Trans-EffNet: A Hybrid Model for Brain Tumor Detection Using EfficientNet and Transformer Encoder

Abdul Haseeb Nizamani
School of Computer Science and Engineering
Central South University
Changsha 410083, China
Email: haseeb@csu.edu.cn

Zhigang Chen
School of Computer Science and Engineering
Central South University
Changsha 410083, China
Email: czg@csu.edu.cn

Ahsan Ahmed Nizamani
School of Computer Science and Engineering
Central South University
Changsha 410083, China
Email: ahsan.official@csu.edu.cn

Mughair Aslam Bhatti
School of Information and Communication Engineering
Hainan University
Hainan, China
Email: mughairbhatti@nnu.edu.cn

Hauwei Ma
University Guangzhou
Guangzhou, China
Email: mkmhw@gzhu.edu.cn

Wencai Du
The University of Saint Joseph
Macao
Email: george.du@usj.edu.mo

2024 IEEE Smart World Congress (SWC) | 979-8-3315-2086-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/SWC62898.2024.00260

***Abstract*—Accurate classification of brain tumors from MRI is critical for effective diagnosis and treatment. In this study, we introduce Trans-EffNet, a hybrid model combining pre-trained EfficientNet architectures with a transformer encoder to enhance brain tumor classification accuracy. By leveraging EfficientNet's deep CNN capabilities for localized feature extraction and the transformer encoder for capturing global contextual relationships, our model improves the identification of intricate tumor characteristics. Fine-tuned with ImageNet-derived weights and utilizing extensive data augmentation, Trans-EffNet was validated on both multi-class and binary datasets. Trans-EffNetB1 achieved 99.49% accuracy on the multi-class dataset, while Trans-EffNetB2 recorded 99.83% accuracy on the binary dataset, with perfect precision, recall, and F1-Score. These results underscore Trans-EffNet's robustness and potential as a significant advancement in brain tumor detection and classification.**

***Index Terms*—Brain tumor detection, EfficientNet, Transformer, Classification, Magnetic Resonance Imaging.**

I. Introduction

Brain tumors is a critical global health concern, leading to significant mortality and severely impacting quality of life. These tumors arise from the uncontrolled growth of abnormal brain cells, often resulting in cancerous conditions [1]. The incidence of brain tumors is increasing, making early detection and treatment essential. Currently, brain cancer ranks as the tenth leading cause of death worldwide [2]. Various types of brain tumors exist, including malignant gliomas and benign meningiomas and pituitary tumors, adding complexity to diagnosis and treatment [3]. Gliomas, the most common primary brain tumors, have the highest mortality rates due to their aggressive nature [4], [5]. Glioblastoma, a severe form of glioma, can evade the immune system, complicating treatment and leading to a poor prognosis [6]. Accurate and timely tumor identification is crucial for improving survival rates.

Advancements in science and technology have significantly enhanced medical imaging techniques, offering detailed insights into health conditions such as bone abnormalities, tissue damage, and blood circulation, thereby enabling precise medical evaluations [7]. The digitization of medical imaging has transformed diagnostic practices through sophisticated image processing methods [8], [9]. MRI provides a non-invasive approach to obtaining high-resolution brain images, which are crucial for the accurate diagnosis and assessment of brain tumors [10]. However, the traditional manual segmentation of MRI images for brain tumors is labor-intensive and prone to human error, which can lead to misdiagnoses and suboptimal treatment strategies [11].

Machine learning (ML) has significantly advanced computer-aided diagnostic (CAD) systems in medical imaging, utilizing techniques like Support Vector Machines (SVM), Decision Trees (DT), and k-Nearest Neighbour (KNN) for data classification [12]. These methods rely on feature engineering, a manual process of selecting and refining data features to enhance model accuracy, which can be labor-intensive and prone to errors, especially with large datasets [14]. Inaccuracies in medical diagnostics can lead to incorrect diagnoses and impact patient outcomes. To

2993-396X/24/$31.00 ©2024 IEEE
DOI 10.1109/SWC62898.2024.00260

overcome these challenges, there is a growing shift towards deep learning techniques that automate feature extraction, enabling CAD systems to learn directly from raw data, thus improving efficiency, accuracy, and reliability in medical image analysis.

Deep learning models have become a robust alternative in computer-aided diagnostic (CAD) systems, outperforming traditional machine learning methods that depend on manual feature engineering [16]. CNNs have demonstrated exceptional performance in the analysis of medical images, particularly in detecting brain tumors [17]. These networks automatically extract crucial details from images, facilitating precise classification without the need for handcrafted features [18]. CNNs streamline the process by learning directly from the data, which enhances both the efficiency and accuracy of medical diagnostics.

Vision Transformers (ViTs) are a groundbreaking architecture distinct from traditional CNNs, demonstrating exceptional performance in brain tumor diagnosis [20]. Leveraging attention mechanisms, ViTs capture intricate visual patterns by modeling long-range dependencies and integrating global and local image features for precise classification [21], [22]. Though effective, there's still a need to enhance the accessibility, accuracy, and efficiency of brain tumor diagnosis. Current research is focused on refining deep learning techniques, including transformers, to revolutionize diagnostic methods, improve patient outcomes, and accelerate cancer detection and classification [19].

Deep learning approaches face challenges, including a lack of labeled data, observer variability in diagnosis, overfitting, and interpretability requirements [23]. Gathering diverse, well-annotated datasets and using strategies like cross-validation, regularization, and data augmentation can mitigate these challenges. Interpretable techniques, such as saliency maps and attention maps, can help the medical community accept and trust deep learning predictions. Addressing these issues can improve the accuracy and reliability of brain tumor identification models, enhancing diagnostic precision and patient care.

A. Key Contributions and Novelty

This study presents a novel approach for brain tumor classification, leveraging advanced deep learning techniques to improve the accuracy and efficiency of MRI-based diagnoses. Using an open-source Kaggle dataset containing four tumor types (glioma, no tumor, meningioma, pituitary) for multi-class classification and tumor/no-tumor for binary classification, we applied transfer learning with pre-trained EfficientNet models enhanced by a transformer encoder. This approach fine-tunes EfficientNets for precise classification and utilizes the transformer encoder's self-attention mechanisms to capture subtle distinctions between tumor types.

Key contributions include:

- Advanced data augmentation and image processing techniques, including spatial transformations and CLAHE, to enhance MRI image clarity.
- A hybrid model combining EfficientNet's feature extraction with transformers' contextual awareness to recognize patterns and interactions between image regions.
- Comprehensive evaluation using metrics such as accuracy, precision, recall, and F1-score, with results visualized through confusion matrices.
- Grad-CAM for interpretability, highlighting key MRI regions influencing model decisions, ensuring transparency in identifying tumorous areas.

II. Proposed Method for MRI Classification

The first step in implementing the Trans-EffNet model involves data preprocessing, including cropping, resizing, applying data augmentation, and enhancing image contrast with CLAHE. After preprocessing, brain MRI data is classified using a fine-tuned model that combines EfficientNet variants (B0 to B3) with a transformer model. Figure 1 presents the block diagram of our proposed approach.

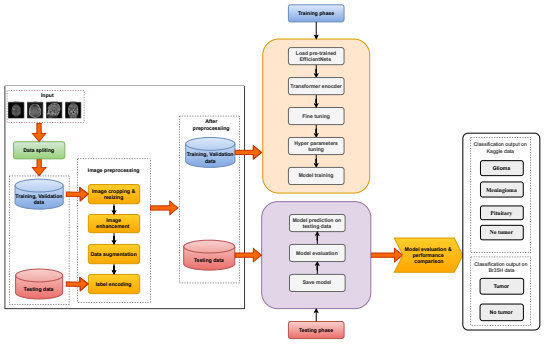

Fig. 1. Block Diagram of Proposed Model.

A. Preprocessing and Data Augmentation

The brain MRI datasets initially contained noise and irrelevant regions, requiring preprocessing for accurate brain tumor classification. Preprocessing included resizing images from $512 \times 512 \times 3$ to $224 \times 224 \times 3$ to fit the Trans-EffNet models, and employing cropping to focus on regions of interest, reducing computational costs while retaining essential features. Random shuffling ensured effective training on unsorted data. CLAHE was utilized to enhance image quality and control noise, improving contrast through tile-based histogram equalization with bilinear interpolation. Data augmentation techniques such as rotation, horizontal flipping, and cropping with resizing increased dataset size and variability, aiding in preventing overfitting. The dataset was split into training and testing subsets, with augmentation applied only to the training data. Class labels were assigned as follows: glioma (0), no tumor (1), meningioma (2), and pituitary tumor (3). Figure 2 shows the preprocessing workflow, from original images to augmented versions.

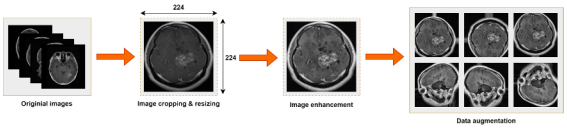

Fig. 2. Preprocessing the input data, from left to right, original images, image cropping and resizing, image enhancement and data augmentation.

B. Proposed Trans-EffNet Model

In 2019, [24] introduced EfficientNets, a groundbreaking advancement in deep CNN architecture, particularly noted for their efficiency in tasks such as ImageNet image classification and transfer learning. EfficientNets depart from traditional methods by employing a compound scaling strategy that systematically adjusts the network's width, depth, and resolution using scaling coefficients. This approach results in a family of models, EfficientNetB0 through B7, that excel in classification and segmentation while minimizing computational expense.

EfficientNet models range from B0, with 237 layers and 5.3 million parameters, to B7, with 813 layers and 66 million parameters. They feature mobile inverted bottleneck convolution (MBConv) layers, a shared component with MobileNetV2 and MnasNet, which supports automatic image normalization. Our research uses EfficientNet variants B0 to B3 for brain tumor classification to balance accuracy with computational requirements. Larger models like B4 to B7, despite their increased depth and parameter count, can overfit and demand substantial computing power. Therefore, we focus on EfficientNetB0 to B3 to maintain efficiency and manage computational load effectively.

As illustrated in Figure 3, our approach integrates EfficientNet architectures with transformer models in a framework we call Trans-EffNet, designed for brain tumor detection in MRI scans. This involves fine-tuning pre-trained EfficientNet models (B0 to B3) initially trained on the ImageNet dataset and adapting them to the MRI brain tumor context. We preserve the convolutional layers and introduce a Global Average Pooling (GAP) layer, followed by a dropout layer, and replace the output with a softmax-activated layer for multi-class classification. For binary classification tasks, a sigmoid activation function is employed. By integrating transformers, we enhance our model's ability to analyze complex MRI data patterns, combining the efficiency of EfficientNets with the deep learning power of transformers, thereby advancing medical image analysis in brain tumor classification.

III. Experimental Setting and Results

A. Evaluation metrics

This section presents four important metrics that are necessary for assessing the model's capability to categorize different forms of brain tumors: accuracy (Γ), precision (Ψ), recall (Θ), and F1-Score (Φ). In order to determine the total efficacy, accuracy (Γ) evaluates the percentage of right predictions. Accuracy (Ψ) represents the percentage of correct predictions (α) out of all positive predictions, while recall (Θ) computes the percentage of correct positives found out of all real positives. In a balanced way, the F1-Score (Φ) combines recall and accuracy. The variables α, β, γ, and δ represent positive and negative results, false positives and false negatives, respectively.

$$\Gamma = \frac{\alpha + \beta}{\alpha + \beta + \gamma + \delta} \tag{1}$$

$$\Psi = \frac{\alpha}{\alpha + \gamma} \tag{2}$$

$$\Theta = \frac{\alpha}{\alpha + \delta} \tag{3}$$

$$\Phi = \frac{2 \times \Psi \times \Theta}{\Psi + \Theta} \tag{4}$$

B. Dataset and Experimental Setup

The dataset used in this study was sourced from Kaggle to evaluate and analyze the performance of the proposed models [25], [26]. The primary dataset consists of 3,264 MRI images, divided into four classes: glioma, no tumor, meningioma, and pituitary tumors. These images are split between training and testing sets to ensure a robust evaluation of model performance. Additionally, our proposed model also utilized a binary MRI dataset from Kaggle, which categorizes images into tumor and no tumor classes. This dataset comprises a total of 3,000 images, with a balanced distribution between training and testing sets. Table 1 presents the class-wise distribution of both datasets, while Figure 4 showcases sample images from each dataset, providing a visual representation of the diversity within the data.

TABLE I
CLASS-WISE DISTRIBUTION OF BOTH DATASETS.

Datasets	Tumor Types	Original data	Train set	Test Set
Kaggle	Glioma	926	826	100
	No tumor	510	395	115
	Meningioma	927	822	105
	Pituitary	901	827	74
	Total	**3264**	**2870**	**394**
Br35H	Tumor	1500	1200	300
	No tumor	1500	1200	300
	Total	**3000**	**2400**	**600**

The experimental hardware configuration comprises a 3.40 GHz central processing unit (CPU), a 64-gigabyte memory and an NVIDIA GeForce RTX 3090 Ti graphics processing unit (GPU). Keras and TensorFlow served as the foundational frameworks for the model development. The Adam optimizer was selected because of its consistency in improving performance. The datasets were separated into training and testing sets, 80% were allocated to the training and remaining 20% for the testing set.

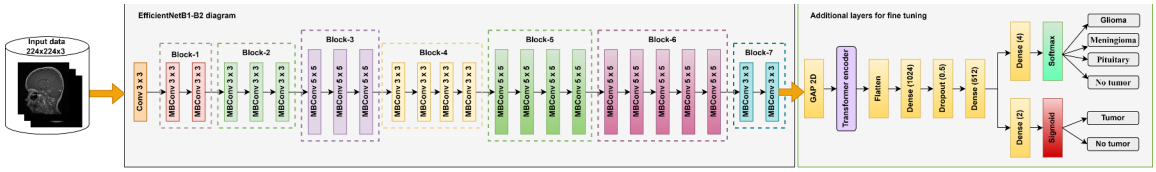

Fig. 3. Diagram of proposed Trans-EffNet models.

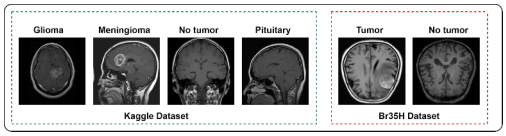

Fig. 4. Sample images of both datasets.

C. Hyper-Parameters Setting

The training process for our model involved careful fine-tuning of various parameters, including optimizers, batch size, learning rate, epochs, and loss functions. We utilized the Adam optimizer with an initial learning rate of 0.001 and categorical cross-entropy as the loss function for both binary and multi-class classification tasks. The model was trained with a batch size of 32 over 50 epochs, incorporating a dropout rate of 0.2 to prevent overfitting. Additionally, 10% of the dataset was reserved for validation to ensure robust monitoring of the model's performance during training. Key hyperparameters included an input shape of (224, 224, 3), a learning rate decay factor of 0.3, and early stopping with a patience of 5 epochs. These settings were crucial in optimizing the model's ability to classify brain tumors effectively.

D. Performance Comparison on Both Datasets

This section explains the evaluated performance of Trans-EffNets on two datasets: a multi-class brain tumor dataset (glioma, no-tumor, meningioma, and pituitary) and the BR35H dataset (tumor and no-tumor).

Kaggle Dataset: The outcomes of the Trans-EffNet models' performance on unseen data are shown in the testing results. After being fine-tuned using a transformer encoder, the EfficientNetB1 model demonstrated outstanding performance on all criteria. The accuracy rates for the glioma, no-tumor, meningioma, and pituitary classifications were 99.10%, 100%, 99.01%, and 100%, respectively. This performance shows the model's capacity to correctly identify different kinds of brain tumors using MRI data. Figure 5 shows the confusion matrix, which shows how well the model classified each type of tumor. Figure 6, which shows the model's training and validation performance, contains loss and accuracy curves. Trans-EffNetB1 achieved a validation accuracy of 99.69% and a training accuracy of 99.96%. Trans-EffNetB1 achieved a test accuracy of 99.49%, precision of 99.51%, recall of 99.57%, and F1-Score of 99.54%, surpassing other models.

BR35H Dataset: This dataset classifies brain MRI images into tumor and no tumor categories. Table3 summarizes the

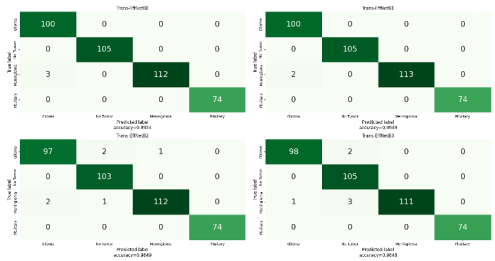

Fig. 5. Confusion matrix of Trans-EffNets on Kaggle dataset.

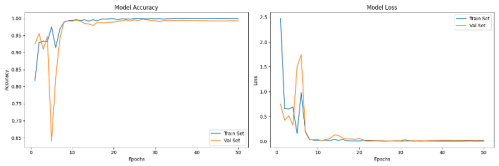

Fig. 6. Accuracy and loss curve of Trans-EffNetB1.

performance metrics for each Trans-EffNet model on both the Kaggle and BR35H datasets, including training, validation, and testing accuracy, as well as precision, recall, and F1-score. Trans-EffNetB2 achieved the highest accuracy on the BR35H dataset, with a test accuracy of 99.83% and perfect precision, recall, and F1-scores of 100%. The confusion matrix in Figure 7 shows minimal misclassifications, demonstrating the model's robustness in handling ambiguous cases. The accuracy and loss curves in The figure 8 demonstrate consistent performance with small variations.

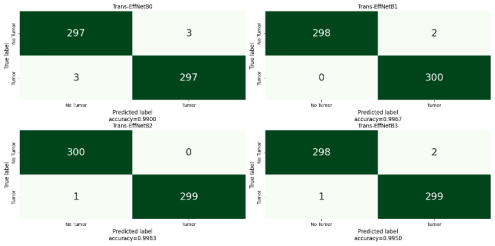

Fig. 7. Confusion matrix of Trans-EffNets on Br35H dataset.

E. Performance Comparison with State-of-the-Art Methods

Evaluating our model against established methods validates its effectiveness and highlights its strengths. For a compre-

TABLE II
PERFORMANCE OF TRANS-EFFNETS ON BOTH DATASETS BASED ON TRAINING, VALIDATION, AND TESTING ACCURACY, ALONG WITH PRECISION, RECALL, AND F1-SCORE.

Dataset	Model	Train Acc(%)	Val Acc(%)	Test Acc(%)	Precision(%)	Recall(%)	F1-Score(%)
Kaggle	Trans-EffNetB0	99.90	99.80	99.24	99.27	99.35	99.30
	Trans-EffNetB1	99.96	99.69	99.49	99.51	99.57	99.53
	Trans-EffNetB2	99.56	99.01	98.48	98.58	98.60	98.58
	Trans-EffNetB3	99.54	99.00	98.48	98.61	98.63	98.48
BR35H	Trans-EffNetB0	99.91	99.17	99.00	99.00	99.00	99.00
	Trans-EffNetB1	99.91	99.58	99.67	100.00	100.00	100.00
	Trans-EffNetB2	99.95	99.58	99.83	100.00	100.00	100.00
	Trans-EffNetB3	99.50	99.58	99.50	100.00	100.00	99.00

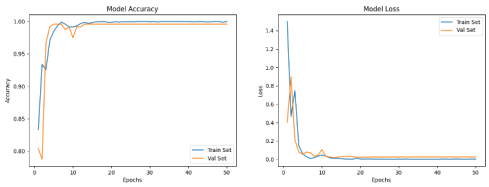

Fig. 8. Accuracy and loss curve of Trans-EffNetB2.

hensive evaluation, we used key metrics such as accuracy, precision, recall, and F1-Score.

Our model employs a hybrid approach, integrating EfficientNet with a Transformer encoder to enhance learning from limited data samples. The performance of our model, particularly the EfficientNetB1 and B2 variants with a Transformer, is outstanding, achieving a test accuracy of 99.49% on the Kaggle dataset for multi-class classification and 99.83% on the BR35H dataset for binary classification.

In comparison to existing methods, our model significantly outperforms previous approaches. For instance, on the Kaggle dataset, traditional CNN-based models reported accuracies of 98.12% [26], 98.10% [27], and 95.70% [28], while our EfficientNetB1 with Transformer achieved a superior accuracy of 99.49%. Similarly, on the BR35H dataset, other methods like GoogleNet, CapsNet, Random Forest, and CNN reported accuracies of 97.10% [29], 90.89% [30], 86.00% [31], and 84.19% [32], respectively. In contrast, our EfficientNetB2 with Transformer model achieved an exceptional accuracy of 99.83%.

Our hybrid Trans-EffNet model, which integrates EfficientNet and Transformer encoder, enhances feature extraction and captures long-range dependencies, overcoming CNN limitations and enabling strong performance on smaller datasets. Our model's computational efficiency suits clinical settings with limited resources, achieving high precision and recall across all tumor types, demonstrating robustness and reliability. It sets a new benchmark in brain tumor classification, showing promise for clinical diagnostics.

F. Grad-CAM Visualization

Grad-CAM (Gradient-weighted Class Activation Mapping) visually illustrates how Trans-EffNetB1 and B2 identify brain tumor categories from MRI images [33]. The Gradients of the predicted class to the feature maps indicate each neuron's significance. These feature maps are weighted by the gradients, aggregated, and passed through a ReLU activation function, resulting in a heatmap that emphasizes regions contributing significantly to the model's predictions. The heatmap is overlaid on the original MRI image, showing the tumor regions relevant to the prediction.

Figures 9 demonstrate Grad-CAM visualizations for each brain tumor category, with the top row displaying feature attention maps, the middle row showing original MRI images, and the bottom row showing model predictions. This provides an intuitive understanding of the model's decision-making process.

TABLE III
PROPOSED MODELS COMPARISON WITH STATE-OF-THE-ART METHODS ON KAGGLE AND BR35H DATASETS.

Dataset	Study	Method	Accuracy (%)
Kaggle	[26]	PDCNN	98.12
	[27]	PatchResNet	98.10
	[28]	WBM-DLNets	95.70
	[20]	RanMerFormer	98.86
	Our	**Trans-EffNEtB1**	**99.49**
Br35H	[29]	deep CNN-SVM	97.10
	[30]	CapsNet	90.89
	[31]	Random Forest	86.0
	[32]	CNN	84.19
	Our	**Trans-EffNEtB2**	**99.83**

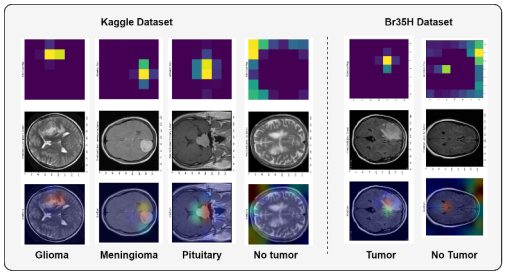

Fig. 9. Grad-Cam Visualization tumor types in Kaggle and Br35H Datasets.

IV. Conclusion

The proposed approach employs rigorous experimental settings, incorporating advanced preprocessing and data augmentation techniques. These enhancements have enabled the Trans-EffNet models to achieve outstanding performance metrics, surpassing existing benchmarks in brain tumor classification. Specifically, Trans-EffNetB1 performed exceptionally well on the multi-class dataset, achieving accuracy, precision, recall, and F1-score metrics of 99.49%, 99.51%, 99.57%, and 99.53%, respectively. For the binary classification dataset, Trans-EffNetB2 demonstrated superior performance with accuracy, precision, recall, and F1-score metrics of 99.83%, 100%, 100%, and 100%, respectively. respectively. These results highlight the versatility and robustness of the Trans-EffNet models across both binary and multi-class classification tasks. While the Trans-EffNet models demonstrate strong performance on the MRI datasets, in future work we will explore its scalability to larger, more diverse datasets and investigate further optimization techniques to enhance computational efficiency in real-world clinical applications.

References

[1] C. P. Davis, "Cancer causes, types, treatment, symptoms signs," MedicineNet, September 2019.

[2] World Health Organization, "Cancer," www.who.int/news-room/fact-sheets/detail/cancer.

[3] U. A. Bhatti, M. Huang, H. Neira-Molina, S. Marjan, M. Baryalai, H. Tang, G. Wu, and S. U. Bazai, "Mffcg – multi feature fusion for hyperspectral image classification using graph attention network," Expert Systems with Applications, vol. 229, 2023.

[4] M. Saeed, I. A. Halepoto, S. Khaskheli, and M. Bushra, "Optimization and efficiency analysis of deep learning based brain tumor detection," Mehran University Research Journal of Engineering and Technology, vol. 42, 2023.

[5] N. H. Grieg, L. G. Ries, R. Yancik, and S. I. Rapoport, "Increasing annual incidence of primary malignant brain tumors in the elderly," Journal of the National Cancer Institute, vol. 82, 1990.

[6] J. Yoo, S. J. Yoon, K. H. Kim, I. H. Jung, S. H. Lim, W. Kim, H. I. Yoon, S. H. Kim, K. S. Sung, T. H. Roh, J. H. Moon, H. H. Park, E. H. Kim, C. O. Suh, S. G. Kang, and J. H. Chang, "Patterns of recurrence according to the extent of resection in patients with idh-wild-type glioblastoma: a retrospective study," Journal of Neurosurgery, vol. 137, 2022.

[7] U. A. Bhatti, M. Huang, D. Wu, Y. Zhang, A. Mehmood, and H. Han, "Recommendation system using feature extraction and pattern recognition in clinical care systems," Enterprise Information Systems, vol. 13, 2019.

[8] U. A. Bhatti, M. Huang, H. Wang, Y. Zhang, A. Mehmood, and W. Di, "Recommendation system for immunization coverage and monitoring," Human Vaccines and Immunotherapeutics, vol. 14, 2018.

[9] Liu, and K. Zhang, "Applications of machine learning in medicine: Current trends and prospects," 2023.

[10] S. A. Nawaz, J. Li, U. A. Bhatti, M. U. Shoukat, and A. Mehmood, "Advance watermarking algorithm using surf with dwt and dct for ct images," vol. 192, 2020.

[11] U. A. Bhatti, Z. Yu, J. Chanussot, Z. Zeeshan, L. Yuan, W. Luo, S. A. Nawaz, M. A. Bhatti, Q. U. Ain, and A. Mehmood, "Local similarity-based spatial–spectral fusion hyperspectral image classification with deep cnn and gabor filtering," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, 2022.

[12] I. Njeh, L. Sallemi, M. B. Slima, A. B. Hamida, S. Lehericy, and D. Galanaud, "A computer aided diagnosis 'cad' for brain glioma exploration," 2014.

[13] S. K. Shil, F. P. Polly, M. A. Hossain, M. S. Ifthekhar, M. N. Uddin, and Y. M. Jang, "An improved brain tumor detection and classification mechanism," vol. 2017-December, 2017.

[14] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," 2014.

[15] P. Jyothi and A. R. Singh, "Deep learning models and traditional automated techniques for brain tumor segmentation in mri: a review," Artificial Intelligence Review, vol. 56, 2023.

[16] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and classification using machine learning: a comprehensive survey," Complex and Intelligent Systems, vol. 8, 2022.

[17] Y. Zhang, L. Deng, H. Zhu, W. Wang, Z. Ren, Q. Zhou, S. Lu, S. Sun, Z. Zhu, J. M. Gorriz, and S. Wang, "Deep learning in food category recognition," Information Fusion, vol. 98, 2023.

[18] A. Karaman, D. Karaboga, I. Pacal, B. Akay, A. Basturk, U. Nalbantoglu, S. Coskun, and O. Sahin, "Hyper-parameter optimization of deep learning architectures using artificial bee colony (abc) algorithm for high performance real-time automatic colorectal cancer (crc) polyp detection," Applied Intelligence, vol. 53, 2023.

[19] F. Zulfiqar, U. I. Bajwa, and Y. Mehmood, "Multi-class classification of brain tumor types from mr images using efficientnets," Biomedical Signal Processing and Control, vol. 84, 2023.

[20] J. Wang, S. Lu, S. Wang, and Y. Zhang, "RanMerFormer: Randomized vision transformer with token merging for brain tumor classification," Neurocomputing, vol. 573, 2024

[21] A. A. Akinyelu, F. Zaccagna, J. T. Grist, M. Castelli, and L. Rundo, "Brain tumor diagnosis using machine learning, convolutional neural networks, capsule neural networks and vision transformers, applied to mri: A survey," 2022.

[22] S. Ali, J. Li, Y. Pei, R. Khurram, K. ur Rehman, and T. Mahmood, "A comprehensive survey on brain tumor diagnosis using deep learning and emerging hybrid techniques with multi-modal mr image," 2022.

[23] Y. Jiang, Y. Zhang, X. Lin, J. Dong, T. Cheng, and J. Liang, "Swinbts: A method for 3d multimodal brain tumor segmentation using swin transformer," Brain Sciences, vol. 12, 2022.

[24] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," vol. 2019-June, 2019.

[25] S. Bhuvaji et al., "Brain tumor classification (mri) dataset," Available online: https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri, 2020.

[26] T. Rahman and M. S. Islam, "Mri brain tumor detection and classification using parallel deep convolutional neural networks," Measurement: Sensors, vol. 26, 2023.

[27] T. Muezzinoglu, N. Baygin, I. Tuncer, P. D. Barua, M. Baygin, S. Dogan, T. Tuncer, E. E. Palmer, K. H. Cheong, and U. R. Acharya, "Patchresnet: Multiple patch division–based deep feature fusion framework for brain tumor classification using mri images," Journal of Digital Imaging, vol. 36, 2023.

[28] M. U. Ali, S. J. Hussain, A. Zafar, M. R. Bhutta, and S. W. Lee, "Wb-mdlnets: Wrapper-based metaheuristic deep learning networks feature optimization for enhancing brain tumor detection," Bioengineering, vol. 10, 2023.

[29] S. Deepak and P. M. Ameer, "Brain tumor classification using deep cnn features via transfer learning," Computers in Biology and Medicine, vol. 111, 2019.

[30] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on mri images and coarse tumor boundaries," vol. 2019-May, 2019.

[31] M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T. L. Jones, T. R. Barrick, F. A. Howe, and X. Ye, "Supervised learning based multimodal mri brain tumour segmentation using texture features from supervoxels," Computer Methods and Programs in Biomedicine, vol. 157, 2018.

[32] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in World Congress on Medical Physics and Biomedical Engineering 2018, L. Lhotska, L. Sukupova, I. Lackovic, and G. S. Ibbott, Eds. Singapore: Springer Nature Singapore, 2019, pp. 183–189.

[33] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," International Journal of Computer Vision, vol. 128, 2020.