Abstract. The human brain, the integral part of the Central Nervous System, plays a crucial role in controlling and coordinating essential biological functions in the human body. However, its convoluted structure and the diversity of cell types make it vulnerable to various disorders, including brain tumors. Brain tumors, the abnormal proliferation of cells within brain tissue, present challenges in diagnosis due to their origin, variability in size, shape and aggressiveness. Among the most diagnosed types are gliomas, meningiomas and pituitary tumors. Early and accurate detection of these tumors is crucial for effective treatment planning. Magnetic Resonance Imaging (MRI) has emerged as widely used imaging technique for brain tumor detection and segmentation due to its high resolution and multi-modal imaging of detailed brain structures. In this paper, we implemented DenseNet121 based transfer learning model for classifying tumour classes and obtained 98.75% accuracy in the testing.

Keywords: Brain Tumor Detection, DenseNet121, Multiclass Tumour Classification, Biomedical Image Analysis

1 Introduction

The brain and spinal cord (together known as the Central Nervous System (CNS)) regulate many biological tasks such as organizing, analyzing, making decisions, giving orders, and integrating **[1]**. The human brain is a complex entity due to its intricate physical structure. Central nervous system (CNS) disorders such as stroke, infections, brain tumors, and migraines pose significant challenges in terms of diagnosis, assessment, and the development of effective treatment strategies.

In earlier days, diagnosis of brain tumors, which are nothing but abnormal expansion of brain cells, presented significant problems for neuropathologists and radiologists due to various limitations of imaging techniques and complexities of tumor characteristics.

1.1 Classification of tumors

Brain tumours are distinguished by the uncontrolled and aberrant proliferation of cells in brain tissue, which results in the disruption of normal brain function **[2]**. Brain tumours can occur at any age and may originate within the brain tissue itself, exhibiting a wide variety of shapes, sizes, and characteristics. Brain malignancies are classified into two categories: Primary tumours and secondary malignancies. Primary tumours arise within the brain tissue, whereas secondary malignancies metastasize from other regions of the body and enter the brain tissue via the circulation **[3].** Malignant brain tumours can be fatal, including the types glioma and meningioma, if identified in early stages. We can classify tumours into three major types, and they are:

Glioma tumour: Gliomas originate from glial cells. Glial cells support neurons in the central nervous system. They can vary in severity from low grades to extensive forms, including types like astrocytoma, oligodendroglioma, and glioblastoma. These

tumours can often be malignant and can impact the functions of the brain depending on their location. Glioma tumours are among the most common and deadly brain tumours.

Meningioma tumour: Meningiomas arise from the meninges, which are the protective layers of the brain and spinal cord. The growth of these tumours is slow in nature, and often these are benign in nature, but can occasionally be atypical or malignant. They may compress nearby brain tissue and result in neurological symptoms based on their size and location in the human brain.

Pituitary tumour: Pituitary tumours originate from the pituitary gland, which is located at the base of the brain. Most are benign and may be hormone-producing, meaning they function or are non-functioning. These tumours can result in multiple symptoms, either by disrupting hormone production or by creating pressure on adjacent brain structures such as the optic nerves.

Magnetic Resonance Imaging (MRI):
MRI is a diagnostic technology that uses magnetic and radio frequency fields to image the body tissues and monitor body chemistry **[4]**. It is a widely used medical imaging technique to diagnose and monitor brain abnormalities, including brain tumours. Unlike CT scans, it provides images with very high resolution with excellent soft tissue contrast, which makes it ideal for visualizing brain structures and pathological changes. Magnetic resonance imaging (MRI) brain tumour detection is a difficult and error-prone manual process **[5]**. Combining MRI images with deep learning models can make tumour detection easier. MRI often scans multiple modalities (e.g., **T1, T2, FLAIR**). These multiple modalities offer different tissue contrasts. This multimodal input enables deep learning models to differentiate tumours from healthy tissue more accurately. Also, MRI scans are 2D/3D grayscale images with very high resolution, which is why deep learning models learn spatial patterns and morphological features of brain tumours.

2 Literature Review:

Accurately identifying various types of brain tumors using Magnetic Resonance Imaging (MRI) remains a significant challenge in modern medical image analysis because of the complexity and variability of tumor characteristics. The identification of brain tumors needs an efficient image segmentation procedure to detect the exact type of the tumor. The manual method is error-prone and laborious. Computer-aided Diagnosis (CAD) enhances the precision of patient diagnosis, leading to improved early detection and ultimately enhancing the patient's quality of life **[6][7].**

Traditional techniques of brain tumor detection are primarily dependent on manual analysis of MRI scans by radiologists or classical image processing techniques such as thresholding, region growing, edge detection, and morphological operations. These processes need deep knowledge and domain expertise, as these are manual processes. Also, these were often restricted by variability in tumor appearance, size, type, and location of the tumor. These processes were deficient with noisy data or low-contrast images and were highly dependent on handcrafted features. For these reasons, the accuracy of the detection was low and also led to generalization.

As the traditional methods became ambiguous, various techniques, moreover, CAD-assisted techniques, started evolving. Machine Learning was one of the techniques. In 2017, Vani et al. employed the Support Vector Machine (SVM) method to categorize brain cancers into two categories: tumors present and tumors absent, achieving an 80% accuracy **[8]**. From Zaw's study, researchers got to know the efficiency of decision trees and naïve Bayes algorithms in the segmentation of brain tumors. The classification findings demonstrate that decision trees provide superior performance compared to naïve Bayes, with an accuracy of 94% **[9]**. Also, algorithms such as Random Forests (RF), K-Nearest Neighbours (KNN) were implemented to classify brain tumor types after extracting features from MRI images. These ML models are based on extracted features like texture, shape, or intensity-based aspects (e.g., GLCM: Gray Level Co-occurrence Matrix, LBP: Local Binary Pattern). In comparison with traditional techniques, ML models offered automation and much higher accuracy, but these models still lag behind because of the quality of extracted features and are incapable of end-to-end learning.

Afshar et al. **[10]** evaluated traditional CNN-based transfer learning methods against a novel Capsule Network (CapsNet) for tumor type classification, noting that while CNNs performed well, CapsNet better preserved spatial hierarchies, especially in scenarios with limited data augmentation. Swati et al. **[11]** used a fine-tuned Inception-v3 model to extract high-level features from MR images for classifying glioma, meningioma, and pituitary tumors, achieving over 98% accuracy with improved sensitivity and precision across classes. Sitaula and Aryal **[12]** introduced a hybrid approach where a Convolutional Neural Network (CNN) was used for deep feature extraction and a Support Vector Machine (SVM) for classification, significantly outperforming standalone CNN classifiers and demonstrating strong generalization across different brain tumor types. Mishra et al. **[13]** proposed a hybrid architecture using VGG19 for deep feature extraction and a non-linear SVM classifier, reaching up to 99.78% accuracy, thus showing the effectiveness of combining deep learning representations with classical machine learning classifiers. Basthikodi et al. **[14]** integrated Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) for handcrafted feature extraction, followed by PCA for dimensionality reduction and SVM for classification, reporting a notable 96.03% accuracy, an F1-score around 96%, and strong generalization across multiple tumor types. Most recently, Girinath et al. **[15]** developed a dual-path framework using a custom 23-layer CNN and a fine-tuned VGG16 model, which utilized transfer learning on limited MRI data, achieving 94.5% test accuracy and surpassing traditional CNN-only approaches in both binary and multi-class tumor detection tasks.

Recent developments in hybrid architectures combining Convolutional Neural Networks along Vision transformers, Swin transformers, and U-Net framework have significantly advanced brain tumour analysis tasks, particularly in MRI-based classification and segmentation. CNN-transformer hybrids like BiTr-Unet **[16]** merge CNN encoder-decoder modules with transformer blocks, effectively capturing global context and achieving outstanding Dice scores 0.93 in the BraTS 2021 segmentation challenge. Similarly, architectures such as TransUNet and TransBTS integrate transformer encoders with CNN based U-Net decoders, thereby ensuring precise tumour localization while capturing long range contextual dependencies **[17]**.

The Swin Transformer has further enhanced these approaches by efficiently modelling both local and global spatial dependencies through shifted window attention. Swin-UNETR and Swin-Unet3D architectures use this hierarchical attention mechanism, attaining better segmentation accuracy and boundary definition compared to conventional CNN approaches **[18][19]**.

In classification tasks, Vision Transformer-based hybrid models, such as RViT (Rotation-Invariant ViT), demonstrated improved accuracy of 98.6% through rotation-aware patch embedding techniques **[20]**. Hybrid models like ResViT combine CNN based MRI synthesis pretraining with ViT fine-tuning, achieving robust performance of 98.5% even on limited datasets **[21]**.

Further enhancing clinical interpretability, TransXAI employs a hybrid CNN-ViT approach integrated with explainable AI techniques, providing saliency heatmaps and transparent decision-making processes in glioma segmentation **[22]**. Additionally, advanced U-Net hybrids like MWG-UNet++ integrate transformers and GAN based augmentation, significantly improving segmentation consistency and precision **[23]**.

These hybrid approaches collectively demonstrate substantial advantages over conventional CNN methods by effectively addressing tumour heterogeneity, data scarcity, and generalization challenges in medical imaging.

3 Methodology:

In this paper, we're aiming to detect brain tumours using Magnetic Resonance Imaging (MRI) in 4 distinct classes, which are Meningioma, Glioma, Pituitary, and No Tumour. To detect an MRI image into some distinct, distinguished class, we need advanced deep learning techniques to obtain a significant confidence level from the model. In this section, we describe a step-by-step procedure to obtain the proposed model, the **DenseNet121 architecture with an optimized Callback function.** Convolutional Neural Networks are one of the prominent techniques to obtain better results for all image-related prediction tasks such as image detection, recognition, or classification. Further, reasons to select this specific method are described in detail, along with the post-processing steps to train the model.

3.1 Dataset:

The dataset is curated from the internet, such as the Kaggle dataset, the BRATS 2019 datasets, and different medical organizations' archives. The obtained image data is separated into 4 distinct labelled classes- Meningioma, Glioma, Pituitary, and No Tumour.

We apply different data handling and manipulating tasks through pandas and numpy libraries. The main reason to do such is that the data we gathered is generally cluttered or simply unorganized. We structure the unorganized image data according to the same feature perfectly and prepare the complete dataset ready for testing and training purposes.

3.2 Preprocessing of Data:

Effective preprocessing is a crucial precursor to deep learning, particularly in medical image classification, where input data can be noisy, imbalanced, or irregular in shape and resolution. In this work, we carefully prepared a publicly available MRI brain tumour dataset to classify images into four categories: Glioma Tumour, Meningioma Tumour, Pituitary Tumour, and No Tumour. Below, we describe each preprocessing stage in detail.

Label Encoding and One-Hot Encoding. The Brain MRI dataset consists of 4 classes: Glioma, Meningioma, Pituitary, and Glioma. These classes are symmetrically loaded to the model. These classes are transformed into labels mapped as discrete integers (0,1,2,3). **[24]**

$$\text{Class Mapping: } \begin{cases} \text{Glioma Tumour} \rightarrow 0 \\ \text{Meningioma Tumour} \rightarrow 1 \\ \text{Pituitary Tumour} \rightarrow 2 \\ \text{No Tumour} \rightarrow 3 \end{cases}$$

The process is as follows to convert the numerical labelling to one-hot encoding so that we can use the Categorical- Cross Entropy Function. The process can be shown as:

$$y_i = [0 \quad \cdots \quad 1 \quad \cdots \quad 0]^T, \quad y_i \in \{0,1\}^4, \quad \text{with } y_i[i] = 1$$

Dataset Balancing: In most cases, the obtained dataset is imbalanced as the classes aren't spread across the dataset equally. For our dataset, the images we collected for Glioma, Meningioma, Pituitary and No Tumour aren't equal in number, which creates differences and is less accurate than the balanced model. We have different approaches to balance the dataset as deleting the excess datasets for a particular class, creating duplicate values for the lesser attributes, or data augmentation. Here we use the imbalance learning approach to balance the dataset. To mitigate this imbalance and prevent biased learning, Random OverSampling was applied using the RandomOverSampler class from the **imblearn** library.

$$D = \{(x_i, y_i)\}_{i=1}^{N}, \quad x_i \in image\ path,\ y_i \in category$$

Image Resizing: As a requirement of the DenseNet architecture, we need to fix a constant output of shape 224×224×3. Formally, for an original image $I' \in R^{H*W*C}$, resizing transforms it to:

$$I_{RGB} = GrayToRGB(I), \qquad I_{RGB} \in R^{224*224*3}$$

This ensures uniform dimensionality and compatibility with pretrained weights initialized from ImageNet.

Pixel Normalization: The pixel intensity values for each p [0, 225] for each channel [Red, Green, Blue] are normalized in the scale of [0,1] to maintain ease of use while applying linear transformation.

$$p_{norm} = \frac{P}{255}$$

As described in **[25]**, this will accelerate the gradient descent process and avoid numerical instability. The data points get a central distribution around [0,1].

Data Augmentation: Data augmentation techniques like random rotations, flips, elastic deformations, and Gaussian noise are used to improve model generalization **[25]**. To prevent overfitting and enrich the training set, a series of data augmentation transformations were applied to each training image. Let I be the original image. The augmented image I is generated by a composite transformation:

$$\bar{I} = T_{rotate} \circ T_{zoom} \circ T_{shear} \circ T_{shift} \circ T_{flip}$$

Where,

T_{rotate}: $random\ rotation\ in\ range\ [-10^{o},\ 10^{o}]$

T_{zoom}: $random\ zoom\ with\ zoom\ factor\ z \in [0.9, 1.1]$

T_{shear}: $shearing\ with\ angle\ \theta \in [-10^{o},\ 10^{o}]$

T_{shift}: $horizontal\ and\ vertical\ shifts\ up\ to\ \pm 10\%\ of\ image\ size$

T_{flip}: $random\ horizontal\ flip\ with\ probability\ p = 0.5$

3.3 Training and Testing Split:

The total dataset is splitted into three parts as Training, Testing and Validation. 80 percent of the dataset is training set, rest 10 percent is validation and 10 percent is testing set.

Let NNN be the total number of samples. Then:

$$|D_{train}| = 0.8N, \qquad |D_{validation}| = 0.1N, \qquad |D_{test}| = 0.1N$$

3.4 DenseNet121 Model:

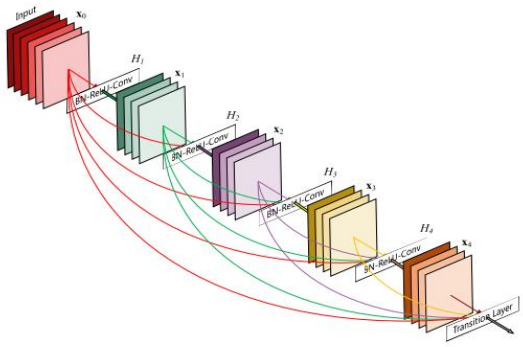

Fig. 1. DenseNet121 Overview

As described in **[26]** DenseNet is a cutting-edge DL model, known for its densely connected layers, making it efficient and powerful in classifying gliomas,

meningiomas, pituitary tumours, and non-tumour cases from MRI and by leveraging its unique feature-propagation mechanisms.

DenseNet connects each layer directly to all subsequent layers in the network. Unlike traditional CNNs, where each layer feeds into the next sequentially, DenseNet enables all layers to share information, optimizing feature extraction and gradient flow.

$$x_1 = H_1([x_0, x_1 \ldots\ldots x_n])$$

where, x1 represents the output of layer l; Hl represents the composite functions, i.e., Batch Normalization (BN), ReLU activation, and convolution; and [x0, x1.....xn] is the concatenation of all preceding layer outputs. This connectivity reduces redundancy, ensures efficient parameter usage, and improves gradient flow, leading to better classification performance. **[27]**

Dense blocks: Each dense block procedure consists of batch Normalization, ReLU Activation, and convolution with a 1X1 filter. This is followed by another operation of ReLU activation and convolution with a 3 x 3 filter, where the matrix values are interconnected. Otherwise, the matrix values should be merged. The purpose of this procedure is to act as a bottleneck, effectively reducing the number of parameters and computations in the model. The value is multiplied by 24, and the dense 4 block undergoes a convolution process that is multiplied by 16.**[26].** This model captures the tumour-specific attributes like shape and texture well and reuses in the next layers due to the connected dense networks that travel from each layer.

Transition layers*:* The transition layer, depicted in Figure 5(B) is situated between the two compact blocks. During the convolution and pooling process, the feature size undergoes alterations. The transition layer consists of convolution operations using 2 x filters and 1 x 1 average pooling with strides of 2. The DenseNet121 design consists of three transition layers, specifically transition layer 1, transition layer 2, and transition layer 3**[26]**

GAP layers: Include a GlobalAveragePooling2D layer and a Dropout layer. Calculating the mean value of each feature map and combining them into a single tensor1D is the aim of this layer. The SoftMax activation function should then be added as the DenseNet121 model's last layer.

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{c} e^{z_j}}$$

where pi is the probability of the input belonging to class i; zi represents the logit score for class i; and C indicates a number of classes (gliomas, meningiomas, pituitary tumors, and non-tumors). Values that are close to 0, 1, and 2 are calculated by this layer. Meningioma is classified as having a value of 1. The type of Pituitary tumour is identified by choice 2, while the type of Glioma is identified by option 0. The rest of the others are classified as 3 or No Tumour.

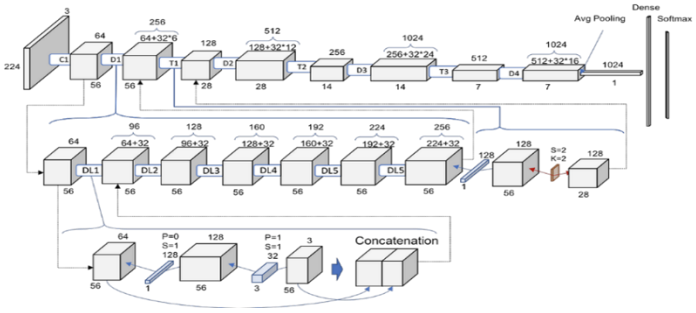

Fig. 2. DenseNet121 architecture

3.5 Weights and Loss function:

For weights, we use ImageNet **[28]** as pretrained weights. ImageNet is a database consisting large-scale ontology of images which is built upon the backbone of the WordNet structure. ImageNet aims to populate the majority of WordNet's 80,000 synsets with an average of 500-1000 clean and full-resolution images. This will result in tens of millions of annotated images organized by the semantic hierarchy of WordNet. The first prescribed model of ImageNet had 12 subtrees with 5247 synsets and 3.2 million images in total.

Fig. 3. ImageNet

For the loss function, we use the 'sparse_categorical_crossentropy' loss function defined as,

$$H(p, q) = -\sum_{x \in classes} p(x) \log q(x)$$

True probability distribution (one-shot)

Your model's predicted probability distribution

Categorical cross-entropy measures how well the predicted probabilities of each class align with the actual target labels. Its primary purpose is to evaluate a classification model's performance by comparing the model's predicted probabilities for each class with the actual class labels. The representation of the correct class is 1 and the incorrect one is 0.

3.6 Early Stopping and Save The Best Model:

For the early stopping method, we've used the loss obtained from the Validation Set as a parameter. Let's say we've obtained a certain result over some epoch, and the validation loss is being decreased through the next 5 epochs, the training will be stopped.

'Save the best model' as a model training step refers to the fact that, for each epoch, we verify if the validation accuracy obtained is better than the previous epoch or not. If it is not better than the previous epoch, we simply drop the epoch and save the previous model as the best model. If the validation accuracy is increased, the current epoch will be saved as the best model, and the next epoch will be started from the current epoch.

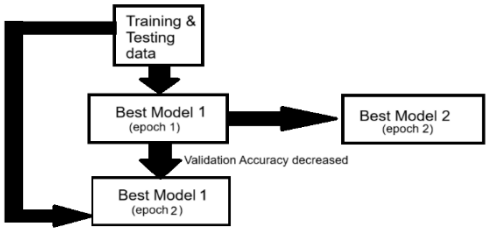

Fig. 4. MODEL WORKFLOW

4 Results

This section refers to the results and discussion portion of the actions performed on a brain MRI results image dataset. We've performed all the procedures on VSCode software that runs on Windows 11.

First, we've tested the dataset over 20 epochs with 3 input channels of RGB. The metrics for the performance analysis of the model are as follows,

1. Accuracy = (TRP + TRN) / TRP + FAP + TRN + FN) * 100
2. Precision = (TRP) / (TRP +FAP) *100
3. Sensitivity= TRPR = TRP / (TRP + FN)
4. Specificity= SPC = TRN / (FAP + TRN)
5. Precision= PPV = TRP / (TRP + FAP)
6. Negative Predictive Value = NPV = TRN / (TRN + FN)
7. False Positive Rate = FAPR = FAP / (FAP + TRN)
8. False Discovery Rate= FDR = FAP / (FAP + TRP)
9. False Negative Rate = FNR = FN / (FN + TRP)
10. Accuracy = ACC = (TRP + TRN) / (P + N)
11. F1 Score = F1 = 2TRP / (2TRP + FAP + FN)
12. Matthews Correlation Coefficient = TRP*TRN - FAP*FN / sqrt((TRP+FAP) *(TRP+F N) * (TRN+FAP) *(TRN+FN)

The results obtained from the training is structured in the following table:

Parameters	Value
Accuracy	0.9857
Specificity	0.9952
Precision	0.9857
Sensitivity (Recall)	0.9857
F1 Score	0.9857
Positive Predictive Value (PPV)	0.9857
Positive Predictive Value (PPV)	0.9857
True Discovery Rate (TDR)	0.9857
Negative Predictive Value (NPV)	0.9952
False Positive Rate (FPR)	0.0048
False Negative Rate (FNR)	0.0143
Matthews Correlation Coefficient (MCC)	0.9940

The confusion matrix obtained from the test results shown below:

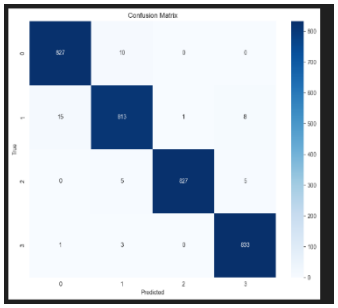

Fig. 5. Confusion matrix

For training and validation phase we plot graphs to estimate the model correctness shown below.

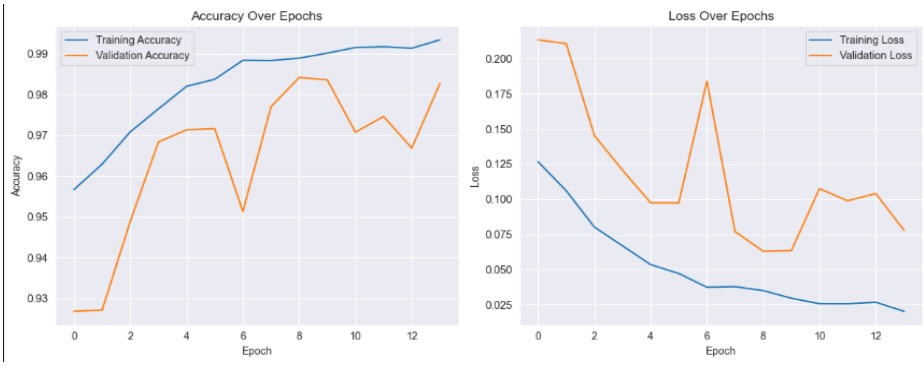

Fig. 6. Accuracy vs loss in training & validation

Fig. 7. Accuracy vs loss in training & validation

5 Conclusion and Discussion

From the training phase we acquire an accuracy of 98.57% with a loss of 0.0202 and same for testing phase we obtain an accuracy of 98.75% with a loss of 0.0182. This procedure that we described has a very significant that can be easily implemented in any hospitals and health surveys to properly detect brain tumour with the provided image. Accurate detection of disease with automation is a sound field of study nowadays. The critical performance of the model can easily be implemented in the next research experiments to implement over a large scale of image data, correctly, effectively and quickly.

References

1. Lee, D. Y. (2015). Roles of mTOR Signaling in Brain Development. Experimental Neurobiology, 24(3), 177–185. https://doi.org/10.5607/en.2015.24.3.177
2. R. Azzarelli, B. D. Simons, and A. Philpott, "The developmental origin of brain tumours: A cellular and molecular framework," Development (Cambridge), vol. 145, no. 10. Company of Biologists Ltd, May 01, 2018. doi: 10.1242/dev.162693.
3. 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, pp. 1–21, Mar. 2021, doi: 10.3390/s21062222.
4. Kasban, Hany. (2015). A Comparative Study of Medical Imaging Techniques. International Journal of Information Science and Intelligent Systems. 4. 37-58.
5. Abdusalomov AB, Mukhiddinov M, Whangbo TK. Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. Cancers (Basel). 2023 Aug 18;15(16):4172. doi: 10.3390/cancers15164172. PMID: 37627200; PMCID: PMC10453020.
6. Iqbal S, Khan MUG, Saba T, Rehman A. Computer-assisted brain tumor type discrimination using magnetic resonance imaging features. Biomed Eng Lett. 2017 Oct 4;8(1):5-28. doi: 10.1007/s13534-017-0050-3.
7. R. Vankdothu and M. A. Hameed, "Brain tumor MRI images identification and classification based on the recurrent convolutional neural network," Measurement: Sensors, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100412.
8. N. Vani, A. Sowmya, and N. Jayamma, "Brain Tumor Classification using Support Vector Machine," International Research Journal of Engineering and Technology, vol. 4, no. 07, pp. 1724 1729, 2017, [Online]. Available: www.irjet.net
9. H. T. Zaw, N. Maneerat, and Y. K. Win, Brain tumor detection based on Naïve Bayes Classification, vol. 19. 2019. 2021, doi: 10.3390/s21062222.
10. P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain tumor type classification via capsule networks," *arXiv preprint arXiv:2004.05778*, Apr. 2020. [Online]. Available: https://arxiv.org/abs/2004.05778
11. Z. N. K. Swati, R. K. S. Yadav, Y. Kwon, B. S. Chong, J. S. Lee, and T. S. Kim, "Brain tumor classification for MR images using transfer learning and fine-tuning," Computerized

Medical Imaging and Graphics, vol. 75, pp. 34–46, Jan. 2020. doi: 10. compmedimag.2019.05.0011016/j.
12. Sitaula, C., & Aryal, S. (2020). A novel brain MRI classification using convolutional neural network and support vector machine. *Multimedia Tools and Applications*, 79, 18269–18294. https://doi.org/10.1007/s11042-019-08319-7
13. V. Mishra, A. Verma, and A. Varma, "A hybrid CNN and SVM approach for brain tumor classification," in *Recent Trends in Communication and Intelligent Systems*, A. Abraham, B. Kaushik, A. Dutta, and A. Choudhury, Eds. Singapore: Springer, 2023, pp. 1–10. doi: 10.1007/978-3-031-43145-6_1
14. M. Basthikodi, M. Chaithrashree, B. M. Ahamed Shafeeq, and A. Prabhu Gurpur "Enhancing multiclass brain tumor diagnosis using SVM and innovative feature extraction techniques," Scientific Reports, vol. 14, no. 1, p. 26023, Oct. 2024, doi: 10.1038/s41598-024-77243-7.
15. Girinath N., Praveen Hari G. S., Dinesh Kumar J. R., Varun, Ganesh Babu C., and Anushree, "An Experimental Analysis of Brain Tumor Detection Using Contemporize CNN Models," *2025 International Conference on Smart Technologies, Communication & Robotics (STCR)*, Sathyamangalam, India, May 2025, pp. 1–8. doi: 10.1109/STCR62650.2025.11019098.
16. Jia H, Cai W, Huang H, Xia Y. BiTr-Unet: a CNN-Transformer combined network for MRI brain tumor segmentation. arXiv preprint arXiv:2109.12271. 2021.
17. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, et al. TransUNet: Transformers make strong encoders for medical image segmentation. Nature Scientific Reports. 2021;11(1). doi:10.1038/s41598-021-85807-4.
18. Hatamizadeh A, Tang Y, Nath V, Yang D, Roth HR, Xu D. Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images. arXiv preprint arXiv:2201.01266. 2022.
19. Wang L, Wang Z, Jin Z, Wu Q, He K, Yang X, et al. Swin-Unet3D: A hybrid Swin transformer U-Net architecture for volumetric medical image segmentation. Frontiers in Neuroscience. 2023;17. doi:10.3389/fnins.2023.1145867.
20. Das S, Banerjee A, Ghosh S, Chakraborty A. RViT: Rotation-invariant vision transformer for brain tumor classification. Frontiers in Neuroinformatics. 2024;18. doi:10.3389/fninf.2024.1414925.
21. Liu X, Yang H, Zhu Y, Zhou Z. ResViT: CNN-based MRI synthesis pretraining with vision transformer fine-tuning for brain tumor classification. arXiv preprint arXiv:2411.12874. 2024.
22. Karthik M, Gupta R, Kumar S, Singh P. TransXAI: Transformer-based hybrid CNN with explainable AI for glioma segmentation. Nature Scientific Reports. 2024;14. doi:10.1038/s41598-024-54186-7.
23. Zhao H, Wang Z, Liu J, Li Y, Shen D. MWG-UNet++: Transformer-enhanced U-Net with GAN augmentation for MRI brain tumor segmentation. Bioengineering. 2025;12(2):140. doi:10.3390/bioengineering12020140.
24. Shah, D., Xue, Z. Y., & Aamodt, T. M. (2022). Label Encoding for Regression Networks. International Conference on Learning Representations (ICLR 2022), Spotlight Presentation. Available as arXiv preprint arXiv:2212.01927
25. H. Kattamanchi, J. S. P, S. R, A. B. K. R, B. A and C. Jagadeesh, "Deep Learning with Explainable AI for Brain Tumor Diagnosis from Multimodal Imaging," 2025 Third International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), Trichy, India, 2025, pp. 1-5, doi: 10.1109/ICAISS61471.2025.11042055.

26. Karthik M, Gupta R, Kumar S, Singh P. TransXAI: Transformer-based hybrid CNN with explainable AI for glioma segmentation. Nature Scientific Reports. 2024;14. doi:10.1038/s41598-024-54186-7.
27. S. Fakheri, M. Yamaghani, and A. Nourbakhsh, "A DenseNet-Based Deep Learning Framework for Automated Brain Tumor Classification," Healthcraft. Front., vol. 2, no. 4, pp. 188-202, 2024. https://doi.org/10.56578/hf020402
28. J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848