# Milestone Progress Report

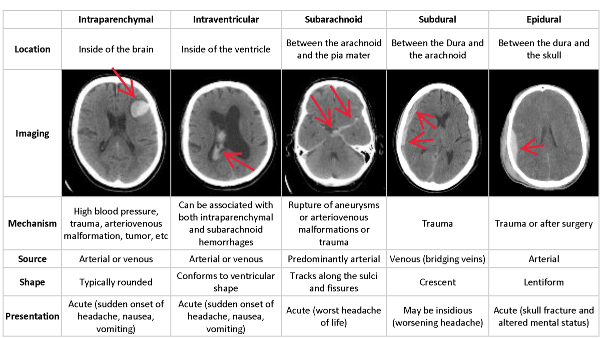
Brain Hemorrhage Classification of CT Scan Images Using CNNs

Josue Antonio Theodore Hattenbach Mara Hubelbank Sahil Sangwan Yinglin Wang
Department of Mathematics, Northeastern University

April 16, 2022


## 1 Abstract

The aim of this project is to train a Convolutional Neural Network (CNN) model for image classification, using Zeta Surgical's labeled CT scan image dataset. When provided with an unlabeled brain CT image, the model should provide an accurate prediction of which class the bleeding belongs to. As shown in the figure below, each image is classified as epidural, intraparenchymal, intraventricular, subarachnoid, subdural, or as belonging to multiple of the previous classes (”multi”).



| | Intraparenchymal | Intraventricular | Subarachnoid | Subdural | Epidural |
|---|---|---|---|---|---|
| Location | Inside of the brain | Inside of the ventricle | Between the arachnoid and the pia mater | Between the Dura and the arachnoid | Between the dura and the skull |
| Imaging | | | | | |
| Mechanism | High blood pressure, trauma, arteriovenous malformation, tumor, etc | Can be associated with both intraparenchymal and subarachnoid hemorrhages | Rupture of aneurysms or arteriovenous malformations or trauma | Trauma | Trauma or after surgery |
| Source | Arterial or venous | Arterial or venous | Predominantly arterial | Venous (bridging veins) | Arterial |
| Shape | Typically rounded | Conforms to ventricular shape | Tracks along the sulci and fissures | Crescent | Lentiform |
| Presentation | Acute (sudden onset of headache, nausea, vomiting) | Acute (sudden onset of headache, nausea, vomiting) | Acute (worst headache of life) | May be insidious (worsening headache) | Acute (skull fracture and altered mental status) |

Moreover, the data is pre-processed to remove ”flagged” samples, which are presumed to be labeled incorrectly or correspond to a corrupted image. It is additionally processed to grayscale format and down-sampled; then, it is fed to a CNN model which uses a sparse categorical cross-entropy loss function, Adam optimizer, and learning rate configured with gradient descent. As of the most recent training/testing iteration, our model has reached a testing accuracy score of 0.989; thus, it is well on its way to a well-generalized model for classification of CT brain scan data.


## 2 Summary of Completed Work

For this project, we distributed the work into the roles of cleaning data (Josue / Teddy), data analysis and visualization (Mara) Data pre-processing and CNN - David, Sahil, analysis and report writeup (all), editing report (Mara), and coming up with questions for Zeta Surgical (Teddy). The group met weekly to discuss progress and collaborate on pieces as needed. Overall, we found that we communicated well as a group, and the results thus far have been promising.

Notably, over 90% of the work for this milestone turned out to be loading in the images and pre-processing them. After loading in the images by class, we converted the colored images to grayscale format. Then, we removed all flagged entries from the combined set. The remaining set was split into training / testing / validation sets using 0.8 / 0.15 / 0.05 ratios. Each of these sets were then down-sampled using a rate of 16. The original images were (512, 512) and after down-sampling, their dimensions are (128, 128).

For our CNN model, the loss function is sparse categorical cross-entropy because our label was a single column with numbers 0 to 5 representing the classes. This also meant that the last dense layer of the model has 6 units. The optimizer we decided to use is Adam after testing out Adagrad, AdaDelta, and RMSProp. We plan on also doing gradient descent to find the optimal learning rate like we did for Lab 4. Due to time constraints, we only looked at the brain window of the CT scans. We may try to incorporate the other windows in our model too. The Keras model breakdown is shown after the results section.

## 3 Results

Using our first version of the CNN model, we saw 0.945 accuracy on the training set and 0.980 accuracy on the testing set. However, the model had 1.8 million parameters and each epoch ran for 20 minutes. We thought it was too excessive, so we simplified our model to only 9 hundred thousand parameters. The new model had greatly improved run-time with each epoch only taking 5 minutes. As expected, the training accuracy after 1 epoch was significantly lower at 0.915. The training accuracy went up subsequently after training multiple epochs. Ultimately, the testing accuracy was 0.989.

```
Model: "sequential_5"
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
conv2d_10 (Conv2D)           (None, 126, 126, 32)      320       
_________________________________________________________________
max_pooling2d_10 (MaxPooling (None, 63, 63, 32)        0         
_________________________________________________________________
conv2d_11 (Conv2D)           (None, 61, 61, 64)        18496     
_________________________________________________________________
max_pooling2d_11 (MaxPooling (None, 30, 30, 64)        0         
_________________________________________________________________
flatten_5 (Flatten)          (None, 57600)             0         
_________________________________________________________________
dense_10 (Dense)             (None, 16)                921616    
_________________________________________________________________
dropout_5 (Dropout)          (None, 16)                0         
_________________________________________________________________
dense_11 (Dense)             (None, 6)                 102       
=================================================================
Total params: 940,534
Trainable params: 940,534
Non-trainable params: 0
_________________________________________________________________
```

# 4 Questions for Zeta Surgical

1. What is the difference between the 2 subdural label files?
2. What does “Gold Standard” mean in context of the labeling state?
3. Where are the Intraventricular labels?
4. Do we classify the multi class?