Classifying Brain Images on MRI Sequencing

Ge Cheng (gecheng@stanford.edu), Michael Fan (mifan@stanford.edu) and Menglin Brown (mnbrown@stanford.edu)

Stanford University CS229

Abstract

MRI imaging has proven useful in studying various forms of brain tumors and cancers. MRI sequence is often mislabeled or not labeled and the datasets generally lack the metadata necessary in order to derive the MRI Sequence. In this paper, we explore supervised machine learning techniques involving Convolutional Neural Networks (CNNs) to predict five MRI sequence (T1 Pre/ T1 Post/T2/Flair/Others; Figure 1) used to capture an input MRI image based on a sample of labeled MRI images.

Dataset and Features

- We processed five datasets found on TCIA (The Cancer Imaging Archive). These datasets pooled to around 1 million images total stored in the .dcm format.
- We obtain the correct label for each image from scraping a 'SeriesDescription' field from the .dcm metadata and applying some custom business logic afterwards to categorize it as one of the five possible labels: T1, T1C, T2, Flair, Other (Figure 1).
- We also obtain a linear feature known as the "z-index" that we scraped from an 'ImagePosition' field from the image metadata that we use as a secondary input to multi-input models
- For the CNN methods, we resize all images to 128x128 with scikit-image's resize function to be used as inputs to neural networks. The resize function also helped to normalize the pixel values to a float between 0 and 1 as different labs used different data types.

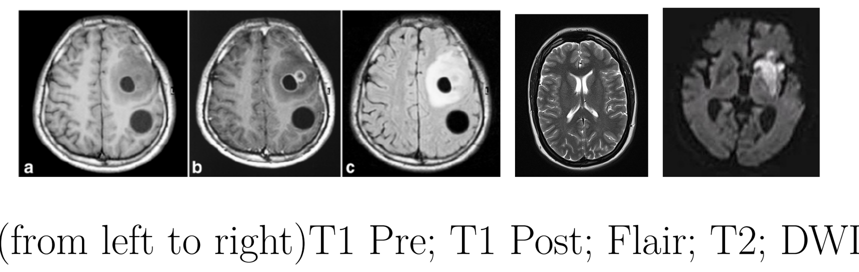

Figure 1:(from left to right)T1 Pre; T1 Post; Flair; T2; DWI (other)

Models – Classical Methods

- Gaussian Naive Bayes: Assuming the probability distribution of a feature given class C_k follows a Gaussian distribution, the classifier is formed by pairing this probability distribution with MAP decision rule.

$$p(X_i = x|Y == y_k) = \frac{1}{\sqrt{2\pi\sigma_{ik}^2}} e^{-\frac{1}{2}(\frac{x-\mu_{ik}}{\sigma_{ik}})^2} \quad (1)$$

- Multilayer Perceptron Neural Network (MLP): MLP trains iteratively by taking partial derivatives of the loss function with respect to parameters in each layer.

Models – CNN Methods

1. VGG16: showed that increasing network depth can benefit accuracy.
2. ResNet: introduced "shortcut" connections that can skip one or more layers
3. InceptionNetV2: shown to achieve high performance at relatively low computational cost.
4. Simple CNN: seven-layer network with 2 sets of alternating convolutional/pooling layers followed by a flatten layer and then 2 fully connected layers. Used ReLU activation function for the convolutional layers.

Experiments

Phase I. trained the models described above.
Phase II. We conduct the following experiments to boost the winning models in Phase I:

1. Multi-input Model: Added an additional input: a "z-index" which measures how far the image slice is from the "top" of the brain.
2. Multi-binary classification: Trained 5 binary classification models with the Simple CNN architecture. Each model classifies an input image with label 0/1 denoting if it is of class $c \in T1, T1C, T2, Flair, Other$.
3. Ensemble: Combine four winning models (ResNet152V2/InceptionResNetV2/VGG16/Simple CNN) and apply means rule to make soft voting predictions.

Results – Phase I

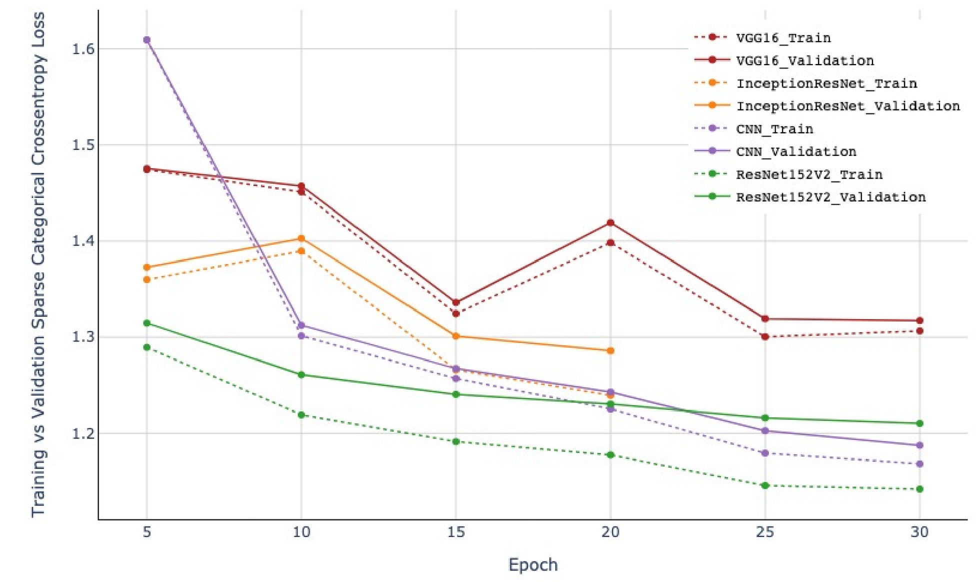

Figure 2:Phase I Loss vs Epochs

Model	Best Hyper-params	Val Acc
ResNet152V2	epoch30-bs32-layer0	0.6896
InceptionNetV2	epoch20-bs32-layer0	0.6121
VGG16	epoch20-bs32-layer3	0.6322
EfficientNetB7	epoch30-bs32-layer0	0.2467
Simmple NN	hidden-layer-size=(50, 20)	0.3960
Simple CNN	L1Filter32-L2Filter64-pool3-kernel4	0.7148

Table 1:Phase I Model Summary

* bs32 = batch size 32, L1 = 1st convolutional layer, L2 = 2nd convolutional layer

Results – Phase II

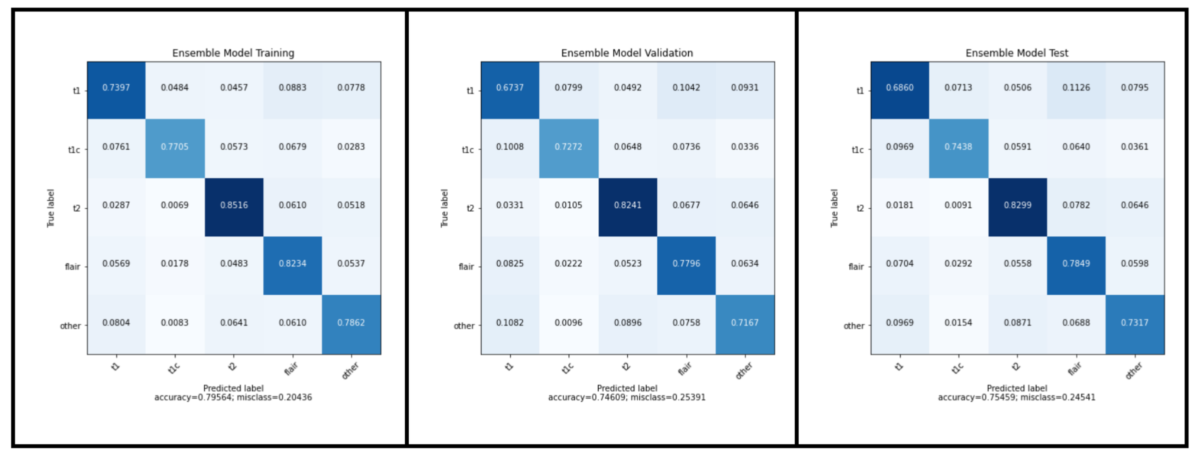

Figure 3:Ensemble Model Confusion Matrices

Discussion

- The unfrozen models needed many training epochs to match the accuracy of the frozen models that had not been trained on any of our MRI data
- Simple CNN yields better results than pre-trained models. Transfer learning from models pre-trained for ILSVRC is only moderately effective
- The earlier layers of the network are indeed more "general" and can be effectively transferred to more tasks than the later, more "specific" layers of the network.
- In **Multi-input model**, adding brain image position (z-index) does not improve the best ResNet152 model and make CNN model much worse. The Simple CNN needs more layers after the z-index concatenation learn.
- In **Ensemble**, we were able to produce a model with validation and test accuracies well over 70%. The ensemble model leverages the strength of different models and cover the their own weakness.

Future

1. Test our model on the private MRI datasets hosted by the Gevaert Lab. It will be interesting to see if our models generalize across foreign datasets having been trained on five public datasets. This will also give us an idea of the variability in the MRI image domain.
2. We also intend to train a model using the ResNet152V2 architecture from scratch given enough image data and compute resources.

References

[1] Mri sequence, Nov 2020.

Acknowledgements

The authors would like to thank Ph.D candidate Bryce Allen Bagley, Ph.D Pritam Mukherjee from the Gevaert Lab and our TA Paul Caron for their guidance. We also want to thank all CS229 course staff for instruction during a difficult quarter..