

# Distributed Networks for Brain Tumor Classification Through Temporal Learning and Hybrid Attention Segmentation

Sayeedakhanum Pathan<sup>1\*</sup>, Savadam Balaji<sup>2</sup>

Research Scholar, Department of CSE, Koneru Lakshmaiah Education Foundation,  
Aziznagar, Hyderabad, Telangana, 500075, India<sup>1</sup>

Assistant Professor, Department of CSE – AIML & IoT, VNRVJIET, Hyderabad, Telangana, 500090, India<sup>1</sup>

Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, Aziznagar, Hyderabad, Telangana, 500075, India<sup>2</sup>

**Abstract**—Brain Tumor (BT), which is the progress of abnormal cells in brain surface is categorized into different types based on the symptoms and the affected parts in brain. Classification of BT using Magnetic Resonance Imaging (MRI) is an important and challenging task for BT diagnosis. Various approaches are designed to solve the issues and there are so many inconsistencies in detecting the tumor at early stage. The changes in variability and the complexity of size, shape, location and texture of lesions, automatic detection of BT still results a challenging task in the medical research community. Hence, a proposed Hybrid Attention Temporal Difference Learning with Distributed Convolutional Neural Network-Bidirectional Long Short-Term Memory (HATDL-DCNN-BiLSTM) is developed in this research to detect and classify the BT at beginning stage that enables to improve the survival rate of humans. The proposed model uses Gaussian filter for input image enhancement, Hybrid Attention-VNet segmentation to generate region of interest and solves the computational issues through the attention modules by minimizing the dimensions. The proposed model consumed less memory utilization and increase the training speed globally using the distributed learning mechanism. The features extracted using Hybrid Attention based Efficient Statistical Triangular ResNet (HA-ESTER) supports the classification model to increase the training efficiency more accurately. The proposed HATDL-DCNN-BiLSTM attains higher efficiency by the metrics of accuracy, recall, F1-score, and precision of 98.93%, 99.21%, 97.67%, and 96.17% with training data, and accuracy, recall, F1-score, and precision of 96.34%, 96.51%, 96.33%, and 96.15% with k-fold using BraTS 2019 dataset.

**Keywords**—Brain tumor; magnetic resonance imaging Gaussian filter; hybrid attention-VNet; distributed convolution neural network

## I. INTRODUCTION

Human brain is a vital organ in the physical body because it is responsible for the various governing processes of humans like feeling, memory, responses, vision, motor skills, and breathing [1] [2]. These regulatory functions are significantly disrupted when BT begin to form inside the brain, which arise due to the unfamiliar development of cells in certain brain tissues. BT is primarily categorized as benign or malignant, the benign tumor can't able to diffuse to diverse parts of the brain, so it is considered non-cancerous. However, the malignant tumor is cancerous because it grows uncontrollably and diffuses to various parts of the brain [3]. There are about 200 various

types of BT that can arise in diverse areas of the brain. These types of tumors cause more life-varying impact on affected individual's lives [4]. The symptoms of BT appear when the illness is in the stage of advanced and the early phase of BT doesn't reveal any symptoms to the affected person. This phenomenon is due to the position and small size of the tumor in the early phases [5] [6]. Remembrance issues, deviations in the power of eyesight, unfamiliar actions, misperception, seizures, and stability issues are the indications of BT and sometimes it varies depend on the type and location of tumors [7]. After the surgery, the survival rate of BT patients is 14% but, if it is detected in the early stage, the survival rate increases to 70% [6].

Various imaging procedures, like positron emission tomography (PET), MRI and computed tomography (CT), are employed to scan the complete structure of the brain [8]. Compared to PET and CT, MRI is regarded as a better imaging modality and it is broadly used to recognize and categorize the BT because of its better resolution. Moreover, the MRI is highly useful and important in the domain of radiology, because it offers various alterations between a variety of body's soft tissues [9]. Recently, Machine Learning (ML) algorithms have been combined with the automatic BT detection system in various studies to detect BT from the brain MRI [10]. Many efforts have been made to create very effective and trustworthy methods for automatically classifying BT. Handmade features are utilized in the traditional ML techniques, which limited the robustness of the solution and increase the cost. Nevertheless, occasionally supervised learning methods can exceed the unsupervised learning methods, thus resultant in an overfitted approach, that is not fit for additional large repository [11]. Additionally, conventional ML methods also rely on handmade features, that impose drawbacks on the durability and effectiveness of the solution [12]. Though there are various systems available for recognizing irregularities in brain MRI, yet there is a possibility for improving the performance and making the categorization within an appropriate amount of time. The effort in examining and identifying BT using conventional methods become more difficult owing to the expanding size of medical information [13].

Globally, there is no effective method has been found for segmenting and identifying the BT in recent studies irrespective

of its position, structure, and intensity [14] [15]. Many researchers utilized numerous traditional feature extraction approaches, such as Histogram of Oriented Gradients (HOG), bag of word (BoW), local binary patterns (LBP), gray level co-occurrence matrix (GLCM), and density histogram to extract the relevant features. Though, these methods unsourced to extract the exact features, that are needed for accurate BT detection [16] [17]. The approach in study [18] applies k-means clustering approach, whereas feature reduction and extraction processes depend on the Principal Component Analysis (PCA) approach and Discrete Wavelet Transform (DWT) approach respectively. Lastly, SVM is utilized for classifying the BT. However, these approaches consumed more time to complete the procedure of accurate BT detection as well as classification [18]. Some methods apply manually defined tumor regions for detection of BT, which forbids them from being entirely computerized. Deep Learning (DL) is gaining more popular, because of the facility to extract features automatically. Still, DL consumes lot of processing volume and memory [17]. The experimental results of various BT detection frameworks are yet in the initial phase since several characteristics influence the detection method, like poor localization of tumor, deficiency of training volume, poor quality such as the deficiency of training data, poor tumor localization, low-quality images and features [19]. However the combination of deep learning and transfer learning aims to enhance the accuracy and efficiency of brain tumor diagnosis by leveraging the power of pre-trained models to improve classification and segmentation tasks [20].

The main motivation of the research is to design and develop a model for BT classification using proposed HATDL-DCNN-BiLSTM. The Gaussian filter is subjected to MRI for generating the pre-processing image result and with the pre-processed result, segmentation is done by HA-based VNet model to enhance the detection rate, and furthermore, the important and the suitable features are being extracted by applying HA-ESTR model. The features extracted through individual feature extraction mechanism are concatenated to generate a feature map that helps to obtain better performance in image classification. Fine tuning the model provides better generalizability by minimizing the error value through categorical cross entropy measure. The research contribution is briefly discussed as follows:

1) *Hybrid Attention-VNet (HA-VNet) based segmentation:* The segmentation model is designed through the incorporation of attention models, like HA with the VNet model to extract the color, contrast, texture, and boundary details of image modality, as it helps to increase the detection rate. It highly focuses in extracting key features in both the split and double attention to increase the classification performance.

2) *Hybrid Attention-Efficient Statistical Triangular ResNet (HA-ESTR) based feature:* The feature extraction process performed using HA-ESTR is to extract useful features to minimize the classification error. This feature extraction model is designed through the integration of different feature extraction models that helps to derive the tumor related features including statistical information. Accordingly, this information is more helpful in finding cancerous and non-cancerous category.

3) *Proposed Hybrid Attention Temporal Difference Learning Distributed CNN-BiLSTM (HATDL-DCNN-BiLSTM):* The proposed framework is designed through the incorporation of attention models, and temporal difference learning with distributed deep learning framework. Due to the higher learning capability and facility of proposed model, it shows optimal performance in BT classification using imaging modality. The trained images are coordinated with test images to minimize the loss function such that minimal error value returns best performance.

The subsequent sections of this manuscript are delineated as follows: Section II deals with the traditional methodologies employed in the processes of feature extraction, segmentation, optimization, and classification of Brain Tumor. Section III offers an in-depth explanation of the proposed framework. Section IV and V addresses the results and discussion part respectively. Finally, Section VI encapsulates the entirety of the research, highlighting key findings, implications and future work.

## II. LITERATURE SURVEY

Few traditional approaches used for the classification of BT are reviewed in this section. Rahman, T. and Islam, M.S [1] designed a parallel deep convolutional Neural Network (PDCNN) mechanism for classification of BT. Here local as well as global features were extracted in a parallel way to solve the over fitting issues by applying dropout regularizer with batch normalization. The performance was evaluated with three different BT datasets, and reported efficient and accurate performance by extracting low-level as well as high-level features. However, it failed to use the model with 3D structure for identifying the tumors. Ullah, Net al. [17] designed a unified end-to-end model based on deep learning approach (TumorDetNet) for detecting and classifying BT. The distinctive features were effectively learned and minimized the over fitting issues. The SoftMax layer was used to detect the tumor and their grades. It showed higher accuracy measure, but failed to detect complex tumor types. Z. Atha and J. Chaki [3] designed a semi supervised deep learning model for detecting BT. This model was derived through the integration of unsupervised autoencoder mechanism with supervised classification model. It trains the learning parameter of descriptors for better classification. Accordingly, the instances were created through fuzzy logic using augmented data. This model showed higher accuracy, but failed to detect the tumor with location.

Anantharajan, S.et al. [2] modelled a deep learning framework to detect the tumor at early stage. Here, the images were captured and pre-processed by applying adaptive contract enhancement algorithm (ACEA) with median filter. The segmentation was done using fuzzy c-means model and the features were extracted effectively. It resulted higher accuracy rate, but faced time complexity issues. Saha, P. et al. [3] designed a deep learning network with ensemble model for detecting the tumor into different classes. The deep features were captured by CNN and these features are used to classify the BT types. It showed better performance, but failed to evaluate the analysis with large sized dataset. Mathivanan, S.K.

et al. [21] focused a deep and transfer learning approach for detection of BT accurately using MRI. The model was trained with benchmark datasets and increase the performance by applying image enhancement methods. It showed higher accuracy rate and classifies the tumor into different grades.

Amin, J. et al. [9] designed a random forest classifier to classify BT into three different regions, namely non-enhancing, enhancing and complete tumor. The cross-validation schemes were applied in the model to reduce over fitting issues. This model was feasible in generating segmented results without manual interactions. Asiri, A.A.et al. [4] introduced a dual module mechanism to improve the accuracy and speed of BT detection. The first module was said as image enhancement approach that used a machine learning strategy for normalizing the images and solve the issues, like low contrast and noise. Accordingly, second module applies support vector machine (SVM) to perform segmentation as well as classification of BT. It increased the robustness and generalizability of model.

### A. Challenges

The challenges discovered from the existing BT detection approaches are labelled below.

- In study [6], the suggested approach offered faster processing time and improved accuracy due to the incorporation of an efficient algorithm. However, the suggested method was inefficient in dealing with imbalanced and large-scale datasets.
- The suggested HBTC framework [22] reduced the complexity of inherent. However, it failed to generalize over different datasets and revealed low exactness and toughness.
- The suggested method [23] was not examined with openly available imaging datasets having irregularity features and on the other hand, the suggested method also had the issue of scalability.
- In study [12], the suggested method reduced feature irrelevance, duplication, and dimensionality to identify the important features, however, the feature selection and extraction process consumed more time because of the high computational complexity in managing huge amounts of data.
- The suggested method [24] effectively classifies the brain tumors, however, it failed to do the cropping and rotating process to the data before it was subjected to the classification process.

### B. Problem Statement

In the recent years, BT has grown uncontrollably in different locations across human body, but it mainly causes in brain. When the tumor grows, it starts to increase the pressure inside the brain, which typically affects the brain and cause brain damage and also it threatens human life. Various BT detection methods introduced in the medical research helps the physicians to detect the cancer using MRI, but it undergoes several challenges. The medical professionals and radiologists examine number of MRI slices that results labor intensive and time-consuming process. However, this manual scheme causes

human error and potential leads to delay or misdiagnosis. Hence, it is an urgent requirement in the medical research to design an accurate BT detection method for accurate diagnosis of BT at earlier stage. To accomplish this task, an input MRI is fetched and the image is captured from BraTS dataset. Assume the dataset as  $T$  with  $s$  number of MRI brain images as,

$$T = \{C_1, C_2, \dots, C_s, \dots, C_s\} \quad (1)$$

The input image is represented as  $C$ , which is subjected to the pre-processing stage to increase the quality of image by eliminating the noise. Accordingly, pre-processed result is represented as  $C$ , which is fetched by applying Gaussian filter to the input image. All the input images do have same dimension and so it is required to pre-process the image to provide uniformity in the training process. The image  $C$  is fed to the segmentation phase, which simplify and change the image representation into meaningful factor, such that segmented image result is specified as,  $H$

$$H = \{C\} \quad (2)$$

Segmented image result is fed to feature extraction phase to extract useful and essential features and based on the features, a feature vector is generated and it is noted as,  $A$ , respectively.

Accordingly, the generated feature map is applied to the model, where the tumor grades are classified into various stages, like no tumor, enhancing tumor, non-enhancing tumor, and peritumoral edema and this is accomplished through reduction of error value by applying categorical cross entropy loss function that is specified as,

$$L = -\sum M \log(M) \quad (3)$$

Here,  $M$  denotes predicted value, and  $M$  is the actual ground truth value.

$$M = t = \begin{cases} 0; & \text{if "C" is normal} \\ 1; & \text{else if "C" is non enhancing tumor} \\ 2; & \text{else if "C" is peritumoral edema} \\ 3; & \text{otherwise "C" is enhancing tumor} \end{cases} \quad (4)$$

Hence, the proposed BT detection model generate accurate results by minimizing the error rate through the loss function.

## III. METHODOLOGY

The HATDL-DCNN-BiLSTM is developed in this research for classification of BT with MRI modality. Input image is acquired from the BraTS dataset and subject to pre-processing phase, in which Gaussian filter is applied for enhancing image quality. The pre-processed image is subjected to segmentation phase, where HA based VNet segmentation framework is applied to generate region of interest. Based on the segmented image result, the features connected with the modality is extracted using HA-ESTR that includes different feature extraction models. After extracting the features, BT classification is performed using proposed HATDL-DCNN-BiLSTM. Fig. 1 represents the schematic view of proposed BT classification framework.

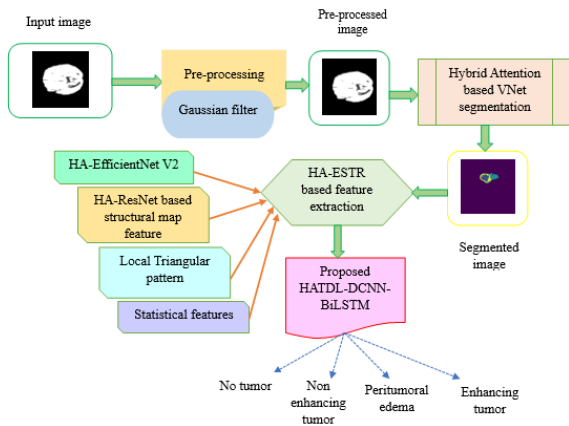


Fig. 1. Schematic representation of proposed HATDL with DCNN-BiLSTM for BT classification.

### A. Gaussian Filter-Based Image Pre-processing

MRI plays an active role in human brain scanning, as MRI provides detailed information about brain soft tissue structure. Due to the better efficiency in detecting the soft tissues of brain, MRI contributes more significant in BT classification. Consider the dataset as  $T$  composed with  $S$  number of MRI, and each image used for the processing is specified as,

$$T = \{C\}; k \in \{1, 2, \dots, s\} \quad (5)$$

Here,  $T$  is the dataset,  $C_k$  denotes  $k^{\text{th}}$  input image with the dimension of  $[224 \times 224]$ , and  $s$  implies total number of images. The brain images are sensitive towards unwanted noise and distortions and hence, pre-processing is needed to remove the noise for increasing greatest quality of image. Gaussian filter [25] is applied to pre-process the input image  $C$  for smoothen image by reducing the noise. Gaussian filter is an essential tool in computer vision and image processing system, and it is considered as an optimal filter to solve imaging issues. It is referred as the linear smoothing filter that selects the weights with respect to shape of Gaussian function. Gaussian filter is a category of low pass filter specifically designed to eliminate the noise subject to normal distribution, and hence it is commonly used in the image processing system. By applying Gaussian filter, the noises are suppressed and smoothed out that further increases the image quality. The Gaussian filter for the image  $C$  with pixel values is expressed as,

$$F(u, v) = \frac{1}{\sqrt{2\pi\sigma}} \exp(-(x + y) / 2\sigma) \quad (6)$$

$$C = F(C) \quad (7)$$

Here,  $\sigma$  denotes the variance of Gaussian filter,  $x$  and  $y$  are the filter kernel, and  $F$  denotes Gaussian filter. The resultant pre-processed image after removing the noise is expressed as,  $C$  with the dimension of  $[224 \times 224]$ , respectively.

### B. Hybrid Attention Based VNet Segmentation

Segmentation process is applied to the image processing task for changing and simplifying image representation into

meaningful form by selecting similar pixels. The pre-processed result  $C$  is applied to segmentation phase, where HA-VNet model is utilized to generate segmentation result. The HA-VNet is designed through the integration of VNet with HA, in which HA is modelled by integrating split attention and double attention model. The major advantage of using HA-VNet is that it allows seamless image segmentation with higher performance and accuracy. The benefit of using attention model in the segmentation is to minimize the spatial and channel dimension of VNet model. The split attention is the computational unit composed of split attention operations and feature map group. On the other hand, the double attention computes pooled features and capture the complex appearances and more efficient in correlating features with each specific location. Accordingly, these benefits are incorporated with the VNet model to provide accurate segmentation results that helps to increase the training process. The structure of VNet [26] [27] contains two different paths, namely compression path (left side) and decompression path (right side). The convolution with suitable padding is performed for exploiting the features from input and reduce the resolution with suitable stride at the end of every stage. The compression path is divided into different phases that operates at varying resolutions such that each stage contains one to three convolutional layers. The VNet takes  $C$  as input, and the input taken by each phase is processed by a residual function used in convolutional layers, and finally output is re through final convolution layer. The HA mechanism is added at the convolutional layer-2 to increase the training efficiency. The HA layer into the VNet takes previous convolution layer output as input of size  $[N \times 6 \times 6 \times 16]$  and resulted output is passed as input to the next convolution layer. Accordingly, the benefit of using VNet architecture is that it offers better convergence than other non-residual learning network, and also it offers  $[5 \times 5]$  size of convolutional layer at each stage. Accordingly, the convolution layer is represented as,

$$X = \sum \left\{ \sum \sum C . E(\alpha, \beta) \right\} + P \quad (8)$$

where,  $X$  denotes the output of convolution operation,  $P$  denotes bias factor,  $C$  is the input,  $E(\alpha, \beta)$  is the weight among the locations  $(\alpha, \beta)$ . The convolutional operations are used to double the feature maps due to the involvement of residual framework and number of increasing feature channels considered in compression path. The convolutional layer aims to minimize the memory required for training procedure and process the input at higher resolution. The increase in the segmentation rate is achieved by reducing the error rate of model, in which the error value is computed through dice loss function that is expressed as,

$$L = \frac{1}{R} \sum \left[ 1 - \frac{2 \sum M . \underline{M}}{\sum M + \sum \underline{M}} \right] \quad (9)$$

Here,  $R$  denotes number of classes,  $r$  specifies number of class samples,  $M$  implies predicted value, and  $\underline{M}$  shows actual

ground truth value. The activation function contains various functions, like rectified function, tanh, and so on.

$$X = f\left(\sum X.E + P\right) \quad (10)$$

Here,  $X$  is the input of activation function,  $E$  and  $P$  shows weight and bias vector. Finally, the rectified linear unit (ReLU) is expressed as,

$$H = \begin{cases} X & ; X \geq 0 \\ \frac{X}{X} & ; X \leq 0 \end{cases} \quad (11)$$

The segmented image result obtained through HA-VNet is represented as  $H$  with the size of  $[N \times 224 \times 224]$ .

- Architecture of Hybrid Attention

The HA [28] is designed through the integration of double attention and split attention mechanism that enable the model to increase the training speed by consuming less memory during training process. Both the attention models are operated

by taking the input with dimension of  $[N \times 6 \times 6 \times 16]$ . Fig. 2 shows the structure of Hybrid Attention model. For any input value, the HA compute the attention map separately through split and double attention and the dimension is adjusted using reshape layer by applying sigmoid function. Assume the input

taken by the HA is  $D$  with size  $[N \times 6 \times 6 \times 16]$  and it is divided into two parts, but the channel number of these parts must be equal. Accordingly, these two parts extracts the boundary information through split attention and double attention. Structure of split attention contains different layers, namely reshape, pooling, dense, and softmax. With split attention, the attention process is enabled across the feature maps. Moreover, the original feature dimension is reduced with reshape layer and the result generated with the respase layer have the size of  $[N \times 6 \times 6 \times 16]$ . On the other hand, the double attention model contains different layers, like convolution, reshape, softmax, and matrix multiplication. Here, the convolution layer performs the convolution operation, whereas reshape layer enable to

reduce the dimension of feature maps into  $[N \times 6 \times 6 \times 12]$ . The output of two different layers is fused by the matrix multiplication layer and finally reshape layer is used to generate the output. Finally, the output of split attention  $d$  and the output of double attention  $d_2$  are fused together and the resultant is specified as,

$$d = \{d + d\}; d \in [N \times 6 \times 6 \times 16] \quad (12)$$

Here, the output of HA mechanism is specified as,  $d$  having the dimension of  $[N \times 6 \times 6 \times 16]$ .

### C. Hybrid Attention Based Efficient Statistical Triangular ResNet Feature Extraction

In this work HA-ESTR feature extraction method is proposed which is formed through the integration of different feature extraction models, like Hybrid Attention based

EfficientNet V2 feature (HA based EfficientNet V2), Hybrid Attention based ResNet structural map (HA based ResNet structural map), Local Triangular pattern (LTrP), and statistical features. Accordingly, the Hybrid Attention based EfficientNet V2 is modeled through the integration of Hybrid attention with standard EfficientNet V2 feature, and also Hybrid Attention based ResNet structural map is formed through the integration of Hybrid attention and structural map with ResNet feature. The HA-ESTR feature extraction model takes the segmentation output ( $H$ ) as input and extract useful features that helps to boost the classification accuracy. The detailed explanation of each feature extraction models is clearly explained as follows.

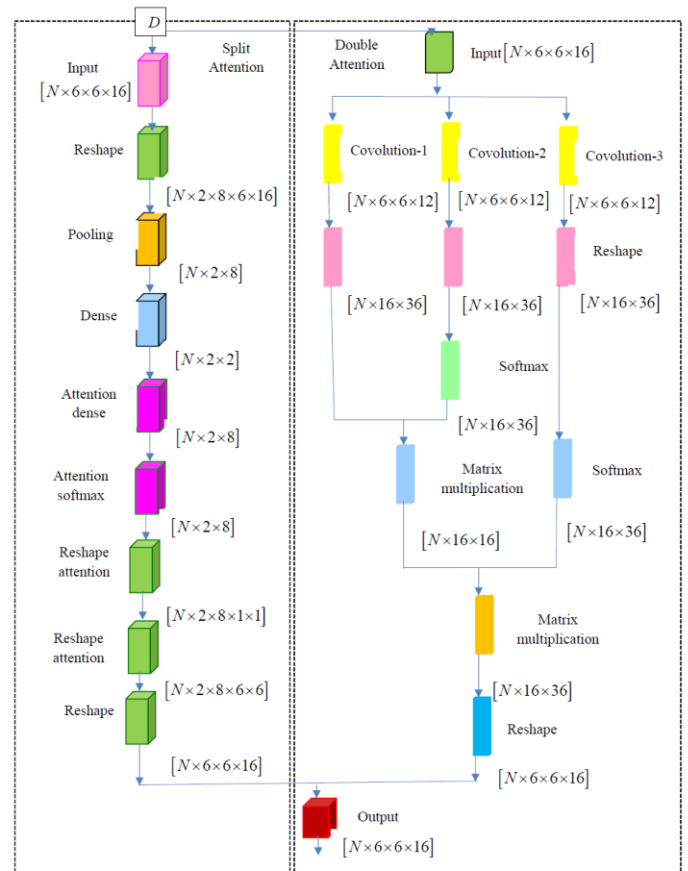


Fig. 2. Architecture of hybrid attention model.

1) Hybrid attention based EfficientNet V2 feature: In this feature is formed by incorporating the attention model features, like HA into the standard EfficientNetV2 feature. The benefit of using this feature model is that it effectively alters the regularization value with respect to image size due to the progressive learning mechanism. It effectively balances the network width, depth, and resolution and helps to increase the performance. The EfficientNetV2 model provides notable enhancement in the training efficiency and offers faster convergence [29] [30]. The EfficientNet V2 is composed with different layers, such as convolution, MBConv, HA, pooling, and fully connected layer. Initially, the input ( $H$ ) with dimension  $[N \times 224 \times 224]$  is passed to the convolution layer with stride  $[3 \times 3]$ , and the output is fed to MBConv layer,

which contains an inverted residual connection followed with the separable convolution. To reduce the overfitting issues, the dropout is used in MBConv block [30]. The HA layer is integrated into the model and the output of HA mechanism is passed to the convolution with stride  $[1 \times 1]$  layer and finally, the pooling and the fully connected layer is utilized to generate the output feature as  $A$  with the size of  $[1 \times 112 \times 112 \times 32]$  such that its dimension is resized into  $[1 \times 32 \times 32]$  for further processing.

2) *Hybrid attention based ResNet structural map feature:* The HA based ResNet structural map feature is designed by integrating the attention mechanism, such as HA into ResNet model with structural map features. The structure of HA model contains different layers, convolution, batch normalization, Maxpooling, flatten, fully connected layer, and LTP. The input  $H$  with size  $[N \times 224 \times 224]$  is passed to the convolution layer, where the HA layer is applied, and followed by flatten and fully connected layer are utilized to reduce the dimension of original features. Each layer contains same number of filters and if the size of feature map halved, then filters used in the model get doubles [31]. The LTP is applied to the texture analysis of the grey scale images and this feature is more superior in computational efficiency and description performance. LTP refers to a descriptor, which describes relationship among a selected pixels in the image with the neighborhood pixels [32]. This feature operates by setting a gray value of certain pixel as  $G$  and the gray value of pixels in the  $[3 \times 3]$  neighborhood as  $G_j (j = 0, 1, \dots, 7)$ . It analyzes the relationship between  $G$  and  $G_j$ , if  $G < G_j$ , the value is set to 0, otherwise the value is set to 1. By doing this process, an 8-bit numbers with 0's and 1's are generated such that the decimal value corresponds to the binary number is referred as the TLP feature value and it is expressed as,

$$A = \sum \chi(G - G_j)2^j ; \chi(a) = \begin{cases} 1 & ; a \geq 0 \\ 0 & ; \text{Otherwise} \end{cases} \quad (13)$$

Here,  $A$  denotes the structural map features. In general, the HA layer is integrated into the ResNet-50 model, and the output is fed to the LTP, which generates the structural map features as,  $A$  with the dimension of  $[1 \times 112 \times 112 \times 32]$  and for easier processing, the original dimension is resized into  $[1 \times 32 \times 32]$ , which is used for further processing. The purpose of extracting this feature is to enhance classification accuracy by modifying the residual block with Rectified Linear Unit (ReLU) activation function. Due to the great representation ability, this feature extraction model is applied in BT classification.

3) *Local triangular pattern feature:* The LTrP is designed using 8-bit binary code and the steps involved in the extraction of LTrP feature is briefly explained as follows. At first, a  $[3 \times 3]$  image matrix is generated for each image and then three neighbour pixels are selected with 600 triangle formation, and these triangles are formed in four different directions [33]. At each direction, the center pixel is referred as the threshold for the remaining neighboring pixels. A binary value of '1' is generated when the threshold value is higher than any of two neighboring pixels, and otherwise the binary value is set to '0'.

Thus, a binary value is created for all the three pixels using the below equation as,

$$\begin{aligned} Nw &= BY(\max[w - N], [w - N], [w - N]) \\ Nw &= BY(\max[w - N], [w - N], [w - N]) \\ Nw &= BY(\max[w - N], [w - N], [w - N]) \\ Nw &= BY(\max[w - N], [w - N], [w - N]) \end{aligned} \quad (14)$$

After finding the values at each direction, the triangle is flipped to 180 and repeat the same process to find the binary code value. Here, each triangle contains four values that includes  $w, N, w$ , and  $w$ . Accordingly,  $w$  is the center pixel, and hence set it as the threshold for  $w_6, N, w_2$ . If  $w$  is higher than any other two neighboring pixel values, that is  $w_6$  or  $N$  or  $w_2$ , then binary value is set to '1', otherwise set to '0'. The process involved in generating the above values are expressed as,

$$\begin{aligned} Nw &= BY(\max[w - w], [N - w], [w - w]) \\ Nw &= BY(\max[w - w], [N - w], [w - w]) \\ Nw &= BY(\max[w - w], [N - w], [w - w]) \\ Nw &= BY(\max[w - w], [N - w], [w - w]) \end{aligned} \quad (15)$$

Finally, the eight coded patterns are generated through the above process using the below equation as,

$$Nw_h = \sum_{h=0}^7 BY(w_h - N) \geq (t+1)/2 \quad (16)$$

$$A = LTP(\alpha, \beta) = \sum \gamma(Nw - N)2^j \quad (17)$$

where,

$$\gamma(X) = \begin{cases} 1, & X \geq 0 \\ 0, & \text{Otherwise} \end{cases} \quad (18)$$

Here,  $BY$  shows binary value,  $LTP$  represents the LTrP features, which is represented as  $A$  with the dimension of  $[1 \times 32 \times 32]$ .

4) *Statistical features:* Some of the statistical features extracted from the segmented image results are mean, standard deviation, variance, median, harmonic mean, geometric mean, and entropy [34] [35].

Mean: homogeneity of brightness of MRI, which is represented as

$$S = \frac{1}{Z} \sum H(z) \quad (19)$$

Here,  $Z$  is total number of pixels,  $S$  refers to mean feature with the dimension of  $[1 \times 1]$ , and  $H(z)$  is the segmented image.

Variance: heterogeneity that strongly connected with standard deviations and it is expressed as,

$$S = \sum \sum (m - S) H(m, n) \quad (20)$$

where,  $S$  refers to the variance feature with the size of  $[1 \times 1]$ .

Standard deviation: shows the difference exists in each observation from the mean value and it is represented as,

$$S = \sqrt{\frac{\sum (H(z) - S)^2}{Z}} \quad (21)$$

Here,  $S$  implies standard deviation feature with size of  $[1 \times 1]$ .

Median: It shows the central tendency of the features and it is generally represented as,  $S$  with the size of  $[1 \times 1]$ .

Harmonic mean: It extracts the average sample size from the number of groups and it is expressed as,

$$S = \frac{z}{\frac{1}{\lambda} + \dots + \frac{1}{\lambda}} \quad (22)$$

Here,  $z$  implies total number of pixels,  $\lambda$  and  $\lambda_z$  are the pixel values, and  $S$  is the harmonic mean features having the size of  $[1 \times 1]$ .

Geometric mean: refers to the appropriate measure, where the value changes exponentially. It is a category of mean with central tendency and it is computed as,

$$S = \sqrt[\lambda]{\lambda, \dots, \lambda} \quad (23)$$

Here,  $\lambda$  refers to the count of pixel values, and  $S$  is the geometric mean feature having the size of  $[1 \times 1]$ .

Entropy: It is used to calculate dissimilarity in MRI and its value will be very high for better performance. It is computed as,

$$S = \sum \sum H(m, n) \log H(m, n) \quad (24)$$

Accordingly, the entropy feature is denoted as,  $S$  with the size of  $[1 \times 1]$ . Finally, the statistical features generated from the segmented image results are expressed as,

$$A = \{S, \dots, S\} \quad (25)$$

where,  $A$  shows the statistical features with the size of  $[1 \times 7]$  and this dimension is resized into  $[1 \times 32 \times 32]$  for

smoothing the training process. The feature map generated using HA-ESTR are represented as,

$$A = \{A_1 \| A_2 \| A_3 \| A_4\} \quad (26)$$

Here,  $A$  is the feature map of HA-ESTR features with the size of  $[1 \times 32 \times 32 \times 4]$ , respectively.

#### D. HATDL-DCNN-BiLSTM Classification

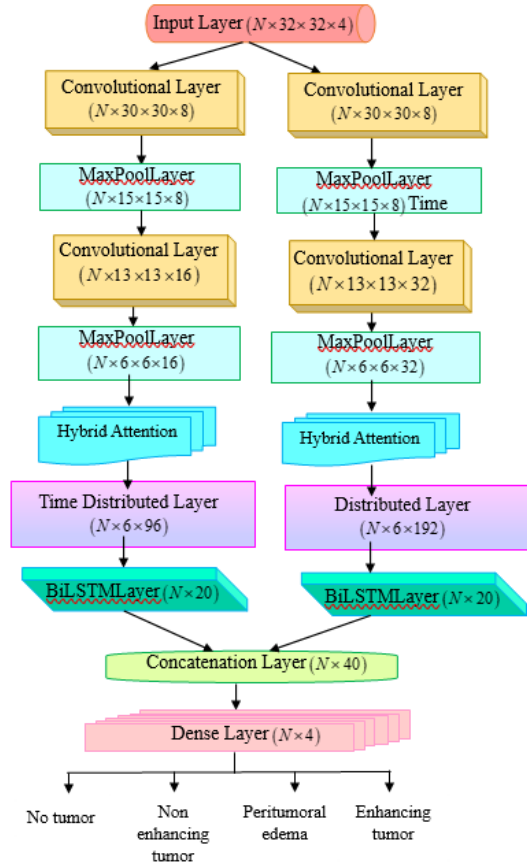


Fig. 3. Architecture of proposed HATDL-DCNN-BiLSTM.

The above Fig. 3 depicts the proposed framework, is modelled by applying the attention models, like split attention and double attention with Temporal Difference Learning (TDL) and CNN-BiLSTM using Distributed training [36] [37]. The benefit of applying this model for classification is due to its efficiency in higher training process. The integration of TDL helps to detect the output more accurately based on the working principle of unsupervised learning model. The distributed model provides completely effective solution with limited number of epochs by allocating the resources in a distributed manner. Hence, it reduces communication overhead and solve the communication delay caused due to the improper allocation of resources. The distributed learning framework partition the dataset and distribute it among the machines, and each data uses same weight factor but trains under various batches and finally, the results are averaged to get global gradient and it also enable to achieve faster training. The proposed HATDL with DCNN-BiLSTM model is made up of various layers, like convolution,

maxpooling, time distributed, Bidirectional, and dense layer [37]. The input layer gets the input by the dimension of  $[N \times 32 \times 32 \times 4]$  and fed to the convolution layer. The model considers two convolution layers that contains the filter size of 64 and 128, and followed by the convolution layer, maxpooling layers are utilized to minimize the dimension size. The aim of convolution and the pooling layer is to filter the incoming information for extracting important features, in which the convolution layer performs convolution operation between input features and the smaller matrices referred as filters or kernels.

Let us assume the input matrix as  $x$ , where  $x \rightarrow (A)$ ,  $y$  refers a kernel matrix, and  $O$  be the result matrix with rows and columns as  $p$  and  $q$ .

$$O[p,q] = x.y[p,q] = \sum \sum y[a,c].x[p-a,q-c] \quad (27)$$

Convolution layer uses ReLU activation function, which is commonly used activation function in the CNN model. The major benefit of using this activation function is that it does not require to activate each neuron at same time, as it converts all the negative value into '0', and due to this reason, ReLU offers more computational efficiency. The ReLU is expressed as,

$$F(J) = J = \max(0, J) \quad (28)$$

The pooling is the sub-sampling model aims to minimize the size of convolution matrix and enable to enhance the robustness of framework. After the maxpooling layer performs the operations, HA layer is applied and it generates the features with the dimension of  $[N \times 6 \times 6 \times 16]$ . Accordingly, the flatten or time distributed layer is utilized to normalize the features from the dimension  $[N \times 6 \times 6 \times 16]$  into  $[N \times 9 \times 96]$ , and also the BiLSTM layer reduce the feature size into  $[N \times 20]$ . Each unit in the BiLSTM network contains a memory cell as well as three gates, namely forget, input and output gate and they are used to maintain the flow of information. The input passed to the input gate is expressed as,

$$K = \sigma(I.[g, Y] + Q) \quad (29)$$

Here,  $\sigma$  denotes activation function,  $Y$  shows the current input that is obtained from the output of flatten layer,  $I$  shows weight matrix,  $Q$  represents bias vector, and  $g$  represents previous hidden state. The sigmoid output is represented as,

$$w = \sigma(I.[g, Y] + Q) \quad (30)$$

where,  $w$  indicates sigmoid output.

$$\mathbb{W} = \tanh(I.[g, Y] + Q) \quad (31)$$

Here,  $\mathbb{W}$  shows sigmoid output. The output of new cells state is specified as,

$$W = K.W + w.\mathbb{W} \quad (32)$$

where,  $W$  shows new cells state and the output obtained through the output gate is expressed as,

$$N = \sigma(I.[g, Y] + Q) \quad (33)$$

$$g = N.\tanh(W) \quad (34)$$

Here,  $\sigma$  shows sigmoid function,  $Y$  is the input at time  $c$ ,  $g$  indicates hidden state at time  $c$ ,  $\tanh$  refers hyperbolic tangent function,  $I$  and  $Q$  are the weight and bias vector. The output receives from two BiLSTM layers are fused together by concatenating the features with the size of  $[N \times 40]$ , which undergoes to dense layer that creates classification result having the dimension of  $[N \times 4]$  through sigmoid activation function. Accordingly, the accurate classification is accomplished by minimizing the error rate through loss function. Here, loss function considered to reduce the error value is accomplished using categorical cross entropy measure that is expressed in Eq. (3). Hence, the model generates the output of four different grades, as normal case, non-enhancing tumor, peritumoral edema, and indicates enhancing tumor.

#### IV. RESULTS

This section illustrates results of proposed HATDL-DCNN-BiLSTM by varying two different datasets using the performance metrics.

##### A. Experimental Setup

The proposed HATDL with DCNN-BiLSTM model is simulated in the PYTHON tool with windows 10 OS, intel core processor, and 16 GB RAM.

##### B. Dataset Description

This experimentation is done using BraTS 2018 dataset [38], and BraTS 2019 dataset [39]. BraTS 2018 dataset [38] contains multimodal scans available at four different modalities, namely T1, post contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 fluid Attenuated Inversion Recovery (T2-FLAIR). The BraTS 2019 dataset [39] contains multimodal scans and they are available at four different modalities as similar that of above dataset. Also, both the dataset contains four different labels, namely grade-0 as no tumor, grade-1 as non-enhancing tumor, grade-2 as peritumoral edema, and grade-3 as enhancing tumor.

##### C. Evaluation Metrics

The effectiveness of the proposed framework is measured with metrics, namely accuracy, F1-score, precision, and recall.

Accuracy: It shows the percentage of accurately predicted samples to the total number of samples predicted.

$$A = \frac{G + G}{G + G + H + H} \quad (35)$$

Here,  $A$  is the accuracy,  $G$  and  $G$  are the true positive and the true negative, whereas  $H$  and  $H$  are the false positive and the false negative.



Precision: It shows the number of true positives towards total number of true and false positives.

$$P = \frac{G}{G + H} \quad (36)$$

Recall: It refers to the percentage of total number of relevant instances found in the data sample.

$$R = \frac{G}{G + H} \quad (37)$$

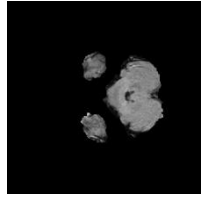
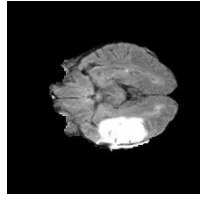

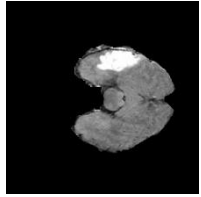
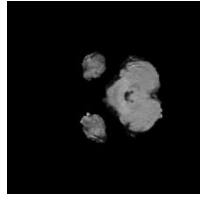
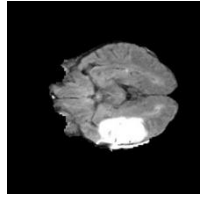

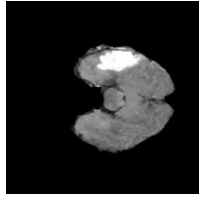
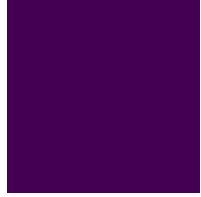


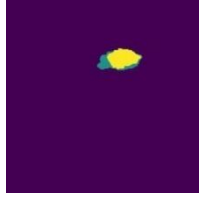
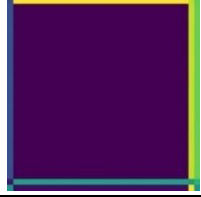

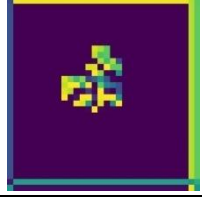

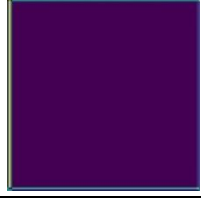

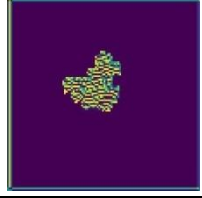

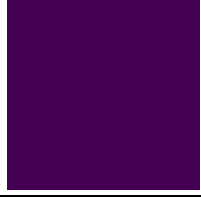
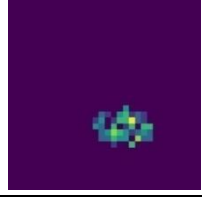
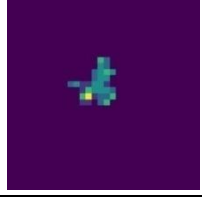
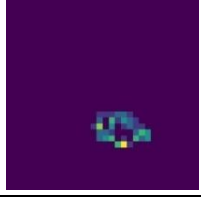
F1-score: It is the weighted average value of precision and recall measure.

$$F = 2 * \frac{P \times R}{P + R} \quad (38)$$

Here,  $P$  is the precision,  $R$  shows recall, and  $F$  is the F1-score.

#### D. Sample Image Results

This section presents the image results collected in BT classification with two different datasets. Fig. 4 shows the sample image results collected using BraTS 2018 dataset. The sample images of the proposed methods that undergoes each phase are captured and are briefly explained in the below figure. The imaging modality belongs to each class, like normal, enhancing tumor, non-enhancing tumor and peritumoral edema are captured and shown below.

Input image				
Pre-processed image				
Segmented image				
HA based EfficientNetV2 feature				
HA based ResNet structural map feature				
LTP feature				

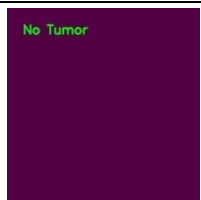
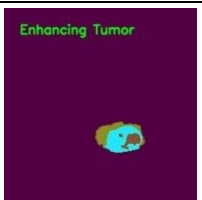


<b>Classification</b>				
-----------------------	-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------

Fig. 4. Sample image results observed by BraTs 2018 dataset.


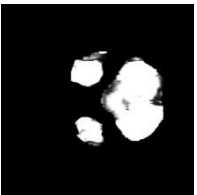



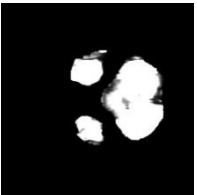


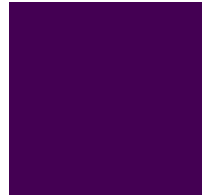




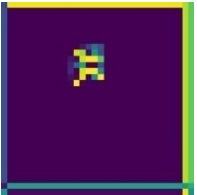


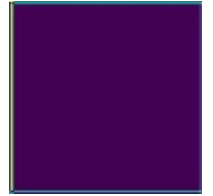

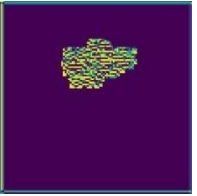
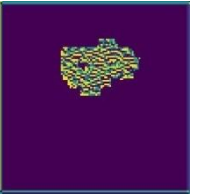

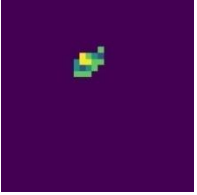


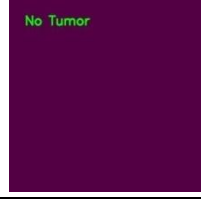

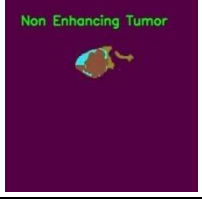

<b>Input image</b>				
<b>Pre-processed image</b>				
<b>Segmented image</b>				
<b>HA based EfficientNetV2 feature</b>				
<b>HA based ResNet structural map feature</b>				
<b>LTP feature</b>				
<b>Classification</b>				

Fig. 5. Sample image results obtained by BraTs 2019 dataset.

The Fig. 5 depicts the sample image results captured by BraTS 2019 dataset. For each grade of BT classification, the images are collected using proposed HATDL with DCNN-BiLSTM, and the collected samples images are clearly represented in the below figure with each classification stage.

E. Performance Analysis

This section explains the performance analysis of HATDL-DCNN-BiLSTM by varying the training percentage of k-fold with two different datasets.

1) *Analysis based on BraTS 2018 dataset:* Fig. 6 depicts the analysis done with BraTS 2018 dataset with training data. Fig. 6 (a) shows accuracy measure. With 80% of training data, the accuracy observed by the HATDL-DCNN-BiLSTM by varying the epoch from 20, 40, 60, 80, and 100 is 87.12%, 88.72%, 91.36%, 94.75%, 97.55%. The analysis made by F1-score is depicted in Fig. 6 (b). For 80% training data, the F1-score computed by HATDL-DCNN-BiLSTM for epoch 20 is 87.11%, epoch 40 is 88.71%, epoch 60 is 91.35%, epoch 80 is 94.73%, and epoch 100 is 97.54%. The performance analysis observed using precision is illustrated in Fig. 6 (c). The precision measured by HATDL-DCNN-BiLSTM at 80% training data for epoch 20 is 87.01%, epoch 40 is 89.11%, epoch 60 is 91.21%, epoch 80 is 95.77%, and epoch 100 is 96.70%. Fig. 6 (d) depicts analysis of recall measure. The recall observed by the proposed HATDL-DCNN-BiLSTM for 80% training data by varying epoch from 20 to 100 is 87.21%, 88.32%, 91.49%, 93.72%, and 98.38%, respectively. The epoch 20 refers that 20 times, the model is trained with 20 iterations, and epoch 40 means that 40 times the model is trained with 40 iterations, and is similar for remaining epochs. Hence, increasing the value of epoch, the performance of HATDL-DCNN-BiLSTM increases effectively due to the increasing volume of trained data. If training data is 60% used, the remaining data are used for testing process and so the performance degrades. The training data of 80% refers that 80% data are allowed in the training procedure and only remaining 20% is considered for testing, and this could allow the model to improve the performance. Hence, increasing the training percentage increases the model performance.

Fig. 7 depicts the analysis done with BraTS 2018 dataset by varying k-fold. Fig. 7 (a) illustrates accuracy measure. Accuracy obtained at k-fold of 10 by the proposed HATDL-DCNN-BiLSTM with epoch 20 is 88.96%, epoch 40 is 91.18%, epoch 60 is 94.18%, epoch 80 is 95.74%, and epoch 100 is 96.62%. The analysis described by F1-score is shown in Fig. 7 (b). The F1-score computed by HATDL-DCNN-BiLSTM at 10 k-fold value by varying the epoch from 20 to 100 is 88.95%, 91.62%, 93.92%, 95.72%, and 96.75%. The performance analysis made using precision is illustrated in Fig. 7 (c). The precision obtained by HATDL-DCNN-BiLSTM with k-fold 10 for epoch 20, 40, 60, 80, and 100 is 88.81%, 90.47%, 93.47%, 95.47%, and 95.91%. Fig. 7 (d) illustrates analysis of recall measure. The recall observed by HATDL-DCNN-BiLSTM with k-fold of 10 for epoch 20 is 89.09%, 40 is 92.79%, 60 is 94.36%, 80 is 95.98%, and 100 is 97.61%, respectively. For K-fold of 8, the entire dataset is partitioned into eight sets and the

model is trained for eight times to analyze the performance. The dataset is divided into different sets with respect to k-fold and model is trained based on the value of k-fold such that it simulates the performance to increase the efficiency. As, increasing the k-fold systematically increases the number of times that the model could be trained, which boost the performance of proposed HATDL-DCNN-BiLSTM.

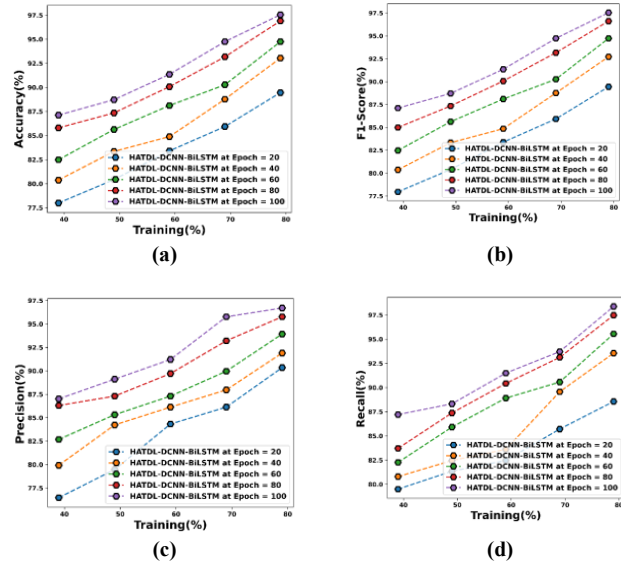


Fig. 6. Performance analysis-BraTS 2018 dataset by varying training percentage, a) accuracy, b) F1-score, c) precision, d) recall.

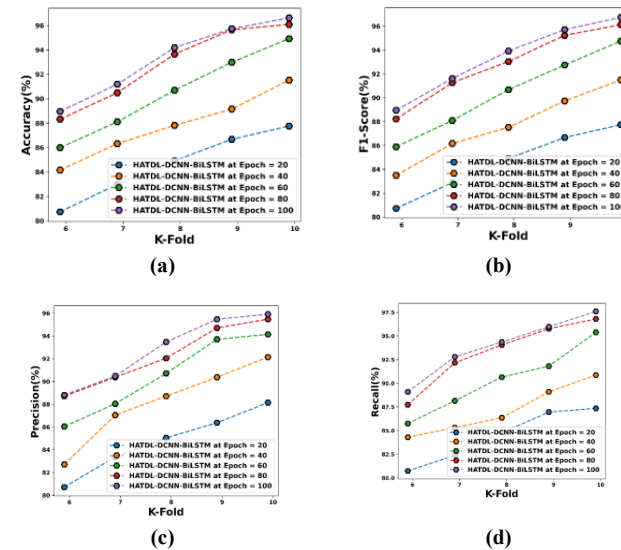


Fig. 7. Performance analysis with BraTS 2018 dataset by varying k-fold, a) accuracy, b) F1-score, c) precision, d) recall.

2) *Analysis based on BraTS 2019 dataset:* Fig. 8 illustrates the performance analysis carried out with BraTS 2019 dataset varying by training data. Fig. 8 (a) illustrates accuracy measure. Accuracy obtained by HATDL-DCNN-BiLSTM for 80% training data by varying the epoch 20 to 100 is 86.38%, 89.44%, 91.88%, 94.14%, and 98.93%. Fig. 8 (b) depicts F1-score

metric. The F1-score obtained by proposed HATDL-DCNN-BiLSTM with 80% training data for epoch 20 is 86.36%, epoch 40 is 89.43%, epoch 60 is 91.87%, epoch 80 is 94.13%, and epoch 100 is 97.67%. The precision used to analyze the performance is illustrated in Fig. 8 (c). With 80% training data, the precision of HATDL-DCNN-BiLSTM by varying the epoch from 20 to 100 is 85.51%, 89.04%, 91.89%, 93.53%, and 96.17%. Fig. 8 (d) illustrates the recall measure. The recall of proposed HATDL-DCNN-BiLSTM at 80% training data for epoch 20, 40, 60, 80, and 100 is 87.24%, 89.82%, 91.84%, 94.74%, and 99.21%.

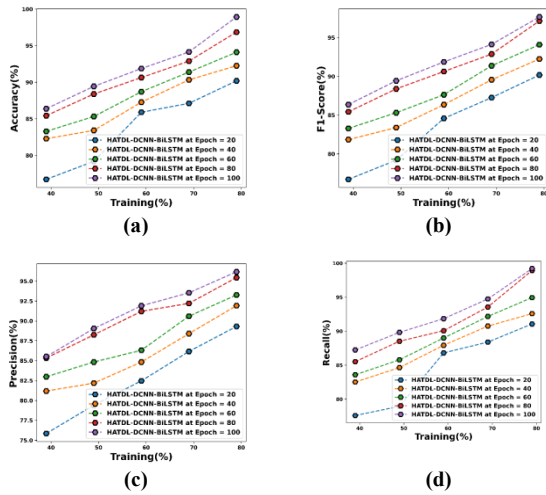


Fig. 8. Performance analysis with BraTS 2019 dataset by varying training data, a) accuracy, b) F1-score, c) precision, d) recall.

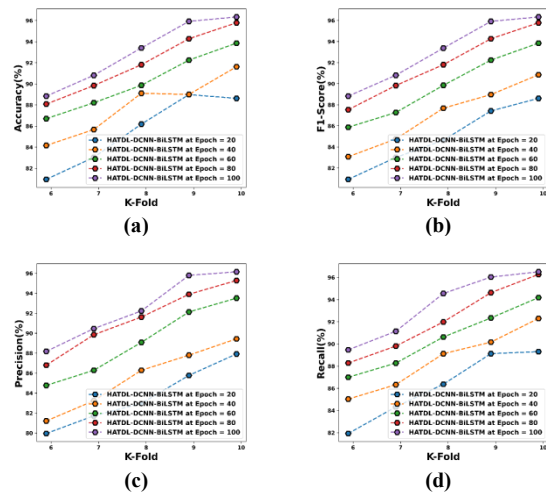


Fig. 9. Performance analysis with BraTS 2019 dataset by varying k-fold, a) accuracy, b) F1-score, c) precision, d) recall.

Fig. 9 shows the performance analysis done with BraTS 2019 dataset with k-fold. Fig. 9 (a) depicts accuracy measure. At k-fold 10, the accuracy observed by HATDL-DCNN-BiLSTM for epoch 20 is 88.84%, 90.81%, 93.40%, 95.93%, and 96.34%. Analysis done by varying F1-score is depicted in Fig. 9 (b). For k-fold 10, the F1-score measured by HATDL-DCNN-BiLSTM by varying the epoch from 20 to 100 is

88.82%, 90.80%, 93.38%, 95.92%, and 96.33%. The precision used to analyze the performance is shown in Fig. 9 (c). The precision obtained by proposed HATDL-DCNN-BiLSTM for k-fold 10 with epoch 20 is 88.18%, epoch 40 is 90.46%, epoch 60 is 92.23%, epoch 80 is 95.80%, and epoch 100 is 96.15%. Fig. 9 (d) illustrates the recall measure. With 10 k-fold value, the recall computed by HATDL-DCNN-BiLSTM by varying the epoch from 20 to 100 is 89.47%, 91.15%, 94.55%, 96.03%, and 96.51%.

### F. Comparative Methods

To evaluate the performance, the proposed HATDL-DCNN-BiLSTM is compared with the existing techniques, like parallel deep convolutional neural network (PDCNN) [11], unified deep learning model (TumorDetNet) [17], Semi-Supervised Brain Tumor Classification Network (SSBTCNet) [10], Ensemble Deep Neural Support Vector Machine (EDN-SVM), and DCNN-BiLSTM [36] [37].

## V. DISCUSSION

This section elaborates the comparative discussion of proposed HATDL-DCNN-BiLSTM model with two different datasets.

1) *Analysis with BraTS 2018 dataset:* Fig. 10 illustrates the comparative analysis of proposed HATDL-DCNN-BiLSTM model with existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods using BraTS 2018 dataset in terms of training percentage. For training percentage of 80%, the proposed HATDL-DCNN-BiLSTM model obtained the accuracy of 97.55%, which is superior than existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods by 12.80%, 18.59%, 8.38%, 16.06% and 6.45% respectively. The F1-Score of the proposed HATDL-DCNN-BiLSTM model achieves the result of 12.73%, 18.61%, 9.34%, 16.05%, and 6.44% higher than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM approaches for 80% of training percentage. The precision of the HATDL-DCNN-BiLSTM model attains the result of 96.70%, which is improved over the existing PDCNN, by 11.41%, EDN-SVM by 16.33%, TumorDetNet by 5.74%, SSBTCNet by 14.62%, and DCNN-BiLSTM by 4.70% for 80% training data. Finally, the recall of the HATDL-DCNN-BiLSTM model is 98.38%, which is improved to 14.05% with PDCNN, 20.80% with EDN-SVM, 12.74% with TumorDetNet, 17.46% with SSBTCNet, and 8.16% with DCNN-BiLSTM with 80% training data.

Fig. 11 show the comparative results of the proposed HATDL-DCNN-BiLSTM model with existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods using BraTS 2018 dataset in terms of k-fold. The HATDL-DCNN-BiLSTM model achieves the accuracy of 96.62%, which shows that the accuracy is superior over the existing PDCNN by 5.46%, EDN-SVM by 12.71%, TumorDetNet by 3.75%, SSBTCNet by 7.01%, and DCNN-BiLSTM by 2.09% for k-fold of 10 respectively. The HATDL-DCNN-BiLSTM model achieves the F1-score of 96.75%, which shows that the F1-score of the HATDL-DCNN-BiLSTM model is 6.07%, 12.84%, 3.90%, 7.32%, and 2.24% advanced

than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM approaches respectively. The precision of the HATDL-DCNN-BiLSTM model for k-fold of 10 is 95.91%, which increased 4.62% with PDCNN, 11.91% with EDN-SVM, 2.53% with TumorDetNet, 5.66% with SSBTCNet, and 1.14% with DCNN-BiLSTM. Moreover, for k-fold of 10, recall of the HATDL-DCNN-BiLSTM model is 97.61%, which is higher than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods by 7.50%, 13.77%, 5.26%, 8.95%, and 3.34% respectively.

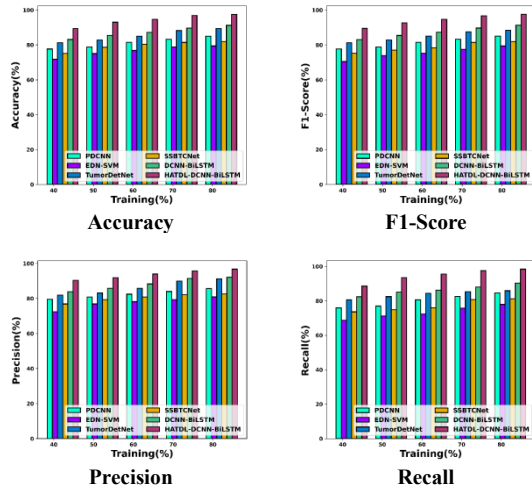


Fig. 10. Comparative analysis of HATDL-DCNN-BiLSTM with training percentage using BraTS 2018 dataset.

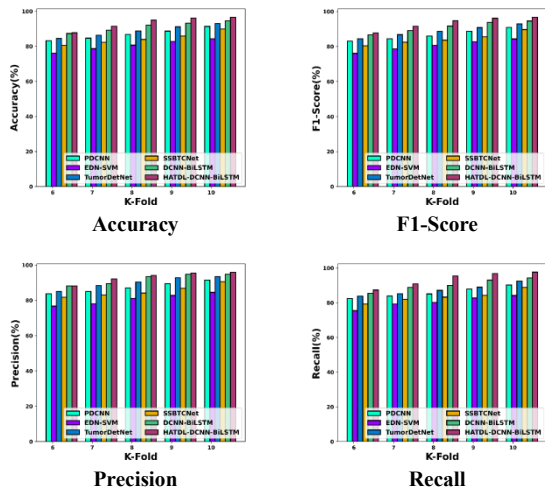


Fig. 11. Comparative analysis of HATDL-DCNN-BiLSTM with k-fold using BraTS 2018 dataset.

2) *Analysis with BraTS2019 dataset:* Comparative analysis of proposed HATDL-DCNN-BiLSTM model with existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods using BraTS 2019 dataset in terms of training percentage is shown in Fig. 12. The accuracy of the proposed HATDL-DCNN-BiLSTM model for 80% of training percentage is 98.93 which is compared with existing schemes, that reveals the improvement of 12.63% with PDCNN, 17.78%

with EDN-SVM, 11.31% with TumorDetNet, 13.82% with SSBTCNet, and 6.39% with DCNN-BiLSTM respectively. The F1-Score of the HATDL-DCNN-BiLSTM model at 80% training percentage is 97.67%, which is higher than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM techniques by 11.52%, 16.75%, 10.18%, 12.76%, and 6.25% respectively. The HATDL-DCNN-BiLSTM model attains the precision of 96.17%, which depicts that it is higher than the conventional PDCNN by 9.48%, EDN-SVM by 14.19%, TumorDetNet by 8.44%, SSBTCNet by 10.52%, and DCNN-BiLSTM by 3.53% for training percentage of 80%. Finally, the recall of the proposed HATDL-DCNN-BiLSTM model is 13.53%, 19.24%, 11.91%, 14.96% and 8.91% superior than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM approaches respectively.

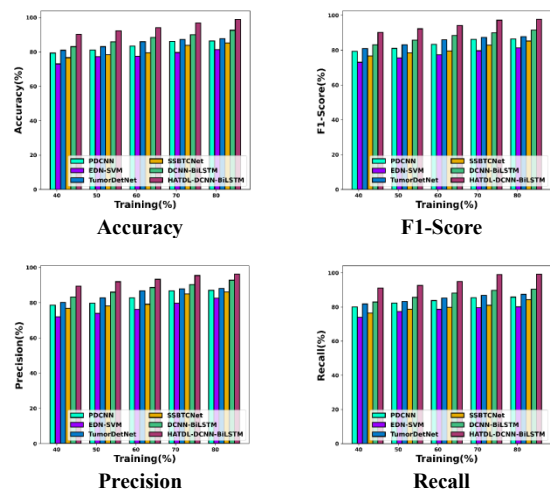


Fig. 12. Comparative analysis of HATDL-DCNN-BiLSTM with training percentage using BraTS 2019 dataset.

Fig. 13 illustrates the comparative outcomes of proposed HATDL-DCNN-BiLSTM model with existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods using BraTS 2019 dataset in terms of k-fold. For k-fold 10, accuracy of HATDL-DCNN-BiLSTM model is 96.34%, which is higher than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM methods by 7.31%, 14.71%, 3.93%, 10.20%, and 1.69% respectively. The HATDL-DCNN-BiLSTM model attains the F1-Score of 96.33%, which shows it is superior than the existing PDCNN by 7.31%, EDN-SVM by 14.72%, TumorDetNet by 3.93%, SSBTCNet by 10.21%, and DCNN-BiLSTM by 1.69% for k-fold 10 respectively. The precision of proposed HATDL-DCNN-BiLSTM is 7.27%, 14.15%, 3.56%, 10.71% and 1.19% advanced than the existing PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, and DCNN-BiLSTM approaches respectively. The recall of the HATDL-DCNN-BiLSTM model for k-fold of 10 is 96.51%, which increased 7.36% with PDCNN, 15.28% with EDN-SVM, 4.30% with TumorDetNet, 9.70% with SSBTCNet, and 2.19% with DCNN-BiLSTM respectively.

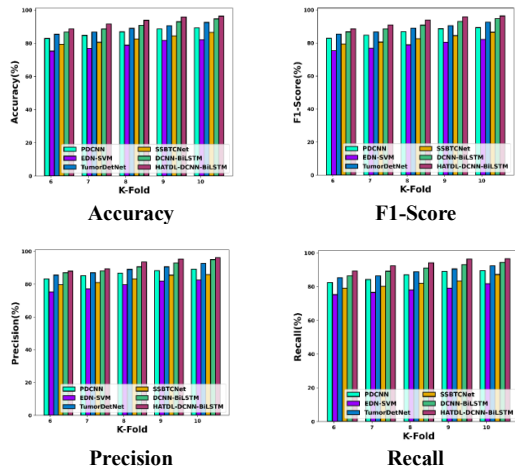


Fig. 13. Comparative analysis of HATDL-DCNN-BiLSTM with k-fold using BraTS 2019 dataset.

3) *Comparative discussion:* Table I illustrates the comparative discussion of HATDL-DCNN-BiLSTM with other existing methods using BraTS 2018 dataset. The existing methods evaluated for BT classification obtains inaccurate results with less accuracy. The PDCNN method [11] failed to use the model with 3D structure for identifying the tumors. Also, the TumorDetNet designed in [17] higher accuracy measure, but failed to identify the tumor grades. The SSBTCNet for BT classification failed to detect the tumor types exactly [10]. The EDN-SVM developed for the tumor classification provide higher performance but results

computational complexity issues. The above stated issues are resolved through the proposed HATDL-DCNN-BiLSTM using the distributed learning mechanism of deep neural network model. The model effectively reduces the overfitting issues and offers higher performance in terms of the evaluation measures. The proposed HATDL-DCNN-BiLSTM framework minimizes the computational issues globally through the extraction of optimal features. The features extracted through different models have varying dimensions and these varying dimensional features are resized to have same dimension for all the features such that this process minimizes the computational complexity issues effectively. The proposed HATDL-DCNN-BiLSTM value is measured by altering the training percentage and k-fold value. With training percentage, the proposed obtained higher value of 97.55%, 97.54%, 96.70%, and 98.38% for accuracy, F1-score, precision, and recall. The proposed HATDL-DCNN-BiLSTM showed 96.62%, 96.75%, 95.91%, and 97.61% for accuracy, F1-score, precision, and recall and these values are measured by BraTS 2018 dataset.

Table II illustrates the comparative discussion of HATDL-DCNN-BiLSTM with other existing methods using BraTS 2019 dataset. With this dataset, the proposed HATDL-DCNN-BiLSTM obtained the accuracy, F1-score, precision, and recall as 98.93%, 97.67%, 96.17%, and 99.21 by changing the training percentage. On the other hand, proposed HATDL-DCNN-BiLSTM obtained accuracy, F1-score, precision, and recall as 96.34%, 96.33%, 96.15%, and 96.51% with k-fold value.

TABLE I. COMPARATIVE DISCUSSION OF HATDL-DCNN-BiLSTM WITH BRATS 2018 DATASET

Methods		PDCNN	EDN-SVM	TumorDetNet	SSBTCNet	DCNN-BiLSTM	Proposed HATDL-DCNN-BiLSTM
TP=80%	Accuracy (%)	85.07	79.42	89.38	81.89	91.26	<b>97.55</b>
	F1-Score (%)	85.11	79.38	88.42	81.87	91.24	<b>97.54</b>
	Precision (%)	85.67	80.91	91.15	82.56	92.15	<b>96.70</b>
	Recall (%)	84.56	77.92	85.85	81.20	90.36	<b>98.38</b>
K-fold=10	Accuracy (%)	91.34	84.33	92.99	89.84	94.59	<b>96.62</b>
	F1-Score (%)	90.88	84.32	92.98	89.67	94.58	<b>96.75</b>
	Precision (%)	91.48	84.48	93.48	90.48	94.81	<b>95.91</b>
	Recall (%)	90.29	84.17	92.48	88.87	94.35	<b>97.61</b>

TABLE II. COMPARATIVE DISCUSSION OF HATDL-DCNN-BiLSTM USING BRATS 2019 DATASET

Methods		PDCNN	EDN-SVM	TumorDetNet	SSBTCNet	DCNN-BiLSTM	Proposed HATDL-DCNN-BiLSTM
TP=80%	Accuracy (%)	86.43	81.33	87.73	85.25	92.60	<b>98.93</b>
	F1-Score (%)	86.42	81.31	87.72	85.21	91.56	<b>97.67</b>
	Precision (%)	87.05	82.52	88.05	86.05	92.78	<b>96.17</b>
	Recall (%)	85.79	80.13	87.40	84.38	90.38	<b>99.21</b>
K-fold=10	Accuracy (%)	89.29	82.16	92.55	86.51	94.71	<b>96.34</b>
	F1-Score (%)	89.28	82.15	92.54	86.50	94.70	<b>96.33</b>
	Precision (%)	89.16	82.55	92.73	85.86	95.01	<b>96.15</b>
	Recall (%)	89.41	81.76	92.35	87.15	94.39	<b>96.51</b>

## VI. CONCLUSION

This research designs a proposed framework named HATDL-DCNN-BiLSTM for detection and classification of BT with MRI. This method uniquely addresses the issues in BT classification through the extraction of important and essential features by applying the attention mechanisms. It shows considerable advancements in efficiency and contrast for differentiating the tumor grades. The extracted features enable the model to increase the training speed that in turn enable to generate more reliable classification results. The transfer learning model introduced in the mechanism provides more beneficial activities, like faster training, less data requirements, higher learning rate, less training time and enhanced generalization. Due to the pre-trained models located in the transfer learning, the tasks to be executed are quickly learned and prevents the overfitting issues. It showed outstanding performance using the metrics, like accuracy, recall, F1-score, and precision of 98.93%, 99.21%, 97.67%, and 96.17% with training data. Also, the proposed scheme measured higher performance in terms of accuracy, recall, F1-score, and precision of 96.34%, 96.51%, 96.33%, and 96.15% using k-fold by considering the BraTS 2019 dataset. However, the proposed method has some limitations. Model performance is dependent on data quality and size, so further improvements can be brought about by using larger diverse datasets for better generalization.

The future direction of research would be the consideration of a hybrid optimization algorithm to improve efficiency in the network so as to give better classification results during BT classification. For example, combining genetic algorithm and particle swarm optimization can enhance the feature selection and the optimization of parameters for achieving an even more accurate and stable classification result.

## REFERENCES

- [1] E. Schulz, and S.J. Gershman, "The algorithmic architecture of exploration in the human brain", *Current opinion in neurobiology*, vol.55, pp.7-14, 2019.
- [2] S. Anantharajan, S. Gunasekaran, T. Subramanian, and R. Venkatesh, "MRI brain tumor detection using deep learning and machine learning approaches", *Measurement: Sensors*, vol.31, pp.101026, 2024.
- [3] P. Saha, R. Das, and S.K. Das, "BCM-VEMT: Classification of brain cancer from MRI images using deep learning and ensemble of machine learning techniques", *Multimedia Tools and Applications*, vol.82, no.28, pp.44479-44506, 2023.
- [4] A.A. Asiri, T.A. Soomro, A.A. Shah, G. Pogrebna, M. Irfan, and S. Alqahtani, "Optimized Brain Tumor Detection: A Dual-Module Approach for MRI Image Enhancement and Tumor Classification", *IEEE Access*, vol.12, pp.42868-42887, 2024.
- [5] A.N. Mavrakis, E.F. Halpern, F.G. Barker, R.G. Gonzalez, and J.W. Henson, "Diagnostic evaluation of patients with a brain mass as the presenting manifestation of cancer" *Neurology*, vol.65, no.6, pp.908-911, 2005.
- [6] S.M. Malakouti, M.B. Menhaj, and A.A. Suratgar, "Machine learning and transfer learning techniques for accurate brain tumor classification", *Clinical eHealth*, vol.7, pp.106-119, 2024.
- [7] S. Deepa, J. Janet, S. Sumathi, and J.P. Ananth, "Hybrid optimization algorithm enabled deep learning approach brain tumor segmentation and classification using MRI", *Journal of Digital Imaging*, vol.36, no.3, pp.847-868, 2023.
- [8] M. Yasmin, S. Mohsin, M. Sharif, M. Raza, and S. Masood, "Brain image analysis: a survey", *World Applied Sciences Journal*, vol.19, no.10, pp.1484-1494, 2012.
- [9] J. Amin, M. Sharif, M. Raza, and M. Yasmin, "Detection of brain tumor based on features fusion and machine learning", *Journal of Ambient Intelligence and Humanized Computing*, pp.1-17, 2024.
- [10] Z. Atha and J. Chaki, "SSBTCNet: Semi-Supervised Brain Tumor Classification Network", *IEEE Access*, vol. 11, pp. 141485-141499, 2023.
- [11] 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.
- [12] M. Wageh, K. Amin, A. D. Algarni, A. M. Hamad and M. Ibrahim, "Brain Tumor Detection Based on Deep Features Concatenation and Machine Learning Classifiers With Genetic Selection," *IEEE Access*, vol. 12, pp. 114923-114939, 2024.
- [13] M.A. Talukder, M.M. Islam, M.A. Uddin, A. Akhter, M.A.J. Pramanik, S. Aryal, M.A.A. Almoyad, Hasan, and K.F. Moni, "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning", *Expert systems with applications*, vol. 230, pp.120534, 2023.
- [14] R.L. Siegel, K.D. Miller, and A. Jemal, "Cancer statistics, 2019", *CA: a cancer journal for clinicians*, vol.69, no.1, pp.7-34, 2019.
- [15] N. Abiwinanda, M. Hanif, S.T. Hesaputra, A. Handayani, and T.R. Mengko, "Brain tumor classification using convolutional neural network", *In World Congress on Medical Physics and Biomedical Engineering 2018: Prague, Czech Republic*, vol.1, pp.183-189, Springer Singapore, 2019.
- [16] E.S.A. El-Dahshan, H.M. Mohsen, K. Revett, and A.B.M. Salem, "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm", *Expert systems with Applications*, vol.41, no.11, pp.5526-5545, 2014.
- [17] N. Ullah, A. Javed, A. Alhazmi, S.M. Hasnain, A. Tahir, and R. Ashraf, "TumorDetNet: A unified deep learning model for brain tumor detection and classification.," *Plos one*, vol.18, no.9, pp.0291200, 2023.
- [18] F.P. Polly, S.K. Shil, M.A. Hossain, A. Ayman, and Y.M. Jang, "Detection and classification of HGG and LGG brain tumor using machine learning", *International Conference on Information Networking (ICOIN)*, pp. 813-817, 2018.
- [19] 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.
- [20] S.K. Mathivanan, S. Sonaimuthu, S. Murugesan, H. Rajadurai, B.D. Shivahare, and M.A. Shah, "Employing deep learning and transfer learning for accurate brain tumor detection", *Scientific Reports*, vol.14, no.1, pp.7232, 2024.
- [21] 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.
- [22] A. Vidyarthi, R. Agarwal, D. Gupta, R. Sharma, D. Draheim, and P. Tiwari, "Machine learning assisted methodology for multiclass classification of malignant brain tumors", *IEEE Access*, vol.10, pp.50624-50640, 2022.
- [23] M.Z. Khaliki, and M.S. Başarslan, "Brain tumor detection from images and comparison with transfer learning methods and 3-layer CNN", *Scientific Reports*, vol.14, no.1, pp.2664, 2024.
- [24] G. Deng, and L.W. Cahill, "An adaptive Gaussian filter for noise reduction and edge detection", *IEEE conference record nuclear science symposium and medical imaging conference*, pp.1615-1619, October 1993.
- [25] A. Abdollahi, B. Pradhan, and A. Alamri, "VNet: An end-to-end fully convolutional neural network for road extraction from high-resolution remote sensing data", *IEEE Access*, vol.8, pp.179424-179436, 2020.
- [26] X. Liu, R. Yin, and J. Yin, "Attention V-Net: a modified V-Net architecture for left atrial segmentation", *Applied Sciences*, vol.12, no.8, pp.3764, 2022.

- [27] H. Huang, Z. Zuo, B. Sun, P. Wu, and J. Zhang, "DSA-SOLO: Double split attention SOLO for side-scan sonar target segmentation", *Applied Sciences*, vol.12, no.18, pp.9365, 2022.
- [28] R.S. Devi, V.R. Kumar, and P. Sivakumar, "EfficientNetV2 Model for Plant Disease Classification and Pest Recognition", *Computer Systems Science & Engineering*, vol.45, no.2, 2023.
- [29] M.M. Hasan, N. Alfaz, M.A.M. Alam, A. Rahman, H.M. Shakhawat, and S. Rahman, "Detection of parkinson's disease from t2-weighted magnetic resonance imaging scans using efficientnet-v2", *International Conference on Computer and Information Technology (ICCIT)*, IEEE, pp.1-6, December 2023.
- [30] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and De P. Geus, "Malicious software classification using transfer learning of resnet-50 deep neural network", *IEEE international conference on machine learning and applications (ICMLA)*, pp.1011-1014, December 2017.
- [31] H. Zhou, and G. Yu, "Research on pedestrian detection technology based on the SVM classifier trained by HOG and LTP features", *Future Generation Computer Systems*, vol.125, pp.604-615, 2021.
- [32] R. Arya, and E.R. Vimina, "Local triangular coded pattern: A texture descriptor for image classification", *IETE Journal of Research*, vol.69, no.6, pp.3267-3278, 2023.
- [33] K.D. Kharat, V.J. Pawar, and S.R. Pardeshi, "Feature extraction and selection from MRI images for the brain tumor classification", *International Conference on Communication and Electronics Systems (ICCES) IEEE*, pp.1-5, October 2016.
- [34] A. Chaddad, and R.R. Colen, "Statistical feature selection for enhanced detection of brain tumor", *In Applications of Digital Image Processing XXXVII, SPIE*, Vol.9217, pp.260-267, September 2014.
- [35] M. Aspri, G. Tsagkatakis, and P. Tsakalides, "Distributed training and inference of deep learning models for multi-modal land cover classification", *Remote Sensing*, vol.12, no.17, p.2670, 2020.
- [36] M. Marjani, M. Mahdianpari, and F. Mohammadimanesh, "CNN-BiLSTM: A Novel Deep Learning Model for Near-Real-Time Daily Wildfire Spread Prediction", *Remote Sensing*, vol.16, no.8, pp.1467, 2024.
- [37] M. Méndez, M.G. Merayo, and M. Núñez, "Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model", *Engineering Applications of Artificial Intelligence*, vol.121, pp.106041, 2023.
- [38] BraTS 2018 dataset, "https://www.med.upenn.edu/sbia/brats2018/data.html", accessed on October 2024.
- [39] BraTS 2019 dataset, "https://www.med.upenn.edu/cbica/brats2019/data.html", accessed on October 2024.