Benchmarking CNN and Cutting-Edge Transformer Models for Brain Tumor Classification Through Transfer Learning

Md Saiful Islam Sajol[1] and A S M Jahid Hasan[2]

[1]Dept. of CSE, Louisiana State University Louisiana
[2]Dept. of ECE North South University Dhaka

July 23, 2024

Proceedings of 2024 IEEE 12th International Conference on Intelligent Systems (IS)

Benchmarking CNN and Cutting-Edge Transformer Models for Brain Tumor Classification Through Transfer Learning

Md Saiful Islam Sajol
Dept. of CSE
Louisiana State University
Louisiana, USA
Email: msajol1@lsu.edu

A S M Jahid Hasan
Dept. of ECE
North South University
Dhaka, Bangladesh
Email:jahid.hasan12@northsouth.edu

***Abstract*—Brain tumor is a serious disease that can lead to fatal consequences. Moreover, there are different types of brain tumor with different progression rate and severeness. Thus, brain tumor classification is an essential task in medical diagnosis. Convolutional Neural Network (CNN) based deep learning methods shown great potential brain tumor classification using Magnetic Resonance Imaging (MRI) scans of brain. Recent development in transformer based deep learning models have shown promising outcome in image classification task. However, these models have not yet much explored in the application of brain tumor classification. In this paper, we employ seven advanced models which utilizes transformers or mimics the self-attention operation of transformers through their design, for brain tumor classification. To assess their performance for this particular task, five conventional CNN based methods have also been applied and compared with these models. To make our assessment more coherent and comprehensive we perform the comparison for four different datasets. The results indicate that transformer-based advanced models do not provide a distinct advantage over conventional CNN-based models. The CNN-based ResNet-50 performs well overall, especially with smaller datasets. For larger datasets, transformer-based models generally perform better than CNN-based models, although the difference is not significant.**

***Keywords*—Brain tumor; deep learning; CNN; transformer**

I. Introduction

Brain is the most vital organ in the body. Each and every function of the body is controlled by the brain. Brain tumor is one of the most deadly diseases of human brain. They emerge due to uncontrolled and abnormal growth of brain cells. Based on their growth rate they can be either malignant (cancerous) or benign (non-cancerous). Malignant tumors are life-threatening as one can assume, but even benign tumors can pose danger towards patients by pressurizing and damaging parts of brain. More than a million people are living with a brain tumor diagnosis while about ninety thousand people are being diagnosed with brain tumors each year [1]. Though there are over hundreds of types of tumors related to brain and central nervous system, the most common types are glioma, meningioma and pituitary adenoma. They account for nearly three-fourths of all brain tumors detected [2]. However, the severity and treatment for each of the types of tumor varies. Hence, proper detection and classification of brain tumors is crucial for receiving proper treatment.

Typically, brain tumors are diagnosed using scans of brain derived from X-rays, Ultrasonography, computer tomography (CT), positron emission tomography (PET) Magnetic Resonance Imaging (MRI), etc. These scanned images are then examined by trained and experienced expert medical professionals to confirm the diagnosis. However, recent advancement of deep learning in object detection and classification within images has also found application in brain tumor classification task. In most cases, the deep learning technique applied for brain tumor classification is essentially some variant of a Convolutional Neural Network (CNN). Many other works have used popular CNN based deep learning architecture for brain tumor classification and evaluated their performance. However, in literature we can find various works reporting inconsistent results for these methods. Therefore, a coherent comparison among different methods for brain tumor classification is necessary. Furthermore, as the domain of deep learning and computer vision is a rapidly advancing field, recently introduced transformer based advanced models have emerged that are not yet studied for brain tumor classification.

In this study, we apply the latest and most advanced transformer based deep learning techniques that have demonstrated effectiveness in other image classification tasks. To ensure consistency in our comparison, we evaluate these techniques across multiple brain tumor datasets rather than relying on a single dataset. Several conventional and widely used CNN based deep learning models for brain tumor classification are also evaluated, to compare their performance against the advanced models. Direct transfer learning is employed instead of fine-tuning the models to assess the raw performance of these models. The classification of brain tumors through the application of transfer learning offers significant advantages, particularly because there are only a limited number of datasets with labeled brain tumor MRI scans, and those available are typically quite small.

II. Related Works

A significant number of research has been conducted to classify brain tumors using deep learning methods. Deep CNN has been applied for brain tumor classification [3]. A three class brain tumor dataset has been used here containing MRI of glioma, meningioma and pituitary tumors. Intensity normalization and contrast enhancement is performed as preprocessing of MRI data. The model achieves an accuracy of 94.74%. The use of transfer learning in deep learning models has also been explored for brain tumor classification in several literature. GoogleNet architecture has been modified by changing the last three layers and then transfer learning is applied in [4]. Initially this modified GoogleNet is pre-trained on ImageNet dataset which serves as the source domain and then later with a three class brain tumor MRI dataset which is the target domain. Preprocessing of MRI data is done using min-max normalization and resizing. This model achieves 98% accuracy in overall classification. In another study, the authors address a binary classification problem to determine whether a tumor is benign or malignant [5]. They utilized three pretrained models—U-Net, AlexNet, and VGG-16—along with ResNet 50 using transfer learning. The performance of these models was compared using various classifiers, including Gaussian, SVM linear, SVM sigmoid, Adaboost, ELM, Decision Tree, and Softmax. The results indicated that ResNet50 with transfer learning outperformed the other pretrained models, achieving an accuracy of 96.8%. Pretrained ResNet 50 and Inception V3 models, originally trained on the ImageNet dataset, were applied to two different datasets using three optimizers: Adam, Nadam, and RMSProp [6]. The first dataset comprises four classes (glioma, meningioma, pituitary, and no tumor), while the second is a binary dataset (tumor and no tumor). The findings indicate that ResNet 50 optimized with Nadam achieves the highest accuracy of 97.68% on the four-class dataset. For the binary dataset, ResNet 50 optimized with Adam performs best, achieving an accuracy of 99.83%. Another paper investigates the performance of five models—Xception, DenseNet201, DenseNet121, ResNet152V2, and Inception-ResNetV2—on two brain tumor MRI datasets [7]. One dataset consists of three classes (glioma, pituitary, and meningioma), while the other includes four classes (the mentioned classes plus healthy brain scans). The preprocessing stages involve finding contours, calculating extreme points, and cropping. The results reveal that the Xception model performs best, achieving 99.67% accuracy on the three-class dataset and 95.87% accuracy on the four-class dataset. Another study also found the Xception model to be the top-performing model with an accuray of 98.75% among others, including ResNet50, InceptionV3, VGG16, and MobileNet, using the previously mentioned four-class dataset [8].

A number of researches have adopted hybrid deep learning approaches for brain tumor classification. One such study utilized Improved Ant Colony Optimization (IACO) to optimize the hyperparameters of ResNet 50 [9]. They compared it several other architectures and evolutionary optimization methods. Proposed method achieves an accuracy of 98.69%, outperforming others. One other study [10] employed AlexNet to extract features from brain MRIs, which were subsequently classified using BayesNet, Sequential Minimal Optimization (SMO), Naïve Bayes (NB), and Random Forest (RF) classifiers. Preprocessing was done using an anisotropic filter. The findings demonstrated that the Random Forest (RF) classifier achieved the highest accuracy among the classifiers evaluated. In [11], the authors implemented patch division using three different patch sizes alongside a pretrained ResNet 50 model to extract features from MRI images. Three feature selectors—neighborhood component analysis (NCA), Chi2, and ReliefF—were employed to generate six feature vectors. Classification was performed using k-Nearest Neighbors (kNN) and enhanced by Iterative Hard Majority Voting (IHMV) to improve classification performance. The proposed method achieved an impressive overall accuracy of 98.1%. Another study [12] employed the Enhanced Chimpanzee Optimization Algorithm (EChOA) to select features, aiming to minimize feature dimensionality. Feature classification was conducted using ResNet-152 in conjunction with a softmax classifier. Preprocessing steps included applying median filtering and dilation techniques to data sourced from the three class dataset. Evaluation of the method yielded a notable accuracy of 98.85%. Ensemble method is also explored using five pretrained models—AlexNet, VGG-16, GoogleNet, ResNet 18, and ResNet 50, featuring a Majority Voting algorithm to determine the predicted class [13]. This ensembled algorithm (MajVot) significantly improved average accuracy across three datasets by 3.60%, 2.84%, 1.64%, 4.27%, and 1.14% compared to AlexNet, VGG16, ResNet18, GoogleNet, and ResNet50 individually.

Some recent introduced methods such as graph neural networks and transformers have also been integrated to brain tumor classification. One such work proposes a graph convolutional neural network (GCNN) for tumor classification using four-class tumor images [14]. The images are preprocessed with a two-dimensional Gaussian filter. Five different network configurations are compared, and the GCNN with Dropout (DO) or Batch Normalization (BN) using a Gaussian Adjacency Matrix performs the best. A hybrid transformer-enhanced convolutional neural network (TECNN)-based model is proposed for brain tumor classification [15], where the CNN is used for local feature extraction and the transformer employs an attention mechanism to extract global features. The experimental results of the model the three class dataset achieves an average accuracy of 99.1%. A deep CNN integrated with spatial attention is explored in [16]

III. Methods

In this paper, we utilize advanced deep learning models, including Vision Transformer (ViT), Swin Transformer V2, ConvNext V2, Convolutional Vision Transformer (CvT), EfficientFormer, Pyramid Vision Transformer V2 (PVTv2), and

MobileViT V2. These models are very recently introduced and have not been explored in the context of brain tumor classification. Except for ConvNext V2, which is a CNN-based method, all the models are transformer-based. However, ConvNext V2 is designed to emulate the self-attention characteristics of a transformer. Therefore, we will consider it as part of the transformer-based advanced methods. To assess their efficacy in brain tumor classification, we also compare these models with widely used CNN based models such as ResNet-50, VGG-16, GoogleNet, MobileNet, and EfficientNetB0. Since these conventional models are extensively covered in the literature, this section will briefly discuss the advanced models only.

A. Vision Transformer

Transformers are the most prominent deep learning methods used in Natural Language Processing (NLP). Vision Transformers (ViT) represent a Transformer model specifically tailored for computer vision applications. Unlike many other methods, ViT does not integrate the self-attention mechanism of Transformers with Convolutional Neural Networks (CNNs); instead, it utilizes a pure Transformer architecture. ViT addresses the 1-D token embedding input requirement of Transformers by converting 2-D images into several image patches and flattening them into a sequence of 1-D input token embeddings. This approach requires significantly fewer resources than CNNs while achieving comparable results in image classification.

B. Swin Transformer V2

Swin Transformers serve as a general-purpose backbone for computer vision, replacing the multi-head self-attention module of traditional Transformers with a shifting window module. The Swin Transformer V2 introduces several improvements over its predecessor. The key differences include the adoption of residual post-normalization instead of pre-normalization, the replacement of dot-product attention with scaled cosine attention, and the implementation of a log-spaced continuous relative position bias approach in place of the previous parameterized approach. These enhancements enable the model to scale up capacity more easily and transfer more effectively across different window resolutions.

C. ConvNext V2

ConvNext represents an evolution of the ResNet-50 architecture, designed to achieve performance comparable to that of Swin Transformers. The latest iteration, ConvNext V2, incorporates the Fully Convolutional Masked Autoencoder (FC-MAE) framework, which enhances the self-supervised learning capabilities inherent in transformer-based models. Additionally, ConvNext V2 addresses the issue of feature collapse through the implementation of Global Response Normalization (GRN). This architecture demonstrates superior accuracy in ImageNet classification, surpassing the performance of other contemporary models.

D. Convolutional Vision Transformer

The Convolutional Vision Transformer (CvT) architecture combines convolutional operations with Vision Transformers (ViTs) to leverage the strengths of both approaches. The first component, known as Convolutional Token Embeddings, adjusts the token feature dimension and the number of tokens. The second component, Convolutional Projection, replaces the position-wise linear projection used in Multi-Head Self-Attention (MHSA) with depth-wise separable convolutions. This architecture achieves higher accuracy with fewer parameters compared to other state-of-the-art Transformer-based models.

E. EfficientFormer

Dimension-consistent design and latency driven slimming method are introduced to ViTs to create a faster Transformer model. This model maintains high accuracy while significantly reducing latency, making it highly suitable for mobile. devices.

F. Pyramid Vision Transformer V2

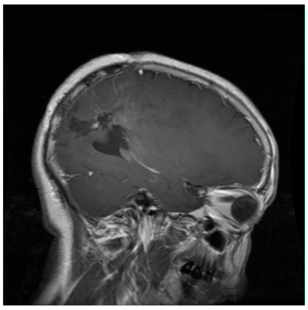
The Pyramid Vision Transformer (Pvt) employs a simpler network design while delivering performance comparable to Swin Transformers. This network progressively reduces the feature map size at each layer before passing it into a transformer encoder. In Version 2, enhancements include a linear complexity attention layer, overlapping patch embedding, and a convolutional feed-forward network.

G. MobileVit V2

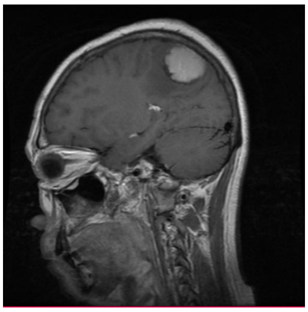
Like EfficientFormer, MobileViT is a lightweight, low-latency model designed for mobile device applications. MobileViT V2 improves on its predecessor by introducing a separable self-attention method with linear complexity, making it faster.

IV. Dataset

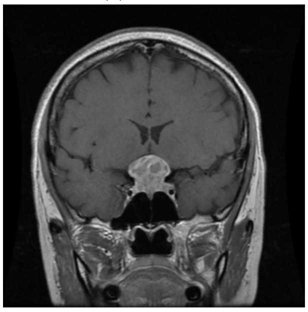
To make our performance comparison more coherent and reliable, we have used four different datasets for our work. Among them, one is the most widely used dataset for brain tumor classification, popularly known as Figshare dataset [17]. It contains total 3,064 T1-weighted contrast-enhanced MRI scans of human brains collected from total 233 patients. There are total three classes of images: (i) glioma (1406 slices), (ii) meningioma (708 slices), and (iii) pituitary (930 slices). The second dataset is newly published and contains a total of 24,410 images [18]. In addition to the three tumor types mentioned in the previous dataset, it includes an extra class labeled ”no tumors”. We call this the ”Akbar” dataset. The third dataset is also a four class dataset with the same four classes as the previous one [19]. The dataset contains 3,260 T1-weighted contrast-enhanced images with cleaning and augmentation performed on them. This dataset is known as the Sartaj dataset. The last one is also a three class dataset containing 3,064 MRI images [20]. We call this the Hemant dataset. The last three datasets are named after their respective



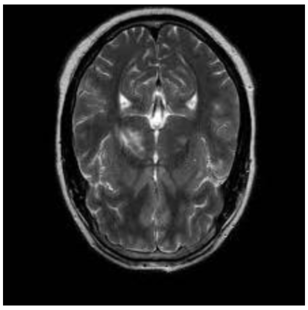
(a) Glioma



(b) Meningioma



(c) Pituitary



(d) No Tumor

Fig. 1: Sample image of each class from Akbar Dataset

primary uploaders. Figure 1 shows sample from each class taken from Akbar dataset.

V. Experiments and Results

The models for brain tumor classification are implemented using PyTorch, a popular Python-based deep learning framework. For the three smaller datasets, the training and test data are split with a ratio of 80% to 20%, while the larger Akbar dataset is split with a ratio of 95% to 5%. A learning rate of 0.0001 and a batch size of 32 are used consistently across all datasets. The experiments have been conducted on a 3.20 GHz Intel Xeon Silver 4216 CPU with 32 GB RAM and Tesla V100-SXM2 GPU.

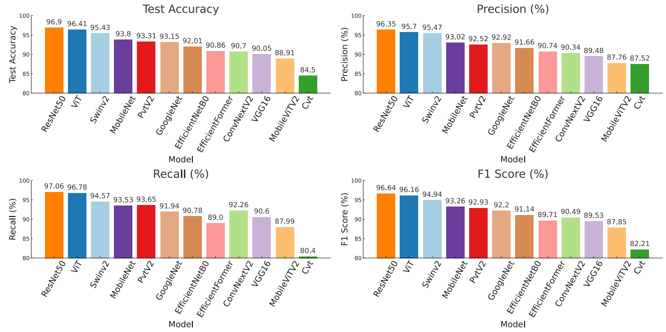
Due to the large number of parameters in many of the models used, we employ direct transfer learning without performing any fine-tuning. This approach provides a clearer understanding of the raw performance of each model. Therefore, we compare the results of each model at the first epoch to serve our purpose.

For each of the model and dataset we compare four metrics: accuracy, precision, recall and F1-score. These metrics are defined as:

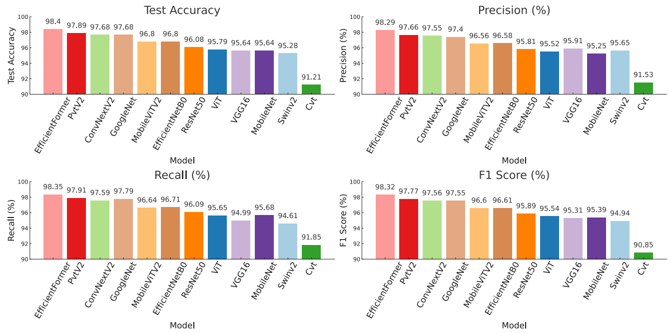
$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$precision = \frac{TP}{TP + FP} \tag{2}$$

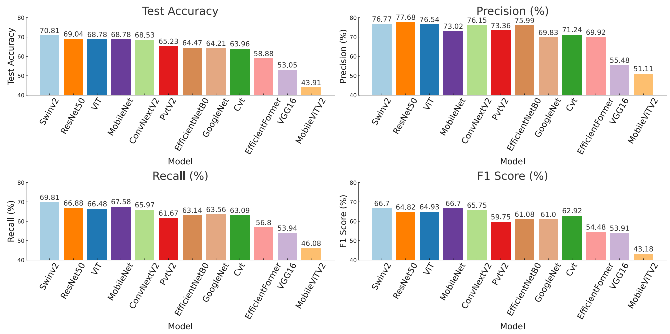
$$recall = \frac{TP}{TP + FN} \tag{3}$$



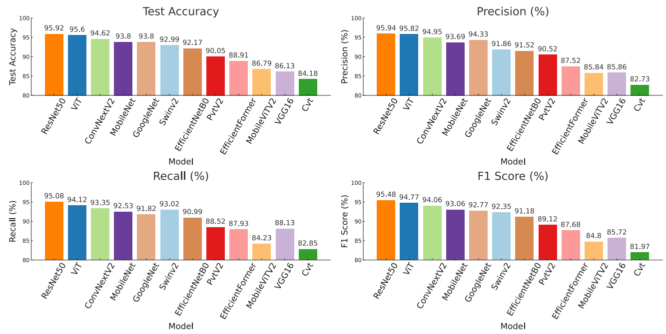
(a) Dataset: Figshare



(b) Dataset: Akbar



(c) Dataset: Sartaj



(d) Dataset: Hemant

Fig. 2: Comparison of model metrics across different datasets.

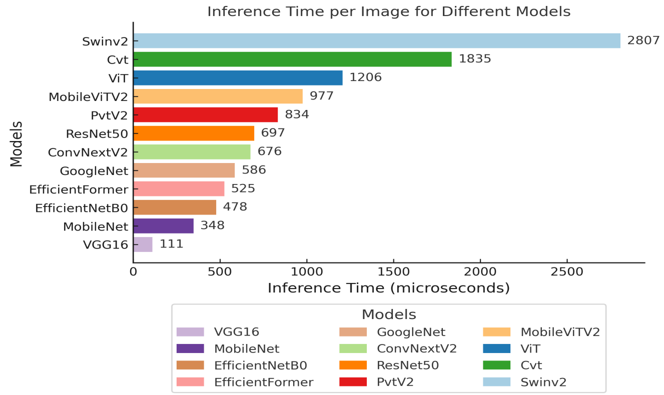


Fig. 3: Inference Time for different models

$$F1 = 2\frac{precision \times recall}{precision + recall} \tag{4}$$

where, TP and TN denote true positive and true negative, while FP and FN stand for false positive and false negative respectively.

Figure 2 illustrates the performance metrics for each model across various datasets. The results reveal that ResNet excels on the three smaller datasets, securing top positions across all metrics for both the Figshare and Hemant datasets, which are three-class datasets. For the four-class Sartaj dataset, ResNet surpasses most other methods in all metrics. Among the smaller datasets, ViT follows closely, ranking second across all metrics for the three-class datasets. In the four-class Sartaj dataset, ViT consistently ranks within the top five for all metrics. Conversely, Swin Transformer V2 shows strong performance on the four-class Sartaj dataset, achieving first or second place in all metrics. When comparing the results for the much larger four-class Akbar dataset, we observe that MobileVit V2 outperforms all other models across all metrics, followed by EfficientNetB0. Evaluating the performance of trasnformer based advanced models against CNN based traditional models indicates that advanced models deliver better performance in comparison to traditional models in general over this dataset. However, the difference within the performance metrics of the different models are negligible. Table I presents the model performance metrics across different datasets.

Figure 3 illustrates the inference time per image for the models used in this work, applied to the Sartaj dataset. The figure shows that VGG-16 is the fastest model, while Swin Transformer V2 is the slowest. Comparing the other models, we observe that advanced models are generally slower compared to traditional models, with the exceptions of ConvNext V2 and EfficientFormer. This indicates that Transformer-based models are typically more computationally intensive and complex, primarily due to their multi-head self-attention (MHSA) module, which increases computational complexity. As EfficientFormer is a light-weight and low latency transformer based model and ConvNext V2 is purely CNN based model, they perform faster than the other advanced models.

TABLE I: Model performance metrics across different datasets

Model	Metric	Figshare	Akbar	Sartaj	Hemant
Swin V2	Accuracy	95.43	95.28	70.82	92.99
	Precision	0.9547	0.9565	0.7677	0.9186
	Recall	0.9457	0.9461	0.6981	0.9302
	F1-Score	0.9494	0.9494	0.6670	0.9235
ViT	Accuracy	96.41	95.79	68.78	95.60
	Precision	0.9570	0.9552	0.7654	0.9582
	Recall	0.9678	0.9565	0.6648	0.9412
	F1-Score	0.9616	0.9554	0.6493	0.9477
ConvNext V2	Accuracy	90.70	97.62	68.53	94.62
	Precision	0.9034	0.9755	0.7615	0.9495
	Recall	0.9226	0.9759	0.6597	0.9335
	F1-Score	0.9049	0.9756	0.6575	0.9406
Cvt	Accuracy	84.50	91.21	63.96	84.18
	Precision	0.8752	0.9153	0.7124	0.8273
	Recall	0.8040	0.9185	0.6309	0.8285
	F1-Score	0.8221	0.9085	0.6292	0.8197
EfficientFormer	Accuracy	90.86	98.40	58.88	88.91
	Precision	0.9074	0.9829	0.6992	0.8752
	Recall	0.8900	0.9835	0.5680	0.8793
	F1-Score	0.8971	0.9832	0.5448	0.8768
PvtV2	Accuracy	93.31	97.89	65.23	90.05
	Precision	0.9252	0.9766	0.7336	0.9052
	Recall	0.9365	0.9791	0.6167	0.8852
	F1-Score	0.9293	0.9777	0.5975	0.8912
MobileViTV2	Accuracy	88.91	96.80	43.91	86.79
	Precision	0.8776	0.9656	0.5111	0.8584
	Recall	0.8799	0.9664	0.4608	0.8423
	F1-Score	0.8785	0.9660	0.4318	0.8480
ResNet-50	Accuracy	96.90	96.08	69.04	95.92
	Precision	0.9635	0.9581	0.7768	0.9594
	Recall	0.9706	0.9609	0.6688	0.9508
	F1-Score	0.9664	0.9589	0.6482	0.9548
VGG-16	Accuracy	90.05	95.64	53.05	86.13
	Precision	0.8498	0.9591	0.5548	0.8586
	Recall	0.9060	0.9499	0.5394	0.8813
	F1-Score	0.8953	0.9531	0.5391	0.8572
MobileNet	Accuracy	93.80	95.64	68.78	93.80
	Precision	0.9302	0.9525	0.7302	0.9369
	Recall	0.9353	0.9568	0.6758	0.9253
	F1-Score	0.9326	0.9539	0.6670	0.9306
GoogleNet	Accuracy	93.15	97.68	64.21	93.80
	Precision	0.9292	0.9740	0.6983	0.9433
	Recall	0.9194	0.9779	0.6356	0.9182
	F1-Score	0.9220	0.9755	0.6100	0.9277
EfficientNetB0	Accuracy	92.01	96.80	64.47	92.17
	Precision	0.9166	0.9658	0.7599	0.9152
	Recall	0.9078	0.9671	0.6314	0.9099
	F1-Score	0.9114	0.9661	0.6108	0.9118

From the above discussions, it is evident that no single model consistently outperforms the others across all datasets when considering the mentioned metrics. Transformer based advanced models do not demonstrate clear superiority over conventional models. In fact, ResNet-50 performs exceptionally well on smaller datasets and is significantly faster than the advanced models. However, for larger datasets, transformer based advanced models generally show somewhat better performance than CNN based models, although the differences are minimal. Therefore, it can be inferred that the raw performance of conventional CNN based models is not necessarily inferior to advanced transformer based models in brain tumor classification with transfer learning.

VI. Conclusion

Brain tumor classification is crucial due to the severity of the disease and the necessary proper treatment. Application of deep learning methods on the MRI scans of brain for tumor detection has shown promise. Different deep learning models employed on different dataset have led to varying results. Moreover, the effect of some recently developed models for image classification have not yet been explored for tumor classification task. This paper performs experiments of seven transformer based models, and five traditional CNN based deep learning models for brain tumor classification task using four datasets. Only direct transfer learning has been of the models are used without any fine tuning to assess the raw performance of each model. The findings from results demonstrate no clear advantage of transformer based models over CNN based conventional ones. ResNet-50 particularly gives good results in terms of performance metrics and inference time. However, transformer based methods in general shows slightly superior performance over traditional ones when larger datasets are concerned. The recent popularity of transformer models may encourage researchers to use them for brain tumor classification. However, our results demonstrate that CNN-based methods remain a strong choice for this task. Their performance is comparable to transformer-based models, and they are generally faster. This work can serve as a reference for others who intend to develop highly accurate deep learning models specially tailored towards brain tumor classification task.

References

[1] Ostrom, Q., Patil, N., Cioffi, G., Waite, K., Kruchko, C. & Barnholtz-Sloan, J. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013–2017. *Neuro-Oncology*. **22**, iv1-iv96 (2020,10), http://dx.doi.org/10.1093/neuonc/noaa200

[2] [Online] "Brain Tumor Types" Available: https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor/brain-tumor-types

[3] Ayadi, W., Elhamzi, W., Charfi, I. & Atri, M. Deep CNN for Brain Tumor Classification. *Neural Processing Letters*. **53**, 671-700 (2021,1), http://dx.doi.org/10.1007/s11063-020-10398-2

[4] Deepak, S. & Ameer, P. Brain tumor classification using deep CNN features via transfer learning. *Computers In Biology And Medicine*. **111** pp. 103345 (2019,8), http://dx.doi.org/10.1016/j.compbiomed.2019.103345

[5] Kumar, S., Choudhary, S., Jain, A., Singh, K., Ahmadian, A. & Bajuri, M. Brain Tumor Classification Using Deep Neural Network and Transfer Learning. *Brain Topography*. **36**, 305-318 (2023,4), http://dx.doi.org/10.1007/s10548-023-00953-0

[6] Ahmmed, S., Podder, P., Mondal, M., Rahman, S., Kannan, S., Hasan, M., Rohan, A. & Prosvirin, A. Enhancing Brain Tumor Classification with Transfer Learning across Multiple Classes: An In-Depth Analysis. *BioMedInformatics*. **3**, 1124-1144 (2023,12), http://dx.doi.org/10.3390/biomedinformatics3040068

[7] Asif, S., Zhao, M., Tang, F. & Zhu, Y. An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning. *Multimedia Tools And Applications*. **82**, 31709-31736 (2023,2), http://dx.doi.org/10.1007/s11042-023-14828-w

[8] Saleh, A., Sukaik, R. & Abu-Naser, S. Brain Tumor Classification Using Deep Learning. *2020 International Conference On Assistive And Rehabilitation Technologies (iCareTech)*. pp. 131-136 (2020)

[9] Mehnatkesh, H., Jalali, S., Khosravi, A. & Nahavandi, S. An intelligent driven deep residual learning framework for brain tumor classification using MRI images. *Expert Systems With Applications*. **213** pp. 119087 (2023,3), http://dx.doi.org/10.1016/j.eswa.2022.119087

[10] Sarkar, A., Maniruzzaman, M., Alahe, M. & Ahmad, M. An Effective and Novel Approach for Brain Tumor Classification Using AlexNet CNN Feature Extractor and Multiple Eminent Machine Learning Classifiers in MRIs. *Journal Of Sensors*. **2023** pp. 1-19 (2023,3), http://dx.doi.org/10.1155/2023/1224619

[11] Muezzinoglu, T., Baygin, N., Tuncer, I., Barua, P., Baygin, M., Dogan, S., Tuncer, T., Palmer, E., Cheong, K. & Acharya, U. PatchResNet: Multiple Patch Division–Based Deep Feature Fusion Framework for Brain Tumor Classification Using MRI Images. *Journal Of Digital Imaging*. **36**, 973-987 (2023,2), http://dx.doi.org/10.1007/s10278-023-00789-x

[12] Athisayamani, S., Antonyswamy, R., Sarveshwaran, V., Almeshari, M., Alzamil, Y. & Ravi, V. Feature Extraction Using a Residual Deep Convolutional Neural Network (ResNet-152) and Optimized Feature Dimension Reduction for MRI Brain Tumor Classification. *Diagnostics*. **13**, 668 (2023,2), http://dx.doi.org/10.3390/diagnostics13040668

[13] Tandel, G., Tiwari, A., Kakde, O., Gupta, N., Saba, L. & Suri, J. Role of Ensemble Deep Learning for Brain Tumor Classification in Multiple Magnetic Resonance Imaging Sequence Data. *Diagnostics*. **13**, 481 (2023,1), http://dx.doi.org/10.3390/diagnostics13030481

[14] Ravinder, M., Saluja, G., Allabun, S., Alqahtani, M., Abbas, M., Othman, M. & Soufiene, B. Enhanced brain tumor classification using graph convolutional neural network architecture. *Scientific Reports*. **13** (2023,9), http://dx.doi.org/10.1038/s41598-023-41407-8

[15] Aloraini, M., Khan, A., Aladhadh, S., Habib, S., Alsharekh, M. & Islam, M. Combining the Transformer and Convolution for Effective Brain Tumor Classification Using MRI Images. *Applied Sciences*. **13**, 3680 (2023,3), http://dx.doi.org/10.3390/app13063680

[16] Islam Sumon, R., Bhattacharjee, S., Hwang, Y., Rahman, H., Kim, H., Ryu, W., Kim, D., Cho, N. & Choi, H. Densely Convolutional Spatial Attention Network for nuclei segmentation of histological images for computational pathology. *Frontiers In Oncology*. **13** (2023,5), http://dx.doi.org/10.3389/fonc.2023.1009681

[17] Cheng, J. brain tumor dataset. (figshare,2017), https://figshare.com/articles/dataset/brain-tumor-dataset/1512427/5

[18] Akbar tumor Dataset. *Roboflow Universe* . (2024,3), https://universe.roboflow.com/akbar-chgxr/tumor-bahib, visited on 2024-06-03

[19] Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S. & Kanchan, S. Brain Tumor Classification (MRI). (Kaggle,2020), https://www.kaggle.com/dsv/1183165

[20] Kumar, H. Brain Tumor Classification MRI Images — kaggle.com. (https://www.kaggle.com/datasets/jarvisgroot/brain-tumor-classification-mri-images/data,0), [Accessed 24-06-2024]