

# Development of a Novel Machine Learning Framework For Brain Tumors Using Explainable Ai And Deep Learning

**Rashi Bhave**

School of Computer Engineering & Technology, Dr  
Vishwanath Karad MIT World Peace University  
Pune, India  
rashibhave@gmail.com

**Yati Kale**

School of Computer Engineering & Technology, Dr  
Vishwanath Karad MIT World Peace University Pune,  
India  
yatikale@gmail.com

**Yuvraj Khare**

School of Computer Engineering & Technology, Dr  
Vishwanath Karad MIT World Peace University Pune,  
India  
yuvrajkhare01@gmail.com

**Dr. Sharmishta Desai**

School of Computer Engineering & Technology, Dr  
Vishwanath Karad MIT World Peace University Pune,  
India  
sharmishta.desai@mitwpu.edu.in

**Abstract**—Majorly in cancers related to Central Nervous System, 85% to 90% of all initial central nervous system (CNS) tumors are brain tumors. Tumors can be detected using multiple state-of-the-art technologies. We studied 6251 MRI scans which contained images for Glioma, Meningioma, Pituitary tumors and no tumor. We built a CNN (Convolutional Neural Networks) model which was iterated for 5 epochs, with overall test accuracy of 83.1%, a training loss of 0.44 and a validation loss of 0.41 for 1200 test images. 225 images were correctly classified as glioma and 195 as meningioma. There are 248 images for no tumor and 342 images for pituitary. 190 images were classified incorrectly. We further used the Explainable AI model LIME to understand the prediction of the CNN model.

**Keywords**—CNN, Deep Learning, Brain Tumor

## I.

### INTRODUCTION

Between 2018 and 2020, there were 40 lakh cases of cancer and 22.54 lakh people died as a result of the disease in India [1]. In 2020, was responsible for nearly 10 million deaths worldwide or nearly one in every six deaths. [2]. Majorly in cancers related to Central Nervous System, 85% to 90% of all initial central nervous system (CNS) tumors are brain tumors[3]. Tumors can be detected using multiple state-of-the-art technologies such as CT scan, MRI(Magnetic resonance imaging), PET (Positron emission tomography) Scan etc. Out of all the options, MRI images have been proven to create a more detailed picture [3].

Researchers have developed multiple Deep Learning (DL) algorithms to analyze MRI images and to predict the occurrence of diseases within the patients. Even though these algorithms are able to successfully detect the occurrence of disease, say ‘cancer’, they are not capable of explaining how an input image gave the predicted output. Neural Networks within

DL have many layers connected or intertwined, hence we are unsure of how each individual neuron interacts with the others to produce the final result. This is why DL is termed a ‘Black Box’.

As AI advances, humans are challenged to comprehend and retrace how the algorithm arrived at a result. Explainability can help developers ensure that the system is operating as intended, may be essential to fulfil regulatory standards, or may be required to provide persons impacted by a decision, the opportunity to contest or amend the decision [3]

We aim to contribute the following to the paper:

1. We want to build an explanation-driven framework with a multi-input DL model that used LIME to provide a detailed description of the results.
2. We have achieved an accuracy of (83.1%) in detecting brain tumors from MRI images

## II. METHODS

### A. Datasets

The initial dataset was obtained from Jun Cheng at the Southern Medical University in Guangzhou, China, who works in the school of biomedical engineering. The dataset was utilized by Jun Cheng in two research. The dataset includes 3064 images from 233 patients.[18][19]

The source for the second dataset is Kaggle [20][23]. The 3264 Three Channel Color (RGB Formatted, but Displayed in Grayscale) T1, T2, and Flair Weighted MRI Brain Scans that make up this dataset are mixed. These scans, like the first dataset, either display gliomas, meningiomas, or pituitary tumors, but they also include pictures of normal brains.

Both datasets contain images from the top axial and sagittal planes.

**B. Data Pre-Processing**

Keeping the bulk of the variance in the original dataset while reducing the number of characteristics in a dataset, we performed dimensionality reduction on the dataset [5][22][24]. all images were resized from 512 x 512 to a fixed size of 256 x 256 before inputting them into the DL algorithm. In order to improve prediction accuracy, we also used one hot encoding to transform categorical data into a format that DL algorithms could use. We used data augmentation to add new data points in the latent space of the original data in order to expand the dataset.

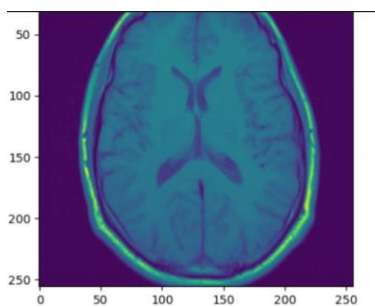


Fig. 1. Pre-processed data

**C. Different types of images in the dataset**

**Normal:** The color and intensity of the brain parenchyma are normal. The ventricular system and cisternal spaces are normal. There is no evidence of an intracranial lesion or a distinct vascular abnormalities. The visible paranasal sinuses, calvarium, and orbits appear ordinary.[7].

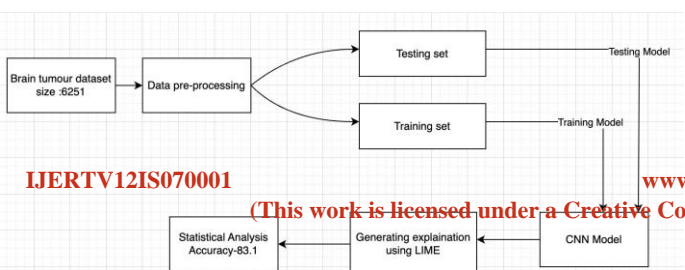
**Glioma tumor:** Gliomas have thick, erratic enhancing borders around their hemorrhagic, focally necrotic cores. A single, relatively large, irregularly shaped lesion that typically develops in the white matter serves as the tumors' typical visual representation [8].

**Meningioma tumor:** In general, imaging studies can quickly distinguish between two main macroscopic forms. The appearance of rounded, clearly defined dural masses has been compared to that of a fried egg (the most common presentation). large areas of dural thickening. [9]

**Pituitary tumor:** Granular cell tumors are uncommon tumors that develop in the pituitary stalk. They have a distinctive appearance that sets them apart from other inflamed or infiltrative stalk processes [10].

**III. PROPOSED FRAMEWORK**

The architecture of our framework uses a dataset with MRI images of brain tumors, these images are processed and split into training and testing datasets. Convolutional Neural



Network(CNN) receives the training set as input. which generated a result. The result is forwarded to the XAI algorithm LIME ( local interpretable model-agnostic explanations).

Fig.2. Architecture

The suggested DL model utilizes a dataset of 6251 MRI images. These images are then split into training and testing sets. The training dataset is fed into the multi-input CNN model, and accuracy is determined. This data is further imposed in LIME for an explanation.

In all hidden layers, the CNN is enforced with ReLU activation. The calculation of ReLU is straightforward when compared to the input value and zero value. In addition, depending on whether an input is positive or negative, ReLU has a derivative of either 0 or 1. Implementing LIME requires this ReLU functionality.

**IV. RESULTS**

Once the data is split, the training set is fed into the CNN and the accuracy of the model is evaluated along with the number of wrong predictions. The below curves are the conventional method to depict the accuracy and validation result.

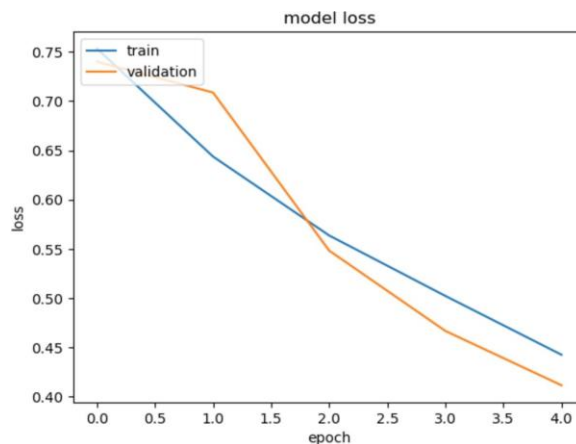


Fig.3(A). Results

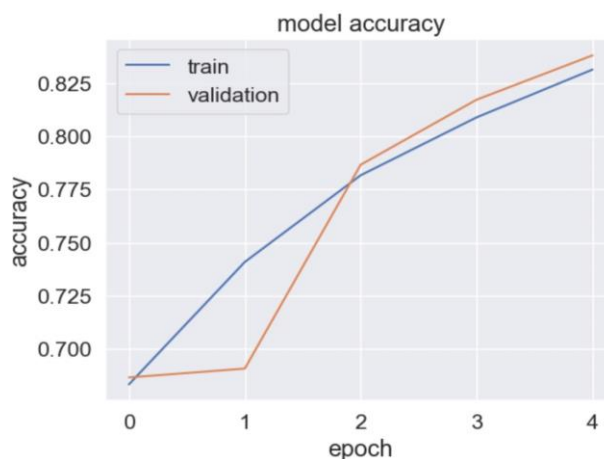


Fig.3(B). Results

The above images depict a training loss of 0.44 and an accuracy of 83.1%.

A. CNN

The model was iterated for 5 epochs, with overall test accuracy of 83.1%, a training loss of 0.44 and a validation loss of 0.41 for 1200 test images. 225 images were correctly classified as glioma and 195 as meningioma. There are 248 images for no tumour and 342 images for pituitary. 190 images were classified incorrectly.

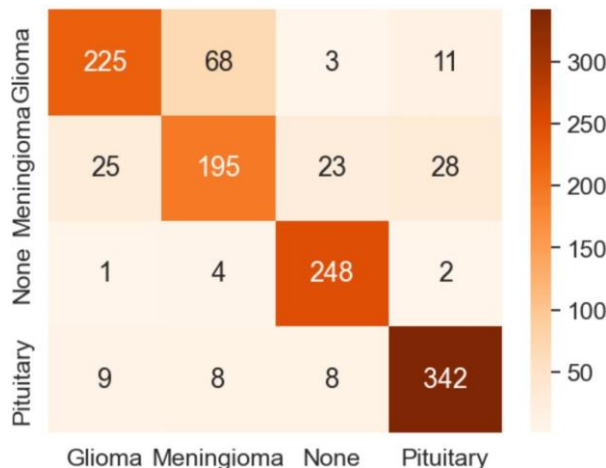


Fig.4.Confusion Matrix

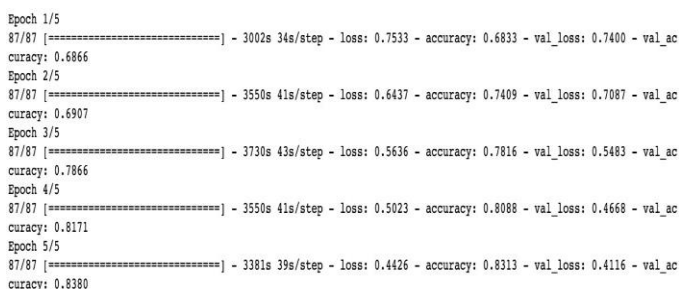


Fig.5. Epochs

B. LIME

There are 150 perturbations utilized in all. Super pixels act as columns and perturbations as rows in a matrix created from a random set of ones and zeros. Based on 0 and 1, the super pixel is either turned off or on. Figure 4 generated 80 superpixels, whereas figure 5 generated 62. The representation shown above is a binary vector showing whether a group of connected super pixels is present or not. The equation for the same is,  $X'=\{0,1\}^p$  where p' is the number of super pixels considered. This is derived by an image segmenting algorithm eg. slic. Here, the mapping function mapped 1 to let the super pixel be as it is in the original image and 0 to grey out the super pixel indicating that it is missing

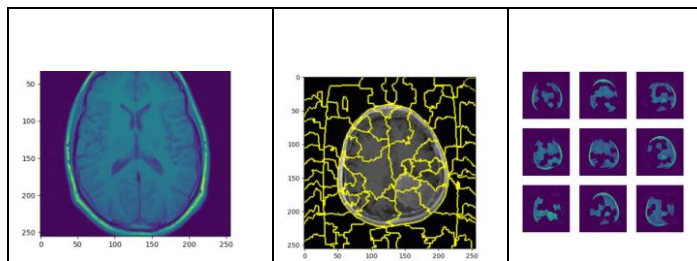


Fig.6. LIME for tumour detected

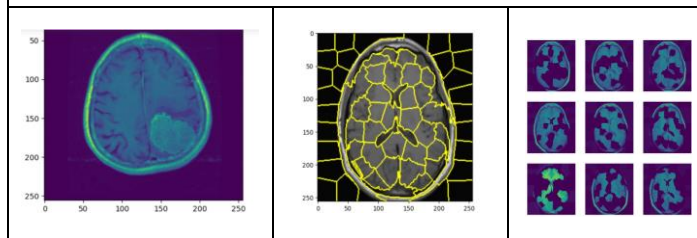


Fig.7. LIME for tumor not detected

V. DISCUSSION

A. Comparison of the Proposed Models vs Traditional Methods

In DL, for MRI images, CNN with LSTM(Long short-term memory) using VGGNet is one of the approaches used. In studies related to that, accuracy was up to 93% [11]. LSTM has been recently introduced with deep learning. Previous studies focused on machine learning methods along with LSTM. Although these models were successfully able to predict the tumor they were not able to give an explanation as to why the result was predicted. The proposed method gives a visual explanation for the predicted outcome.

The most effective used algorithm of Deep Learning for brain tumors is CNN using ResNet. When CNN models are tested on augmented data, they are proven to give an astonishing accuracy of up to 99.90% [12]. The algorithm has proven to be very effective with some limitations. When we try to reproduce the result of a given image, it may not give the same output twice. The proposed method allows us to regenerate an output.

Another library which is widely used for XAI is SHAP(Shapley Additive explanations). Even though SHAP has proven to be effective in generating XAI results, it is mainly used for explaining the functioning of the entire model rather than a single prediction. For predictions that require explanation and visual representations of a single image, LIME is very intuitive [13]

Table.1 Brain tumor detection using other state-of-the-art models

Methods	Accuracy	XAI
CNN [12]	98.9%	NO
CNN with LSTM [11]	93%	NO
CRF-RNN [14]	92.3	NO
ANN [15]	65.21%	NO
Ensemble Transfer Learning and Quantum Variational Classifier [16]	90.91	NO
Proposed Method	83.1	YES

VI. CONCLUSION AND FUTURE DIRECTION

In the proposed model we were able to achieve an accuracy of 83.1% in predicting different types of tumors in MRI images. We used 2 datasets and a total of 6251 images. The method used CNN and LIME to predict the tumor and tried to explain the prediction, thus making it reproducible. In our model, we were able to run only 5 epochs due to processing restrictions. This may have caused model underfitting. In our method, we have tried to solve the problem of the 'black box'. Currently, there are limited studies in the field of XAI in brain tumors. In future, models which give higher accuracy can be fed to XAI models to gain a better understanding.

REFERENCES

[1] India reported over 40 lakh cancer cases and 22.54 lakh deaths between 2018 to 2020, the government told Lok Sabha on Friday. <https://www.indiatoday.in/india/story/india-cancer-cases-cancer-death-3-years-govt-report-1911983-2022-02-11>

[2] Cancer is a leading cause of death worldwide <https://www.who.int/>

[3] Brain Tumor: Statistics <https://www.cancer.net/cancer-types/brain-tumor/statistics>

[4] Explainable AI <https://www.ibm.com/in-en/watson/explainable-ai>

[5] Dimensionality reduction techniques you should know in 2021 <https://towardsdatascience.com/11-dimensionality-reduction-techniques-you-should-know-in-2021-dcb9500d388b>

[6] One-hot Encoding <https://deepchecks.com/glossary/one-hot-encoding/#:~:text=One%2Dhot%20encoding%20in%20machine,categorical%20ata%20in%20machine%20learnin> g.

[7] Normal brain MRI <https://radiopaedia.org/cases/normal-brain-mri-5?lang=us>

[8] Glioblastoma Multiforme: A Review of its Epidemiology and Pathogenesis through Clinical Presentation and Treatment <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5563115/>

[9] Meningioma. <https://radiopaedia.org/articles/meningioma#:~:text=Meningiomas%20are%20extra%2Daxial%20tumors,are%20presumed%20to%20be%20located>.

[10] Primary Tumors of the Pituitary Gland: Radiologic-Pathologic Correlation <https://pubs.rsna.org/doi/10.1148/rg.2021200203>

[11] CNN-LSTM: Cascaded Framework For Brain Tumour Classification [https://www.researchgate.net/publication/330710656\\_CNN-LSTM\\_Cascaded\\_Framework\\_For\\_Brain\\_Tumour\\_Classification](https://www.researchgate.net/publication/330710656_CNN-LSTM_Cascaded_Framework_For_Brain_Tumour_Classification)

[12] DACBT: deep learning approach for classification of brain tumours using MRI data in IoT healthcare environment <https://www.nature.com/articles/s41598-022-19465-1#:~:text=In%20the%20design%20of%20the,augmentation%20and%20transfer%20learning%20methods>

[13] LIME vs. SHAP: Which is Better for Explaining Machine Learning Models? <https://towardsdatascience.com/lime-vs-shap-which-is-better-for-explaining-machine-learning-models-d68d8290bb16>

[14] Brain Tumour Detection Using Convolutional Neural Networks and Deconvolution <https://ieeexplore.ieee.org/document/9640818>

[15] Brain tumor detection from MRI images using deep learning techniques <https://iopscience.iop.org/article/10.1088/1757-899X/1055/1/012115/pdf>

[16] A New Model for Brain Tumor Detection Using Ensemble Transfer Learning and Quantum Variational Classifier <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9023211/>

[17] Explanation-Driven Deep Learning Model for Prediction of Brain Tumour Status Using MRI Image Data <https://www.frontiersin.org/articles/10.3389/fgene.2022.822666/full>

[18] Cheng J, Huang W, Cao S, Yang R, Yang W, et al. (2015) Correction: Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. PLOS ONE 10(12):e0144479.

[19] Cheng J, Yang W, Huang M, Huang W, Jiang J, et al. (2016) Retrieval of Brain Tumors by Adaptive Spatial Pooling and Fisher Vector Representation. PLOS ONE 11(6): e0157112

[20] Kaggle.com. 2021. Brain Tumor Classification (MRI) Dataset. [online] Available: <https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri/discussion/154260>

[21] figshare. 2021. brain tumor dataset. [online] Available: [https://figshare.com/articles/dataset/brain\\_tumor\\_dataset/1512427](https://figshare.com/articles/dataset/brain_tumor_dataset/1512427)

[22] Sharmishta Desai and S.T.Patil, 2018. Boosting Decision Trees for Prediction of Market Trends. Journal of Engineering and Applied Sciences, 13: 552-556

[23] DESAI, Sharmishta. Dynamic Induction Model for Student's Behavior Analysis. International Journal of System & Software Engineering.[S.I.], p. 13-19, jun. 2019. ISSN 2321-6017.

[24] Shingte, Kshitija and Chaudhari, Anuja and Patil, Aditee and Chaudhari, Anushree and Desai, Sharmishta, Chatbot Development for Educational Institute (June 6, 2021). Available at SSRN: <https://ssrn.com/abstract=3861241> or <http://dx.doi.org/10.2139/ssrn.3861241>