**Image Classification Brain Tumor**

# **Bhushan Suryawanshi** [**DSC680-T302 Applied Data Science (2215-1)**](https://cyberactive.bellevue.edu/webapps/blackboard/execute/courseMain?course_id=_500232_1)

# Bellevue University

**Abstract:**

Images are used in various fields to make the problem easier to understand. Image processing techniques are most widely used in medical imaging to identify the affected area through an X-ray, computed tomography scan(CT scan), MRI scan(Magnetic resonance images). These images are used to detect, identify, and locate infections, abnormal growths from the human body. Heart diseases, Cancer, Brain tumor, Blood clotting, these are some of the abnormalities that can be found by medical imaging techniques. We can use different machine learning techniques to classify different types of brain tumors by using MRI. The Convolutional Neural Network (CNN) is a class of deep learning neural networks that are highly effective with image classifications.

**Introduction:**

Every year, the number of patients with brain tumors is increasing. There are two classes of brain tumors, primary and secondary tumors. Primary tumors have several types; one of the frequently found is a meningioma type. It is very challenging to locate, detect, and select the infected tumor portion in the brain from the MRI (Magnetic resonance images). This tedious and time-consuming job requires radiologists and medical field experts. The accuracy of this task is mainly subject to the experience and expertise of the person performing this task. So, if we use a machine learning model to perform this task, it will help to overcome the shortcomings of the person involved in performing this task. So, I think if we can automate this process of classifying the tumors by using machine learning algorithms, it will improve the accuracy of the results and cost due to the expertise required.

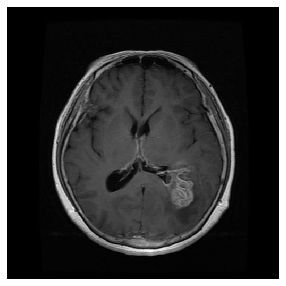
**Data Details:**

The brain tumor dataset contains 2870 training images. The dataset includes the three kinds of tumor images. The three tumors have the following distribution of data. -

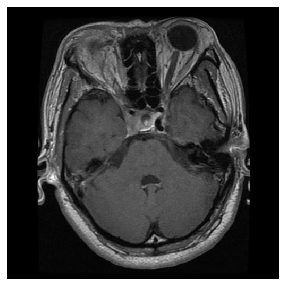
1. 822 Meningioma images. The meningioma type of tumors seen near the top-outer part of the brain. Meningioma is slowly growing noncancerous tumors that cause seizures and visual problems. This type of tumors accounting for 37.6% of all tumors, and 53.3% of all non-malignant tumors.



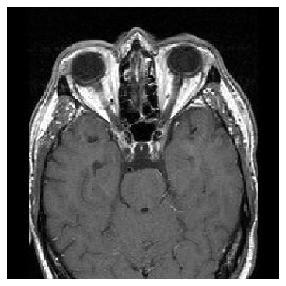
1. 826 Glioma images. Glioma is an abnormal growth in glial cells present around the neurons in the brain. Gliomas make up 81% of malignant brain tumors in adults.



1. 827 pituitary tumor images. Pituitary tumors grow in pituitary glands that affect body functions. Some pituitary tumors result in too many of the hormones that regulate important functions of your body. Some pituitary tumors can cause your pituitary gland to produce lower levels of hormones.



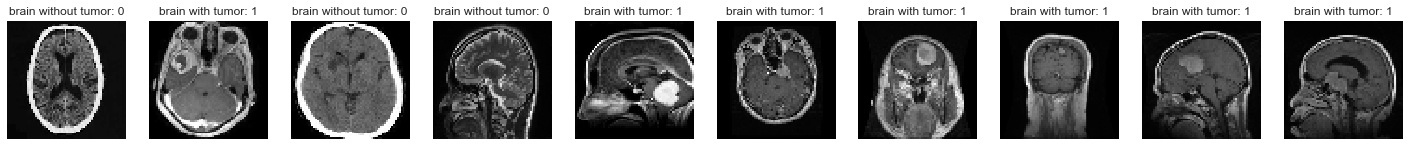
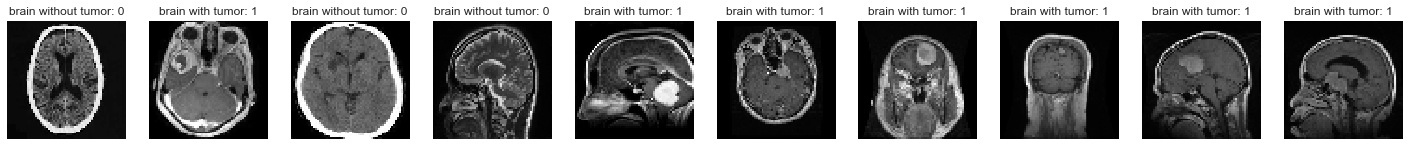
1. 395 No tumor images.



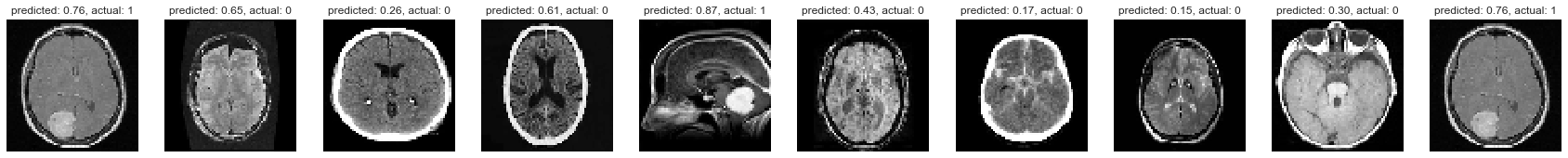
*Dataset Source - https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri*

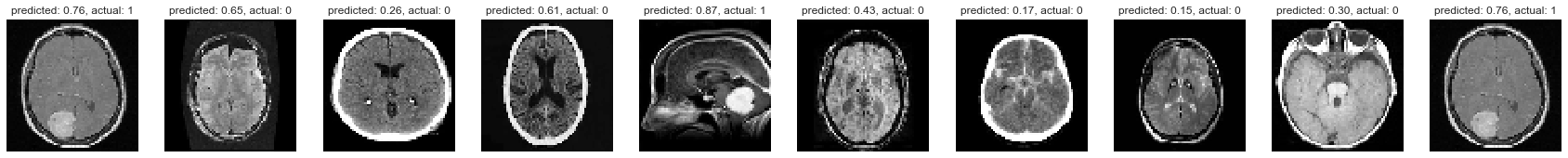
**Approach:**

I am using grayscale images hence, the size of each image is smaller compared to using RGB. Following are some sample images of the brain tumor.

To start with, I used binary classification to predict tumors. The data used for this model is from the no\_tumor and meningioma\_tumor class. The data is first converted to grayscale and then resized to 64 x 64 pixel size. The model uses 4 Conv2D and 4 max\_pooling layers followed by flattening and dropout layers. The accuracy of the binary model was above 80%. This model is further tuned with hyper-parameters like batch size and epochs.

The best model shows a validation accuracy of more than 82%. The use of binary classification here is to verify the data is good for model building. Following are some of the results from the test data.





As seen in the dataset we have 4 different classes. Hence using multiclass classification model to find the accuracy with different types of tumors. For the multiclass classification, the image resolution of 128 x 128 pixels is used. To make the most of our few training examples, we will "augment" them via a number of random transformations, so that our model would never see twice the exact same picture. This helps prevent overfitting and it generalizes the model better. Keras helps with ImageDataGenerator which can be used for –   
1. Random transformation and normalization.  
2. Instantiate generators of augmented image batches

ImageDataGenerator(

featurewise\_center=False,

samplewise\_center=False,

featurewise\_std\_normalization=False,

samplewise\_std\_normalization=False,

zca\_whitening=False,

rotation\_range=0,

zoom\_range = 0,

width\_shift\_range=0,

height\_shift\_range=0,

horizontal\_flip=True,

vertical\_flip=False)

In our case, we will use a very small CNN with few layers and few filters per layer, alongside augmentation and dropout. Dropout helps reduce overfitting, by preventing a layer from seeing twice the same pattern, thus acting in a way analogous to data augmentation. This approach gave us a validation accuracy of around 73%. The model uses 20 epochs with a batch size of 20.

**Conclusion:**

For image processing, the Convolution neural networks are the best approach. By using the Convolution neural network, we have created a model to classify brain tumors from MRI images. The validation accuracy of the model is approximately 80 percent for binary classification. In the case of multiclass classification having augmentation 73% accuracy is achieved on test data. The precision of the model can be further improved by using more image processing techniques like image segmentation. The overfitting of the model can be reduced by using more data along with some more image processing techniques like segmentation or hyperparameter tuning.

**References:**

[1] <https://docs.scipy.org/doc/scipy/reference/generated/scipy.ndimage.median_filter.html>

[2] <https://www.pyimagesearch.com/2016/04/11/finding-extreme-points-in-contours-with-opencv/>

[3] <https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/>

[4] [https://www.analyticsvidhya.com/blog/2020/01/fundamentals-deep-learning-activation-functions-when-to-use-them/#:~:text=The%20ReLU%20function%20is%20another,neurons%20at%20the%20same%20time](https://www.analyticsvidhya.com/blog/2020/01/fundamentals-deep-learning-activation-functions-when-to-use-them/" \l ":~:text=The ReLU function is another,neurons at the same time).

[5] Brain Tumor Detection Using K-Means Clustering Algorithm: V.Surudhi, K.Sanjana, R.Saravanan, G.Santhosh, S.Kirubha <https://pdfs.semanticscholar.org/b70e/4a5455a4531ca650272474d49d29e5e1da5c.pdf>

[6] The Complete Beginner’s Guide to Deep Learning: Convolutional Neural Networks and Image Classification: Anne Bonner: <https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb>

[7] QUICK BRAIN TUMOR FACTS: National Brain Tumor Society: <https://braintumor.org/brain-tumor-information/brain-tumor-facts/>

[8] Pei, L., Vidyaratne, L., Rahman, M.M. *et al*  Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. *Sci Rep* 10, 19726 (2020). https://doi.org/10.1038/s41598-020-74419-9

[9] Seetha J, Raja S. S. Brain Tumor Classification Using Convolutional Neural Networks. Biomed Pharmacol J 2018;11(3).

[10] An Hoai Truong, Viktoriia Sharmanska, Clara Limbӓck-Stanic, Matthew Grech-Sollars, Optimization of deep learning methods for visualization of tumor heterogeneity and brain tumor grading through digital pathology, Neuro-Oncology Advances, Volume 2, Issue 1, January-December 2020, vdaa110, https://doi.org/10.1093/noajnl/vdaa110