MB-MSAT-Net: A Multi-Branch Multi-Scale Attention Framework with EFO and TabNet for Accurate and Interpretable Brain Tumor Classification

Jayaprakasam Kandasamy[1] , Mahendran Gandhi[2*]

[1] Department of Biomedical Engineering, Sethu Institute of Technology, Kariyapatty 626115, India
[2] Department of Electronics and Communication Engineering, Syed Ammal Engineering College, Ramanathapuram 623502, India

Corresponding Author Email: gmahendran@syedengg.ac.in

ABSTRACT

In recent times, many people have been affected by a brain tumour which increases the mortality rate day by day. To reduce this mortality, Brain tumor classification is performed for an earlier diagnosis and treatment planning for patients with neurological conditions. In this proposed system, an advanced brain tumor classification is presented by integrating hybrid methodologies such as Multi-Branch Multi-Scale Attention Transformer Network (MB-MSAT-Net) for feature extraction, Electric Fish Optimization (EFO) for feature selection, and TabNet for classification. The proposed MB-MSAT-Net is designed to capture spatial and contextual information at multiple scales that effectively extract discriminative features from medical imaging data. The EFO technique is an optimization method that is used to select relevant and significant features to improve model performance and minimise computational complexity. At last, the classification is done by TabNet which is used to classify the tumor types based on the selected features. This enhanced classification achieved higher accuracy and transparency. The result was evaluated using publicly available brain tumor datasets that were used to validate both the proposed and conventional methods. This proposed hybrid model attained a better classification accuracy and robustness in brain tumor classification than the existing methods. For future enhancement, this method can hold greater promising tool to automate a brain tumor diagnosis for medical professionals.

1. INTRODUCTION

In recent years, brain tumors have been a major deadly disease occurring in large numbers with an enormous mortality rate [1]. These tumors are abnormal growths of cells in the brain. They can lead to neurological symptoms such as headaches, seizures, cognitive dysfunction, and motor impairments. These tumors are classified into two types, benign and malignant, and they are also categorized based on their cellular structure, location, and aggressiveness [2, 3]. The main cause of these tumors is not fully understood, but some of the reasons behind them include genetic mutations, environmental factors, and family history.

To overcome these tumors, early detection is vital which can improve patient outcomes significantly to enable timely treatment and intervention. However, brain tumors prediction is performed manually in traditional methods which consumes more time and is error-prone. Advanced computational methods are used to predict the brain tumor types that can streamline diagnosis, improve accuracy, and personalized treatment planning.

Mostly, Image-based brain tumor detection plays a significant role among the various detection methods. Because image detection has a non-invasive behaviour and the ability to capture high-resolution images of brain structures. A few

Image modalities like Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET) are common methods used to visualize and diagnose brain tumors [4]. However, among the three modalities, MRI is the most used imaging modality because of its superior soft-tissue contrast. It can be able to capture detailed images of brain tissue, and the absence of harmful ionizing radiation to attain a safer and more effective modality in tumor.

The process of brain tumor detection using imaging data typically involves several key steps: Preprocessing that involves skull stripping, normalization, and resizing. Feature Extraction is used to extract a relevant feature to represent the tumor's texture, shape, intensity, and spatial relationships within the image [5]. Next, Classification is used to classify the tumor as benign or malignant or to identify specific tumor types. To process a feature extraction and classification, machine learning (ML) or deep learning (DL) models like convolutional neural networks (CNNs) and transformer-based models are employed [6, 7]. It has shown exceptional performance in complex medical image classification to learn data hierarchical representations.

Moreover, many existing approaches failed to consider the interpretability in a clinical setting where decisions need to be explained to medical professionals [8]. The previous method

also provided an inefficiency in feature selection where a large number of extracted features often leads to redundancy that increases complexity and overfitting. Some Feature selection techniques like statistical methods do not address the complex relationships between features effectively. It caused a lack of adaptive and intelligent feature selection to hinder the optimization of the model. It seems that current methods have several limitations that limit their practical utility in real-world medical diagnosis.

To address all these limitations, this research presented an advanced brain tumor classification with three innovative models. Firstly, the Multi-Branch Multi-Scale Attention Transformer Network (MB-MSAT-Net) is used as a feature extraction that has multiple convolutional blocks with varied kernel sizes to capture both local and global features. It also has an attention mechanism at different scales where MB-MSAT-Net is used to focus on the most critical regions in brain images. Secondly, Electric Fish Optimization (EFO) is used for selecting features that are inspired by the electric fish character. It minimised the computational complexity and mitigated overfitting by improving the model's ability. Finally, TabNet is employed for classification to handle tabular data by offering high interpretability. This TabNet contains a decision tree mechanism to process an accurate prediction for medical applications. This research attained greater accuracy, reduced computational demands, and enhanced transparency in brain tumor diagnosis for healthcare professionals.

2. RELATED WORK

Islam et al. [9] developed a multifractional Brownian motion (mBm) which is a stochastic model that used to extract multifractal features and contains an enhanced AdaBoost algorithm for patient-independent tumor segmentation. Experimental results demonstrate an improved performance using the BRATS2012 dataset with superior segmentation robustness and accuracy.

Saeedi et al. [10] presented an Inception-v3 and DenseNet201 models for feature extraction. These Concatenated features are classified using a softmax classifier. This method achieved a higher accuracy of 99.34% and 99.51%, respectively. Gumaei et al. [11] explored a Regularized Extreme Learning Machine (RELM) to classify brain tumor. The result showed that this method utilised a new public dataset to improve classification accuracy from 91.51% to 94.23% than an existing approach.

Bibi et al. [12] developed an InceptionV4 model for precise and efficient brain tumor classification. It used a 7,022 MRI dataset that categorised tumours into three classes. The result validated a higher accuracy of 98.7% with greater computational efficiency in medical decision-making. To address sensitivity to background variations in MRI images, Afshar et al. [13] presented a modified Capsule Network (CapsNet). By including tumor boundary data, the architecture is used to enhance the classification significantly by handling MRI image variability. Farzamnia et al. [14] implemented a contourlet transform and whale optimization model that was used to enhance a self-organizing map for a brain benign or malignant classification. This method attains a higher classification accuracy of 98.5% by maintaining computational efficiency with reliable diagnostic support for medical practitioners.

Zaitoon and Syed [15] presented a hybrid DL model for brain tumors using the BraTS dataset. This method had a DBT-CNN and RU-Net2+ model to attain 99% accuracy in classification and segmentation. The proposed model attained an accurate survival rate with a revolutionizing patient care and diagnostic automation. Rahman et al. [16] proposed a CNN-based random graph generation (CNNBCN) model that has a modified activation functions with an accuracy of 95.49% in brain tumor classification.

Bhimavarapu et al. [17] explored a semi-supervised learning approach (SSBTCNet) model that combined both Autoencoders and supervised networks to classify brain tumor. The proposed method enhanced with fuzzy-logic-based data augmentation that attains a higher accuracy with effective robustness and efficiency. Ramprasad et al. [18] presented a model called secured brain tumor classification network (SBTC-Net) for MRI-based brain tumor classification. This model uses secure image watermarking and transfer learning for MRI image processing. Gómez-Guzmán et al. [19] implemented an InceptionV3 using a dataset of 7,023 MRIs with an accuracy of 97.12% for classification in brain tumor.

In their study, Kokkalla et al. [20] developed a deep dense inception residual network using customized Inception ResNet v2 model. This method attains a highest accuracy of 99.69% than other models.

Kesav and Jibukumar [21] proposed a region-based CNN (RCNN) that combines a Two-Channel CNN and bounding box detection for classification. The RCNN achieves 98% accuracy with reduced execution time, effectively handling Glioma, Meningioma, and Pituitary tumors. Wankhede et al. [22] presented a Transfer learning with CNN architectures like ResNet50-152 for brain tumor classification using open-source datasets. The model achieves up to 96% accuracy using pre-trained weights.

Ali et al. [23] developed an attention-based UNET model that includes VGG layers in UNET for accurate segmentation. Results on BRATS'20 dataset shows that the model achieves high dice coefficients (up to 0.90) across tumor subtypes. Albalawi et al. [24] presented a federated learning-based CNN model for medical image processing. This model uses VGG16 for brain tumor localization.

In their study, Hong et al. [25] proposed a 3D-Feature Map Reconstruction Network (FRN)-ResNet model for brain tumor analysis. The FRN-ResNet model achieves higher accuracy by considering spatial details in diagnosing tumors. Hencya et al. [26] proposed using the Xception model to detect brain tumors. Also, the attention-based layers added in the learning model to process the more relevant features selectively.

Ali et al. [27] developed pre-trained GoogleNet, ShuffleNet, and NasNet-Mobile with ML classifiers of KNN, SVM, and LDA for brain tumor detection. Using MRI images of four tumor types, ShuffleNet with SVM achieved the best results with 98.40% accuracy.

Krishnasamy and Ponnusamy [28] developed hybrid FCN-ResNet and SegNet-MobileNet for classification. These models achieved high accuracies of 93.9% and 91.3% for two different publicly available datasets. Zahid et al. [29] used differential evolution and particle swarm optimization to find optimal feature vectors for brain tumor classification. This method achieved a speedup of 25.5x in prediction time by maintaining 94.4% accuracy. It attains significant computational efficiency and is also a viable approach for faster and more efficient tumor detection.

Krishnan et al. [30] introduced a Rotation Invariant Vision Transformer (RViT) designed with a rotated patch

embeddings with 98.6% of accuracy. The model achieves rotation invariance to enhance its robustness in detecting brain tumors. Rahman et al. [31] presented a dilated parallel deep CNN (PDCNN) to handle gridding artefacts and extract detailed features from MRI images. Using multiple dilation rates, the model attains both coarse and fine details with an accuracy of 98.67%.

Ullah et al. [32] proposed transfer learning (TL) based models that were fine-tuned to process a classification task by maintaining top-tier performance. The TL-based InceptionResNetV2 achieved the best performance with an accuracy of 98.91% accuracy to attain an automated medical diagnostic. Wang et al. [33] presented a new module called RanMerFormer to reduce the computational complexity of classification. This module can be combined with vision transformers (ViT) to increase computational efficiency. It removes redundant tokens in transformers and uses randomized vector functional links for swift training.

A hybrid CNN-SVM model is proposed by Bansal et al. [34] for multi-class classification. The CNN extracts features and SVM ensures high classification with an accuracy of up to 99%. This approach attains the potential of hybrid methods by improving diagnostic accuracy and speed. Tummala et al. [35] validated an ensemble of ViT models using MRI scans. It works by multi head attention technique to increase the feature learning capacity of the model. This method achieved a test accuracy of 98.2% for Kaggle dataset images.

Also, Cinar et al. [36] developed a hybrid of UNet and DenseNet121 models for tumor detection. The model focuses on tumor sub-regions and achieved superior results on the BRATS 2019 dataset in terms of memory requirements and inference times.

Haque et al. [37] proposed a model called NeuroNet19 which integrates VGG19 with an Inverted Pyramid Pooling Module (iPPM) for multi-scale feature extraction. Compared to U-Net models, the pyramid network achieves a higher accuracy of 97.86%. Stephe et al. [38] presented an Osprey Optimization Algorithm-based DL model (OOA-DL) for brain tumor classification. Initially, this model uses MobileNetV2 for feature extraction. Then, Osprey Optimization is used for feature selection. Finally, a Graph Convolutional Network is applied for classification.

The DeepTumorNet model is proposed by Raza et al. [39] for multi-class brain tumor categorization. This model is built using a modified GoogLeNet architecture to increase classification accuracy. In the modified GoogLeNet, the last five layers are replaced with 15 new ones using leaky ReLU activations. Results show that DeepTumorNet achieves 99.67% accuracy for validation sets. Haque et al. [40] presented a ViT model paired with a DCGAN-based data augmentation technique. It achieved a 99.33% accuracy that reduces training loss and enhances robustness with an advanced tumor diagnosis.

Nag et al. [41] used a TumorGANet that combines ResNet50 and GANs for feature extraction and data augmentation with 99.53% accuracy in brain tumor classification.

Hosny et al. [42] implemented an ensemble model that has seven DL models for deeper feature learning and brain tumors classification. Sahu et al. [43] presented a Cumulative Learning (CL) model and Multi-Rated New Loss (MRNL) that integrates DropOut, DropBlock, and Modified RandAugment, respectively. The method balanced the data limitations and increased 99.70% of accuracy effectively.

A hybrid model is proposed by Yoon [44] to classify the brain tumor. At first, the adaptive Wiener filtering hybrid with neural networks for preprocessing. Then, the SVM classification is used to achieve a 98.9% accuracy with high sensitivity.

3. PROPOSED MODEL

The proposed model has presented a feature extraction, feature selection, and classification for brain tumor. In the feature extraction stage, the MB-MSAT-Net processes input brain MRI images. MB-MSAT-Net uses multi-branch convolutional architecture and attention mechanisms to capture spatial and contextual information across multiple scales. Then, the extracted features are optimized using EFO. EFO selects the most discriminative features in order to reduce the dimensionality and improves computational efficiency. Finally, the refined features are fed into a TabNet classifier. This classifier accurately categorizing the input images into glioma, pituitary, meningioma, or no tumor. The overall workflow is given in Figure 1.
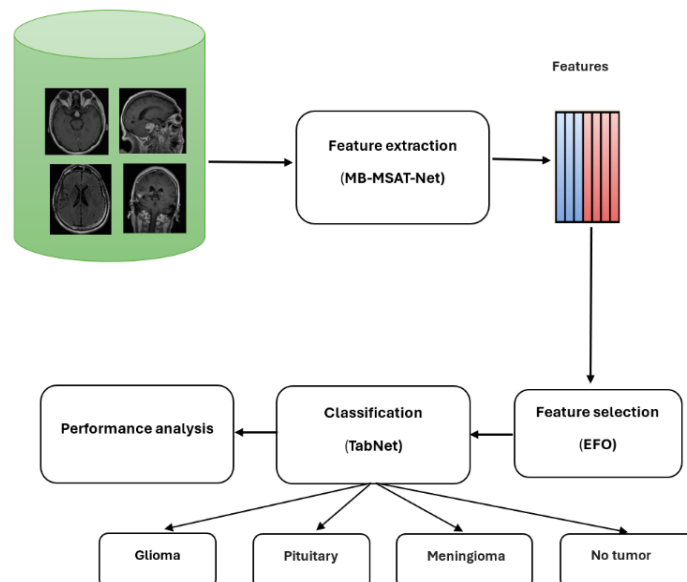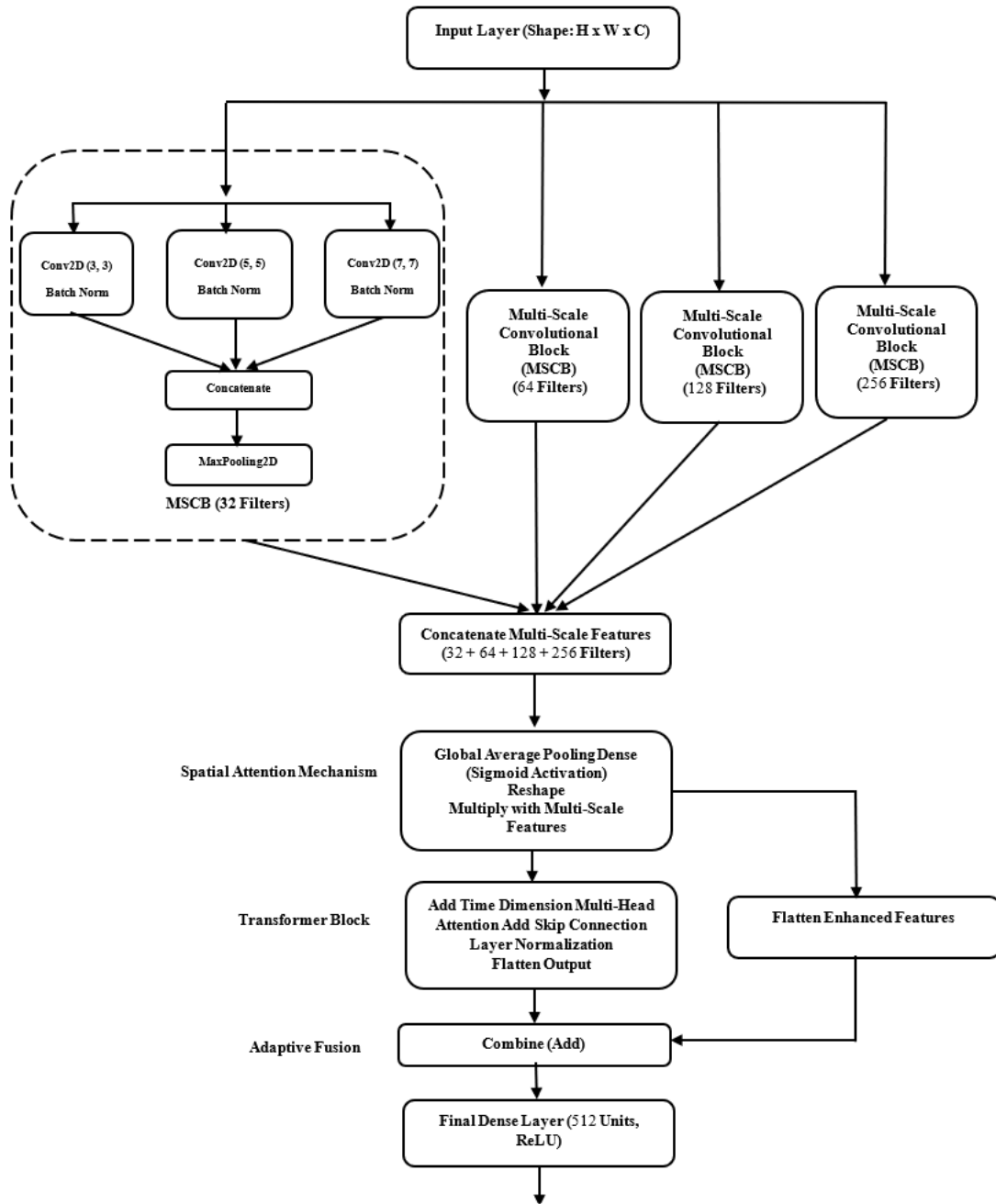

Figure 1. Proposed system

Figure 2. MB-MSAT-Net architecture

3.1 MB-MSAT-Net model

The MB-MSAT-Net DL architecture is used as a multi-branch and multi-scale architecture. Also, to improve feature learning capacity, the MB-MSAT-Net model is hybridised with attention mechanisms and transformer modules to capture both local and global features from MRI images.

Multi-Branch Multi-Scale Architecture: it has several branches with various kernel sizes like 3×3, 5×5, etc. These processes are used to extract both fine-grained details and contextual features to handle different tumor structures.

Collaboration among Branches: extracted Features of various branches are fused to merge both local and global information. The branches work in parallel to capture complementary features.

Attention Mechanisms and Transformers: A spatial attention mechanism prioritizes relevant regions of the image. It mainly focuses on important areas like tumor boundaries.

The Transformer modules are used to capture long-range dependencies and model complex global relationships across the image.

The architecture is shown in Figure 2.

MB-MSAT-Net Architecture

Here, the input layer fetched an image with dimensions H×W×C respectively. where H denotes the height, W indicates the width and C represents the number of channels. This layer used to prepare the input data for further processing by the network.

$$input: X \in R^{H \times W \times C} \tag{1}$$

Multi-Scale Convolutional Blocks

The multi-scale convolutional block applies multiple convolution operations with varying kernel sizes (e.g., 3×3, 5×5, 7×7) to capture features at different spatial scales. This helps the network learn fine-grained details (using small

kernels) and broader contextual information (using larger kernels). For each scale i, the convolution operations are performed, and the results are concatenated to capture multi-scale features. For a kernel of size $k \times k$ with FFF filters, the convolution operation is defined as:

$$conv_i(X) = conv2D(X, F, k \times k) \qquad (2)$$

where, X is the input feature map, F is the number of filters, and $k \times k$ is the kernel size. After applying convolutions with different kernel sizes, the outputs from all scales are concatenated:

$$Scale_i = concatenate(conv_1(X), conv_2(X), \ldots, conv_n(X)) \qquad (3)$$

MaxPooling is then applied to reduce spatial dimensions:

$$Scale_i = MaxPool2D(scale_i) \qquad (4)$$

Thus, the output from the multi-scale block is:

$$Scale_i = MaxPool2D \left(concatenate(conv_1(X), conv_2(X), \ldots, conv_n(X))\right) \qquad (5)$$

Spatial Attention Mechanism

This mechanism is used to focus on significant spatial regions in the feature map by assigning higher weights to relevant areas. This is achieved using a learned attention map that is generated from the input feature map. The mechanism amplifies the features in important spatial locations while suppressing irrelevant ones. After the multi-scale fusion is passed to a Global Average Pooling (GAP) layer is expressed as:

$$gap = GlobalAvgPool2D(multi\ scale\ features) \qquad (6)$$

Uisng a fully connected layer and the attention weights are the GAP output is given as.

$$Attention\ map = \sigma(dense(gap)) \qquad (7)$$

where, σ indicates a sigmoid.

The attention map is then reshaped and multiplied with the feature map:

$$enhanced\ features = multi - scale\ features \times attention\ map \qquad (8)$$

where, $\times$ indicates an element-wise multiplication.

Transformer block

The transformer block is used to attain a long-range dependency within the feature map using multi-head self-attention. It is used to understand relationships among distant regions of the input to process a complex task. Initially, the enhanced feature map is flattened with an expression:

$$flattened\ features = flatten(enhanced\ features) \qquad (9)$$

These are passed with an expand dimension operation to present a sequence dimension that is given in equation below.

$$transformer\_input = expland\_dims(flattened\ features, axis = 1) \qquad (10)$$

$$transformer_{input} \in R^{1 \times C \times H \times W} \qquad (11)$$

Now, the attention mechanism is used to process an input and then the multi-head self-attention is applied. It used to compute an attention for each part of the sequence based on the other parts:

$$Self_{attention}(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (12)$$

where, Q, K, and V are queries, keys, and values, and d_k is the dimension of the keys. The transformer block output is passed through a Layer Normalization layer, and a skip connection is added:

$$mha_{output} = LayerNormalization(transformer\ input + Self_{attention}) \qquad (13)$$

The transformer block output is expressed as:

$$mha_{output} = Flatten(mha_{output}) \qquad (14)$$

Adaptive Feature Fusion

After obtaining the features from both the CNN (multi-scale) and the transformer (long-range dependencies), these features are fused adaptively to create a comprehensive feature representation. This step ensures that both local and global information are combined effectively. The features from CNN and transformer are flattened and combined:

$$conv_{features} = Flatten(enhanced\ features) \qquad (15)$$

$$adaptive_{fusion} = conv_{features} + mha_{output} \qquad (16)$$

where, the Add () operation denotes element-wise addition.

Final Dense Layer

This layer can aggregate the fused features for a final output. It attained class probabilities for classification tasks or regression values. The fused features are forwarded through a Dense layer with a ReLU activation to learn a non-linear transformation:

$$final_{output} = Dense(adpative_{fusion}, 512, activation = 'relu') \qquad (17)$$

The above expression provides extracted features and then the softmax layer processes a classification.

3.2 EFO model

This model is a metaheuristic algorithm that is inspired by the electrolocation mechanism of electric fish [45]. These electric fields can navigate into murky waters, can also detect objects, and interact with their surroundings. Based on this biological feature, this method processes both exploration (global search) and exploitation (local search) mechanisms. It helps to balance diversity and convergence in the search to attain an optimal solution. The EFO model is explained with objectives and mathematical formulations are given as

follows:

Initialization

Initially, the population of candidate solutions is generated randomly within the search space bounds. Every candidate solution is used to represent a potential solution to overcome an optimization problem. The i-th candidate solution at iteration t is expressed as:

$$X_i(t) = [x_{i,1}(1), x_{i,2}(1), \ldots \ldots x_{i,D}(1)] \quad (18)$$

where, D indicates the dimensionality of the problem and $x_{i,j}(t)$ indicates a j th variable of the i-th solution.

$$x_{i,j}(0) = x_j^{min} + r(x_j^{max} - x_j^{min}) \quad (19)$$

where, x_j^{min} and x_j^{max} represents bounds for the j-th variable and also R indicates uniform random number in [0,1].

Passive Electrolocation (Exploration)

Passive electrolocation is used to detect an external electric field without generating new signals. This enables fish to broadly sense their environment. In EFO, passive electrolocation corresponds to global exploration, where new solutions are generated to probe unexplored regions of the search space. This phase prevents premature convergence where new solutions are generated by perturbing existing ones:

$$X_i(t + 1) = X_i(t) + \alpha R_i(t) \quad (20)$$

where, α denotes the Control parameter regulating perturbation magnitude and $R_i(t)$ indicates Random vector for perturbation.

The random vector is defined as:

$$R_i(t) = U.\left(X_j(t) - X_k(t)\right) \quad (21)$$

where, U denotes a Uniform random distribution. $X_j(t)$ and $X_k(t)$ indicates a random solution in population.

Active Electrolocation (Exploitation)

Active electrolocation is used to emit electric signals and analyse distortions caused by objects. It helps the fish to refine their perception. In EFO, this exploitation mainly focused on promising regions of the search space. It is used to improve the quality of the solution by refining the best solutions. The local refinement is performed as:

$$X_i(t + 1) = X_i(t) + \beta.(X_{best}(t) - X_i(t)) \quad (22)$$

where, β indicates a scaling factor and $X_{best}(t)$ represents the best solution.

Fitness-Based Frequency Calculation

This calculation was used to validate the fish's ability to modulate electric signal strength based on environmental feedback. The frequency $f_i(t)$ for the i-th solution is expressed in equation below.

$$= \frac{f_i(t)}{\frac{1}{1+\exp(-\gamma.(fit_{best}(t)-fit_i(t)))}} \quad (23)$$

where, $fit_{best}(t)$ indicates the best fitnes solution, $fit_i(t)$ denotes a i-th fitness solution and Γ denotes the scaling parameter.

Evaluation and Selection

Here, the fitness of every candidate solution is validated using the objective function $f(X)$ that is expressed as:

$$f(X_i(t)) = Objective\ Function\ Value \quad (24)$$

Selection rule:

$$X_i(t + 1) = \begin{cases} X_i(t + 1), if f(X_i(t + 1)) \le f(X_i(t)) \\ X_i(t) \qquad\qquad\quad otherwise \end{cases} \quad (25)$$

It repeats its iterations until a termination condition is met, whereas maximum number of iterations T or attain a desired fitness threshold.

3.3 EFO-based feature selection

```
Pseudocode for EFO-Based Feature Selection
# Initialize parameters and population
population_size = 50 # Number of candidate solutions
num_features = len(features) # Total number of features
max_iterations = 100 # Maximum number of iterations
alpha = 0.5 # Control parameter for exploration
beta = 0.5 # Scaling factor for exploitation
gamma = 0.1 # Sensitivity parameter

# Step 1: Initialize population randomly
population = initialize_population(population_size,
num_features)

# Step 2: Evaluate fitness for each solution
fitness_values = evaluate_fitness(population)

# Step 3: Iterate for a maximum number of iterations
for t in range(max_iterations):
    # Step 4: Exploration (Passive Electrolocation)
    for i in range(population_size):
        random_solution = random_selection(population)
        perturbation = alpha * (random_solution -
population[i])
        new_solution = population[i] + perturbation
        fitness_new = evaluate_fitness([new_solution])
        if fitness_new < fitness_values[i]:
            population[i] = new_solution
            fitness_values[i] = fitness_new

    # Step 5: Exploitation (Active Electrolocation)
    for i in range(population_size):
        best_solution = select_best_solution(population,
fitness_values)
        attraction = beta * (best_solution - population[i])
        population[i] = population[i] + attraction

    # Step 6: Update frequencies based on fitness
    frequencies = update_frequencies(fitness_values,
gamma)

    # Step 7: Evaluate and select the best solution
    best_solution = select_best_solution(population,
fitness_values)

    # Step 8: Stopping criteria (e.g., max iterations or desired
fitness)
```

```
    if stopping_condition_met(fitness_values):
        break
# Return the best feature subset
 return best_solution
```

The feature optimisation is used to improve the classification accuracy and reduce computational complexity. It helps to choose the specific and most relevant features for an accurate classification. In this work, the EFO model is used for feature selection to attain an effective classification performance. It is particularly suitable for high-dimensional medical image feature selection due to its unique balance between exploration and exploitation. The EFO focuses on navigating a large search space and avoiding local optima. With the comparison of Genetic Algorithms (GA) [46] and Particle Swarm Optimisation (PSO) [47], EFO used the natural electrolocation mechanism of electric fish to refine solutions

adaptively and select the most discriminative features for classification.

EFO-based feature optimisation selects the most relevant features by reducing the data dimensionality. This EFO model is used to increase its performance and reduce computational complexity.

3.4 TabNet for brain tumor classification

TabNet is a DL architecture designed to handle tabular data efficiently. Compared to other models, TabNet can automatically select important features and generate explainable decisions. In this work, the TabNet is used for classification based on the optimized features from EFO.

TabNet Architecture

TabNet uses a novel architecture that combines decision trees with DL model for both local and global data that is given in Figure 3.
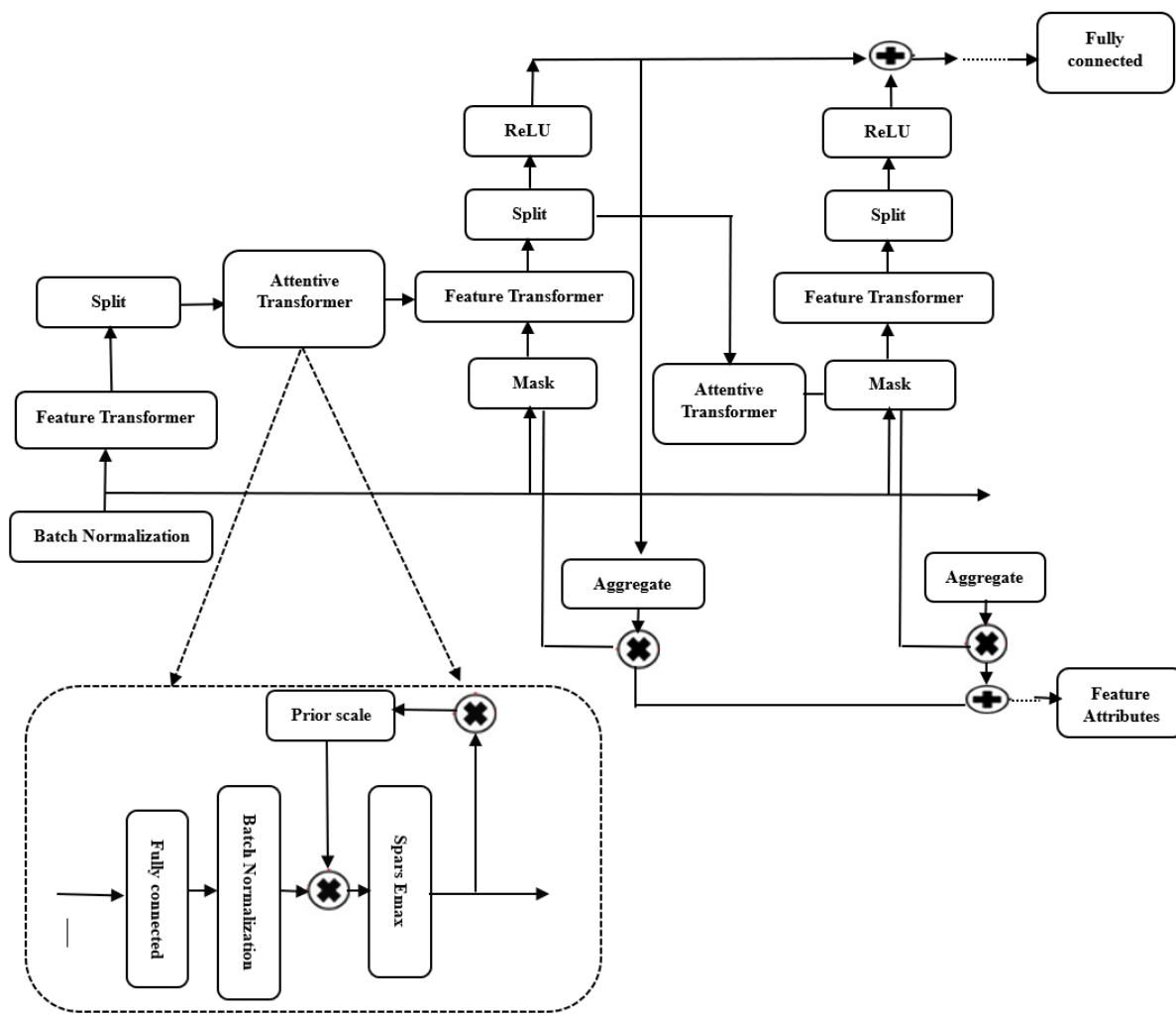

Figure 3. TabNet architecture

This architecture consists of a series of decision steps. In each step, a D-dimensional vector is executed by a Feature Transformer Module for classification. This module contains multiple layers for accurate learning. This learned knowledge is shared with other connections for final decision making. To handle non-linearity, the module consists of a Gated Linear Unit as an activation function. In addition, the residual connections are used to reduce network variations. The multi-layer design of the block increases feature selection and optimizes the network's parameter efficiency. The overall

architecture is given in Figure 3.

Input Layer and Embedding

Initially, the TabNet embedds the raw input features. For each input feature, the architecture applies an embedding layer to transform the raw data into dense vectors. This transformation is used for the model to capture feature relationships effectively.

Let the input data be represented as:

$$X = [x_1, x_2, ..., x_D] \tag{26}$$

where, D is the number of features in the dataset. These features are passed through an embedding layer, where each feature x_i is embedded into a dense vector e_i. If E is the embedding matrix, then the transformation can be written as:

$$e_i = f_{embed}(x_i), for\ all\ i = 1,2,\dots,D \qquad (27)$$

The embedded vectors are concatenated to form the embedding of the entire input:

$$X_{embed} = [e_1, e_2, \dots, e_D] \qquad (28)$$

Attention Blocks

The attention mechanism in TabNet is a sparse mechanism, meaning at each decision step, the model selects only a subset of features to focus on. Each attention block consists of two key components: Sparse Attention and Decision Layer.

Sparse Attention Mechanism

The sparse attention mechanism is implemented using the following steps:

- The input embeddings are first passed through a shared feature transformer which produces query (Q), key (K), and value (V) vectors:

$$Q = W_Q X_{embed}, K = W_K X_{embed}, V = W_V X_{embed} \qquad (29)$$

where, W_Q, W_K, W_V represents the learnable weight matrices for the query, key, and value transformations, respectively.

- The attention scores are calculated using below expression:

$$Attention\ (Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (30)$$

where, d_k indicates key vector's dimensionality.

The attention scores decide which features (or parts of the input) are important at each decision step. In TabNet, the attention is sparse, meaning only a small subset of features is attended to at each decision step.

Masking for Sparse Attention

A mask is applied to ensure the attention mechanism focuses only on a limited subset of features at each decision step:

$$m_t = softmax(QK^T).mask(t) \qquad (31)$$

where, $mask(t)$ is a learned mask that helps the network select which features to focus on at step t.

Decision Layer

After the attention mechanism, the selected features are passed through a decision layer to make predictions. This layer involves a feed-forward neural network (FFNN) applied to the attended features. The decision layer is defined as:

$$z_t = \sigma(W_t X_{attended} + b_t) \qquad (32)$$

where, $X_{attended}$ represents the attended features. σ as ReLU activation. W_t and b_t are learnable parameters for the t-th decision layer. This step helps refine the feature representation and prepares the model for the final output layer.

Update and Aggregation

The output of the decision layer is passed through the update and aggregation mechanism. The model updates its parameters and aggregates the attended features across all decision steps.

This is done using the following equation:

$$X_{updated} = X_{attended} + z_t \qquad (33)$$

where, $X_{attended}$ are the features selected by the attention mechanism, and z_t are the decision layer outputs.

Output Layer

The final output is generated through a fully connected layer that has a softmax activation which is given as follows:

$$y = softmax\left(W_{out} W_{updated} + b_{out}\right) \qquad (34)$$

where, W_{out} and b_{out} are the weights and bias for the output layer and y is the output vector. It gives the class probabilities.

4. RESULT AND DISCUSSION

To validate the MB-MSAT-Net, the data set is collected from the Mendeley Data Repository (https://data.mendeley.com/datasets/w4sw3s9f59/1). The dataset contains labelled MRI images of brain tumors. The dataset includes four distinct classes: glioma, meningioma, no tumor, and pituitary that used for both training set and the testing set. The training dataset consists of 1321 glioma images, 1339 meningioma images, 1595 no-tumour images, and 1457 pituitary images. The testing dataset is composed of 300 glioma images, 306 meningioma images, 405 no-tumour images, and 300 pituitary images. The visualization of the dataset images is shown in Figure 4.
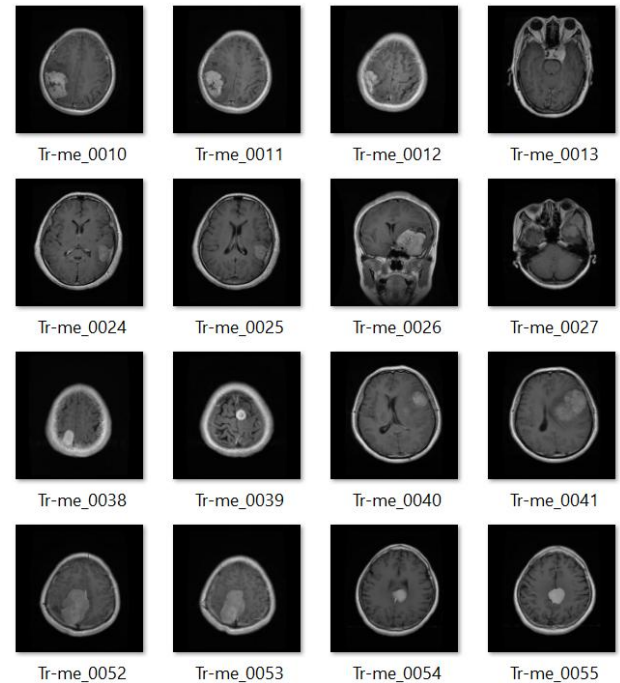

Figure 4. Dataset visualization

The MB-MSAT-Net model is coded in Python and simulated using IDLE 3.12 version. The packages like TensorFlow 3.10 are installed to implement model layers. The metrics like Sensitivity, Accuracy, F1 Score and Precision are calculated for evaluation. It can be calculated as follows:

$$Accuracy = \frac{TP+TN}{(TP + TN+FP + FN')} \qquad (35)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN}')} \tag{36}$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP}')} \tag{37}$$

$$\text{F1 score} = 2.\frac{\text{Precision} \cdot \text{Recall}}{(\text{Precision} + \text{Recall})} \tag{38}$$

where, TN is a True Negatives, TP is True Positives, FN is False Negatives and FP is False Positives. The optimization is carried out over 200 iterations. For each iteration, the fitness of each solution is evaluated by training the model using selected features with its classification error on the test set. The fitness function computed the error score using the categorical cross-entropy loss. Following the EFO optimization, the TabNet classifier is used as the final model. The features selected by EFO are used to train the TabNet model. It uses decision trees and attention mechanisms to model the relationships between the features. The TabNet classifier is trained using a batch size of 256, patience of 5 epochs, and

early stopping to prevent overfitting. The virtual batch size of 128 is used to stabilize training.

The EFO-based feature selection is shown in Figure 5 where the fitness value steadily decreases. Initially, the fitness value fluctuates as the optimizer explores various features. Over time, it converges toward a stable and lower classification error where it refines the feature set and increases feature selection effectively.

The performance of the model is given in Table 1. The proposed MB-MSAT-Net model achieves a higher accuracy of 99.2%. Also, EFO supports optimising feature selection in a better way, and TabNet also integrates tabular data effectively.

The confusion matrix of the proposed model is given in Figure 6. The large values along the main diagonal (294 for glioma, 304 for meningioma, 405 for no tumor, and 297 for pituitary) indicate that the model performs very well at correctly classifying instances into their respective classes. These are the *true positives* for each class.
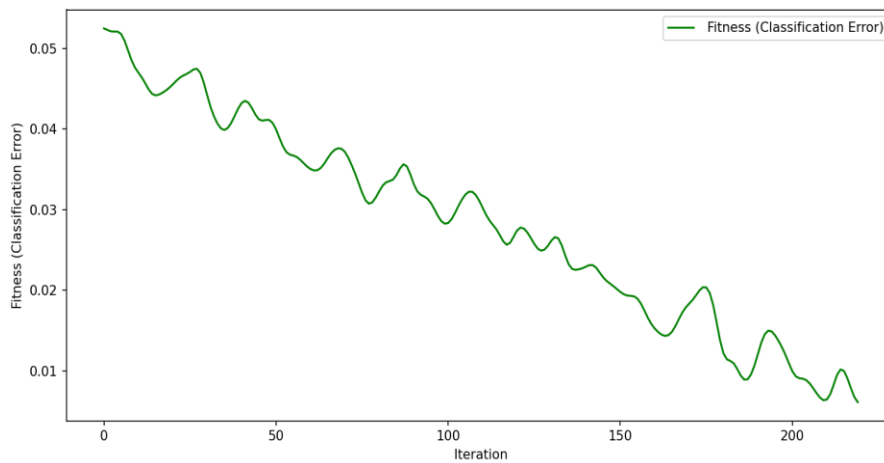

Figure 5. Feature selection fitness plot

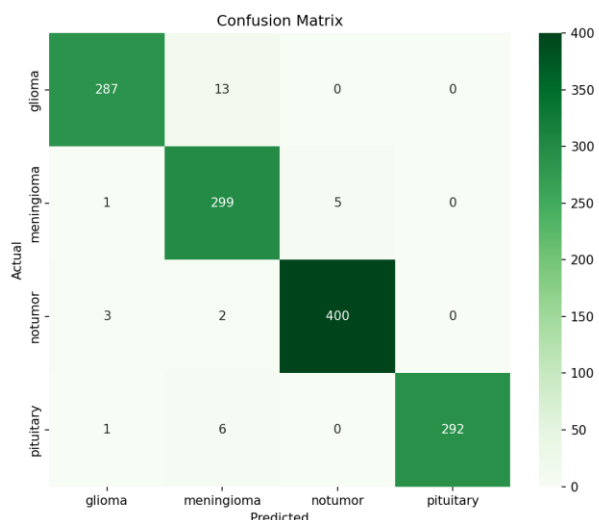**Table 1.** Ablation study of the model

| Model	Class	Precision	Recall	F1-Score	Accuracy
MB-MSAT-Net + EFO+ TabNet	glioma	99.7%	98%	98.8%	99.2
	meningioma	97.7%	99.7%	98.7%	
	notumor	99.8%	100%	99.9%	
	pituitary	100%	99.7%	99.9%	
MB-MSAT-Net (Multi-Branch + Multi-Scale, No EFO)	glioma	99.3%	97.3%	98.3%	98.7
	meningioma	97.1%	99.3%	98.2%	
	notumor	99.5%	100%	99.8%	
	pituitary	100%	99.0%	99.5%	
MB-MSAT-Net (Multi-Branch only, No Multi-Scale, No EFO)	glioma	98.3%	95.7%	97.0%	97.6
	meningioma	93.4%	98.0%	95.7%	
	notumor	98.8%	98.8%	98.8%	
	pituitary	100%	99.7%	99.9%	
MB-MSAT-Net (Multi-Scale only, No Multi-Branch, No EFO)	glioma	97.6%	95.0%	96.3%	97.4
	meningioma	92.9%	98.0%	95.4%	
	notumor	99.0%	98.8%	98.9%	
	pituitary	100%	99.3%	99.7%	
Simple CNN + TabNet, No EFO	glioma	95.5%	90.7%	93.0%	95.5
	meningioma	88.8%	94.1%	91.4%	
	notumor	97.6%	98.8%	98.2%	
	pituitary	99.7%	98.6%	99.2%	
Simple cnn+catboost	glioma	94.2%	87.1%	90.5%	93.2
	meningioma	82.9%	92.1%	87.3%	
	notumor	96.8%	98.0%	97.4%	
	pituitary	99.6%	94.3%	96.9%	
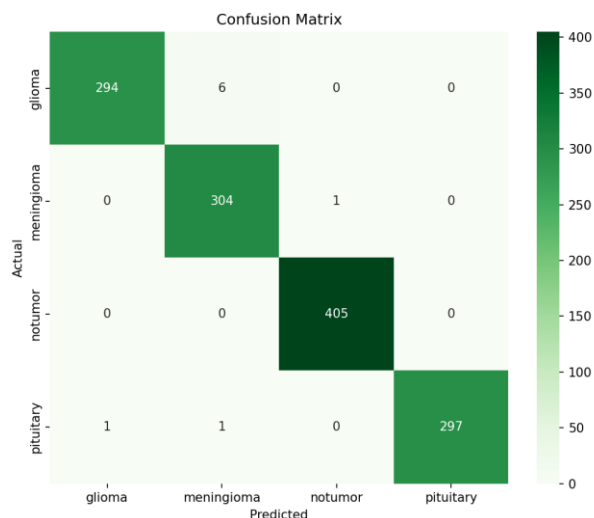
(a) MB-MSAT-Net + EFO+ TabNet
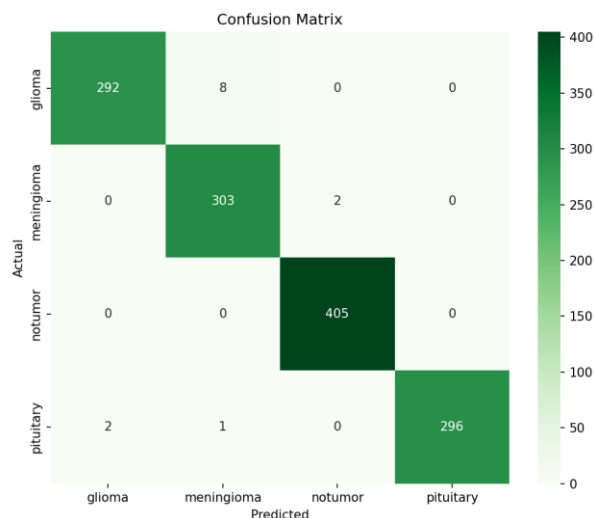
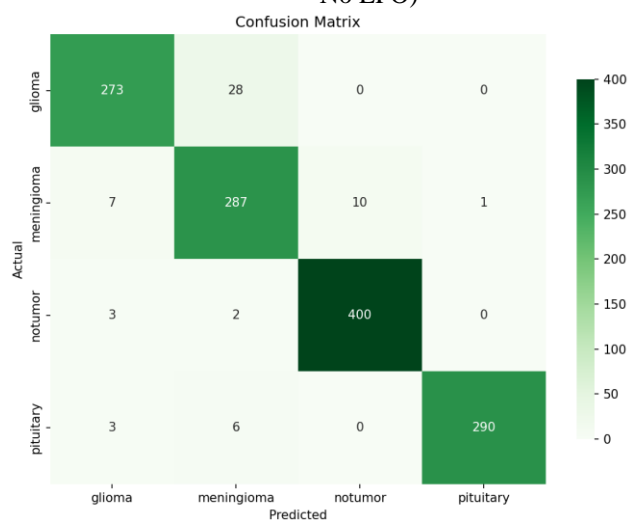(b) MB-MSAT-Net (Multi-Branch + Multi-Scale, No EFO)

(c) MB-MSAT-Net (Multi-Branch only, No Multi-Scale, No EFO)
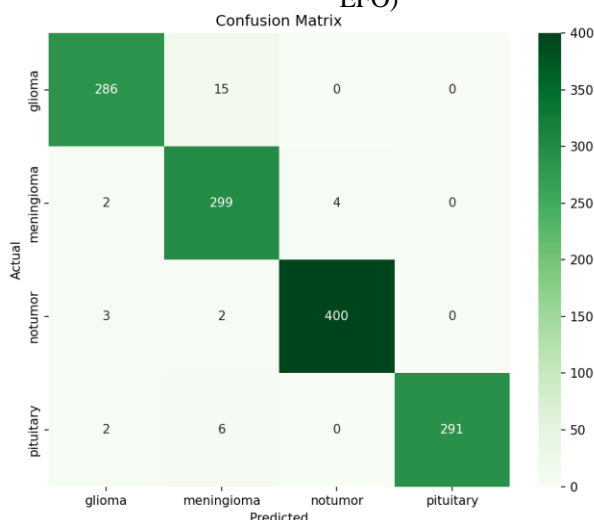
(d) MB-MSAT-Net (Multi-Scale only, No Multi-Branch, No EFO)
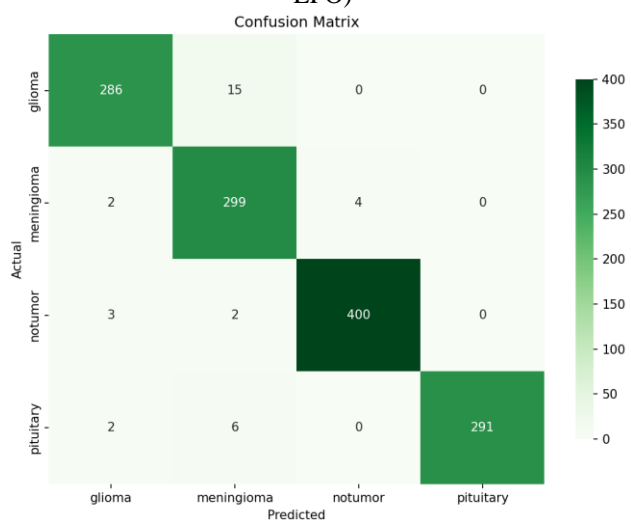
(e) Simple CNN + TabNet, No EFO

(f) CNN +XGBoost

Figure 6. Confusion matrix analysis of the models

The ablation study provides the impact of various components integrated in the proposed EFO-based MB-MSAT-Net architecture. It highlighted the performance of the multi-branch, multi-scale, and EFO components. In the ablation study, the 'Multi-Branch Only' and 'Multi-Scale Only' configurations are used to simplify the full model by reducing a few key modules where both maintain functional architectures. But, it occurs a performance degradation due to the loss of important feature extraction capabilities. In Multi-Branch Only, the model captures various local features but lacks multi-scale resolution. In Multi-Scale Only, it captures broad spatial context but loses feature diversity. It shows the

need to hybrid both modules for optimal performance in brain tumor classification.

Below is an analysis based on the provided results:

MB-MSAT-Net + EFO + TabNet

This is the full-fledged model which incorporates multi-branch, multi-scale processing, EFO, and TabNet. It achieves the best overall performance across all classes, with an impressive accuracy of 99.2% and F1 scores nearing or exceeding 99% for all tumor types. The inclusion of EFO is used to optimize the feature set which enhances precision and recall.

MB-MSAT-Net (Multi-Branch + Multi-Scale, No EFO)

Removing EFO slightly reduces accuracy to 98.7%. It proves the importance of EFO in fine-tuning the feature representation. However, the multi-branch and multi-scale components still deliver strong results with balanced precision, recall, and F1 scores across all classes.

MB-MSAT-Net (Multi-Branch only, No Multi-Scale, No EFO)

When the multi-scale component is excluded, the performance declines further with an accuracy of 97.6%. The F1 scores for glioma and meningioma classes drop more significantly. This version highlights the importance of multi-scale features in improving model performance.

MB-MSAT-Net (Multi-Scale only, No Multi-Branch, No EFO)

Similarly, excluding the multi-branch component and relying solely on multi-scale processing results in an accuracy of 97.4%. The absence of a multi-branch design limits the model's ability to integrate diverse feature representations. It showed that the multi-branch complements multi-scale processing by attaining a higher feature set.

Simple CNN + TabNet, No EFO

Without the multi-branch or multi-scale components, the accuracy drops to 95.5%. Although the model still benefits from TabNet's tabular data integration. But it cannot capture hierarchical and spatial features results in lower precision and recall rates.

Simple CNN + XGBoost

This configuration, with a simpler CNN backbone and XGBoost, exhibits the lowest performance, with an accuracy of 93.2%. The reduced capability to model complex interactions between features leads to significant drops in precision and recall for glioma and meningioma.

Figure 7 shows the ROC plot of the MB-MSAT-Net model. The True Positive Rate (TPR) represents the proportion of actual positives that are correctly identified. A TPR of 1 means all actual positives are correctly classified. The False Positive Rate (FPR) represents the proportion of actual negatives that are incorrectly classified as positives. An FPR of 0 means no actual negatives are misclassified. The curves for "notumor" and "pituitary" are very close to the top-left corner. It denotes the near-perfect performance for these classes. They both have an AUC of 1.0. The curve for "meningioma" is also very close to the top-left corner, with an AUC of 0.99 which indicates excellent performance of the model.The curve for "glioma" is slightly lower, with an AUC of 0.98, but it still represents very good performance.

Table 2 presents a direct comparison of the proposed model with existing state-of-the-art models under identical experimental settings. All models are assessed on the dataset using the same preprocessing steps, training parameters. Compared to all models and recently proposed models, the MB-MSAT-Net achieves the highest accuracy of 99.2%. This architecture's ability to extract different and multi-resolution features, coupled with optimal feature selection.

The analysis of the feature selection capability of EFO with other optimizers is given in Table 3. The EFO-based feature selection achieves the highest classification accuracy of 99.2% when compared to GA (96.7%) and PSO (97.5%). The convergence rate is represented as a numerical value to compare how fast the algorithm reaches a stable optimal solution. EFO has the highest convergence rate (0.9) among other optimizers. EFO converges quickly to the optimal solution.
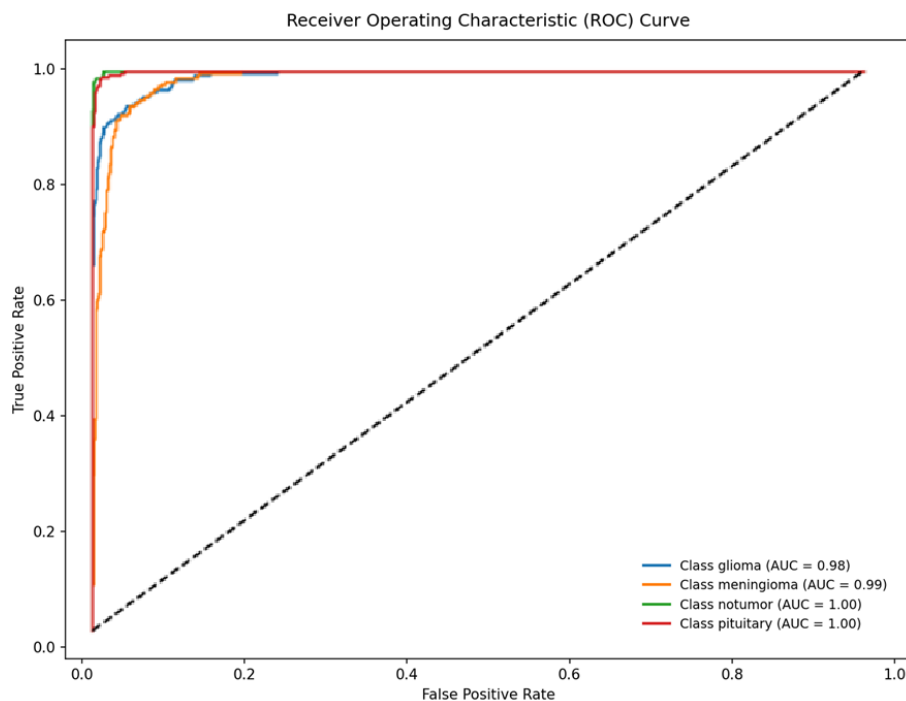

Figure 7. Receiver Operating Characteristic (ROC) analysis

Table 2. Comparison with other models

Author(s)	Method	Accuracy
Bhimavarapu et al. [17]	SSBTCNet	92.4%
Ramprasad et al. [18]	SBTC-Net	91.5%
Gómez-Guzmán et al. [19]	CNN, EfficientNet B1	90.8%
Kokkalla et al. [20]	Deep Inception ResNet v2	91.0%
Kesav and Jibukumar [21]	RCNN based model	88.3%
Wankhede et al. [22]	Inception v3	89.5%
Ali et al. [23]	UNET with pre-trained VGG19	90.0%
Albalawi et al. [24]	ResNet50, ResNet152	91.8%
Hong et al. [25]	3D FRN-ResNet	89.6%
Hencya [26]	Xception	90.4%
Ali [27]	ShuffleNet with SVM	87.2%
Krishnasamy and Ponnusamy [28]	FCN+ResNet	88.0%
Zahid et al. [29]	PCA + DRNN	84.5%
Krishnan [30]	Rotation Invariant Vision Transformer (RViT)	92.0%
Rahman [31]	Dilated Parallel Deep Convolutional Neural Network (PDCNN)	89.7%
Ullah [32]	Inception GoogLeNet	97.8%
Wang et al. [33]	RanMerFormer	97.9%
Bansal et al. [34]	CNN+SVM	97.5%
Tummala et al. [35]	Ensemble ViT models	98.0%
Cinar et al. [36]	Hybrid DenseNet121-UNet model	97.7%
Haque et al. [37]	NeuroNet19	97.6%
Stephe et al. [38]	OOA-DL	97.4%
Raza et al. [39]	DeepTumorNet	97.8%
Haque et al. [40]	DCGAN	97.6%
Nag et al. [41]	TumorGANet	98.4%
Hosny et al. [42]	GoogLeNet, Xception, MobileNetV2, ResNet50V2 Ensemble	98.3%
Sahu et al. [43]	CLA + MRNL	97.2%
Yoon [44]	Wiener Filtering + SVM	97.0%
Proposed	MB-MSAT-Net with EFO (accuracy of 99.2%)	99.2%

Table 3. Comparison of feature selection performance

Optimization Algorithm	Accuracy	Convergence Rate
EFO	99.2%	0.9
GA (Genetic Algorithm)	96.7%	0.8
PSO (Particle Swarm Optimization)	97.5%	0.6

Table 4. Statistical analysis of the model

Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	p-value (paired t-test)
Proposed MB-MSAT-Net + EFO + TabNet	99.7	98.0	98.8	99.2	-
SSBTCNet	92.4	90.7	91.5	92.4	0.0001
SBTC-Net	91.5	89.2	90.3	91.5	0.0003
CNN, EfficientNet B1	90.8	89.4	90.0	90.8	0.0002
Deep Inception ResNet v2	91.0	89.8	90.4	91.0	0.0001
RCNN-based model	88.3	85.7	86.9	88.3	0.0015
ResNet50	91.8	90.5	91.1	91.8	0.0004
GoogLeNet, Xception, MobileNetV2, ResNet50V2 Ensemble	98.3	97.6	97.9	98.3	0.0002
TumorGANe	98.4	97.9	98.1	98.4	0.0001
MB-MSAT-Net with EFO	99.7	98.0	98.8	99.2	-

Table 5. Computational complexity analysis of MB-MSAT-Net

Metric	MB-MSAT-Net	Ensemble	Deep Inception ResNet v2	TumorGANet	ResNet50 + XGBoost	Simple CNN
Training Time	13.5 hours	14 hours	12.8 hours	13.8 hours	6.3 hours	4.2 hours
Inference Time	0.45 seconds	0.48 seconds	0.42 seconds	0.46 seconds	0.25 seconds	0.18 seconds
Memory Consumption	4.8 GB	5 GB	4.5 GB	4.8 GB	2.5 GB	1.5 GB

The statistical result of a paired t-test on the performance between the MB-MSAT-Net + EFO + TabNet model and the baseline methods is presented. The obtained results are given in Table 4. A p-value of less than 0.05 indicates that the performance difference between the two models is statistically significant. The p-values for all comparisons are below 0.05, which denotes that the performance improvements of the MB-MSAT-Net + EFO + TabNet model over the baseline models

are statistically significant.

The computational analysis of the MB-MSAT-Net is given in Table 5. The computational complexity of MB-MSAT-Net is higher than simpler models like Simple CNN and ResNet50 + XGBoost, However, these trade-offs are acceptable for real-time applications with their highest accuracy.

The performance of the model for varying clinical conditions is given in Table 6. The outcomes denote that MB-MSAT-Net maintains strong accuracy even under noisy or low-resolution conditions. It is observed that only a moderate reduction in performance for extreme noise or very low resolution. In misdiagnosis risk assessment, the model performance is evaluated with an additional layer of uncertainty. It simulates a scenario where the model's predictions are flagged for misdiagnosis risk. The model shows a slight drop in performance.

The Grad-CAM outputs of MB-MSAT-Net are given in Figure 8. It denotes that the model successfully learned to localize pathology in the brain MRIs. For healthy scans, it shows low activation. For scans with tumors, it points to the tumor which influences its diagnostic classification. To analyse the feature importance, the SHAP (SHapley Additive exPlanations) feature importance plot is generated as shown in Figure 9. These plots identify specific radiomics features like GLRLM_Feature_04828 as the most critical drivers of the model's predictions. The highest values of this feature strongly contribute to the "positive" prediction.

To analyse model generalization and real-time applicability, the MB-MSAT-Net is applied for the other brain tumor datasets like BraTS 2020, Kaggle and the TCIA brain tumor dataset. The measured results are given in Table 7. The MB-MSAT-Net show better results in terms of all metrics for different datasets. This cross-validation proves the model's ability to work in different datasets.

Table 6. Performance of the model for varying clinical conditions

Test Scenario	Accuracy (%)	Precision	Recall	F1-Score
Original Test Set	99.2	0.997	0.998	0.997
Low-Quality Image (Gaussian Noise $\sigma = 0.01$)	97.8	0.974	0.976	0.975
High-Noise Image (Gaussian Noise $\sigma = 0.05$)	96.5	0.962	0.961	0.961
Very High Noise (Gaussian Noise $\sigma = 0.1$)	93.4	0.937	0.938	0.937
Resolution 112×112	98.4	0.985	0.986	0.985
Low Resolution 56×56	94.6	0.948	0.946	0.947
Misdiagnosis Risk Assessment (Model Uncertainty)	92.5	0.925	0.923	0.924

Table 7. Performance of MB-MSAT-Net for other datasets

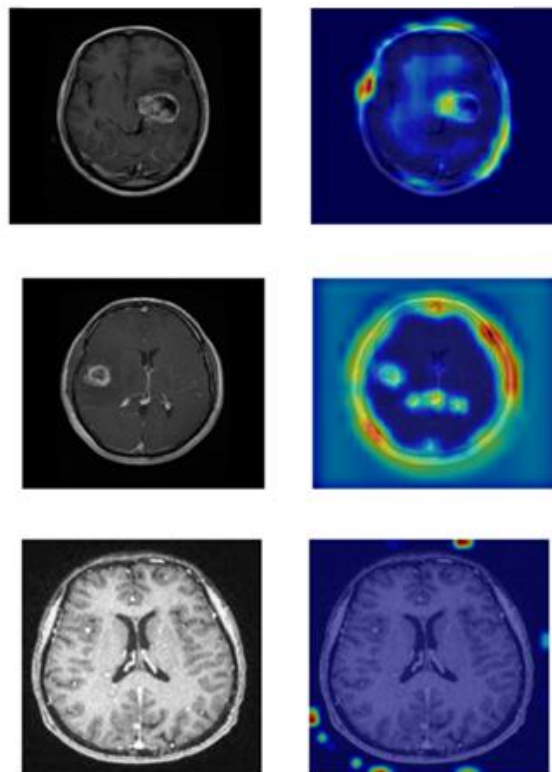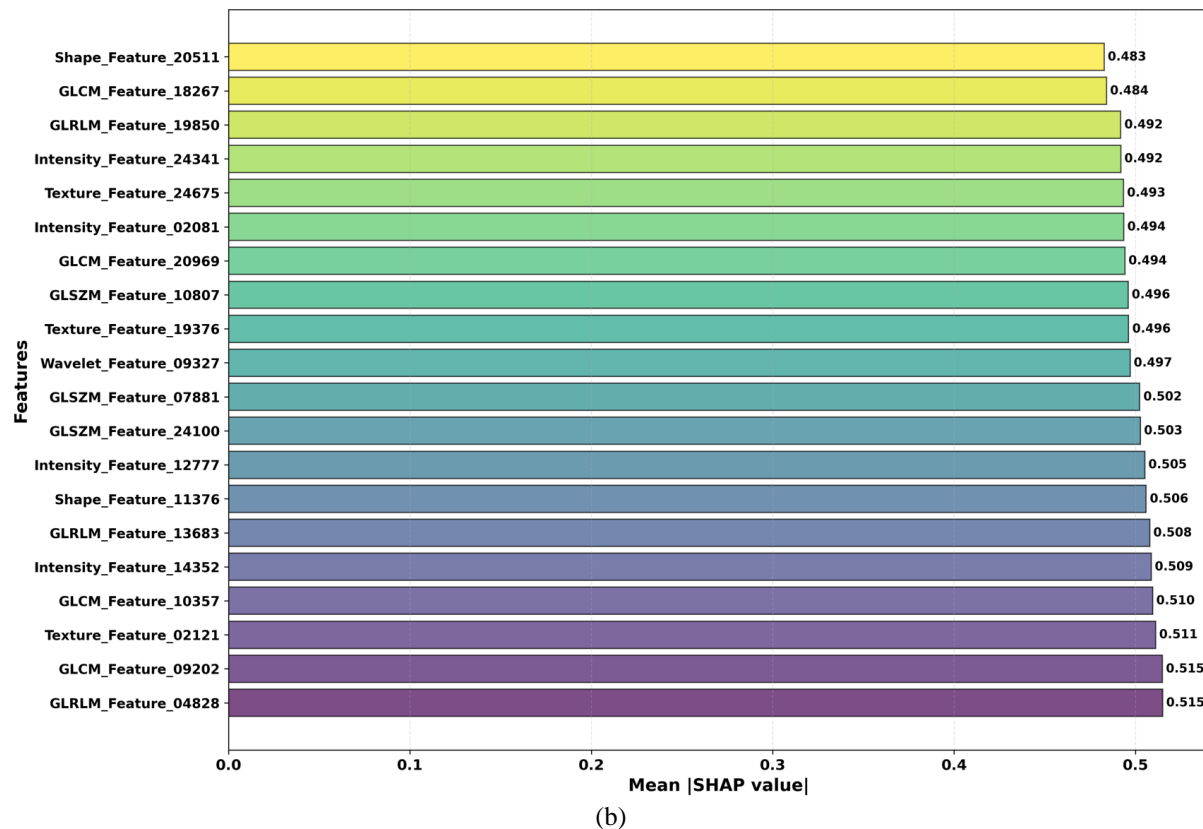| Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
BraTS 2020 Dataset	99.2	99.7	98.0	98.8
Brain MRI Kaggle Dataset	98.5	98.0	97.5	97.8
TCIA brain tumor dataset	99.1	99.0	98.5	98.7

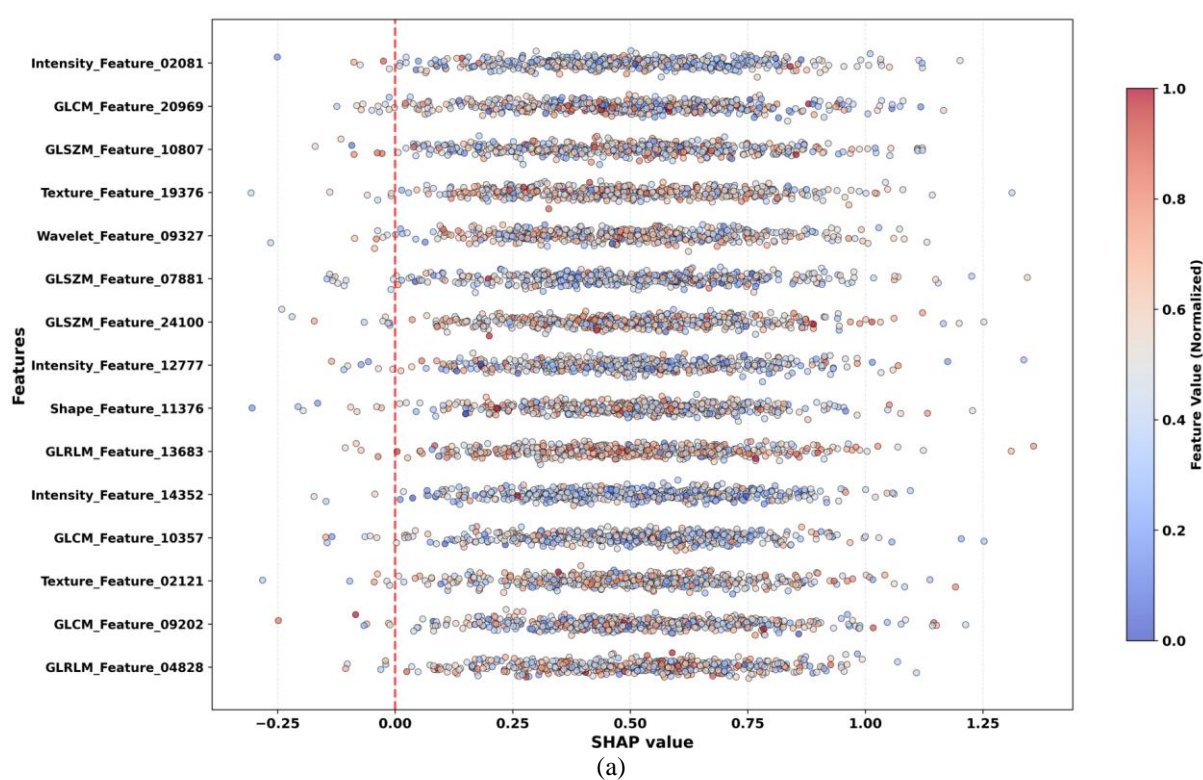**Figure 8.** Grad-CAM visulization of the model

Figure 9. (a) SHAP summary (b) Top 20 features

5. CONCLUSION

The proposed MB-MSAT-Net for brain tumor classification integrates the MB-MSAT-Net, EFO, and TabNet for medical image analysis. By using the strengths of advanced feature extraction and optimized feature selection, the framework significantly increases the accuracy and efficiency.

Experimental results on publicly available brain tumor datasets confirm that this integrated approach outperforms traditional methods. Future work could explore additional optimization techniques and extend the approach to other medical imaging tasks. In addition, different MRI scanning parameters and multiple image modalities will be applied to increase detection reliability.

REFERENCES

[1] About Brain Tumors: A Primer for Patients and Caregivers. https://www.abta.org/secure/about-brain tumors-a-primer.pdf, accessed on 2015.

[2] Siegel, R.L., Miller, K.D., Fuchs, H.E., Jemal, A. (2022). Cancer statistics, 2022. CA: A Cancer Journal for Clinicians, 70(1): 7-30. https://doi.org/10.3322/caac.21590

[3] Mansur, Z., Talukdar, J., Singh, T.P., Kumar, C.J. (2024). Deep learning-based brain tumor image analysis for segmentation. SN Computer Science, 6(1): 42. https://doi.org/10.1007/s42979-024-03558-x

[4] Wang, S., Summers, R.M. (2012). Machine learning and radiology. Medical Image Analysis, 16(5): 933-951. https://doi.org/10.1016/j.media.2012.02.005

[5] Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., Shoaib, M. (2020). A deep learning model based on concatenation approach for the diagnosis of brain tumor. IEEE Access, 8: 55135-55144. https://doi.org/10.1109/ACCESS.2020.2978629

[6] Naeem, A.B., Senapati, B., Zaidi, A. (2025). Enhancing brain tumor detection from MRI-based images through deep transfer learning models. AI, 6(12): 305. https://doi.org/10.3390/ai6120305

[7] Babar, N.A., Lateef, J., Syed, S., Dietlmeier, J., O'Connor, N.E., Raupp, G.B., Spanias, A. (2025). Brain tumor classification in MRI scans using edge computing and a shallow attention-guided CNN. Biomedicines, 13(10): 2571. https://doi.org/10.3390/biomedicines13102571

[8] Rastogi, D., Johri, P., Donelli, M., Kumar, L., Bindewari, S., Raghav, A., Khatri, S.K. (2025). Brain tumor detection and prediction in MRI images utilizing a Fine-Tuned transfer learning model integrated within deep learning frameworks. Life, 15(3): 327. https://doi.org/10.3390/life15030327

[9] Islam, A., Reza, S.M., Iftekharuddin, K.M. (2013). Multifractal texture estimation for detection and segmentation of brain tumors. IEEE Transactions on Biomedical Engineering, 60(11): 3204-3215. https://doi.org/10.1109/TBME.2013.2271383

[10] Saeedi, S., Rezayi, S., Keshavarz, H., Niakan Kalhori, S.R. (2023). MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Medical Informatics and Decision Making, 23(1): 16. https://doi.org/10.1186/s12911-023-02114-6

[11] Gumaei, A., Hassan, M.M., Hassan, M.R., Alelaiwi, A., Fortino, G. (2019). A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. IEEE Access, 7: 36266-36273. https://doi.org/10.1109/ACCESS.2019.2904145

[12] Bibi, N., Wahid, F., Ma, Y., Ali, S., Abbasi, I.A., Alkhayyat, A. (2024). A transfer learning-based approach for brain tumor classification. IEEE Access, 12: 111218-111238. https://doi.org/10.1109/ACCESS.2024.3425469

[13] Afshar, P., Plataniotis, K.N., Mohammadi, A. (2019). Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, pp. 1368-1372.

https://doi.org/10.1109/ICASSP.2019.8683759

[14] Farzamnia, A., Hazaveh, S.H., Siadat, S.S., Moung, EG. (2023). MRI brain tumor detection methods using contourlet transform based on time adaptive self-organizing map. IEEE Access, 11: 113480-113492. https://doi.org/10.1109/ACCESS.2023.3322450

[15] Zaitoon, R., Syed, H. (2023). RU-Net2+: A deep learning algorithm for accurate brain tumor segmentation and survival rate prediction. IEEE Access, 11: 118105-118123. https://doi.org/10.1109/ACCESS.2023.3325294

[16] Rahman, T., Islam, M.S., Uddin, J. (2024). MRI-based brain tumor classification using a dilated parallel deep convolutional neural network. Digital, 4(3): 529-554. https://doi.org/10.3390/digital4030027

[17] Bhimavarapu, U., Chintalapudi, N., Battineni, G. (2024). Brain tumor detection and categorization with segmentation of improved unsupervised clustering approach and machine learning classifier. Bioengineering, 11(3): 266. https://doi.org/10.3390/bioengineering11030266

[18] Ramprasad, M.V.S., Rahman, M.Z.U., Bayleyegn, M.D. (2023). SBTC-net: Secured brain tumor segmentation and classification using black widow with genetic optimization in IoMT. IEEE Access, 11: 88193-88208. https://doi.org/10.1109/ACCESS.2023.3304343

[19] Gómez-Guzmán, M.A., Jiménez-Beristaín, L., García-Guerrero, E.E., López-Bonilla, O.R., Tamayo-Perez, U.J., Esqueda-Elizondo, J.J., Palomino-Vizcaino, K., Inzunza-González, E. (2023). Classifying brain tumors on magnetic resonance imaging by using convolutional neural networks. Electronics, 12(4): 955. https://doi.org/10.3390/electronics12040955

[20] Kokkalla, S., Kakarla, J., Venkateswarlu, I.B., Singh, M. (2021). Three-class brain tumor classification using deep dense inception residual network. Soft Computing, 25(13): 8721-8729. https://doi.org/10.1007/s00500-021-05748-8

[21] Kesav, N., Jibukumar, M.G. (2022). Efficient and low complex architecture for detection and classification of Brain Tumor using RCNN with Two Channel CNN. Journal of King Saud University-Computer and Information Sciences, 34(8): 6229-6242. https://doi.org/10.1016/j.jksuci.2021.05.008

[22] Wankhede, D.S., Shelke, C.J., Shrivastava, V.K., Achary, R., Mohanty, S.N. (2024). Brain tumor detection and classification using adjusted InceptionV3, AlexNet, VGG16, VGG19 with ResNet50-152 CNN Model. EAI Endorsed Transactions on Pervasive Health & Technology, 10(1): 1-8. https://doi.org/10.4108/eetpht.10.6377

[23] Ali, T.M., Nawaz, A., Rehman, A.U., Ahmad, R.Z., Javed, A.R., Gadekallu, T.R., Chen, C., Wu, C.M. (2022). A sequential machine learning-cum-attention mechanism for effective segmentation of brain tumor. Frontiers in Oncology, 12: 873268. https://doi.org/10.3389/fonc.2022.873268

[24] Albalawi, E., TR, M., Thakur, A., Kumar, V.V., Gupta, M., Khan, S.B., Almusharraf, A. (2024). Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor. BMC Medical Imaging, 24(1): 110. https://doi.org/10.1186/s12880-024-01261-0

[25] Hong, J., Huang, Y., Ye, J., Wang, J., et al. (2022). 3D

FRN-ResNet: An automated major depressive disorder structural magnetic resonance imaging data identification framework. Frontiers in Aging Neuroscience, 14: 912283. https://doi.org/10.3389/fnagi.2022.912283

[26] Hencya, F.R., Mandala, S., Tang, T.B., Zahid, M.S.M. (2023). A transfer learning-based model for brain tumor detection in MRI images. Jurnal Nasional Teknik Elektro, 12(2): 1-12. https://doi.org/10.25077/jnte.v12n2.1123.2023

[27] Ali, R., Al-Jumaili, S., Duru, A.D., Uçan, O.N., Boyaci, A., Duru, D.G. (2022). Classification of brain tumors using MRI images based on convolutional neural network and supervised machine learning algorithms. In 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, pp. 822-827. https://doi.org/10.1109/ISMSIT56059.2022.9932690

[28] Krishnasamy, N., Ponnusamy, T. (2023). Deep learning-based robust hybrid approaches for brain tumor classification in magnetic resonance images. International Journal of Imaging Systems and Technology, 33(6): 2157-2177. https://doi.org/10.1002/ima.22974

[29] Zahid, U., Ashraf, I., Khan, M.A., Alhaisoni, M., Yahya, K.M., Hussein, H.S., Alshazly, H. (2022). BrainNet: optimal deep learning feature fusion for brain tumor classification. Computational Intelligence and Neuroscience, 2022(1): 1465173. https://doi.org/10.1155/2022/1465173

[30] Krishnan, P.T., Krishnadoss, P., Khandelwal, M., Gupta, D., Nihaal, A., Kumar, T.S. (2024). Enhancing brain tumor detection in MRI with a rotation invariant Vision Transformer. Frontiers in Neuroinformatics, 18: 1414925. https://doi.org/10.3389/fninf.2024.1414925

[31] Rahman, T., Islam, M.S., Uddin, J. (2024). MRI-based brain tumor classification using a dilated parallel deep convolutional neural network. Digital, 4(3): 529-554. https://doi.org/10.3390/digital4030027

[32] Ullah, N., Khan, J.A., Khan, M. S., Khan, W., Hassan, I., Obayya, M., Negm, N., Salama, A.S. (2022). An effective approach to detect and identify brain tumors using transfer learning. Applied Sciences, 12(11): 5645. https://doi.org/10.3390/app12115645

[33] Wang, J., Lu, S.Y., Wang, S.H., Zhang, Y.D. (2024). RanMerFormer: Randomized vision transformer with token merging for brain tumor classification. Neurocomputing, 573: 127216. https://doi.org/10.1016/j.neucom.2023.127216

[34] Bansal, S., Jadon, R.S., Gupta, S.K. (2024). A robust hybrid convolutional network for tumor classification using brain MRI image datasets. International Journal of Advanced Computer Science & Applications, 15(4): 576-584. https://doi.org/10.14569/IJACSA.2024.0150459

[35] Tummala, S., Kadry, S., Bukhari, S.A.C., Rauf, H.T. (2022). Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling. Current Oncology, 29(10): 7498-7511. https://doi.org/10.3390/curroncol29100590

[36] Cinar, N., Ozcan, A., Kaya, M. (2022). A hybrid DenseNet121-UNet model for brain tumor segmentation from MR Images. Biomedical Signal Processing and Control, 76: 103647.

https://doi.org/10.1016/j.bspc.2022.103647

[37] Haque, R., Hassan, M.M., Bairagi, A.K., Shariful Islam, S.M. (2024). NeuroNet19: An explainable deep neural network model for the classification of brain tumors using magnetic resonance imaging data. Scientific Reports, 14(1): 1524. https://doi.org/10.1038/s41598-024-51867-1

[38] Stephe, S., Nivedita, V., Karthikeyan, B., Nithya, K., Sikkandar, M.Y. (2024). Enhancing brain tumor detection and classification using osprey optimization algorithm with deep learning on MRI images. Journal of Intelligent Systems & Internet of Things, 12(1): 33-44. https://doi.org/10.54216/JISIoT.120103

[39] Raza, A., Ayub, H., Khan, J.A., Ahmad, I., S. Salama, A., Daradkeh, Y.I., Javeed, D., Ur Rehman, A., Hamam, H. (2022). A hybrid deep learning-based approach for brain tumor classification. Electronics, 11(7): 1146. https://doi.org/10.3390/electronics11071146

[40] Haque, M., Paul, S.K., Paul, R.R., Islam, N., Rashidul Hasan, M.A.F.M., Hamid, M. (2023). Improving performance of a brain tumor detection on MRI images using DCGAN-based data augmentation and Vision Transformer (ViT) approach. In GANs for Data Augmentation in Healthcare, pp 157-186. https://doi.org/10.1007/978-3-031-43205-7_10

[41] Nag, A., Mondal, H., Hassan, M.M., Al-Shehari, T., Kadrie, M., Al-Razgan, M., Alfakih, T., Biswas, S., Bairagi, A.K. (2024). TumorGANet: A transfer learning and generative adversarial network-based data augmentation model for brain tumor classification. IEEE Access, 12: 103060-103081. https://doi.org/10.1109/ACCESS.2024.3429633

[42] Hosny, K.M., Mohammed, M.A., Salama, R.A., Elshewey, A.M. (2025). Explainable ensemble deep learning-based model for brain tumor detection and classification. Neural Computing and Applications, 37(3): 1289-1306. https://doi.org/10.1007/s00521-024-10401-0

[43] Sahu, A., Das, P.K., Paul, I., Meher, S. (2024). A hybrid deep learning framework for automatic detection of brain tumours using different modalities. IEEE Transactions on Emerging Topics in Computational Intelligence, 9(2): 1216-1225.
https://doi.org/10.1109/TETCI.2024.3442889

[44] Yoon, S. (2025). Brain tumor classification using a hybrid ensemble of Xception and parallel deep CNN models, Informatics in Medicine Unlocked, 54: 101629. https://doi.org/10.1016/j.imu.2025.101629

[45] Yilmaz, S., Sen, S. (2020). Electric fish optimization: A new heuristic algorithm inspired by electrolocation. Neural Computing and Applications, 32(15): 11543-11578. https://doi.org/10.1007/s00521-019-04641-8

[46] Khan, A.H., Sarkar, S.S., Mali, K., Sarkar, R. (2022). A genetic algorithm based feature selection approach for microstructural image classification. Experimental Techniques, 46(2): 335-347. https://doi.org/10.1007/s40799-021-00470-4

[47] Bala, I., Karunarathne, W., Mitchell, L. (2025). Optimizing feature selection by enhancing particle swarm optimization with orthogonal initialization and crossover operator. Computers, Materials and Continua, 84(1): 727-744. https://doi.org/10.32604/cmc.2025.065706