Brain Tumor Detection

Machine Learning Classification of MRI Scans

Nicholas Beninato - Department of Computer Sciences - Undergraduate
Sebastian Murrell - Department of Computer Sciences - Undergraduate
Garrison Waugh - Department of Computer Sciences - Undergraduate

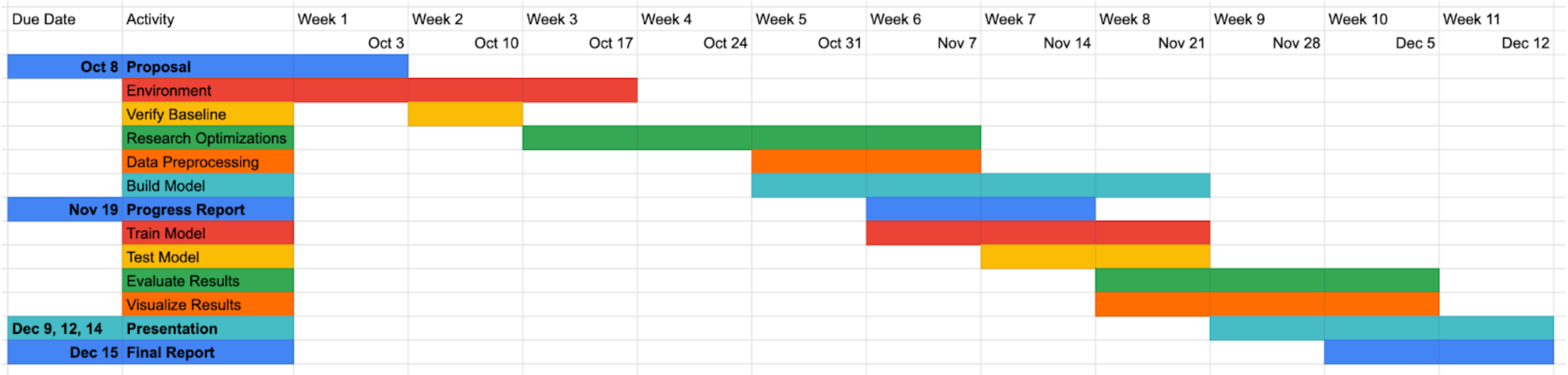
Problem Statement and Tasks

The goal of this project is to create a CNN for brain tumor identification.

- Performing multi-class classification on human brain MRI images
- Image labels are: glioma, meningioma, pituitary, and no tumor (control).
- Initial model (progress report) was binary classification of tumor / no tumor
- Final model accurately classifies images into the four labels

Problem Statement and Tasks

Here is our Gantt chart for our timeline of completing subtasks for the projects



Dataset

Data comes from three datasets:

- Brain Tumor Dataset - Jun Cheng, Southern Medical University, Guangzhou, China
- Brain Tumor Classification - @sartajbhuvaji on Kaggle
- Brain Tumor Detection - @ahmedhamada0 on Kaggle

Initial images in .mat, .jpeg, and .png file formats. Preprocessing consisted of stacking and reshaping images. Resulting dataset contains 7500+ images (1.5 GB).

Dataset Preprocessing

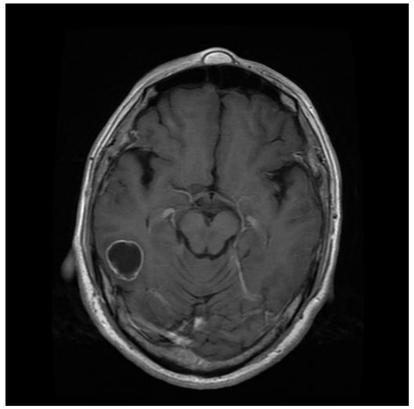
As our data came from various sources, we had to standardize it

- Converting all .mat files into JPEG images
- Stacking all single channel greyscale images to have the same dimaneions as color
- Using OpenCV to resize all images to 150 x 150 x 3
- Flattening images and saving to CSV with the first column as the label

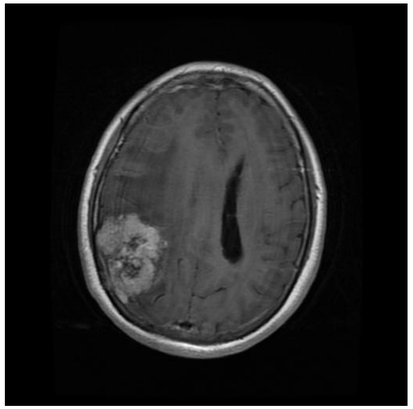
Dataset cont.

Label	# Samples	Percentage
No Tumor	2000	25.5%
Meningioma	1645	21.0%
Pituitary	1831	23.4%
Glioma	2352	30.1%

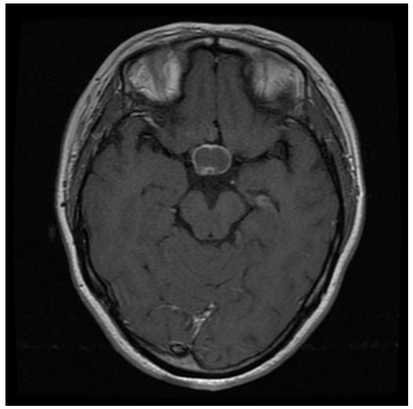
Dataset cont.



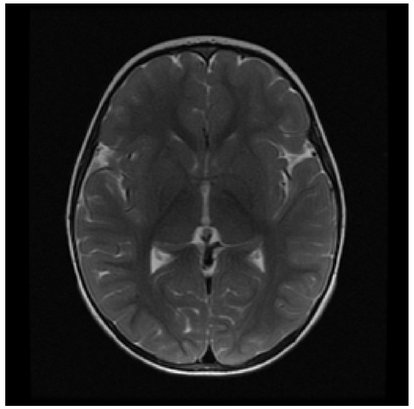
Glioma Tumor Image



Meningioma Tumor Image



Pituitary Tumor Image



Healthy Brain Image

Related Work

We took inspiration from an example on Kaggle which utilized transfer learning

- Transfer learning allows us to use the progress that a model has already made on Computer Vision benchmarks, and fine tune that for our own model.
- There was also a recent study that shows that deep learning models, like the one we worked on, could perform better at identifying brain tumors than neuroradiologists.

Method

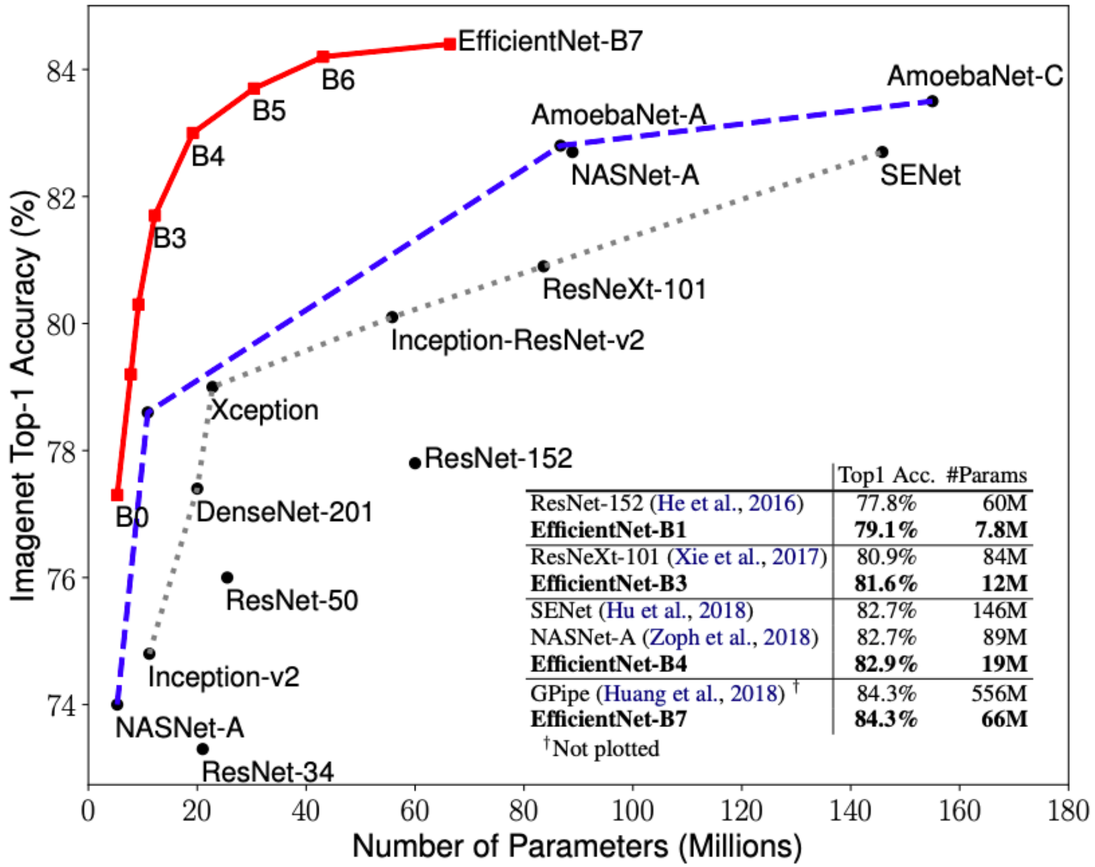
Transfer Learning - EfficientNetB0

```
tf.keras.applications.EfficientNetB0(
    include_top=True,
    weights="imagenet",
    input_tensor=None,
    input_shape=None,
    pooling=None,
    classes=1000,
    classifier_activation="softmax",
    **kwargs
)
```

Pre-trained on ImageNet

#Params: 5.3 Million

Comparison to similar models



	Top1 Acc.	#Params
ResNet-152 (He et al., 2016)	77.8%	60M
EfficientNet-B1	**79.1%**	**7.8M**
ResNeXt-101 (Xie et al., 2017)	80.9%	84M
EfficientNet-B3	**81.6%**	**12M**
SENet (Hu et al., 2018)	82.7%	146M
NASNet-A (Zoph et al., 2018)	82.7%	89M
EfficientNet-B4	**82.9%**	**19M**
GPipe (Huang et al., 2018) †	84.3%	556M
EfficientNet-B7	**84.3%**	**66M**

†Not plotted

Method cont.

Convolutional Neural Network

Transfer learning base

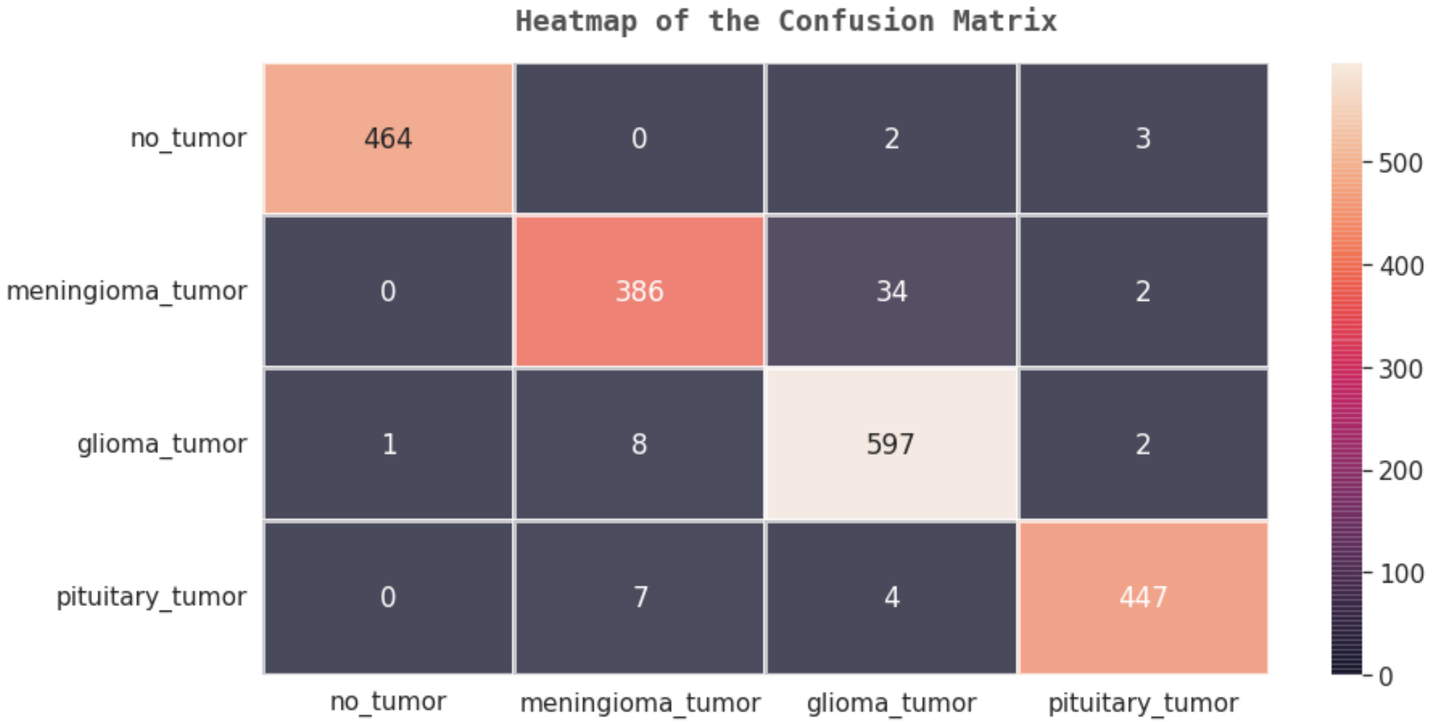
```
model = effnet.output
model = keras.layers.GlobalAveragePooling2D()(model)
model = keras.layers.Dropout(rate=0.5)(model)
model = keras.layers.Dense(4, activation='softmax')(model)
model = keras.models.Model(inputs=effnet.input, outputs=model)
model.compile(loss='categorical_crossentropy', optimizer='Adam',
metrics=['accuracy'])
```

Additional Layers:

Avg Pooling, 50% Dropout, Softmax

#Params: ~4 million

Results


Heatmap of the Confusion Matrix
no_tumor
meningioma_tumor
glioma_tumor
pituitary_tumor
464
0
2
3
0
386
34
2
1
8
597
2
0
7
4
447
no_tumor
meningioma_tumor
glioma_tumor
pituitary_tumor
500
400
300
200
100
0

Discussion

Baseline

	precision	recall	f1-score	support
0	0.98	0.96	0.97	93
1	0.96	1.00	0.98	51
2	0.98	0.98	0.98	96
3	1.00	1.00	1.00	87
accuracy			0.98	327
macro avg	0.98	0.98	0.98	327
weighted avg	0.98	0.98	0.98	327

Testing Accuracy

	precision	recall	f1-score	support
0	1.00	0.99	0.99	469
1	0.96	0.91	0.94	422
2	0.94	0.98	0.96	608
3	0.98	0.98	0.98	458
accuracy			0.97	1957
macro avg	0.97	0.97	0.97	1957
weighted avg	0.97	0.97	0.97	1957

Lower accuracy likely results from our larger dataset.

Thank you