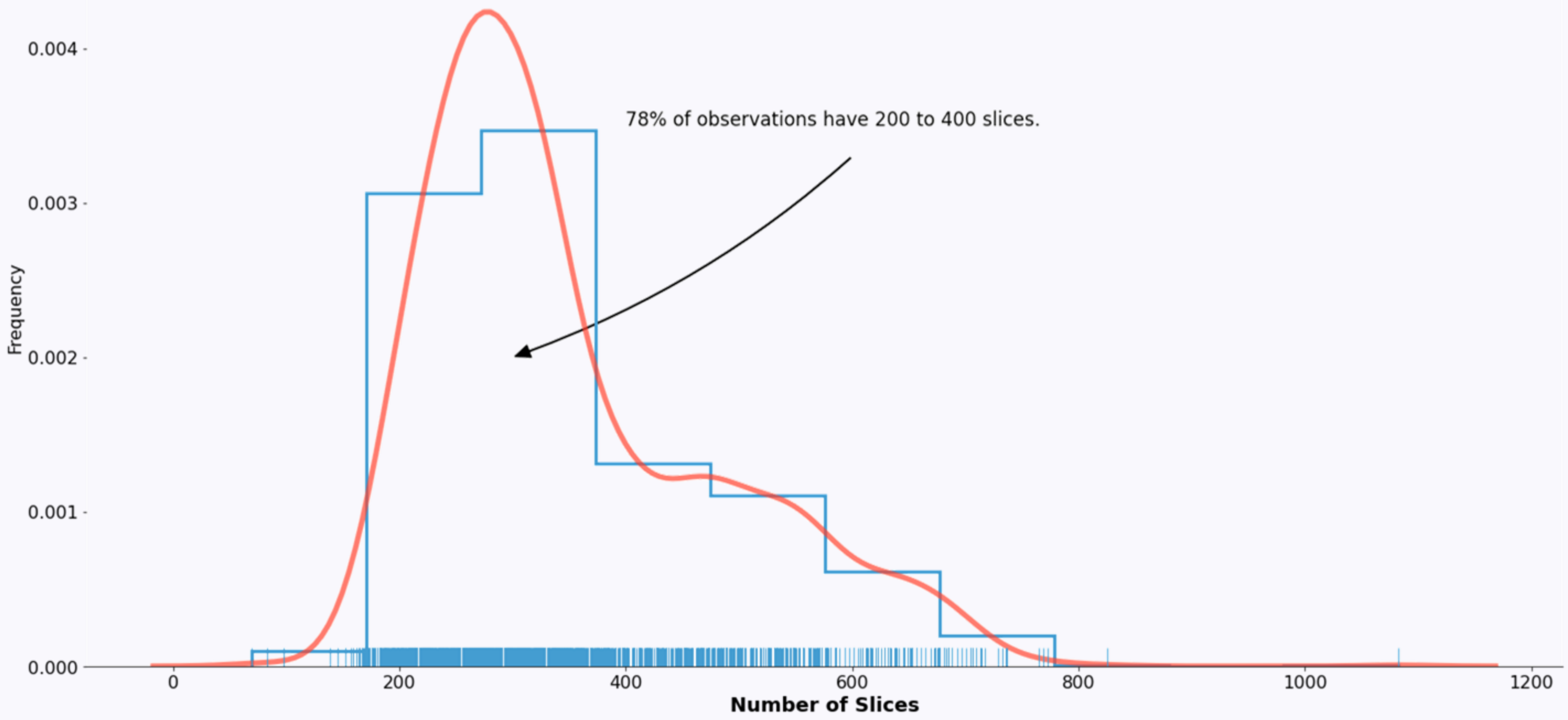
1. Extreme imbalanced sample distribution



01
02
03
04
05
06

Challenges

2. Imbalanced train image file slices distribution



01
02
03
04
05
06

Challenges

3. Small segmentations for initial provision

Methodology

01 Stage 1 preprocessing: HU scaling

02 Stage 1 model: UNet-EfficientNet-B5

03 Stage 2 preprocessing 1: Vertebrae image centering

04 Stage 2 preprocessing 2: Cropping voxel

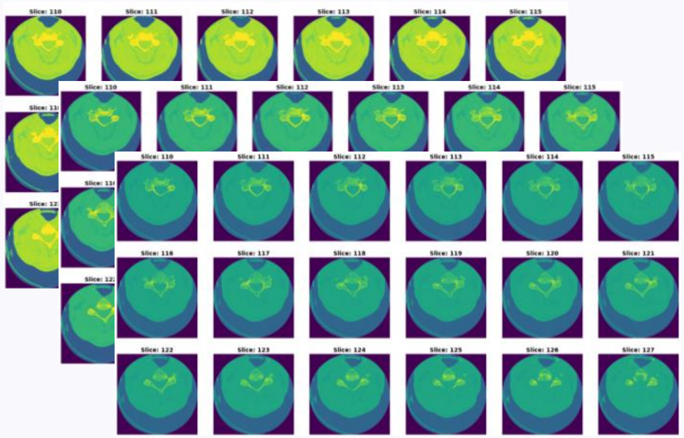
05 Stage 2 model: CrackNet-BiLSTM

Methodology

Stage 1 preprocessing: **HU scaling / windowing**

Extract clearer bone image from provided dataset

- Window = 1800 & level = 400; **spine bone**
- Window = 2800 & level = 600; **head and neck temporal bone 1**
- Window = 4000 & level = 700; **head and neck temporal bone 2**



Methodology

Stage 1: UNet + EfficientNet-B5

- Semantic segmentation
- Classifies the bone image for each vertebra

UNet

traditional and widely used model for medical imaging

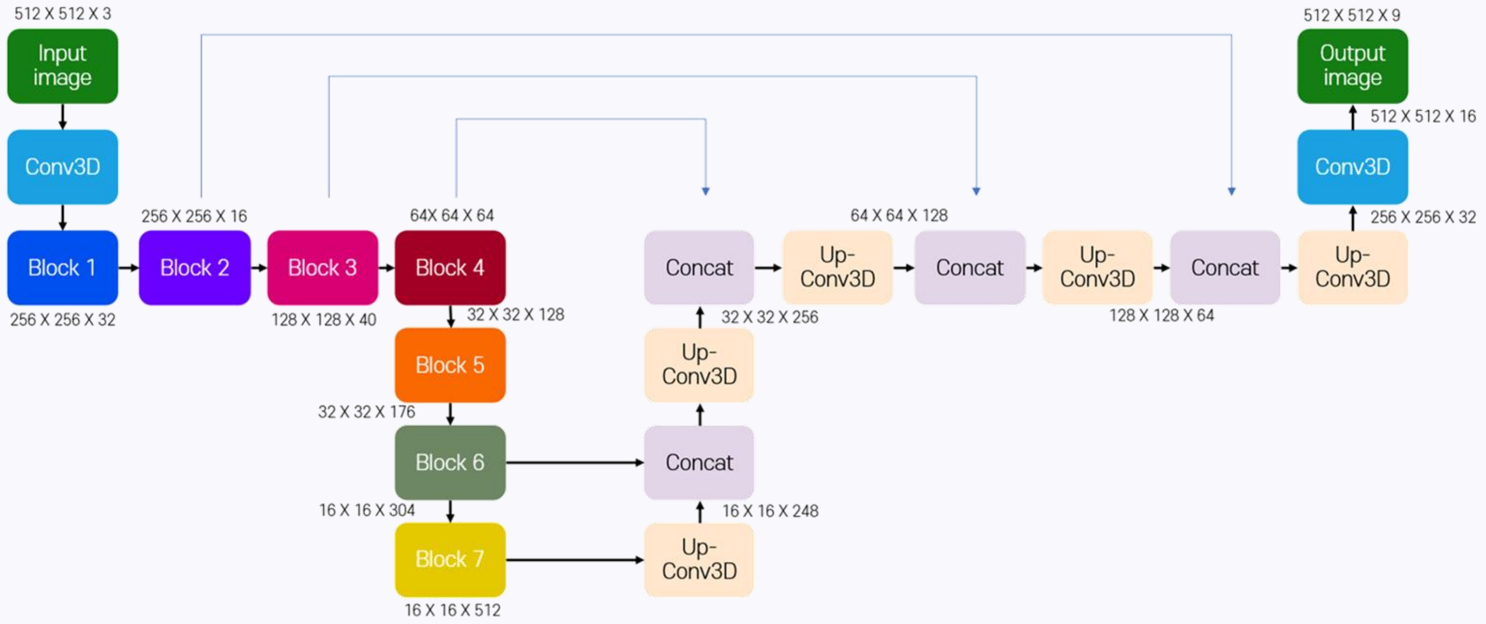
EfficientNet-B5

large enough parameters to process the CT images without overfitting

Methodology

Stage 1: UNet + EfficientNet-B5

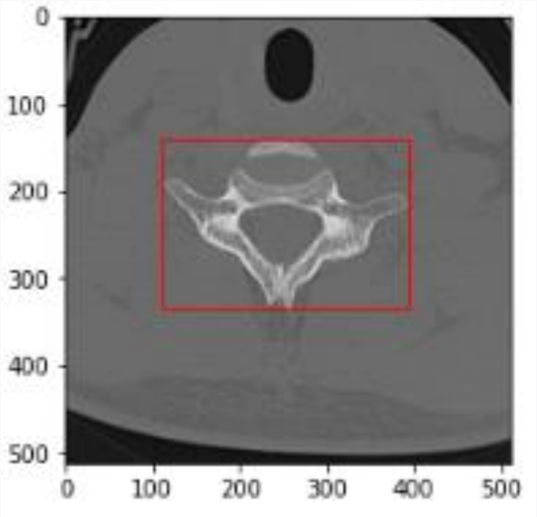
- Encoder: EfficientNet = create the representation of features at different levels
- Decoder: UNet = combines the features and generates a prediction as a segmentation mask

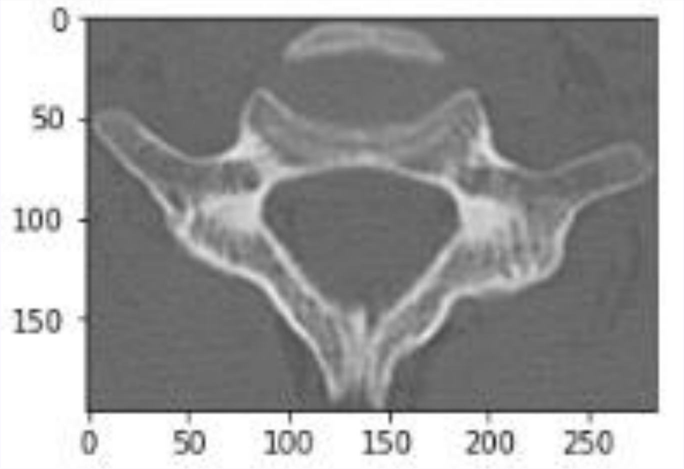


Methodology

Stage 2 preprocessing 1: Vertebrae image centering

- Create bounding boxes around the vertebrae
- Focus the images with Yolov5

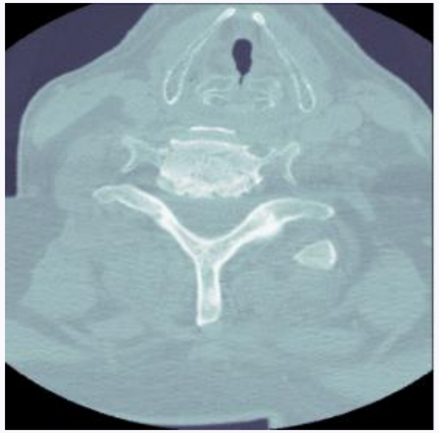


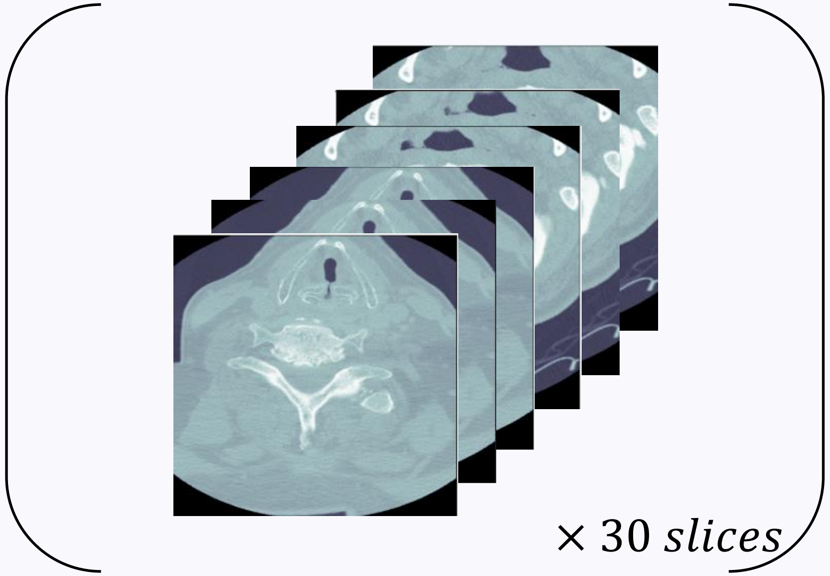


Methodology

Stage 2 preprocessing 2: Cropping voxel

- Take average slice for each vertebra = 30 slices
- Combine the cervical bones
- Combination: utilizing the Yolov5 cropped images





$\times 30\ slices$

01
02
03
04
05
06

Methodology

Stage 2: CrackNet + BiLSTM

- Detects the bone fracture with each vertebra combined images

CrackNet

simply finds the crack and the characteristics of the bone fracture

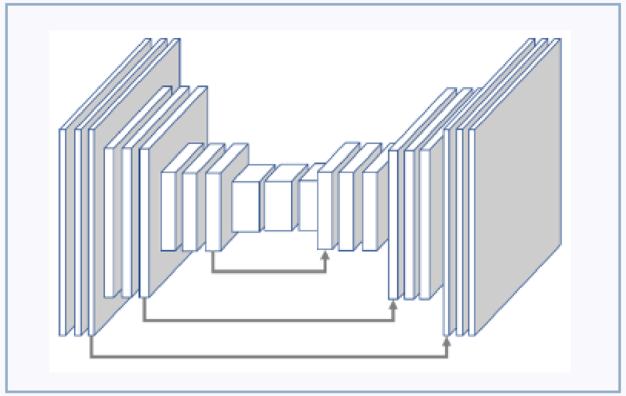
LSTM

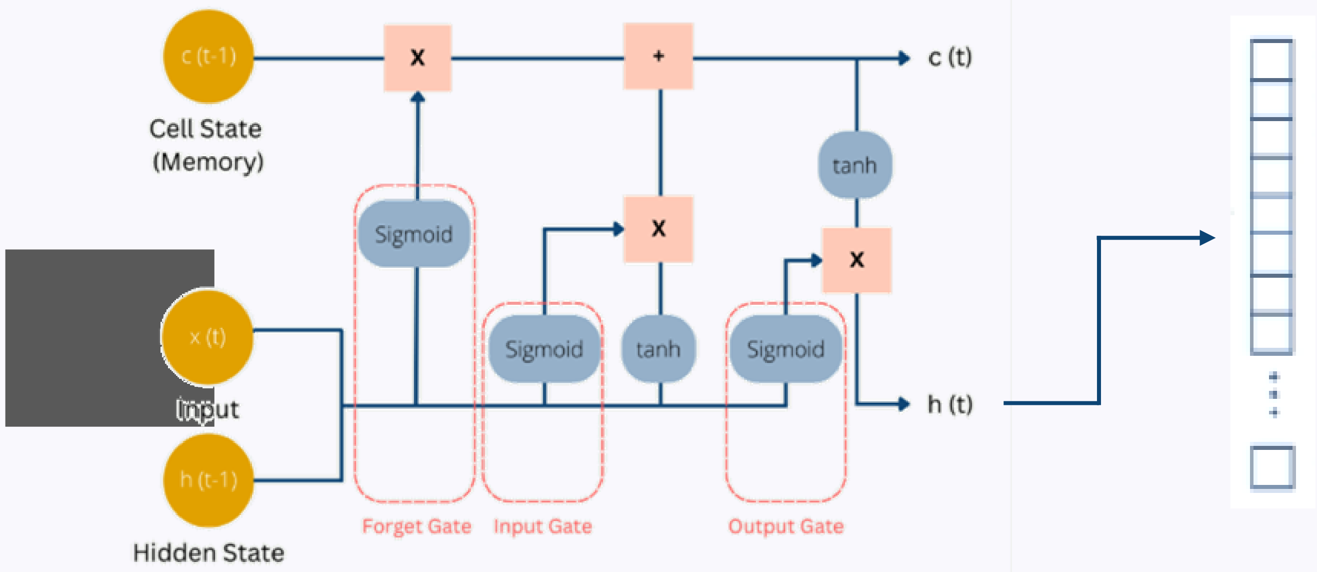
classifies the fractured bones. The correlations between fracture and neighboring cervical spine is high

01
02
03
04
05
06

Methodology

Stage 2: **CrackNet + BiLSTM**





01
02
03
04
05
06

Experiment Results

1. Binary weighted log loss
   - RSNA competition metrics

$$BCE = -\frac{1}{N}\sum_{i=0}^{N} y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)$$

2. Jaccard index
   - Used only for stage 1 training
   - False Positive scaling

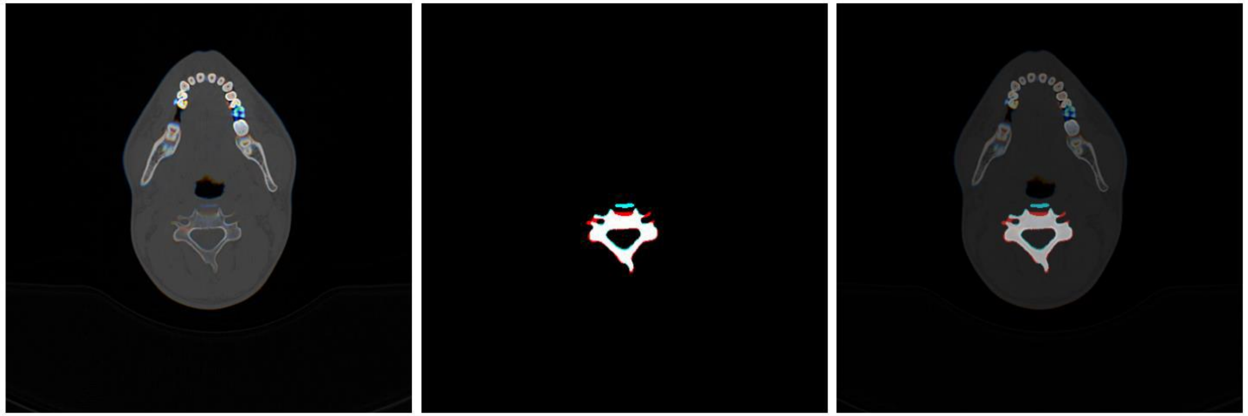
$$J = \frac{|A \cap B|}{|A \cup B|}$$

3. Dice coefficient
   - Precision calculation

$$DICE = \frac{2|A \cap B|}{2|A| + |B|}$$

Experiment Results

Stage 1 Model	Accuracy	Dice precision
UNet-EfficientNet with Windowing	0.999	1.000
UNet-EfficientNet with 2.5D	0.980	1.000
3D ResNet-101	0.82	0.52
EfficientNet-V2	0.953	-



Experiment Results

Stage 2 Model	Loss	Accuracy
CrackNet-LSTM without YOLOv5	0.21±0.01	0.943
CrackNet-LSTM with YOLOv5	0.20±0.01	0.949
YOLOv5	0.64	0.94
ResNet-50 without Cropping Voxel	1.6	0.4
Simple CNN-LSTM	0.32	0.87

Experiment Results

Stage 1+2 Model	Accuracy
UNet-EfficientNet + CrackNet-LSTM	0.949
UNet-EfficientNet + YOLOv5	0.94
UNet-EfficientNet + ResNet-50 no cropping voxel	0.4
UNet-EfficientNet + Simple CNN-LSTM	0.87
AlexNet + LSTM	0.71
ResNet + LSTM	0.71
VGGNet + LSTM	0.84
CNN + SVM	0.70
ResNet-50 + BLSTM-256	0.792

Conclusion

Lots of increase in the accuracy
Will help radiologists on missing detections

0.2 ± 0.01
Binary weighted log loss + Jaccard index

100%
Dice precision

94.9%
Accuracy

CONTACTS

1. **Doyeon Kim [Presenter] - D.Kim19@student.xjtlu.edu.cn**
2. Xujia Ning - Xujia.Ning21@student.xjtlu.edu.cn
3. Kaicheng Liang - Kaicheng.Liang21@student.xjtlu.edu.cn
4. Yi Ni - Yi.Ni21@student.xjtlu.edu.cn
5. Duan Wang - Duan.Wang21@student.xjtlu.edu.cn
6. Mingyuan Li - Mingyuan.Li21@student.xjtlu.edu.cn
7. Yichuan Wang - Yichuan.Wang21@student.xjtlu.edu.cn
8. Erick Purwanto - Erick.Purwanto@xjtlu.edu.cn
9. Ka Lok Man - Ka.Man@xjtlu.edu.cn