

Detection of Intracranial hemorrhage and its subtypes

Dipam Vasani
dvasani@iu.edu

Prajakta Phadke
pphadke@iu.edu

Siddhesh Mirjankar
smirjank@iu.edu

Biraj Parikh
bparikh@iu.edu

Abstract—Intracranial hemorrhage is one of the top 5 fatal diseases in the US. In this paper, we aim to automate the process of detecting these bleeds in order to relieve overworked radiologists and flag patients in need of immediate attention.

Keywords—hemorrhage, radiology, image classification, transfer learning, bag of tricks

I. INTRODUCTION

Hemorrhage in the brain (Intracranial Hemorrhage) is one of the top five fatal health problems. The hemorrhage causes bleeding inside the skull (typically known as cranium). It accounts for approximately 10% of strokes in the U.S thus diagnosing it quickly and efficiently is of utmost importance.

Intracranial Hemorrhage can have many consequences depending on the location, size, type of bleed in the brain. The relationship is not that simple to interpret. A small hemorrhage might be equally fatal as a big one if it is in a critical location. If a patient shows acute hemorrhage symptoms such as loss of consciousness, then an immediate diagnosis is required to carry out the initial tests but the procedure is very complicated and often time-consuming. The delay in diagnosis and treatment can have severe effects on patients, especially the ones in need of immediate attention. The solution proposed in this project aims to make this process efficient by automating the preliminary step to identify the location, type, and size of the bleed, the doctors can provide a better diagnosis and a faster one as the work of identification is done by the algorithm. Also, the algorithm treats every scan with the same scrutiny, reducing human error and learning more as it goes.

Thus, the challenge is to build an algorithm to detect acute intracranial hemorrhage and its subtypes. This project will be a step towards identifying each of these relevant attributes for a better diagnosis.

II. BACKGROUND AND RELATED WORK

Deep learning has been used extensively in medical imaging in the last decade whether it is detecting diabetes using retinopathy or pneumonia using Chest X-ray. Image segmentation has been used for color coding scans which is then used for applications like automatic measurement of organs, cell counting, or simulations based on the extracted boundary information. A recent paper by the Qure.ai team [1] shows the use of deep learning for classifying MRIs. Their approach included the use of CNNs as well as NLP techniques on the corresponding medical reports. The dataset included 313,318 anonymous CT scans from several centers in India. One major problem they faced was that of 3D scans. The mean age of the dataset was

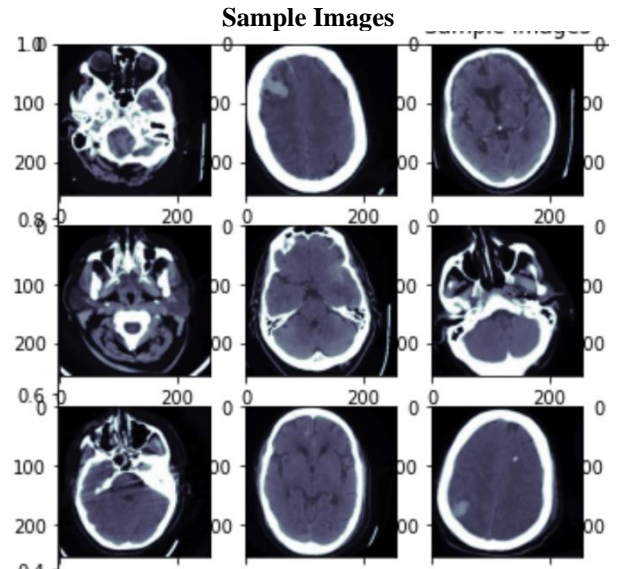
43.4 years and 42.87% of the population was female. They managed to achieve an AUC of 0.94 ± 0.3 on all the sub categories. Other interesting papers include Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning.[2], Extracting 2D weak labels from volume labels using multiple instance learning in CT hemorrhage detection.[3] In the latter, they used bags of 2D slices and calculated the losses per slice. These losses were then max pooled together to calculate the gradient and update parameters of the model. However, they used only about 600 images for this project.

III. PROGRESS SO FAR

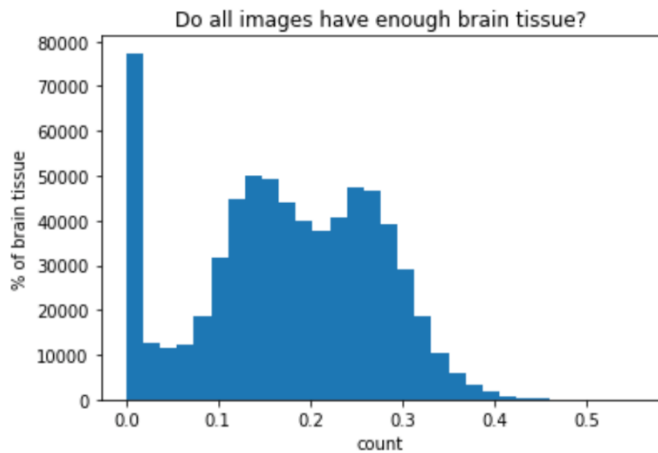
This section explains in steps the work done till date.

A. Data Cleaning

The data is given in the form of DICOM files. A DICOM file consists of metadata and image data packed into a single file. The information in the metadata consists of various tags. It also removes confidential information about patients to maintain the integrity. We have 194082 images in the dataset. We start by converting the metadata into a dataframe since the DICOM format is slow to use. We then view some of the images which are in the form of slices from top of the brain to the bottom of the brain as shown.



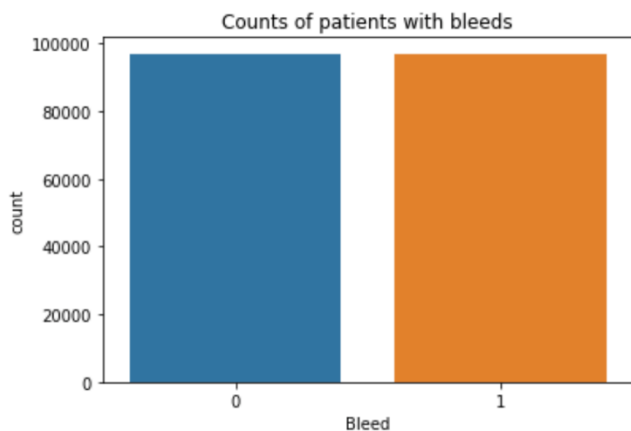
Sometimes the slice can be completely blank (a slice before or after the brain). To remove such slices, we use a column called “img_pct_window”. It tells us the percentage of brain pixels in the slice. We remove images with less than 20% value in that column.



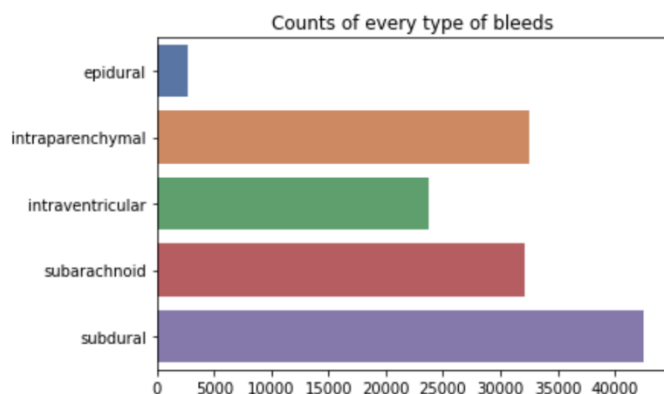
We also fix the column Rescale_Intercept, resize and crop the images to (256,256) and save them as “.jpg”.

B. Data Exploration

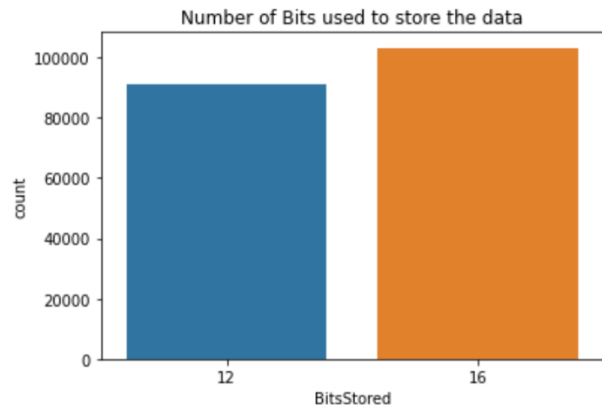
We look at the distribution of the data.



We see that we have a pretty balanced dataset when it comes to images with and without bleeds. We also check the counts of the different subtypes of bleeds.

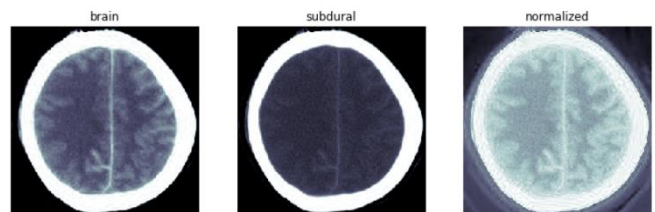


We see that “subdural” is the most common type of bleed present. We then move onto analyzing the metadata. One interesting column in the “bits stored” column which indicates the number of bits used to store the data. It has two distinct values: 12 and 16.



This might indicate that the data might have come from two different organizations. It is usually better for deep learning models that the data comes from same distribution but, we will move forward with this data.

Finally, the images can be viewed in many different windows.



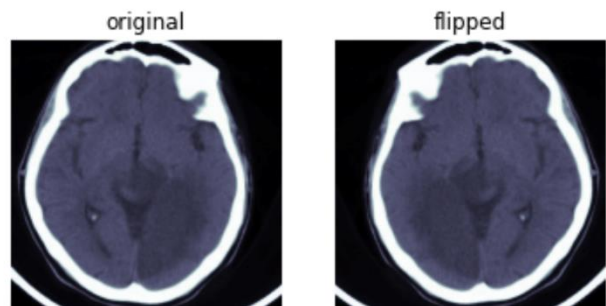
However, this is for the human perception. A neural network accepts floating point data and does not require any windowing.

C. Data Augmentation

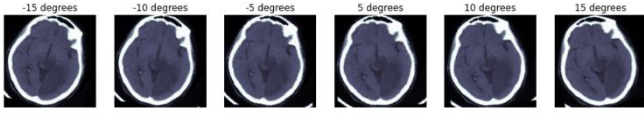
The first step to train a good deep learning model is to have lots of data. However, this is not always possible. Hence, we do small random transformations on the data that do not change what is inside the image(for the human eye) but changes the pixel values. Models trained with Data Augmentation generalize better.

Below are some of the transformations which can be used for our data.

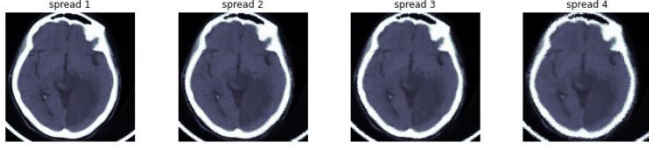
Flip: The bleed can occur on any side of the brain. Hence, flipping gives the model more cases to consider.



Rotation: The slices do not always appear straight. They might have different orientations. Hence, we randomly rotate few images in range of (-15, 15) degrees.



Blur: During an MRI, if the patient moves, the MRI might become blur. Hence, we apply blurring effect on some images. The blur effect is increased sequentially left to right.



D. Modelling

Recent research has shown that metadata can prove useful in image classification. Our initial goal was to use a combination of metadata and images for the task. We try to classify the hemorrhages solely using the metadata to gauge its usefulness. We find that even after using a robust model like Random Forest, we get an accuracy of 50%. Hence, we decide to discard the metadata and focus on the images itself.

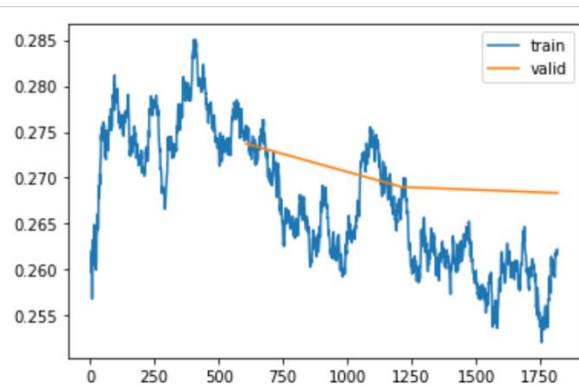
For our baseline model, we start with a resnet18 model. Transfer learning has shown good results when it comes to image classification. It also reduces training time and resource consumption significantly.

We use Leslie Smith's learning rate finder [5] to find a good learning rate. The batch size has been set to 256. We randomly sample 20% of the data as validation set. For hyperparameter tuning, we use Leslie Smith's one cycle training policy [4].

In this method, we initially freeze the the pretrained part of the model and only train the new head. However, we do update the BatchNorm layers of the resnet to maintain mean = 0 and std deviation = 1 of each layer. We are able to achieve ~89% accuracy with and without pretraining in just 5 epochs.

| epoch | train_loss | valid_loss | error_rate | accuracy | time |
|-------|------------|------------|------------|----------|-------|
| 0 | 0.268112 | 0.266647 | 0.111629 | 0.888371 | 11:24 |
| 1 | 0.264794 | 0.265100 | 0.110753 | 0.889247 | 11:26 |
| 2 | 0.262578 | 0.265090 | 0.110959 | 0.889041 | 11:20 |

Plot of the loss is as shown below:



IV. REVISED RESEARCH PLAN

We plan to conduct the following experiments in the coming 4 weeks:

1. Multilabel classification: (Week 1)

We will take the images which have bleeds in them and classify them into every subcategory present.

2. Different architectures: (Week 2)

Resnet34, Resnet50, Resnet101, Densenet121, Alexnet

3. Progressive image resizing: (Week 3)

Research has shown that training a model on smaller sized images and then gradually increasing the size during the train process helps the model learn better.

We intend to use sizes 64, 128, 224, 256

4. Mixed Precision Training: (Week 3)

We will attempt to reduce the training time with this method. In this method, we perform some operations using 16 bit data and others using 32 bit data (weight update which requires higher precision).

5. Testing and deployment: (Week 4)

We have data from a separate company Qure.ai. We will show this data to the model only once to see how well it performs. We also plan to deploy our model on a simple website where one can upload an MRI and get results.

REFERENCES

- [1] Sasank Chilamkurthy, Rohit Ghosh, Swetha Tanamala, Pooja Rao and Qure.ai team. Development and Validation of Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans.
- [2] Weicheng Kuo, Christian Häne, Pratik Mukherjee, Jitendra Malik, and Esther L. Yuh, Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning.
- [3] Samuel W. Remedios, Zihao Wu, Camilo Bermudez, Cailey I. Kerley, Snehashis Roy, Mayur B. Patel, John A. Butman, Bennett A. Landman, and Dzung L. Pham, "Extracting 2D weak labels from volume labels using multiple instance learning in CT hemorrhage detection".
- [4] Leslie N. Smith, "A disciplined approach to neural network hyperparameters: Part 1 -- learning rate, batch size, momentum, and weight decay".
- [5] Leslie N. Smith, "Super-Convergence: Very Fast Training of Neural Networks Using Large Learning Rates".