Brain Tumor Classification with Semi-Supervised Learning & Contrastive Pretraining

Shanmukha Bodapati, Arya Sawant, Sarthak Gaur, Mehak
The University of Texas at Dallas
{shanmukha.bodapati, arya.sawant, sarthak.gaur, mehak.lnu}@utdallas.edu

***Abstract*—Semi-Supervised Learning has been increasing in popularity for image classification due to not needing as many labels. In this paper, we perform Image Classification on MRI scans with brain tumors using semi-supervised learning. We also pretrain the model using contrastive learning to improve accuracy on new images. We present data preprocessing, model architecture, and results.**

I. Introduction

SEMI-supervised learning (SSL) is extremely important in the field of medical image classification, and is being utilized more often by doctors to develop an initial understanding of a prognosis or confirming a diagnosis. With SSL, the task becomes much simpler due to not requiring labeled data, which is generally difficult to acquire for medical image scans without the guidance of a medical expert.

For our project, we apply SSL with contrastive pretraining for Brain Tumor Classification. We tested multiple algorithms but chose the most feasible with the resources at our disposal.

We first utilized the Fixmatch [1] algorithm developed by researchers at Googlefix which provides a SSL approach to generate pseudo-labels on weakly-augmented unlabeled images. [1] It trains the model on strongly-augmented images (same original images as the weakly-augmented which had a high-confidence prediction). [1] Our process for testing this model included data preprocessing, modifying files to use archived libraries as well as adding our custom image dataset, and calling the shell commands which would create and run the model. However, after much work and testing, we realized that the average runtime for a Fixmatch algorithm is several days and requires a high amount of memory even when using multiple GPUs. Unfortunately, we do not have the resources to run the model to its completion. Therefore, we searched for other methods that could accomplish the task for this project and we found contrastive learning.

Contrastive learning is a self-supervised training method to configure models to recognize unlabeled data by comparing it to similar data points. [2] The Fixmatch algorithm compares strongly-augmented images to weakly-augmented images. It is comparing the structure of the weakly-augmented images to the structure of the respective strongly-augmented images. For our project, this is done by maximizing the similarity of two images and minimizing two loss functions.

II. Model Architecture

A. Dataset and Data Preprocessing

Our dataset consists of MRI scans of the human brain collected from 3 different datasets: figshare, SARTAJ, and Br35H. [3] It contains 7023 images with a training split of 5712 images and testing split of 1311 images. There are four classes: glioma, meningioma, pituitary, and no tumor.

We preprocessed the data by cropping the images to remove noise and resizing the images. MRI scans contain a black background which constitutes noise, so we cropped the images to remove marginal pixels. These pixels are found after converting the image to grayscale, then creating a margin threshold for the image and eroding/dilating it using the threshold. We then calculate the pixel boundaries of the area without the noise and crop the image. Finally, since the data consisted of images of different sizes, we resized all the images to be of size 256 x 256 pixels.

The training data was split into an unlabeled set consisting of 286 images (5% of the training data) and a labeled set consisting of 5426 images (95% of the training data).

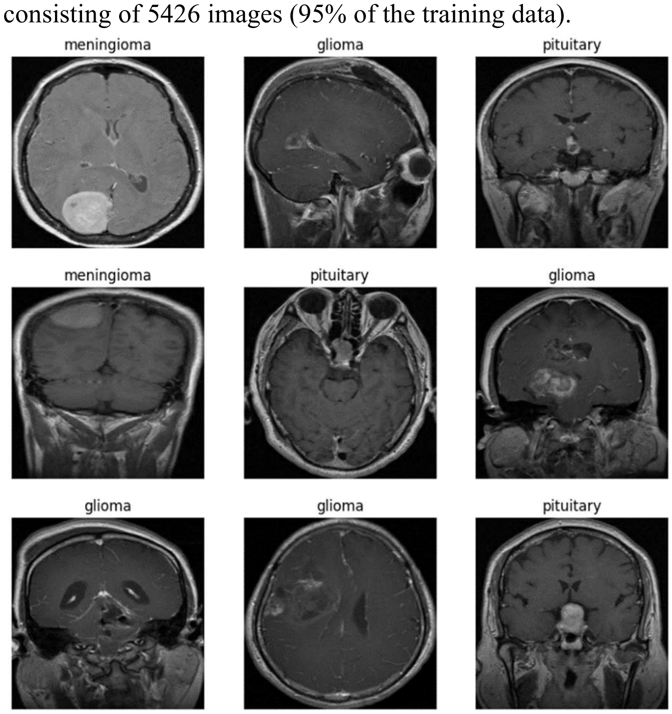

Fig. 1. Some examples of images from the dataset along with their classes.

B. SSL Model with Contrastive Pretraining

The model includes an encoder to encode representations of the images using Conv2D layers as well as an augmenter that performs two augmentations. A random horizontal flip is performed along with a random vertical and horizontal shift. Augmentations are weak-augmentation changing 13% of the image and strong-augmentation changing 50% of the image.

The model also includes a non-linear projection head which improves the quality of the encoder's representations. A linear probing layer is added which is used as a performance metric.

We use two metrics to measure the performance of the model and minimize objective functions for training. The first is contrastive accuracy based on the SimCLR method. SimCLR allows the SSL model to learn important representations and similarities of our data based on the types of data augmentations and the nonlinear transformation of the representation of the image. [4] The SimCLR metric measures the proportion of the representations of an image being more like the augmented version to the representations of images being more like other images. [5] Another metric used was linear probing accuracy which is computed as accuracy of a logistic regression classifier trained on the encoder's features.

Accuracies are calculated using both performance metrics and SparseCategoricalAccuracy. We minimize the InfoNCE (Information Noise-Contrastive Estimation) loss function and the NT-Xent (Normalized Temperature-Scaled Cross Entropy) function. The loss is computed by assuming each image has its own class, finding two augmentations on the image, comparing the image's representation with every possible pair, and calculating categorical cross-entropy as a temperature-scaled cosine similarity.

A test step is added to the model which predicts the classes on given testing data and is performed only after training each epoch. The model is also finetuned using the Adam optimizer to improve the performance of the initial model. The initial and finetuned models are trained for 20 epochs.

C. Results

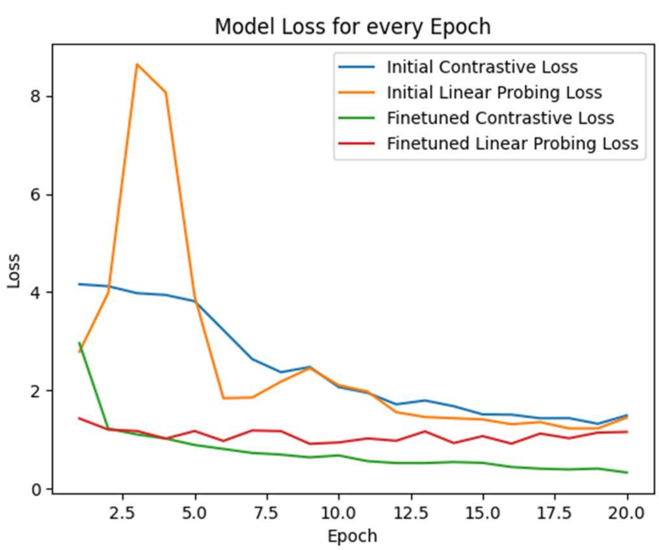

Fig. 2. Contrastive and linear probing losses for the initial and finetuned model across 20 epochs.

We see that the loss improves significantly after training for 20 epochs and as expected, the finetuned model losses are much lower than the initial model losses.

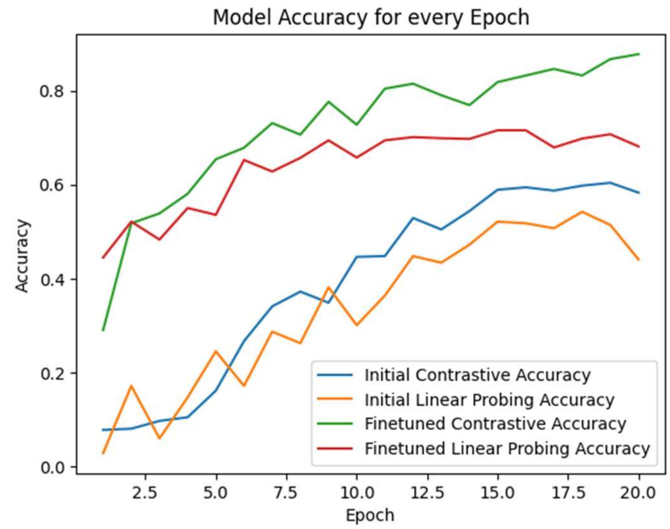

Fig. 3. Contrastive and linear probing accuracies for the initial and finetuned model across 20 epochs.

We see that the training accuracy also improves significantly after training for 20 epochs. There is a large performance improvement for the finetuned model from the initial model. The peak training accuracy reaches **87.76%** for the finetuned model but only **60.37%** for our initial model. We saw a peak testing accuracy of **54.61%** for the initial model and a peak testing accuracy of **71.55%** on our finetuned model, indicating a big performance jump after finetuning.

We are satisfied with our results and the performance metric improvement from the initial model to the finetuned model.

III. Contributions

Everyone in our group contributed equally to the project including research, code, report, and presentation.

References

[1] Kihyuk Sohn and David Berthelot and Chun-Liang Li and Zizhao Zhang and Nicholas Carlini and Ekin D. Cubuk and Alex Kurakin and Han Zhang and Colin Raffel. FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence. *arXiv preprint arXiv:2001.07685*, 2020.

[2] Tiu, E. (2021, January 8). *Understanding contrastive learning*. Medium. Retrieved May 4, 2023, from https://towardsdatascience.com/understanding-contrastive-learning-d5b19fd96607

[3] Msoud Nickparvar. (2021). *Brain Tumor MRI Dataset* [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/2645886

[4] Ting Chen and Simon Kornblith and Mohammad Norouzi and Geoffrey Hinton. A Simple Framework for Contrastive Learning of Visual Representations. *arXiv preprint arXiv:2002.05709*, 2020

[5] Béres, A. (2021, April 24). *Keras documentation: Semi-supervised image classification using Contrastive Pretraining with simclr*. Keras. Retrieved May 4, 2023, from https://keras.io/examples/vision/semisupervised_simclr/