AI-driven Cervical Spine Fracture Detection from CT Scans

Md Mushfiqur Rahman Chowdhury, Soujatya Sarkar, Kriti Khare, Krishna Kumar Dixit

Medical Imaging EE-XXX [Instructor: Dr. Sneha Singh]

भारतीय प्रौद्योगिकी संस्थान मंडी
Indian Institute of Technology Mandi

Work Outline

Project Roles and Equipments

Project Roles

- **Md Mushfiqur Rahaman Chowdhury -** Objective Selection, Literature Survey and Report Writing
- **Kriti Khare -** Dataset Preparation and Augmentation.
- **Soujatya Sarkar -** Data Augmentation and Coding Implementation.
- **Krishna Kumar Dixit -** Methodology; Theoretical Support for the Biology and Mechanisms behind Cervical Spine and CT Scan.

Equipments and Resources

- Resources Used: 1 NVIDIA RTX A6000 48GB GPU (from Cheetah Server)
- Time taken to complete the project: 1.5 months (completed in last week of May)

The Spine

A Biological Overview

Spinal Cord. The spinal cord is a long, cylindrical structure made up of nervous tissue that extends from the base of the brain (medulla oblongata) to the lower back.

Cervical Spine. The cervical spine, located in the neck, consists of seven vertebrae labeled C1 through C7.

- C1 (Atlas): Supports the skull, and allows nodding motion.
- C2 (Axis): Features the odontoid process (dens), allowing rotational movement of the head.
- C3-C7: Provide structural support and flexibility

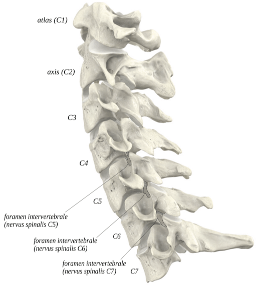

Fig. 1. The Cervical Spine

Cervical Spine Fracture

A Brief Introduction

Cause

- Trauma: Motor vehicle accidents, falls, sports injuries, or violent incidents.
- Osteoporosis: Weakening of the bones makes them more susceptible to fractures.
- Cancer: Metastasis to the spine can weaken vertebrae.

Symptoms

- Neck pain: Severe and persistent pain, especially when moving the neck.
- Neurological symptoms: Numbness, weakness, or paralysis in the limbs.
- Swelling and bruising: Around the neck area

Objective as Stated

Problem Statement

AI-driven Cervical Spine Fracture Detection from CT Scans of the Human Spine

- Implementing 3D Vision Transformer to enhance the accuracy and efficiency of fracture detection.
- Evaluate the potential impact of the developed model on clinical decision-making by measuring the reduction in diagnosis time.
- Contribute to the advancement of AI in medical imaging by providing radiologists and doctors with a reliable tool for early and accurate fracture detection.

State-of-the-art

What has already been done?

Qishen Ha - Gold Medal Solution - Kaggle

- Model architecture: For segmentation - EfficientNet v2s + UNet model; For classification - 2.5D CNN + LSTM classification. [We are only considering classification]
- Best Model - Loss: 0.2047.

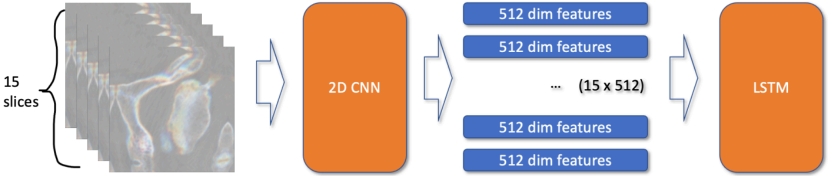

Fig. 2. Architecture of 2.5D CNN + LSTM architecture.

State-of-the-art

What has already been done?

Yusuke Uchida - Silver Medal Solution - Kaggle

- Model architecture: For segmentation - MONAI 3D UNet; For classification - 2.5D CNN (EfficientNetV2-L) + Transformer encoder. [We are only considering classification]
- Best Model - Loss: 0.4140.

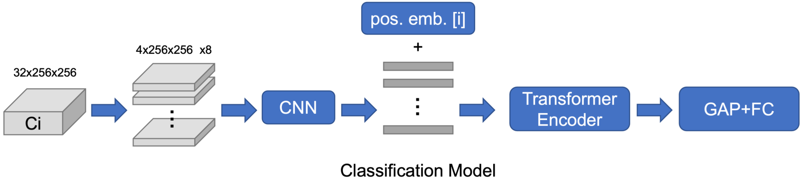

Fig. 3. Architecture of 2.5D CNN + Transformer encoder.

State-of-the-art

What has already been done?

Arunodhayan - Bronze Medal Solution - Kaggle

- Model architecture: For segmentation - MONAI 3D UNet; For classification - EfficientNet-V2-S classifier. [We are only considering classification]
- Best Model - Loss: 0.4982.

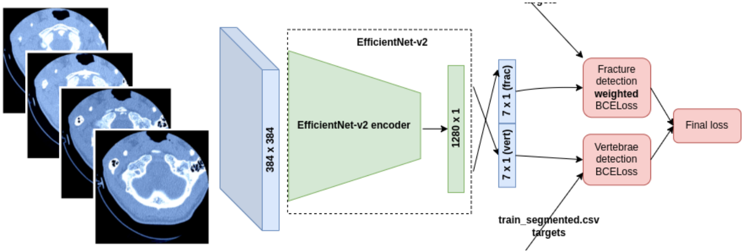

Fig. 4. Architecture of EfficientNet-V2-S classifier.

Methodology

The Dataset

We screened the 2019 patient studies from the full RSNA Cervical Spine Dataset of size 189 GB to extract 12 GB of data consisting of 86 studies for which segmentations were already given in the dataset.

Total CT Scan Dataset Split:

- 70 CT scans (Train)
- 14 CT scans (Val)
- 2 CT Scans (Test)

Also, only 210 slices are required per study as they represent the cervical spine [C1-C7] (according to NiFTi file metadata).

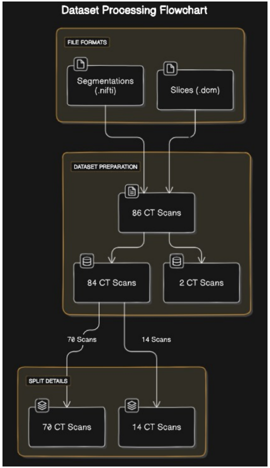

Fig. 5. Dataset Preparation Pipeline.

Methodology
Data Augmentations

Every CT Scan study of a patient is a folder of DICOM files containing all the slices –> (Folder name: StudyInstanceUID; Files: 1.dcm, 2.dcm, 3.dcm, ...).

For each patient study, we had a single NiFTi (.nii) file containing segmentations in sagittal orientation for each corresponding DICOM slice, but index starts from 0. Hence, DICOM[n] ==> NiFTi[n-1].

First, We make sure to convert the NiFTi file to axial orientation, as DICOMs are in axial orientation. Then, the augmentation we do is we create a '.npy' file for each slice which contains a 6-channel image ==> [DICOM(n-2), DICOM(n-1), DICOM(n), DICOM(n+1), DICOM(n+2), NiFTi(n-1)]. We only keep the odd slices as every even image will be redundant data.

This arrangement helps us to get a volumetric approach of our fractured area, helping in prediction in Attention based networks.

Methodology

Proposed Architecture

Applying the 3D Vision Transformer on the resulting 105 npy files generated from the 210 slices present in each CT Scan. We used Binary Cross-Entropy Loss to evaluate our model.

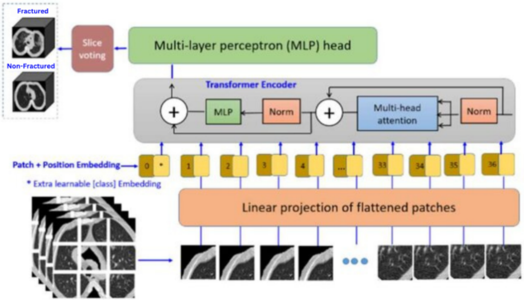

Fig. 6. Architecture of 3D Vision Transformer.

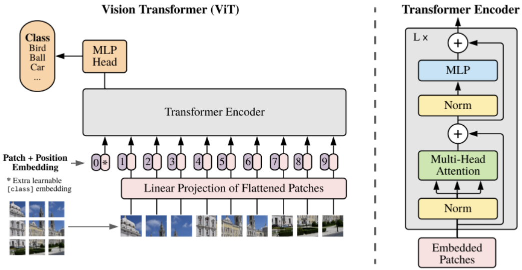

Fig. 7. Architecture of Vision Transformer.

Comparative Study

Model	Train Loss	Val Loss
2.5D CNN + LSTM classification	0.2117	0.2047
2.5D CNN + Transformer encoder	0.4049	0.4140
3D ViT B2 [on 1/25 data]	0.4169	0.4158
3D ViT B1 [on 1/25 data]	0.4893	0.4879
EfficientNet-V2-S classifier	0.4812	0.4982

Ablation Study

Model	Train Loss	Val Loss	Train Accuracy	Val Accuracy
3D ViT [Batch - 1; LR - 3e-4; S - 0.0]	0.4893	0.4879	90.2899	90.4176
3D ViT [Batch - 1; LR - 5e-4; S - 0.1]	0.7179	0.8186	78.6929	75.069
3D ViT [Batch - 2; LR - 3e-4; S - 0.1]	0.4169	0.4158	93.6419	94.4136

Loss Curves

Model 1 Loss
Model 2 Loss
Model 3 Loss
Train Loss
Val Loss
Loss
Epoch
0
25
50
75
100
125
150
175
200

Future Direction

What can make our Model good enough to be the SOTA on the whole dataset

For now, we have only trained the model on a limited dataset, i.e., the model has been trained on just 86 patients overall, but we have a total data of 2000 patients whose segmentation maps are not provided in the original dataset.

- Due to the huge extent of this dataset, we expect to create an excellent algorithm which can beat the present SOTA provided we create a good segmentation map-generating network.
- The challenges we faced were restricted access to top-notch GPUs, hence we had to drop the plan of developing a segmentation map generation network.
- Larger dataset can help this model converge better and provide better performance.

AI-driven Cervical Spine Fracture Detection from CT Scans

Thank you!