Hemorrhage Detection
Individual Report - Swetha

Introduction:

The workflow of the project was laid out as follows:

The first week was set for the proposal of the project, the EDA analysis and Preprocessing of the images.

The next week was utilized for researching relevant methods to build models. Model 1 & Model 2 were built for the prototype data and Model 3 & Model 4 were built for the actual data.

These models were then ensembled to make classification and predictions, respectively.

The third week was used to assign proper evaluation metrics to test our models.

The fourth week was set aside for creating the presentation and to make sure everything is working properly.
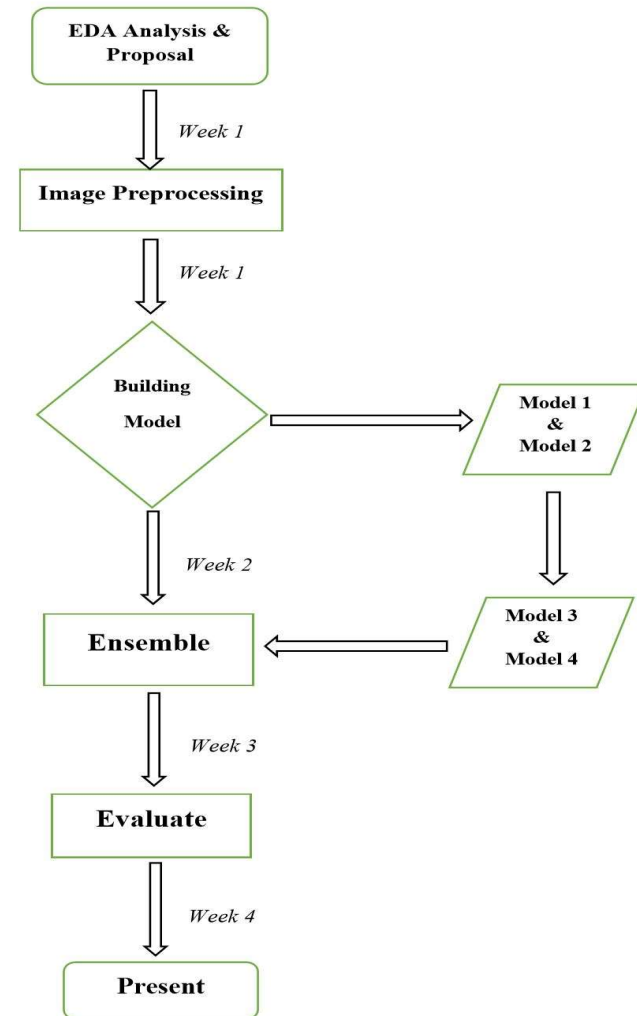
Description of my work:

(a) This was my first time working with images, let alone DICOM images. The problem by itself is quite challenging and took some time for me to understand the entire workflow. With an immense help from my teammate, I was able to investigate her parts of analysis and get a grasp of the concepts. To aid with the preprocessing, I wanted to create masks based on Hounsfield units to differentiate the brain window from the bone window. I was able to create masks, which I then fine-tuned using erosion and dilation techniques to remove tiny features like pulmonary vessels or noise. For this, I created a function called make_mask which will:

1. Standardize the pixel value by subtracting the mean and dividing by the standard deviation
2. Identify the proper threshold by creating 2 KMeans clusters comparing centered on foreground (soft tissue/bone) and background.
3. Using Erosion and Dilation, which has the net effect of removing tiny features like pulmonary vessels or noise.
4. Identify each distinct region as separate image labels
5. Using bounding boxes for each image label to identify which ones represent brain and which ones represent "everything else"
6. Create the masks for bone.
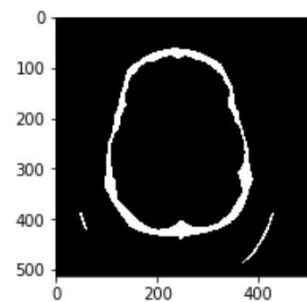7. Apply mask onto the original image to erase voxels outside of the brain.

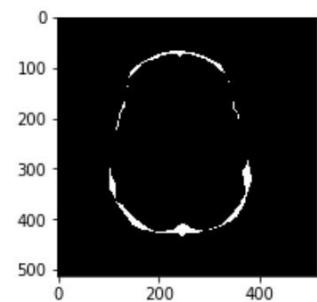The results of this process are as follows:
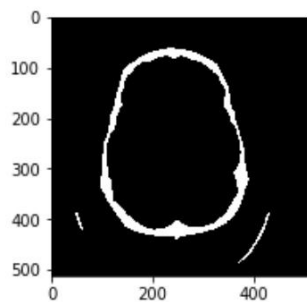
Mask of bone
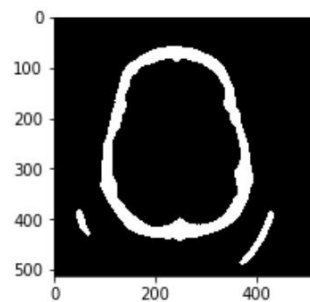
Dilation: Increasing the size of the mask

Close: Filling in the holes of the mask

Erosion: Removing pixels around the edges

Although the code I used to create these masks was not used in the preprocessing of our model, it was still crucial in understanding of biomedical image analysis. There were some issues with this function, and the results were not functioning as desired.
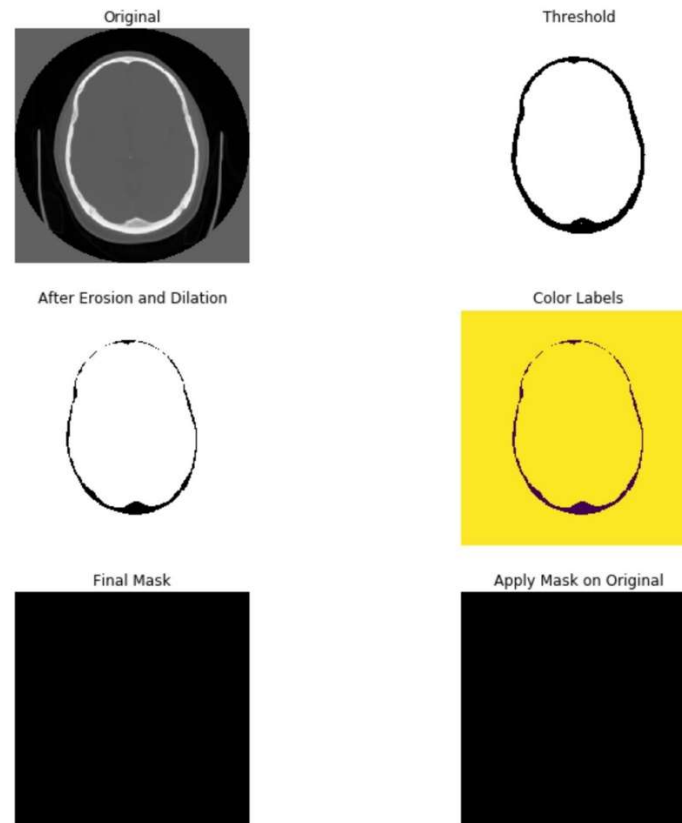
(b) Based on the code generated for the prototype model, I was able to use the same code to replicate Model_4 with inception_V3. Although, I still could not get the code to run without any errors, with the help of my teammate debugging the code, we were able to run the model successfully.

Portion of my work:

My work majorly involved understanding the preprocessing of the images and trying to build the final Model *(Model_4)* to make multilabel classification.

Results:

We were able to create a model that correctly classifies the type of hemorrhage, even an MRI scan has more than one type of hemorrhage.

Percentage of Code:

Most of the code I used to preprocess the images was not used in deployment, and Model_4 was essentially adding Inception_v3 as the pre-trained network instead of vGG16. To quantify the numbers based on the metrics given, less then 5 percent of the total code.

References:

We investigated various resources such as Kaggle Discussion Blogs, Published Article[1], Published Article[2], Deep Learning for Medical Image Analysis and Online courses to make sure we attain sufficient background in regards to Biomedical image analysis. As this was a Kaggle competition, we also investigated the solutions of second position and fifth position, both of which were extremely complicated.