Computer Vision Using Neural Networks and UNET

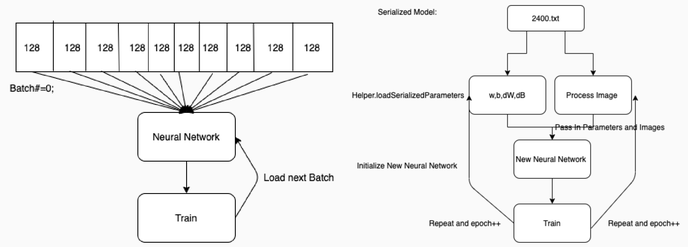
Javin Liu, Supervisor: Professor Joseph Vybihal
McGill University

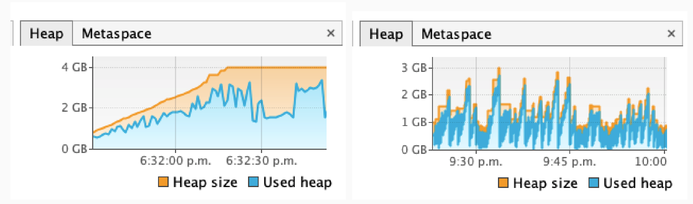
1.Introduction

Vision is one of the hottest topics in computer science today. It is very critical in robotics and in environments where we need to understand the surroundings. For this project, we have a data set of pictures that belong in 4 categories: open, hall, stairs, and room. We wanted to improve the accuracy as well as the speed of the training and prediction. The training time was cut by half by using multithreaded matrix multiplication and the pre-processing portion of our code has also been improved to be able to scale to larger and bigger datasets.

2. Loading images and training in chucks

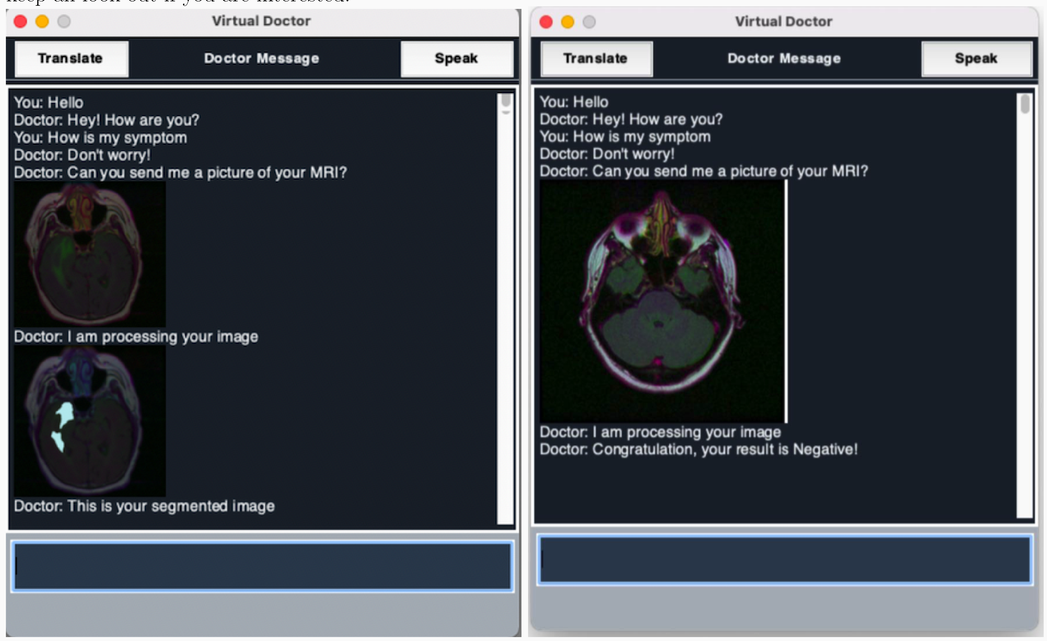
The previous year's code process all the images at once, and it loads the images one by one from the disk and then stores it. The benefit of this technique is that it is time-efficient. By loading all the images in one go, we only need to process the image once. However, we run out of heap space when we try to process more and more images. Hence, itis nearly impossible to scale the previous implementation of the code to work with larger data sets. To combat this issue, the code was changed to import the images in batches and then the neural network is trained on these batches. We de-reference the variables that previously held the imported images and rely on the garbage picker to free up space. This problem was solved by loading images and training in batches, as shown in the figure below. For larger data sets, our batch technique effectively keeps the heap usage under 3 GB while the algorithm runs out of heap if we load all images at once. This helps our algorithm scale to bigger data sets. Transfer learning techniques were borrowed so that the model can be trained on multiple datasets as well as be able to continue training if the algorithm stopped inexpertly.





3. Bonus Section!!!!

Another interesting vision task in computer science is image segmentation. We worked in a group of 4: Yintao Tai (y.tai-6@sms.ed.ac.uk) University of Edinburgh, Tuofu Li(andy7093li@gmail.com) University of California, Davis, and Yuxuan Tian(yuxuan.tian@mail.mcgill.ca) Mcgill University, and me (javin.liu@mail.mcgill.ca) on a tumor segmentation task. We are currently in the process of publishing our paper. Our paper title is "Brain Tumor Segmentation with Attention-based U-Nets". Please keep an look out if you are interested.



Abstract of UNET Paper

Brain Tumors are a hazardous type of tumor, and they build pressure inside the skull when they grow, which can potentially cause brain damage or even death. Attention mechanisms are widely adopted in state-of-the-art deep learning architectures for both computer vision and neural translation tasks since they tend to enhance networks' ability to capture spatial and channel-wise relationships. We offer an attention-based image segmentation model that outlines the brain tumors in Magnetic Resonance Imaging (MRI) scans if present. In the paper, we integrated two types attention modules into the commonly used segmentation model, U-Net, on resolving the problem of concatenating unnecessary information into the decoder blocks in which it attempts to locate the tumor boundaries. Our research clearly shows the application of attention mechanism in U-Net, incorporates the attention modules, improved the performance in Brain Tumor segmentation task. The model is delivered on an app with additional text to speech and chatbot features provided. In the end, we were able to achieve an IoU of 85.42%.