© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MULTI HEAD ATTENTION-BASED LSTM AND GRADIENT-WEIGHTED CLASS ACTIVATION MAPPING FOR BRAIN TUMOR DETECTION USING MRI IMAGES

PRATHIMA DEVADAS[1*] and G. MATHIVANAN[2]

[1,2]Department of Computer science and Engineering
[1,2]Sathyabama Institute of Science and Technology
Chennai, India
[1]prathima.dev20@outlook.com
[2]mathivanan.it@sathyabama.ac.in
*Corresponding author: Prathima Devadas

ABSTRACT

The presented of attention-based architectures in medical imaging has ushered in a novel era of precision diagnostics, mainly for the identification and classification of brain tumors. This research developing a novel knowledge distillation method that employs a tripartite attention mechanism within transformer encoder models to identify various brain tumor types utilizing magnetic resonance imaging (MRI). This study offerings a unique method for brain tumor identification, integrating Multi-Head Attention-based Long Short-Term Memory (MHA-LSTM) networks with Gradient-weighted Class Activation Mapping (Grad-CAM). The MHA-LSTM design utilizes multi-head attention to capture complex spatial-temporal relationships across consecutive MRI slices, enabling the model to focus on the most critical features. Grad-CAM is incorporated to provide visual explanations by highlighting key regions contributing to the model's predictions, improving both interpretability and clinical relevance. Experimental results demonstrate that the suggested technique surpasses conventional LSTM models in terms of accuracy, sensitivity, and specificity. Moreover, the Grad-CAM visualizations offer valuable insights into the model's decision-making process, fostering better understanding and facilitating future clinical validation. This method presented a robust and interpretable solution for brain tumor identification, advancing the application of deep learning in medical imaging.

Keywords: *Brain Tumor Detection, Deep Learning, Grad-CAM, Interpretability, Long Short-Term Memory, Magnetic resonance imaging, Medical Imaging, Multi-Head Attention, Neural Networks, Spatiotemporal Modeling.*

1. INTRODUCTION

The human brain, which contains numerous complex neural networks, is the most innovative organ in the body. These neurons are interconnected and work together in response to instructions since other regions of the body. Brain tumors increase when these neurons begin to function abnormally. In 2018, approximately 9.6 million people died since brain tumors, making them the second leading cause of death worldwide. Brain tumors are responsible for approximately one in every ten deaths globally [1]. Brain tumours come in two varieties: primary and secondary. Additionally, brain tumours can be classified as either healthy (non-cancerous) or malignant (cancerous). Benign tumours hardly ever spread throughout the body, whereas malignant tumours can kill people. Examples of brain cancers include pituitary tumors, gliomas, and meningiomas [2]. The most frequent primary brain tumour between them is glioma. These malignancies originate in the brain's glial cells and are classified into four grades, which indicate their aggressiveness.

Slow growth and limited dissemination are characteristics of meningiomas, which are healthy tumours that originate since the meninges tissue layer. On the other hand, pituitary tumours are less frequent and usually benign, developing on the pituitary gland [3]. The symptoms that manifest depend on the location, size, and shape of brain tumours. Brain tumours frequently cause headaches, seizures, nausea, vomiting, and trouble thinking or speaking. The majority of techniques for diagnosing brain malignancies combine biopsies with imaging tests like CT and MRI scans, depending on the tumor's kind and stage. Chemotherapy, radiation therapy, and surgery are additional therapeutic options [4]. The purpose of treatment is to reduce symptoms, manage tumor growth, and increase survival. Regardless of whether specific risk factors have been recognized, brain tumors can affect people of all ages and genders. Ionising radiation exposure, genetic susceptibility, and several hereditary disorders are risk factors. Though, the specific etiology of most brain tumors remains unknown [5]. Additionally, advancements in medical imaging and diagnostic techniques have increased the chance of a successful intervention by enabling the earlier detection of brain tumours.

In this research highlights that MRI image-based brain tumor diagnosis is a critical medical image task requiring precise and interpretable models. This effort proposes a unique method that combines an MH-LSTM of improved feature extraction with Grad-CAM for comprehension. Combining these methods advances automated tumour detection processes by growing diagnostic accuracy and providing that visual observations to the decision-making process.

1.1 Motivation of this research

The motivation for presented a "Multi-head Attention-based LSTM and Gradient-weighted Class Activation Mapping of Brain Tumor Detection utilizing MRI Images" stems since the pressing essential for accurate, effectual, and interpretable diagnosis tools in medicinal imaging. Brain tumors pose important challenges due to their complexity and variability, making early diagnosis critical for effective treatment. By integrating multi-head attention mechanisms with LSTM and employing Grad-CAM for visualization, the suggested method goals to enhance feature extraction, develop classification accuracy, and presents interpretable insights into model decisions, thereby supporting clinicians in creation additional informed diagnoses.

1.2 Contribution of the research

- To improve a unique method for brain tumor identification that integrates MHA-LSTM networks with Grad-CAM.
- The MHA-LSTM design uses multi-head attention to capture complex spatial-temporal relationships across successive MRI slices, enabling the technique to focus on the most important features.
- Lastly, experimental findings determine that the suggested technique outclasses standard LSTM models in words of accuracy, sensitivity, and specificity.

The remainder of this work is divided into the following sections: The literature is reviewed in Section 2, the suggested methods are explained in Section 3, the results and discussion are proposed in Section 4, and future research directions are concluded in Section 5.

2. SURVEY

This survey explores advancements in brain tumor diagnosis using MRI images, focusing on innovative techniques and methodologies that enhance accuracy and efficiency. By analyzing recent improvements in imaging technologies and ML algorithms, this research goals to present a complete instant of state-of-the-art methodologies for the promptly and precise diagnosis for brain tumors.

Several ML and DL methods, including as Naive Bayes, Multi-Layer Perceptrons (MLP), Random Forest, Support Vector Machines (SVM), and k-Nearest Neighbours (k-NN), were presented by Kumar et al. [6] for the recognition and segmentation of brain tumours. Remarkably, traditional SVM succeeded the highest classification accuracy of 92.4%.

The authors Khan et al. [7] offer a DL technique for the classification for brain tumours by analysing MRI data, which might be useful for medical professionals. Preprocessing, k-means clustering for brain tumour segmentation, and MRI data classification by benign/malignant using a fine-tuned VGG19 model are the three primary components of the suggested approach. Furthermore, the idea of synthetic data intensification is presented to increase the quantity of accessible data for classifier training, leading to greater classification accuracy. The suggested method was tested in a controlled environment utilizing the BraTS 2015 benchmark data sets. The outcomes establish that the suggested approach is additional accurate and more effective when associated to the previously published state-of-the-art methods.

Rasool et al. [8] utilise CNN and deep learning to classify data in two separate ways. The first approach uses an unsupervised support vector machine (SVM) for pattern organization and a pre-trained convolutional neural network (CNN) (like SqueezeNet) for feature extraction. In the second method, a refined SqueezeNet is integrated into the supervised soft-max classifier. To estimate the efficiency of the proposed technique, brain MRIs were used to analyse 396 normal brain images, 926 pituitary cancer images, 1937 glioma tumour images, and 926 meningioma tumour images. The optimised SqueezeNet model obtained a 96.5% accuracy rate, according to the experiment data. However, when SqueezeNet was used as a feature extractor in combination with an SVM classifier, the recognition accuracy increased to 98.7%.

Almalki et al. [9] employed MRI in conjunction with a ML strategy to rapidly identify the severity of brain tumours (glioma, meningioma, no tumour, pituitary) using MRI. MRI Gaussian and nonlinear scale features are widely employed to recover images with typical imaging characteristics, including texture, local binary patterns, orientation gradient histograms, etc., since they can tolerate rotation, scaling, and noise problems. In order to get fine details, each magnetic resonance imaging (MRI) scan is converted into several tiny 8 × 8-pixel MR pictures for the features. By selecting the most robust features according to variance, they can then divide them into four hundred features with a Gaussian and four hundred with a nonlinear scale. They then hybridise these features with each MRI to address memory concerns. At last, the proposed hybrid feature vector's performance is tested utilizing traditional ML classifiers. To ensure the suggested method is accurate, it is tested using a publicly accessible catalogue of brain MRI images. From what they can see, the model that was trained using support vector machines had the best classification accuracy of 95.33% while using very little computing time.

A multi-level attention mechanism is proposed by Shaik et al. [10] in their work on the recognition of brain tumours. Spatial and cross-channel attention are combined in the proposed multi-level attention network (MANet) to prioritise the cancer region and monitor the cross-channel temporal correlations in the semantic feature sequence obtained since the Xception backbone. The BraTS and Figshare benchmark datasets are utilized to assess the suggested technique's presentation. According to the experimental results, merging cross-channel attention with spatial attention enhances generalisability and yields better performance with fewer model parameters. They fared better than many other models in the cancer recognition challenge utilising our proposed MANet, which had an optimum accuracy of 96.51% on Figshare and 94.91% on BraTS'2018 datasets.

2.1 Research Gap

Despite improvements in MRI image-based brain tumour diagnosis, high accuracy and reliability are still difficult to achieve since tumour shapes, sizes, and locations vary, and more reliable algorithms that perform well over a variety of datasets are required. This highlights the need for improved methods integrating advanced AI techniques to enhance diagnostic precision and clinical utility.

2.2 Limitation of Existing System

- Many existing systems lack high accuracy, particularly in differentiating amongst tumor varieties and results, which can lead to potential misdiagnosis.
- The presentation of current models seriously depends on high-quality, labeled MRI datasets, which are often limited and lack universal standardization.
- Deep learning systems are prone to overfitting to training data, diminishing their capability to generalize efficiently to new or unseen cases.

3. PROPOSED SYSTEM

In this proposed section, to described a unique technique for brain tumor identification is described, integrating MHA-LSTM networks with Grad-CAM. The MHA-LSTM design employs multi-head attention to capture complex spatial-temporal relationships across successive MRI slices, enabling the method to focus on the most important features. Grad-CAM is incorporated to provide visual explanations by highlighting key regions contributing to the model's predictions, thereby enhancing both interpretability and clinical usefulness. Figure 1 illustrates the block diagram.

Journal of Theoretical and Applied Information Technology
31st May 2025. Vol.103. No.10
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

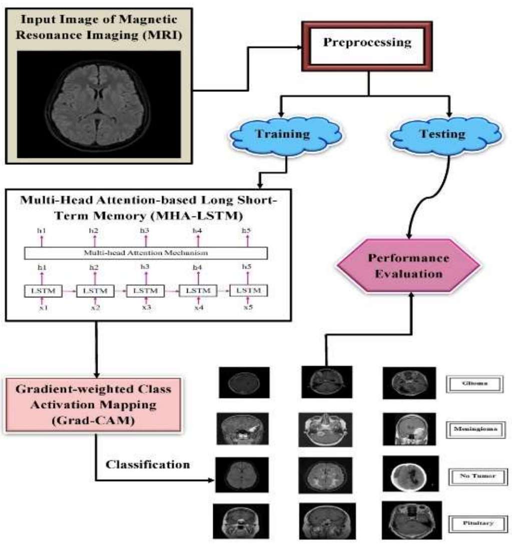

Figure 1: Architecture of block diagram

3.1. Magnetic Resonance Imaging (MRI) Dataset

This dataset contains 1,311 carefully selected high-resolution brain MRI scans, used to classify and identify brain tumors. The classifications for the MRI scans include "pituitary," "glioma," "meningioma," and "no tumor." Particularly useful tools for creating and assessing ML methods for the automated diagnosis and categorisation of brain tumors are convolutional neural networks (CNNs) [11]. Meningioma (306 images), Pituitary (300 images), Glioma (300 images), and No Tumor (405 images) are the categories into which the 1,311 photographs are separated. Figure 2 displays some pictures from the Brain MRI dataset. Figure 3 shows the distribution of MRI Scans of brain tumor classification.

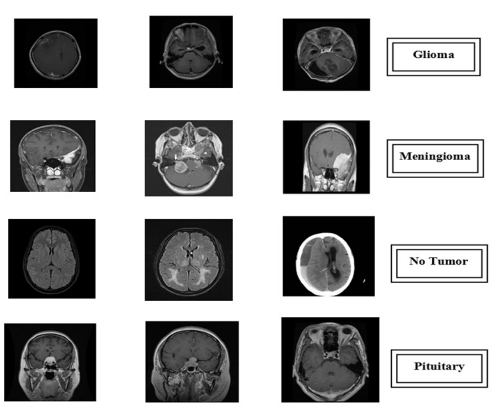

Figure 2: Brain MRI dataset images

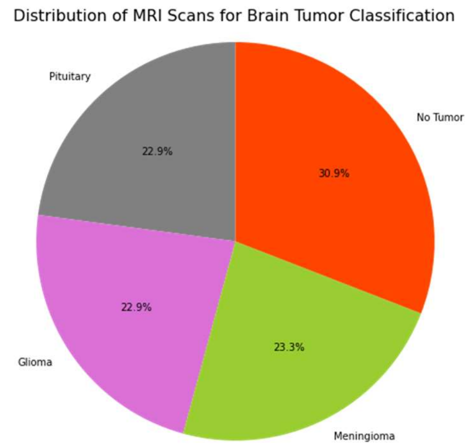

Figure 3: Distribution of MRI Scans for brain tumor classification

3.2 Preprocessing

The training and testing sets of our dataset were separated during the data preprocessing workflow. Using PyTorch's data loaders, we employed batch processing with a total batch size of 32 to facilitate effective model training. As an organizational step, we resized each image to a fixed dimension of 64x64 pixels to ensure consistency in image proportions. To improve the model's presentation, we incorporated data augmentation methods [12]. Specifically, with a probability of 0.5, a significant transformation was applied using PyTorch's Random Horizontal Flip function. Through a ability to learn since together the original and horizontally flipped photos, this augmentation enhanced the variety of the dataset and enhanced the model's capacity for generalisation. During training, the pictures were randomly flipped horizontally with a 50% chance. All of these preprocessing procedures worked together to improve our brain tumour detection model's accuracy.

3.3 Multi-Head Attention-based Long Short-Term Memory

The MHA-LSTM approach associations the strengths of LSTM networks and Multi-Head Attention (MHA) to effectively model spatial and temporal dependencies in MRI images [13]. Figure 4 shows the architecture of MHA-LSTM.

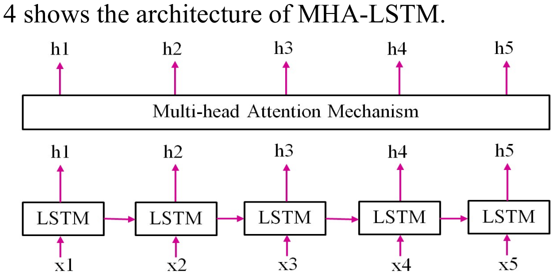

Figure 4: MHA-LSTM architecture

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

It goals to increase the organization and diagnosis for brain tumors by leveraging the following features:

1. LSTM Component: Long-term dependencies can be learned using LSTMs, a type of RNN. In the context of MRI images, LSTMs process the extracted spatial features sequentially, capturing contextual and temporal information across slices or regions of interest within the images.

2. Multi-Head Attention Component: The method can focus on multiple input segments simultaneously due to the MHA mechanism, which was integrated into the Transformer design. This mechanism enhances the model's ability to identify critical features by assigning different attention weights to various regions or features of the MRI images.

3. Feature Fusion: The MHA module operates on the features extracted since the MRI images, emphasizing the most relevant ones. The fused features are then passed to the LSTM to learn temporal relationships. This combined architecture ensures that the model captures both spatial relationships (via attention) and temporal dependencies (via LSTM).

4. Brain Tumor Detection: After MHA-LSTM processing, a fully connected layer and a softmax classifier are utilized to categorise brain tumours and identify their type or presence.

Multi-Head Attention:

Let the input feature sequence from the MRI image be $X \in \mathbb{R}^{n\times d}$, where n is the sequence length (e.g., slices or patches), and d is the feature dimension.

$$Attention(Q,K,V) = soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

Where $Q = XW_Q, K = XW_K, V = XW_V$, $W_Q, W_K, W_V \in \mathbb{R}^{d\times d_k}$

Multi-Head Attention: Combine multiple attention heads:

$$MHA(X) = Concat(head_1, head_2, ..., head_h)W_O \quad (2)$$

Where $head_i = Attention(Q,K,V)$, $W_O \in \mathbb{R}^{hd_k\times d}$ is a learned projection matrix.

LSTM:

The output of the MHA is fed into an LSTM for sequential modeling:

LSTM State Update:

$$\begin{aligned}
f_t &= \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right) \\
i_t &= \sigma\left(W_i h_{t-1} + U_i x_t + b_i\right) \\
o_t &= \sigma\left(W_o h_{t-1} + U_o x_t + b_o\right) \\
\tilde{C}_t &= \tanh\left(W_c h_{t-1} + U_c x_t + b_c\right) \\
C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\
h_t &= o_t \odot \tanh\left(C_t\right)
\end{aligned} \quad (3)$$

Wherever f_t, i_t, o_t are forget, input, and output gates, C_t is the cell state, and h_t is the hidden state, x_t is the input at time t, and W, U, b are learned parameters.

Combined MHA-LSTM Output:

The output from the LSTM is transferred to a fully connected layer.

$$y = soft\max\left(W_h h_t + b_h\right) \quad (4)$$

Where y is the predicted probability distribution for brain tumor classes.

3.4 Grad-CAM

The well-liked explainable AI (XAI) technique Grad-CAM develops the interpretability of DL models in the field of medical imaging, primarily in MRI-based brain tumour diagnosis [14]. Grad-CAM gives doctors the ability to determine that areas for the input image are greatest crucial to the methods prediction, whether the model is focused on tumour identification or classification by importance.

The results of the target class that enter a convolutional neural network's (CNN) final convolutional layer are utilized to accomplish this. These gradients generate a coarse localization map, which emphasizes critical areas in the MRI scan. Grad-CAM is especially valuable in the medical field as it provides insights into model decisions, fostering trust and aiding in the diagnostic process.

The Grad-CAM algorithm computes the class-specific importance weights $\left(\alpha_k^c\right)$ for the feature maps of the final convolutional layer $\left(A_k\right)$ by global average pooling the gradients of the output class score $\left(y^c\right)$ with deference to the feature map activations:

$$\alpha_k^c = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A_k^{i,j}} \quad (5)$$

Here, Z is the total number of spatial locations (i,j) in the feature map. The class activation map $L_{Grad-CAM}^c$ is then computed as:

$$\left(L_{Grad-CAM}^c\right) = \mathrm{Re}\,LU\left(\sum_k \alpha_k^c A_k\right) \quad (6)$$

The ReLU ensures that only the features positively correlated with the target class are visualized. This heatmap, overlaid on the original MRI image, effectively aids in identifying tumor regions.

Advantages of proposed method

The proposed technique for combining Multi-Head Attention-based LSTM and Grad-CAM offers significant advantages of brain tumor diagnosis in MRI images. The Multi-Head Attention mechanism enhances feature extraction by focusing on critical regions, thereby improving the LSTM's ability to capture temporal and spatial dependencies. Grad-CAM further aids interpretability by provided that visual clarifications of the model's predictions, allowing clinicians to identify tumor regions with greater confidence and trust. This method ensures robust detection while fostering transparency in medical AI requests.

4. RESULT AND DISCUSSION

4.1 Experimental setup

The CNN was developing using Python on a computer equipped with an 8GB GTX 1060 GPU, 32GB of RAM, and a 12th-generation i9 processor. This configuration was used to forecast brain tumors.

4.2 Performance Metrics

Positive Predictive Value (PPV) is a metric utilized to calculate how well a model predicts favorable events. It signifies the proportion of true positive forecasts between all the model's favorable forecasts.

$$PPV = \frac{True\ Positivies}{True\ Positives + False\ Positives} \quad (7)$$

Negative Predictive Value (NPV) is a metric utilized to assess the probability that a person receiving a negative test result truthfully does not have the disease.

$$NPV = \frac{True\ Negatives}{True\ Negatives + False\ Negatives} \quad (8)$$

Accuracy is a measure that estimates the percentage of accurate forecasts to whole forecasts, thereby assessing the whole correctness of a model.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ \mathrm{Pr}\,edictions} \quad (9)$$

Execution time refers to the total time taken by a model or procedure to process input data and generate an output. It is an important metric for calculating the efficiency of a system.

$$Execution\ Time = End\ Time - Start\ Time \quad (10)$$

4.3 Comparative Methods

Parallel Deep Convolutional Neural Network (PDCNN) [15]: A new PDCNN topology that utilizes batch standardization and dropout regularization to address overfitting while extracting both local and global information.

Convolutional Neural Network (CNN) [16]: CNNs have been employed to address these issues, as they have proven highly effective in medical images processing. To maximize the identification for brain tumors, several CNN architectures were implemented and evaluated.

Convolutional Neural Network- Particle Swarm Optimization (CNN-PSO) [17]: Our approach employs the PSO procedure to identify the CNN hyperparameter configuration that performs best. These optimized limitations are then used to classify the CNN architectures.

4.3.1 PPV Analysis

Table 1: PPV Analysis for proposed method

Types	PDCNN	CNN	CNN-PSO	Proposed
Pituitary	76.65	87.55	89.44	93.54
Meningioma	81.87	83.56	81.11	94.17
Glioma	66.18	81.91	84.23	94.91
No Tumor	79.27	92.17	91.87	95.67

The PPV analysis presented in Figure 5 and Table 1 demonstrates a clear advantage of the proposed model across all tumor types. The PPV for

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the suggested technique surpasses that of other models (PDCNN, CNN, and CNN-PSO) in detecting Pituitary, Glioma, Meningioma, and No Tumor. For instance, the suggested method achieves a PPV of 93.54% for Pituitary tumors, compared to 76.65%, 87.55%, and 89.44% for PDCNN, CNN, and CNN-PSO, respectively. Similar developments are observed for other tumor types, highlighting the suggested model's superior accuracy in positive predictions and its improved diagnostic reliability.

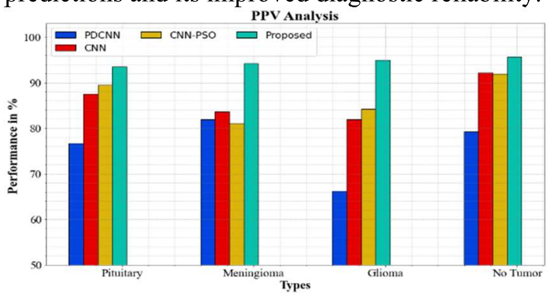

Figure 5: PPV Analysis for proposed method

4.3.2 Negative Predictive Value (NPV) Analysis

Table 2: NPV Analysis for proposed method

Types	PDCNN	CNN	CNN-PSO	Proposed
Pituitary	65.45	88.16	87.43	95.16
Meningioma	87.23	87.34	91.11	96.34
Glioma	85.45	80.33	94.34	95.91
No Tumor	81.23	78.45	79.34	96.77

The Negative Predictive Value (NPV) analysis offered in Figure 6 and Table 2 establishes that the suggested method outclasses other approaches (PDCNN, CNN, and CNN-PSO) in accurately identifying negative cases (No Tumor). For all tumor categories, containing Pituitary, Glioma, Meningioma, and No Tumor, the suggested approach persistently succeeds the highest NPV. For example, the suggested model's NPV for Pituitary tumor diagnosis is 95.16%, significantly higher than that of PDCNN (65.45%) and other models, underscoring its superior competence to correctly detect non-tumor cases.

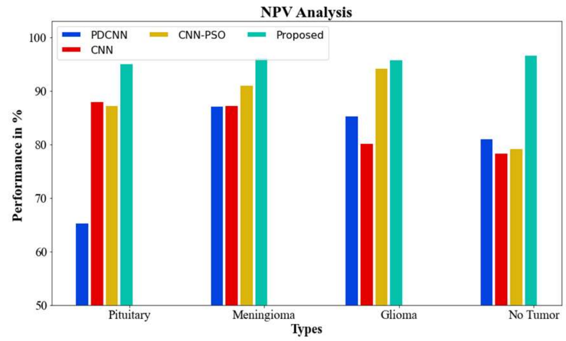

Figure 6: NPV Analysis for proposed method

4.3.3 Accuracy Analysis

Table 3: Accuracy Analysis for proposed method

Types	PDCNN	CNN	CNN-PSO	Proposed
Pituitary	75.46	84.34	72.34	93.44
Meningioma	81.11	85.12	82.26	96.87
Glioma	84.23	89.44	82.92	95.92
No Tumor	89.25	92.33	91.54	98.91

The accuracy analysis in Figure 7 and Table 3 show that the suggested technique outclasses PDCNN, CNN, and CNN-PSO across all tumor types. For Pituitary, Meningioma, Glioma, and No Tumor, the suggested technique succeeds the highest accuracy, with notable enhancements such as 93.44% for Pituitary and 98.91% for No Tumor. This identifies that the suggested model provides superior organization presentation, making it additional reliable for brain tumor detection.

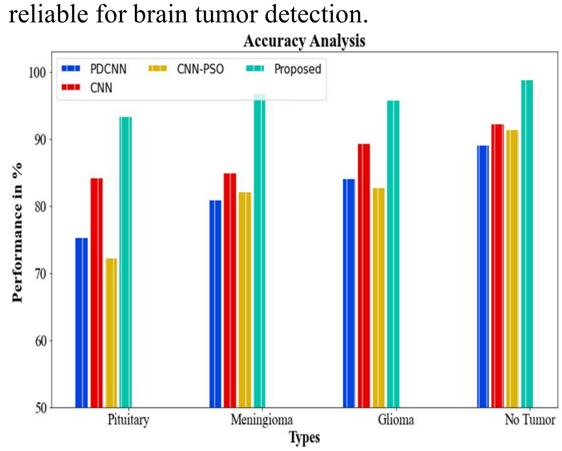

Figure 7: Accuracy Analysis for proposed method

4.3.4 Execution Time Analysis

Table 4: Execution Time Analysis for proposed method

Types	PDCNN	CNN	CNN-PSO	Proposed
Pituitary	14.667	10.145	7.345	1.765

© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Meningioma	13.876	11.332	9.221	3.443
Glioma	16.987	13.456	8.861	2.876
No Tumor	17.876	14.876	8.991	5.567

The study presented in Figure 8 and Table 4 on Execution Time establishes that the proposed model significantly outclasses PDCNN, CNN, and CNN-PSO in terms of processing speed. Across all tumor categories, including Pituitary, Meningioma, Glioma, and No Tumor, the suggested method constantly achieves the lowest execution times. For example, in Pituitary tumor detection, the proposed method completes the task in just 1.765 seconds, compared to 14.667 seconds for PDCNN. This highlights the efficiency of the suggested technique, creation it highly suitable for faster, real-time requests.

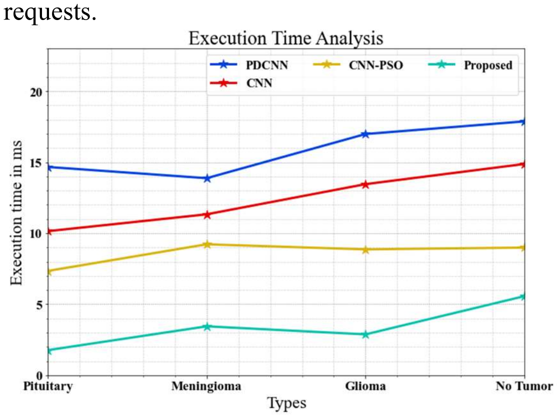

Figure 8: Execution Time Analysis for proposed method

4.4 Training and Testing Validation

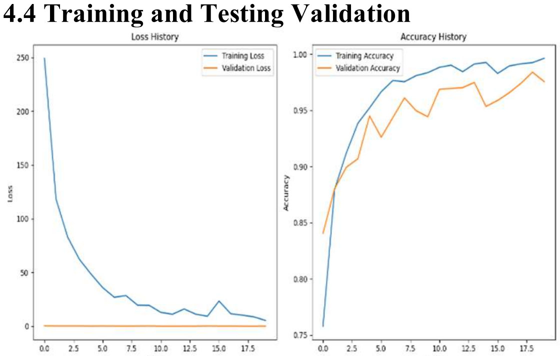

Figure 9: Training and testing validation

Figure 9 demonstrates that the brain tumor detection models are calculated based on training-validation loss and training-validation accuracy to assess their presentation. A model with high training and validation accuracy can correctly classify various types of brain tumours, while one with low training and validation loss shows that the model is learning effectively without overfitting. Maintaining a balance amongst these metrics ensures the model's reliability and generalization for real-world applications.

4.5 Ablation study

The ablation study evaluates the efficiency of individual mechanisms in the suggested model, of brain tumor detection utilizing MRI images. This comparative analysis examines three configurations: the standard LSTM model, MHA-LSTM, and MHA-LSTM combined with Grad-CAM. The results reveal a reliable improvement in performance as additional components are integrated into the architecture. The standard LSTM model succeeds an average accuracy of 85%, serving as a strong baseline. Incorporating Multi-head Attention into the LSTM architecture raises the accuracy to 95.46%, showcasing the competence of attention mechanisms to increase feature extraction by focusing on relevant spatial-temporal regions in MRI sequences. Further incorporation of Grad-CAM boosts the model's interpretability and classification presentation, achieving an accuracy of 98.91%. This growth underscores Grad-CAM's role in providing localized explanations of decision-making though also guiding the network to focus on critical tumor regions during training. This ablation study validates the synergistic effect of joining attention mechanisms and explainability tools in the proposed model, making it a robust and interpretable solution of brain tumor diagnosis in MRI scans.

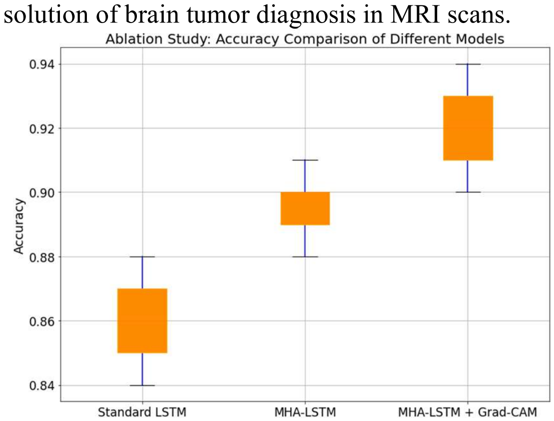

Figure 10: Ablation study accuracy comparison of different models

Table 5: Comparison of the proposed method with other models

Authors	Classifier	Year	Accuracy
Zeineldin et al. [19]	Multimodal CNN	2022	94.2%
Xue et al. [20]	Multimodal CNN	2022	93.0%
Zahoor et al. [18]	Res-BRNet (CNN)	2024	98.22%
Our Model	MHA-LSTM	-	98.91%

© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The suggested method, using the MHA-LSTM architecture, establishes superior presentation in brain tumor classification associated to existing models. In contrast to Zeineldin et al. (2022) and Xue et al. (2022), who reported accuracies of 94.2% and 93.0%, respectively, using multimodal CNNs, Zahoor et al. (2024) succeeded an accuracy of 98.22% using the Res-BRNet (CNN), according to the comparative table. In contrast, our model achieves an outstanding accuracy of 98.91%, surpassing all these methods. Figure 11 and Table 5 illustrate how effectively the MHA-LSTM captures intricate spatiotemporal correlations, creation it a viable option for precise brain tumor diagnosis.

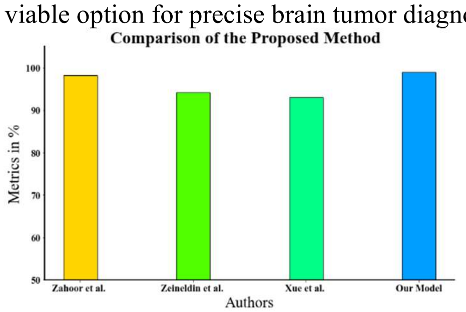

Figure 11: Comparison of the proposed method with other models

Limitations

Though the Multi-Head Attention-based LSTM (MHA-LSTM) joined with Grad-CAM offers significant growths in brain tumor detection, it has certain limitations. First, the method's presentation heavily depends on the quantity and constancy of MRI data, with any noise or artifacts in the images potentially affecting accuracy. Furthermore, the computational complication of the MHA-LSTM model, mainly with multiple attention heads and Grad-CAM visualizations, can lead to longer training and inference times connected to simpler models. Lastly, although Grad-CAM develops interpretability, it may not always offerings perfect or clinically actionable explanations, necessitating further refinement for real-world medical requirements.

5. CONCLUSION AND FUTURE WORK

In conclusion, the suggested Multi-Head Attention-based LSTM, combined with Grad-CAM, presents a robust and interpretable technique of brain tumor diagnosis utilizing MRI images. The integration of multi-head attention increases the model's capability to capture complex spatial and contextual relations within MRI scans, while the LSTM component efficiently models sequential dependencies in the data. Grad-CAM further aids in visualizing and interpreting the decision-making process, offering valuable insights for clinical applications. The experimental results determine the methods superior presentation in words of accuracy and reliability, making it a promising tool of assisting radiologists in the early and precise detection for brain tumors. Future work could explore extending this framework to multiclass tumor organization and incorporating additional modalities of medical imaging to develop diagnostic abilities.

DECLARATION

Data Availability: Data will be made available on request.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

Ethical Approval: The declaration is "Not Applicable"

REFERENCE

[1]. T.A. Sadoon, M.H. Ali, "Deep learning model for glioma, meningioma and pituitary classification," *Int. J. Adv. Appl. Sci*, pp. 2252, 8814. https://doi.org/10.11591/ijaas.v10.i1.pp88-98

[2]. O.R. Devi, C.S. Bindu, E.S. Kumar, (November 2022). "Identification and Classification of Brain Tumor from MRI Using Transfer Learning Approach," *In Proceedings of the 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS),* Greater Noida, India, 4–5. https://doi.org/10.1109/ICCCIS56430.2022.10037223

[3]. P. Nagaraj, V. Muneeswaran, L.V. Reddy, P. Upendra, M.V. Reddy, (May 2020). "Programmed multi-classification of brain tumor images using deep neural network," *In Proceedings of the 2020 4th international conference on intelligent computing and control systems (ICICCS),* Madurai, India, pp. 13–15. https://doi.org/10.1109/ICICCS48265.2020.9121016

[4]. L.B. Jans, M. Chen, D. Elewaut, F.V.D. Bosch, P. Carron, P. Jacques, R. Wittoek, J.L. Jaremko, N. Herregods, (2021). "MRI-based synthetic CT in the detection of structural lesions in patients with suspected sacroiliitis: Comparison with MRI," *Radiology*, vol. 298, pp. 343–349. https://doi.org/10.1148/radiol.2020201537

[5]. I. Wahlang, A.K. Maji, G. Saha, P. Chakrabarti, M. Jasinski, Z. Leonowicz, E. Jasinska, (2022). "Brain magnetic resonance imaging classification using deep learning architectures

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

with gender and age," *Sensors*, vol. 22, pp. 1766. https://doi.org/10.3390/s22051766

[6]. G. Kumar, P. Kumar, D. Kumar, (December 2021). "Brain tumor detection using convolutional neural network," *In Proceedings of the International Conference on Mobile Networks and Wireless Communications (ICMNWC),* Tumkur, Karnataka, India, 3–4, pp. 1–6. https://doi.org/10.1109/ICMNWC52512.2021.9688460

[7]. A.R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, Z. Mehmood, "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification," *Microsc. Res Tech,* vol. 84, pp. 1389–1399, 2021. https://doi.org/10.1002/jemt.23694

[8]. M. Rasool, N.A. Ismail, A. Al-Dhaqm, W.M.S. Yafooz, A. Alsaeedi, (2022). "A novel approach for classifying brain tumours combining a squeezenet model with svm and fine-tuning," *Electronics,* vol. 12, pp. 149. https://doi.org/10.3390/electronics12010149

[9]. Y.E. Almalki, M.U. Ali, W. Ahmed, K.D. Kallu, A. Zafar, S.K. Alduraibi, M. Irfan, M.A.A. Basha, H.A. Alshamrani, A.K. Alduraibi, (2022). "Robust Gaussian and Nonlinear Hybrid Invariant Clustered Features Aided Approach for Speeded Brain Tumor Diagnosis," *Life,* vol. 12, pp. 1084. https://doi.org/10.3390/life12071084

[10]. N.S. Shaik, T.K. Cherukuri, "Multi-level attention network: Application to brain tumor classification," *Signal Image Video Process,* vol. 16, pp. 817–824, 2022. https://doi.org/10.1007/s11760-021-02022-0

[11]. https://www.kaggle.com/datasets/shreyag1103/brain-mri-scans-for-brain-tumor-classification

[12]. B. B. Gupta, A. Gaurav, & V. Arya, (2024). "Deep CNN based brain tumor detection in intelligent systems," *International Journal of Intelligent Networks,* vol. 5, pp. 30-37. https://doi.org/10.1016/j.ijin.2023.12.001

[13]. F. Hasan, & H. Huang, (2023). "MALS-Net: A multi-head attention-based LSTM sequence-to-sequence network for socio-temporal interaction modelling and trajectory prediction," *Sensors,* vol. 23, no. 1, pp. 530. https://doi.org/10.3390/s23010530

[14]. T. R, M., V, V. K., & S. Guluwadi, (2024). "Enhancing brain tumor detection in MRI images through explainable AI using Grad-CAM with Resnet 50," *BMC Medical Imaging*, vol. 24, no. 1, pp. 107. https://doi.org/10.1186/s12880-024-01292-7

[15]. T. Rahmanand, M.S. Islam, (Apr. 2023)."MRIbraintumordetectionandclassification using parallel deep convolutional neural networks Measurement," *Sensors,* vol. 26, Art. no. 100694. https://doi.org/10.1016/j.measen.2023.100694

[16]. R. Martínez-Del-Río-Ortega, J. Civit-Masot, F. Luna-Perejón, & M. Domínguez-Morales, (2024). "Brain Tumor Detection Using Magnetic Resonance Imaging and Convolutional Neural Networks. Big Data and Cognitive Computing," vol. 8, no. 9, pp. 123. https://doi.org/10.3390/bdcc8090123

[17]. R. Ibrahim, R. Ghnemat, & Q. Abu Al-Haija, (2023). "Improving Alzheimer's disease and brain tumor detection using deep learning with particle swarm optimization," *AI,* vol. 4, no. 3, pp. 551-573. https://doi.org/10.3390/ai4030030

[18]. M.M. Zahoor, S.H. Khan, T.J. Alahmadi, T. Alsahfi, A.S.A. Mazroa, H.A. Sakr, S. Alqahtani, A. Albanyan, B.K. Alshemaimri, (2024). "Brain tumor MRI classification using a novel deep residual and regional CNN," *Biomedicines,* vol. 12, pp. 1395. https://doi.org/10.3390/biomedicines12071395

[19]. R.A. Zeineldin, M.E. Karar, O. Burgert, F. Mathis-Ullrich, "Multimodal CNN networks for brain tumor segmentation in MRI: A BraTS 2022 challenge solution," (September 2022). *In Proceedings of the International MICCAI Brainlesion Workshop,* Singapore, 18 Springer: Berlin, Germany, 2022; pp. 127–137. https://doi.org/10.1007/978-3-031-33842-7_11

[20]. J. Xue, Y. Yao, Y. Teng, (2022). "Multi-modal Tumor Segmentation Methods Based on Deep Learning: A Narrative Review," *In Quantitative Imaging in Medicine and Surgery; AME Publishing Company,* Hong Kong, China.