HIGH SCHOOL EDITION
Journal of Student Research
Volume 11 Issue 2 (2022)

Computer Aided Diagnosis of Gliomas Using Machine Learning Classification Algorithms

Shreya Singh

Centerville High School, Centerville, TX, USA

ABSTRACT

The application of machine learning approaches to diagnose gliomas from magnetic resonance image data is becoming increasingly common as machine learning algorithms are being refined and inexpensive, high power computing resources are readily available. Most current glioma detection techniques involve a biopsy which is invasive and often has several long-lasting negative impacts. Many different classification techniques have been tried for image classification of scans containing gliomas including traditional models such as K-nearest neighbors and Random Forest. Deep learning models, specifically convolutional neural networks and pre-built models, have also been applied to the situation. This study applied 2 traditional machine learning models, Support Vector Machine and logistic regression classifier, to a publicly available dataset containing over 1200 brain scans of healthy and diseased patients. After normalization, 854 images were used to train the two models and 367 images were used to test the models. They were then evaluated on the parameters of: area under the receiver operating characteristic curve and sensitivity to determine which model performed better. Both models had similar AUC scores, but the SVM model had much higher sensitivity. Considering the possibly fatal ramifications of incorrectly diagnosing a patient who has been infected with glioma, it was determined that the SVM was better equipped to handle the classification task. Computer aided diagnosis of the tumor will hopefully be able to increase the survival rate of people diagnosing with gliomas.

Introduction

Gliomas are a category of intra-axial brain tumor that infect the glial cells that surround and support neurons (John Hopkins Medicine, n.d.). They are dangerous because tissue which is infected with a glioma is able to mix with the normal, unaffected cells of the brain often leading to adverse effects. Gliomas are one of the most commonly diagnosed cancers of the brain among the human population. One third of all brain tumor diagnoses are some form of glioma (John Hopkins Medicine, n.d.). There are many types of gliomas. Gliomas can occur in many parts of the brain including in: astrocyte cells, oligodendrocyte cells, and ependymal cells. A specific type of glioma is named based on which cells have been infected. Regrettably, people infected with gliomas will usually experience many excruciatingly painful symptoms such as severe headaches, intense nausea and vomiting, blurred vision and routine seizures (John Hopkins Medicine, n.d.). Unfortunately, the five year relative survival rate for adults over the age of forty five years old is under ten percent (American Cancer Society, 2020). The relative survival rate drops rapidly as the age at first diagnosis increases, adults over the age of 60 years old only have a five year survival rate of a little bit over 5 percent (American Cancer Society, 2020). Unfortunately, receiving a diagnosis of glioma is almost always fatal. There is the possibility that the extremely low survival rate for people who have the disease may partly be attributed to the difficulty of diagnosing it due to the fact that current detection techniques of the tumor require dangerously invasive biopsy techniques that often leave many lasting and harmful repercussions, such as frequent spells of dizziness, nausea and seizures. Computer aided diagnosis techniques will be useful in the noninvasive and accurate diagnosis of gliomas in

patients. These techniques can be even more accurate than human diagnoses (Bae et. al, 2020) of cancerous tissue due to gliomas.

Literature Review

As machine learning and computer vision algorithms are becoming increasingly available and efficient, the use of these algorithms for medical imaging applications are becoming commonplace. The basic premise of all computer vision algorithms is as follows: an algorithm is applied to a dataset (i.e. the past 6 months of magnetic resonance imaging scans collected by a radiology lab) and some insight provided by the data (i.e. benign vs malignant tumors). The algorithm learns from the data provided to make predictions on scans it has not seen before (Erikson et. al, 2017). In recent years many researchers have developed algorithms to classify scans as cancerous or not cancerous with varying success (Table 1). One of the recurring issues in trying to diagnose a patient with glioma is that doctors and researchers are unable to determine a cause for glioma. This means that gliomas are generally detected once the cancerous cells have accumulated in a large enough quantity that an individual starts to experience symptoms and informs his/her doctor about them. A lack of a clear indicator of what to look for in scans has caused difficulty for radiologists and machine learning algorithms to diagnose gliomas.

A certain amount of success has been found in using the compound O6-methylguanine-DNA-methyl-transferase (MGMT) promoter's methylation status as an indicator for any gliomas that have formed around astrocytes (Perkuhn et. al, 2018; Kansas et. al, 2017).While many different algorithms have been used, they all primarily intended to decrease the frequency of a false negative being reported. Such an incorrect diagnosis could end up with unintentionally lethal results. It was also a priority to reduce false positives as such reports required additional resources for further testing. It is important to note that a direct comparison of the attempts listed below is not recommended as all of these algorithms have not been tested on the same dataset before, but these attempts can be used as an indicator for the relative success of the types of algorithms that are available for use for computer-aided diagnosis of glioma patients.

Table 1. Comparison of machine learning algorithms used to classify scans

Attempted by	Model Used	Accuracy
Kanas et. al, 2017	Random Forest	70.2%
Kanas et. al, 2017	K- Nearest Neighbors (k =5)	73.6%
Bae et. al, 2020	Human readers	77.4%
Shin et. al, 2021	ResNet-50[1]	88.9%
Brunese et. al, 2020	Ensemble Learning	>99%[2]

[1] ResNet- 50 is a pre-trained convolutional neural network

[2] Trained on over 111, 205 images

Objective

The purpose of this research study is to provide an efficient implementation of a computer vision algorithm that has a higher accuracy and higher sensitivity than the current models available for the diagnosis of scans containing glioma infected tissue. It is hypothesized that one of the machine learning models tested will achieve higher accuracy and higher sensitivity than the current models created for the purpose of diagnosing glioma in patients.

Methods

Data Acquisition

The dataset used in this study, Brain Tumor Classification (MRI), was acquired from the community published data platform www.kaggle.com. Originally, Navoneel Chakrabarty and Swati Kanchan collected scans from multiple different sources through Google Images and Dr. Shashikant N. Ubhe of Chandrabhaga Clinic and Nursing Home reviewed them to ensure the images' correct ground-truth labels. This dataset was originally put together for a university computer science project and was afterwards publicly released on www.kaggle.com for use by others.

The dataset contains over 3000 scans of pituitary, meningioma, glioma, and non-cancerous patients. For the purpose of this study, only 826 scans of the glioma patients and 395 scans of non-cancerous patients were used. Scans of various sizes were used with dimensions ranging from: 225 pixels by 225 pixels to 512 pixels to 512 pixels. Images of the brain were taken in the axial, sagittal, and coronal planes to allow for maximum coverage of the head and brain region.

Data Pre-Processing

Prior to being fed into the machine learning algorithms, the image data was processed using the Pandas and NumPy libraries (Figure 1). First, all images were given a label of 0 (representing non-cancerous) or 1 (representing glioma infected tissue) to identify their target label. All the images were then cropped to an uniform size of 175 pixels by 175 pixels (Figure 2). All images were then converted to NumPy arrays to allow for numeric transformations to be applied. It was then necessary to flatten the three dimensional images into two dimensional images, so that it could be used in computer vision algorithms. To reduce the complexity of the algorithm required, the final transformation made to the images was to divide all red-blue-green values by 255 to convert all images to grayscale images with colors represented by an intensity value from 0 to 1. Lastly, seventy percent of the images (854) were randomly assigned to the training set (data available for the computer to learn from) and thirty percent of the images (367) were put into the testing data set, which will be used later to evaluate the model.

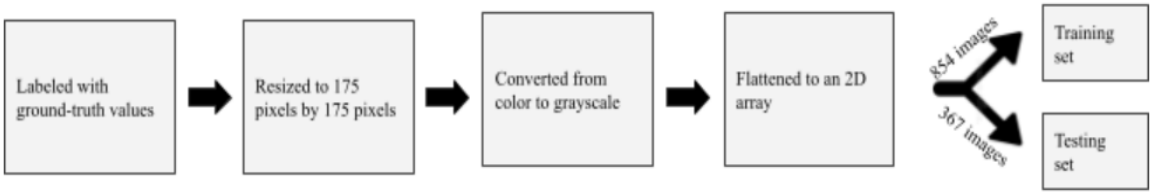

Figure 1. Flowchart of the steps that were taken during pre-processing of the dataset

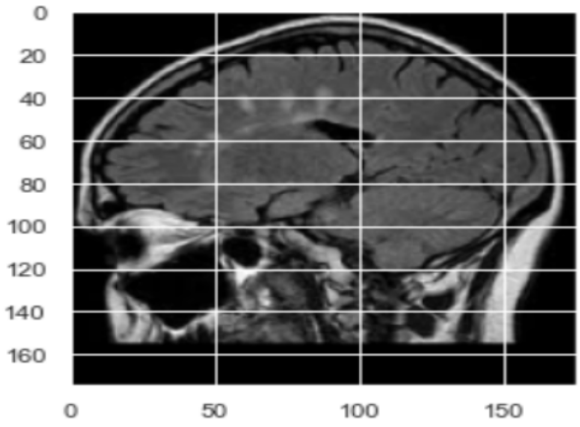

Figure 2. Patient scan after transformations

Models Used

Both of the models created for testing were built using the sci-kit learn library and the Python programming language in a Jupyter notebook. The first model used on the data was a Support Vector Machine (SVM) utilizing a linear kernel. An SVM works by finding the hyperplane that maximizes the gap between the positive and negative class (Figure 3).

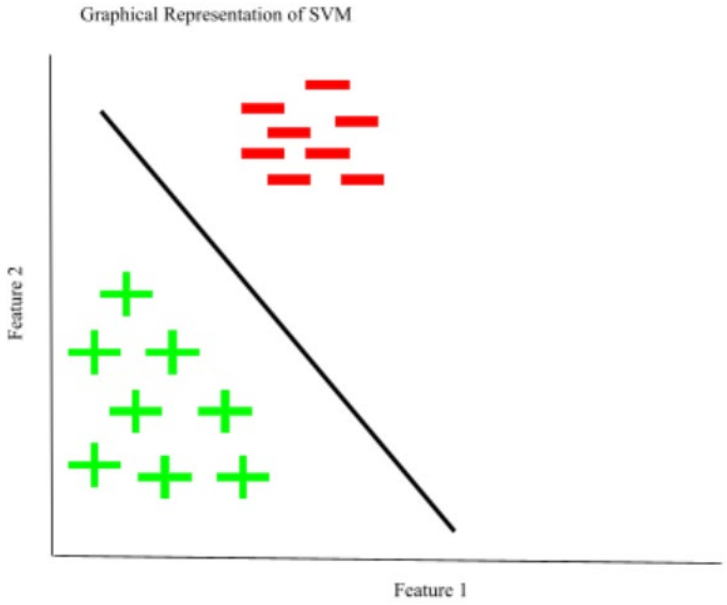

Figure 3. Graphical Representation of SVM implementation on a two dimensional surface. Green plus signs (Positive class) is separated from red negative signs (negative class) by the decision boundary, bolded above.

Different possibilities for the kernel shape were tested including polynomial, squared exponential kernel, sigmoid, and a precomputed option. In the end, a linear kernel was decided upon because the SVM utilizing a linear kernel required the least training time when compared to the SVM's training time when using the non-linear kernel options.

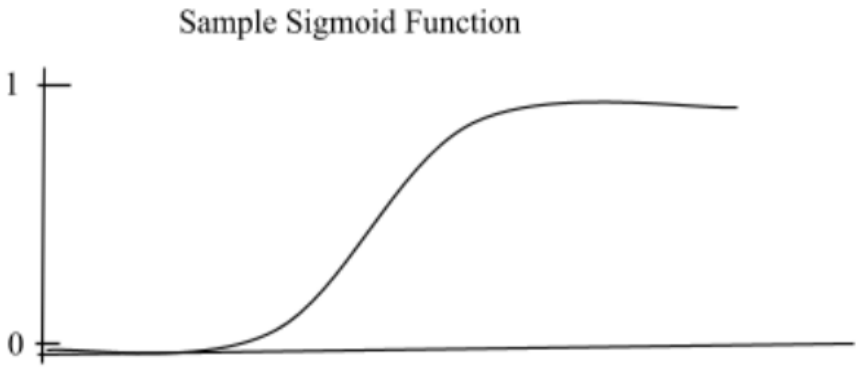

Figure 4. Shape of a general logistic regression curve (sigmoid function)

Another model that was tested on the dataset was a logistic regression model (Figure 4). In this case, the logistic regression model will use a sigmoid function ($p = 1/ (1 + e^{-(b_0+b_1x_1+b_2x_2+b_3x_3+\ldots+b_nx_n)})$) to estimate the probability of a scan containing cells affected with glioma. If the probability is greater than 0.5, it will be classified in the 1 class (as a case of glioma); otherwise the algorithm will classify it in the 0 class (as a case of no tumor).

Model Evaluation

Both of the models were trained on the 854 images in the training dataset and tested on 367 images. Then, each model was evaluated on the basis of the area under the receiver operating characteristic curve and sensitivity (Equation 1). Sensitivity quantifies how well the model is able to avoid classifying an image containing glioma infected tissue as a non-cancerous case.

Equation 1: Used to quantify the sensitivity of a machine learning model

$$\frac{True\ Positives}{True\ Positives\ +\ False\ Negatives}$$

Results

After being tested on the testing image dataset of 367 images, the SVM model had an area of 0.8756 under the receiver operating characteristic curve (AUC score) while the logistic regression had an AUC score of 0.8684. A direct comparison of the 2 models is graphed in Figure 4. Additionally, the Support Vector Machine was able to classify 245 of the testing points correctly as a form of glioma (true positive), 5 glioma containing scans were incorrectly classified as not cancerous scans (false negative), 67 scans of non-cancerous patients were correctly classified (true negative), while 50 non-cancerous patients were incorrectly classified with a diagnosis of glioma (false positive). These results are summarized in the confusion matrix below (Figure 5).

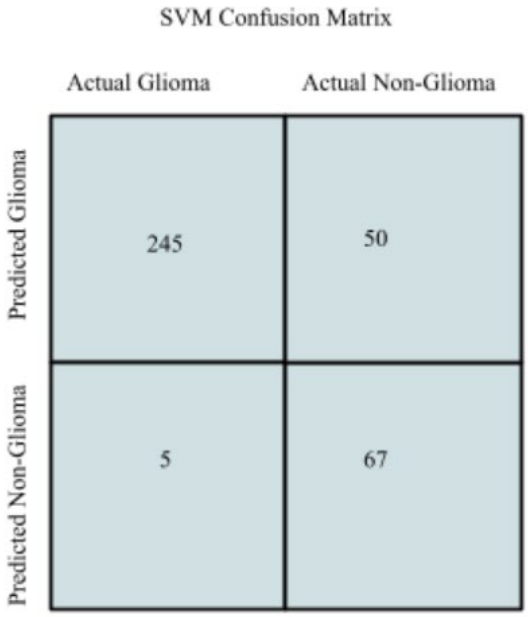

Figure 5. Confusion matrix detailing the computer's classification of all images in the testing dataset when trained using a SVM

The logistic regression classifier was able to classify correctly 220 of the images containing glioma. It correctly classified 81 of the non-tumorous scans. This model incorrectly classified 30 scans containing glioma infected tissue as non-cancerous scans, and it also misclassified 36 non-cancerous scans as containing glioma infected tissue. These results are summarized in the confusion matrix below (Figure 6).

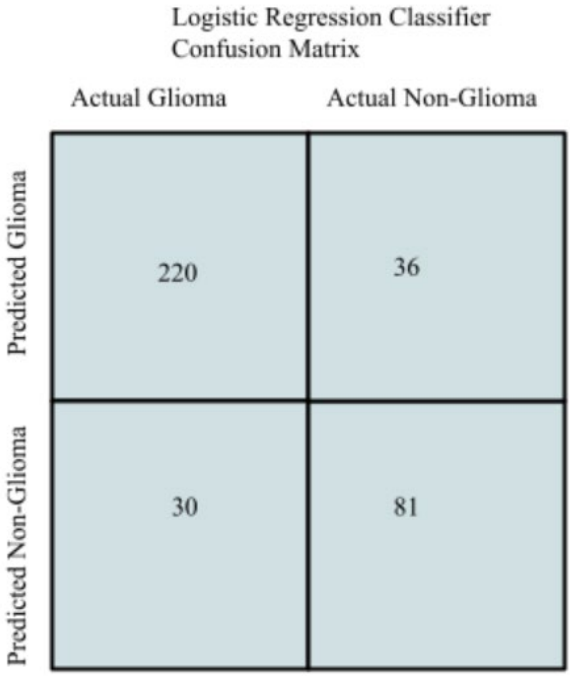

Figure 6. Confusion matrix detailing the computer's classification of all images in the resting dataset when trained using a logistic regression classifier.

Calculating the sensitivity of both models from Equation 1 yields 98% as the sensitivity for the model classifying scans using a SVM and a sensitivity of 88% for the model classifying scans using the logistic regression classifier. All the values for evaluation metrics used to compare the two models are summarized in Figure 7.

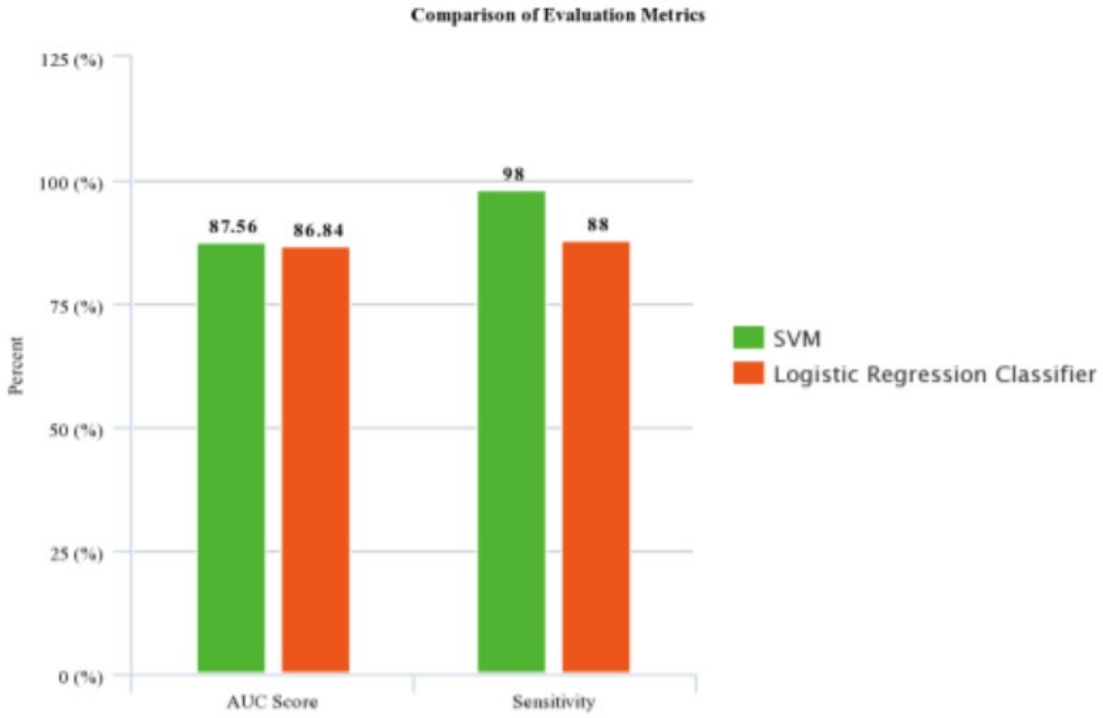

Figure 7. Summary of the metrics used to compare the two machine learning models.

Discussion

Both the SVM and the logistic regression classifier had similar AUC scores at .8756 for the SVM trained classifier and .8684 for the Logistic Regression Classifier. This metric may misleadingly create the impression that the two models are equally successful at classifying the 367 images they were tested on, but it is necessary to remember the use case for each of these models. It is blatantly apparent that the SVM algorithm is a very conservative model when compared to the logistic regression classifier. It has a ten percent higher sensitivity than the logistic regression classifier. This means that the SVM model is better at identifying all the patients that do have glioma. With a sensitivity of 98% , almost every scan containing glioma infected tissue will be marked as such. Considering that the practical use case for this model is for diagnosing a potentially fatal disease, it would be better to classify the patients' scans using the SVM model, so that a case of glioma is not missed. An incorrect diagnosis of being told that a patient does not have the disease when the patient does in fact have the diagnosis can end up costing human lives.The SVM trained for this task was able to achieve 98% sensitivity and an accuracy of close to 90%. This model was coded and tested on a Windows Surface Laptop (Generation 2) without using expensive and complicated computing resources, such as cloud-hosted virtual machines, and was able to correctly identify almost every case of glioma in the scans it was tested on. This shows that a SVM trained classifier is a very promising algorithm to use for the accurate and non-invasive diagnosis of patients with gliomas. The algorithm does its job well, as noted by the sensitivity and AUC score, and it can be used without requiring a massive technological resource investment by radiology departments.

Conclusion

This implementation of a traditional machine learning algorithm (SVM) shows that it is absolutely possible to use machine learning to successfully diagnose gliomas. Both models had higher accuracies than the 77.4% accuracy of humans classifying scans in a 2020 experiment (Bae et. al, 2020) and further testing of this algorithm can be done by comparing the model's classification to humans' performance when classifying scans. A possibility for further improvement of the classifier is to test the effect of applying dimensionality reduction to the images prior to feeding the data into the algorithms. Dimensionality reduction would reduce the number of features of the dataset and thereby allow for a reduction in training time of the algorithm and reduce the need for high power compute resources to run the model. A limitation of this study was a lack of high power compute resources with sufficient central processing unit size and storage space. More complex algorithms could be tested for this task by hosting a virtual machine on cloud hosting services such as Google Cloud Platform or Amazon Web Services. With higher power computing resources it may be possible to train a convolutional neural network (CNN) or use one of the many pre-trained CNNs available, such as ResNet-50 or GoogLeNet (Cinar & Yildirim, 2020) for this task.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

Works Cited

Johns Hopkins Medicine. (n.d.). *Gliomas.* *https://www.hopkhttps://www.hopkinsmedicine.org/health/conditions-and-diseases/gliomasinsmedicine.org/health/conditions-and-diseases/gliomas*

American Cancer Society. (2020, May 5). *Survival Rates for Selected Adult Brain and Spinal Cord Tumors.* https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/detection-diagnosis-staging/survival-rates.html#written_by

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine Learning for Medical Imaging. *RadioGraphics*, *37*(2), 505–515. https://doi.org/10.1148/rg.2017160130

Bae, S., An, C., Ahn, S. S., Kim, H., Kim, S. W., Park, J. E., Kim, H. S., & Lee, S.-K. (2020). Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: Model development and validation. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-68980-6

Perkuhn, M., Stavrinou, P., Thiele, F., Shakirin, G., Mohan, M., Garmpis, D., Kabbasch, C., & Borggrefe, J. (2018). Clinical evaluation of a multiparametric deep learning model for glioblastoma segmentation using heterogeneous magnetic resonance imaging data from clinical routine. *Investigative Radiology*, *53*(11), 647–654. https://doi.org/10.1097/rli.0000000000000484

Kanas, V. G., Zacharaki, E. I., Thomas, G. A., Zinn, P. O., Megalooikonomou, V., & Colen, R. R. (2017). Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma. *Computer Methods and Programs in Biomedicine*, *140*, 249–257. https://doi.org/10.1016/j.cmpb.2016.12.018

Shin, I., Kim, H., Ahn, S. S., Sohn, B., Bae, S., Park, J. E., Kim, H. S., & Lee, S.-K. (2021). Development and validation of a deep learning–based model to distinguish glioblastoma

from solitary brain metastasis using conventional MR images. *American Journal of Neuroradiology*, *42*(5), 838–844. https://doi.org/10.3174/ajnr.a7003

Brunese, L., Mercaldo, F., Reginelli, A., & Santone, A. (2020). An ensemble learning approach for brain cancer detection exploiting radiomic features. *Computer Methods and Programs in Biomedicine*, *185*, 105134. https://doi.org/10.1016/j.cmpb.2019.105134

Çinar, A., & Yildirim, M. (2020). Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Medical Hypotheses, 139, 109684. https://doi.org/10.1016/j.mehy.2020.109684

Chakrabarty, N. & Kanchan, S. (2020). Brain Tumor Classification (MRI) (Version 2). [Data set]. https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri