Review of Computer Engineering Research
2023 Vol. 10, No. 3, pp. 110-121
ISSN(e): 2410-9142
ISSN(p): 2412-4281
DOI: 10.18488/76.v10i3.3500

Creating a deep learning model using a Swin Transformer and tree growth optimisation to classify brain tumour

K. Sankar[1]
V. Gokula Krishnan[2+]
S. Sendil Kumar[3]
P. Pushpa[4]
B. Prathusha Laxmi[5]

[1]*CSE Department, CVR College of Engineering, Hyderabad, Telangana, India.*
Email: sankarkrish78@gmail.com
[2]*Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, Tamil Nadu, India.*
Email: gokul_kris143@yahoo.com
[3]*Department of EEE, S A Engineering College, Chennai, Tamil Nadu, India.*
Email: sendilkumar2009@gmail.com
[4]*Department of AIDS, Rajalakshmi Institute of Technology, Kuthambakkam, Chennai, Tamil Nadu, India.*
Email: ppushpacse88@gmail.com
[5]*Department of AIDS, R.M.K. College of Engineering and Technology, Gummidipoondi, Tamil Nadu, India.*
Email: prathushalaxmi@gmail.com

(+ Corresponding author)

Article History
Received: 27 July 2023
Revised: 22 September 2023
Accepted: 3 October 2023
Published: 16 October 2023

Keywords
Adaptive weighted frost filter
Brain tumor
Classification
Swin transformer
Deep learning
Tree growth optimization.

ABSTRACT

The brain, which has billions of cells, is the largest and most complex organ in the human body. A brain tumor is the primary malignant intracranial tumor of the central nervous system that develops most frequently. They are frequently found too late for effective therapy. The use of minimally invasive procedures is necessary to make a diagnosis and monitor a tumor of the central nervous system's response to therapy. There exist three distinct classifications of tumors, namely benign, premalignant, and malignant. This study concentrated on using deep learning to identify brain tumors (BT) using normal or abnormal brain pictures. Numerous methodologies have been employed to augment the quality of images, encompassing image smoothing and noise restoration procedures. The present study employs the proposed Adaptive Weighted Frost filter as it has been identified as the optimal approach for noise reduction in BT photographs. The Swin Transformer technology is employed for the purpose of classifying the BT. The efficiency of the Tree Growth Optimization (TGA) model for Swin transformer hyper parameter tweaking has been evaluated in this work. Before using our unique BT dataset for extensive experimental comparisons, medical specialists carefully examined it down to the pixel level. The predicted model achieved the greatest F1 score of 99.82% and the maximum accuracy, recall, and 100%, respectively.

Contribution/Originality: This work focused on using normal or abnormal brain images to identify brain tumors (BT) using deep learning. To improve image quality, image smoothing, and noise restoration, numerous strategies have been employed. In this paper, the effectiveness of the Tree Growth Optimization (TGA) model for fine-tuning Swin transformer hyper-parameters has been assessed. The best classification accuracy ever was reached with the suggested model, coming in at 99.89%.

1. INTRODUCTION

The brain is a vital organ since it regulates the rest of the Central Nervous System (CNS). The connection between the brain and the bone marrow completes the CNS. The human brain is responsible for controlling all bodily

motions. It takes in data from a variety of sources, processes it, and then instructs the body on what to do next. The human brain is the body's primary control center [1]. Brain tumors are tumors that develop confidentially in the brain and are detected by the tissues directly underneath the skull or the brain [2]. Survival rates for people with brain tumors are higher if the cancer is caught early [3]. Differentiating benign (noncancerous) brain tumors from malignant (cancerous) ones is possible. Brain tumors that aren't malignant tend to have a clear border. They are localized and do not invade other organs. Even benign tumors have the potential to cause serious health problems by pressing on vulnerable brain areas [4]. Rapid metastasis of malignant tumors to other brain regions may worsen the BT patient's situation [5]. It is also possible to do a surgery called brain angiography, which allows doctors to visualize the veins in the brain that carry blood to a brain tumor [6]. Finding brain tumors and keeping tabs on their development can be a tedious and error-prone manual process. Figure 1 shows the flow of brain tumor detection using a simple DL model [7].

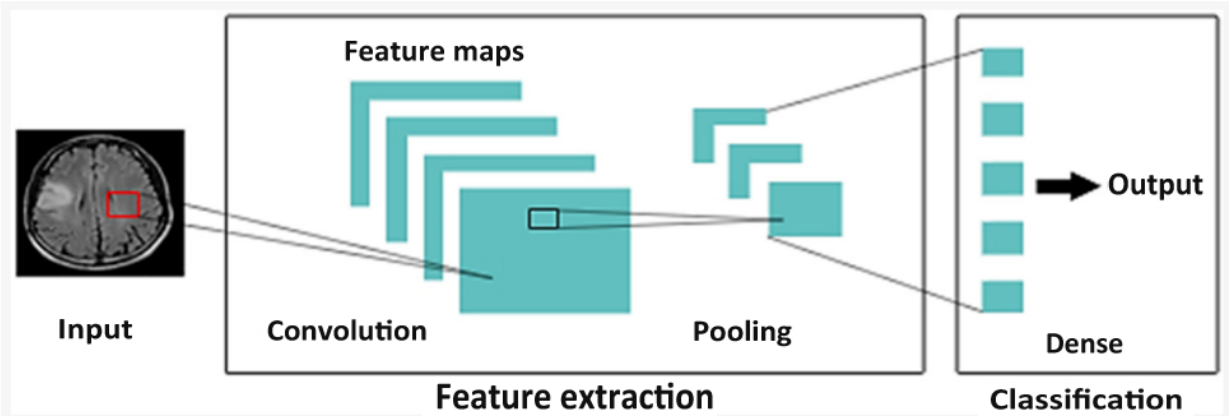

Figure 1. Workflow for detection of brain tumor using simple CNN.

Deep learning has supplanted traditional methods of machine learning [8]. Pre-trained deep learning models are appealing because they may save developers time and effort by eliminating the requirement for a sizable training dataset [9]. When compared to more conventional methods, deep learning-based systems rely heavily on the data used in training and require far less pre-processing [10]. Using a hierarchical framework and beginning with basic characteristics, deep learning is characterized as a method of learning at several levels. Magnetic Resonance Imaging (MRI) images of the brain are an actual means of diagnosing BT [11]. All of these recent breakthroughs in computer-based treatments and the interpretation of brain MRI data have one thing in common: they were all made likely by the extensive adoption of (CNNs) and deep learning (DL). CNN has had a significant impact on a variety of industries, including digital image processing, computer vision, and artificial intelligence. CNN is the most widely used method for image analysis because of its multi-layered design, despite the fact that various deep learning algorithms have been developed over the years [12].

In this paper,

- The Frost filter is cleverly modified in our suggested approach to eliminate noise from BT images.
- The brand-new categorization technique Swin Transformer combines Swin-B with two resolutions gained through pre-training and Swin-T obtained through steady training. Pre-training is used in the segmentation assignment to aid the model and enhance the precision of our investigations.
- Tree Growth Algorithm (TGA), a recently developed metaheuristic optimization method, is suggested for hyper parameter tuning in Swin Transformer.

The remaining part of the article is organized as follows: the related works of the proposed research are encompassed within Section 2. Proposed/anticipated works can be found in Section 3. Section 4 provides access to

the simulation results and analysis for each case study. Finally, Section 5 presents the concluding remarks of the article.

2. RELATED WORKS

Mahmud, et al. [13] provide a CNN architecture for precise brain cancer diagnosis using magnetic resonance (MR) images. This study also evaluates and contrasts the proposed architecture with a wide range of models, including ResNet-50 and V3. To assess the performance of each model, we looked at loss and AUC (area under the curve). We found that the proposed model outperformed both the alternatives and our earlier efforts. On a dataset with an accuracy of 93.3%, an AUC of 98.43%, a recall of 91.19%, and a loss of 0.25, we tested the CNN model. We discover that the proposed model is successful for the early diagnosis of various forms of brain cancer by comparing it to the others.

Ali, et al. [14] present a deep learning network (DLNets) feature optimization method using magnetic resonance imaging (MRI). In this study, sixteen pre-trained deep neural networks are used to calculate the attributes. Classification performance is measured using a cost function based on the support vector machine (SVM) and metaheuristic optimization algorithms (MOA); Harris Hawks optimization algorithm (HHOA), butterfly algorithm (BOA), whale optimization algorithm (WOA), grey wolf optimization algorithm (GWOA), bat algorithm (FA), and firefly algorithm (FFA). To identify the optimal network, we employ a deep-learning network selection strategy. We use the best features of deep neural networks to train the SVM perfectly. The suggested WBM-DLNets technique is tested on a publicly available dataset. The accuracy is greatly enhanced when utilizing the selected features as opposed to those gathered using the entire collection of deep features. The best accuracy of 95.7% is attained by DenseNet-201-GWOA and Net-b0-MOA. The WBM-DLNets approach is evaluated against the published literature's outcomes.

To better distinguish between low-grade and high-grade gliomas, as proposed by Tandel, et al. [15], three datasets were constructed using of T1-weighted fluid-attenuated inversion recovery (T1FLAIR) MRI sequences. For this tumor classification study, we used five widely used CNNs: GoogleNet and ResNet50. To provide more reliable and superior results than any single DL model, an ensemble technique was devised that takes into account the majority vote of more than five models. The five-fold approach was used for both the training and the testing phases. Maximum test accuracies of 98.88%, 97.98%, and 94.75% were attained, showing an improvement in accuracy of 4.17 percent compared to T1-Weighted-MRI data and 0.91 percent compared to T2-Weighted-MRI data. Therefore, FLAIR-MRI data is extremely important for diagnosing many forms of brain tumors. The suggested ensemble method (MajVot) outperformed AlexNet, Visual Geometry Group (VGG16), ResNet18, GoogleNet, and ResNet50 on average accuracy for all three datasets by 3.60 percent, 2.84 percent, 1.64 percent, 4.27 percent, and 1.14 percent.

Brain cancer is detected using a deep network optimized with stochastic gradient descent (SGD) in a paper by Asad, et al. [16]. Using the ResNet-50 model with the Kaggle brain-tumor dataset, which is open to the public, we conduct multi-classification of brain tumors. The method obtained a training and testing accuracy of 99.82 and 99.5%, respectively. The experimental data backs up the argument that the suggested model is an improvement over the current method and has potential for use in the treatment of other diseases.

Anand, et al. [17] describe a deep learning model for the diagnosis of brain cancers that makes use of an ensemble. The weighted ensemble classification model is built on the Convolution Neural Network (CNN) model without augmentation, the Transfer Learning VGG19 model, and the CNN model with augmentation. The optimum weights from the three distinct feature spaces are blended using a grid search. The necessary data for modeling, which consists of 3929 MRI pictures from 110 individuals with lower-grade gliomas, was contributed by the Cancer Genome Atlas (TCGA). Overfitting can be minimized by using an ensemble model, which integrates multiple models that have acquired knowledge of different facets of the data. When it comes to diagnosing brain cancers, the proposed ensemble model performs better than any of the three individual models in terms of accuracy, precision, and F1-score.

Radiologists may utilize the proposed model as a second opinion aid when making a diagnosis of brain cancer using MRI data.

This study by Özkaraca, et al. [18] creates a new modular model by retaining the benefits of well-known transfer learning techniques while avoiding their drawbacks. Kaggle was scoured for publicly available images of brain tumors. We used two distinct methods of partitioning the data to train the model. In the initial phase, we allocated 80% of the MRI image dataset for the purpose of training, while the remaining 20% was reserved for testing. The utilisation of tenfold cross-validation was also implemented. When compared with other, more established transfer learning methods, the innovative deep learning model performed better in classification but took more time to analyze the same MRI dataset.

3. PROPOSED WORKS

3.1. Adaptive Weighted Frost Filter for Quantum Noise Elimination

The unequal impact of photons on the screen is what causes the noise. This noise is typically created during the acquisition of images due to low count photons. Given that the degree of quantum strike unevenness is strongly correlated with the presence of noise, the Poisson distribution is the optimum choice to depict this noise due to its signal-dependent character. For a sufficiently large number of quanta shared per pixel, electronic noise can be expressed using the Gaussian model. With very little change that is difficult to perceive with the naked eye, quantum noise impacts every pixel of an image. The modified Frost filter is employed in this study to remove quantum noise from pictures of brain tumors. A popular noise-reduction filter is the frost filter. The computer vision (CV) in the current filter is defined as the ratio of the local standard deviation (SD) to the local mean (average) of the deformed image, as shown in Equation 1. The issue with the current filters is that they cannot create the correlation between pixels in an image. There is a very low chance that the core pixel will be distorted if the nearby pixels are almost connected. As a result, the filtered image won't give satisfying results.

$$DN = \sum_{nxn} k\alpha e^{-\alpha|t|} \quad (1)$$

$$\text{Where } \alpha = \left[\frac{4}{n\bar{\sigma}^2}\right]\left[\frac{\sigma^2}{ni^2}\right]$$

DN= “Digital number.

K= “Normalization continuous”.

$\bar{I}$= “Local weighted mean intended by using Equation 2”

σ^2 = “Local variance”.

$\bar{\sigma}$ = “Image coefficient of difference value”.

$$|t| = |X - X_o| + |Y - Y_o|$$

n = moving kernel size, where $|X - X_o|$ modulates the difference between the mean in the x-axis, whereas |Y-Y_o| modulates the change amid the mean in the y-axis. The Frost filter is modified to take a weighted local average of the corrupted picture to fix this issue. The proposed adaptive weighted filter uses a weighted average of the local window. Equation 2 is used to calculate the Weighted average of a 5×5 window selected empirically. Evidently, a filter that takes into account how neighboring pixels inside the moving window are correlated will produce noticeably superior-enhancing outcomes.

In order to replace the center pixel, the projected improved Adaptive Weighted Frost filter uses a weighted sum of values from an N × N window. The weighting factor also falls when the correlation between the central and remaining pixels does, and so on. Equation 2 with a 5×5 window size elaborates the modified Frost filter. Using an adaptive weighted frost filter, prepare digital images of brain tumours.

Equation 2 presents the weighted sum formula that is suggested.

$$weighted\_sum^k = \sum_{i=-2,i\neq 0}^{2}\left(\frac{w^{i*}|PIX^i|}{5}\right), -2 \le i \le 2 \quad (2)$$

Where,

$$W_i = \begin{cases} 2, & i = -1, +1 \\ 1, & i = -2, +2 \end{cases}$$

The weighted mean will be computed using the suggested weighted window.

Gray scale values are typically closer together if there is a smaller lag interval between two pixels. The correlation between adjacent pixels is then created by giving more weight to pixels with a lag distance of 1 and less weight to pixels with a lag distance of 2. As a result, the strong correlation of connected pixels is attained, which is crucial in this case because the nearby pixels greatly aid in the restoration of the core pixel of the deteriorated image. As a result, the weighted mean of the filter window has been used in place of the output pixel value.

3.2. Framework Description for BT Classification

To assist in identifying early BT, this work used Swin Transformer to conduct some BT research. Swin Transformer has three phases:

However, the transformer was initially developed for Natural Language Processing (NLP) and is used to analyze sequences of natural language; thus, CNN directly uses the picture as a matrix for processing. As a convolutional neural network (CNN), it is not very convenient for extracting features from images.

3.2.1. Patch Embedding

Patch division is used to convert RGB maps into patch blocks. In this instance, the patch size is 4 4 times the corresponding RGB stations, giving a size of 4 4 3 = 48. The predicable dimensions required to create a feature matrix are present in the processed patches.

3.2.2. Patch Merging

The feature matrix created in the previous phase is split into four 2 × 2 size windows, with the corresponding positions of each window being combined. The four featured media are then concatenated.

3.2.3. Mask

Only the continuous part of the window will execute self-attention, thanks to the way the mask is made. The right place the original window in matrix. Shift size and window size are related to one another using the following formula:

$$s = \left\lfloor \frac{\omega}{2} \right\rfloor \tag{3}$$

Where s is shift size, ω is the window scope.

Since this component of the mask matrix is not related to the section in the unique matrix, it must be divided at this time. Areas can be sliced in two different directions, horizontally and vertically, respectively, using the ranges. According to the window partition, the labelled mask matrix is broken up into sub-blocks of the window size. $\left\lfloor \frac{H}{\omega} \right\rfloor$ rows and $\left\lfloor \frac{H}{\omega} \right\rfloor$ columns proportionally, then combine the dimensions for the quantity and the size. The unique matrix mask must be cut into narrow windows so that they may be measured in window units; hence, this division is required.

3.3. Algorithm Design

3.3.1. First and Second Stages

We finished image processing in the first stage. The RGB brain tumor scan image is first divided into a sum of distinct patches, as is typical of transformer architectures. Each patch in the Swin Transformer configuration is 4 × 4 and has a dimension of 4×4×3 since each pixel has RGB three channel values. A linear embedding layer then transforms each patch into a C-dimensional feature matrix.

The block of the Swin Transformer is the second stage. Swin Transformer, like the majority of CNN designs, captures deep properties by stacking a number of blocks. In this study, four repeated attention blocks were utilized to teach visual features. The projected, processed patches are placed in the designated area. Prior to sending the input feature to Swin Transformer Block, we used linear embedding to partition it into C dimension. A multi-head self-attention shift window-based Swin Transformer block is included in multi-head self-attention (MSA). A shift window-based MSA and two levels of Multi-Layer Perception (MLP) are included in the Swin Transformer block. A layer specification (LN) layer comes before each connection. The Patch Merging operation then begins by stitching patches that are immediately in the 2 × 2 range. This makes the number of patch blocks $H/8 \times W/8$ and the feature dimension 4C. Following the same linear embedding process used in stage 1, 4C is compressed into 2C and sent into the Swin Transformer blocks. A multi layered same features mapping resolution as a traditional convolutional network is produced by the combination of these building elements.

3.3.2. Self-Attention in Non-Overlapped Windows

Due to its quadratic complexity in terms of token count, global calculation is not suitable for many vision applications that require large token sets with high-resolution pictures. It is instead possible to model well by computing within a limited window. In these areas, the photographs are regularly and non-overlappingly segmented. The assumption that each window has M × M patch increases the cost of the global MSA modules and the windows based on $h \times \omega$ patches photos.

$$\Omega(MSA) = 4h\omega C^2 + 2(h\omega)^2 C \quad (4)$$

$$\Omega(W - MSA) = 4h\omega C^2 + 2M^2 h\omega C \quad (5)$$

In this case, the former is proportional to the sum of patches hw, whereas the latter follows a linear path (with a default value of 7) when M is held constant. Scalability is an advantage of window-based self-attention over global self-attentive computing's high cost due to the massive hardware it requires.

3.3.3. Shifted Window Partitioning in Successive Blocks

Cross-window connections are added because the modelling capabilities of non-overlapping windows are severely constrained by the lack of information flow between them. The method involves two Swin Transformer blocks that alternate between Window-MSA and Shifted Window-MSA. The perceived range of view is expanded thanks to the shifted window division approach, which establishes a bridge between neighboring non-overlapping windows in the top layer. The Swin Transformer building blocks after the redesign are determined in the following way:

$$\hat{z}^l = W - MSA\left(LN(z^{l-1})\right) + z^{l-1},$$
$$z^l = MLP\left(LN(\hat{z}^l)\right) + \hat{z}^l,$$
$$\hat{z}^{l+1} = SW - MSA\left(LN(z^l)\right) + z^l,$$
$$z^{l+1} = MLP\left(LN(\hat{z}^{l+1})\right) + \hat{z}^{l+1}, \quad (6)$$

Where $\hat{z}^l$ and z^l signify the features of the *(S) W-MSA* unit and the *MLP* unit for block, correspondingly; *SW-MSA* are multi-head self-attention procedures that use normal and configurations, correspondingly.

3.3.4. Multi-Head Self-Attention

Migration from the transformer to dream is accomplished using a multi-head attention method. The precise formula is given below.

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad \left(Q,K,V \in R^{M^{2+d}}\right) \quad (7)$$

Where B, the relative parameter, is presented in a manner similar to that of the position embedding in the transformer. The size of d, which balances the sizes of QK^T and B, is the dimension that corresponds to each head.

Calculation of Q, K, and V: For the information from the incoming window, the relevant values for the inquiry, key, and value are gotten after a linear layer.

3.3.5. Third Stage

This research successfully classified data using a Swin Transformer, and it also describes how to utilize extract features.

The third step incorporates the downstream processes of classification and segmentation. Our research split participants into two groups: those with and those without nodules. The number of classifications is the classification mission's output dimension, and softmax is used to calculate the likelihood of a given classification. This lays the groundwork for feature extraction in images for the hyper parameter job.

3.4. Architecture Variants

We used the Swin Transformer as our basic model to achieve a model size and computational difficulty comparable to ViT-B and DeiT-C. Due to limitations in our experimental setup, we were only able to train on a tiny portion of Swin-T and Swin-S duplicates. Swin-T and Swin-S models have 0.25, 0.5, and 2 times increased computational complexity. Swin-T is comparable to ResNet-50 (DeiT-S) in terms of difficulty, and Swin-S to ResNet-101. The window size has a default value of M = 7. Every MLP has an expansion layer of 4, and every experiment's head is d = 32. The following are the architectures' hyper-parameters for the model versions:

- $Swin-T: C = 96, layer\ statistics = \{2, 2, 6, 2\}.$
- $Swin-S: C = 96, layer\ statistics = \{2, 2, 18, 2\}.$
- $Swin-B: C = 128, layer\ statistics = \{2, 2, 18, 2\}.$

Where C is the first stage's hidden layers' channel number.

3.5. Loss Function

The classifying mission's loss function is as follows:

$$L = \frac{1}{N}\sum_i L_i = -\frac{1}{N}\sum_i \sum_{c=1}^{M} y_{ic} \log\left(p_{ic}\right) \quad (8)$$

Where, y_{ic}. is the symbolic function and M is the number of categories (0 or 1). Take 1 if category 2 is the true category of sample I otherwise, take 0., p_{ij} .is the anticipated probability that category 2 is the true category of sample i.

Equation 8, which is pixel-classified, serves as the loss function for the hyper parameter tuning.

3.6. Tree Growth Optimization

This work has assessed the effectiveness of the Tree Growth Optimization (TGA) model for hyper parameter tuning in the Swin transformer. 2018 saw the launch of the Tress Growth Algorithm, a new metaheuristic approach inspired by nature developed by Cheraghalipour and his colleagues (TGA). The algorithm's intensification and diversification are the two stages of the TGA. Typically, intensification works by starting over with high-quality solutions or by changing the selection criteria to include qualities for these solutions. This permits the prime trees, which meet light absorption requirements, to compete for the role of food supply. At this point, the modernist strategy makes sure that the trees are moving in the direction of a superior food supply. This indicates that throughout the intensification stage, we are just aiming for the local or global optimum. The other trees compete with one another to absorb light and spread to new or uninhabited areas during the diversification phase. It is possible to strike a balance between intensification and diversification by modifying the parameters. Several common and engineering challenges were used to test the effectiveness and performance of the TGA. Four categories of trees are distinguished in woods. Due to favorable growing conditions, the first is referred to as the best tree group N1, and the second is

known as the competition for the light tree category. In the latter category, some trees slant their approaches towards the best trees in an effort to catch the sun. The replacement or removal group is also known as group N3, the third group. This group's major objective is to replace weak and unhealthy trees with new ones. The reproduction phase then follows, during which the last group, N4, is referred to as the reproductive group.

The following steps are used to explain the TGA principle: In the first, a random number generator creates the initial population of trees. Each tree's fitness value is determined, and the trees in the first group, N1, are ordered in climbing order based on the fitness values. The following formula is used to generate the new tree in N1.

$$X_i^{t+1} = \frac{X_i^t}{\theta} + rX_i^t \tag{9}$$

That θ is a reduction in energy use on the part of the tree as a result of ageing, increasing growth, and a lack of available food. The location of the solution (tree) in iteration t is indicated by a random number, r, which is spread among the values [0, 1], X_i^t. In the following formula, the distance between the best nearby trees is calculated after the N2 trees are moved into the second group in the second phase.

$$D_i = \left(\sum_i^{N_1+N_2}\left(X_{N_2}^t - X_i^t\right)^2\right)^{\frac{1}{2}} \tag{10}$$

The present tree then makes a move to win and get to the light as the competition for light intensifies. As a result, the following formula is used to determine the linear combination of the two nearby trees:

$$y = \lambda T_1 + (1 - \lambda)T_2 \tag{11}$$

Where the y used to regulate the effect of the closest tree is called.

Following is a mathematical description of how this tree moved between two nearby trees at different angles.

$$X_{N_2}^{t+1} = X_{N_2}^t + \alpha_i y \tag{12}$$

Where y represents the distribution angle, with a value between 0 and 1.

The weakest N3 trees are eliminated from the third group, and their replacements are produced using the following method:

$$N_3 = N = N_1 - N_2 \tag{13}$$

Last but not least, the fourth group's new trees are created around the best trees using the mask operator. The population is then expanded by including the most recent N4 trees. Based on the fitness value of the built-in population, savings are made. As a fresh population for the subsequent iteration, the best N trees are then represented. The algorithm's steps will be carried out again and again until the halting condition is satisfied. In the end, a better solution is found in the best global tree.

4. RESULT AND DISCUSSION

A freely accessible contrast-enhanced (CE)-MRI dataset [19] was used in this investigation. The 233 patients included in the dataset are from scans gathered during a 5-year period (2005-2010). There are a total of 3062 MRI pictures in the collection, representing 3 kinds of BT and 233 individuals. The pixels in the 2D volumes that make up the pictures in the collection are 0.49 mm2 in size, and the resolution is 512 pixels on a side. Three expert radiologists also hand-drawn a perimeter around the tumor area on the MRI scans. The .mat files used to represent the dataset on figshare measure 512 pixels on the square. Figure 2 shows the sample images of the dataset.

The purpose of this research was to accurately categorize three distinct Benchtop (BT-MRI) image types. The CE-MRI dataset was classified using a combination of models and the projected model, with 30% of the dataset utilized for testing and 70% for training. The outcomes were accomplished using computer resources, namely a Matlab environment, an i5 processor, and 8 Giga Bytes (GB) of Random Access Memory (RAM). The classification performance of deep learning networks may be evaluated in a number of ways. In the Swin Transformer procedure, a confusion matrix is often employed for classification. A confusion matrix is used for the computations. In Table 1, we can see the outline of the confusion matrix that was used in this study.

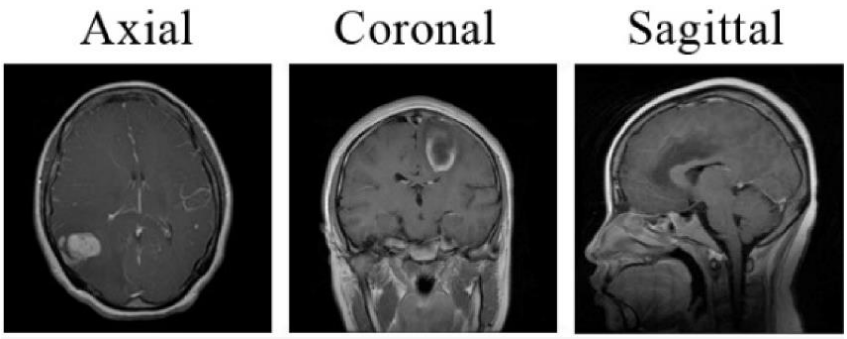

Figure 2. Sample images of three different planes of the MRI images.

Table 1. General confusion matrix.

Predicated classes				
		Meningiomas	**Pituitary**	**Gliomas**
Actual class	Meningiomas	PMM	PPM	PGM
	Pituitary	PMP	PPP	PGP
	Gliomas	PMG	PPG	PGG

Here, PGG stands for the dataset that projected glioma, PMG for the photos from the dataset that were meningiomas but were mistakenly forecasted as gliomas, and PPG for the images from the dataset that were pituitaries but were incorrectly predicted as gliomas. PGM stands for the photos from the dataset that were actually gliomas but were incorrectly identified as meningiomas, while PMM stands for the images from the dataset that were meningiomas but were identified as meningiomas. The datasets that were actually pituitary and incorrectly predicted as meningioma are designated as PPM; the images from the datasets that were actually glioma and incorrectly predicted as pituitary are designated as PGP; the images from the datasets that were actually meningioma and incorrectly predicted as pituitary are designated as PMP; and the datasets that were really pituitary and correctly predicted as pituitary are designated as PPP. Additionally, the following definitions include accuracy, precision, recall, and F1-score.

4.1. Performance Metrics

How many instances it correctly predicts out of an entire set of data serves as a gauge of a classification system's precision.

Equation 14 and Figure 1 depict the calculation used to determine precision:

$$Acc = \frac{TP+TN}{TP+FP+TN+FN} \tag{14}$$

A true negative (TN) is anything that is detrimental and is predicted adversely, whereas a false negative (FN) is something that is beneficial but is predicted negatively. True positive, abbreviated TP, refers to an approximation of the actual quantity of information. False positivity refers to something that is detrimental yet incorrectly diagnosed as positive.

Accuracy (optimistic anticipated values) may be defined as the fraction of forecasts that fall into the all-positive category. Equation 15 displays the equation for the accuracy value:

$$prec = \frac{TP}{TP+FP} \tag{15}$$

The recall metric is intended to be in-between the sum of positive class components and the sum of true positive (TP) outcomes. Equation 16 provides a method for determining the level of recall:

$$Recall = \frac{TP}{TP+FN} \tag{16}$$

The F1 score is a combination of the recall and accuracy scores. Equation 17 demonstrates how to compute the F1 score:

$$F1 - measure = \frac{2*prec*recall}{prec*recall} \quad (17)$$

Additionally, after the suggested model had been trained, the score performance values were assessed and are shown in Table 2.

Table 2. Confusion matrix of projected perfect.

Based classes				
Actual class		**Meningiomas**	**Pituitary**	**Gliomas**
	Meningiomas	211	00	01
	Pituitary	00	278	00
	Gliomas	01	01	426

Table 3. Experimentation consequences of pretrained representations.

Model	Accuracy	Recall	Precision	F1-store
DBN	96.8%	97.66%	96.6%	97.66%
CNN	97.37%	98.66%	97.33%	97.93%
GAN	98.46%	99.66%	98.42%	98.53%
Proposed model	99.89%	100%	99.76%	99.82%

In the analysis of accuracy, deep belief networks (DBN) achieved 96.8%, CNN achieved 97.37%, generative adversarial networks (GAN) achieved 98.46%, and the proposed model achieved 99.89%. In the analysis of Recall DBN achieved 97.66%, CNN achieved 97.33%, GAN achieved 99.66%, and the proposed model achieved 100%. Figure 3 and Table 3 present the graphical analysis and numerical analysis of the proposed model.

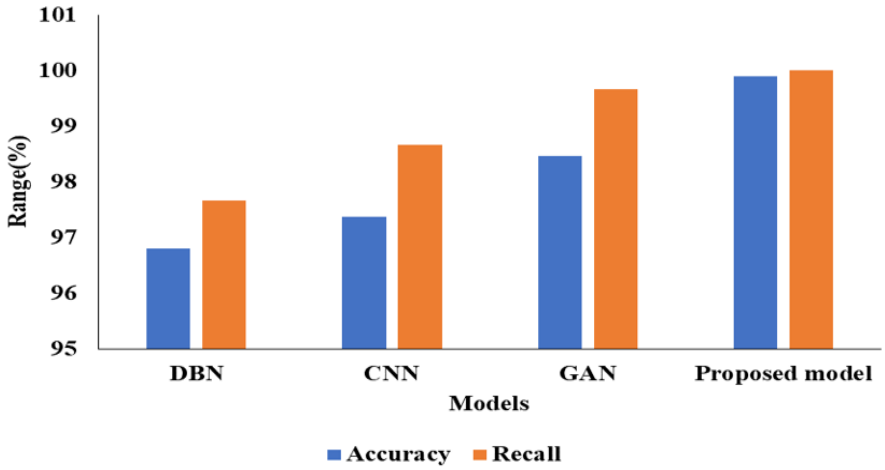

Figure 3. Analysis of accuracy & Recall.

In the analysis of Precision, DBN achieved 96.6%, CNN achieved 97.33%, GAN achieved 98.42%, and the proposed model achieved 99.76%. In the analysis of F1-Score, DBN achieved 97.66%, CNN achieved 97.93%, GAN achieved 98.53%, and the proposed model achieved 99.82%. Figure 4 and Table 3 present the graphical analysis and numerical analysis of the proposed model.

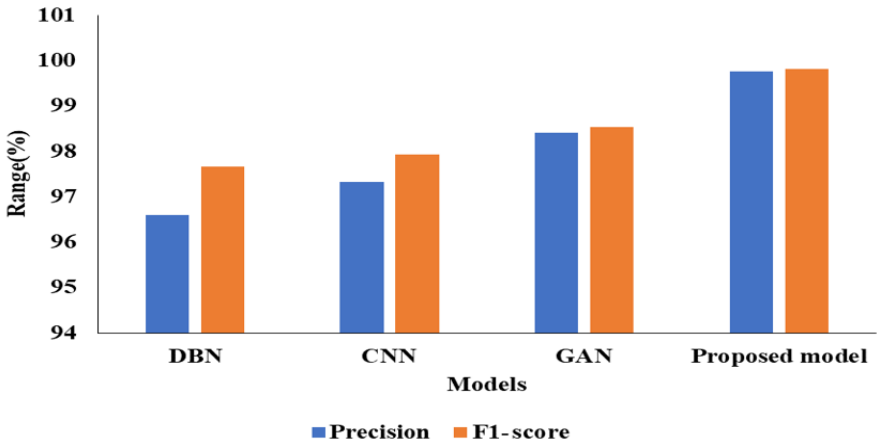

Figure 4. Analysis of precision & F1-score.

5. CONCLUSION

An abnormal growth of brain or nervous system cells is called a brain tumor. A tumor can develop from brain tissue. Another risk is brain tumours or tumours in the tissues surrounding the brain. The brain's surface membranes and the pituitary and pineal glands are nearby structures. The present study employs the adaptive weighted frost filter technique for the purpose of image pre-processing, with the specific objective of mitigating quantum noise. The utilization of pre-processed BT images yielded improved outcomes, which offer advantages for further procedures such as BT categorization and optimization, among other potential applications. In this study, separable convolution and control of Swin Transformer layers are used to find a balance between the rise in computational complexity and the improvement in brain tumour classification accuracy. This is a side effect of full self-attention learning in the transformer. TGA optimization is used to perform hyper parameter tuning in Swin Transformer. With the suggested model, the best classification accuracy ever was attained, at 99.89%. This paper's next work will combine patient photos across planes to expand the dataset size and provide insights about tumor types that are difficult to see from only one plane.

Funding: This study received no specific financial support.
Institutional Review Board Statement: The Ethical Committee of the Saveetha Institute of Medical and Technical Sciences, India has granted approval for this study (Ref. No. SIMATS/RSH/2023/741).
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: The conception and design part of this paper, K.S.; the data collection and preparing the manuscript, V.G.K.; the analysis and interpretation of results, S.S.K.; the simulation work, P.P.; formatting and revising the paper as per the editorial comments, B.P.L. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] A. H. Khan *et al.*, "Intelligent model for brain tumor identification using deep learning," *Applied Computational Intelligence and Soft Computing*, vol. 2022, pp. 1-10, 2022. https://doi.org/10.1155/2022/8104054

[2] C. L. Choudhury, C. Mahanty, R. Kumar, and B. K. Mishra, "Brain tumor detection and classification using convolutional neural network and deep neural network," presented at the 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE. https://doi.org/10.1109/ICCSEA49143.2020.9132874, 2020.

[3] N. Alturki *et al.*, "Combining CNN features with voting classifiers for optimizing performance of brain tumor classification," *Cancers*, vol. 15, no. 6, p. 1767, 2023. https://doi.org/10.3390/cancers15061767

[4] P. Sharma, M. Diwakar, and S. Choudhary, "Application of edge detection for brain tumor detection," *International Journal of Computer Applications*, vol. 58, no. 16, pp. 21-25, 2012. https://doi.org/10.5120/9366-3820

[5] M. Siar and M. Teshnehlab, "Brain tumor detection using deep neural network and machine learning algorithm," presented at the 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE). IEEE. https://doi.org/10.1109/ICCKE48569.2019.8964846, 2019.

[6] P. Sapra, R. Singh, and S. Khurana, "Brain tumor detection using neural network," *International Journal of Science and Modern Engineering* vol. 1, no. 9, pp. 2319-6386, 2013.

[7] M. A. Amou, K. Xia, S. Kamhi, and M. Mouhafid, "A novel MRI diagnosis method for brain tumor classification based on CNN and Bayesian optimization," *Healthcare*, vol. 10, no. 3, p. 494, 2022. https://doi.org/10.3390/healthcare10030494

[8] A. R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, and Z. Mehmood, "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification," *Microscopy Research and Technique*, vol. 84, no. 7, pp. 1389-1399, 2021. https://doi.org/10.1002/jemt.23694

[9] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, "A deep learning model based on concatenation approach for the diagnosis of brain tumor," *IEEE Access*, vol. 8, pp. 55135-55144, 2020. https://doi.org/10.1109/access.2020.2978629

[10] K. Muhammad, S. Khan, J. Del Ser, and V. H. C. De Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 507-522, 2020. https://doi.org/10.1109/tnnls.2020.2995800

[11] M. F. Alanazi *et al.*, "Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model," *Sensors*, vol. 22, no. 1, p. 372, 2022. https://doi.org/10.3390/s22010372

[12] M. Yaqub *et al.*, "State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images," *Brain Sciences*, vol. 10, no. 7, p. 427, 2020. https://doi.org/10.3390/brainsci10070427

[13] M. I. Mahmud, M. Mamun, and A. Abdelgawad, "A deep analysis of brain tumor detection from mr images using deep learning networks," *Algorithms*, vol. 16, no. 4, p. 176, 2023. https://doi.org/10.3390/a16040176

[14] M. U. Ali, S. J. Hussain, A. Zafar, M. R. Bhutta, and S. W. Lee, "WBM-DLNets: Wrapper-based metaheuristic deep learning networks feature optimization for enhancing brain tumor detection," *Bioengineering*, vol. 10, no. 4, p. 475, 2023. https://doi.org/10.3390/bioengineering10040475

[15] G. Tandel, A. Tiwari, O. Kakde, N. Gupta, L. Saba, and J. Suri, "Role of ensemble deep learning for brain tumor classification in multiple magnetic resonance imaging sequence data," *Diagnostics*, vol. 13, no. 3, p. 481, 2023. https://doi.org/10.3390/diagnostics13030481

[16] R. Asad, S. U. Rehman, A. Imran, J. Li, A. Almuhaimeed, and A. Alzahrani, "Computer-aided early melanoma brain-tumor detection using deep-learning approach," *Biomedicines*, vol. 11, no. 1, p. 184, 2023. https://doi.org/10.3390/biomedicines11010184

[17] V. Anand *et al.*, "Weighted average ensemble deep learning model for stratification of brain tumor in MRI images," *Diagnostics*, vol. 13, no. 7, p. 1320, 2023. https://doi.org/10.3390/diagnostics13071320

[18] O. Özkaraca *et al.*, "Multiple brain tumor classification with dense CNN architecture using brain MRI images," *Life*, vol. 13, no. 2, p. 349, 2023. https://doi.org/10.3390/life13020349

[19] J. Cheng, "Brain tumor dataset. Figshare. Dataset," 2017. https://doi.org/10.6084/m9.figshare.1512427.v5