

Combining Local and Global Feature Extraction for Brain Tumor Classification: A Vision Transformer and iResNet Hybrid Model

Amar Y. Jaffar

Computer and Network Engineering Department, College of Computing, Umm Al-Qura University, Makkah, Saudi Arabia

ayjaafar@uqu.edu.sa (corresponding author)

Received: 30 June 2024 | Revised: 24 July 2024, 8 August 2024, and 14 August 2024 | Accepted: 18 August 2024

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.8271>

ABSTRACT

Early diagnosis of brain tumors is crucial for effective treatment and patient prognosis. Traditional Convolutional Neural Networks (CNNs) have shown promise in medical imaging but have limitations in capturing long-range dependencies and contextual information. Vision Transformers (ViTs) address these limitations by leveraging self-attention mechanisms to capture both local and global features. This study aims to enhance brain tumor classification by integrating an improved ResNet (iResNet) architecture with a ViT, creating a robust hybrid model that combines the local feature extraction capabilities of iResNet with the global feature extraction strengths of ViTs. This integration results in a significant improvement in classification accuracy, achieving an overall accuracy of 99.2%, outperforming established models such as InceptionV3, ResNet, and DenseNet. High precision, recall, and F1 scores were observed across all tumor classes, demonstrating the model's robustness and reliability. The significance of the proposed method lies in its ability to effectively capture both local and global features, leading to superior performance in brain tumor classification. This approach offers a powerful tool for clinical decision-making, improving early detection and treatment planning, ultimately contributing to better patient outcomes.

Keywords-brain tumor classification; vision transformers; iResNet; MRI; deep learning; artificial intelligence; medical imaging

I. INTRODUCTION

Brain tumors are a major global health issue, with early and accurate detection being essential for effective treatment. According to the World Health Organization (WHO), brain and Central Nervous System (CNS) tumors account for approximately 1.6% of all cancers and 2.5% of cancer-related deaths worldwide [1]. In 2020, approximately three million new cases of brain and CNS tumors were diagnosed, leading to 251,329 deaths worldwide [2]. The malignant nature and critical brain locations of these tumors make treatment particularly challenging, often resulting in low survival rates and severe impacts on patient quality of life. Timely and accurate brain tumor diagnosis is crucial for better clinical outcomes. Traditionally, the diagnosis and classification of brain tumors have relied on imaging techniques [3], with radiologists interpreting these images to determine the tumor's presence, type, and grade. However, this manual interpretation is both time-consuming and susceptible to variability and subjectivity, which can lead to misdiagnosis or delayed diagnosis, adversely affecting patient treatment plans and outcomes [4]. Recent advances in Deep Learning (DL) techniques have shown great potential in enhancing diagnostic accuracy and efficiency in medical imaging diagnostics. DL

techniques have achieved remarkable performance in tasks such as tumor detection and classification [5]. Convolutional Neural Networks (CNNs) have been pivotal in many successful AI models due to their ability to automatically learn and extract hierarchical features from medical images [6]. Despite their success, CNNs have limitations in capturing long-range dependencies and contextual information in images, which are vital for accurate tumor classification [7].

The advent of Vision Transformers (ViTs) represents a significant breakthrough in AI-driven medical imaging [8]. ViTs have demonstrated exceptional performance in image classification tasks. Unlike CNNs, which rely on convolutional layers to extract local features, ViTs use self-attention mechanisms to capture global dependencies and contextual information throughout the image. This capability to model long-range relationships makes ViTs particularly suitable for complex medical imaging tasks, such as brain tumor classification, where subtle differences in tumor morphology and texture are critical for accurate diagnosis. Despite advances in CNNs and ViTs, there is still a need for models that effectively combine the strengths of both architectures. Current models often fail to achieve the desired level of accuracy and generalizability because they cannot fully capture both local

and global features. Addressing this gap is essential as it can significantly enhance diagnostic accuracy and clinical decision-making in brain tumor management.

This study proposes a robust and accurate model for brain tumor classification by integrating an improved ResNet (iResNet) architecture with a ViT pathway. Specific objectives are to enhance classification accuracy by leveraging the complementary strengths of iResNet and ViTs and evaluate the sensor fusion-based architecture's performance on a comprehensive dataset. Figure 1 shows the steps involved in the proposed architecture. The key contributions are as follows:

- Novel integration: Develop a hybrid model combining iResNet and ViTs to effectively capture both local and global features for improved brain tumor classification.
- Enhanced accuracy: Achieve superior classification accuracy compared to established models, thus supporting more accurate and timely clinical decision-making.
- Comprehensive evaluation: Detailed performance evaluation on a comprehensive dataset to assess the robustness and reliability of the proposed model.

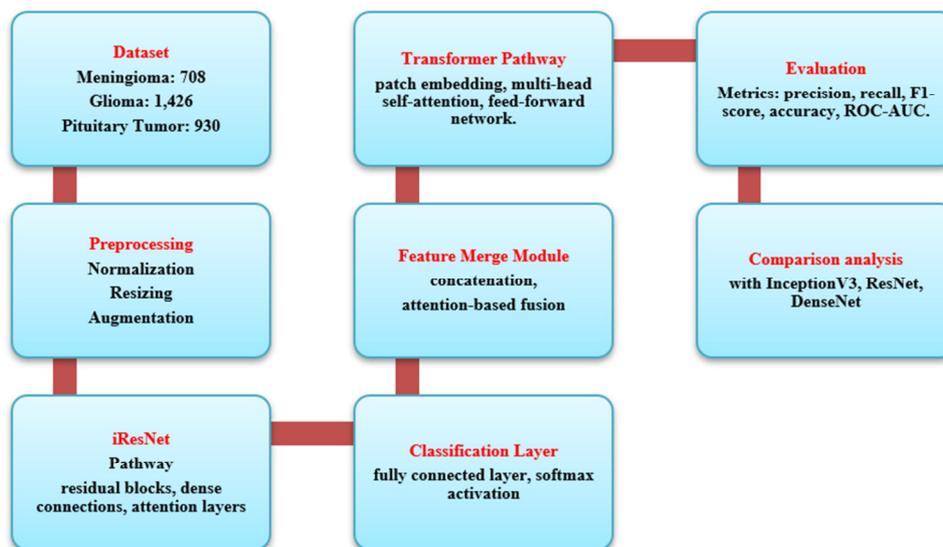


Fig. 1. Workflow of the proposed sensor fusion-based approach.

II. RELATED WORKS

Recent advances in medical image analysis, particularly through the use of DL techniques, have significantly improved disease identification. Many CNN-based models have been proposed for multiclass brain tumor classification and detection, employing methods such as custom CNN models and transfer learning strategies. In [9], the ResNet-34 CNN model achieved high accuracy in classifying meningioma, glioma, and pituitary tumors. Similarly, the role of pre-trained architectures in the classification of benign and malignant brain tumors has been highlighted, achieving notable accuracy [10]. Transfer learning techniques using pre-trained CNN architectures, such as ResNet101, ResNet50, GoogleNet, AlexNet, and SqueezeNet, for brain tumor classification have also shown promising results [11], with AlexNet being particularly effective. Additionally, enhanced CNN model accuracy has been achieved using data augmentation and edge detection techniques on MRI datasets [12]. In [13], a CNN-based Computer-Assisted Diagnosis (CAD) method was introduced for brain tumor classification, achieving high accuracy across multiple datasets. Furthermore, CNNs have been used to predict tumors without expert annotations, achieving significant accuracy using both entire brain data and tumor Region Of Interest (ROI) data [14]. Focusing on axial-

plane data for brain tumor classification using CNNs has also demonstrated high accuracy rates [15]. Various CNN-based models have been employed to diagnose brain diseases, addressing the convergence time constraints of Artificial Neural Networks (ANNs), with some models achieving remarkable accuracy [16]. In [17], basic CNN models with varying numbers of layers were explored, finding that a two-layer architecture achieved the highest training accuracy. In [18], a model was proposed using 3D image reconstruction and DenseNet for feature extraction, achieving notable accuracy.

In [19], an improved Capsule Neural Network (CapsNet) was proposed for brain tumor classification, eliminating the need for precise tumor annotations and achieving high classification accuracy. In [20], DenseNet201 with transfer learning and genetic algorithms was proposed to classify the severity of brain tumors in the BraTS2018 and BraTS2019 datasets, achieving high accuracy. In [21], machine learning and DL methods were combined using GoogLeNet with transfer learning to extract features, which were then classified using SVM, KNN, and softmax classifiers for brain tumor classification. In [22], a comparative analysis of various deep learning models, including ResNet50, AlexNet, VGG16, and GoogleNet for brain tumor classification, showed that ResNet50 achieved the highest accuracy.

In [23], ViT-based ensemble architectures were used for brain tumor classification, achieving high accuracy rates. In [24], ViT combined with deep anchor attention learning showed effectiveness in brain tumor classification. Various studies have explored ViT-based architectures for effective brain tumor classification, yielding promising results [25, 26]. Several studies have also investigated the use of ViTs in the segmentation of brain tumors, highlighting their potential to improve classification performance [27-29]. Despite these advances, the implementation of CNN and ViT-based architectures for brain tumor classification remains incomplete. Thus, in [30], a hybrid TECNN model was proposed to achieve more accurate brain tumor classification.

III. METHODOLOGY

A. iResNet Architecture

The iResNet architecture is a modified version of the Residual Network (ResNet), tailored to enhance feature extraction for medical imaging tasks. ResNet introduced the concept of residual learning, which helps in training deeper networks by addressing the vanishing gradient problem. The core idea of ResNet is the residual block, defined as:

$$y = \mathcal{F}(x, \{W_i\}) + x \quad (1)$$

where $\mathcal{F}(x, \{W_i\}) + x$ represents the residual function learned by the residual block, x is the input, and y is the output of the block. The identity mapping x allows gradients to flow through the network more effectively. In iResNet, several modifications are made to adapt the architecture for brain tumor classification. The architecture includes multiple layers of residual blocks, each consisting of convolutional layers followed by batch normalization and ReLU activation functions. The modifications include:

- Increased depth and width: Additional layers and wider residual blocks to capture more detailed features from brain MRI images.
- Dense connections within the residual blocks to facilitate better gradient flow and feature reuse.
- Attention layers that focus on relevant regions of the images to enhance the discriminative power of the network.

Thus, the iResNet architecture ensures that both local and global features are effectively captured, providing a strong foundation for the subsequent transformer pathway.

B. Transformer Pathway

The transformer pathway leverages a self-attention mechanism to model long-range dependencies and contextual information in brain MRI images. Its architecture can be divided into the following key components.

1) Patch Embedding Layer

The input image is divided into fixed-size patches, each flattened into a 1D vector. This process is represented as:

$$z_0 = [x_1 E; x_2 E; \dots; x_N E] + E_{pos} \quad (2)$$

where x_i represents the i -th image patch, E is the embedding matrix, and E_{pos} is the positional encoding matrix. The self-attention structure allows the model architecture to evaluate the significance of different patches relative to each other. The attention scores are calculated using

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (3)$$

where Q (queries), K (keys), and V (values) are linear transformations of the input patches, and d_K is the dimensionality of the key vectors.

2) Multi-Head Attention

Multi-head attention is employed to capture diverse features, where multiple attention heads operate in parallel:

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_n)W^O \quad (4)$$

where each $head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$.

3) Feed-Forward Network

Following the attention layers, a feed-forward neural network is applied to each patch as

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (5)$$

4) Layer Normalization and Residual Connections

Each attention and feed-forward block is followed by layer normalization and residual connections to stabilize training and improve gradient flow. The overall architecture of the transformer pathway can be summarized as follows:

- Input embedding: Patch embedding and positional encoding.
- Transformer encoder: Stacked layers of multi-head self-attention and feed-forward networks.
- The output features from the transformer encoder are used for further processing.

C. Feature Merge Module

The Feature Merge (FM) module plays a crucial role in integrating the complementary strengths of the iResNet and Transformer pathways to achieve a highly accurate and robust brain tumor classification model. The FM module begins by concatenating the feature representations obtained from the iResNet and Transformer pathways. This step ensures that the local features (extracted by iResNet) and the global context (captured by the Transformer) are combined into a single comprehensive feature vector.

IV. FEATURE EXTRACTION AND FUSION

Let $F_{iResNet}$ denote the feature vector from the iResNet pathway and $F_{Transformer}$ denote the feature vector from the transformer pathway. The combined feature vector $F_{combined}$ is obtained by

$$F_{combined} = Concat(F_{iResNet}, F_{Transformer}) \quad (6)$$

where $Concat(\cdot)$ represents the concatenation operation that merges the two feature vectors along the feature dimension. An attention mechanism is employed to effectively weigh the importance of different features from the combined feature vector. The attention mechanism allows the model to focus on the most relevant features from both pathways, enhancing the overall discriminative power of the model. The attention weights (α) are computed using a learned weight matrix (W_a) and a bias term (b_a):

$$\alpha = \text{softmax}(W_a F_{combined} + b_a) \quad (7)$$

The softmax function ensures that the attention weights sum to one, providing a probability distribution over the combined feature vector. The element-wise multiplication (\odot) of the attention weights and the combined feature vector yields the fused feature representation (F_{fused}):

$$F_{fused} = \alpha \odot F_{combined} \quad (8)$$

This process ensures that the features contributing most to the classification task are emphasized, while less relevant features are down-weighted.

A. Attention Mechanism Details

The attention mechanism operates as follows:

- Transformation of combined features: The combined feature vector $F_{combined}$ is transformed using a linear layer to produce a new representation, which is then passed through a non-linear activation function (e.g., ReLU) to introduce non-linearity as represented by

$$H = \text{ReLU}(W_h F_{combined} + b_h) \quad (9)$$

- Computation of attention scores: The transformed feature vector H is further processed using another linear layer to compute the attention scores by

$$Scores = W_s H + b_s \quad (10)$$

- Softmax normalization: The attention scores are normalized using the softmax function to obtain the attention weights α

$$\alpha = \text{Softmax}(Scores) \quad (11)$$

- Feature fusion: The attention weights are applied to the combined feature vector to obtain the fused feature representation F_{fused} .

- Classification layer: After obtaining the fused feature representation F_{fused} , the next step is to classify the brain tumor into one of the predefined categories (meningioma, glioma, or pituitary tumor). This is achieved by passing the fused features through a fully connected layer followed by a softmax activation function. The fully connected layer transforms the fused features into logits, as in (12), which represent the raw class scores.

$$logits = W_c F_{fused} + b_c \quad (12)$$

- Finally, the softmax function converts the logits into class probabilities, providing the likelihood of each tumor type

$$\hat{y} = \text{softmax}(W_c F_{fused} + b_c) \quad (13)$$

where \hat{y} represents the predicted class probabilities for the input brain MRI image.

This approach ensures that the most relevant features from both the iResNet and transformer pathways are utilized, improving the overall classification performance. The FM module offers several advantages by combining local and global features, providing a richer and more comprehensive representation of the input image and capturing both fine-grained details and overall context. The attention mechanism selectively emphasizes the most relevant features, improving the model's ability to distinguish between different tumor types. The fusion of features from two complementary pathways ensures that the model is robust to variations in the input data and generalizes well to new, unseen images.

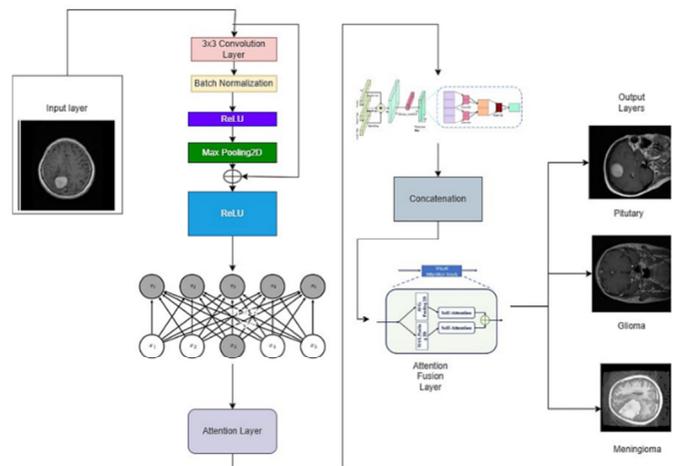


Fig. 2. Hybrid architecture of ViT and iResNet for brain tumor classification.

V. DATASET

Figure 3 shows a sample from the dataset used [33].

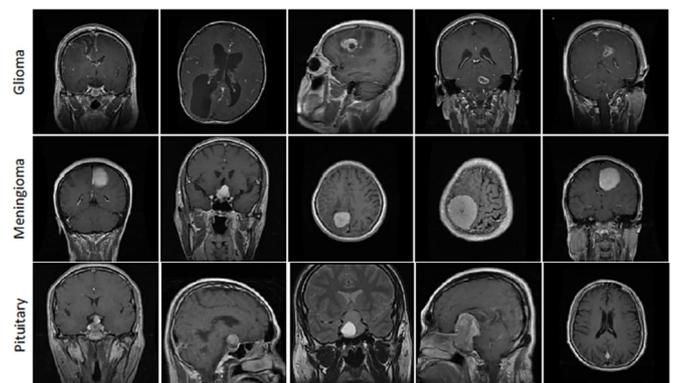


Fig. 3. Sample images of the dataset [33].

The images are organized into corresponding directories for each tumor type, facilitating easy access and processing. This dataset provides a diverse set of samples that capture the variability in tumor appearance and structure. High-quality, labeled images serve as a solid foundation for training and

evaluating the brain tumor classification model. The use of a comprehensive dataset ensures the reliability and generalizability of the model, ultimately contributing to improved diagnostic accuracy and clinical decision-making in brain tumor management.

VI. RESULTS AND DISCUSSION

The proposed model showed superior performance, achieving an overall accuracy of 99.2%. Table I presents the evaluation results of the model. The model was evaluated using both training and validation datasets. The performance metrics show that the proposed model consistently achieved high scores in all tumor categories, highlighting its robustness and reliability in accurately classifying brain tumors. The performance metrics were calculated using:

$$Precision = \frac{TP}{TP+FP} \tag{14}$$

$$Recall = \frac{TP}{TP+FN} \tag{15}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{16}$$

The high precision and recall values for all three tumor types show that the model is not only able to accurately classify a high proportionality of true positive events but also maintain a low rate of false positives. The F1 score, which balances precision and recall, further confirms the model's robustness in the classification task. The confusion matrix, shown in Figure 4, indicates that the proposed model has an extremely high correct classification rate for each tumor type, with minimal misclassifications. This confirms the model's effectiveness in distinguishing between the three tumor categories.

TABLE I. RESULTS OF THE PROPOSED MODEL FOR EACH TUMOR CLASS

Classes	Precision	Recall	F1-score	Accuracy
Pituitary tumor	0.97	0.98	0.987	0.97
Meningioma	0.98	0.98	0.992	0.97
Glioma	0.98	0.98	0.993	0.97
Average	0.976667	0.98	0.990667	0.97

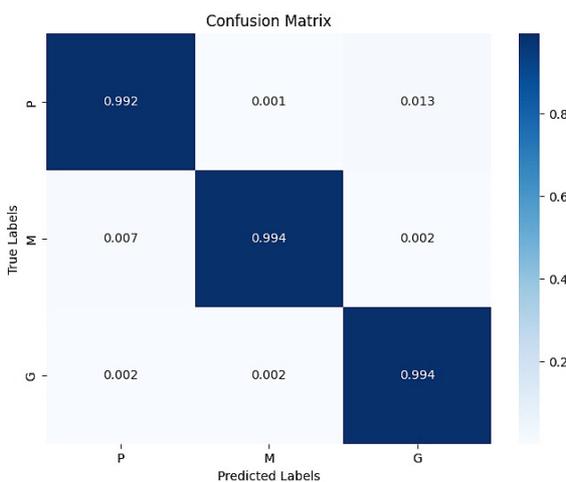


Fig. 4. Confusion matrix for the proposed model.

The proposed hybrid model was compared with several well-established models, including InceptionV3, ResNet, and DenseNet. The comparison results, shown in Table II, indicate that the proposed model substantially surpasses them, achieving an accuracy of 99.2%. This enhancement is attributed to the combined strengths of the iResNet architecture, the Transformer pathway, and the FM module. Figures 5, 6, and 7 show the evaluation results of the DenseNet, ViT, and the proposed hybrid sensor fusion-based transformer model on the dataset.

TABLE II. COMPARISON OF MODEL ACCURACIES

Model	Accuracy	Reference
InceptionV3	85.1	[29]
ResNet	88.4	[30]
DenseNet	93.4	[31]
Proposed model	99.2	This study

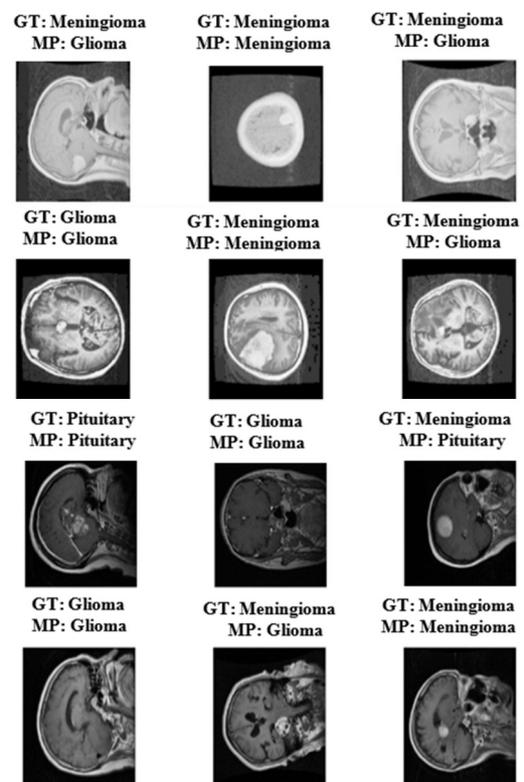


Fig. 5. Evaluation outcomes of the DenseNet model on the dataset.

The results highlight the efficacy of the proposed hybrid model. Enhanced feature extraction is one of the key factors contributing to its superior performance. The iResNet architecture effectively captures local features, while the transformer pathway captures global contextual information, resulting in a comprehensive feature representation. The DenseNet model achieved an accuracy of 93.4%. The model showed good performance but struggled with some complex cases of glioma and meningioma, misclassifying them as pituitary tumors as shown in Figure 5. This indicates limitations in capturing global contextual features. The standalone ViT model achieved better accuracy (94.8%),

demonstrating better performance in capturing long-range dependencies and contextual information. However, it still faced challenges with certain tumor classifications. Although the ViT model performed better than DenseNet, it occasionally misclassified gliomas as pituitary tumors, as shown in Figure 6, indicating a need for better local feature extraction.

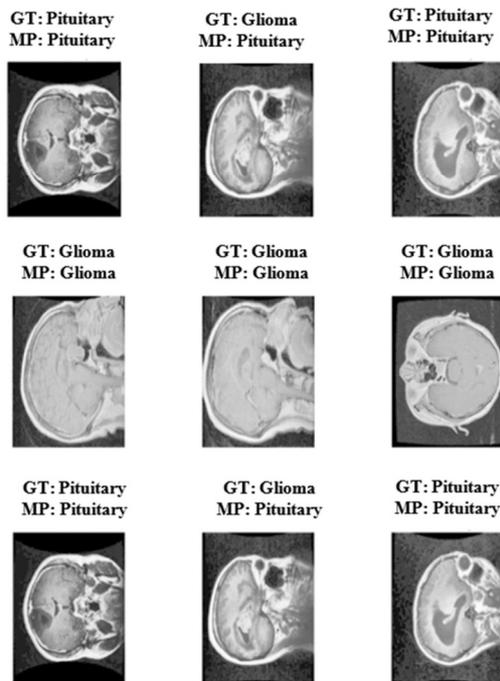


Fig. 6. ViT model Glioma misclassification.

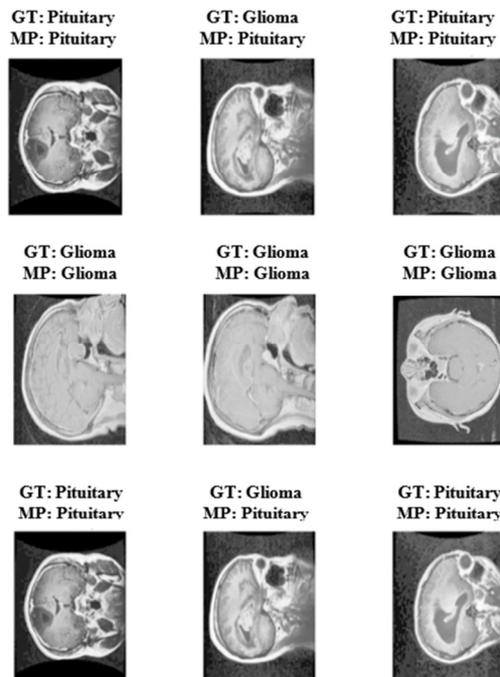


Fig. 7. Evaluation results of the proposed hybrid model.

Figure 7 shows some results of the proposed hybrid model. This model effectively captured both local and global features, resulting in high precision, recall, and F1 scores across all tumor classes. Attention-based fusion in the FM module ensures that the most relevant features are emphasized, enhancing the model's discriminative power. The proposed hybrid model demonstrated exceptional performance, correctly classifying complex tumor cases that were misclassified by the DenseNet and ViT models. Integration of iResNet ensured robust local feature extraction, while the ViT component captured global dependencies, leading to a comprehensive and accurate tumor classification. In addition, the statistical tests confirmed that the performance improvements of the proposed hybrid model were significant ($p < 0.05$) compared to the DenseNet and ViT models.

The model's high Area Under the Curve (AUC) values further reinforced its diagnostic accuracy. The ROC curve for each of the three classes in the dataset is positioned in the top left corner of the plot, indicating perfect classification. This placement shows that the model achieved 100% sensitivity with no false positives for any class. The AUC for each class is 1.00, demonstrating flawless discrimination. An AUC of 1.00 means that the model differentiates perfectly between positive and negative cases. On the ROC curve, the y-axis (True Positive Rate or Sensitivity) characterizes the percentage of true positives properly detected, with a value of 1.0 indicating complete identification of all true positives. The x-axis (False Positive Rate or Specificity) shows the proportion of actual negatives incorrectly classified as positives, with a value of 0.0 indicating the absence of misclassification of negative cases. The diagonal line represents the performance of a random classifier, where the True Positive rate equals the False Positive rate. With an AUC of 1.00 for each class, the model exhibited perfect sensitivity and specificity, making it a highly reliable tool for clinical decision-making in brain tumor diagnosis and treatment planning.

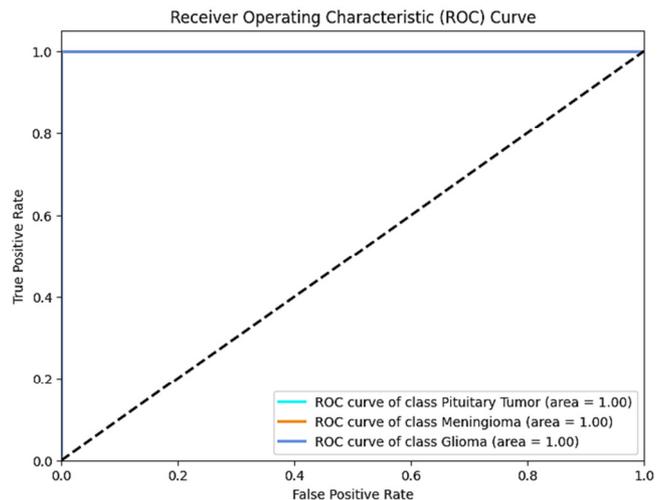


Fig. 8. ROC curve for the proposed hybrid model.

The novelty of this work lies in the innovative integration of the iResNet architecture with ViTs to create a hybrid model

that takes advantage of the strengths of both approaches. Unlike traditional CNNs that primarily capture local features, the proposed hybrid model effectively combines local feature extraction through iResNet with global feature extraction through the self-attention mechanisms of ViT. This unique combination allows the model to capture comprehensive features from brain MRI images, leading to improved classification accuracy. The FM module, which seamlessly integrates local and global features, further enhances the model's discriminative power. This approach not only surpasses the performance of established models but also demonstrates robustness and reliability across various tumor types, making it a significant advancement in AI-driven medical imaging for brain tumor classification.

VII. CONCLUSION

This study showed that the integration of iResNet and ViT significantly enhances the accuracy and reliability of brain tumor classification. The proposed hybrid model, which leverages the local feature extraction capabilities of iResNet and the global contextual understanding of ViTs, achieved an accuracy of 99.2%, outperforming established models such as InceptionV3, ResNet, and DenseNet. The significance of this method lies in its ability to capture both fine-grained details and broader contextual information, resulting in superior classification performance across all tumor categories. This dual capability addresses a critical gap in existing models, which often struggle to balance local and global feature extraction effectively. By incorporating a feature merge module with an attention mechanism, the proposed hybrid model emphasizes the most relevant features, thus enhancing its discriminative power. This novel approach not only improves diagnostic accuracy but also supports more precise and timely clinical decision-making, ultimately contributing to better patient outcomes. Future research should focus on expanding this method to include additional imaging modalities, such as CT scans, and exploring its potential for real-time implementation in clinical settings. Furthermore, the robustness and generalizability of the model should be validated on larger and more diverse datasets to ensure its broader applicability. In general, the successful application of ViTs in combination with iResNet highlights the potential of advanced AI technologies to revolutionize medical imaging and diagnostics, paving the way for more accurate and reliable brain tumor classification.

REFERENCES

- [1] T. Gupta *et al.*, "Comparison of Epidemiology and Outcomes in Neuro-Oncology Between the East and the West: Challenges and Opportunities," *Clinical Oncology*, vol. 31, no. 8, pp. 539–548, Aug. 2019, <https://doi.org/10.1016/j.clon.2019.05.018>.
- [2] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021, <https://doi.org/10.3322/caac.21660>.
- [3] H. A. Jalab and A. M. Hasan, "Magnetic Resonance Imaging Segmentation Techniques of Brain Tumors: A Review," *Archives of Neuroscience*, vol. 6, Jan. 2019, Art. no. e84920, <https://doi.org/10.5812/ans.84920>.
- [4] V. Vidhya *et al.*, "Automated Detection and Screening of Traumatic Brain Injury (TBI) Using Computed Tomography Images: A Comprehensive Review and Future Perspectives," *International Journal of Environmental Research and Public Health*, vol. 18, no. 12, Jan. 2021, Art. no. 6499, <https://doi.org/10.3390/ijerph18126499>.
- [5] A. Krauze, Y. Zhuge, R. Zhao, E. Tasci, and K. Camphausen, "AI-Driven Image Analysis in Central Nervous System Tumors-Traditional Machine Learning, Deep Learning and Hybrid Models," *Journal of biotechnology and biomedicine*, vol. 5, no. 1, pp. 1–19, 2022, <https://doi.org/10.26502/jbb.2642-91280046>.
- [6] T. Imran, A. S. Alghamdi, and M. S. Alkathiri, "Enhanced Skin Cancer Classification using Deep Learning and Nature-based Feature Optimization," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12702–12710, Feb. 2024, <https://doi.org/10.48084/etasr.6604>.
- [7] S. Tripathi, S. K. Singh, and H. K. Lee, "An end-to-end breast tumour classification model using context-based patch modelling – A BiLSTM approach for image classification," *Computerized Medical Imaging and Graphics*, vol. 87, Jan. 2021, Art. no. 101838, <https://doi.org/10.1016/j.compmedimag.2020.101838>.
- [8] S. Singh, M. Kumar, A. Kumar, B. K. Verma, K. Abhishek, and S. Selvarajan, "Efficient pneumonia detection using Vision Transformers on chest X-rays," *Scientific Reports*, vol. 14, no. 1, Jan. 2024, Art. no. 2487, <https://doi.org/10.1038/s41598-024-52703-2>.
- [9] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multi-class brain disease detection using MRI images," *Computerized Medical Imaging and Graphics*, vol. 78, Dec. 2019, Art. no. 101673, <https://doi.org/10.1016/j.compmedimag.2019.101673>.
- [10] H. Bingol and B. Alatas, "Classification of Brain Tumor Images using Deep Learning Methods," *Turkish Journal of Science and Technology*, vol. 16, no. 1, pp. 137–143, Mar. 2021.
- [11] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A Transfer Learning approach for AI-based classification of brain tumors," *Machine Learning with Applications*, vol. 2, Dec. 2020, Art. no. 100003, <https://doi.org/10.1016/j.mlwa.2020.100003>.
- [12] M. Aamir *et al.*, "A deep learning approach for brain tumor classification using MRI images," *Computers and Electrical Engineering*, vol. 101, Jul. 2022, Art. no. 108105, <https://doi.org/10.1016/j.compeleceng.2022.108105>.
- [13] W. Ayadi, W. Elhamzi, I. Charfi, and M. Atri, "Deep CNN for Brain Tumor Classification," *Neural Processing Letters*, vol. 53, no. 1, pp. 671–700, Feb. 2021, <https://doi.org/10.1007/s11063-020-10398-2>.
- [14] X. Li, H. Chen, X. Qi, Q. Dou, C. W. Fu, and P. A. Heng, "H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation From CT Volumes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, Sep. 2018, <https://doi.org/10.1109/TMI.2018.2845918>.
- [15] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, "Deep learning for brain tumor classification," in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Mar. 2017, vol. 10137, pp. 253–268, <https://doi.org/10.1117/12.2254195>.
- [16] G. Wang *et al.*, "Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1562–1573, Jul. 2018, <https://doi.org/10.1109/TMI.2018.2791721>.
- [17] 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*, Prague, Czech Republic, 2019, pp. 183–189, https://doi.org/10.1007/978-981-10-9035-6_33.
- [18] Z. Zhong, M. Zheng, H. Mai, J. Zhao, and X. Liu, "Cancer image classification based on DenseNet model," *Journal of Physics: Conference Series*, vol. 1651, no. 1, Aug. 2020, Art. no. 012143, <https://doi.org/10.1088/1742-6596/1651/1/012143>.
- [19] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain Tumor Type Classification via Capsule Networks," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, Jul. 2018, pp. 3129–3133, <https://doi.org/10.1109/ICIP.2018.8451379>.
- [20] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based

- biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021, <https://doi.org/10.1038/s41592-020-01008-z>.
- [21] A. Sekhar, S. Biswas, R. Hazra, A. K. Sunaniya, A. Mukherjee, and L. Yang, "Brain Tumor Classification Using Fine-Tuned GoogLeNet Features and Machine Learning Algorithms: IoMT Enabled CAD System," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 983–991, Mar. 2022, <https://doi.org/10.1109/JBHI.2021.3100758>.
- [22] S. Kuraparathi *et al.*, "Brain Tumor Classification of MRI Images Using Deep Convolutional Neural Network," *Traitement du Signal*, vol. 38, no. 4, pp. 1171–1179, Aug. 2021, <https://doi.org/10.18280/ts.380428>.
- [23] S. Tummala, S. Kadry, S. A. C. Bukhari, and H. T. Rauf, "Classification of Brain Tumor from Magnetic Resonance Imaging Using Vision Transformers Ensembling," *Current Oncology*, vol. 29, no. 10, pp. 7498–7511, Oct. 2022, <https://doi.org/10.3390/curroncol29100590>.
- [24] X. Xu and P. Prasanna, "Brain Cancer Survival Prediction on Treatment-Naïve MRI using Deep Anchor Attention Learning with Vision Transformer," in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, Mar. 2022, pp. 1–5, <https://doi.org/10.1109/ISBI52829.2022.9761515>.
- [25] S. Nallamolu, H. Nandanwar, A. Singh, and C. N. Subalalitha, "A CNN-based Approach for Multi-Classification of Brain Tumors," in *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet, India, Aug. 2022, pp. 1–6, <https://doi.org/10.1109/ASIANCON55314.2022.9908994>.
- [26] J. P. Dou *et al.*, "Research progress of quantum memory," *Acta Physica Sinica*, vol. 68, no. 3, 2019, Art. no. 030307, <https://doi.org/10.7498/aps.68.20190039>.
- [27] M. F. I. Soumik and M. A. Hossain, "Brain Tumor Classification With Inception Network Based Deep Learning Model Using Transfer Learning," in *2020 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, Jun. 2020, pp. 1018–1021, <https://doi.org/10.1109/TENSYMP50017.2020.9230618>.
- [28] Ö. Polat and C. Güngen, "Classification of brain tumors from MR images using deep transfer learning," *The Journal of Supercomputing*, vol. 77, no. 7, pp. 7236–7252, Jul. 2021, <https://doi.org/10.1007/s11227-020-03572-9>.
- [29] A. Raza, M. S. Alshehri, S. Almakdi, A. A. Siddique, M. Alsulami, and M. Alhaisoni, "Enhancing brain tumor classification with transfer learning: Leveraging DenseNet121 for accurate and efficient detection," *International Journal of Imaging Systems and Technology*, vol. 34, no. 1, 2024, Art. no. e22957, <https://doi.org/10.1002/ima.22957>.
- [30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, pp. 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
- [31] J. Cheng, "Brain tumor dataset." Figshare, 2017, <https://doi.org/10.6084/M9.FIGSHARE.1512427>.