Brain Tumor Classification

Ruar Podushkin

***Abstract* — This paper explores the application of transfer learning to the problem of classifying mri images of brain tumors. A pretrained version of resnet18 was used, and the retrained model classified images from the testing set with an accuracy of over 99%.**

I. Introduction

According to the National Brain Tumor Society, 700,000 Americans are currently living with a brain tumor, and around 88,970 more will be diagnosed in 2022 [1]. Though most brain tumors are benign, malignant tumors have a low survival rate of just 36%. It is important that robust methods of identifying brain tumors are developed, so that prophylactic measures can be taken against the related diseases.

This report will focus on the process of developing a model to identify and classify brain tumors based off of MRI scans using transfer learning in the Pytorch library [2]. Given an MRI scan, the model should be able to tell with a high degree of certainty whether the image contains no tumor, a pituitary tumor, a meningioma, or a glioma. If a high rate of correct classification is achieved, such a model could be utilized in the medical field to quickly identify MRI scans of interest, which would then be verified by medical professionals.

The following content will be divided into three parts. The first will discuss works relevant to the development of the model. It will also give an explanation about the dataset and preprocessing steps taken. The second section will provide details about the model and the various experiments conducted to improve model performance. The final results of the model will be shown as well. The last section will discuss key takeaways from the process.

II. Dataset and Background

A. Related Works

Transfer learning is a technique in machine learning used for applying stored knowledge on a particular task to a different, but related problem [3]. Resnet18 is a neural network architecture that has performed exceptionally well at image classification tasks [4]. In this report, transfer learning is used on a version of resnet18 that was fully trained on the task of general image classification. The structure and weights of the model are kept generally the same, but the last layer's weights are modified through training on the mri images. Since classifying mri images is similar to what resnet18 was trained to do, the process produces accurate results in a much shorter time frame compared to training the entire model from random weights.

B. Dataset

The dataset used in this report is called the Brain Tumor MRI Dataset and was found on Kaggle [5]. The dataset is composed of 7022 images separated into a train and a test folder each with 4 classes: no tumor, glioma, meningioma, and pituitary. Training and test counts for each class are compared in figure 1. It can be seen that there is a slight data imbalance, with slightly more images being found in the dataset under the no tumor class. The test set for each class was in the range of 17-20%. Since the classes were generally balanced and contained an appropriate train/test split, no steps were taken to prune the number of images in a class.

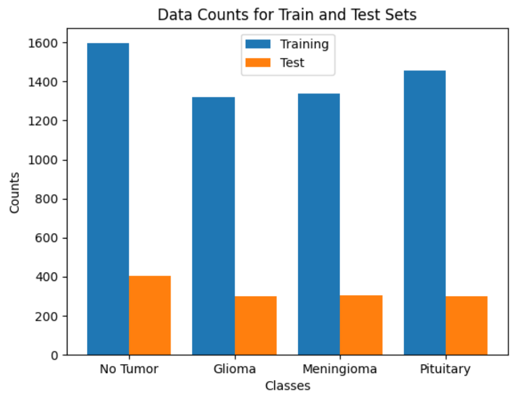

Fig. 1: Chart for the data counts

One possible issue of note is that almost every single image in the no tumor class is taken at a particular top-down angle, whereas other classes have much more variation in the angle. The difference is shown in figure 2, where it can be seen that images from the no tumor class are taken in what appears to be exclusively a top-down angle and at a particular depth, whereas the meningioma images have a much wider variety of angles and depth, with multiple images being taken in profile. The concern was that the model might learn to classify images into the no tumor class based on undesirable characteristics, such as the angle at which the mri image was taken.

The dataset was processed using a method recommended in the Kaggle dataset description. This was done because there was a discrepancy between the size of some images, and also because many images had a significant amount of unused space. Extreme points of each image were found after applying various erosions and blurs. Relevant content was cropped, and all images were changed to a uniform size. Finally, a random horizontal and vertical flip was applied to each image. A small number of images were cropped incorrectly but the processing step worked well overall.

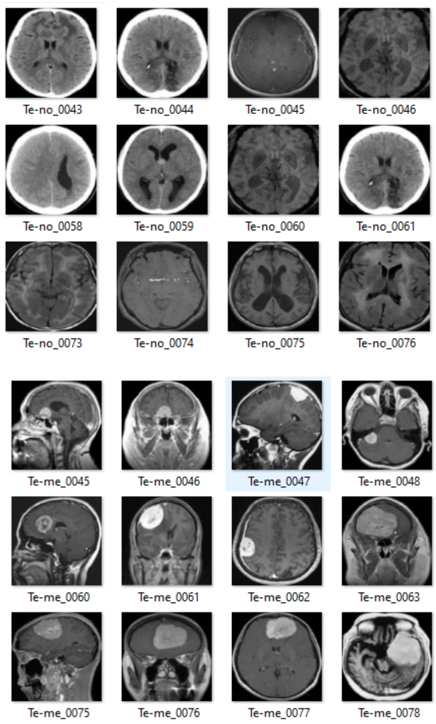

Fig. 2: Comparison of no tumor and meningioma classes

```
No Tumor:  79.75 / 20.25
Glioma:  81.49 / 18.51
Meningioma:  81.40 / 18.60
Pituitary:  82.93 / 17.07
```

Fig. 3: Train / test splits for each class

III. Methods and Experiments

Multiple experiments were conducted to determine the best model parameters. The first experiment managed to achieve a very high accuracy, so only a minor improvement was seen with hyperparameter alterations.

A. First Experiment

All resnet18 model weights were frozen up to the last layer. Stochastic gradient descent was used as the initial optimizer. The learning rate was set to 0.001 and the momentum was set to 0.9. Standard cross entropy was used as the loss function. The model was trained on the dataset for 24 epochs over the course of 10 minutes. The training accuracy and loss were 0.9984 and 0.0055 respectively. The testing accuracy and loss were 0.9925 and 0.0378. The model achieved smooth improvements in accuracy up to epoch 5, where both metrics began to oscillate slightly. The graphs for each metric can be seen in figure 4.

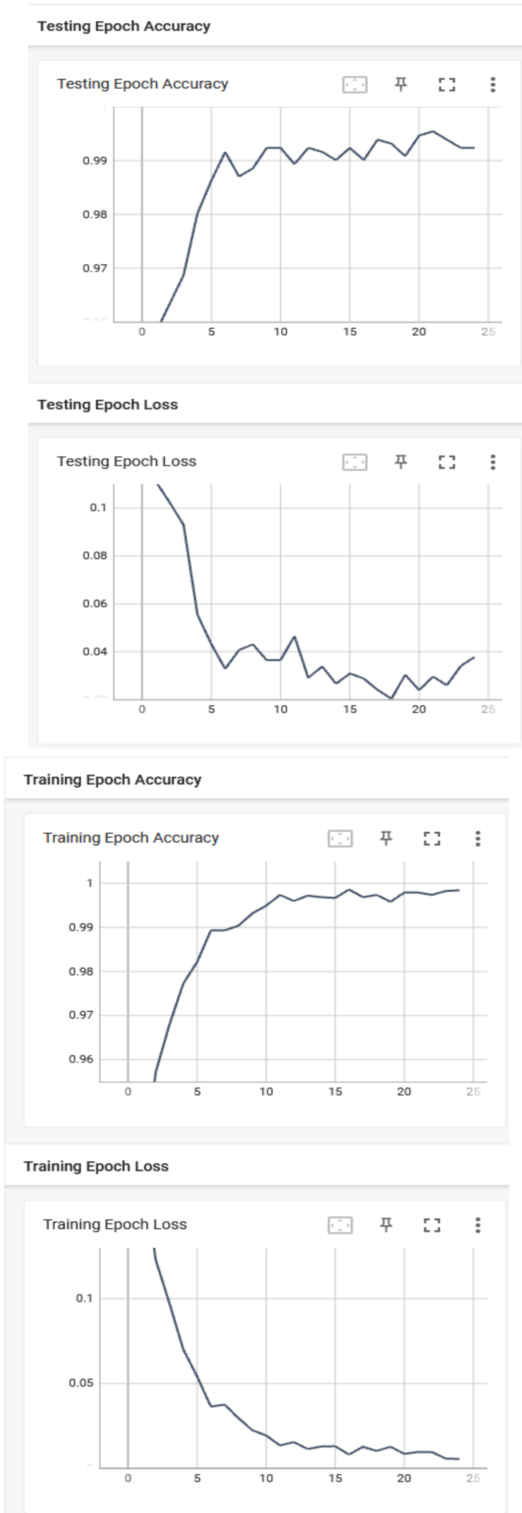

Fig. 4: Results of experiment 1

B. Other Experiments

Multiple attempts were made to increase the accuracy of the model. NAdam and stochastic gradient descent optimizers were tested with changes made to the learning rate and momentum, a loss function called label smoothing cross entropy was tried, and a learning scheduler was implemented to decrease learning rate at epoch intervals. A model with slight improvement was found while using the NAdam optimizer with a learning rate of 0.001 with a standard cross entropy loss function. The model achieved a testing accuracy of 0.9931 and a testing loss of 0.03446. Figure 5 shows the smoothed graphs for each metric compared to each other.

IV. Conclusion

I started this project with expectations of getting decent results, but the effectiveness of resnet greatly exceeded my expectations. There is a possibility that even better results are possible by cleaning up the data which were cropped incorrectly, and I plan to attempt another round of training after removing the few erroneous elements.

Many practical lessons were learned throughout the process, and this project made me really curious about other possible applications of deep learning and transfer learning. I am strongly considering solving a different problem with machine learning for my senior project.

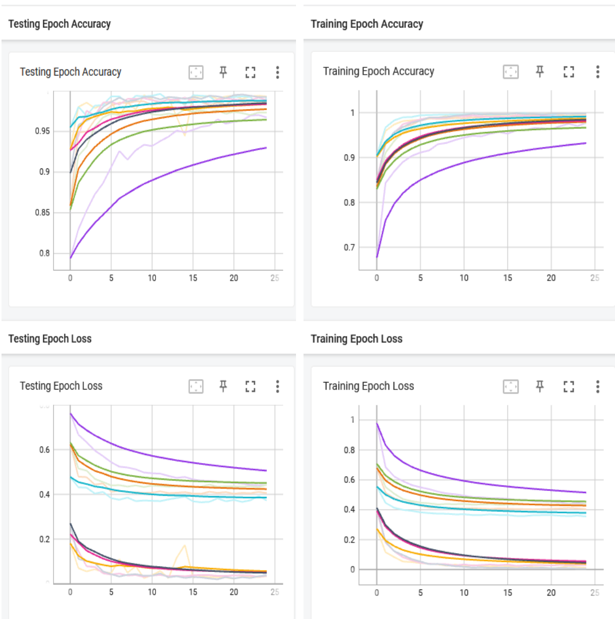

Fig. 5: Comparison of experimental results

References

[1] https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/

[2] https://pytorch.org/

[3] https://arxiv.org/pdf/1911.02685.pdf

[4] https://arxiv.org/pdf/1512.03385.pdf

[5] https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset