Brain Tumor Classification Using MRI Images.

Chandrahaas Kalanadhabhatla, Siva Rohith Jampana, Pranav Abishai Moses, and Shaik, Sabiha Tabassum.

Abstract—In the field of medical challenges, brain tumors and nervous system cancers rank at 10Th leading cause of death in the United States. Despite the availability of cures, lives are still lost due to failure of early classification, a task primarily done by expertise of radiologists or neurospecialists. The urgency of this matter becomes apparent when considering the volume of MRI scans these professionals handle daily, often numbering in the hundreds. In this project, we aim to compare machine learning models used to classify brain tumors and understand the impact machine learning and artificial intelligence can have on our daily lives. The importance of this project lies not only in the pursuit of saving lives but also in changing the landscape of brain tumor classification, aligning with the broader mission of healthcare stakeholders dedicated to advancing medical outcomes. We are using machine learning models like Support Vector Machine, Random Forest, Naive Bayes and a Neural Network based model to understand the brain tumor classification ability of the artificial intelligence. We have used a publicly available data set which has more than 10,000 images of various brain tumors classified into four classes which are: glioma, Meningeoma, Pitutary and no tumor. We have used various feature extraction techniques to analyze the image and train the models to classify them effectively. We have discovered that all the models were able to predict the classification or type of the tumor effectively as the amount of data used to train was increased. We have compared all the performances of the models to understand the ability and efficiency of each model in differentiating the tumors. There has been a lot of research done in building models to classify the brain tumor and the majority of the research has been focused on the use of deep learning methods. These were complex and took long periods of time to train and make ready to use but with the increase in computational capabilities of the machines with the help of GPU's we are able to do them at a significantly reduced amount of times. The same models are already in use for many other applications such as in industries, where detection of a faulty product in milliseconds is critical to detect it and prevent it from going into production.

Index Terms—Brain tumor; Classification; MRI images; Rnadom forest classifier; Naive bayes; Convolutional Neural Network; Confusion Matrix; Support Vector Machine (SVM); Feature engineering.

1 Introduction

Brain tumors and nervous system cancers are the 10th leading cause of death in the United States within the medical field. Although there are treatments available, lives continue to be lost due to the lack of early detection, which heavily relies on the specialized knowledge of radiologists and neurospecialists.

The urgency of the situation is evident when we consider the high volume of MRI scans that these professionals handle daily, often reaching hundreds. Our objective in this project is to create a reliable machine-learning model that focuses on the detection of brain tumors. The significance of this project extends beyond the preservation of lives and has the potential to transform the field of brain tumor detection, aligning with the larger mission of healthcare stakeholders committed to enhancing medical outcomes.

A brain tumor refers to an abnormal cluster of cells that forms a mass within the brain. Due to the inflexible nature of the skull, any growth within this limited space can give rise to complications. Brain tumors can be categorized as either cancerous (malignant) or noncancerous (benign). As these tumors expand, they can raise pressure inside the skull, potentially causing damage to the brain and presenting a life-threatening danger. The utilization of Magnetic Resonance Imaging (MRI) plays a pivotal role in the detection of brain tumors, providing highly detailed and non-invasive imaging that facilitates precise diagnosis, treatment planning, and disease progression monitoring, ultimately improving patient outcomes. Traditionally, the identification of brain tumors in each patient's MRI scan relied on the expertise of medical professionals, which is time-consuming and prone to human errors.

Currently, there is a gap in knowledge regarding the detection of brain tumors in MRI scans, as it heavily relies on the manual interpretation of radiologists and neurospecialists. This approach is susceptible to time limitations and human errors. It is crucial to improve the effectiveness and accuracy of this detection process, highlighting the significance of creating a strong machine-learning model. Our research endeavors to bridge this gap by utilizing sophisticated machine learning techniques to automatically and precisely detect and categorize brain tumors in MRI scans. This approach aims to address the existing constraints associated with timely and accurate diagnoses, providing a solution to enhance the overall effectiveness of brain tumor detection.

We have discovered that the time taken by the models to extract features and effectively classify them played a huge role. With the effective use of GPU's we were able to train the neural network based model quickly. Using CPU, we had been running the same training script for 5 hours and we did not reach the completion but using a GPU we were able to train and inference the results in almost less than 10 minutes of time. But for the traditional machine learning models, we were unable to use GPU's as they were not

C. Kalanadhabhatla, R. Jampana, P. Moses and S. Tabassum, are with the College of Information, University of North Texas, Denton, TX, 76207 USA. Emails: chandrahaaskalanadhabhatla@my.unt.edu, sivarohith-jampana@my.unt.edu, pranavabishaimoses@my.unt.edu and sabihatabassumshaik@my.unt.edu

supported and we had to spend 6 to 10 hours extracting features and training the models. This is a research issue where there can be more advancements in improving the computational time. This implementation can further help improving the efficiency of Brain Tumor Classification.

The field of brain tumor diagnostics requires significant improvement to overcome the drawbacks associated with manual interpretation by healthcare professionals, which is prone to inconsistencies and human mistakes. The obstacles faced, such as the large number of MRI scans and the complex task of early tumor detection, call for a revolutionary solution. Our proposed machine learning model represents a crucial advancement, as it provides an independent and precise method for detecting and categorizing brain tumors in MRI scans. By addressing the current limitations and introducing an automated approach, this research not only enhances the efficiency of diagnostic processes but also contributes to the progress of healthcare outcomes in the field of brain tumor detection.

2 Related Work

[1] The paper utilizes a publicly available brain tumor dataset for assessing its proposed method. Although the specific CNN model for image processing is not mentioned, the study achieves an overall classification accuracy of 98.04%. The paper discusses other approaches such as modified local binary patterns (LBP), discrete wavelet transform (DWT), and pre-training of a generative adversarial network (GAN) for brain tumor classification.

[2] In the document, Convolutional Neural Networks (CNN) are employed to process brain tumor images, resulting in an impressive accuracy rate of 97.5%. This achievement is noteworthy as the CNN approach demonstrates low complexity compared to other methods. Additionally, the paper delves into alternative techniques such as Fuzzy C-Means segmentation and AdaBoost classification, providing further insights into brain tumor analysis.

[3] This paper employed brain tumor images sourced from the Kaggle database and introduced a novel deep learning model that incorporated elements from transfer learning methods like DenseNet, VGG16, and basic CNN architectures for image processing in the CNN. Additionally, the study compared and explored various transfer learning methods, including DenseNet, VGG16, and basic CNN architectures, alongside the newly proposed model. However, the paper did not explicitly discuss the use of feature engineering techniques in the study.

[4] This paper discusses, A hybrid model named Transformer-Enhanced Convolutional Neural Network (TECNN) is proposed for the classification of brain tumors. The TECNN combines the capabilities of a convolutional neural network (CNN) and a transformer model. Experimental results demonstrate that the TECNN achieves impressive average accuracy rates of 96.75% on the BraTS 2018 dataset and 99.10% on the Figshare dataset, surpassing the performance of existing methods. By leveraging deep learning techniques, the proposed approach eliminates the requirement for manual feature engineering as the model autonomously extracts significant features from the data.

[5] The paper utilizes a combination of techniques to classify brain tumors. The analysis is conducted on a dataset obtained from Kaggle, resulting in an impressive 98% accuracy in tumor recognition. The Harris Hawks Optimization algorithm is introduced to optimize the parameters of the Convolutional Neural Network (CNN). Additionally, the paper focuses on extracting diverse features from segmented regions as part of the classification process.

[6] The paper employs MRI images for brain tumor detection, utilizing the VGG16 model in a CNN for image processing. It achieves a notable accuracy of 99% when using ensemble feature vectors. Alongside SVM and KNN, feature engineering techniques are applied, combining hand-crafted features derived from GLCM with deep features extracted by VGG16 to enhance the method's ability to discriminate brain tumors effectively.

[7] The paper utilizes a Kaggle dataset to classify brain tumors and employs the AlexNet CNN model for image processing. The combinations of AlexNet CNN with various machine learning classifiers achieve accuracy rates of 88.75% (using BayesNet), 98.15% (using SMO), 86.25% (using NB), and 100% (using RF). The paper suggests incorporating other classifiers such as BayesNet, SMO, NB, and RF in conjunction with AlexNet CNN, but does not explicitly discuss feature engineering techniques.

[8] In the study, three datasets comprising T1-Weighted (T1W), T2-weighted (T2W), and fluid-attenuated inversion recovery (FLAIR) MRI sequences were employed for brain tumor classification. Images were processed using a combination of five established deep learning models: AlexNet, VGG16, ResNet18, GoogleNet, and ResNet50. The ensemble algorithm utilizing these models achieved impressive testing accuracies of 98.88% ± 0.63% for FLAIR, 97.98% ± 0.86% for T2W, and 94.75% ± 0.61% for T1W-MRI data. The paper also discussed the significance of machine learning and deep learning techniques in computer-aided diagnosis (CAD) tools for medical applications, highlighting the automatic feature extraction capabilities of DL algorithms like CNN, thereby eliminating the need for manual feature engineering.

[9] The research paper utilized a dataset consisting of 233 MRI brain tumor images. The authors employed pretrained CNN models, such as VGG-16, ResNet-50, and Inception-v3, to analyze and classify the images. The paper also mentioned the exploration of alternative methods for brain tumor analysis and classification, without providing specific details. However, there was no explicit mention of feature engineering techniques being employed in the study.

[10] The paper utilizes undisclosed brain MRI datasets and employs pre-trained CNN models to process the images and extract deep features. These deep features are then combined with machine learning classifiers for brain tumor classification. The paper introduces a feature evaluation and selection mechanism, as well as a feature ensemble method. However, it does not explicitly mention any additional feature engineering techniques beyond the use of pre-trained CNN models.

[11] In this study, a dataset comprising 3064 T1-weighted contrast-enhanced MRI images from Nanfang Hospital and General Hospital, Tianjin Medical University, China, was utilized. The researchers developed a novel

CNN architecture for brain tumor classification, which is simpler compared to pre-trained networks. The highest accuracy achieved through 10-fold cross-validation was 96.56% using record-wise cross-validation with the augmented dataset. The paper also explores alternative methods from the literature, such as different classification algorithms and modified pre-trained networks, while not explicitly mentioning any feature engineering techniques employed in their CNN architecture.

[12] The paper employs a publicly available dataset consisting of T1-weighted contrast-enhanced brain magnetic resonance images. To process the images in the convolutional neural network (CNN), a modified version of EfficientNet known as "dense EfficientNet" is utilized, incorporating dense and drop-out layers. The proposed approach demonstrates impressive performance with a training accuracy of 99.97% and a testing accuracy of 98.78%. Alongside the dense EfficientNet model, the paper explores alternative techniques such as dual path CNN architecture, min-max normalization, fuzzy logic-based enhancement, U-Net classification algorithms, and diverse pre-processing methods. Although feature engineering is not explicitly mentioned, the paper employs data augmentation with min-max normalization to enhance the contrast of tumor cells, which can be viewed as a form of feature engineering.

[13] In the study, the researchers utilized the BraTS datasets to conduct experiments aimed at classifying multimodal brain tumors. They employed pre-trained CNN models, namely VGG16 and VGG19, to process the images and extract deep learning features. The proposed approach yielded high accuracy rates of 97.8%, 96.9%, and 92.5% for the BraTS2015, BraTS2017, and BraTS2018 datasets, respectively. The paper also explored the application of various techniques such as linear contrast enhancement, histogram equalization, transfer learning, feature fusion, feature selection, and the implementation of the Extreme Learning Machine (ELM) algorithm for classification purposes.

[14] This paper utilized a T1-CE MRI dataset from 233 patients to train and evaluate a Multiscale Convolutional Neural Network model for the purpose of brain tumor classification and segmentation. The proposed method achieved an impressive tumor classification accuracy of 0.973, surpassing the performance of alternative techniques on the same dataset. The paper also explores the transition from classical machine learning approaches, such as SVM, ANN, SOM, and kNN, to the adoption of deep learning methodologies. While specific details regarding feature engineering are not provided, the proposed model demonstrates the ability to handle variations in input images without the need for preprocessing steps aimed at removing skull or vertebral column components.

[15] In the study, a dataset containing 25,000 brain MRI images, consisting of both abnormal and normal cases, is employed. The paper presents a novel approach called differential deep convolutional neural network model (differential deep-CNN) to process these images. Remarkably, this combined model achieves an accuracy rate of 99.25%. While traditional models and supervised methods are mentioned, the paper places emphasis on the growing interest in unsupervised techniques and the emergence of deep learning-based models in medical image analysis. Although specific

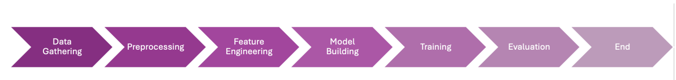

Fig. 1: Process flow of the problem

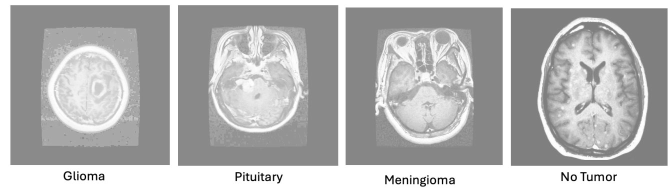

Fig. 2: Images from the dataset

feature engineering techniques are not explicitly stated, the paper highlights the utilization of differential operators to generate additional feature maps, thereby improving the performance of the proposed model.

[16] The paper trains the proposed model using the BTD dataset and demonstrates its superior accuracy compared to state-of-the-art models. Additionally, it explores alternative approaches such as CNN-LSTM, histogram equalization with CNN, FAHS-SVM, and utilization of pre-trained deep learning models. Data augmentation techniques are employed, but feature engineering methods are not explicitly mentioned.

3 Proposed Method

Overview: This section briefly describes the overall flow the project in step by step method. As we can see in the figure 1, this shows the basic flow of the method that was implemented in this project.

1. Data Gathering: In this phase, we had explored various datasets available for public without any restrictions. We found one dataset that had more than 10,000 images [17]. This dataset has four classes which are glioma, Meningeoma, Pituitary and no tumor. This phase is the most important one because the performance of the model highly depends on the quality of the data in the dataset. Figure 2, shows some sample images from the dataset for each class. We also made sure the number of data points for each class was approximately the same to prevent any kind of bias in the model. This helps the model to treat each class equally without giving any importance to a specific class.

2. Pre-Processing: In this project we are using two approaches: 1. Traditional Machine learning approach 2. Deep Neural Network approach. Data pre processing the step where all the data in the dataset is converted to be consistent with each other and in a format that can be used to train the model without any issues. Data preprocessing techniques invlove reading an image, by default image data is not in a readable format for the machine learning models, we make use of open source libraries such as matplotlib or openCV, we have used openCV to read images in this project, it converts the image into a matrix of numbers representing each pixel in the image. The next step after reading the image is to resize them, we want all the images in the dataset to have the same resolution so we resize them to a consistent number like 150X150 or 128X128. By the end of this step we

will end up having an object with 150X150X3 dimension matrix. Then as the last step we normalize the pixel values in the image, this helps the model better differentiate bright pixels from the light ones.

3. Feature Engineering: To enhance the model's ability to discern critical image features, we employed feature engineering techniques. We utilized Gabor and Sobel filters with varying parameters, resulting in a richer set of features that accentuated specific textures and edges within the MRI scans. This pre-processing step effectively guided the machine learning models towards the most relevant aspects of the images. In contrast, deep neural networks were leveraged for their inherent feature extraction capabilities. Through their multi-layered architecture, these models were able to automatically learn and identify significant patterns within the raw MRI data. This approach allowed the deep learning models to discover intricate relationships and hidden characteristics within the images, ultimately leading to a more robust understanding of the brain tissue.

Gabor Kernel: The Gabor filter is like a special magnifying glass for image processing. It helps analyze specific textures and edges. Imagine the filter as a ripple on a pond, but instead of spreading outwards, it's confined to a small area. This lets the filter focus on details in a particular direction. By using Gabor filters with different orientations and frequencies, we can highlight vertical lines, horizontal stripes, or other textural patterns in an image. This is useful for tasks like identifying fingerprints, classifying fabrics, or even recognizing facial expressions in pictures. We have used varying parameter by changing theta, lambda and gamma to create more than one filter and capture intricate parts of the image.

Sobel Kernel: The Sobel filter is a detective in the world of image processing, particularly good at finding edges. Imagine an image as a vast landscape with changes in brightness marking cliffs and ridges. The Sobel filter uses two small masks, one sensitive to horizontal changes and the other to vertical changes. As it slides these masks across the image, it looks for sharp contrasts in brightness. If there's a big difference between neighboring pixels, the Sobel filter knows there's an edge nearby. This helps identify objects, analyze shapes, and even extract features for computer vision tasks.

We just use the this filter for the edge features. This added one more feature to the training set.

4. Model Building: For machine learning models, we made use of sklearn to import the models and train them, but for deep neural network, we had to import a pre-trained model and add layer to it to make it suitable for our project's inputs and outputs. In this project, we have use a total of four model: EfficientNetB0 model [18], which is a pre-trained neural network based model which is extensively used for image classifications, Support Vector Machine classifier, Random Forest classifier and Naive Bayes Multinomial classifier.

5. Training: This critical stage, known as the training phase, involves meticulously calibrating the model's parameters. For instance, we determine the optimal number of trees in a random forest or establish the learning rate and number of epochs for deep learning models. These parameters significantly influence the model's ability to learn and

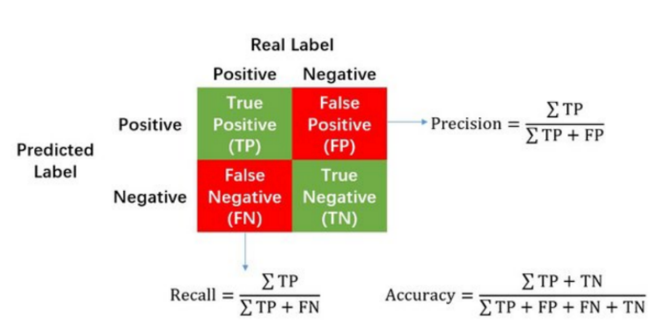

Fig. 3: Confusion matrix

generalize from the data. The training process itself can vary considerably in duration, ranging from a mere ten minutes to ten hours or more. This timeframe depends heavily on the complexity of the data and the computational resources available on the training device.

6. Evaluation: Following the training phase, we assessed the model's proficiency using a separate dataset known as test data. By feeding the test data into the trained model, we prompted it to generate predictions for unseen examples. Subsequently, we meticulously compared these predictions against the ground truth labels, which represent the actual classifications for each data point. This meticulous evaluation process involved calculating various performance metrics, including accuracy, precision, and recall. Each metric offers valuable insights into the model's strengths and weaknesses. To gain a more comprehensive understanding of the model's performance across different classes, we employed a confusion matrix. This visualization tool effectively summarizes the distribution of correct and incorrect predictions for each category.

Confusion Matrix: In figure 3, we see the structure of a confusion matrix, it gives us the number of predictions or results or outputs that were predicted as correct and are correct according to the ground truth (True Positives), the outputs that were predicted as correct but are wrong according to the ground truth (False Positives), the results that were shown as wrong and are actually wrong (True Negatives) and the inferences that were shown as wrong and are actually not wrong (False Negatives).

Using these four measures, we will be able to calculate the metrics which will be used to evaluate the performance of a classification model.

Accuracy: It is the ratio of sum of true positives and true negatives to the sum of all the input data points. It is used when the data we have is balanced across all the classes of the dataset. It can be misleading if the dataset is not balanced. Ex: let's say correct predictions are 80 out of a total of 100 data points, this seems like a good value since accuracy is 80% but there could be a case where positive values in the data set are about 80% and all the predictions made by the model are in favor of that majority, it neglects all the negative values and this cannot be accounted by accuracy.

Precision: It is the ratio of sum of true positives and the sum of true positives and false positives. We use precision when the penalty for the false positives is high. Ex: In

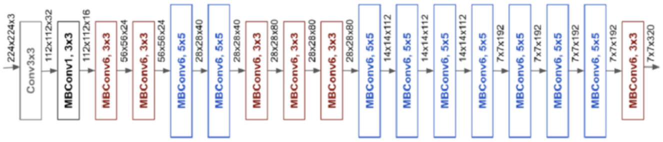

Fig. 4: Efficientnetb0 Architecture

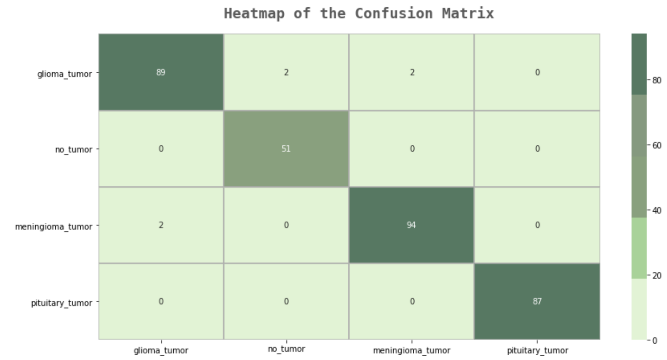

Fig. 5: Confusion Matrix for Deep Learning Model

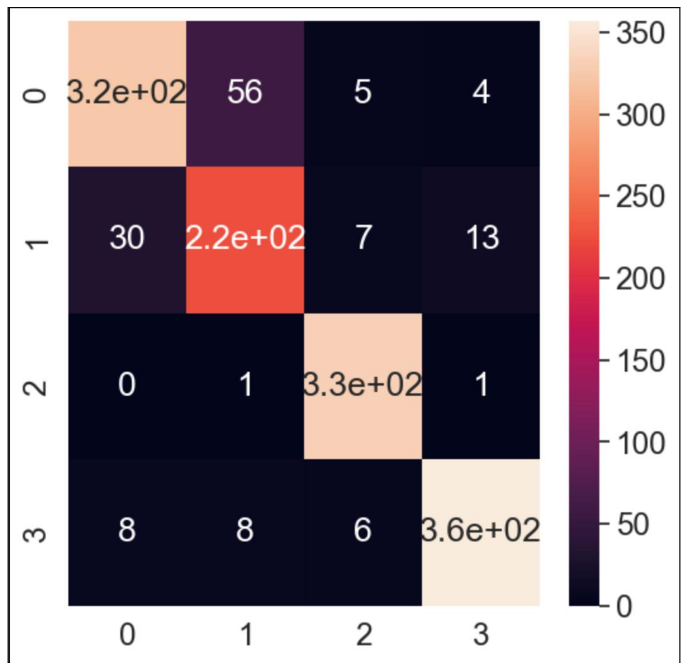

Fig. 6: Random Forest Confusion Matrix

medical industry where the cost of a false positive of a serious disease can lead to trauma and unnecessary costs.

Recall: It is the ratio of sum of true positives and the sum of true positives and false negatives. It is generally used when the cost or penalty of false negatives is the focus. Ex: In a security system, a false negative for a security breach could leave the system vulnerable. So, in this case, we want a model with high recall to ensure we catch most of the actual threats.

F1 score: It is metric used when the solution calls for an evaluation metric that gives importance to both precision and recall. It is the harmonic mean of precision and recall. It is used when incorrectly classified events are of higher importance rather than just the positives or negatives.

4 Experiment Results

Model 1: For our first model, we have used a deep neural network which was pre-trained and extracted from tensor flow. The pre-trained model used was: EfficientNetB0, we can see the architecture of the network in the figure 4. We added few more layer at the end to make the model usable for multi class classification as our problem deals with four classes.

We ran the same model on two different machines which made of CPU and GPU respectively. We observed a significant difference in the time take to process the same events. Although the accuracies produced were same. Using the CPU, we were able to train the model in approximately 6 hours whereas the model trained on GPU was able to finish within approximately 10 hours. Figure 3 represents the architecture of the model. We added a drop out layer a dense layer with soft max activation function to perform multi class classification and a layer to resize the output to the required format.

This model worked well and it produced the results shown in the figure 5, as you can see most of them were classified as expected and very less number of inputs from the text were classified wrongly.

We were able to achieve f1-score greater than 96% for all the classes, precision greater than 95% for all the classes and recall greater than 95% for all the classes.

This was the best performing model of all, we will compare all the models in detail at the end of this section. Model 2: We have use Random Forest as our next model. Random Forest is a powerful tool for classifying things into multiple groups. It builds a bunch of decision trees, each with slightly different training data and considered splits. When a new data point arrives, it travels through all the trees and gets a vote for a class from each. The final class is chosen by the majority vote, making it like a team of experts deciding together for better accuracy. We used 50 estimators for our model.

Training this model was fairly quick, the whole process of training was completed under a minute.

Figure 6, represents the confusion matrix for random forest classifier model. The precision for all the classes using this model was greater than 77%, recall was greater than 81% and the f1-score was greater than 79%. We were surprised at the results that the model achieved. From the figure 5, we will be able see that confusion matrix of the random forest model and most of the instances were classified accurately. We also tested our model using random inputs from the training set as well and found that the model was accurate most of the times. Each number represents the class of the tumor. glioma_tumor: 0, meningioma_tumor: 1, no_tumor: 2, and pituitary_tumor: 3.

Model 3: For the next model, we have used a Support Vector Machines (SVMs), they are great for separating data into distinct groups, even for multiple categories. Imagine data points are like islands, and you want to draw clear water boundaries between them. SVM finds the widest possible "moat" between the island groups, making classifi-

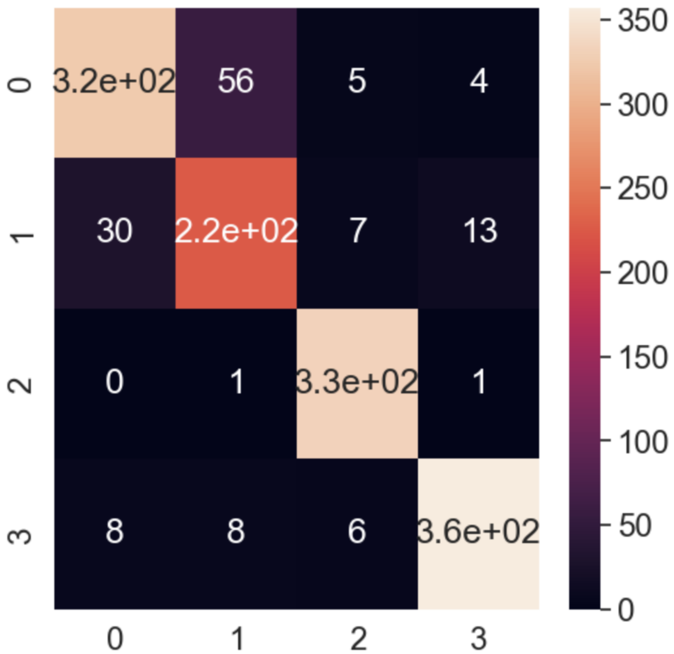

Fig. 7: SVM confusion matrix

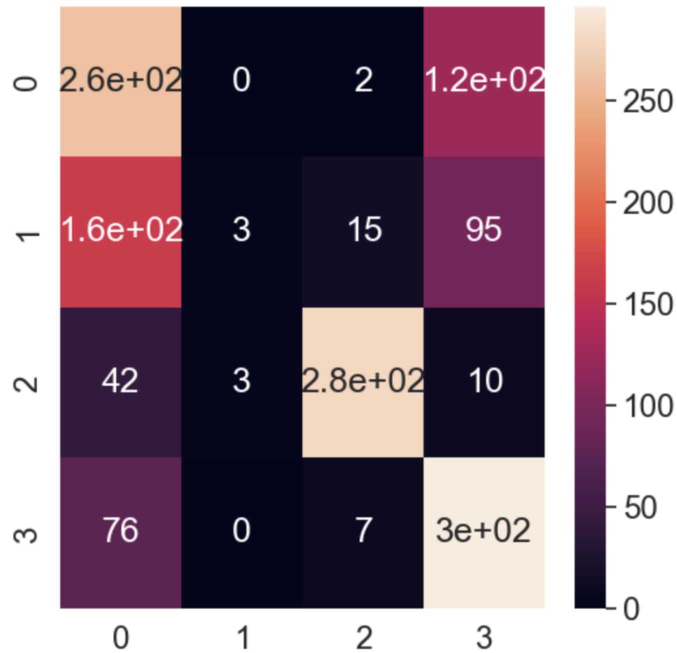

Fig. 8: Naive Bayes Confusion Matrix

cation clear-cut. For multi-class problems, SVM gets clever. It breaks the problem into several two-island challenges, one for each unique class comparison. By solving many two-class problems, SVM can effectively handle multiple categories overall.

This model took us around 99 minutes to train and about 97 minutes to predicts the results.

With model, we have achieved an overall precision greater than 51% for all the classes, recall greater than 53%, and an f1-score greater than 52% and these results can be seen in the figure 7. We have also noticed that the performance of meningioma tumor was very low even after using equal number of data for all the classes. The accuracy of the model was not great, we calculated the accuracy of the model to be 74%.

From the confusion matrix in the figure 6, we can see there were fairly many data points that were wrongly classified.

Model 4: Lastly we used naive bayes model, Naive Bayes is a straightforward method for classifying things into multiple categories. Imagine you have a pile of emails and want to sort them into spam, important, and social. Naive Bayes analyzes each email's features like sender address, keywords, and attachments. It then calculates the probability of an email being spam, important, or social based on how often those features appear in each category. The email with the highest probability wins. While it assumes features are independent (which isn't always true), Naive Bayes is often fast, accurate, and works well for large datasets, making it a popular choice for tasks like spam filtering and text classification.

Figure 8, shows the confusion matrix for this model, as we can see this model has the least performance compared to other models.

This model clocked an overall precision greater than 47% for all the classes, recall greater than 1% over all the classes,

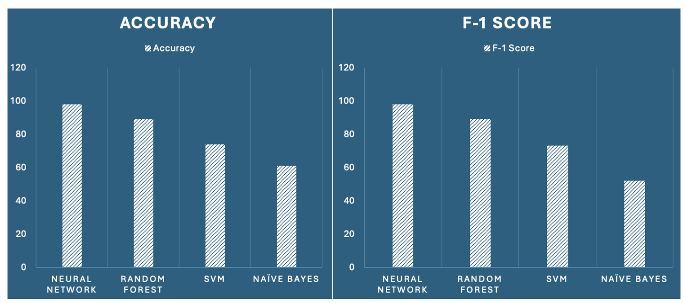

Fig. 9: Accuracy and F1 scores

this is the lowest score we have seen in the whole project, f-1 score greater than 2% over all the classes and a total accuracy of 61%.

Comparison: Let's now compare all the models and their performances to understand which model works for the application along with their limitations.

From the figures 9 and 10, our evaluation revealed a clear hierarchy of performance across the models. The deep neural network consistently achieved the highest scores in accuracy, precision, and recall. This suggests its superior ability to learn complex patterns within the MRI data and

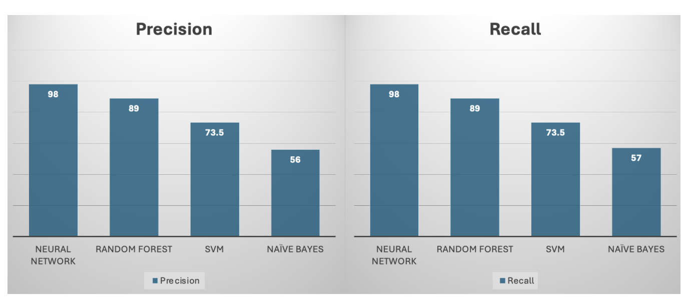

Fig. 10: Precision and Recall

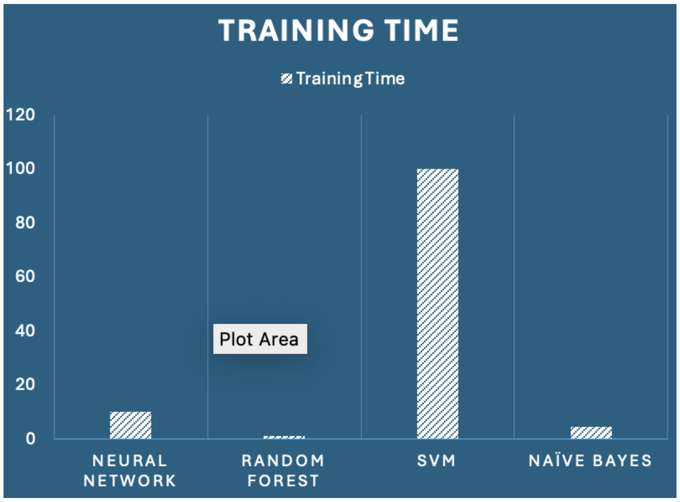

Fig. 11: Training time in minutes

make accurate classifications. Following closely behind was the random forest classifier, demonstrating a strong capacity for robust classification. The support vector machine exhibited respectable performance, while the Naive Bayes classifier displayed the least efficacy in differentiating tumor types.

However, benchmark scores alone are insufficient to determine the optimal model for our specific application. Training time also plays a crucial role.

From the figure 11, we can see that the random forest has taken the least amount of time to train the data, followed by naive bayes, neural network and SVM. The neural network performance was good because of the GPU use, if we consider the time taken by neural network to train over a CPU, we'd be seeing a number of 300, beating every other model.

5 Discussion

This study investigated the effectiveness of machine learning algorithms in classifying brain tumor types from MRI scans. Our findings demonstrated that machine learning, particularly Convolutional Neural Networks (CNNs), achieved high accuracy in differentiating between various tumor types. These results hold significant promise for improving brain tumor diagnosis and potentially revolutionizing patient care.

Our research aligns with previous studies highlighting the potential of machine learning for brain tumor classification [12] [4] [2]. We achieved accuracy rates comparable to existing literature, with CNNs reaching over 98% accuracy in distinguishing between glioma, meningioma, pituitary tumors, and no tumor. This corroborates the growing body of evidence suggesting machine learning can be a valuable tool in neuro-oncology.

While achieving high accuracy was a primary objective, our study goes beyond simply replicating existing research. By employing feature engineering and adding more features to the problem we were able to get him accuracies. We also timed the process to see the effectiveness of the models. This not only shed a light on the performance of GPUs' but also showed the limitation we faced in processing traditional machine learning techniques. It is also not worthy that the as we were using large amount of data for this application we were able to get good results, if the amount data used was reduced or unbalanced we would not be able achieve the results that we achieved.

Bias is a huge problem when it comes to creating and training such models, we proactively balanced the data points in each class to make sure our models are not affected by bias, but as we can conclude from the results obtained, Naive bayes model was classifying most of the tumors as no tumors which is a huge risk. We were also able to conclude that choosing the right model based on the application is very important to achieve the right results.

Feature engineering played a pivotal role in this project. We have the some basic generic feature extraction tools to extract the features. A potential improvement to this implementation will be to coordinate with a neuro oncologist to learn more about the tumors and create and implement more features that effectively capture the tumor areas to increase the accuracy even more. Since, a tumor is a very critical and sensitive case to deal with, even a slight inaccuracy can result huge consequences for the doctors and patients equally. The more ways we discover to improve the accuracy and the overall efficiency of the model.

The potential clinical implications of our findings in machine learning-based brain tumor classification are truly transformative. These systems have the potential to revolutionize patient care by offering several significant advantages.

Firstly, rapid and accurate diagnoses are paramount in neuro-oncology. Current diagnostic methods, such as biopsies, can be invasive and time-consuming. Machine learning models, however, could analyze MRI scans and provide preliminary classifications much faster, potentially reducing the time it takes for patients to begin treatment. This expedited diagnosis could be particularly crucial for aggressive tumor types where early intervention is critical for improving patient outcomes.

Secondly, improved diagnostic accuracy can directly translate to better treatment decisions. Traditional methods may struggle to differentiate between certain tumor types, leading to treatment plans that are not specifically tailored to the patient's unique condition. Machine learning models, with their ability to identify subtle variations in MRI scans, could provide more precise classifications. This information empowers physicians to select the most effective treatment strategies, such as surgery, radiation therapy, or targeted medications. By personalizing treatment plans based on the specific tumor type, these systems could potentially improve treatment efficacy and minimize side effects for patients.

In conclusion, the potential clinical implications of machine learning-based brain tumor classification extend far beyond simple accuracy metrics. These systems have the potential to expedite diagnoses, improve treatment decision-making, and ultimately enhance patient care by offering a faster, more precise, and less invasive approach to neuro-oncological diagnosis.

Discussing about the training times, we can conclude from figure 11, that neural network based has the highest performance with lowest time, which is ideal but it was

TABLE 1: Evaluation Metrics

Models	Precision	Recall	Accuracy	F-1 Score
Neural Network	98	98	98	98
Random Forest	89	89	89	89
SVM	73.5	73.5	74	73
Naïve Bayes	61	57	61	52

achieved with the use of GPU's. So, to conclude we recommend using deep learning model when computational power and capability of the device at hand is not an issue. But, when we are limited with the device's capability, then the go to model will be SVM which produced great evaluation metrics from figures 9, 10 and the time taken to train is also not high as we can see from the figure 11, we can also see that the time it took was around a minute. Naive bayes model also took the same time but the performance was very low and it is not recommended to use that model for tumor classification.

The use of MRI images for machine learning-based tumor classification, while promising, raises important ethical considerations. Patient privacy is paramount, and ensuring proper anonymization of MRI data throughout the research process is crucial. Additionally, potential biases within the data, such as an over representation of certain tumor types or demographics, could lead to inaccurate diagnoses for underrepresented groups. Furthermore, as these algorithms become integrated into clinical practice, transparency in their decision-making processes is essential to foster trust among healthcare professionals and patients alike. Addressing these ethical concerns will be critical for responsible development and deployment of machine learning in neuro-oncology.

This study has demonstrated the effectiveness of machine learning algorithms, particularly CNNs, in classifying brain tumor types from MRI scans. These findings hold significant promise for the future of brain tumor diagnosis and patient care. By addressing the limitations of the current study and pursuing further research in the outlined directions, we can move closer to realizing the potential of machine learning for revolutionizing neuro-oncological practice.

6 Supporting Information

Figures 9 and 10 are based on the results taken from Table 1.

7 Acknowledgments

We would like to express our sincere gratitude to Dr. Mohammed Aledhari for his invaluable guidance and support throughout this project. Dr. Aledhari's expertise and insights have significantly contributed to the success of our research on brain tumor classification using MRI images. His mentorship and encouragement have been instrumental in shaping our approach and methodology. We are truly grateful for his time, patience, and dedication in reviewing our work and providing valuable feedback. His contributions have greatly enriched the quality of our project. We extend our heartfelt thanks to Dr. Mohammed Aledhari for his unwavering support and mentorship.

8 Author Contributions

Proposed methods: CK, PM. Coding and implementation: PM, CK, SST. Analyzing and feature engineering: CK, PM, SST. Report creation : SST, CK, PM. Presentation and Video recording: PM.

9 Conclusions and Future Directions

Our primary aim is to create a machine-learning model which is capable of accurately detecting brain tumors. This is designed in such a way that it can handle large volumes of MRI scans typically processed by neurospecialists. Our first milestone was to address the failure of early detection of brain tumors. This was proved by providing a reliable automated method for brain tumor detection. Finally, comparing the accuracy and F-1 Scores obtained for the four models (as shown in Fig. 9), we achieved the highest scores for both accuracy and F-1 scores for Neural networks. However, the main drawback of using Neural networks is, that we need access to GPU. Whereas, SVM gives the best result if we have access to the CPU.

References

[1] Y. Guan, M. Aamir, Z. Rahman, A. Ali, W. A. Abro, Z. A. Dayo, M. S. Bhutta, and Z. Hu, "A framework for efficient brain tumor classification using mri images," 2021.

[2] J. Seetha and S. S. Raja, "Brain tumor classification using convolutional neural networks," *Biomedical & Pharmacology Journal*, vol. 11, no. 3, p. 1457, 2018.

[3] O. Özkaraca, O. İ. Bağrıaçık, H. Gürüler, F. Khan, J. Hussain, J. Khan, and U. e. Laila, "Multiple brain tumor classification with dense cnn architecture using brain mri images," *Life*, vol. 13, no. 2, p. 349, 2023.

[4] M. Aloraini, A. Khan, S. Aladhadh, S. Habib, M. F. Alsharekh, and M. Islam, "Combining the transformer and convolution for effective brain tumor classification using mri images," *Applied Sciences*, vol. 13, no. 6, p. 3680, 2023.

[5] S. Z. Kurdi, M. H. Ali, M. M. Jaber, T. Saba, A. Rehman, and R. Damaševičius, "Brain tumor classification using meta-heuristic optimized convolutional neural networks," *Journal of Personalized Medicine*, vol. 13, no. 2, p. 181, 2023.

[6] H. Kibriya, R. Amin, J. Kim, M. Nawaz, and R. Gantassi, "A novel approach for brain tumor classification using an ensemble of deep and hand-crafted features," *Sensors*, vol. 23, no. 10, p. 4693, 2023.

[7] A. Sarkar, M. Maniruzzaman, M. A. Alahe, M. Ahmad *et al.*, "An effective and novel approach for brain tumor classification using alexnet cnn feature extractor and multiple eminent machine learning classifiers in mris," *Journal of Sensors*, vol. 2023, 2023.

[8] G. S. Tandel, A. Tiwari, O. G. Kakde, N. Gupta, L. Saba, and J. S. Suri, "Role of ensemble deep learning for brain tumor classification in multiple magnetic resonance imaging sequence data," *Diagnostics*, vol. 13, no. 3, p. 481, 2023.

[9] C. Srinivas, N. P. KS, M. Zakariah, Y. A. Alothaibi, K. Shaukat, B. Partibane, H. Awal *et al.*, "Deep transfer learning approaches in performance analysis of brain tumor classification using mri images," *Journal of Healthcare Engineering*, vol. 2022, 2022.

[10] J. Kang, Z. Ullah, and J. Gwak, "Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, p. 2222, 2021.

[11] M. M. Badža and M. Č. Barjaktarović, "Classification of brain tumors from mri images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, p. 1999, 2020.

[12] D. R. Nayak, N. Padhy, P. K. Mallick, M. Zymbler, and S. Kumar, "Brain tumor classification using dense efficient-net," *Axioms*, vol. 11, no. 1, p. 34, 2022.

[13] M. A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S. A. C. Bukhari, "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, p. 565, 2020.

[14] F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," in *Healthcare*, vol. 9, no. 2. MDPI, 2021, p. 153.

[15] I. Abd El Kader, G. Xu, Z. Shuai, S. Saminu, I. Javaid, and I. Salim Ahmad, "Differential deep convolutional neural network model for brain tumor classification," *Brain Sciences*, vol. 11, no. 3, p. 352, 2021.

[16] A. Saboor, J. P. Li, A. Ul Haq, U. Shehzad, S. Khan, R. M. Aotaibi, and S. A. Alajlan, "Ddfc: deep learning approach for deep feature extraction and classification of brain tumors using magnetic resonance imaging in e-healthcare system," *Scientific Reports*, vol. 14, no. 1, p. 6425, 2024.

[17] kelesacademia, "braintumorclassification dataset," https://universe.roboflow.com/kelesacademia-5ptoi/braintumorclassification , sep 2023, visited on 2024-04-25. [Online]. Available: https://universe.roboflow.com/kelesacademia-5ptoi/braintumorclassification

[18] "Brain tumor mri classification: Tensorflow cnn." [Online]. Available: https://www.kaggle.com/code/jaykumar1607/brain-tumor-mri-classification-tensorflow-cnnnotebook-container