Received: March 13, 2024. Revised: April 11, 2024.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Global Self-Attention Module Based Convolutional Neural Network for Brain Tumor Classification

Preeti Sadanand Topannavar[1]* **Dinkar M. Yadav[2]** **Varsha Bendre[3]**

[1]*School of Electrical and Communication Engineering, JSPM University, Pune, India*
[2]*Department of Electronics and Telecommunication,*
S. N. D. College of Engineering and Research Center, Yeola, India
[3]*Department of E&TC, Pimpri Chinchwad College of Engineering, Pune, India*
* Corresponding author's Email: topannavarp@gmail.com

Abstract: The early analysis of brain tumors plays a significant role in enhancing treatment and patient survival rates. The accurate diagnosis is critical to creating treatment plans that extend the lifetime of affected individuals. The existing methods have limitations like less accuracy and not able to learn global features. To overcome this issue, Global Self-Attention Module based Convolutional Neural Network (GSAM based CNN) is proposed for brain tumor classification. The Figshare dataset is preprocessed through data augmentation which is utilized to enhance the data size, and the median filter is employed to eliminate noise in an input image. The preprocessed images are given to Adaptive Kernel Fuzzy C Means (AKFCM) with Otsu thresholding for segmentation process. After segmentation, the Sine Cosine Reptile Search Algorithm (SCRSA) is employed for feature selection. Then, the features are provided to GSAM based CNN for brain tumor classification. The proposed model achieves better result on Figshare dataset on the metrics of accuracy, precision, sensitivity, specificity and f1-score, with values of 99.83%, 99.65%, 99.79%, 99.78% and 99.71%, correspondingly, when compared to the existing methods like Comprehensive Learning Elephant Herding Optimization based K-Nearest Neighbor (CLEHO based KNN) and Parallel Deep CNN (PDCNN) with data augmentation.

Keywords: Convolutional neural network, Global self-attention module, Histogram of oriented gradients, Otsu thresholding, Sine cosine reptile search algorithm.

1. Introduction

An abnormal growth of cells in the brain that is often classified as benign or malignant, is a brain tumor. The various types of brain tumors are, gliomas obtained from the glial cells, meningiomas arising from the meninges, and metastatic tumor, which are spread from the body to the brain [1]. Brain tumor symptoms are detected through image processing and the incorporation of Machine Learning (ML) in medical images [2]. Early diagnosis by identifying the abnormal patterns involves precise segmentation and classification, which aids in identifying the neurological disorders [3]. Enhancing patient survival rates contributes to improved treatments through pivotal early-stage diagnosis outcomes [4]. Therapeutic interventions increase the chances of successful treatment and recovery with more effective, timely interventions [5]. The methods of imaging like Computed Tomography (CT), Positron Emission Tomography (PET), and Magnetic Resonance imaging (MRI), require major pieces of information to be gathered about the tumor, such as its type, size, shape, and location for an accurate diagnosis [6, 7]. MRI is the most utilized method in clinical settings, and is favourable because of its superior attributes of detailed and non-invasive nature [8].

The brain tumor detection from MRI images is a challenging task because of the requirements of trademark [9]. The accuracy of brain tumor diagnosis is amplified through the process of classification. Nowadays, the methods based on Deep Learning

(DL) are utilized in medical image classification [10]. Moreover, it has a probability for obtaining noises such as salt & pepper, Gaussian and speckle from the MRI images [11]. Therefore, the noise removal technique is significant for diminishing the noise [12]. The feature selection plays a main part in the classification, because it reduces the estimation time and enhance the classification performance [13]. The DL application produces an ideal solution because it extracts prominent features from the image, better than manually extracted features [14]. The newly implemented techniques for classification based on DL necessitate masked images for the predictable result [15]. Definitely, those labels sustenance to handle the learning procedure at the classification stage, however this procedure takes huge computational time [16]. The main contributions of the research are given below:

- The GSAM based CNN proposed for brain tumor classification is utilized to learn global features and avoid vanishing gradient issues.
- The sine cosine operator is implemented in traditional RSA that avoids local minima trapping issues through a solution search space which selects the best features from segmented regions.
- The performance of the proposed GSAM based CNN is evaluated based on performance metrics of accuracy, precision, sensitivity, specificity, and f1-score which are classified into meningioma, glioma and pituitary tumor.

The remaining part of paper is arranged as follows: Section 2 provides literature review. Section 3 describes the proposed methodology, while Section 4 provides results and discussion, and section 5 gives the conclusion.

2. Literature Review

This section describes some of the recent literature work based on brain tumor classification using ML and DL techniques.

Talukder [17] implemented an effective DL for classifying brain tumor through reconstruction and fine-tuning. The classification process integrated preprocessing, reconstruction and fine-tuning to improve model effectiveness. The reconstruction was adapted through considering augmentation for resolving overfitting issues. Moreover, to reduce the computational speed, the image was rapidly regularized based on configuration that assisted to reimplement the augmentation procedure. The fine-tuning was utilized to include layers through enhanced architecture which classify tumors, while augmentation resolved an overfitting issue. The implemented model needed huge number of computational memory and sources because of its deep architecture.

Mishra and Kaur [18] developed a texture and deep features-based brain tumor classification through Magnetic Resonance Images (MRI). The CLEHO was deployed to classify MRI images. The texture features were extracted through Gabor 2D filter and deep features were removed over DenseNet121 CNN. Furthermore, normalization was done through three techniques of Mean Median Absolute Deviation (MMAD), Z-score and Tanh. The KNN was employed to classify the MRI images. The developed model enhanced the classification accuracy, however, computational cost was high because of calculating distance among data points for every training sample.

Rahman and Islam [19] introduced a PDCNN for classification of MRI brain tumors. The PDCNN was utilized to extract local and global features from dual parallel steps and deal by overfitting issues through employing dropout with Batch Normalization (BN). The input image was resized and grayscale assisted to diminish the complexity, then data augmentation was used to enhance the dataset. The parallel pathway was generated through integrating dual synchronized DCNN with various window size which enabled to learn the global and local data. However, it was incapable to manage complex data due to its vanishing gradient problems.

Abd El-Wahab [20] suggested a Brain Tumor Classification using Fast Convolution Neural Network (BTC-fCNN). The developed model was employed to diminish the computational cost, parameters and processing time. The average pooling layer was applied to resolve the issue of overfitting because of unnecessary optimization. The suggested model utilized a sparse categorical cross-entropy loss function and Adam optimizer through primary finishing procedure. The suggested model overcame the overfitting issue because of unnecessary parameters but was computationally intensive and struggled with noisy data.

Swaraja [21] presented an integrated feature optimum selection through transfer learning for brain tumor segmentation and detection. After preprocessing, segmentation was utilized to segment the best features which separate one class from another through incorporating handcrafted and DL features. These features are optimized through Particle Swarm Optimization (PSO) algorithm. Lastly, these features are classified through ensemble classifier. For feature extraction, GoogleNet was employed with CNN for extracting DL features. It diminished the number of data for training and

enhanced the accuracy through pre-trained techniques. However, this model was unable to augment certain features of brain images, therefore leading to misclassification.

Ejaz Ul Haq [22] introduced a Deep CNN (DCNN) for brain tumor classification using BRATS2018 and Figshare dataset. The GoogleNet architecture are expanded and provided two various forms to attain multi-scale features from MRI for recognition and classification of brain tumor. The cascaded pooling layer was developed in two GoogleNet structures, that integrates features and reduce the complexity. The data augmentation was applied for brain tumor classification issues and it is considered as enhancement. However, it unable to detect and classify various sub-regions automatically.

Ajay S. Ladkat [23] suggested a Deep Neural Network-based tumor segmentation by mathematical model. Each slice of 3D image was improved by mathematical model which was transferred through 3D attention U-net for produce segmented result. It generated accurate tumor pixel segmentation from 3D brain images. The feature extraction was performed as main criterion. It enhances the human lifespan and diminishes death rate through huge accuracy and less complication rates. However, it has inadequate memory, false tolerance and necessitated huge computational time.

Ayesha Jabbar [24] developed a hybrid Caps-VGGNet model for brain tumor multi-grade segmentation and detection. The Caps-VGGNet incorporated the CapsNet and VGGNet through including VGGNet layers. It was used to optimize through accelerated and stochastic optimizer which user gradient and provides effective gradient descent in precision and speed. The aim was to enhance the generalization and produce precise detection through localization. The Caps-VGGNet enhanced the patient result and accuracy of tumor detection. However, interpretability was inadequate due to the difficult nature of VGGNet and CapsNet.

3. Proposed methodology

In this research, GSAM based CNN is proposed for brain tumor classification. The Figshare dataset is utilized in this research, comprising 3064 images with three different classes. The preprocessing is done by using data augmentation which is utilized to enhance the data size, while the median filter is employed to eliminate noise in an input image. The preprocessed images are given to AKFCM with Otsu thresholding for the segmentation process. After segmentation, the SCRSA is employed for feature selection. Then, the features are provided to GSAM based CNN for brain tumor classification. Fig. 1 is a block diagram of the proposed method.

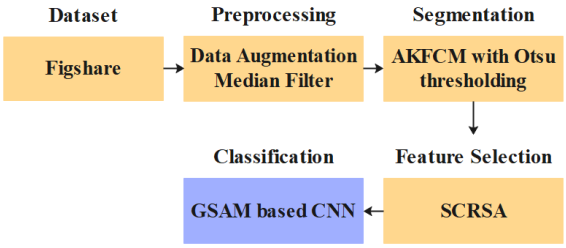

Figure. 1 Block diagram of proposed methodology

Table 1. Dataset positional views

Positional Views	Classes			Total
	Meningioma	Glioma	Pituitary	
Axial	209	494	291	994
Coronal	268	437	319	1024
Sagittal	231	494	320	1046

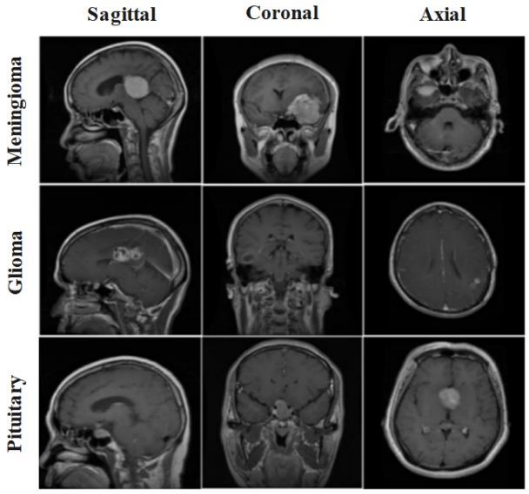

Figure. 2 Dataset sample images

3.1 Dataset

The Figshare dataset [25] which contains 3064 Contrast-Enhanced T1-weighted (CE-T1 MRI) brain images of 233 patients is utilized in this research. It comprises three tumor types, namely, 708 slices of meningioma, 1426 slices of glioma and 930 slices of pituitary tumor. Table 1 presents the MRI images of the dataset in three positional views. Fig. 2 shows the sample images.

3.2 Preprocessing

The data preprocessing is a procedure of exchaning raw data to a suitable form where the data taken from various sources contains inadequate data. For further examination, the dataset is essential for augmentation and filtration.

3.2.1. Data augmentation

Data augmentation is a technique that increases the size of dataset images by making small changes or modifications in the existing data. Various modifications are made to maximize the size of the dataset. By utilizing data augmentation, chances of over-fitting are reduced and the performance of the model is maximized. A range rotation, translation of height and weight, range of brightness, and vertical and horizontal flips are performed. Then, the data augmentation size of the dataset is increased by 4 times i.e., 12,256 images.

3.2.2. Median filter

The median filter is a procedure of nonlinear technique that eliminates noise from the MRI brain tumor images. It is processed by changing the image pixel, and adjusting each value by median value of neighboring pixel. The pixel is estimated through separating pixel from the adjacent form into a scientific order, and adjusting the pixel that is taken as the average pixel value. It efficiently eliminates noise without diminishing the sharpness of image, as presented in Eq. (1).

$$\hat{f}(x,y) = median\{g(s,t)\}, \quad where\ (s,t) \in S_{xy} \tag{1}$$

Where, S_{xy} is a group of coordinates in a rectangular image that comprises center at (x,y). The $\hat{f}(x,y)$ is a restored image, $g(s,t)$ is a calculated and corrupted area under S_{xy}. By utilizing the median filter, the noise is removed in an input image, replacing each pixel with a median value in its local adjacent. After pre-processing, the brain images are given as input to AKFCM with Otsu thresholding for brain tumor segmentation.

3.3 Segmentation

The preprocessed images are segmented through an AKFCM with Otsu thresholding technique. In the segmentation process, centroids used prove to be highly efficient, compared to other clustering techniques. These centroids are artificially initialized to ensure an efficient performance. The segmented process involves several steps for clustering tumors as shown as below:

Step 1: The primary cluster centroid, convergence condition and total number of cluster and adjacent pixels are set in filter window of the image.

Step 2: The adaptive weighting mean factor of the filtering image is calculated.

Step 3: The efficient membership function is estimated.

Step 4: Reduce the objective function through clustering centroids and best membership function.

Step 5: The absolute value is calculated.

Step 6: The tumor segment deploys clustering in the best way with an efficient membership degree rule.

In the clustering step involves applying Otsu thresholding to the input image for effective segmentation of tumor regions. Discriminate image analysis is involved to automatically utilize Otsu thresholding technique by considering the histogram shapes. The thresholding technique initializes the clustered image for brain tumor segmentation and comprises the foreground region f_r and background region b_r. In this technique, a discrete probability density function is utilized to construct a normalized histogram which is presented in Eq. (2).

$$p_r(r_q) = \frac{s_q}{s}, q = 0,1,2,\dots,h-1 \tag{2}$$

Where, h is the maximum intensity level for clustered image, s denotes the cluster image's total number of pixels, s_q is the practical pixel value and intensity level is denoted as r_q. The prior threshold value is a midpoint in-between, to access, minimum and maximum intensity values for clustered brain tumor image. The optimal threshold value is 0^* which increases the class variance σ^2, which is presented in Eqs. (3) and (4).

$$\sigma^2 = e_0, e_0 = \sum_{q=0}^{0-1} p_q(r_q) \tag{3}$$

Here,

$$\sigma^2(0^*) = \frac{arg}{0<o<h-1} max\sigma^2(0) \tag{4}$$

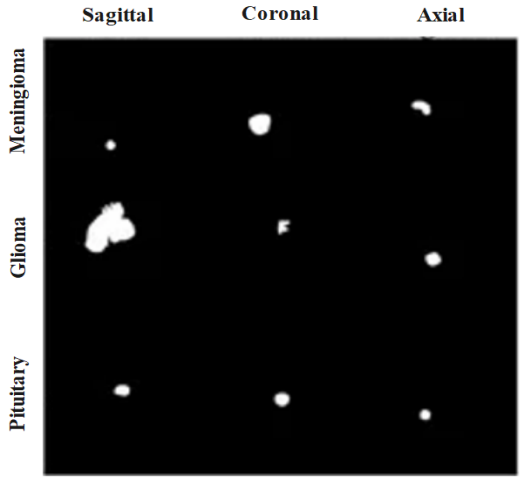

Figure. 3 Segmented images

If 0 is considered in the starting point of the threshold value, then f_r indicates the level of the pixel is $[0,1,\dots 0,-1]$, where b_r pixel level is set in the image $[0,0+1,\dots h-1]$. The segmented images are given to SCRSA for selecting the best features. Fig. 3 presents the segmented images.

3.4 Feature selection

The optimization techniques are utilized to find the best features among the probable combinations of feature subset, thereby minimizing the computational complexity. The traditional RSA solves different optimization problems. Yet, for solving huge dimensional nonconvex optimization issues, traditional RSA has certain limitations like slow convergence speed, huge executional complexity and local minima trapping. Hence, to overcome these problems, certain modifications are implemented in the traditional RSA algorithm. The sine operator is implemented in the traditional RSA so as to provide the global exploration capability. Hence, including sine operator in RSA avoids the local minima trapping through leading a search of the solution space. To improve the exploration phase, sine operator is involved in walking phase of traditional RSA. The sine operator in traditional RSA avoids the local minima issue through the search of solution space. The mathematical formula for sine operator in RSA is presented in Eq. (5).

$$x_{jk}(\tau+1) = Best_k(\tau) + r_1 \times \sin(rand) \times \left| r_2 \times Best_k(\tau) - x_{jk} \right| \quad for\ \tau \le \frac{T}{3} \tag{5}$$

Where, r_1, r_2 and $rand$ are randomly selected numbers among 0 and 1, x_{jk} is a present position and $Best_k$ is a best solution. Levy flight is a random process which follows levy distribution function. The numerical formula of levy flight is presented in Eq. (6).

$$levy = 0.01 \times \frac{u}{v^{\frac{1}{\zeta}}} \tag{6}$$

Where, u and v is a probability distribution which is mathematically presented in Eqs. (7), (8) and (9).

$$u \sim (0, \sigma_u^2), v \sim (0, \sigma_v^2) \tag{7}$$

$$\sigma_u = \frac{\delta(1+\zeta) sin\frac{\pi\zeta}{2}}{\delta\left[\frac{1+\zeta}{2}\right] \zeta * 2^{\zeta-\frac{1}{2}}} \tag{8}$$

$$\sigma_v = 1 \tag{9}$$

Here, δ is a standard gamma function. ζ is a substantial parameter in levy fight which defined jump size. The minimum score of ζ outcomes in little random stages enables solution candidates for searching field closest to acquired solution, thereby enhancing the exploitation ability. This enhanced exploitation assures global convergence. The numerical expression of levy fight for updating location in the last phase of SCRSA is presented in Eq. (10).

$$x_{jk}(\tau+1) = Best_k(\tau) + randn \times levy \oplus \left(x_{jk} - Best_k(\tau)\right), for\ \tau \le T\ and\ \tau > 3\frac{T}{4} \tag{10}$$

Where, $\oplus$ is an entry-wise multiplication and $randn$ is the uniformly distributed random number. The traditional RSA suffers from huge time complexity and less convergence speed. The SCRSA improves the global exploration, effectiveness, convergence speed, with lesser time complexity.

3.5 Classification

The selected features are classified through GSAM based CNN model which is deployed to learn global features. The CNN contains four parts: convolution (conv), Self-Attention (SA), Fully Connected (FC) and softmax layers. Before utilizing SA layer, conv layers are used to perform with inputs. The CNN layer employs $2D$ conv layers in the time and frequency domain of the model, and this is useful to diminish variance of time-frequency due to transformation invariance of the conv process. Moreover, CNN layer is employed as positional encoding to input and obtains respective position data due to the input. In this research, CNN layer has 2 conv layers with 32 filters, a $2D$ max pooling, and 2 conv layers with 64 filters. The CNN layer contains liner layer which converts input frames to their respective dimension, and the next 4 self-attention and 3 FC layers. SA layer is utilized for capturing a long range data of input frames through its effective attention mechanism, and FC layer mapped features to a much separable space. The last 2 layers are linear and softmax. Linear layer maps dimensions of each input frame to a vocab size of dictionary. Next, each frame develops probability distribution through the softmax layer. Fig. 4 presents the process of GSAM based CNN.

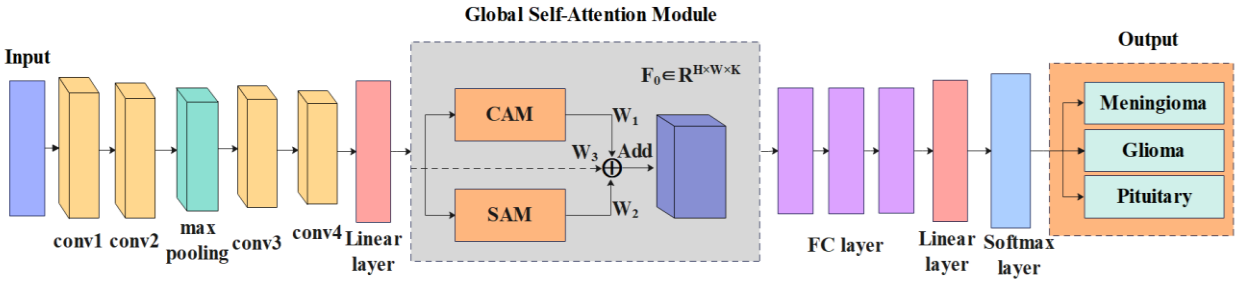

Figure. 4 Process of GSAM based CNN

3.5.1. Global self-Attention module

The GSA is established for learning much relevant and discriminative features from brain images. The self-attention allows interaction between spatial locations; in this research, GSAM implements two self-attention modules, Channel Attention Module (CAM) and Spatial Attention Module (SAM) for allowing communication between various channels and spatial locations. Lastly, feature maps are attained through every module that is incorporated through the input feature map, by an integration approach for attaining global attention feature map.

3.5.2. Channel attention module (CAM)

The aim of attention is to define feature inter-dependency through the dimensions of the channel and implement CAM for estimating implication of each channel. Feature input map $F_1 \in R^{H\times W\times K}$ is given to 3 CAM. In initial branch, Global Max Pooling (GMP) is processed through 1×1 Conv, and Reshape Operation (RO) is used for acquiring query tensor $Q_{ch} \in R^{1\times K}$. Then, Global Average Pooling (GAP) and 1×1 Conv is implemented through the RO, where the transpose process is employed for acquiring key tensor $K_{ch} \in R^{K\times 1}$. Next, Q_{ch} and K_{ch} are multiplied through softmax activation for acquiring attention weights $F'_{ch} \in R^{K\times K}$ and the numerical expression is presented in Eq. (11).

$$F'_{ch} = \sigma(K_{ch} \otimes Q_{ch}) \tag{11}$$

Where, σ is a softmax and $\otimes$ is a cross-product. In the next branch, 1×1 Conv block is presented next by RO for producing value tensor $V_{ch} \in R^{HW\times K}$. Then, V_{ch} and F'_{ch} are multiplied next by RO for acquiring CAM feature tensor $F''_{ch} \in R^{H\times W\times K}$ which is formulated in Eq. (12).

$$F''_{ch} = V_{ch} \otimes F'_{ch} \tag{12}$$

Lastly, feature map F''_{ch} is accumulated element-wise through F_1 for yielding F_{ch} that is global channel wise attention map which is formulated in Eq. (13).

$$F_{ch} = F''_{ch} \odot F_1 \tag{13}$$

Where, $\odot$ is a process of element wise multiplication, in which, every 1×1 Conv block has swish activation and normalization layer.

3.5.3. Spatial attention module (SAM)

The optic cup and disc areas in brain images of the affected patient majorly differs in size. Therefore, SAM is implemented for determining connection among each location within the feature map which is respective for SA. The SAM considers $F_1 \in R^{H\times W\times K}$ as being fed to the three branches. In initial and final branch, 1×1 Conv through K' channels followed by RO are presented for obtaining query tensor $Q_{sp} \in R^{HW\times K'}$ and the value tensor $V_{sp} \in R^{HW\times K'}$, where $K' = K/2$. Then, branch has 1×1 Conv through K' channels by RO and transfer procedure for obtaining key tensor $K_{sp} \in R^{K'\times HW}$. Next, multiplication among Q_{sp} and K_{sp} is managed through softmax activation for acquiring spatial attention weights $F'_{sp} \in R^{HW\times HW}$, as expressed in Eq. (14).

$$F'_{sp} = \sigma\left(Q_{sp} \otimes K_{sp}\right) \tag{14}$$

Then, F'_{sp} is multiplied with V_{sp} that is next by RO and 1×1 Conv block with K channels for obtaining spatial wise attention feature tensor $F''_{sp} \in R^{H\times W\times K}$ which is expressed in Eq. (15).

$$F''_{sp} = Conv\left(F'_{sp} \otimes V_{sp}\right) \tag{15}$$

Where, V_{sp} is a value tensor for spatial attention. Lastly, F_{sp} is attained through processing element wise accumulation among F''_{sp} and F_1 that is expressed in Eq. (16).

$$F_{sp} = F_{sp}'' \odot F_1 \quad (16)$$

Where, F_{sp}'' is a spatial wise attention feature tensor and $\odot$ is a process of element wise multiplication.

3.5.4. Fusion strategy

The results of feature map from every module are combined through actual feature maps F_1 with weighted sum by weights W_1, W_2 and W_3 for attaining global feature map $F_O \in R^{H \times W \times K}$. It enables model for emphasizing on tumor regions while ignoring inappropriate regions through catching numerous class-specific features which is expressed in Eq. (17).

$$F_O = W_1 F_{ch} \oplus W_2 F_{sp} \oplus W_3 F_1 \quad (17)$$

Where, W_1, W_2 *and* W_3 are the learnable scalar weights, F_{ch} and F_{sp} are a feature map of channel and spatial attention, F_1 and F_o are actual and output feature map, correspondingly. Then, the FC, linear and softmax layer are employed to classify the images into meningioma, glioma and pituitary tumor.

4. Experimental analysis

The GSAM based CNN is stimulated through Python with a system configuration of 16GB RAM, intel core i7 processor and windows 10 operating system. The accuracy, precision, sensitivity, specificity, f1-score, Dice Score Coefficient (DSC), Structural Similarity Index (SSIM), Intersection-Over-Union (IoU) and Mean IoU (MIoU) are employed to estimate the GSAM based CNN performance which is mathematically shown in Eqs. (18-26).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (18)$$

$$Precision = \frac{TP}{TP+FP} \quad (19)$$

$$sensitivity = \frac{TP}{TP+FN} \quad (20)$$

$$Specificity = \frac{TN}{TN+FP} \quad (21)$$

$$F1-score = 2 \times \frac{Precision \times sensitivity}{Precision+sensitivity} \quad (22)$$

$$DSC = \frac{2 \times TP}{(TP+FP)+(TP+FN)} \quad (23)$$

$$SSIM = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (24)$$

$$IoU = \frac{TP}{TP+FP+FN} \quad (25)$$

$$MIOU = \frac{1}{k+1}\sum_{i=0}^{k} \frac{TP}{FN+FP+TP} \quad (26)$$

Where, TP, TN, FP and FN respectively denote True Positive, True Negative, False Positive and False Negatives, μ_x and μ_y are the average of x and y, σ_x^2 and σ_y^2 are the variance of x and y, σ_{xy} is the covariance of x and y.

4.1 Quantitative and qualitative analysis

The quantitative and qualitative analysis of GSAM based CNN performance is evaluated based on accuracy, precision, sensitivity, specificity and f1-score. The Recurrent Neural Network (RNN) is incapable to handle lengthy data due to the vanishing gradient problems. The Long-Short Term Memory (LSTM) requires maximum training data and is slow for training on large datasets. The proposed GSAM based CNN is utilized to learn global features and avoid the vanishing gradient issues.

Table 2 and Fig. 5 show the performance of AKFCM with Otsu thresholding on Figshare dataset. The performances of FCM, AKFCM and Otsu

Table 2. Performance of segmentation

Method	DSC	SSIM	IoU	MIOU
FCM	0.9163	0.9155	0.9152	0.9146
AKFCM	0.9284	0.9276	0.9273	0.9264
Otsu thresholding	0.9376	0.9357	0.9348	0.9343
AKFCM with Otsu thresholding	0.9568	0.9562	0.9556	0.9551

Figure. 5 Performance of segmentation

Table 3. Performance of feature selection

Method	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
WOA	92.69	92.61	92.57	92.51	92.48
GOA	93.74	93.65	93.61	93.57	93.52
RSA	95.51	95.43	95.48	95.36	95.33
SCRSA	96.86	96.77	96.81	96.73	96.65

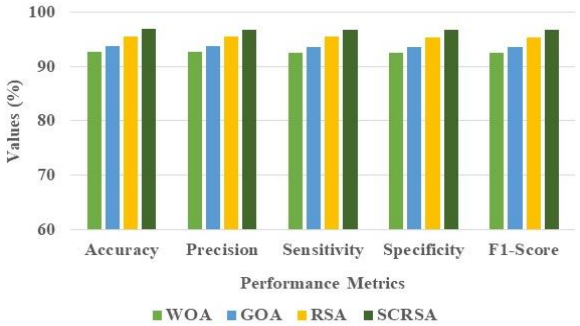

Figure. 6 Performance of feature selection

thresholding are measured and matched with that of the AKFCM with Otsu thresholding. The obtained result shows that the AKFCM with Otsu thresholding achieves superior results by utilizing DSC, SSIM, IoU and MIOU values about 0.9568, 0.9562, 0.9556 and 0.9551, correspondingly, while compared with other techniques.

Table 3 and Fig. 6 exhibit the SCRSA's outcomes on the Figshare dataset. The Whale Optimization Algorithm (WOA), Grasshopper Optimization Algorithm (GOA) and RSA performances are matched with SCRSA. The attained results illustrate that the SCRSA achieves superior results with accuracy, precision, sensitivity, specificity and f1-score values of about 96.86%, 96.77%, 96.81%, 96.73% and 96.65%, correspondingly, in contrast to the other techniques.

Table 4 and Fig. 7 present the GSAM's results on Figshare dataset. The Residual Attention Module (RAM), CAM and SAM performances are matched with GSAM. The attained results signify that GSAM achieves better results with accuracy, precision, sensitivity, specificity and f1-score values of about 98.61%, 98.47%, 98.56%, 98.52% and 98.49%, respectively, in contrast to the previous techniques.

Table 4. Performance of attention module

Method	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
RAM	94.75	94.68	94.71	94.63	94.56
CAM	96.87	96.75	96.83	96.72	96.64
SAM	97.56	97.51	97.48	97.39	97.35
GSAM	98.61	98.47	98.56	98.52	98.49

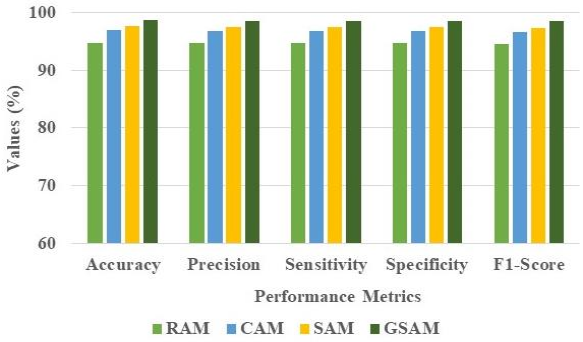

Figure. 7 Performance of attention module

Table 5. Performance of classification

Method	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
RNN	96.59	96.43	96.51	96.38	96.46
LSTM	97.76	97.71	97.68	97.42	97.35
CNN	98.67	98.51	98.63	98.56	98.48
GSAM based CNN	99.83	99.65	99.79	99.78	99.71

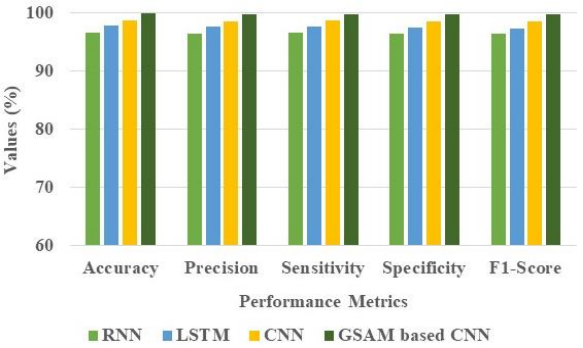

Figure. 8 Performance of classification

Table 5 and Fig. 8 demonstrate the GSAM based CNN's performance on Figshare dataset. The performances of GSAM with RNN, LSTM and CNN are matched with that of the GSAM based CNN. The attained results evidence that GSAM based CNN achieves better results through respective accuracy, precision, sensitivity, specificity and f1-score values of 99.83%, 99.65%, 99.79%, 99.78% and 99.71% more preferable than the comparative techniques.

4.2 Comparative analysis

This section demonstrates the comparative analysis of GSAM based CNN outcomes on performance metrics of accuracy, precision, sensitivity, specificity and f1-score, as shown in Table 6. The existing techniques of [17-24] are utilized for evaluate the ability of the classifier. The GSAM based CNN achieves better results than existing methods.

Table 6. Comparative analysis

Dataset	Methods	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Figshare	ResNet50V2 [17]	99.68	99.49	99.78	N/A	99.64
	CLEHO based KNN [18]	98.97	99.53	98.91	99.33	99.42
	PDCNN with data augmentation [19]	97.60	97.00	97.00	N/A	97.00
	BTC-fCNN [20]	98.86	98.72	98.83	98.77	99.41
	MSVM based optimum features [21]	94.9	98.7	99.5	99.7	N/A
	DCNN [22]	97.86	96.50	96.77	96.18	N/A
	Proposed GSAM based CNN	**99.83**	**99.65**	**99.79**	**99.78**	**99.71**
BRATS2018	DCNN [22]	96.98	95.09	95.97	95.61	N/A
	Proposed GSAM based CNN	**98.39**	**98.14**	**97.73**	**97.41**	**97.26**
BRATS2019	Mathematical model with 3D attention U-net [23]	98.90	99	98	N/A	98.50
	Caps-VGGNet [24]	98.9	N/A	98.1	98.7	98.1
	Proposed GSAM based CNN	**99.26**	**99.12**	**98.63**	**98.85**	**98.79**
BRATS2020	Caps-VGGNet [24]	99.6	N/A	98.5	99	98.4
	Proposed GSAM based CNN	**99.71**	**99.27**	**98.74**	**99.19**	**98.68**

4.3 Discussion

The advantages of GSAM based CNN and drawbacks of the existing methods are discussed in this section. The ResNet50V2 [17] needs advanced approaches to enhance the model's performance. The CLEHO based KNN [18] requires huge number of data for training. The PDCNN with data augmentation [19] is incapable of managing complex data due to its vanishing gradient problems. The BTC-fCNN [20] enhanced the required learning time to improved number of learning parameters. The MSVM based optimum features [21] was unable to augment certain features of the brain images. The DCNN [22] unable to detect and classify various sub-regions automatically. The Mathematical model with 3D attention U-net [23] has inadequate memory, false tolerance and necessitated huge computational time. In Caps-VGGNet [24] interpretability was inadequate due to the difficult nature of VGGNet and CapsNet. The GSAM based CNN overcomes limitations of the existing model. The GSAM based CNN is deployed to learn global features and avoid vanishing gradient issues.

5. Conclusion

This paper suggests a GSAM based CNN for brain tumor classification. The Figshare dataset is preprocessed through data augmentation which is utilized to enhance the data size, while the median filter is employed to eliminate noise in an input image. The preprocessed images are given to AKFCM with Otsu thresholding for the segmentation process. After segmentation, the SCRSA is employed for feature selection. The SCRSA enhances the global exploration, efficiency, convergence speed, with a lesser time complexity. Then, features are provided to GSAM based CNN for brain tumor classification. The GSAM based CNN achieves more robust results on the Figshare dataset on the metrics of accuracy, precision, sensitivity, specificity and f1-score values of 99.83%, 99.65%, 99.79%, 99.78% and 99.71%, respectively, in contrast to the existing methods. In future, various deep learning methods can be used for the more effective performance.

Notation

Notation	Description
S_{xy}	Group of coordinates
$\hat{f}(x,y)$	Restored image
$g(s,t)$	Calculated and corrupted area
h	Maximum intensity level
s	Total number of pixels
s_q	Practical pixel value
r_q	Intensity level
σ^2	Class variance
r_1, r_2 and $rand$	Randomly selected numbers between 0 and 1
x_{jk}	Current position
$Best_k$	Best solution
u and v	Probability distribution
δ	Standard gamma function
ζ	Substantial parameter in levy fight
$\oplus$	Entry-wise multiplication
$randn$	Uniformly distributed random number
σ	Softmax function

$\otimes$	Cross-product
$\odot$	Process of element wise multiplication
Q_{sp}	Query tensor
V_{sp}	Value tensor
K_{sp}	Key tensor
F'_{sp}	Spatial attention weights
F''_{sp}	Spatial wise attention feature tensor
W_1, W_2 and W_3	Learnable scalar weights
F_{ch} and F_{sp}	Feature map of channel and spatial attention
F_1 and F_o	Actual and output feature map
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
μ_x and μ_y	Average of x and y
σ_x^2 and σ_y^2	Variance of x and y
σ_{xy}	Covariance of x and y

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd and 3rd author.

References

[1] M. Rizwan, A. Shabbir, A.R. Javed, M. Shabbir, T. Baker, and D. Al-Jumeily Obe, "Brain Tumor and Glioma Grade Classification Using Gaussian Convolutional Neural Network", *IEEE Access*, Vol. 10, pp. 29731-29740, 2022.

[2] H. Kibriya, R. Amin, J. Kim, M. Nawaz, and R. Gantassi, "A Novel Approach for Brain Tumor Classification Using an Ensemble of Deep and Hand-Crafted Features", *Sensors*, Vol. 23, pp. 4693, 2023.

[3] J. Kang, Z. Ullah, and J. Gwak, "MRI-Based Brain Tumor Classification Using Ensemble of Deep Features and Machine Learning Classifiers", *Sensors*, Vol. 21, pp. 2222, 2021.

[4] M.I. Sharif, M.A. Khan, M. Alhussein, K. Aurangzeb, and M. Raza, "A decision support system for multimodal brain tumor classification using deep learning", *Complex & Intelligent Systems*, Vol. 8, No. 4, pp. 3007-3020, 2022.

[5] K. Yuan, G. Chi, Y. Zhou, and H. Yin, "A novel two-stage hybrid default prediction model with k-means clustering and support vector domain description", *Research in International Business and Finance*, Vol. 59, pp. 101536, 2022.

[6] M. Aamir, Z. Rahman, Z.A. Dayo, W.A. Abro, M.I. Uddin, I. Khan, A.S. Imran, Z. Ali, M. Ishfaq, Y. Guan, and Z. Hu, "A deep learning approach for brain tumor classification using MRI images", *Computers and Electrical Engineering*, Vol. 101, pp. 108105, 2022.

[7] Ö. Polat, Z. Dokur, and T. Ölmez, "Brain tumor classification by using a novel convolutional neural network structure", *International Journal of Imaging Systems and Technology*, Vol. 32, No. 5, pp. 1646-1660, 2022.

[8] N. Kesav, and M.G. Jibukumar, "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*, Vol. 34, No. 8B, pp. 6229-6242, 2022.

[9] S.A. Nawaz, D.M. Khan, and S. Qadri, "Brain tumor classification based on hybrid optimized multi-features analysis using magnetic resonance imaging dataset", *Applied Artificial Intelligence*, Vol. 36, No. 1, pp. 2031824, 2022.

[10] S. Ahmad, and P.K. Choudhury, "On the Performance of Deep Transfer Learning Networks for Brain Tumor Detection Using MR Images", *IEEE Access*, Vol. 10, pp. 59099-59114, 2022.

[11] A. Veeramuthu, S. Meenakshi, G. Mathivanan, K. Kotecha, J.R. Saini, V. Vijayakumar, and V. Subramaniyaswamy, "MRI brain tumor image classification using a combined feature and image-based classifier", *Frontiers in Psychology*, Vol. 13, pp. 848784, 2022.

[12] S. Saeedi, S. Rezayi, H. Keshavarz, and S.R.N. Kalhori, "MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques", *BMC Medical Informatics and Decision Making*, Vol. 23, pp. 16, 2023.

[13] S. Shanthi, S. Saradha, J.A. Smitha, N. Prasath, and H. Anandakumar, "An efficient automatic brain tumor classification using optimized hybrid deep neural network", *International Journal of Intelligent Networks*, Vol. 3, pp. 188-196, 2022.

[14] M. Siar, and M. Teshnehlab, "A combination of feature extraction methods and deep learning for brain tumour classification", *IET Image Processing*, Vol. 16, No. 2, pp. 416-441, 2022.

[15] S.M. Vijithananda, M.L. Jayatilake, B. Hewavithana, T. Gonçalves, L.M. Rato, B.S. Weerakoon, T.D. Kalupahana, A.D. Silva, and K.D. Dissanayake, "Feature extraction from MRI ADC images for brain tumor classification using machine learning techniques", *Biomedical Engineering Online*, Vol. 21, pp. 52, 2022.

[16] S. Alsubai, H.U. Khan, A. Alqahtani, M. Sha, S. Abbas, and U.G. "Mohammad, Ensemble deep learning for brain tumor detection", *Frontiers in Computational Neuroscience*, Vol. 16, pp. 1005617, 2022.

[17] M.A. Talukder, M.M. Islam, M.A. Uddin, A. Akhter, M.A.J. Pramanik, S. Aryal, M.A.A. Almoyad, K.F. Hasan, and M.A. Moni, "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning", *Expert Systems with Applications*, Vol. 230, pp. 120534, 2023.

[18] H.K. Mishra, and M. Kaur, "Classification of brain tumour based on texture and deep features of magnetic resonance images", *Expert Systems*, Vol. 40, No. 7, p. e13294, 2023.

[19] T. Rahman and M.S. Islam, "MRI brain tumor detection and classification using parallel deep convolutional neural networks", *Measurement: Sensors*, Vol. 26, pp. 100694, 2023.

[20] B.S. Abd El-Wahab, M.E. Nasr, S. Khamis, and A.S. Ashour, "BTC-fCNN: Fast Convolution Neural Network for Multi-class Brain Tumor Classification", *Health Information Science and Systems*, Vol. 11, pp. 3, 2023.

[21] K. Swaraja, K. Meenakshi, H.B. Valiveti, and G. Karuna, "Segmentation and detection of brain tumor through optimal selection of integrated features using transfer learning", *Multimedia Tools and Applications*, Vol. 81, No. 19, pp. 27363-27395, 2022.

[22] E.U. Haq, H. Jianjun, K. Li, H.U. Haq, and T. Zhang, "An MRI-based deep learning approach for efficient classification of brain tumors", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 14, pp. 6697–6718, 2023.

[23] A.S. Ladkat, S.L. Bangare, V. Jagota, S. Sanober, S.M. Beram, K. Rane, and B.K. Singh, "Deep neural network-based novel mathematical model for 3D brain tumor segmentation", *Computational Intelligence and Neuroscience*, Vol. 2022, 2022.

[24] A. Jabbar, S. Naseem, T. Mahmood, T. Saba, F.S. Alamri, and A. Rehman, "Brain tumor detection and multi-grade segmentation through hybrid caps-VGGNet model", *IEEE Access*, Vol. 11, pp. 72518 – 72536, 2023.

[25] Figshare dataset link: https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset (Accessed on 23/01/2024).