Skull Fracture detection using Maximum Intensity Projections

By Yash D. Kopulwar (EE18B039)

Abstract –The traumatic brain injuries especially intracranial haemorrhages are associated with the presence of skull fracture. But they are not easy to detect accurately. Three-dimensional display helps a lot in medical imaging for analysing various organs in the human body. Yet, many a times medical practitioners study the slices of the 3D scan, which is a long process and not so accurate. Also, for automatic fracture detection the input size is very large if it is a three-dimensional image. So, we are first going to pre-process the CT scan image by thresholding. Then, we will convert the three-dimensional images to two-dimensional images using Maximum Intensity Projections and this will be given as input to a Depp Learning network for detection of the fractures.

I. INTRODUCTION

A skull fracture is a break in one or more of the eight bones that form the cranial portion of the skull. This usually occurs due to severe head injury or head trauma. Skull fractures are often associated with traumatic injury to the brain. The brain can be directly damaged due to the injury to the brain tissues, bruising or bleeding. The indirect affects include blood clot which can block blood flow to the brain cells or can cause intracranial pressure on the brain. These injuries can lead to life-threatening complications. Therefore, quick and accurate detection of skull fracture is very important.

This project aims to develop an MIP based method to detect and segment the skull fractures by using Deep Learning models. A similar technique was used to detect Pulmonary Nodules in the lungs [1].

II. DATASET

The dataset was acquired from Physionet [2]. The dataset includes CT scans for patients diagnosed with intracranial haemorrhage with the following types: Intraventricular, Intraparenchymal, Subarachnoid, Epidural and Subdural. Each CT scan includes about 30 slices with 5mm slice thickness. The dimension of images from axial view is 512x512.

III. METHOD

1. SKULL SEGMENTATION

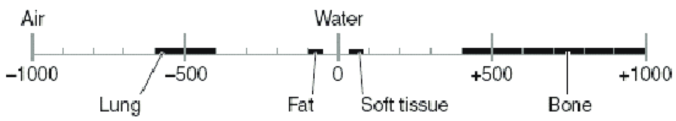
To remove the irrelevant regions like cloths, machine, soft tissues etc. we need to segment out the skull part. This will be done with the 3D CT image. The 3D CT image is made up of cubic voxels which are equivalent to pixels in a 2D image. The CT scanners use a set of software algorithm to calculate the amount of x-ray radiation absorbed by every element of any body organ. Each of these elements is represented by a voxel. The voxels are assigned a numerical value called CT-number.

The scale of CT-number is not so useful as it is set according to the machine which is recording the values. We need to scale it to some useful and interpretable values. Human body is mostly made up of water. So, it is better to use the Hounsfield Unit (HU) instead of CT-number.

The Hounsfield Unit (HU) is a relative quantity measurement of radio density with respect to the attenuation value of distilled water (at standard temperature and pressure). This CT-numbers are compared to the attenuation value of the water and displayed on the scale of arbitrary units. The HU scale is applied to the CT numbers by scaling all the attenuation constants according to the following linear transformation:

$$HU = \frac{\mu - \mu_{water}}{\mu_{water}} \times 1000$$

Here, the Hounsfield Unit for water is zero and -1000 HU for air. The HU for bones can reach up to 1000 HU, 2000 HU for dense bones, and more than 3000 HU for metals like steel and silver [3]. The Hounsfield scale is represented as a grayscale 3D image where the denser body organs appear brighter because of more x-ray absorption compared to less dense tissues. Fig. 1 shows the Hounsfield values of some body organs with respect to air and water.



Bone	+400 → +1000
Soft tissue	+40 → +80
Water	0
Fat	−60 → −100
Lung	−400 → −600
Air	−1000

Figure 1: Hounsfield values of some body organs

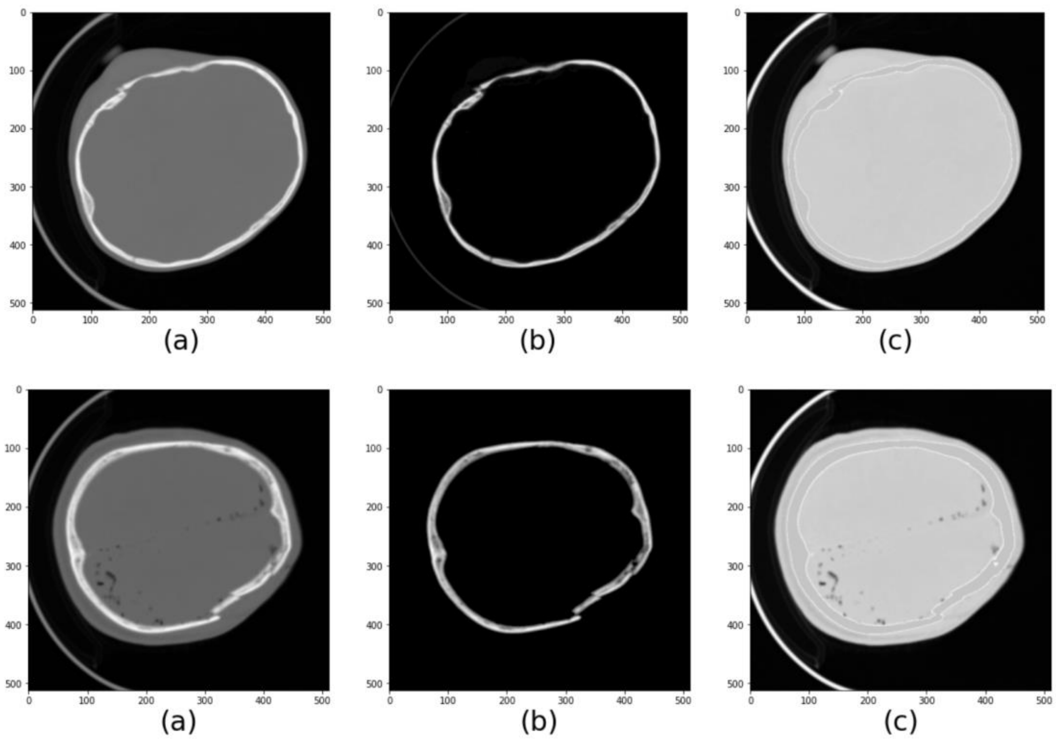


Figure 2: Here are two examples of segmentation by thresholding using Hounsfield values. (a) Leftmost images are of a slice of the CT image of the skull. (b) The middle images show voxels with HU values from 300 to 2000. (c) The rightmost image show voxels with HU < 300. (There is no distinction in the soft tissues because the CT image was downsampled)

Now, as we know the specific values for every body organ, we can separate it in the CT image by selecting a specific window of Hounsfield values. This is called as Windowing or grey-level mapping. Windowing will change the appearance of the image to highlight a specific type of structure. All this will be done to the 3D CT image. Fig. 2 shows windowing using Hounsfield units. This example was done on single 2D slices of CT images of two different skulls. The leftmost images (a) are of normal CT scan which contains both soft tissue and bones. The middle image (b) has HU values between 300 and 2000. Therefore, the skull bones and the fractures are clearly visible. The rightmost images (c) which has HU values less than 300, shows the soft tissues and some part of the CT scanning machine.

2. MAXIMUM INTENSITY PROJECTION

The CT images are three-dimensional. The traumatic brain injuries are checked mostly by analysing several axial slices of the CT scan. It is not easy to estimate the total damage done to the skull by looking only at the slices. And checking these many slices takes lots of time. The 3D rending of the CT scan also takes too much time and resources. Therefore, the radiologists use Maximum Intensity Projection (MIP) (formerly known as Maximum Activity Projection [4]) to project the 3D image on a 2D plane.

To understand MIP, we can take the example in Fig. 3. For average projection, an average of all the slices of the CT scan is taken and considered as the final image. But in MIP, for each x-y coordinate only the pixel with the highest Hounsfield number along the z-axis is projected for the final image. This helps to observe the dense structures in the scan. Either the whole image can be used to make the MIP image or several MIP images can be made by using slabs of various thicknesses. Slab is a part of the 3D image which consists of consecutive frames in the z-direction. This becomes useful while analysing high resolution CT scans.

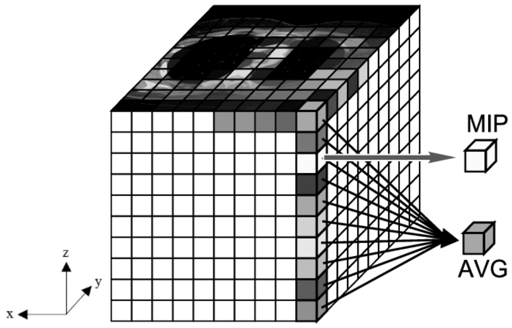


Figure 3: Difference between averaging (AVG) and creating MIP of a CT image. [5]

In this project, we are dealing with CT scans of skull. If we apply MIP on whole image, the intensity of the bones in the jaw and neck will interfere with the final image. The fractures in the skull will not be properly visible. As there will be some high intensity voxels below the fracture somewhere in the whole scan, they will show up in place of the fracture in the final MIP image. To prevent this, we need to look specifically to the calvaria, which is the top part of the skull.

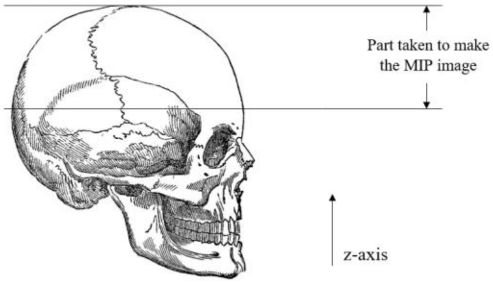


Figure 4: The upper part (Calvaria) was taken to detect the fractures

For detecting a wide range of fracture sizes, we need to check MIP images of different slab thickness like 1 mm, 5 mm, 10 mm and 15 mm. It also helps reducing the false positives. Fig. 5 shows the MIP images (5mm) without and with windowing. We can see that the fracture is clearly visible from starting to the end.

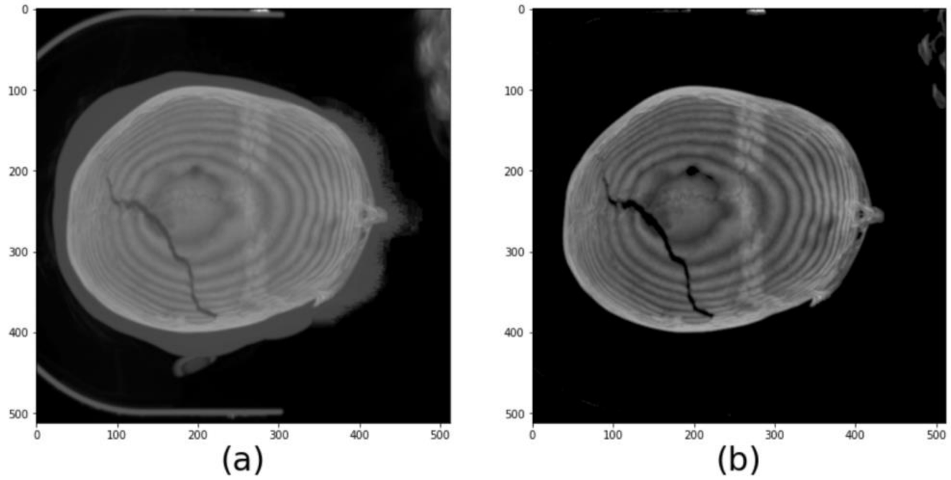


Figure 5: (a) MIP with no windowing (b) MIP after applying windowing

3. MIP-BASED CONVOLUITONAL NETWORK

As shown in [1], we can use the U-net [6] deep learning model for detection in the CT images. It consists of an encoding part and a decoding part. The input to the model will be the MIP images with different slab thicknesses. Each image will be sent separately to the U-net network and then will be merged.

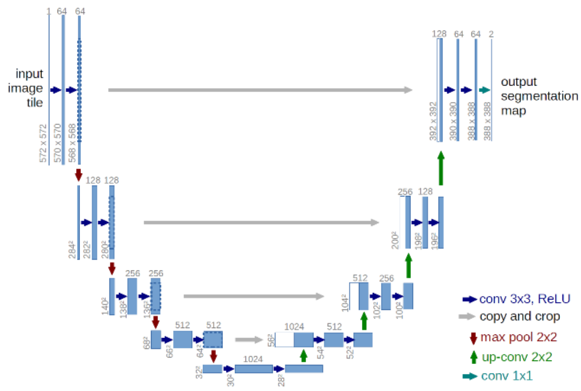


Figure 6: The architecture of U-net

The encoding part consists repeated blocks of two convolution layers and max pooling. The convolution layer has 32-component feature map at the starting and then it doubles after every application. Batch normalization and ReLU is used after every layer. The decoding part is like reverse of the encoding part. It consists of up-sampling layers followed by concatenation with the corresponding feature map from the encoding period. The decoding part has 128-component feature map in the starting and it halves after each block.

As said earlier, all the outputs of the U-net network from different slab thicknesses will be merged. The contours of these different outputs are combined to refine the final output.

IV. FUTURE WORK

The data used here was downsampled by a large extent. In high resolution images we can alter the slab parameter to get required MIPs. We can use 4-connected neighbourhood operator [7] to separate the unrelated tissues from the bone tissue. The 4-connected neighbourhood operator can be applied in 3 dimensions instead of applying only on 2D slices. This will be beneficial in high resolution 3D CT images.

Instead of simple MIP, different techniques can be used like curved-slab MIP or rectangular-slab MIP. The detection method can also be compared to the segmentation method [8].

V. REFERENCES

[1] Zheng S, Guo J, Cui X, Veldhuis RNJ, Oudkerk M, van Ooijen PMA. Automatic Pulmonary Nodule Detection in CT Scans Using Convolutional Neural Networks Based on Maximum Intensity Projection. IEEE Trans Med Imaging. 2020 Mar;39(3):797-805. doi: 10.1109/TMI.2019.2935553. Epub 2019 Aug 15. PMID: 31425026.

[2] Hssayeni, M. (2020). Computed Tomography Images for Intracranial Hemorrhage Detection and Segmentation (version 1.3.1). PhysioNet. https://doi.org/10.13026/4nae-zg36.

[3] https://www.ncbi.nlm.nih.gov/books/NBK547721/

*[4] Wallis JW, Miller TR, Lerner CA, Kleerup EC (1989). "Three-dimensional display in nuclear medicine". IEEE Trans Med Imaging. **8** (4): 297–303. doi:10.1109/42.41482. PMID 18230529.*

[5] softMip: A Novel Projection Algorithm for Ultra-Low-Dose Computed Tomography - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Maximum-intensity-projection-and-AVG-algorithms_fig1_51397092 [accessed 18 Dec, 2021]

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation." pp. 234-241.

[7] Digital Image processing 3rd edition, Rafael C. Gonzalez & Richard E. Woods

[8] Shan, Wei & Guo, Jianwei & Mao, Xuewei & Zhang, Yulei & Huang, Yikun & Wang, Shuai & Li, Zixiao & Meng, Xia & Zhang, Pingye & Wu, Zhenzhou & Wang, Qun & Liu, Yaou & He, K.L. & Wang, Yongjun. (2021). Automated Identification of Skull Fractures With Deep Learning: A Comparison Between Object Detection and Segmentation Approach. Frontiers in Neurology. 12. 687931. 10.3389/fneur.2021.687931.