BRIAN TUMOR DETECTION USING ML

ST CLAIR COLLEGE

GROUP MEMBERS

Rechel Thomas Rebello

Priyal Hetalkumar Patel

Karan Boghara

Mahima Gupta

Navjot Kaur

PROF: UMAIR DURANI

BRIAN TUMOR DETECTION USIML

ST CLAIR COLLEGE

GROUP MEMBERS

Rechel Thomas Rebello

Priyal Hetalkumar Patel

Karan Boghara

Mahima Gupta

Navjot Kaur

PROF: UMAIR DURANI

Index

**1. INTRODUCTION …………………………………………………......3**

**2.MOTIVATION …………………………………………………......3**

**3.DATASET …………………………………………………...…3**

**4.METHODOLOGY ……………………………………………………...3**

**5.MODELLING ……………………………………………………...4**

**6.RESULTS …………………………………………………....…5**

**7. DISCUSSION ………………………………………………...…….5**

**8.DEPLOYMENT ………………………………………………….........5**

**9.CONCLUSION ……………………………………………………….6**

**10.CONTRIBUTION**

**11.LINK FOR GITHUB**

**12.REFERENCES**

. **INTRODUCTION**

Brain tumours are growths of abnormal cells in the brain that can become malignant and life-threatening due to the changes they cause to the vital structures of the brain. There are more than 120 different types of tumours that can develop in the brain, and symptoms may include headaches, seizures, difficulty thinking or speaking, changes in personality, anxiety, depression, disorientation, fatigue, abnormal eye movements, numbness or tingling on one side of the body, and more. Diagnosing a brain tumour usually involves a neurological exam, brain scans, and a biopsy if it can be done safely. Frontal lobe tumours may cause behavioural and emotional changes, impaired judgment, motivation or inhibition, impaired sense of smell or vision loss, paralysis, and more. [1][2][3]

According to the provided search results, there are over 130 different types of brain tumours. The most common types of brain tumours in adults include glioblastoma multiforme (GBM), gliomas, meningiomas, and pituitary tumours. Meningiomas are the most common primary brain tumours in adults.[4][5 ]

2.MOTIVATION

The motivation for utilizing machine learning models in brain tumour detection is to improve patient outcomes by enabling early detection and treatment. The accurate and timely diagnosis of brain tumours is crucial for effective treatment and improving patient survival rates. By utilizing machine learning models to analyze large volumes of medical images, medical professionals can detect brain tumours accurately, reducing the risk of misdiagnosis and improving patient outcomes.

***The development of a reliable and accurate brain tumour detection system using machine learning models can also improve the efficiency and cost-effectiveness of healthcare services***. Automated image analysis can reduce the time and resources required for manual image analysis, enabling medical professionals to focus on providing quality patient care.

The vast amount of data generated by medical imaging technology contains hidden patterns and correlations that can be challenging for human experts to detect. Machine learning models can process large volumes of data and extract features that indicate the presence of brain tumours accurately. By utilizing these models in conjunction with data analytics and visualization tools,

medical professionals can gain valuable insights into patient health and make informed clinical decisions.

In conclusion, the motivation for utilizing machine learning models in brain tumour detection is to improve patient outcomes, increase efficiency, and reduce healthcare costs. By leveraging the power of machine learning models and data analytics, we can transform medical imaging data into valuable insights that enable medical professionals to provide high-quality patient care.[6][7][8]

3.**DATASET**

**The provided image dataset aims to classify brain tumours into four distinct categories: Glioma, Meningioma, Pituitary, and No-tumour. The dataset contains a total of 3264 images, with each category having a different number of images. The Meningioma category consists of 937 images, Glioma has 926 images, Pituitary has 901 images, and the No-tumour category has 500 images. (**Figure 1**)below**

**It is important to note that while Swati Kanchan provided a portion of the data, the No-tumour category was sourced from Navoneel Chakrabarty's Kaggle dataset. However, the dataset was later made readily available on Kaggle.** [9**]**[10**] The data required cleaning and standardization to be ready for use in machine learning models. With the dataset now cleaned and standardized, researchers and practitioners can use it to train and test different machine learning models for the accurate classification of brain tumours into their respective categories. The availability of the dataset on Kaggle also presents an opportunity for the research community to collaborate and share findings, leading to improved diagnosis and treatment of brain tumours.**

Text

Description automatically generated

*Figure 1:distribution of class*

**4.METHODOLOGY**

**Firstly, duplicate images were identified and removed from each class, including Glioma, Meningioma, and Pituitary, as well as a fourth class for No-tumour. *Figure 2***

Chart, bar chart

Description automatically generated

*Figure 2:count distribution of tumour classes*

Next, the images were pre-processed by assigning each image a new name based on its class, and the images were stored in a new folder named 'raw'. Additionally, three new columns were added to the dataset, including Class, noise, and brisquescore, to check the quality of the data. Also discarded images having noise above 60.

Chart, scatter chart

Description automatically generated*Figure 3:Count distribution on noise an brisquescore*

To further analyze the dataset, exploratory data analysis was performed on a CSV file that was created from the image descriptions. This analysis involved examining the distribution of the data, identifying any outliers or anomalies, and visualizing the data using graphs and charts.

Chart, scatter chart

Description automatically generated

*Figure 4:distribution of height and width*

Overall, the data pre-processing step was crucial in ensuring the quality and accuracy of the dataset, and the exploratory data analysis provided valuable insights that will inform the development of machine learning models for brain tumour detection.

**5.MODELLING**

**Random Forest(Multiclass)**

Random Forest is an effective machine learning algorithm that can be used for brain tumour detection in image datasets. The images were resized to a fixed size of 100x100 pixels flattened the input array X from a 3D array (containing multiple 2D arrays of image data) into a 1D array. Chart, treemap chart

Description automatically generated

*Figure 5 contusion metrix for random forest*

The confusion matrix shows that the model is not effectively classifying instances into the correct classes and is making many incorrect predictions. This is reflected in the very low accuracy score of 0.0029895, which means that the model is correctly predicting the class of the target variable only 0.3% of the time.

**Binary classifier**

Random forest classification for n tumour and pituitary class.

Chart, treemap chart

Description automatically generated

*Figure 6 Confusion metric for no\_tumor vs pituitary\_tumor*

The model's accuracy is 0.190. The confusion matrix demonstrates that the model accurately predicted no\_tumor only once but misclassified all other no\_tumor cases as pituitary\_tumor. All pituitary\_tumor cases, on the other hand, were classed as no\_tumor.

Random forest classification for n tumour and meningioma class.

Chart

Description automatically generated

*Figure 7 confusion metrix for no\_tumor vs meninigioma\_tumor*

The model's accuracy is 0.194. The confusion matrix demonstrates that the model predicted meningioma tumour accurately just three times but misclassified all other meningioma tumour cases as no tumour. All no tumour cases, on the other hand, were classified as meningioma tumour.

Random forest classification for n tumour and glioma class.

Chart

Description automatically generated

*Figure 8 confusion metrix for no\_tumors vs glioma tumor*

The model's accuracy is 0. 206.The confusion matrix demonstrates that the model accurately predicted glioma\_tumor only once, but incorrectly predicted no tumour in all other glioma tumour occurrences. All no tumour cases, on the other hand, were classified as glioma tumour.

**CNN(Multiclass)**

The purpose of a Convolutional Neural Network (CNN) is to analyze visual imagery and recognize patterns in images to identify objects, classes, and categories. CNNs are specifically designed to learn directly from data and are particularly useful for image classification and recognition.

Images with size (250,250) was imputed in batch of 16 to train the data

70% of the images will be used for training the model, while the remaining 30% will be used for evaluating the model's performance.

Table

Description automatically generated

*Figure 9 CNN classification report*

Chart

Description automatically generated

*Figure 10.CNN classification matrix*

The report provides an evaluation of a classification model's performance for four different classes: glioma\_tumor, meningioma\_tumor, no\_tumor, and pituitary\_tumor.

Precision measures the accuracy of positive predictions made by the model, while recall measures the ability of the model to identify all the positive cases in the dataset. F1-score provides a balanced measure of both precision and recall, and accuracy measures the overall accuracy of the model.

The precision for glioma\_tumor was perfect, while the recall was low. This indicates that the model correctly identified all positive cases for glioma\_tumor but missed many actual positive cases. For meningioma\_tumor and pituitary\_tumor, the precision was good, but the recall was relatively low. The recall for no\_tumor was perfect, indicating that the model correctly identified all negative cases.

The F1-score was highest for meningioma\_tumor and no\_tumor, indicating that the model's performance was relatively balanced for these classes. For glioma\_tumor and pituitary\_tumor, the F1-score was relatively low, reflecting the low recall for these classes.

The overall accuracy of the model was moderate, with an accuracy of 61%. The macro average precision was good, but the macro average recall and F1-score were relatively low, indicating that the model's performance was affected by class imbalance.

The weighted average precision and recall were relatively good, but the weighted average F1-score was lower than the macro average F1-score. This suggests that the lower performance for glioma\_tumor and pituitary\_tumor classes had a greater impact on the overall weighted F1-score.

Chart, line chart

Description automatically generated

*Figure 11precision recal curve*

The precision-recall curve has a threshold of 0.5, which means that any point with a predicted probability more than or equal to 0.5 is positive, while any point with a forecasted probability less than 0.5 is negative.

Overall, the report suggests that the model performed well for some classes but struggled with class imbalance and identifying all positive cases for certain classes.

**Binary classification for each class**

**Glioma**

**Table

Description automatically generated**

*Figure 12classification report for no tumor vs glioma*

The precision for the glioma\_tumor class is 0.78, meaning that 78% of the samples predicted as glioma\_tumor were actually glioma\_tumo. The recall for glioma\_tumor is 0.21, meaning that only 21% of actual glioma\_tumor samples were correctly classified as glioma\_tumor. Similarly, the precision for no\_tumor is 0.56, meaning that 56% of the samples predicted as no tumor were actually "no\_tumor", and the recall for no tumor is 0.94, meaning that 94% of actual no tumor samples were correctly classified as no\_tumor. Overall, the classifier has an accuracy of 0.59, but the F1-score for both classes is relatively low, indicating that there is room for improvement in the classifier's performance.

**Pituitary**

**Table

Description automatically generated**

*Figure 13 classification report for no tumor vs pituitary tumor*

The classifier was used to classify instances in this report between "no\_tumor" and "pituitary\_tumor" and has an accuracy of 0.88. Both classes have good precision and recall, showing that the classifier is correctly classifying cases. Overall, the classifier appears to be functioning well.

**Meningioma**

**Table

Description automatically generated**

Fi*gure 14 classification report for no tumor vs meningioma*

The precision for the meningioma tumor class is 0.93, meaning that 93% of the samples predicted as meningioma\_tumor were actually meningioma tumor The recall for meningioma tumor is 0.61, meaning that only 61% of actual meningioma tumor samples were correctly classified as meningioma tumor. Similarly, the precision for no tumor is 0.69, meaning that 69% of the samples predicted as no tumor were actually no tumor, and the recall for no tumor is 0.95, meaning that 95% of actual no tumor samples were correctly classified as no\_tumor.

The classifier has an accuracy of 0.77, which means that 77% of the samples were correctly classified by the model. The F1-score for "meningioma\_tumor" is 0.74, and the F1-score for "no\_tumor" is 0.80.

.**RESULTS**

RANDOM FOREST

Multiclass

Accuracy score:**0.0029**

Pituitary vs no tumor

Accuracy score**: 0.190**

Meningioma vs no tumor

Accuracy score: **0.194**

Glioma vs no tumor

Accuracy score**: 0.206**

CNN

Multiclass

Accuracy score: **0.882**

Glioma vs no tumor

Accuracy score: **0.585**

Meningioma vs no tumor

Accuracy score:**0.772**

Pituitary vs no tumor

Accuracy score: **0.882**

**7. DISCUSSION**

After conducting a brain tumour detection analysis using two different models, it was found that both models perform well, but one has higher accuracy than the other.

The first model achieved an accuracy of 0.0029 while the second model achieved an accuracy of 0.882The second model has a higher accuracy rate, making it the best model for this analysis.

Overall, the results of this analysis suggest that the second model is the most effective tool for accurately detecting brain tumours. This information can be useful for medical professionals in making accurate diagnoses and improving patient outcomes.

**8.DEPLOYMENT**

After completing the brain tumour detection project and selecting the best model i.e CNN with accuracy of, deployment of model was done using Streamlit.

To deploy the application, a basic Streamlit script was loaded the CNN model and provided users with an interface for uploading brain scans to be analysed. later on additional functionality was added to the application, such as the ability to display the model's prediction and confidence level for each scan.

Once the application was complete, it is hosted on a amazon web server, which allowed us to make it accessible to users from anywhere with an internet connection. The application was thoroughly tested to ensure that it was functioning correctly and that it could handle many users and data inputs without crashing or slowing down.

**9.CONCLUSION**

The machine learning models developed using the Brain Tumor Classification MRI dataset have the potential to improve the accuracy and speed of brain tumor diagnosis and reduce the workload of radiologists. Further research is needed to investigate the potential of combining multiple imaging modalities, such as MRI and PET scans, for more accurate and reliable brain tumor diagnosis.

**10.CONTRIBUTION**

The Brain Tumor Classification MRI dataset was analysed by a team consisting of Karan, Rachel, Mahima, Priyal, and Navjot. Karan and Rechel conducted exploratory data analysis (EDA) on the dataset to understand the distribution of data and gain valuable insights into the features that distinguish each class. Mahima applied random forest classification algorithms to the dataset, focusing on multiclass classification, while Priyal and Navjot applied random forest classification algorithms for binary classification between tumor and no tumor. Karan contributed by developing a convolutional neural network (CNN) model that could classify all four classes, while Rechel developed individual CNN models for each class. Their combined efforts resulted in a comprehensive analysis of the dataset, utilizing different machine learning algorithms to achieve high accuracy in brain tumor classification. The team's contributions have the potential to improve the accuracy and efficiency of brain tumor diagnosis and treatment.

**11.****LINK FOR GITHUB**

<https://github.com/rechelrebello/Brain-tumor-detection>

**12.REFERENCES**

[1]<https://www.hindawi.com/journals/ijbi/2017/9749108/>

[2]<https://www.pennmedicine.org/news/news-releases/2022/december/ai-enables-the-largest-brain-tumor-study>

[3]<https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor>

[4]<https://mayfieldclinic.com/pe-braintumor.htm>

[5]<https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084>

[6]<https://www.cancerresearchuk.org/about-cancer/brain-tumours/types>

[7]<https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/about/types-of-brain-tumors.html>

[8]<https://www.pennmedicine.org/updates/blogs/neuroscience-blog/2018/november/what-are-the-most-common-types-of-brain-tumors>

[9]<https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>

[10]https://github.com/sartajbhuvaji/brain-tumor-classification-dataset