

# Detecting and Classifying Brain Tumors using a variety of Machine Learning Methods

1<sup>st</sup> James Sanii, 20111876

*Dept. Name of Organization: ECE*

*Name of Organization Queen's University*

Date: 2021/12/11

Course: ELEC872

**Abstract**—It is estimated that 24, 530 adults in the United States will be diagnosed with a brain tumor this year [1]. The analysis of an MRI scan to extract information about a brain tumor is time consuming and requires a high level of technical skill only possessed by neurologists.

The goal of this study was to investigate the viability of a tumor detection and classification system to automate tumor diagnosis. The tumor detection system would first analyze if a tumor were present then distinguish the type of tumor as either glioma, meningioma or pituitary. The automated system would take in MRI image data, perform data augmentation and feature extraction on the images before outputting if a tumor is present and if present, the type of tumor.

The data set the system was trained on contained MRI slices from various angles and contained 3264 images where 500 images were tumor free, and the rest were evenly split between the three tumor types.

The analysis for the system included testing multiple machine learning algorithms for each problem. The models tested were Support Vector Machine, Random Forest, K-Nearest Neighbors and a Convolutional Neural Network was also tested for tumor classification.

The system was able to correctly detect if a tumor was present with 99 percent accuracy with no images with a tumor being incorrectly classified. The model used for tumor detection was K-Nearest Neighbors. The system struggled to identify differences between the different types of tumors with the best model being Random Forest with an accuracy of 68 percent. The system struggled to identify the differences between pituitary and glioma tumors but was able to identify meningioma tumors with a high degree of accuracy.

**Index Terms**—Tumor Detection, Support Vector Machine, Random Forest, K-Nearest Neighbors, Convolutional Neural Network

## I. INTRODUCTION

A brain tumor occurs due to an abnormal growth of cells that have multiplied in an uncontrolled manner. Cancerous brain tumours are an aggressive disease with a low 5-year survival rate of about 36 percent [1]. Early detection of a brain tumor is critical to maximize expected survival rate [1]. Sadly, the analysis on MRI data is very time consuming and requires a large amount of technical expertise. This is not a major issue in developed countries, but third world nations struggle to have the required medical staff to support the required treatment [2]. The goal of this project is to create a machine learning model that can detect tumors at a rate

similar to a neurologist. If this is achieved, another model will be created to attempt to determine the type of tumor that was detected. Similar to a neurologist is defined as 99 percent accuracy for tumor detection and 95 percent accuracy on tumor classification. The model will aim to detect the differences between a glioma, meningioma and pituitary tumors. This will provide a neurologist with the needed information to create a treatment plan without having to analyze the MRI by hand. By reducing neurologist workload, they can be freed up to do other tasks which could include doing remote diagnose for areas of the world that lack neurologists which would help combat medical care inequality between first and third world nations [2].

## II. LITERATURE REVIEW

Brain tumor segmentation has been a common problem in machine learning for the past 10 years. Tumor segmentation steps can be used as preprocessing for tumor detection and classification problems.

J. Machiraju proposed that Contrast Limited Adaptive Histogram Equalization (CLAHE) works as an effective preprocessing step for tumor detection [3]. It is most effective when images in the data set are low contrast or have varying levels of contrast [3]. Furthermore, unlike Adaptive Histogram Equalization, CLAHE prevents the over enhancement of noise and minimizes the edge shadowing effect by limiting the contrast enhancement [3].

Canny Edge Detection is commonly used in tumor segmentation problems. I. Maiti and M. Chakraborty highlighted how Canny can be used in combination with other preprocessing steps like Contrast Enhancement to segment out a brain tumor [4]. By showing that Canny can be used to segment out a brain tumor it shows that it can be used as a viable preprocessing step for feature extraction for tumor classification and detection problems.

The system proposed by Parthasarathy G, Ramanathan L, Anitha K and Justindhas Y uses Canny Edge Detection algorithm to locate the tumor as well as its source, spread area and age [5]. The model outperformed the existing algorithm and was able to detect and grade tumors effectively. This algorithm further highlights that using Canny Edge Detection for feature extraction as a preprocessing technique is viable

and can greatly help perform a variety of tasks associated with tumor detection and classification.

R. Kumar et al. highlights that using a ResNet 152 as a feature extraction algorithm then running the data through a Random Forest results in 97 percent classification accuracy for determining if a patient is healthy, has Pneumonia or COVID-19 [6]. This shows that with proper feature extraction Random Forest can excel at image analysis problems.

C. Kamarajan et al. uses MRI data and a variety of demographic variables to classify if users have Alcohol Use using a Random Forest (RF) [7]. By accurately classifying users using MRI brain data it shows that RF can be used to classify differences in the brain [7].

E. I. Zacharaki et al. used MRIs as input to develop an Support Vector Machine (SVM) to classify tumor type and tumor grade [8]. It used a SVM with recursive feature elimination in order to extract the best SVM using the minimum number of inputs by removing all features that do not directly improve performance [8]. It also highlights the benefits of using multiple SVMs in a voting scheme in order to improve performance [8].

R. J. Ramteke's and K. Y. Monali's paper showed that K-Nearest Neighbors (KNN) networks can be used successfully to detect abnormalities in medical images [9]. Brain Tumor Detection Using KNN by P. Aiwale and S. Ansari shows that a variety of features can easily be extracted from MRI images than can be processed by a KNN and Lloyd to calculate the area occupied by a brain tumor but for the system to be effective morphological operations to reduce noise must be implemented [10].

S. Das et al paper was able to show that a simple Convolutional Neural Network (CNN) can perform extraordinarily well with the model achieving a 93.33 percent average precision on a similar data set to the one explored in this papers analysis [11]. S. Deepak and P. M. Ameer proposed a paper which highlights the benefits of using transfer learning when determining tumor type. The model uses a pre-trained GoogLeNet to extract features from Brain MRI images before classifying them. The proposed system has a mean classification accuracy of 98 percent which exceeds any previous works on the given data set [12]. These two papers highlight the potential high performance a CNN can offer in a tumor classification problem.

### III. METHOD

The machine learning tumor detection system is separated into 3 different steps, image processing and augmentation, training and optimizing algorithms and testing the algorithms on hidden data.

The first step to processing the images is to load in image as the desired size of 256 by 256 pixels then converting from a color image to a black and white image. The image is converted to black and white since MRI machines create a black and white image by default [13], but the data set had saved the images as a 3 channel RGB images. Therefore, the images can be converted back to black and white to reduce

computational intensity without losing any original source data.

Next a 3x3 median filter was applied to the images in order to reduce the impact of noise over the image. Median filters are able to remove noise while preserving edges which is ideal for this data set. CLAHE was applied to the image to improve image contrast in order to maximize the difference between different parts of the image [14]. CLAHE was used over Adaptive Histogram Equalization (AHE) since AHE can cause the over amplification of noise which can significantly reduce image usability in machine learning applications [14].

Canny Edge Detection was then applied to the images. Edges are important for this task since it highlights where significant changes in brightness occur. If they occur somewhere different than normal or some expected edges are missing these changes can indicate the presence of a tumor.

The final images, that currently contains only edges was then normalized. This was done to make sure each pixel has a value between 0 and 1. The normalization step was used since it generally speeds up convergence for machine learning algorithms.

The training data was then split into training and validation data. The training data was used to train each of the model. The validation data was used to check model performance during the process of optimizing metrics. The testing data was already separated in the database from the training data so additional split was not required.

The models that were chosen were RF, KNN and SVM. A variety of parameters were varied for each model and a summary on how the parameters effected the models can be found in the Results Section.

The resulting models were then used on the validation data to determine how well they performed. Since the data set was unbalanced with 395 images containing no tumor and 2475 that contained a tumor, predicted outputs should be analyzed to make sure the model performs well on each class, not just the majority class. Additionally, the most important aspect of the model is to minimize false negatives since determining a tumor is not present when someone has a tumor would be fatal if the model is used in practice. Therefore, recall will be used as the most important metric for the tumor detection problem. Accuracy will be used as the tie breaking metric when deciding what parameters are optimal for each algorithm.

The second problem of determining tumor type will follow a similar workflow highlighted above for the first aspect of the problem. The difference being that the models will change slightly to handle multiple classes and will have different optimized parameters due to this fact. A CNN will also be tested for the second problem.

The preprocessing for the CNN image data will be the same as above but will skip the Canny Edge Detection since a CNN should be able to detect edges through the convolutional steps. Data augmentation was performed on the images to increase the variety in training data to create a model that generalizes better. The transformations that were applied for data augmentation were rotation of a maximum of 20 degrees,

width and height shift and zooming in on the image. Each of these augmentations were used since they vary the image without making any drastic changes to them. A common data augmentation technique that was not used was vertical and horizontal flipping of the image. This technique was not used since flipping images in a medical application can change the diagnosis. Optimization of the CNN will be explained in the results section.

The data set for the tumor classification is a balanced unlike the data set for tumor detection. The data set contains, 826 images of glioma tumors, 822 of meningioma tumors and 827 of pituitary tumors. Therefore, accuracy will be used to determine the best parameters when optimizing the algorithms listed above. F1-Score will be used as a tie breaker if accuracy is the same for multiple combinations of parameters.

#### IV. RESULTS

The first part of the result section that needs to be covered is the optimization and effect of parameters on models for the tumor detection task.

For RF, the maximum tree depth, split criteria and number of decision trees were varied. Maximum tree depth and number of decision trees varied how well the model is fit to the training data. As the number of trees and/or maximum tree depth increases the model performs better on the training data but did not generalize as well, since the model became overfit. If the number of trees or maximum tree depth were to low the model struggled to perform on the validation data since not enough information was learned to make valuable predictions. The goal is to find the ideal number of trees with ideal depth that it performs well on the training data and performs well on the testing data. The optimal parameters were found to be 20 trees with a max tree depth of 12 which used Gini impurity as the split criteria.

For KNN, number of neighbors, weight function and algorithm to compute nearest neighbor was varied. Number of neighbors controlled the number of clusters chosen for the algorithm. Too many clusters caused the model to not generalize well. Too few clusters resulted in photos being incorrectly identified since they would be added to clusters they did not actually identify with, which meant the model performed worse since they did not learn enough from the given information to distinguish between the two classes. The algorithm and weight function had very little impact on output. The ideal set up was using 3 neighbors with, KD tree as the algorithm and weights being set to distance.

For SVM, polynomial degree and kernel type was varied. Kernel type varied and RBF performed best on task one. For polynomial degree, as the polynomial degree increases it allows for more flexibility in the decision boundary but too flexible can cause the model to be overfit. The challenge is finding a polynomial degree large enough to fit the data but not so large it overfits the model. The final polynomial value for the SVM was 3. The results for the optimized algorithms for the tumor detection problem are seen in Table 1.

TABLE I  
TUMOR DETECTION RESULTS

Machine Learning Algorithms	Metrics			
	Accuracy	Precision	Recall	F1-Score
Random Forest	0.9213	0.9243	0.9723	0.9477
K-Nearest Neighbors	0.9746	0.9666	1.000	0.9830
Support Vector Machine	0.9442	0.9377	0.9896	0.9630

The KNN model performed the best due to having the highest scores in every category. The recall score of 1.000 means that the model had zero false negatives which signifies that all possible tumors were detected by the model. A confusion matrix for the KNN model is shown below in Fig. 1, where 0 represents no tumor detected and 1 represents a tumor was detected.

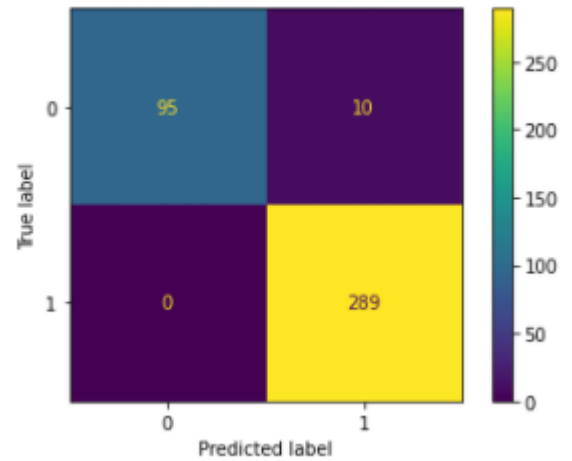


Fig. 1. Confusion Matrix of KNN model predictions

The confusion matrix shows that the model excels at classify when a tumor exists and occasional incorrectly classifies when a tumor is not present. Overall, the model performs very well on both classes and can be used to accurately identify when tumors are present.

For the second task of tumor type prediction the same optimization approach was used for the RF, SVM and KNN. The optimal parameters for the SVM was kernel RBF with a degree 3 polynomial. RF optimal parameters were 2000 trees with a maximum depth of 16, using Gini impurity split criteria. Optimal KNN parameters were 2 neighbors, algorithm set to auto using uniform weights.

Looking at the optimal performance criteria for each model the result for KNN optimal parameters highlighted an issue with the models since there are 3 classes so there should be at least 3 clusters. KNN optimizing to only predict two classes which signified that it is not a viable model for this data set.

For optimization of the CNN the number of layers, size of convolutions and dropout rate was varied. Due to time constraints the CNN was not optimized using a loop to vary parameters. The number of convolution and max pooling steps was increased from 3 to 4 during optimization since the model

was shown to be underfit. I could tell the model was underfit since the validation data was outperforming the training data. Another Dense layer was added after which helped to decrease the time for model convergence. Dropout Layers were then added since the model now showed signs of being overfit due to performing poorly on validation data. Two additional dropout layers were then added, one before and after the first dense layer. The first dropout layer had a magnitude of 0.1 and the second dropout layer had a drop out rate of 0.5. ReLU activation function was used since it is the best activation function to combat the exploding gradient problem. The number of training epoches varied depending on set up, the final model took 60 epoches to train. Model was trained 5 epoches at a time until a convergence was hit. The overall model can be found below in Fig. 2.

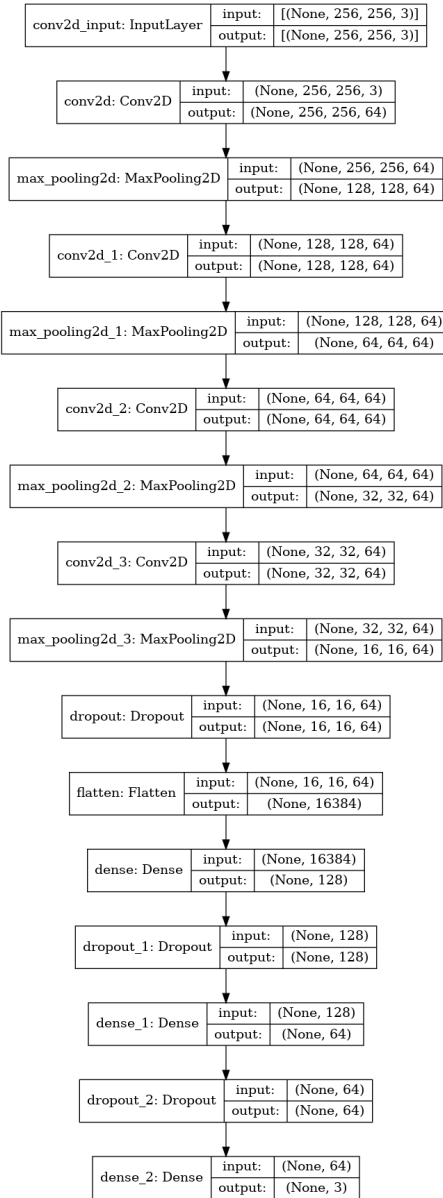


Fig. 2. CNN model used to classify tumors

TABLE II  
TABLE OF TUMOR CLASSIFICATION RESULTS

Machine Learning Algorithms	Metrics			
	Accuracy	Precision	Recall	F1-Score
Random Forest	0.6817	0.7166	0.6877	0.6256
K-Nearest Neighbors	0.5121	0.6843	0.4554	0.4087
Support Vector Machine	0.6298	0.7359	0.6406	0.5812
Convolutional Neural Network	0.6298	0.7877	0.6271	0.5181

The result of each model on the task of tumor classification is shown in Table 2.

RF was the best performing model on the data. This was an unexpected outcome since CNN typically outperform other machine learning methods on image data. This highlights the CNN model could not generalize to the training data very well.

A class-by-class breakdown of the RF is highlighted in the confusion matrix shown in Fig. 3.

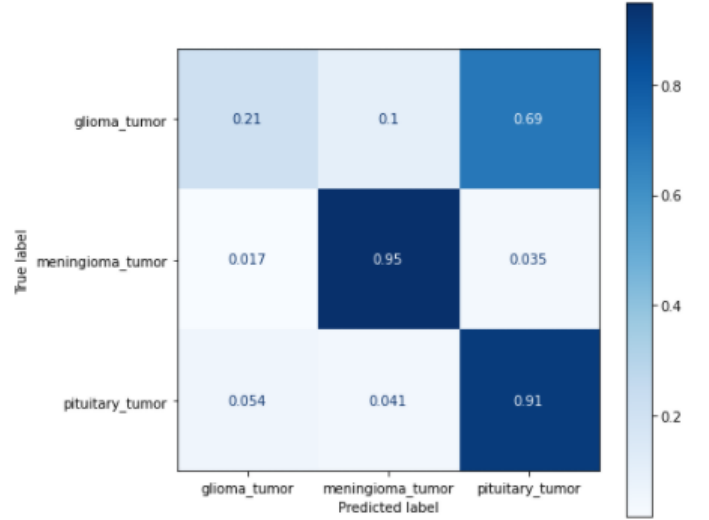


Fig. 3. Random Forest Confusion Matrix

The confusion matrix highlights that the model accurately predicts pituitary and meningioma tumors but struggles to accurately predict glioma tumors. The high error rate on glioma tumors highlights that pituitary tumors and glioma tumors have similar characteristics which cause the RF to confuse glioma tumors as pituitary tumors.

## V. ANALYSIS AND DISCUSSION

Overall, the system excels at identifying when a tumor is present but struggles to distinguish between the type of tumor present. This allows for the system to replace the need for a neurologist to work on tumor detection, but if a tumor was detected a neurologist would still need to be used to establish the type of tumor.

Analyzing the data set, various angles were used for MRI scans. This likely contributed to the struggle of identifying types of tumors since tumor physical features and location are the main two contributors to determining tumor type but by removing the ability to determine tumor type by location the

algorithms could only determine type by using tumor features which contributes to the lower performance of the models compared to using similar models on other MRI data sets.

The next step for this work is to apply the same methodology to a data set that has MRI image data take from a similar angle. This would allow for the models to use location features of the tumor on top of the physical features that are used in the current model. Since MRI scans are volumetric, training a model on only one slice angle is a valid option since every MRI scan should be able to produce a similar slice that can be used as input for model prediction.

If given large scale computational resources, it could be possible to create a tumor classification system using the volumetric scans as input with a 3D CNN model which would likely outperform any model that uses an individual scan as input.

## REFERENCES

- [1] "Brain Tumor - statistics," Cancer.Net, 10-Nov-2021. [Online]. Available: <https://www.cancer.net/cancer-types/brain-tumor/statistics>. [Accessed: 11-Dec-2021].
- [2] Sartaj, "Brain tumor classification (MRI)," Kaggle, 24-May-2020. [Online]. Available: <https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri>. [Accessed: 09-Dec-2021].
- [3] J. Machiraju, "Pathological Brain Tumor Detection Using CLAHE and LS-SVM," TEST Engineering and Management, vol. 82, pp. 11323-11332, 2020.
- [4] I. Maiti and M. Chakraborty, "A new method for brain tumor segmentation based on watershed and edge detection algorithms in HSV colour model," 2012 NATIONAL CONFERENCE ON COMPUTING AND COMMUNICATION SYSTEMS, pp. 1-5, 2012.
- [5] G. Parthasarathy, L. Ramanathan, K. Anitha and J. Y, "Predicting Source and Age of Brain Tumor Using Canny Edge Detection Algorithm and Threshold Technique," Asian Pacific Journal of Cancer Prevention, vol. 20, no. 5, pp. 1409-1414, 2019.
- [6] R. Kumar and et. al., "Accurate Prediction of COVID-19 using Chest X-Ray Images through Deep Feature Learning model with SMOTE and Machine Learning Classifiers," medRxiv, 2020. unpublished
- [7] C. Kamarajan et. al., "Random Forest Classification of Alcohol Use Disorder Using fMRI Functional Connectivity, Neuropsychological Functioning, and Impulsivity Measures," Brain Sciences, vol. 10, no. 2, 2020.
- [8] E. I. Zacharaki et al., "MRI-based classification of brain tumor type and grade using SVM-RFE," 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1035-1038, 2009.
- [9] R. J. Ramteke and K. Y. Monali, "Automatic Medical Image Classification and Abnormality Detection Using KNearest Neighbour," International Journal of Advanced Computer Research, vol. 2, no. 4, pp. 190-196, 2012.
- [10] P. Aiwal and S. Ansari, "Brain Tumor Detection Using KNN," International Journal of Scientific and Engineering Research, vol. 10, no. 12, pp. 187-193, 2019.
- [11] S. Das, O. F. M. R. R. Aranya and N. N. Labiba, "Brain Tumor Classification Using Convolutional Neural Network," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1-5, 2019.
- [12] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," Computers in Biology and Medicine, vol. 111, 2019.
- [13] "MRI, Magnetic Resonance Imaging mayfield brain and amp; spine Cincinnati, Ohio," mayfieldclinic.com. [Online]. Available: <https://mayfieldclinic.com/pe-mri.htm>. [Accessed: 09-Dec-2021].
- [14] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," Computers in Biology and Medicine, vol. 111, 2019.