Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 21 No. 6 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v21i06.54367

PAPER

DATDL-DCNN-BiLSTM: Dual Attention Temporal Difference Learning Based Distributed Deep Learning Model for Brain Tumor Detection

Sayeedakhanum Pathan[1,2](✉), Savadam Balaji[1], S. Sai Anuraghav[3]

[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India

[2]Department of CSE – AIML & IoT, VNRVJIET, Hyderabad, Telangana, India

[3]Department of Computer Science Engineering, Sanjose University, San Jose, CA, USA

pathan.sayeeda@klh.edu.in

ABSTRACT

The brain tumor (BT) is a critical disorder related to neurons characterized through the proliferation growth. The survival rates associated with this disease are steadily declining, primarily due to insufficient early detection and precise diagnosis of BT. The manual inspection often suffers from subjectivity, low accuracy, and inefficiency due to complex tumor shapes. To identify and handle the large sized datasets along with capturing the more subtle tumor features, this study proposes a model named socio-swarm intelligence optimizer (SSIO) enabled dual attention temporal difference learning with distributed convolutional neural network and bidirectional long short-term memory (SSIO-DATDL-DCNN-BiLSTM) framework for identification of BT, focusing on the integration of advanced optimization techniques. The dual attention-AlexNet layer efficient statistical triangular ResNet (DA-ALESTR) provides detailed localization of relevant features. The LayerCAM (layer-wise class activation map) and AlexNet features in the model add more deliberation by capturing high-level semantic patterns and enhancing interpretability. The SSIO optimizer increase the interpretation through adjusting the factors. The results of the proposed SSIO-DATDL-DCNN-BiLSTM demonstrates higher scores on accuracy, recall, F1-score, and precision with 99.56%, 98.27%, 96.80%, and 99.79% with training data, and with k-fold 96.95%, 96.96%, 96.77% and 97.15% using BraTS 2018 dataset.

KEYWORDS

brain tumor (BT), socio-swarm intelligence optimization (SSIO), dual attention, AlexNet, layer-wise class activation map (LayerCAM)

1 INTRODUCTION

Brain tumor (BT) is defined as an abnormal growth within the tissues surrounding the human skull, significantly affecting health. They can be classified as malicious or non-malicious. The main cells originate inside the scalp. Such common types of

Pathan, S., Balaji, S., Sai Anuraghav, S. (2025). DATDL-DCNN-BiLSTM: Dual Attention Temporal Difference Learning Based Distributed Deep Learning Model for Brain Tumor Detection. *International Journal of Online and Biomedical Engineering (iJOE)*, 21(6), pp. 56–75. https://doi.org/10.3991/ijoe.v21i06.54367

Article submitted 2025-01-12. Revision uploaded 2025-03-07. Final acceptance 2025-03-07.

BTs include gliomas and many others. The slow-rate growth develops near the backside of the skull and spinal cord [1]. They can produce various complications related to the specific affected area, which may contain seizures, headaches, vomiting, cognitive changes, and visual disturbances [10]. More specific symptoms can manifest as difficulties in speech and walking, as well as sensory issues. An early BT diagnosis can save lives and reduce the risk of disability. The technique of magnetic resonance imaging (MRI) is employed for identification because of superior imaging capabilities [3]. Currently, various modalities, such as scanning and laser treatments, are utilized for the diagnosis of BTs. Among these modalities, MRI stands out for its ability to deliver detailed images, offering clear insights into the cerebral structure and identifying abnormalities within brain tissues [6]. The application of extensive datasets helps in detection that can be enhanced through the implementation of machine learning and deep learning methods [4]. Various transfer learning models have demonstrated significant effectiveness in this domain [5].

Analyzing brain MRI images presents challenges such as resolution, contrast, motion concerns, etc., which complicates tumor segmentation and the classification of normal versus abnormal images [7]. Image segmentation entails isolating specific regions of interest given the vast amount of data present in each image [11]. Medical examiners often utilize modalities concurrently for the process of segmentation. The earlier approaches predominantly depend on the manual extraction of a limited number of features, a process that is labor-intensive, time-consuming, and subjective, thereby hindering automated analysis [24] [8]. Conventional ML methods rely on handcrafted features, which restrict the robustness of the results, whereas DL techniques automatically extract significant features, leading to markedly improved performance. Nevertheless, the outcomes achieved through deep learning methods are still inadequate, and they do not consistently yield high classification accuracy [9]. The conventional techniques for BT classification posed several challenges. Some of them are listed below:

- The SSBTCNet [3] leverages limited label data to effectively train the model. This leads to poor convergence in deep neural networks with excessive layers, which can lead the model to get stuck into local minima.
- The deep learning approach [5] effectively detects BT, but with limited availability of the large annotated datasets of brain MRI images, it leads to overfitting and reduced model generalization.
- The multiple divisions of patches in the image produced increased computational cost making the model difficult to implement and optimize effectively [6].
- The upgraded images may cause improper enhancement which can lead to distorted tumor features when dealing with noisy or inconsistent data from different sources [7].
- The model [8] demonstrates greater effectiveness in classifying BT. However, it encountered limitations with certain data types. Additionally, the utilization of ResNet50 and GAN resulted in significant computational expenses during both installation and training, rendering the model resource-intensive.

The main motivation behind this work is to design and develop SSIO-enabled DATDL-DCNN-BiLSTM model for precise detection and classification of BT. The SSIO is combined alongside DATDL-DCNN-BiLSTM for the accurate outcomes. Additionally, significant and relevant features are derived through the application of the DA-ALESTR. The extracted information obtained from various methods are combined to create a map, which enhances better outcomes in classification tasks. The optimizer improves generalization by reducing the loss. The main contributions of this study are summarized below:

- **Dual attention-AlexNet layer efficient statistical triangular ResNet (DA-ALESTR) based feature extraction:** The method based on DA-ALESTR for extracting features integrates various approaches to effectively capture a wide range of tumor characteristics, facilitating precise detection and classification. By extracting diverse features, this methodology enhances clinical reliability, enabling the model to achieve robust, interpretable, and accurate localization of tumor features that are essential for effective tumor detection.
- **Socio-swarm intelligence optimization:** The SSIO algorithm combines the social intelligence behavior of gorillas and the swarm mechanism of spider monkeys. The SSIO algorithm provides higher reliability and increases interpretability by gaining the better solution.
- **Socio-swarm intelligence optimization enabled dual attention temporal difference learning distributed CNN-BiLSTM (SSIO-DATDL-DCNN-BiLSTM):** The approach effectively generalizes intricate details that is essential in the classification process. By incorporating the SSIO with the DATDL-DCNN-BiLSTM framework, the model achieves precise classification of BT based on its labels, enhancing accuracy, minimizing overfitting, and reducing errors. In summary, the proposed SSIO-DATDL-DCNN-BiLSTM model significantly improves the efficiency and accuracy of BT classification.

The subsequent sections of this manuscript are delineated as follows: Section 2 deals with literature survey where the previous research will be discussed and identifies the gaps. Section 3 offers an in-depth explanation of the proposed methodology. Section 4 and 5 addresses the results and discussion part respectively. Finally, Section 6 encapsulates the entirety of the research, highlighting key findings, implications, and future work.

2 SURVEY OF THE LITERATURE

Rahman, T., and Islam, M.S. [1] approach involved extracting important information, in parallel, addressing overfitting concerns. The model's efficacy was tested on three distinct BT datasets, yielding accuracy rates of 97.33%, 97.60%, and 98.12% by capturing relevant features. Ullah, N. et al. [2] created a comprehensive approach utilizing a advanced strategy, known as TumorDetNet, for the detection and classification of BT. A SoftMax layer was employed to identify tumors and their respective grades. It demonstrated a high accuracy rate but struggled with complex tumor types. The classification performance for BT was evaluated across six MRI datasets, achieving 99.833% accuracy for distinguishing between dangerous cells and 99.27% for other tumor classifications. Z. Atha and J. Chaki [3] developed a approach known as SSBTCNet for detecting BT. It was constructed by combining an encoder and a classifier framework. The occurences are generated by degrees of truth utilizing modified copies of data. Although it demonstrated a high rate, it was unable to identify the tumor's location. The model was evaluated using various datasets, achieving an accuracy of 96.5%, surpassing other existing methods.

Anantharajan, S. et al. [4] modeled a framework for earlier detection of tumors. In this approach, an adaptive-based contrast enhancement algorithm (ACEA) and segmentation through a fuzzy c-means model. This method encountered issues related to time complexity. The authors proposed a classifier called ensemble deep neural support vector machine (EDN-SVM) to differentiate between abnormal and normal tissues, resulting in notable improvements in accuracy, sensitivity, and specificity, with values of 97.93%, 92%, and 98%, respectively. Mathivanan, S.K. et al. [5] explored

several networks along with transfer methodologies for diagnosis. It gained precise diagnostic accuracy with limited medical imaging data but resulted in overfitting due to small sample sizes. The study utilized the Kaggle database for evaluating brain tumors and showed a superior performance of 99.755% in accuracy. Muezzinoglu, T. et al. [6] introduced an approach (PatchResNet) for classifying brain tumors by dividing MRI images into patch divisions and a ResNet50 for training. Additionally, the framework is self-organized, and capable of selecting optimal predictions but struggles with computational requirements and complexity issues. The experiment employed a public BT image dataset and reported a classification accuracy of 98.10%. Asiri, A.A. et al. [7] designed a model aimed at enhancing the accuracy of BT classification. The image enhancement technique and support vector machines are used to normalize images and validate the performance of tumor classification. However, the high-quality MRI images and the multiple models resulted in computational overhead. The results demonstrate that the approach showed improvements by superior performance. Nag, A. et al. [8] introduced a unique approach (TumorGANet) that uses transfer learning and generative adversarial networks (GANs) for BT classification. The integration of both techniques leads to improved performance in distinguishing tumors, mitigating the issues related to limited MRIs. The experiment is carried out by analyzing 7023 MRI images, and the TumorGANet achieved superior performance by acquiring values such as 99.533%, 100%, 99%, and 0.2% on the evaluation metrics.

From the overall survey, the existing methods were facing the below problems, such as SSBTCNet [3] leveraging limited label data to effectively train the model. This leads to poor convergence in deep neural networks with excessive layers, which can lead the model to get stuck in local minima. The deep learning approach [5] effectively detects BT, but with limited availability of the large annotated datasets of brain MRI images, it leads to overfitting and reduced model generalization. The multiple divisions of patches in the image produced increased computational cost, making the model difficult to implement and optimize effectively [6]. The upgraded images may cause improper enhancement, which can lead to distorted tumor features when dealing with noisy or inconsistent data from different sources [7]. The model [8] demonstrates greater effectiveness in classifying BT. However, it encountered limitations with certain data types. Additionally, the utilization of ResNet50 and GAN resulted in significant computational expenses during both installation and training, rendering the model resource-intensive.

To overcome the above-mentioned limitations, an SSIO-enabled DATDL-DCNN-BiLSTM model is proposed to perform the accurate BT detection and classification. To accurately classify the BT, the SSIO is integrated with DATDL-DCNN-BiLSTM to achieve an efficient result. Furthermore, the important and suitable features are being extracted by applying DA-ALESTR 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 accuracy loss.

3 PROPOSED METHODOLOGY

The primary objective considered for the study is to detect BT utilizing an SSIO-DATDL-DCNN-BiLSTM model. The process begins with a Gaussian model-based preprocessing stage, to enhance the quality of the brain images. Subsequently, the outcome images are transformed to the DA-based VNet segmentation model. Following segmentation, the images are then fed into DA-ALESTR to identify the significant information that contains relevant features, minimizing any potential

loss in accuracy. These features are subsequently allowed in the SSIO-DATDL-DCNN-BiLSTM model, which effectively addresses long-range dependency challenges to classify and grade the BT accurately. The final classification categorizes the output into four categories: no tumor, non-enhancing tumor, enhancing tumor, and peritumoral edema. A brief visualization is illustrated in Figure 1.

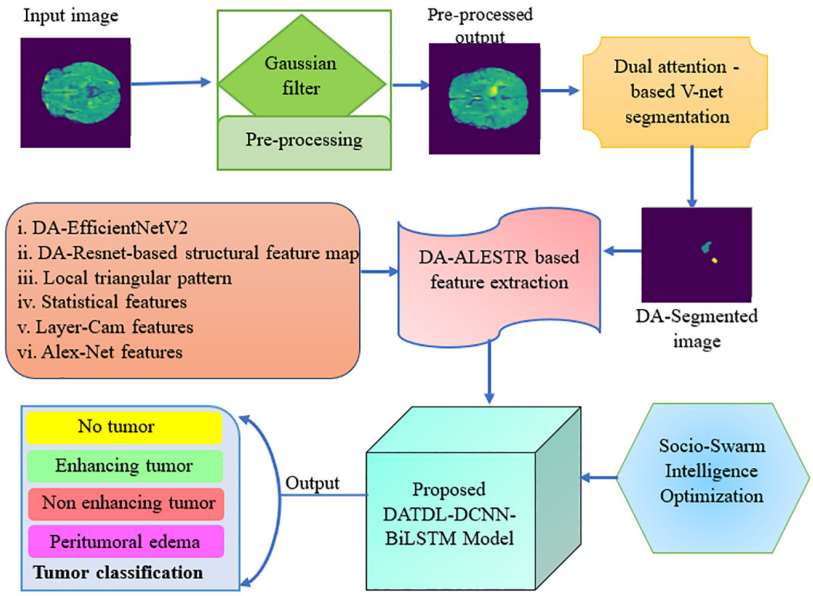

Fig. 1. Diagrammatic illustration of SSIO-DATDL-DCNN-BiLSTM framework

3.1 Preprocess utilizing Gauss model

The MRI images contain different types of distorted noises that affect the image quality. To avoid this distortion, a Gaussian filter is used that reduces the noise and elevates the quality. The resultant enhanced image can be represented in the equation (1) as follows:-

$$B_p^* = G(B_p) \tag{1}$$

Here, B_p^* denotes the Gaussian-based preprocessed image and G denotes the Gaussian filter.

3.2 Dual attention-based VNet segmentation

The pre-processed output B_p^* is utilized, that has the DA-VNet to produce segmented results. The DA-VNet is developed by integrating VNet [28] with a dual attention mechanism (DA), which incorporates both the attentions such as split and double. The split attention functions as a computational unit that consists of split attention operations and grouped feature maps. Conversely, the double attention mechanism evaluates features that are under pool, effectively capturing complicated visuals by improving the correlation of features with specific locations. The VNet receives B_p^* as input, that is employed in the subsequent layers of convolutions, ultimately producing a value using the last layer. The outcome produced by DA-VNet will have $[N \times 224 \times 224]$ dimension.

The two attention models divide the input dimension [$N \times 224 \times 224$] into distinct segments. These models are designed to extract boundary information effectively. The architecture of the split attention model comprises various layers, including reshape layer, pooling, dense layer, and finally the softmax layer. Conversely, the double attention model features a same set of layers, employing additional layers along with matrix multiplication. The layer of convolution executes the operation, while the reshape layer facilitates the reduction of feature map dimensions. The outputs from these two distinct layers are combined through matrix multiplication, resulting in a specific dimension [$N \times 224 \times 224$].

3.3 Dual attention-based AlexNet layer efficient statistical triangular ResNet (DA-ALESTR) feature extraction

Here a feature extraction method known as DA-ALESTR has been developed, which integrates various feature extraction models, including features from AlexNET, layer CAM, dual attention-based EfficientNet V2, dual attention-based ResNet structural maps, local triangular pattern (LTP), and statistical features. The brief explanation of these is provided below:

Dual attention-based EfficientNet V2 features. The DA employed feature is developed by integrating attention based components into the conventional EfficientNetV2 framework. The EfficientNetV2 [19] model significantly enhances its ability to inherent relevant parts such as tumor regions and so on. It consists of various layers, each serving a distinct purpose. At first, the M with dimension [$N \times 224 \times 224$] is directed to the first layer along with defined stride, and outcome is processed through the next layer, that mitigates overfitting through separable convolution. The DA layer incorporated is subsequently routed to the next layer of convolution with a defined stamp. Ultimately, other sub-ordinate layers are employed to produce the first feature L_1 of [$1 \times 112 \times 112 \times 32$] that is re-shaped to [$1 \times 32 \times 32$].

Dual attention-based ResNet structural map features. This architecture comprises various layers, including convolution, batch normalization, max pooling, flattening, a fully connected layer, and the local ternary pattern (LcTP). The LcTP [20] is utilized for texture analysis of grayscale images. It functions by designating a value of gray E_j for a specific pixel and the gray values of neighboring pixels E_i ($i = 0, 1, ..., 7$). It examines the relationship between E_i and E_j values; if the condition $E_i < E_j$ is met, the value is 0, otherwise, it is assigned as 1. Through this method, an 8-bit number made up of 0s and 1s is produced, with the corresponding decimal value representing the values as single digits is considered for LcTP as,

$$L_2 = \sum_{i=0}^{7} \chi(E_i - E_j)2^i;\ \chi(a) = \begin{cases} 1;\ a \geq 0 \\ 0;\ Otherwise \end{cases} \tag{2}$$

Here, L_2 structural map features. χ distribution function that shows the relationship between E_i and E_j. In general, the DA layer is integrated into the ResNet-50 model, and LcTP, that generate L_2, and has dimension [$1 \times 32 \times 32$].

Local triangular pattern-based features. Local triangular pattern is a texture descriptor that compares pixel intensities in a local neighborhood to capture fine textures and edges. LTP [21] works by comparing the initial pixel and its neighbors using a threshold. It encodes textual information by transforming pixel differences into ternary values as 1,0 or –1, depending on whether the value is large, or less than neighboring pixels. The extracted LTP histogram L_3 provides texture features that can be useful for distinguishing different tissues or abnormalities.

$$L_3 = \begin{cases} 1 & ;if\ I_{(x,y)} - I_{(x',y')} > T \\ -1 & ;if\ I_{(x,y)} - I_{(x',y')} < -T \\ 0 & ;otherwise \end{cases} \quad (3)$$

Here, T indicates the threshold value, $I_{(x,y)}$ denotes the intensity of pixels whereas, $I_{(x',y')}$ denotes the intensity of the neighboring pixel. L_3 denotes the LTP features and a dimension of $[1 \times 32 \times 32]$.

Statistical features. Statistical features capture the distribution and variability of pixel intensities, which is essential for distinguishing between normal and abnormal regions of tumors. The statistical feature used for the process are mean, median, variance, standard deviation, harmonic mean, geometric mean, and entropy etc. The features of statistics derived from the image that undergone segmented outcomes are articulated as

$$L_4 = \{N_1, \ldots, N_7\} \quad (4)$$

The statistical characteristics are illustrated with L_4 the specified size $[1 \times 7]$, and the size is adjusted to $[1 \times 32 \times 32]$ for facilitating the precise training of the data.

AlexNet-based features. This uses large kernel sizes in initial layers to capture the aforementioned features. The model uses ReLU for non-linear activation to accelerate the training. The local response normalization (LRN) and dropout regularization are the two key concepts used in AlexNet [22]. The LRN enhances the feature response along with the generalization by creating competition between relevant regions. The dropout regularization controls the challenge of overfitting by randomly assigning activation to zero during training. The max-pool deduces inherent and texture-based length, and finally, the fully connected (FC) layers perform final integration and classification employing a softmax activation.

$$L_5 = soft\ \max(g \cdot M + b) \quad (5)$$

where, g indicates the weights of the network, b indicates the bias L_5 indicates the Alexnet-based feature with the dimension of $[1 \times 32 \times 32]$.

LayerCAM features. LayerCAM is an advanced visualization technique designed to enhance the interpretability of CNNs by generating class activation maps (CAMs) at various hierarchical layers. This approach helps to visualize which area of the image leads to the decision process. LayerCAM generates activation maps across multiple layers of the CNN. This multi-layer approach captures both coarse and fine-grained features. LayerCAM [23] utilizes target class gradient concerning feature maps to determine the importance of each feature map. By combining activation maps from different layers, LayerCAM produces high-quality CAMs that accurately localize object regions, enhancing the performance of prediction tasks. The process of CAM generation is performed by weighting the feature maps along with their importance weight value to generate the layercam feature L_6.

$$L_6 = \sigma\left(\sum_t w_{u,v}^{tc} \cdot S_{u,v}^t\right) \quad (6)$$

where, σ is ReLu activation function, $w_{u,v}^{tc}$ importance weight, $S_{u,v}^t$ Feature map for spatial values u, v.

The final feature map produced by DA-ALESTR after extracting the specific features that are essential for the detection and classification is denoted mathematically as,

$$L = \left\{ L_1 \| L_2 \| L_3 \| L_4 \| L_5 \| L_6 \right\} \tag{7}$$

Here, the feature map of DA-ALESTR is denoted by L and its size is $[1 \times 32 \times 32 \times 4]$, respectively.

3.4 Socio-swarm intelligence optimization

This algorithm leverages the unique characteristics of artificial gorilla troops [14] as well as spider monkeys [16] to enhance BT detection and classification. The new swarm intelligence technique of gorillas and the social foraging behavior of grouped monkeys make the SSIO algorithm more feasible and optimal. The strategic behavior of searching for food in gorilla troops excels in global exploration by diverting complicated positions without the fall of trap. In contrast, the social organization behavior of distributing food by group communication and foraging characteristics in spider monkeys specializes in local exploitation, ensuring precise fine-tuning of solutions. By combining these strengths, the SSIO achieves an optimal balance between exploration and exploitation, crucial for processing complex and high-dimensional data in medical imaging. This synergy improves feature selection and classification accuracy by mitigating the issues of computational complexity and falling to local optima. Such an integrated approach, SSIO, offers significant potential for advanced medical diagnostics and addressing the critical challenges in BT analysis. The various phases involved in this are represented in Figure 2.

A) Initialization: Let us assume the position is demonstrated using the matrix of population z_u. The population can also be represented mathematically in a matrix as,

$$z_u : z_{u,v} = eb_v + r(0,1) \cdot (fb_v - eb_v) \tag{8}$$

$$z = \begin{bmatrix} z_{1,1}, \ldots, z_{1,v}, \ldots, z_{1,J} \\ \vdots\vdots\vdots \\ z_{u,1}, \ldots, z_{u,v}, \ldots, z_{u,l} \\ \vdots\vdots\vdots \\ z_{m,1}, \ldots, z_{m,v}, \ldots, z_{m,l} \end{bmatrix}_{m \times l} \tag{9}$$

where, z_u is the initial vector of the u^{th} solution, l, m indicates the solutions and the search spaces whereas, fb_v and eb_v indicates the primary and secondary bounds and u, v indicates the dimensions.

B) Evaluation of fitness function: The measure of the solutions is calculated for the best fitness is assigned with the metric of accuracy and can be shown as,

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_u \\ \vdots \\ F_m \end{bmatrix}_{m \times 1} = \begin{bmatrix} F(z_1) \\ \vdots \\ F(z_u) \\ \vdots \\ F(z_m) \end{bmatrix}_{m \times 1} \tag{10}$$

where F denotes value of the fitness and the maximum value of the fitness determines the better solution to be selected.

C) **Solution updation phase:** The solutions have been revised to incorporate the characteristics of both specimens. The phases developed with these traits improve the SSIO-DATDL-DCNN-BiLSTM model's capacity for effective training.

If the current iteration i.e q is lessthan the max iteration qmax then the pahse (i) of the artificial gorilla troops [15] gets executed which has three sub-phases and are described below,

Sub phase (i) Evacuation phase: if $r_a < h$, the position update equation according to the random search space can be determined as follows,

$$z_u^{q+1} = (eb_u - fb_u) * w_1 + fb_u \tag{11}$$

where, w_1 is the random variable $\in (0, 1)$ z_u^{q+1} is the randomly chosen solution from the u^{th} dimension.

Sub phase (ii) Co-ordinating movement phase: if $r_a \geq 0.5$, then the position can be updated as

$$z_u^{q+1} = (w_2 - J) * z_w^{(q)} + B \times D \tag{12}$$

where, w_2 is the random variable that belongs to [0, 1] and $z_w^{(q)}$ represents a randomly selected solution and J, B and D implies its transition factors.

Sub phase (iii) Un-familiar territory phase: if $r_a < 0.5$, Then the equation for updating the position as

$$z_u^{q+1} = z_u^q - B \times \left[B \times \left(z_u^{(q)} - z_w^{(q)} \right) + w2 \times \left(z_u^{(q)} - z_w^{(q)} \right) \right] \tag{13}$$

where, w2 is a random variable $\in (0, 1)$. It uses the current position and random position to explore the global search for a new area. This leads to a maximum chance of finding a global solution.

Phase 2: Congregate phase: After performing the movement phase, the congregate phase gets executed.

If $J \geq P$, the position update equation

$$z_u^{q+1} = B \times U \times \left(z_u^{(q)} - z_b^{(q)} \right) + z_u^{(q)} \tag{14}$$

where, P is the parameter and is set to $P = 0.8$. Usually, during the congregate phase, the solution refers to the global best solution, but here additional information based on the history is incorporated that yields the maximum chance of finding a global better solution.

Competitive phase: if $J < P$ then, $z_u^{(q+1)} = z_b - (z_b \times t - z_u^q \times t) \cdot f$ (15)

$$t = 2 \times r_7 - 1 \tag{16}$$

$$f = \alpha \times S \tag{17}$$

where, t is a factor based on the random value $r_7 \in (0, 1)$ and f is the coefficient vector, α and S are the scaling factors. S shows the boundaries affecting under the threshold of 0.5.

D) **Termination:** If the current iteration exceeds the maximum limit q_{max}, the SSIO algorithm concludes and identifies the overall best solution. Consequently, the SSIO algorithm enhances the performance of DATDL-DCNN-BiLSTM through parameter optimization. Figure 2 illustrates the flowchart representation of the SSIO algorithm.

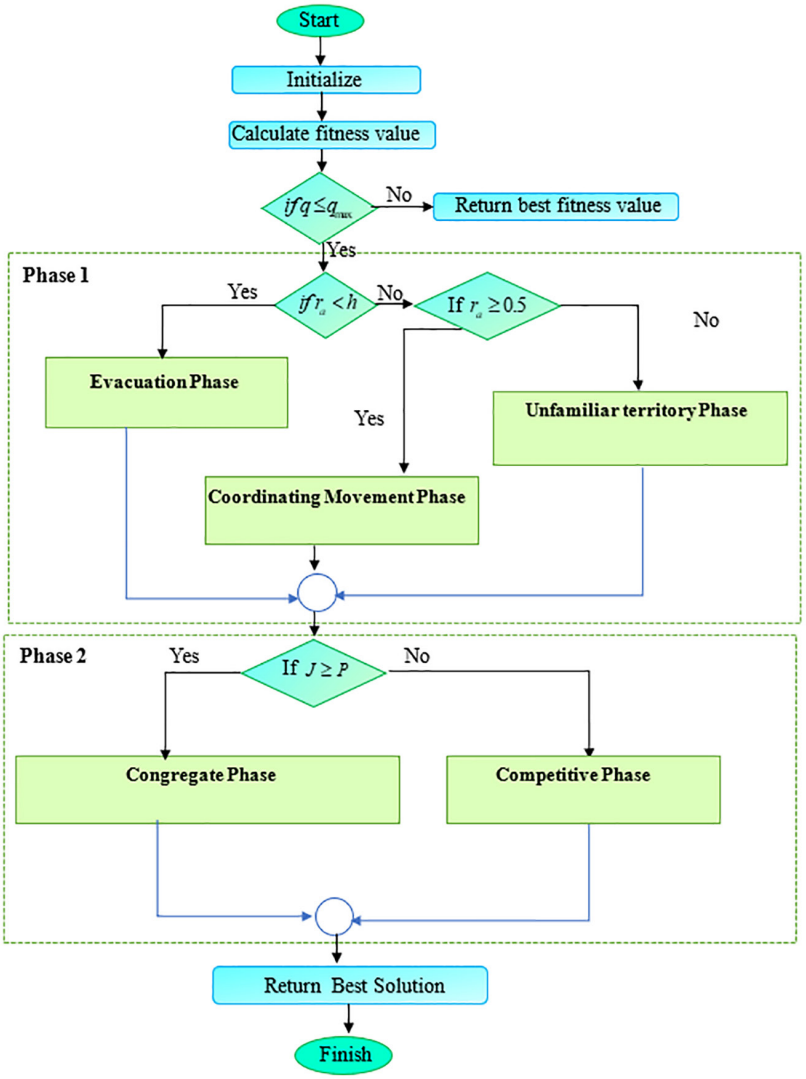

Fig. 2. Flowchart of the SSIO algorithm

3.5 DATDL-DCNN-BiLSTM-based classification framework

This framework incorporates dual attention models, such as split and double techniques, alongside temporal difference learning (TDL) [13] with CNN-BiLSTM [12], utilizing distributed training methodologies. The TDL is a reinforcement learning approach that updates values based on the difference between consecutive predictions. Combining such techniques with DCNN and BiLSTM leads to performing precise detection tasks along with greater efficiency. Figure 3 represents a architectural illustration of the DATDL-DCNN-BiLSTM.The classification method comprises diverse layers, including layer of convolution, max pooling, layer of time distribution, bi-directional, and dense layers. The input receives data from the first layer with specified dimensions [$N \times 32 \times 32 \times 4$] and forwards it to the next layer. The convolutional layer employs the ReLU activation function, to transform all negative values into zero. The model employs two convolution layers (conv2d) followed by pooled layer of max to deduce dimensionality. The pool technique serves as a sub-ordinate model for sampling to reduce the dimension of the matrix while simultaneously

enhancing the robustness of the framework. Following the operations conducted by the max pooling layer, the DA layer is implemented, which produces features with a specified dimension [$N \times 2 \times 2 \times 32$]. Subsequently, a flattened way of layer is employed for normalizing their original size from [$N \times 2 \times 2 \times 32$] to a new one [$N \times 2 \times 64$], while the layers in the BiLSTM further compresses the dimension to [$N \times 20$]. The outputs obtained from the BiLSTM layers are combined by multiplying them and fixing it to [$N \times 20$]. Consequently, precise classification can be achieved by reducing the factor of loss.Thus, the approach produces outputs corresponding to four distinct categories: normal case, non-enhancing tumor, peritumoral edema, and enhancing tumor.

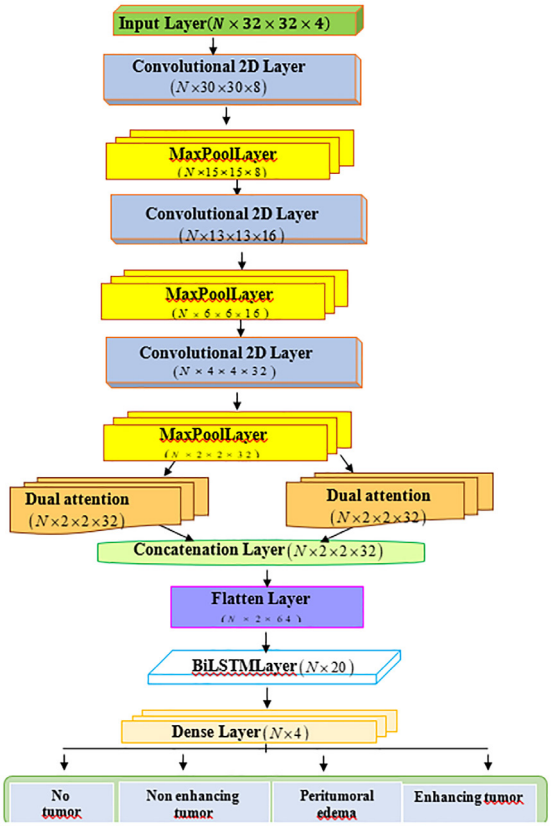

Fig. 3. Architecture of SSIO-DADTDL-DCNN-BiLSTM classification mode

4 EXPERIMENTAL RESULTS

4.1 Experimental setup

The SSIO-DADTL-DCNN-BiLSTM model was implemented using the PYTHON programming environment on a Windows 10 equipped with an IntelCore processor and 16 GBRAM.

4.2 Dataset description

The experiments were conducted through the datasets namely the BraTS 2018 [17] and the BraTS 2019 [18] dataset. Each dataset features four MRI modalities

for each case: native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 fluid attenuated inversion recovery (T2-FLAIR). Expert neuroradiologists have manually segmented these images to delineate tumor sub-regions, including four categorization.

4.3 Sample image results

The Figures 4 and 5 represents the obtained image results based on BraTs 2018 [17] and BraTS2019 [18] dataset corresponding to various phase. Each class has four classes along with its associated imaging modalities, which are captured under different areas.

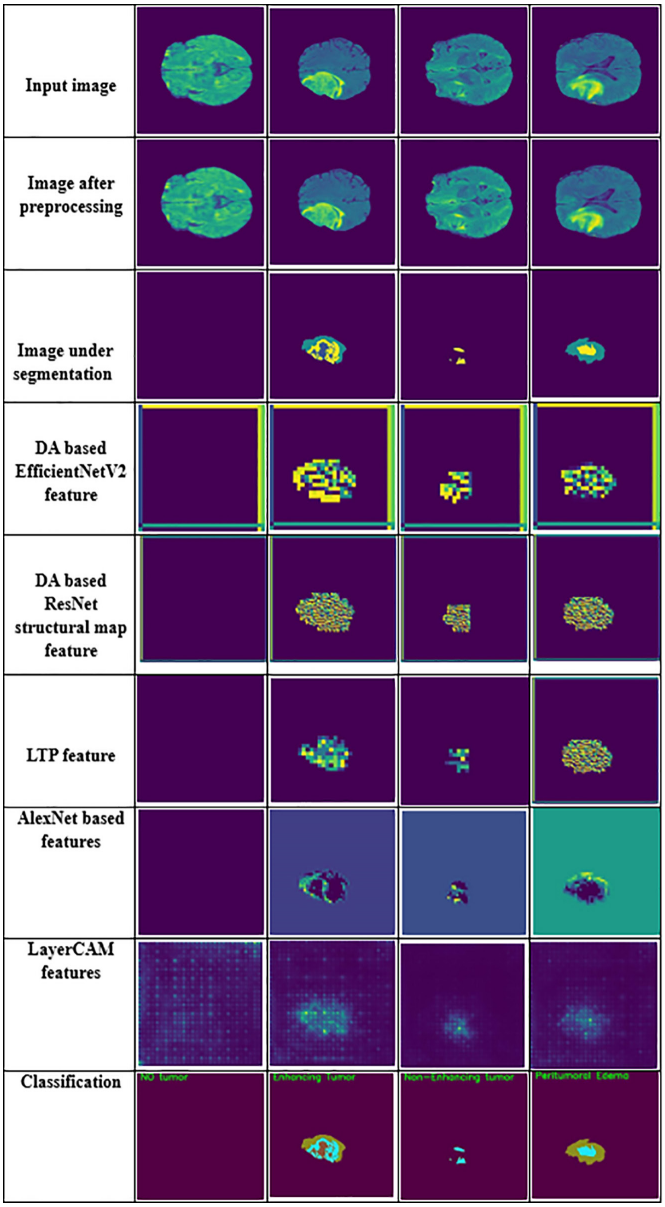

Fig. 4. Sample results of images based on the BraTS 2018 dataset

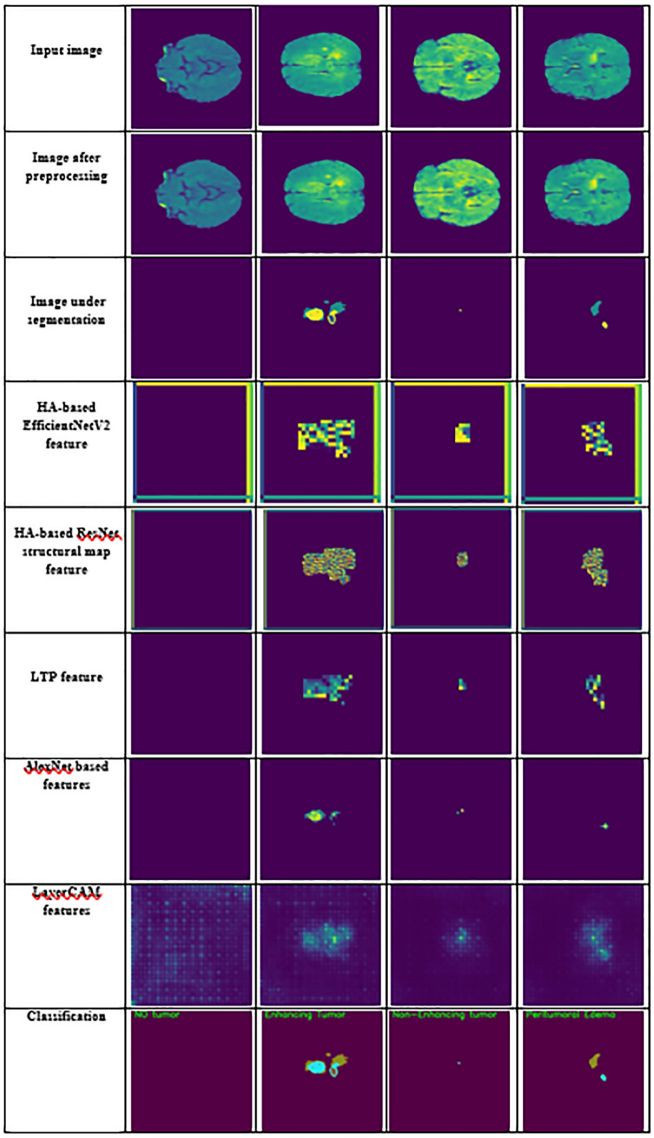

Fig. 5. Sample image results based on the 2019 dataset of BraTS

4.4 Comparative analysis

The SSIO-DATDL-DCNN-BiLSTM model is estimated and compared with existing techniques such as PDCNN, EDN-SVM, TumorDetNet, SSBTCNet, DCNN-BiLSTM, HATDL-DCNN-BiLSTM, AGTO-DATDL-DCNN-BiLSTM, SMO-DATDL-DCNN-BiLSTM, and to evaluate the effectiveness of the proposed model, we adjusted the train percent and the k-fold. The proposed model with the BraTS2018 dataset achieved accuracy, F1-score, precision, and recall values of 99.56%, 98.27%, 96.80%, and 99.79%, respectively, with varying training percentages, and with a k-fold of 10, the model yielded accuracy, F1-score, precision, and recall of 96.95%, 96.96%, 96.77%, and 97.15%, represented in Table 1. The graphical illustration can be seen in Figures 6 and 7. The model with the BraTS2019 dataset, achieve accuracy, F1-score, precision, and recall values of 98.17%, 98.10%, 97.27%, and 97.22%, respectively, with varying training percentages, and with a k-fold of 10, the model yielded accuracy, F1-score,

precision, and recall value of 97.22%, 97.36%, 96.47%, and 98.25%, as presented in Table 2. The same is illustaterd graphically in Figures 8 and 9.

Table 1. Comparison of SSIO-DATDL-DCNN-BiLSTM with the BraTS 2018 dataset

	Methods	PDCNN	EDN-SVM	TumorDetNet	SSBTCNet	DCNN-BiLSTM	HATDL-DCNN-BiLSTM	AGTO-DATDL-DCNN-BiLSTM	SMO-DATDL-DCNN-BiLSTM	Proposed SSIO-DATDL-DCNN-BiLSTM
TP=80%	Accuracy (%)	86.43	81.33	87.73	85.25	92.60	98.93	95.77	97.35	**99.56**
	F1-Score (%)	86.42	81.31	87.72	85.21	91.56	97.67	94.63	96.16	**98.27**
	Precision (%)	87.05	82.52	88.05	86.05	92.78	96.17	94.47	95.32	**96.80**
	Recall (%)	85.79	80.13	87.40	84.38	90.38	99.21	94.80	97.00	**99.79**
K-fold=10	Accuracy (%)	89.29	82.16	92.55	86.51	94.71	96.34	95.52	95.93	**96.95**
	F1-Score (%)	89.28	82.15	92.54	86.50	94.70	96.33	95.51	95.92	**96.96**
	Precision (%)	89.16	82.55	92.73	85.86	95.01	96.15	95.58	95.87	**96.77**

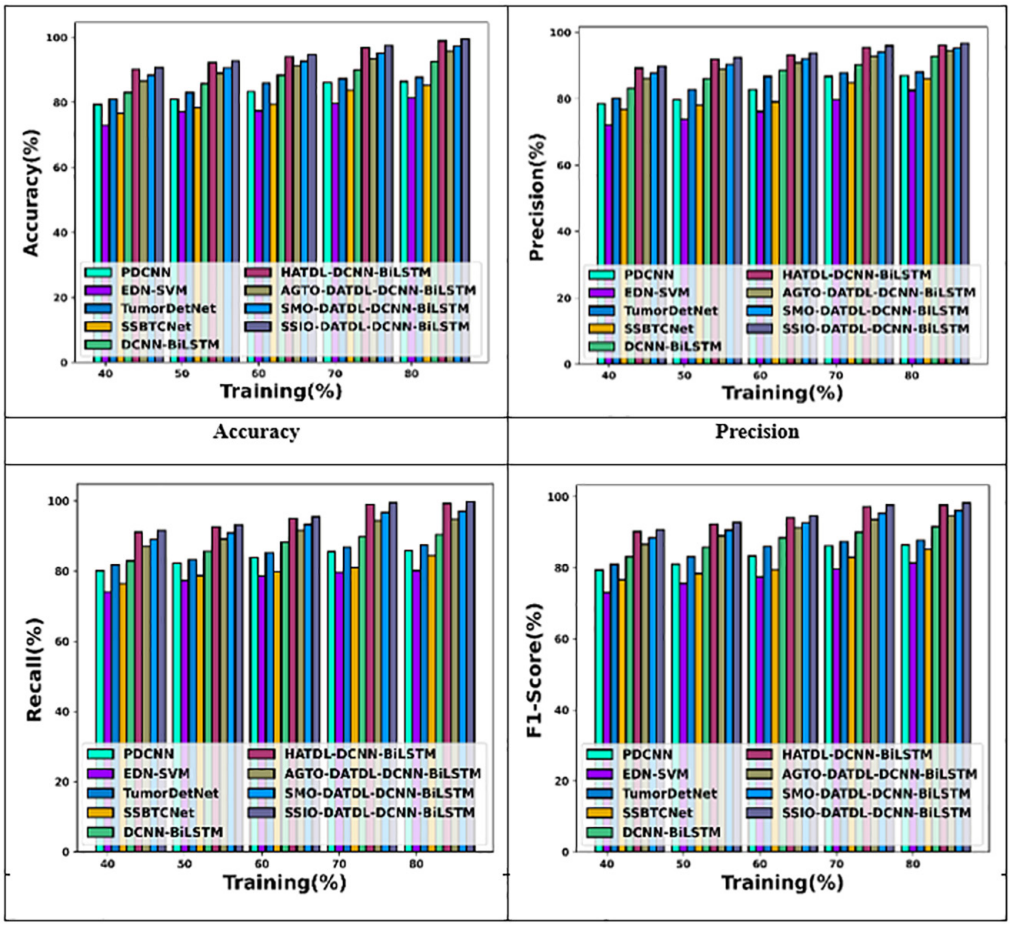

Fig. 6. Comparative analysis of proposed model with training percentage using 2018 dataset of BraTS

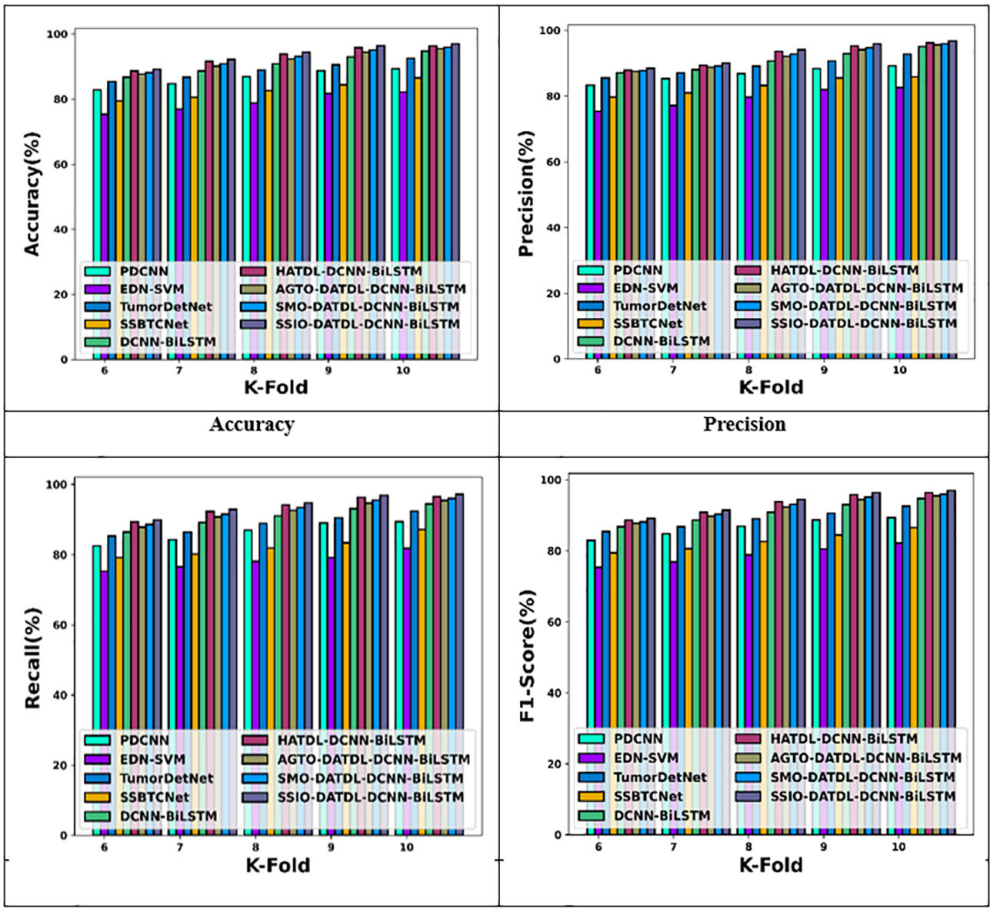

Fig. 7. Comparative analysis of proposed model k-fold using 2018 dataset of BraTS

Table 2. Comparison of SSIO-DATDL-DCNN-BiLSTM with the BraTS 2019 dataset

	Methods	PDCNN	EDN-SVM	TumorDetNet	SSBTCNet	DCNN-BiLSTM	HATDL-DCNN-BiLSTM	AGTO-DATDL-DCNN-BiLSTM	SMO-DATDL-DCNN-BiLSTM	Proposed SSIO-DATDL-DCNN-BiLSTM
TP=80%	Accuracy (%)	85.07	79.42	89.38	81.89	91.26	97.55	94.41	95.98	**98.17**
	F1-Score (%)	85.11	79.38	88.42	81.87	91.24	97.54	94.40	95.97	**98.10**
	Precision (%)	85.67	80.91	91.15	82.56	92.15	96.70	94.43	95.57	**97.27**
	Recall (%)	84.56	77.92	85.85	81.20	90.36	98.38	94.37	96.38	**98.95**
K-fold=10	Accuracy (%)	91.34	84.33	92.99	89.84	94.59	96.62	95.60	96.11	**97.22**
	F1-Score (%)	90.88	84.32	92.98	89.67	94.58	96.75	95.67	96.21	**97.36**
	Precision (%)	91.48	84.48	93.48	90.48	94.81	95.91	95.36	95.63	**96.47**
	Recall (%)	90.29	84.17	92.48	88.87	94.35	97.61	95.98	96.79	**98.25**

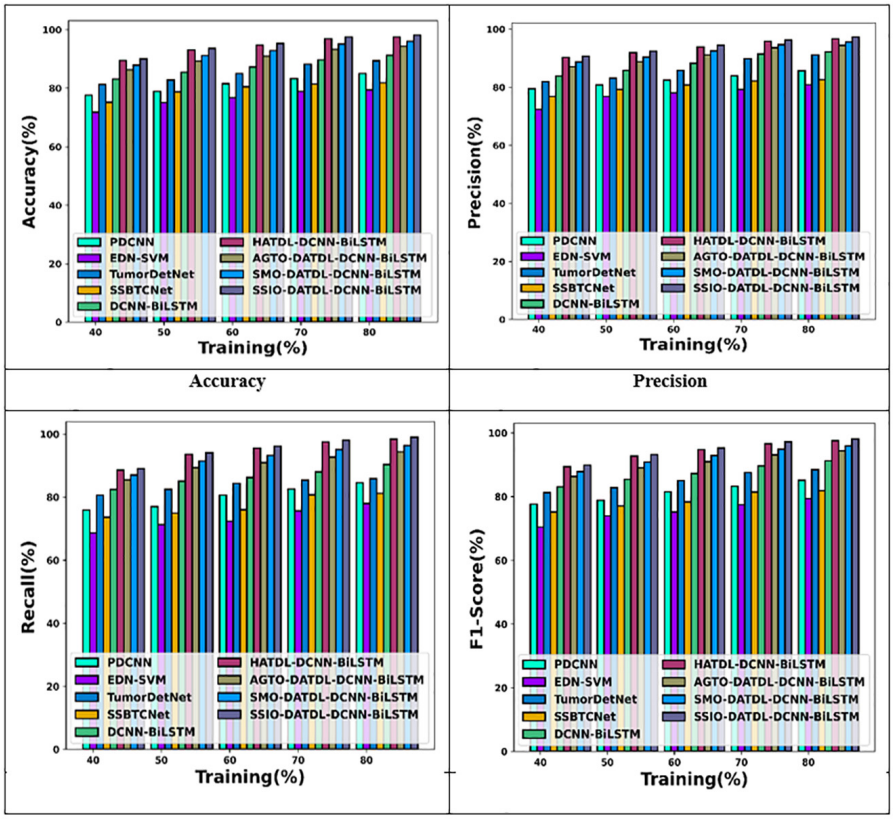

Fig. 8. Comparative analysis of proposed model with training percentage using 2019 dataset of BraTS

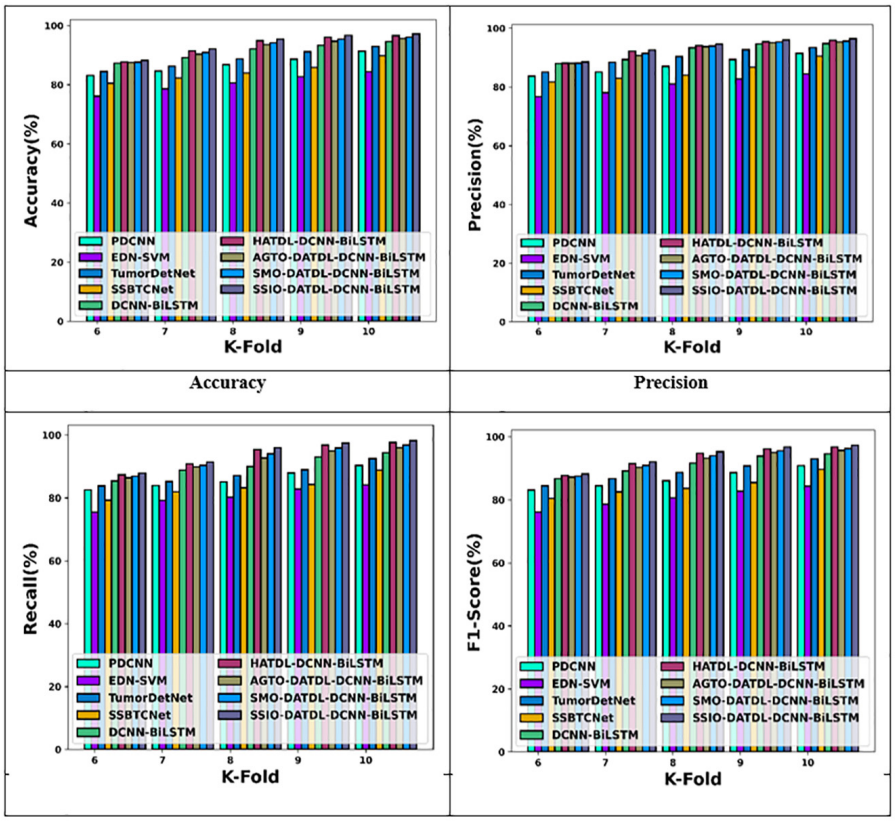

Fig. 9. Comparative analysis of proposed model with k-fold using the BraTS 2019 dataset

5 DISCUSSION

Table 1 offers a detailed comparison of the SSIO-DATDL-DCNN-BiLSTM model alongside other established methods utilizing the 2018 dataset of BraTS for BT classification. The examined existing techniques reveal that they fall short of optimal performance, particularly highlighted by their lower accuracy rates. Notably, the PDCNN method [1] did not utilize a 3D structural model for identifying tumors. In contrast, the TumorDetNet proposed in [2] achieved a commendable accuracy level, it struggled with accurately categorizing tumor grades. Similarly, the SSBTCNet method faced challenges in precisely identifying tumor types [3]. Although the EDN-SVM was designed for tumor classification and demonstrated better overall performance, it encountered issues with computational complexity. Addressing these limitations, the proposed SSIO-DATDL-DCNN-BiLSTM model integrates a distributed learning mechanism within a deep neural network structure. This innovative approach significantly reduces overfitting and improves performance across various metrics. Moreover, the SSIO-DATDL-DCNN-BiLSTM framework effectively manages computational challenges by optimizing feature extraction processes. The model standardizes the diverse dimensions of features derived from different methods, ensuring uniformity and minimizing complexity. To evaluate the effectiveness of the SSIO-DATDL-DCNN-BiLSTM, we adjusted the train percent and the k-fold. The model achieved impressive results on the BraTS2018 dataset, achieving accuracy, F1-score, precision, and recall values of 99.56%, 98.27%, 96.80%, and 99.79%, respectively, with varying training percentages. Furthermore, when using the BraTS2018 dataset, the SSIO-DATDL-DCNN-BiLSTM yielded accuracy, F1-score, precision, and recall of 96.95%, 96.96%, 96.77%, and 97.15%. Table 2 represents the discussion related to the comparison of SSIO-DATDL-DCNN-BiLSTM with the 2019 dataset of BraTS. It compares the proposed method with the above-mentioned existing approaches. When evaluated using the accuracy, F1-score, precision, and recall with the training percentage and k-fold, the proposed SSIO-DATDL-DCNN-BiLSTM attained values of 98.17%, 98.10%, 97.27%, and 97.22% on the 80% training percentage, whereas, on the k-fold of 10, the SSIO-DATDL-DCNN-BiLSTM model attained values of 97.22%, 97.36%, 96.47%, and 98.25%, respectively.

6 CONCLUSION

The proposed SSIO-DATDL-DCNN-BiLSTM provides an innovative solution for BT detection and classification by incorporating sophisticated feature extraction techniques, temporal-based learning, and advanced optimization. The amalgamation of layer CAM and AlexNet features allows for improved interpretability and capturing high-level patterns that are crucial for tumor classification. The dual attention in segmentation and feature extraction leads to better refining of features, ensuring the model focuses on relevant tumor characteristics. Additionally, the SSIO utilizes the unique characteristics of individual algorithms to upgrade the capability and fine-tune complex parameters, resulting in improved convergence and overall performance. This model not only addresses the challenge of overfitting issues but also provides a robust and scalable solution that can significantly support the decision-making process and early detection. The experiments conducted showed superior achievements utilizing the metrics, such as accuracy, recall, F1-score, and precision of 99.56%, 98.27%, 96.80%, and 99.79%, with training data and 96.95%, 96.96%, 96.77%, and 97.15% with k-fold on the BraTS 2018 dataset. The model with the BraTS2019 dataset, achievies accuracy, F1-score, precision, and recall values of

98.17%, 98.10%, 97.27%, and 97.22%, respectively, with varying training percentages, and with a k-fold of 10, the model yielded accuracy, F1-score, precision, and recall values of 97.22%, 97.36%, 96.47%, and 98.25%. The work should focus on future improvements in considering more precise techniques of optimizers, which speed up the network's efficiency, thereby providing accurate BT classification and prediction. The future enhancements could include integrating 3D convolutional networks for spatial analysis, improving real-time detection capabilities, and incorporating multi-modal imaging data for more accurate tumor classification.

7 REFERENCES

[1] T. Rahman and M. S. Islam, "MRI brain tumor detection and classification using parallel deep convolutional neural networks," *Measurement: Sensors*, vol. 26, p. 100694, 2023. https://doi.org/10.1016/j.measen.2023.100694

[2] 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, p. e0291200, 2023. https://doi.org/10.1371/journal.pone.0291200

[3] Z. Atha and J. Chaki, "SSBTCNet: Semi-supervised brain tumor classification network," *IEEE Access*, vol. 11, pp. 141485–141499, 2023. https://doi.org/10.1109/ACCESS.2023.3343126

[4] S. Anantharajan, S. Gunasekaran, T. Subramanian, and R. Venkatesh, "MRI brain tumor detection using deep learning and machine learning approaches," *Measurement: Sensors*, vol. 31, p. 101026, 2024. https://doi.org/10.1016/j.measen.2024.101026

[5] 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. 7232, 2024. https://doi.org/101038/s41598-024-57970-7

[6] T. Muezzinoglu *et al.*, "PatchResNet: Multiple patch division–based deep feature fusion framework for brain tumor classification using MRI images," *Journal of Digital Imaging*, vol. 36, no. 3, pp. 973–987, 2023. https://doi.org/10.1007/s10278-023-00789-x

[7] 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. https://doi.org/10.1109/ACCESS.2024.3379136

[8] A. Nag *et al.*, "TumorGANet: A transfer learning and generative adversarial network-based data augmentation model for brain tumor classification," *IEEE Access*, vol. 12, pp. 103060–103081, 2024. https://doi.org/10.1109/ACCESS.2024.3429633

[9] A. Sarkar, M. Maniruzzaman, M. A. Alahe, and M. Ahmad, "An effective and novel approach for brain tumor classification using AlexNet CNN feature extractor and multiple eminent machine learning classifiers in MRIs," *Journal of Sensors*, vol. 2023, no. 1, 2023. https://doi.org/10.1155/2023/1224619

[10] 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.

[11] J. Liu *et al.*, "IOUC-3DSFCNN: Segmentation of brain tumors via IOU constraint 3D symmetric full convolution network with multimodal auto-context," *Scientific Reports*, vol. 10, no. 6256, 2020. https://doi.org/10.1038/s41598-020-63242-x

[12] 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. https://doi.org/10.3390/rs12172670

[13] 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, p. 1467, 2024. https://doi.org/10.3390/rs16081467

[14] A. G. Hussien, A. Bouaouda, A. Alzaqebah, A. S. Kumar, G. Hu, and H. Jia, "An in-depth survey of the artificial gorilla troops optimizer: Outcomes, variations, and applications," *Artificial Intelligence Review*, vol. 57, no. 9, p. 246, 2024. https://doi.org/10.1007/s10462-024-10838-8

[15] M. A. El-Dabah, M. H. Hassan, S. Kamel, and H. M. Zawbaa, "Robust parameters tuning of different power system stabilizers using a quantum artificial gorilla troops optimizer," *IEEE Access*, vol. 10, pp. 82560–82579, 2022. https://doi.org/10.1109/ACCESS.2022.3195892

[16] H. Sharma, G. Hazrati, and J. C. Bansal, "Spider monkey optimization algorithm," *Evolutionary and Swarm Intelligence Algorithms*, Springer, Cham, vol. 779, 2019. pp. 43–59. https://doi.org/10.1007/978-3-319-91341-4_4

[17] Center for Biomedical Image Computing & Analytics, "Multimodal brain tumor segmentation challenge 2018," 2018. Accessed on: October 2024, [Online]. Available: https://www.med.upenn.edu/sbia/brats2018/data.html

[18] Center for Biomedical Image Computing & Analytics, "Multimodal brain tumor segmentation challenge 2019," 2019. Accessed on: October 2024, [Online]. Available: https://www.med.upenn.edu/cbica/brats2019/data.html

[19] R. S. Devi, V. R. Kumar, and P. Sivakumar, "EfficientNetV2 model for plant disease classification and pest recognition," *Computer Systems Science and Engineering*, vol. 45, no. 2, pp. 2249–2263, 2023. https://doi.org/10.32604/csse.2023.032231

[20] 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. https://doi.org/10.1016/j.future.2021.06.016

[21] 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. https://doi.org/10.1080/03772063.2021.1919222

[22] I. Singh, G. Goyal, and A. Chandel, "AlexNet architecture-based convolutional neural network for toxic comments classification," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 9, pp. 7547–7558, 2022. https://doi.org/10.1016/j.jksuci.2022.06.007

[23] P. T. Jiang, C. B. Zhang, Q. Hou, M. M. Cheng, and Y. Wei, "LayerCAM: Exploring hierarchical class activation maps for localization," *IEEE Transactions on Image Processing*, vol. 30, pp. 5875–5888, 2021. https://doi.org/10.1109/TIP.2021.3089943

[24] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," *Medical Image Analysis*, vol. 43, pp. 98–111, 2018. https://doi.org/10.1016/j.media.2017.10.002

8 AUTHORS

Sayeedakhanum Pathan is a research scholar at the Department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Hyderabad, India. She earned her Bachelor's degree in Computer Science & Engineering in 2009 and her Master's degree in Computer Science & Engineering in 2015. Currently, she is pursuing Ph.D. in Computer Science, with a focus on machine learning and deep learning. Her research interests include deep learning, data science, machine learning. Sayeeda has contributed significantly to her field by publishing numerous

research articles in esteemed journals and conferences (E-mail: pathan.sayeeda@klh.edu.in).

Dr. Savadam Balaji working as Associate Dean Academics, KL University, Hyderabad Campus, is a technocrat with exposure to Industry, Academics and Research. His education is MTech and PhD in cyber forensics, with specific research interest in the field of Cyber Security and IOT. Having more than 28 years of teaching experience in various Engineering Colleges. He is also Principal investigator for AICTE MODROBS project "Design of Embedded and VLSI Excellence Centre", worth Rs. 20 lakhs. Completed many Consultancy Projects and has organized many Conferences, Workshops and FDPs. Presently Guiding eight PhD Scholars under K L University and M.Tech Students in JNTUH and KLU. Published 30+ research papers in many reputed journals, Conferences, one patent and 3 books. Recipient of two prestigious Awards "Education Excellence Award-2020" and "Dr. A.P. J Abdul Kalam Pillars of India-22 Award" (E-mail: balajis@klh.edu.in).

S. Sai Anuraghav is currently a Teaching Assistant at San Jose State University USA and a seasoned software engineer with over 3 years of experience at Tata Consultancy Services in Hyderabad, India. He focuses on using Python, AI/ML, Data Science, SQL, and databases to promote innovation and solve complicated problems. His work ethic is that discipline fosters success, allowing him to constantly improve his abilities, stay goal-oriented, and accept new learning chances. He holds Google's TensorFlow Developer certification and Hackerrank's Intermediate Problem Solving certification. Earning a world rank of 75 on Hackerearth in programming/problem solving, solving the top 500 tasks on LeetCode, a 1900+ competitive rating with knight status, and six study plan badges (E-mail: saianuraghav.savadam@sjsu.edu).