Grenze International Journal of Engineering and Technology, January Issue

A Vision Transformer Deep Learning Model with CLAHE Technique of Image Enhancement for Brain Tumor Classification

Subhash Chand Gupta[1], Shripal Vijayvargiya[2] and Vandana Bhattacharjee[3]
[1-3]Department of CSE, Birla Institute of Technology, Mesra, Ranchi, India
Email: sgupta@bitmesra.ac.in, svijayvargiya@bitmesra.ac.in, vbhattacharya@bitmesra.ac.in

***Abstract*— This paper explores the deep learning-based classification of Brain Tumors using MRI (Magnetic Resonance Imaging) images. Brain tumors are characterized by abnormal and fast-growing tissues in the brain leading to severe health risks and require accurate and timely diagnosis for effective treatment. MRI is non-invasive radiation free technique commonly used to capture high-resolution images of body parts. Manual interpretation of these images is slow process and dependent on radiologists' expertise. Since Deep learning in medical imaging have shown the promising results to automate the interpretation process, we have used in our work a fine-tuned Vision Transformer Model (ViT_B_16) for classification task. We have also modified its MLP head to customize classification into four categories. This model was trained on a publicly available brain MRI dataset from Kaggle [22]. In testing phase, ViT_B_16 achieved accuracy of 97.94% with overall precision, recall and F1-Score 97.91%, 97.76% and 97.79% respectively. The comparison of ViT_B_16 with CNN based models shows the performance superiority of the ViT_B_16 model. This better performance is attributed to vision transformer self-attention mechanism, which captures global image features, unlike CNNs that focus on local spatial features. This highlights the potential of transformer-based architectures for more accurate brain tumor classification.**

***Index Terms*— Brain Tumor, Deep Learning, Medical imaging, ViT_B_16 Model, Vision Transformer.**

I. INTRODUCTION

Brain tumors are abnormal growing tissues in the brain. Their growth is uncontrolled and that can be life-threatening and lead to severe health issues. These tumors affect individuals across all age groups and can be either benign or malignant, with the latter posing significant threats to life and neurological function. Early and precise diagnosis is critical for effective treatment planning. Magnetic Resonance Imaging (MRI) stands out for its non-invasive nature and its ability to produce high-resolution images of internal brain structures. Despite its advantages of invasive nature, the manual interpretation of MRI scans by radiologists is very time-consuming and may be subject to variability and diagnostic error-prone and highly dependent on the radiologist's expertise. It becomes more complicated particularly given the heterogeneous appearance of tumors. To address these challenges, deep learning has emerged as a transformative approach in medical image analysis.

Grenze ID: 01.GIJET.12.1.78

Deep learning models, especially Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs), have shown great promise in automating the detection and classification of brain tumors in MRI scans. These models can learn complex visual patterns directly from raw MRI image data, bypassing the aid of manual feature engineering. Their ability to generalize across different tumor types enables them to distinguish among common categories such as gliomas, meningiomas, and pituitary tumors with high accuracy. These three types of brain tumors can be explained as:

Glioma: A type of tumor that arises from glial cells, which support and protect neurons in the brain. Gliomas are often malignant and aggressive, making them one of the most dangerous brain tumor types.

Meningioma: This tumor develops in the meninges, the protective layers surrounding the brain and spinal cord. Most meningiomas are benign, but their growth can compress nearby brain tissue, leading to neurological symptoms.

Pituitary Tumor: These tumors form in the pituitary gland at the base of the brain. They are typically benign and slow-growing but can affect hormone production and cause vision problems if they enlarge.

Motivation behind the proposed work:

To address the challenges associated with the manual interpretation of MRI scans by radiologists e.g. time-consumption, subjective variability and diagnostic error-prone, researchers have been motivated to automate the MRI images analysis task by leveraging the approach of deep learning-based models. Many researchers have applied the transfer learning using pretrained CNN models. In our work we have proposed a fine-tuned vision transformer (ViT_B_16) based model to increase the performance of the classification task.

Key contributions of the proposed work:

- Applying an image enhancement technique CLAHE (contrast Limited Adaptive Histogram equalization) on the original images of Brain MRI to make tumor regions more distinguishable.
- Modification in classifier head and Fine-tuning of baseline ViT_B_16 vision transformer model to make it capable for brain tumor classification.

II. Literature Review

Vision Transformers (ViTs) based deep learning models have shown a significant potential in improving the efficiency of image classification task. Such, as brain tumor classification, segmentation, and in other areas like remote sensing etc. ViTs are capable to capture long-range dependencies and modelling global context, which are crucial for high-resolution image analysis, while traditional CNN models are lacking to capture the global context.

Dosovitskiy et al. [1] pioneering the application of transformers to vision tasks. Based on this, many research works have explored ViT models for brain tumor classification. Khaniki et al. [2] introduced a Vision Transformer with Selective Cross-Attention and Feature Calibration, which improved the classification accuracy on MRI images. Martín et. al. [3] evaluated multimodal ViTs across many body organs like brain, lung, and kidney tumor data and confirmed robustness of ViT models in medical imaging.

Afshar et al. [4] presented a ViT-iResNet hybrid model, and combining local and global features successfully. Similarly, Tummala et al. [5] combined CNN and Transformer layers, and enhancing MRI-based tumor classification task accuracy. Jia et al. [6] developed BiTr-Unet, combining CNN and Transformer structures for improved tumor segmentation. Swin UNETR by Hatamizadeh et al. [7] utilized Swin Transformers, achieving high performance in semantic segmentation tasks. Bazi et al. [8] utilized ViTs for remote sensing classification, illustrating their flexibility across domains. Khan et al. [9] introduced NEUROSCAN, a transformer-based pipeline for brain tumor detection. Zhang et al. [10] presented GT-Net, a global transformer network effective in multiclass brain tumor classification task. Peiris et al. [11] proposed a Hybrid Window Attention Transformer, demonstrating fine-grained segmentation capabilities. Beyond brain imaging, Zhang et al. [12] applied ViTs for bamboo resource identification, showcasing cross-domain potential. Jiang et al. [13] designed LeMeViT with learnable meta tokens for efficient remote sensing analysis. Wang et al. [14] and Roy et al. [15] applied ViTs to SAR and multimodal fusion in remote sensing, respectively. Dhakshnamurthy et al. [16] used pretrained models for accurate tumor detection. De Benedictis et al. [17] applied topological data analysis with tensor decomposition for enhanced MRI classification.

Thapa et al. [18] applied CLAHE for enhancing and improving retinal image classification. CNN-based methods, like Gupta et al. [19], applied CNN based model in brain tumor classification tasks. Younis et al. [20] combined VGG-16 and ensemble techniques for robust classification. Finally, Nassar et al. [21] proposed a hybrid model by combining multiple deep learning components, obtained high classification accuracy on MRI scans.

III. MRI IMAGE DATASET DESCRIPTION AND PRE-PROCESSING

This section discusses the brain tumor dataset which has been used in this work, and its preprocessing.

A. Image Dataset description

The dataset used in this study was obtained from a publicly available source on Kaggle.com [22]. It consists of MRI brain scan images categorized into two main directories: Training and Testing. Each of these directories contains four subfolders corresponding to the respective tumor classes: No Tumor, Glioma, Meningioma, and Pituitary Tumor. This organized structure facilitates supervised learning by providing clearly labeled data for both the training and the evaluation phases of the deep learning models.

TABLE I: TABULAR REPRESENTATION OF THE IMAGES COUNTS OF DIFFERENT TUMORS CLASSES

SN.	Type of Brain Tumor	# Images for Training	# Images for Testing
1	No Tumor	1595	405
2	Glioma Tumor	1321	300
3	Meningioma Tumor	1339	306
4	Pituitary Tumor	1457	300

B. Pre processing

Before feeding MRI images into classification model, following pre-processing steps have been performed to improve image quality and model performance. These steps are:

- Image Enhancement
- Image Resizing
- Image Normalization

Image Enhancement

Image enhancement plays a vital role in improving the visibility and clarity of critical regions, such as tumor areas, in MRI scans. The primary objectives of contrast enhancement are - to make tumor regions more distinguishable for better visual interpretation and to improve the accuracy and efficiency of subsequent tasks such as classification, segmentation, and object detection. Contrast Limited Adaptive Histogram Equalization (CLAHE) technique has been applied for image enhancement. Unlike global methods, CLAHE enhances contrast locally by operating on small regions of the image, called tiles. It effectively improves the visibility of important features such as tumor boundaries and tissue textures, while limiting the amplification of noise in uniform areas. Additionally, bilinear interpolation is used to blend adjacent tiles, eliminating artificial edges or boundaries between them. This localized enhancement approach helps retain essential details in medical images, thereby supporting more accurate classification results. The figure 1 and figure 2 show the original images and enhanced images respectively after applying the image enhancement technique.

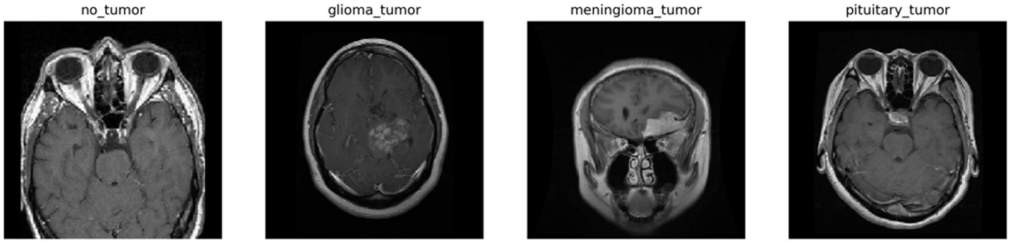

Figure 1. Original images of various type of brain tumors

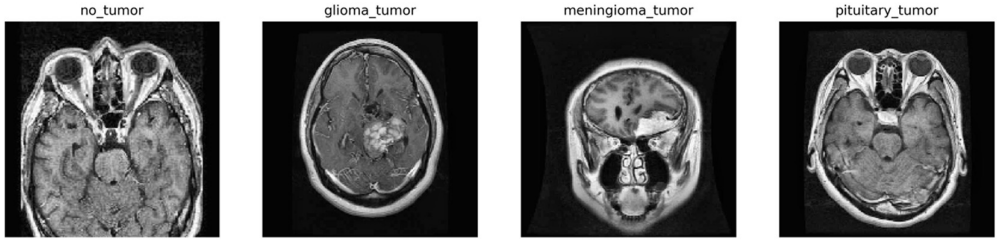

Figure 2. Enhanced Images of brain tumor after applying the CLAHE image processing technique.

Image Resizing

The deep learning models used in this paper is ViT_B_16 (Vision Transformer) pretrained on the ImageNet dataset, which consists of images standardized to a size of 224×224×3. As such, resizing all input MRI images to this dimension is essential to ensure compatibility with these architectures. This model divides the input image into non-overlapping patches of 16×16 pixels. For a 224×224 image, this results in 14×14 = 196 patches, which aligns with the pretrained model's positional embeddings. Deviating from this input size changes the number of patches, causing incompatibility with the learned positional encodings. Therefore, resizing to 224×224 is necessary for proper patch embedding and optimal performance.

Image Normalization

Normalization is a key pre-processing step that adjusts the pixel values of input images to a consistent scale. This standardization improves both the training stability and convergence of deep learning models. In this paper, image normalization is performed using the standard mean [0.485, 0.456, 0.406] and standard deviation [0.229, 0.224, 0.225]. These values are based on the ImageNet dataset statistics.

IV. METHODOLOGY

This section outlines the architectural design of the deep learning model used for image classification, along with the setup for training, validation, and testing. A complete framework of the methodology is shown in the figure 3.

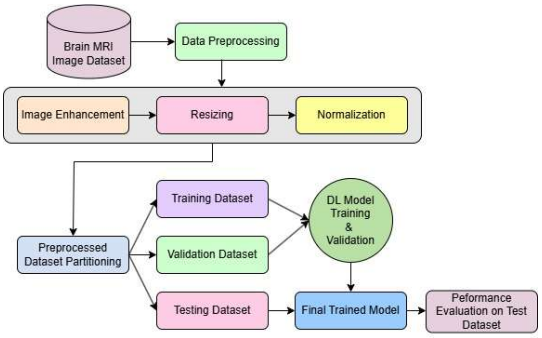

Figure 3: Block diagram of the complete methodology of the work.

A. Architectural Design of ViT_B_16 – Vision Transformer Model:

ViT_B_16 (Vision Transformer Base with 16×16 patch size) a deep learning model that processes images as sequences of patches and leverages self-attention mechanisms to capture global contextual information. Originally, introduced by Google in the paper "An Image is Worth 16x16 Words: Vision Transformers for Image Recognition at Scale" Dosovitskiy et al. [1]. The ViT_B_16 variant refers to the base model (B) with a patch size of 16×16 pixels. ViT uses self-attention mechanisms to process images. It is pretrained model, but in our work, we have fine-tuned the pretrained model with unfreezing the last three encoder layers and modified the MLP classifier head to accommodate only four categories. The figure 4 shows the architectural diagram of this model.

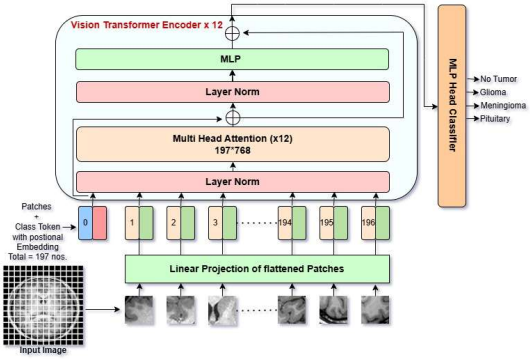

Figure 4. Architectural design of VIT_B_16 Vision Transformer with MLP Head classifier.

The following are the steps involved in the working of Vision transformer:

- Dividing an input image (typically 224×224 pixels) into 196 non-overlapping patches of size 16×16. These Patches are treated as tokens which are similar to words in NLP task.
- Flattening of every patch into a 768-dimensional vector and passing through a learnable linear projection layer.
- Adding of a special classification token called as Class Token (CLS) to the sequence of patches.
- Using the positional embeddings step to retain the spatial positions of patches.
- Processing of the resulting sequence of 197 tokens (one class token and 196 patches tokens) by a stack of 12 transformer encoder layers as shown in the figure 4.

Each encoder layer comprises multi-head self-attention (with 12 heads), layer normalization, and a feed-forward network employing ReLU activation. To preserve gradients during training, residual connections are applied after each attention and feed-forward block. Final encoder layer provides the representing class token as an output which is passed through an MLP head for classification.

Key Features of ViT_B_16 Model:

- Patch Size – 16×16 Pixels: The input image is divided into fixed-size patches of 16×16 pixels, each treated as a separate input token.
- Base Configuration (B): The ViT_B_16 follows the base model architecture, consisting of 12 transformer encoder layers for feature extraction and representation learning.
- Embedding Dimension – 768: Each image patch is flattened and linearly projected into a 768-dimensional embedding vector, corresponding to the total number of pixel values in a patch (16×16×3 for RGB).
- Multi-Head Self-Attention – 12 Heads: The model employs 12 attention heads to simultaneously capture diverse relationships and contextual information across all image patches.
- MLP Head – A Fully connected Neural network with Linear, ReLU and Dropout Layer.

B. Fine-Tuning and Modification of the ViT_B_16 Vision Transformer Model

In this paper, we have used ViT_B_16 (a Vision Transformer)—for training, validation, and testing on our MRI brain tumor image dataset. Since it is pretrained model trained on Imagenet dataset. To effectively reuse this model to our specific classification task, we have applied fine-tuning and architectural modifications. These adjustments enable the ViT_B_16 model to specialize in our medical image classification while retaining the benefits of its pretrained features.

To adapt the pretrained ViT_B_16 model for the classification of brain tumor MRI images, we implemented the following modifications and fine-tuning strategies:

Fine-Tuning the Transformer Encoder:

While the initial layers of the Vision Transformer capture generic visual features, domain-specific learning is essential for medical imaging tasks. Hence, we have unfrozen the last 3 encoder blocks out of the 12 encoder blocks of the base ViT_B_16 model, allowing the model to fine-tune its internal representations based on the characteristics of brain MRI images.

Modification of the Classification Head:

The original ViT_B_16 model is designed to classify images into 1,000 ImageNet categories. To align it with our task, we have replaced the final classification head with a new fully connected layer tailored for four categories. Specifically, the modified classifier includes the following layers:

- A Linear (fully connected) layer mapping 768 feature vector to 64 feature vector
- A ReLU activation function to introduce non-linearity.
- A Dropout layer with a dropout rate of 0.5 to mitigate overfitting.
- A final Linear layer that maps the 64 features vector to 4 output values corresponding to the four brain tumor classes.

These modifications allow the ViT_B_16 model to leverage its strong global context modelling capabilities while adapting to the unique patterns found in MRI scans for multi-class tumor classification.

C. Selection of hyperparameters, loss function and optimization function

In our experimentation we have used the following hyperparameters and utility functions to train our DL models.

- Learning rate (α) = 0.0005.
- Number of epochs = 30 Nos.
- Batch Size (N) = 64
- Loss function = crossEntropyLoss() function from pytorch for multi class classification

- Optimization function: adam() function from pytorch for minimizing the loss function.

Besides this, we have partitioned the training dataset in 80:20 ratio. 80% for training and 20% for validation of the model.

D. Hardware Environment Setup

Deep learning models based on Vision Transformers for complex computer vision tasks like image classification requires substantial computational power and memory. Standard CPU-based machines are typically inadequate for such resource-intensive operations. To meet these requirements, we have utilized the Google Colab Pro cloud platform, which offers access to high-performance hardware accelerators. Colab Pro provides a variety of options, including A100 GPUs, L4 GPUs, T4 GPUs, V5e-1 TPUs, and V2-8 TPUs and System RAM: 83 GB and GPU RAM: 40 GB.

V. Experimental Results of Model Performance

This section presents the experimental evaluation of the ViT_B_16 Vision Transformer—on the MRI brain tumor classification task. The assessment includes quantitative performance metrics, training and validation curves, as well as detailed classification reports and confusion matrices for testing phase.

A. Description of Performance Metrics

To evaluate the effectiveness of the proposed deep learning model for brain tumor classification, we have used widely accepted performance metrics: Accuracy, Precision, Recall, and the F1-Score. The evaluation metrics are computed based on the following classification outcomes for each tumor class:

Table II

TP (True Positive): The model correctly predicts a sample as belonging to the class.	FN (False Negative): The model fails to predict a sample that actually belongs to the class.
FP (False Positive): The model incorrectly predicts a sample as belonging to the class.	TN (True Positive): The model correctly predicts a sample as not belonging to the class.

The performance metrics are defined with their respective equations as follows:

- Accuracy: Number of correct predictions divided by the total number of samples. It can be calculated as:

$$\text{Accuracy} = (TN + TP) / (TP + TN + FP + FN) \quad (1)$$

- Precision: How many predicted positives are actually correct. It is calculated as:

$$\text{Precision} = TP / (TP + FP) \quad (2)$$

- Recall: How many actual positives were correctly predicted. It can be calculated as:

$$\text{Recall} = TP / (TP + FN) \quad (3)$$

- F1-Score: Weighted harmonic-mean of Precision and Recall. It is generally calculated as:

$$\text{F1-Score} = (2*\text{Precision}*\text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

B. Epoch-wise Training and Validation Curves

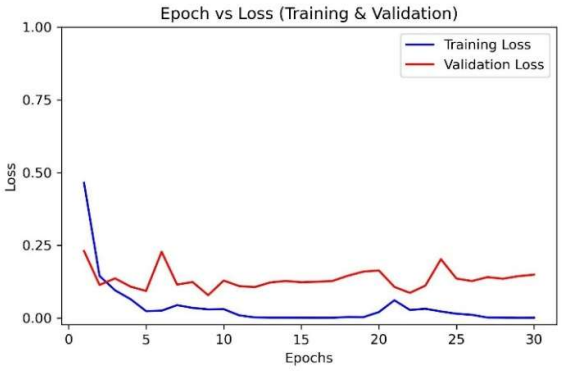

Figure 5. Plot for Epoch vs Loss

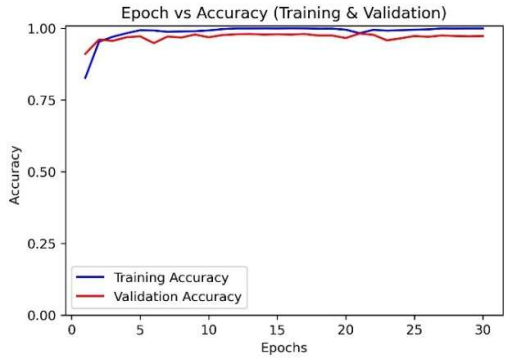

Figure 6. Plot for Epoch vs Accuracy

To observe the learning behavior and convergence of the deep learning models during training, we have generated plots of epoch-wise training and validation loss as well as training and validation accuracy. These visualizations are critical for understanding how well the models generalize to unseen data and whether they are overfitting or underfitting.

In training phase of ViT_B_16 Vision Transformer model, the training curves illustrate how the models prediction error (loss) decreases and accuracy improves over time with each epoch. A consistent decline in validation loss alongside rising accuracy indicates stable learning, while any widening gap between training and validation curves indicates the potential overfitting. Figure 5 and figure 6 show the Epoch-wise Loss and Accuracy curves respectively during Training and Validation for ViT_B_16 Model.

C. Classification Reports and Confusion Matrices and result comparison

Following the completion of 30 training epochs for ViT_B_16 Vision Transformer deep learning model, we have evaluated its performance on both the training and testing datasets. This evaluation generated detailed classification report—including precision, recall, F1-score, and overall accuracy—as well as confusion matrix that provides a class-wise breakdown of correct and incorrect predictions. The report and confusion matrix offer deeper insight into how effectively the model distinguishes between different tumor types and help identify any misclassification trends. The table II shows the classification report and figure 7 shows the confusion matrix for testing phase.

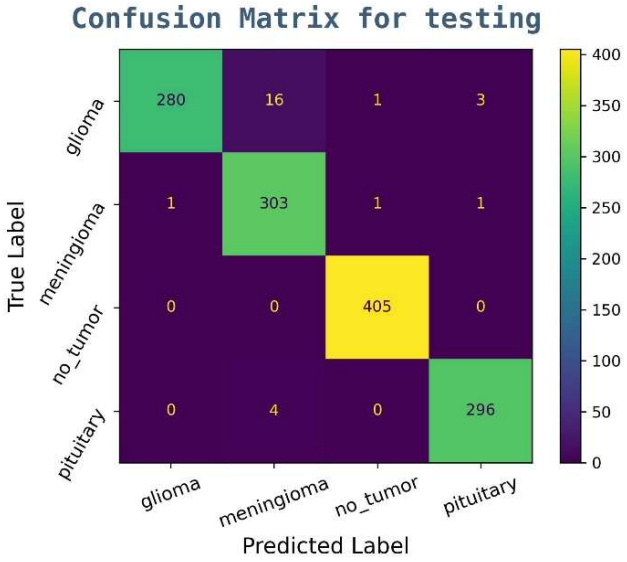

Figure 7.: Confusion matrix for testing of ViT_B_16 Transformer model

TABLE II: TESTING PERFORMANCE REPORTS FOR VIT_B_16 MODEL

S. No.	Type of Tumor	Precision (%)	Recall (%)	F1-Score (%)	Support	Accuracy (%)
1	Glioma	99.64	93.33	96.39	300	97.94 %
2	Meningioma	93.81	99.02	96.34	306	
3	No Tumor	99.51	100.00	99.75	405	
4	Pituitary	98.67	98.67	98.67	300	
5	Overall	97.91	97.76	97.79	1311	

It is observed from the table II, that Vision Transformer model for MRI brain tumor classification has achieved accuracy of 97.94% with overall precision, recall and F1-Score 97.91%, 97.76% and 97.79% respectively. In our work, results obtained from our proposed model (Fine-tuned ViT_B_16) were compared with the various CNN based models like ResNet50, VGG16, AlexNet, Custom CNN model, GoogleNet and Shufflenet etc. The comparison is shown in the table III. The comparison of ViT_B_16 with CNN based models shows the performance superiority ViT_B_16 model. This indicates that the transformer-based architecture is more effective due to self-attention mechanism in capturing complex spatial dependencies and subtle variations in tumor patterns, which are critical in medical image analysis.

TABLE III. COMPARISON OF PERFORMANCE RESULTS.

References #	DL Models	Accuracy (%)
[16]	AlexNet	95.60
[16]	VGG16	97.66
[16]	ResNet50	96.90
[19]	VGG16	96
[20]	CNN	89.5
[20]	VGG16	97.6
[20]	Ensemble Model	91.29
[21]	GoogleNet	95.16
[21]	ShuffleNet	96.84
Proposed work	Fine-tuned ViT_B_16	97.94

VI. CONCLUSION AND FUTURE SCOPE OF THE WORK

This work focused on fine-tuning with modified MLP heads ViT_B_16 Vision Transformer for MRI Scan based brain tumor classification. MRI images were pre-processed to make tumor regions more distinguishable for better visual interpretation and to improve the accuracy and efficiency of subsequent tasks such as classification, segmentation, and object detection. The ViT_B_16 model was trained for 30 epochs and evaluated using standard performance metrics. ViT_B_16 model has achieved accuracy of 97.94% with overall precision, recall and F1-Score 97.91%, 97.76% and 97.79% respectively. The comparison of ViT_B_16 with CNN based models shows the performance superiority ViT_B_16 model. This superior performance can be attributed to the ViT_B_16 model's ability to capture global contextual information via self-attention mechanisms, unlike CNNs that primarily learn local features. These results highlight the potential of transformer-based architectures in medical image analysis. However, further evaluation on larger and more diverse datasets is necessary to validate scalability. Future work may also explore more advanced architectural design to enhance the performance. Combining CNN and transformer features into hybrid models is another promising direction.

Conflicts of Interest: The authors declare no conflicts of interest.

REFERENCES

[1] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale," arXiv preprint arXiv:2010.11929, 2020.

[2] M. A. L. Khaniki, N. Tabassian, S. Vosoughi, A. Mostaar, and A. A. Shojaii, "Brain Tumor Classification Using Vision Transformer with Selective Cross Attention Mechanism and Feature Calibration," arXiv preprint arXiv:2406.17670, 2024.

[3] Ó. A. Martín and J. Sánchez, "Evaluation of Vision Transformers for Multimodal Image Classification: A Case Study on Brain, Lung, and Kidney Tumors," arXiv preprint arXiv:2502.05517, 2025.

[4] P. Afshar, M. Heidari, H. Fathi, B. Fetrat, and K. Nasiri, "Combining Local and Global Feature Extraction for Brain Tumor Classification: A Vision Transformer and iResNet Hybrid Model," Engineering Technology & Applied Science Research, vol. 13, no. 4, pp. 8854–8859, 2023.

[5] S. Tummala, C. Chimakurthi, S. Manthalkar, and S. R. Meka, "Combining the Transformer and Convolution for Effective Brain Tumor Classification Using MRI Images," Current Oncology, vol. 30, no. 3, pp. 2573–2585, 2022.

[6] Q. Jia and H. Shu, "BiTr Unet: A CNN Transformer Combined Network for MRI Brain Tumor Segmentation," arXiv preprint arXiv:2109.12271, 2021.

[7] A. Hatamizadeh, L. V. Le, F. Mahmood, D. Yang, B. Y. Jung, and N. M. Navab, "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images," arXiv preprint arXiv:2201.01266, 2022.

[8] [8] Image Classification," Remote Sensing, vol. 13, no. 3, p. 516, 2021. Available: https://doi.org/10.3390/rs13030516

[9] K. Khan, M. A. Mir, T. S. Rajput, and S. Abbas, "NEUROSCAN: Revolutionizing Brain Tumor Detection Using Vision Transformer," International Journal of Innovative Science and Technology, vol. 4, no. 2, pp. 1–10, 2024.

[10] C. Zhang, Z. Wang, X. Li, and J. Liu, "GT Net: Global Transformer Network for Multiclass Brain Tumor Classification Using MR Images," Biomedical Engineering Letters, vol. 14, pp. 29–40, 2024.

[11] H. Peiris, T. D. Nguyen, C. Rochan, and Y. W. Lui, "Hybrid Window Attention Based Transformer Architecture for Brain Tumor Segmentation," arXiv preprint arXiv:2209.07704, 2022.

[12] J. Zhang, X. Zhang, and Y. Zhang, "An Improved Vision Transformer Network with a Residual Convolution Block for Bamboo Resource Image Identification," Electronics, vol. 12, no. 4, p. 1055, 2023. Available: https://doi.org/10.3390/electronics12041055

[13] W. Jiang, M. Yang, X. Li, W. Zhou, and Z. Feng, "LeMeViT: Efficient Vision Transformer with Learnable Meta Tokens for Remote Sensing Image Interpretation," in Proc. IEEE Int. Conf. Computer Vision (ICCV), 2024, pp. 901–904.

[14] H. Wang, C. Xing, J. Yin, and J. Yang, “Land Cover Classification for Polarimetric SAR Images Based on Vision Transformer,” Remote Sensing, vol. 14, no. 18, p. 4656, 2022. Available: https://doi.org/10.3390/rs14184656
[15] S. K. Roy, M. P. Kundu, S. Bose, A. A. Maulana, and A. Ghosh, “Multimodal Fusion Transformer for Remote Sensing Image Classification,” in Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), 2022, pp. 1234–1237.
[16] V. K. Dhakshnamurthy, M. Govindan, K. Sreerangan, M. D. Nagarajan, and A. Thomas, “Brain Tumor Detection and Classification Using Transfer Learning Models,” Engineering Proceedings, vol. 62, p. 1, 2024. Available: https://doi.org/10.3390/engproc2024062001
[17] S. G. De Benedictis, G. Gargano, and G. Settembre, “Enhanced MRI Brain Tumor Detection and Classification via Topological Data Analysis and Low-Rank Tensor Decomposition,” J. Comput. Math. Data Sci., vol. 13, p. 100103, 2024. [Online]. Available: https://doi.org/10.1016/j.jcmds.2024.100103
[18] N. Thapa, S. Kumari, and J. Lee, “Impact of CLAHE Based Image Enhancement for Diabetic Retinopathy Classification through Deep Learning,” Procedia Computer Science, vol. 214, pp. 197–204, 2023.
[19] M. Gupta, S. K. Sharma, and G. C. Sampada, “Classification of Brain Tumor Images Using CNN,” Computational Intelligence and Neuroscience, vol. 2023, Article no. 2002855, 2023. Available: https://doi.org/10.1155/2023/2002855
[20] A. Younis, L. Qiang, C. O. Nyatega, M. J. Adamu, and H. B. Kawuwa, “Brain Tumor Analysis Using Deep Learning and VGG 16 Ensembling Learning Approaches,” Applied Sciences, vol. 12, no. 14, p. 7282, 2022. Available: https://doi.org/10.3390/app12147282
[21] S. E. Nassar, I. Yasser, H. M. Amer, and M. A. Mohamed, “A Robust MRI Based Brain Tumor Classification via a Hybrid Deep Learning Technique,” The Journal of Supercomputing, vol. 80, no. 2, pp. 2403–2427, 2024. Available: https://doi.org/10.1007/s11227-023-05549-w
[22] https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset/data