Brain Hemorrhage Classification

Joseph Kwilman, Emmanuel Manolios, Aidan Domondon, Alissa Amchentseva, Jake Wallack

December 7, 2022

## Abstract:

*Image classification is the problem of assigning input images into classes, or categories, based on parameters of the image. Computer vision uses concepts in linear algebra like convolutions to identify characteristics of images that it then uses to determine the image’s class. Our mission is to classify brain hemorrhages in a data set of brain CT scans.*

**1. Introduction**

A brain hemorrhage is a type of stroke in which a ruptured blood vessel or burst artery leads to bleeding in the brain [1]. The bleeding can kill off brain cells, making hemorrhages serious medical emergencies. To prevent permanent damage or loss of life, patients need to be treated as fast as possible.

Intracerebral hemorrhages (ICHs), specifically, due to their location, require immediate attention. If they are left untreated, the bleeding will lead to clotting, causing a stroke. Thus, radiologists must also be mindful of this time constraint.

Reading computerized tomography (CT) scans of the brain can reveal what type of hemorrhage the patient has and where in their brain it is. Professionally trained radiologists have to analyze many subtleties in the scan that aren’t always obvious. Furthermore, a radiologist might not always be on standby, and reading the scans can be a complex, time-consuming endeavor.

Assisting radiologists in analyzing these images will improve efficiency, leading to better patient outcomes. Offloading some of the labor from specialists has an effect that is twofold; decreased reliance on radiological specialists–of which there are a limited number– has the potential to both decrease the cost and increase the accessibility of care.

More prospectively, automated classification of CT scans of brain hemorrhages could potentially provide insight into relevant research. Analysis of exactly which features characterize the presence of a brain hemorrhage could lead to the discovery of new indications of hemorrhaging.

Thus, it is our goal to develop a neural network to automatically classify CT scans with the hope of determining the type of brain hemorrhage. We believe that this would greatly improve patient outcomes, increase the accessibility and affordability of care, and potentially further our understanding of brain hemorrhaging.

**2. Related Work**

The use of machine learning in diagnosing medical conditions has shown promising results in recent years. The emergence of deep learning has paved the road for this progress. Neural networks can be trained to recognize patterns that it then uses to make predictions. The increased complexity and abstraction of its layers provide it an advantage over traditional, linear algorithms [2]. Some recent breakthroughs have been in the diagnosis of skin and breast cancer, as well as the grading of diabetic retinopathy[5].

Another interesting example is the use of a 2D convolutional neural network classifier for the diagnosis of intracranial hemorrhage (ICH). Using the 2019-RSNA Brain CT Hemorrhage Challenge, which has over 25,000 CT scans, the authors were able to get ACUs of around 0.95 for acute ICH detection[4].

**3. Formulation and Data Cleaning**

In order to get the data into a proper format for the neural network, there was a considerable amount of data pre-processing and cleaning that took place. There were two main steps we took in order to achieve this goal. First, we went through the individual images, and removed any which were excessively corrupted, or missing a brain. There were a few thousand images which fell in this category, and by removing them, we ensured that the training took place on only valid images.

The next step we took was to remove any labels which did not have a corresponding image. While there were roughly 100,000 valid images, there were roughly 700,000 labels for images. This meant that the vast majority of the labels were incorrect, and needed to be removed. We filtered through all the data, removing all invalid labels. Thus, we were left with roughly 100,000 labels, each of which corresponded to a valid image. We opted to limit our dataset to only use the max-contrast window images. This was done in order to lower the CPU processing power of our model, which was a limited resource.

Finally, it was necessary to downsample the data. This was because the images were so large that we could only load a couple hundred before we ran out of memory. We ran a few different models with different kinds of downsampling. For the neural networks we trained, we downsampled the data by a factor of 16. The first two neural networks were done using 1,750 different images in our training (250 of each class). By downsampling the data, we converted the input from a 2d array of size 512 by 512 to size of 128 by 128. For the final neural network, we used every single image in the dataset for training and validation.

**4. Implementation**

There were a few different neural networks which we implemented.

**Model 1**

For the first model, we implemented a standard neural network using Keras, a useful package for constructing neural networks from TensorFlow. We started with a 1d image array of size 32768. Then, we added two hidden layers. One had 300 nodes and the other had 100 nodes. Finally, we had another densely-connected layer, with a node for each possible classification.

**Model 2**

Next, we created a convolution neural network with Keras. In order for this to properly work, a number of steps had to be taken. First, we started with a 2d array of size 128 by 128. From here, we added four 2d convolution layers. The first layer had an output space dimensionality of 32, and a strides value of (4, 4). The second layer had an output space dimensionality of 64, and a strides value of (2, 2). The final two layers both had an output space dimensionality of 128, and a strides value of (2, 2). Once the convolution layers were added, the data was flattened back into a 1d array. Then, we added two hidden layers, one which had 300 nods, and another which had 100 nodes. Finally, we had another densely-connected layer, which had 8 different nodes, one for each possible classification.

**Model 3**

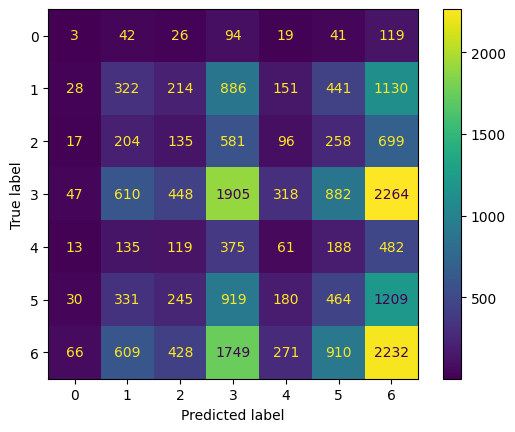
For our final model, we used tensorflow datasets to make training models on data much easier [3]. With tensorflow datasets, the images could be trained without overloading memory. This allowed us to train on the entire dataset, instead of a smaller sample of the dataset. Some additional useful features of tensorflow datasets were the ability to make batches of data and downsize data easily. The actual convolutional neural network used with the dataset was similar to model 2. It expects a 2d array of size 128 by 128. From here, there is a 2d convolution, plus a max pooling, plus a dropout layer. This pattern repeats 3 times. The dropout layers are useful for making sure our model does not overfit. The last layers flatten and condense the results into one of each possible classification.

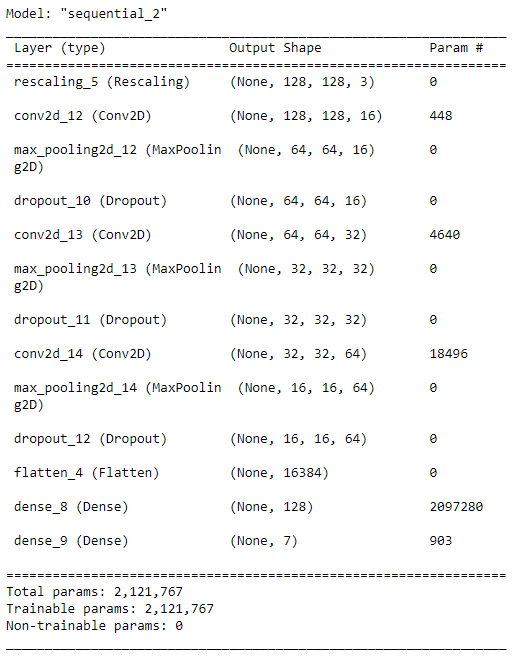
**5. Results**

The first two neural networks were unsuccessful at prediction. They both had a roughly 15% accuracy rating, which was slightly better than randomly guessing. The first neural network would guess almost exclusively normal. The second neural network would guess more evenly between each classification, but was still inaccurate. It was slightly better than the first neural network, but only marginally.

The final neural network was significantly more accurate. It had an accuracy rating of around 60%, which was significantly better than randomly guessing. This was our most successful model, and the additional complexity it contained, along with tensorflow datasets, helped it achieve this result. Figure 2 includes a graph showing the accuracy and loss of the model over 10 epochs of training. You can see how the validation accuracy rises over the training period, up to around 60%. We’ve also included the confusion matrix below. From it we can see that the model favored certain classifications (3-multiple hemorrhages and 6-subdural hemorrhages) over others. This is due to an imbalance in the amount of photos associated with each class. Since there were more photos in these categories, the network was rewarded more often for guessing these classes.







**6. Future Works**

Going forward, our goal is to improve the accuracy of the model’s hemorrhage classification. This goal could come via further data cleaning or in the refinement of the model’s construction. In terms of cleaning the image dataset, the most apparent problem is that not all of the images in the set are clean images of brains. This problem was quite apparent in the images which were classified as normal, but also affected other image classifications. Cleaning the dataset of these poorly taken images is a long and tedious process, but could greatly improve the model’s accuracy. In terms of improving the model itself, one means of doing so would be expanding the inputs of the model to include all four brain windows which are included in the dataset. A foreseeable execution of this model would pass each of the four images through an initial CNN model. The results of this model would be pooled and then sent into a second, separate CNN model to produce a classification. This is a commonly used approach for multi-image classification , but we did not incorporate it into our initial modeling process.

**7. References**

1. “Brain Hemorrhage (Bleeding): Causes, Symptoms, Treatments.” *WebMD*, WebMD, https://www.webmd.com/brain/brain-hemorrhage-bleeding-causes-symptoms-treatments.
2. Burns, Ed, and Kate Brush. “What Is Deep Learning and How Does It Work?” *Enterprise AI*, TechTarget, 29 Mar. 2021, <https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network#:~:text=Deep%20learning%20is%20a%20type,includes%20statistics%20and%20predictive%20modeling> .
3. “Image Classification: Tensorflow Core.” *TensorFlow*, <https://www.tensorflow.org/tutorials/images/classification>.
4. Wang, Xiyue, et al. “A Deep Learning Algorithm for Automatic Detection and Classification of Acute Intracranial Hemorrhages in Head CT Scans.” *NeuroImage: Clinical*, vol. 32, 2021, p. 102785., <https://doi.org/10.1016/j.nicl.2021.102785>.
5. Yu, KH., Beam, A.L. & Kohane, I.S. Artificial intelligence in healthcare. *Nat Biomed Eng* 2, 719–731 (2018). <https://doi.org/10.1038/s41551-018-0305-z>