Accurate brain tumor classification with STN-NAM in ResNet50 using MRI

Preeti Sadanand Topannavar[1], Varsha Sachin Bendre[2], Deepti Khurge[2]
[1]School of Electrical and Communication Engineering, Faculty of Science and Technology, JSPM University, Pune, India
[2]Department of Electronics and Telecommunication, Pimpri Chinchwad College of Engineering, Pune, India

Article Info

Article history:

Received May 17, 2024
Revised Oct 8, 2024
Accepted Nov 19, 2024

Keywords:

Magnetic resonance imaging
Non-local attention mechanism
Normalized median filter
Spatial transformer network
U-Net

ABSTRACT

Brain tumor is an abnormal cell growth that contains malignant and benign cells emerging from numerous cell types within brain. Magnetic resonance imaging (MRI) is utilized for brain tumor classification which provides high-resolution images. However, tumors exhibit different characteristics like shape, location, and size which make it challenging to accurately distinguish among different tumor types and accurately classify them. In this research, spatial transformer network and non-local attention mechanism (STN-NAM) is proposed in ResNet50 to accurately classify tumors. STN transforms spatial information while NAM identifies relationships among normal and lesion areas, which together accurately classify tumors. Initially, images are obtained from Figshare, Brats 2019, and Brats 2020 datasets. These images are pre-processed using a normalized median filter (NMF) to reduce salt and pepper noise. Then, normalization is performed to resize original image to a standard size which assists uniformity in image dimension. U-Net is employed to segment tumor regions and STN-NAM is performed to accurately classify tumors. In comparison to the existing techniques namely, multi-level attention network (MANet), mathematical model with 3D attention U-Net, and convolutional neural network (CNN), the STN-NAM achieves superior accuracy of 98.06%, 99.05%, and 98.66% in Figshare, Brats 2019, and Brats 2020 datasets, respectively.

Corresponding Author:

Preeti Sadanand Topannavar
School of Electrical and Communication Engineering, Faculty of Science and Technology
JSPM University
Pune, Maharashtra, India
Email: topannavarp@gmail.com

1. INTRODUCTION

Recently, brain tumors have become one of the most aggressive diseases which results in a very short life span if not detected at an advanced stage [1]. It is split into two common types: primary and secondary tumors. The primary tumors are typically non-cancerous and formed from the cells of human brain. The secondary tumors spread to the brain along the blood flow from other body parts [2]. Brain tumors are categorized into gliomas, pituitary, and meningiomas. Glioma is established in brain tissues, rather than blood vessels and nerve cells [3], [4]. Meningioma grows on the surface of membrane which covers the brain and surrounds the central nervous system, and pituitary form within the skull [5], [6]. Arising primarily in the spinal cord or brain, gliomas are classified into two grades, containing high-grade gliomas (HGG) and low-grade glioma (LGG). HGG is regarded as more penetrative and destructive and is connected with a life expectancy of nearly two years after diagnosis [7]. Brain tumors are determined by using numerous tests that contain computer tomography (CT) scans, biopsies, magnetic resonance imaging (MRI), and positron emission tomography (PET). Among all tests, MRI is the most common for tumor classification [8], [9]. MRI is primarily utilized to detect and classify cancer in the human body and provides visually relevant information. Also, it has a capability to capture various parameters, contrast quality performance on soft issues, and its capability to image scans in various directions [10]. The segmentation and classification of brain tumors by employing MRI is an essential part of medical treatment [11]. This process provides data associated with its anatomical structure for planning treatment. The segmentation of tumors is helpful for brain modeling and for establishing brain atlases [12], [13]. Brain MRIs are determined by radiologists to diagnose brain tumors and the radiologists training consumes a numerous time and has high economic costs. Also, the less of number of radiologists, lack of time, lack of radiologists experience, and fatigue are the negative field factors. Furthermore, it has observed that amount of radiologists analysis is required for the disease detection recently. Deep learning (DL) techniques acquire greater performance with satisfactory reliability in the diagnosis of brain tumors as they generate substantial results in several classification issues, depending on the imaging approaches [14]. Conventional machine learning (ML) approaches contain a hand-crafted feature extraction technique, out of which the features are extracted from training images before the process of learning starts. This approach requires expert help with extensive knowledge to recognize the significant features. As a result, while working with huge datasets, the classification accuracy of ML-based approaches is limited and prone to errors. Furthermore, DL approaches are proven to be extremely efficient in a narrow range of applications in medical imaging [15]. However, tumors exhibit different characteristics like shape, location, and size, making it difficult to accurately distinguish among different tumor types and classify them. Shaik and Cherukuri [16] suggested a multi-level attention network (MANet) for the classification of brain tumors. The MANet employed pre-trained Xception for extracting the representation of deep semantic features from MRI images. These spatial representations were fed into two consecutive attention modules. The initial attention model learned spatial attention and tumor-specific features. Then, the second attention model enabled the learning of cross-channel attention among spatial features, and assisted in focusing on the feature maps with tumor portions. The MANet was superior not only with respect to performance but also in several parameters. However, MANet suffered from reduced effectiveness due to its heavy dependence on diversity and quality of training data.

Rahman and Islam [17] implemented a parallel deep convolutional neural network (PDCNN) with data augmentation for brain tumor detection and classification. The input images were resized and grayscale transformation which assisted in minimizing the complexity was performed. The PDCNN extracted both local and global features from two parallel phases, and was performed with over-firring issues by employing a batch normalization and dropout regularizer. The parallel pathways were generated by combining two simultaneous DCNN having two window sizes which further aided the model to effectively learn local and global features. Nonetheless, PDCNN was incapable of handling complex data because of its vanishing gradient issues. Ladkat *et al.* [18] presented a mathematical model with 3D attention U-Net for the segmentation of brain tumors. Every slice of 3D image was increased by the presented approach which was then transmitted via 3D attention U-Net to generate a result of the segmented tumor. Feature extraction was employed as a primary criterion and the presented approach provided accurate segmentation of tumor pixels from 3D brain images. The presented approach maximized human lifespan and decreased the death rate with high accuracy and lower complication rates. Nevertheless, the presented approach struggled to capture diverse tumor characteristics due to its attention mechanism which does not possess the capacity to adequately subtle variations in tumor features. Rao and Karunakara [19] introduced a kernel-based support vector machine and social ski driver (K-SVM-SSD) for brain tumor detection and classification. Initially, blur removal was performed by utilizing normalized median filter (NMF) to smoothen the image and enhance its quality. A binomial thresholding was employed to segment the tumor regions. Then, the grey-level co-occurrence matrix (GLCM) and spatial grey-level dependence matrix (SGLDM) were employed to extract the features. The harris hawks optimization was utilized to select the features and finally, KSVM-SSD was performed to classify the brain tumors. The K-SVM-SSD accurately detected and classified the images of brain tumors with precise segmentation and low computation complexity. Nevertheless, K-SVM-SSD suffered from managing high-dimensional data due to the curse of dimensionality. Chattopadhyay and Maitra [20] developed a CNN to detect brain tumors from MRI images. The MRI images with different locations, tumor sizes, image intensities, and shapes were considered for training the CNN effectively. Then, SVM classifier and other activation functions like sigmoid, SoftMax, and RMSProp were utilized to cross-check the developed CNN technique. The developed technique learnt difficult features automatically from multi-modal MRI images. Yet, CNN the struggled with detecting brain tumors from MRI while managing small or subtle tumors, due to limited spatial resolution. In the overall analysis, it is evident that the existing approaches have limitations of being incapable of handling complex data, inaccurate classification due to various characteristics like location and size, and invariance issues due to default matrix sampling technique.

To solve these problems, the spatial transformer network and non-local attention mechanism (STN-NAM) is proposed in ResNet50, so as to accurately classify tumors by managing complex data and invariance issues.

The primary contributions of this research are as: i) U-Net is performed to segment tumor regions by capturing both local and global contexts in brain tumors. It effectively preserves spatial data via a network that assists in complex tumor boundaries and structures, which enables accurate segmentation; ii) ResNet50 is improved by using STN which increases the capability to extract features by enhancing the network's spatial invariances in brain tumors; and iii) NAM increases the model's robustness for distinguishing among different tumors, providing accurate classification results. The rest of the paper is structured as follows: section 2 discusses the proposed method. Section 3 explains the STN-NAM in ResNet50, while section 4 determines the results of the proposed method, and section 5 indicates the conclusion.

2. METHOD

In this research, STN-NAM is proposed in ResNet50 to classify the tumors. Initially, the image is acquired from three benchmark datasets: Figshare, Brats 2019, and Brats 2020 to determine the proposed technique. NMF is employed to reduce noise, enhance the image quality, and then resize to a standard size using normalization. U-Net is used to segment the brain tumor regions and finally, STN-NAM is performed for brain tumor classification. Figure 1 determines the block diagram for the proposed technique.

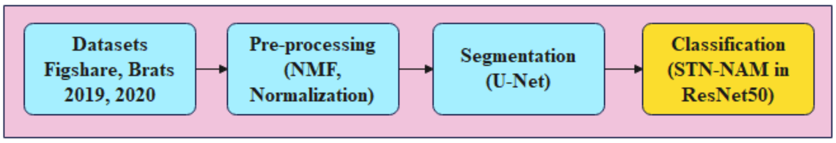

Figure 1. Block diagram for the proposed approach

2.1. Datasets

The proposed technique is analyzed on three benchmark datasets: Figshare [21], Brats 2019 [22], and Brats 2020 [23]. Table 1 indicates the three datasets' descriptions. These obtained images are fed as input to the pre-processing stage for removing the noise.

Table 1. Dataset description

Dataset	Patients	Total scans	Meningiomas	Glioma tumors	Pituitary tumors	LGG	HGG
Figshare	233	3064	708	1426	930	-	-
Brats 2019	355	1340	-	-	-	76	259
Brats 2020	369	1476	-	-	-	110	259

2.2. Pre-processing

After obtaining the image, the pre-processing stage using NMF is performed to remove the salt and pepper noise, and to enhance the image quality. Further, the Gaussian filter also removes the noise and enhances the image but also creates blurred edges due to the smoothening effect of the filter. By utilizing NMF, the corrupted pixel of an image is replaced with a median value which is formulated in (1):

$$mv_i^{(n-1)} = median[I_i^{(n-1)} | i \in W] \tag{1}$$

where, W denotes the window size employed to analyze the median value. Image sequence after the iteration $(n-1)$ is represented by $I_i^{(n-1)}$ and mv indicates the median value. After removing the noise, the image resizing is established utilizing normalization [24] technique that originates in the image being resized to 256×256 which is the standard size, as expressed in (2):

$$I_N(i,j) = \frac{(I^D - I_{Min})(I_{max\_N} - I_{min\_N})}{I_{max} - I_{min}} + I_{min\_N} \tag{2}$$

where, I^D represents the denoised image and I_{max}, I_{min} indicates the max and min intensities. $I_N(i,j)$ denotes the new resized image, and I_{max_N}, I_{min_N} determines the new max and min intensities. At last, the resized image is acquired after the normalization procedure, employed to segment the affected tumor region.

2.3. Segmentation

After pre-processing, the U-Net is used to effectively segment various regions in the brain tumors. U-Net effectively captures both global and local contexts in brain tumors. U-Net preserves spatial data via the network which captures the complex tumor boundaries and structure, as compared to Segnet as it depends on the pooling layers which lead to the loss of spatial data. U-Net integrates contraction (encoder) and expanding paths (decoder) with skip connections which provide fine-grained information from various scales that make accurate delineation of tumor boundaries, even in heterogenous and complex tumor structures. The U-Net's [25] encoder path captures the context of the input image; this path is simply a pipeline of pooling and convolutional layers. Decoder path employs transposed convolutions which enable accurate localization. It has only a stack of max-pooling and convolution layers and there is no fully connected (FC) feedforward layer in U-Net. Various stacked convolutional layers allow network to learn more accurate features from compressed images. The input images are compressed to fit into a representation of latent space.

The U-Net segmentation performance is analyzed by monitoring its pixel error, rand, and warping. It has nested pathway series of dense skip that minimizes the gap among pathways and feature maps. The U-Net architecture is achieved by employing TensorFlow, where it has four convolutional blocks. Every network's convolutional block has two convolutional layers with 3×3 kernel size and zero padding at every layer to manage object dimension shrinkage after applying filters. After each layer, the filter size per convolutional block is varied, wherein the size of the filter increases in the 16th step. Every convolutional block layer is activated by rectified linear unit (ReLU), whereas among these layers, a batch normalization phase is employed. At the encoder network layer, a 2×2 max-pooling layer is utilized after a call function to include a convolutional block for reducing an input image's spatial dimensions. Additionally, max-pooling is employed at the decoder layer and its application is to up-sample the feature map by utilizing indices of memorized max-pooling. Figure 2 shows the U-Net architecture. U-Net is proven to be greatly efficient for the segmentation of brain tumors which facilitates accurate delineation of tumor boundaries via its capability to maintain spatial information and combine at various scales. These segmented images are fed into STN-NAM in ResNet50 to classify the brain tumors.

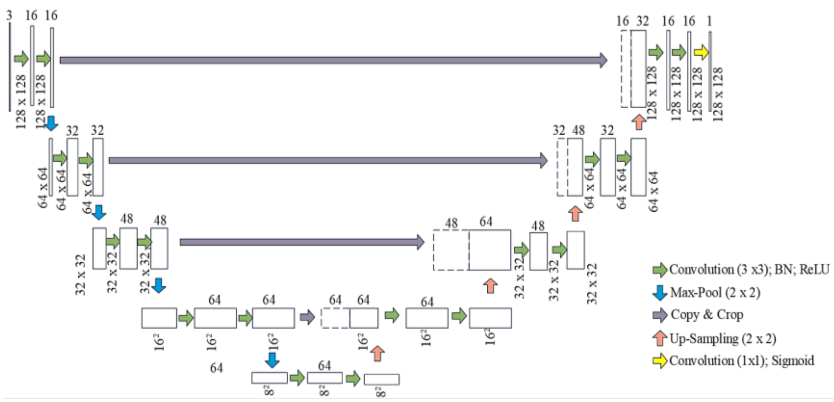

Figure 2. U-Net architecture

3. STN-NAM IN RESNET50

The segmented region portion is fed into ResNet50 with the help of STN-NAM to classify the brain tumors effectively. ResNet-50 has a deeper architecture that allows it to capture complex features, solve vanishing gradient issues, and manage large-scale image classification tasks with accurate classification results in brain tumor images. It is a type of CNN model that establishes the residual learning concept, thereby reducing the vanishing gradient issue. However, ResNet50 suffers from invariance to affine image transformation which is caused by the matrix sampling technique. When an input image transforms rotation and translation, the model's output is varied, also affecting the capability to generalize among various orientations or sizes which leads to invariance. The ResNet50 does not satisfactorily extract the feature space's long-distance correlation because of limited receptive field of convolutional layers. To solve this issue, an STN-NAM is developed in ResNet50 to classify the brain tumors. The developed approach employs MRI slice input in a huge amount of subjects to train the network, where the image features are learned automatically, preventing manual extraction, and then the image is categorized depending on these features for acquiring accurate classification outcomes. Figure 3 represents the architecture of STN-NAM.

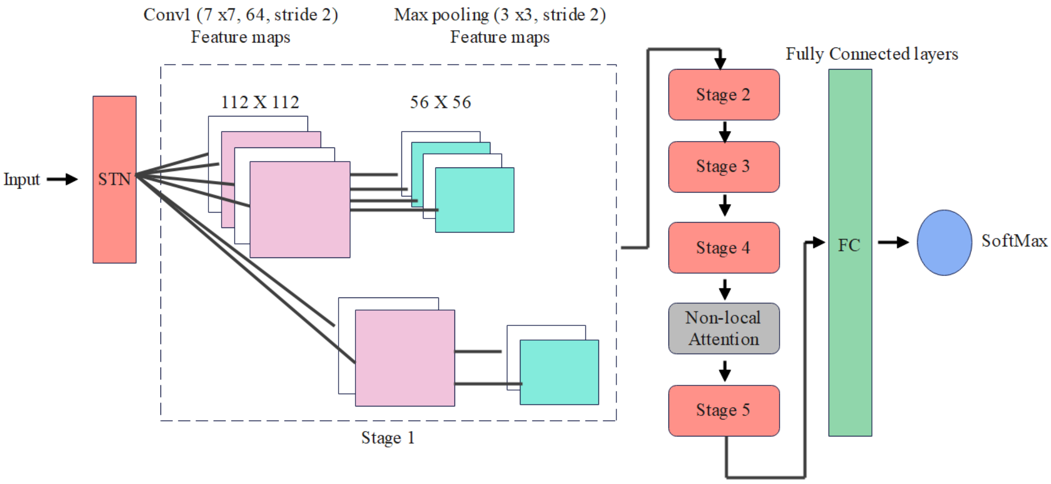

Figure 3. Diagram for STN-NAM in ResNet50

STN is established among input layer and ResNet50 that transforms spatial data in MRI images of tumor patients into other space and keeps the key information which reduces the model's spatial invariance. STN allows the model to generate geometric transformations on input images that align them to a standard scale or orientation. This ensures that the features of tumor are consistently indicated across images which increase the model's capability to classify tumors despite variations in orientation and positions. NAM is presented among 4th and 5th stages of ResNet50 that identifies the relationship among normal areas and lesion areas in the feature space. NAM enables the model to capture long-range dependencies within an input image, which allows it to focus on appropriate image regions for tumor classification. By processing to significant spatial and contextual data, the model determines a subtle tumor feature and distinguishes them from nearby healthy tissues. It generates a more accurate classification of complex tumor structures. A local network is utilized in transformation regression parameter θ and their size is based on the type of transformation utilized. According to predicted transformation parameters, a grid generator is employed to construct the sampling grid. After transformation and sampling, it is a result of point groups in an input image. It is considered that the coordinate of every pixel of input and output image is (x_i^s, y_i^s) and (x_i^t, t_i^t). A function of space transformation T_θ is a 2D affine transformation function. An associated relationship among (x_i^s, y_i^s) and (x_i^t, t_i^t) is expressed in (3):

$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = T_\theta(G_i) = A_\theta \begin{pmatrix} x_i^t \\ t_i^t \\ 1 \end{pmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix} \tag{3}$$

where, s indicates the input feature image's coordinate point, T denotes the output feature image's coordinate point and A_θ represents the local network's output. The sampler in STN employs a feature map's input and sampling grid as input to generate an output. Moreover, n and m transverse each coordinate of the original graph U, and U_{nm}determines pixel values in the original graph U. Then, x_i^s, y_i^s indicate the coordinates of associating point in U graph to be determined at i^{th} point on V, as expressed in (4).

$$V_i = \sum_n \sum_m U_{nm}\, max(0, 1 - |x_i^s - m|)\, max\,(0,1 - |y_i^s - n|) \tag{4}$$

The combination of STN module among input and ResNet50 enables a network to learn automatically how to transfer feature maps, therefore assisting in minimizing the overall cost of network training. The output value is located in the network that represents how to transfer every item of the training data. The NAM is embedded as an element in ResNet50 which learns new weights in transfer learning, hence the pre-trained weights are not necessary due to new modules in brain tumors. The combination of STN increases the capability to extract features by enhancing the network's spatial invariances in brain tumors. The NAM increases the model's robustness for distinguishing among different tumors which provides accurate classification results. Table 2 displays the notation description for the equation.

Table 2. Notation description

Symbols	Description
W	Window size
mv	Median value
I^D	Denoising image
I_{max}, I_{min}	Max and min intensities
$I_N(i,j)$	New resized image
I_{max_N}, I_{min_N}	New max and min intensities
x_i^s, y_i^s	Coordinate of each input image pixel
x_i^t, t_i^t	Coordinate of each output image pixel
s	Input feature image's coordinate point
T	Output feature image's coordinate point
A_θ	Local network's output
U_{nm}	Pixel values in original graph U

4. RESULTS

The proposed STN-NAM is simulated using MATLAB R2020b with 16 GB RAM, Windows 10 operating system, Intel i5 processor, and 6 GB graphics processing unit (GPU). The performance metrics of accuracy, recall, precision, F1-score, dice score coefficient (DSC), intersection over union (IoU), and mean IoU (MIoU) are evaluated by using (5) to (11). Accuracy defines the number of correct predictions divided by an overall number of predictions. Recall determines the number of all true predictions (both negative and positive) by a total number of predictions. Precision evaluates the number of true positives divided by a number of true positives and false positives. F1-score is the combination of both precision and recall. TP indicates true positive, FN represents false negative, TN denotes true negative, and FP determines false positive.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$F1 - Score = \frac{2\times TP}{2\times TP+FP+FN} \times 100 \tag{8}$$

$$DSC = \frac{2\times TP}{(TP+FP)+(TP+FN)} \tag{9}$$

$$IoU = \frac{TP}{TP+FP+FN} \tag{10}$$

$$MIOU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{FN+FP+TP} \tag{11}$$

4.1. Performance analysis

The performance analysis of STN-NAM is presented in Tables 3 to 8. Table 3 displays the segmentation results on the Figshare dataset. The existing techniques like DeepLab, Segnet, and V-Net are compared with U-Net technique. The obtained results show that the U-Net achieves a high DSC of 0.9075 due to encoder-decoder pathways with skip connections which provide an accurate localization of structures while preserving spatial data.

Table 3. Segmentation performance using Figshare dataset

Metrics	DeepLab	Segnet	V-Net	U-Net
DSC	0.8205	0.8344	0.8475	0.9075
IoU	0.8264	0.8475	0.8567	0.9168
MIoU	0.8347	0.8520	0.8905	0.9355

Table 4 indicates the classification analysis on Figshare dataset. The performance of recurrent neural network (RNN), deep neural network (DNN), CNN, and ResNet 50 are compared with proposed STN-NAM in the ResNet50 technique. When compared with these existing techniques, the proposed technique achieves a high accuracy of 98.15% in Figshare dataset. Due to existing techniques having limitations in capturing complex spatial relationships and long-range dependencies in the images of brain tumors, achieving lower performance. The proposed STN-NAM effectively captures spatial variations and extracts appropriate features which increase the accuracy of classification performances.

Table 4. Classification analysis using Figshare dataset

Metrics (%)	RNN	DNN	CNN	ResNet50	STN-NAM in ResNet50
Accuracy	86.50	87.36	90.66	94.89	98.15
Recall	84.26	86.44	88.14	91.38	97.48
Precision	87.05	88.47	89.06	89.95	97.50
F1-score	85.23	84.35	87.96	89.30	98.06

Table 5 determines the segmentation analysis for Brats 2019 dataset. The performances of DeepLab, Segnet, and V-Net are compared with the U-Net architecture. The U-net achieves a high DSC of 0.9247 as compared to the existing techniques due to encoder-decoder pathways with skip connections which provides accurate localization of structures.

Table 5. Segmentation analysis using Brats 2019 dataset

Metrics	DeepLab	Segnet	V-Net	U-Net
DSC	0.8247	0.8311	0.8437	0.9247
IoU	0.8397	0.8364	0.8533	0.9285
MIoU	0.8421	0.8475	0.8667	0.9404

Table 6 represents the classification performance on Brats 2019 dataset. The performance of RNN, DNN, CNN, and ResNet50 are compared with proposed STN-NAM in ResNet50. When compared to these existing approaches, the proposed technique achieves a high accuracy of 99.24% due to it effectively capturing spatial variations and extracts relevant features with a combination of STN and non-local attention, which has accurate classification performances. Table 7 indicates the segmentation performance on Brats 2020 datasets. The performance of DeepLab, Segnet, and V-net are compared with the U-Net technique which achieves a high DSC of 0.9458 compared to existing techniques due to encoder-decoder pathways with skip connections which provides accurate localization of structures.

Table 6. Classification analysis by employing Brats 2019

Metrics (%)	RNN	DNN	CNN	ResNet50	STN-NAM in ResNet50
Accuracy	84.54	83.67	83.53	88.58	99.24
Recall	83.24	84.56	86.67	87.02	98.67
Precision	84.12	85.14	83.58	85.36	99.15
F1-score	82.06	85.57	88.01	87.48	99.05

Table 7. Segmentation performance of Brats 2020 dataset

Metrics	DeepLab	Segnet	V-Net	U-Net
DSC	0.8487	0.8347	0.8514	0.9458
IoU	0.8567	0.8399	0.8658	0.9574
MIoU	0.8677	0.8458	0.8699	0.9605

Table 8 represents the classification analysis on Brats 2020 datasets. The performance of RNN, DNN, CNN, and ResNet50 are compared with the proposed technique. The proposed technique achieves a high accuracy of 99.89% as opposed to RNN, DNN, CNN, and ResNet50 techniques as the proposed approach effectively captures spatial variations and extracts relevant features with a combination of STN and non-local attention, which has accurate classification performances.

Table 8. Classification evaluation utilizing Brats 2020

Metrics (%)	RNN	DNN	CNN	ResNet50	STN-NAM in ResNet50
Accuracy	85.35	86.52	88.96	93.58	99.89
Recall	84.20	85.06	87.58	90.25	99.04
Precision	86.28	86.98	87.24	89.55	99.25
F1-score	85.67	85.04	88.72	89.34	98.66

4.2. Comparative analysis

Table 9 indicates the comparative analysis with existing techniques. The existing techniques like MANet [16], PDCNN [17], mathematical model with 3D attention U-Net [18], K-SVM-SSD [19], and CNN [20] are compared with STN-NAM in ResNet50 on different datasets. In relation to these existing techniques, the proposed approach achieves high accuracy of 98.15%, 99.24%, and 99.89% in Figshare, Brats 2019, and Brats 2020 datasets due to the ResNet50 effectively capturing spatial variations and extracting appropriate features with the help of STN and non-local attention, which achieves accurate classification performances.

Table 9. Comparative analysis with existing methods

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
MANet [16]		96.51	N/A	N/A	N/A
PDCNN [17]		97.60	97.00	97.00	97.00
Proposed method	Figshare	98.15	97.50	97.48	98.06
DNN-3D attention U-Net [18]		98.90	99	98	98.50
Proposed method	Brats 2019	99.24	99.15	98.67	99.05
KSVM-SSD [19]		99.15	99	98.53	98.4
CNN [20]		99.74	N/A	N/A	N/A
Proposed method	Brats 2020	99.89	99.25	99.04	98.66

4.3. Discussion

The advantages of the proposed STN-NAM in ResNet50 and the disadvantages of existing techniques are discussed here. The existing techniques have limitations as follows: MANet [16] suffers from reduced effectiveness due to its heavy dependence on diversity and quality of training data. PDCNN [17] does not handle complex data because of its vanishing gradient issues. The mathematical approach with 3D attention U-Net [18] struggles to capture diverse tumor characteristics due to its attention mechanism which does not possess the capacity to adequately subtle variations in tumor features. CNN [20] struggles with detecting brain tumors from MRI while managing small or subtle tumors due to limited spatial resolution. The proposed technique overcomes these existing limitations. The ResNet50 captures complex features or patterns in brain tumors. The combination of STN increases the capability for extracting network features by enhancing the network’s spatial invariances in brain tumors. NAM increases the model’s robustness for distinguishing among different tumors which provide accurate classification results. By performing these operations, the STN-NAM in ResNet50 achieves a high accuracy of 98.15%, 99.24%, and 99.89% in Figshare, Brats 2019, and Brats 2020 datasets, respectively.

5. CONCLUSION

In this research, the STN-NAM is proposed in ResNet50 to classify the brain tumors accurately. The ResNet50 is constructed with the help of STN and NAM techniques; in that, STN transforms spatial data in MRI images of tumor patients into other space and keeps key information that reduces the model’s spatial invariance. NAM identifies the relationship among normal and lesion areas in the feature space which provides accurate classification of brain tumor. Mish activation function is utilized for replacing ReLU in traditional ResNet-50 to address local information loss issues by effectively extracting the feature’s space long-distance correlation in brain tumors. Through this process, the proposed approach achieves a high accuracy of 98.15%, 99.24%, and 99.89% on three datasets as opposed to MANet, U-Net, and CNN. In the future, the another improved DL methods like long short term memory (LSTM), RNN, or gated recurrent unit (GRU) will be used to capture temporal relationships to improve the classification outcomes.

ACKNOWLEDGEMENTS

This work was supported by the JSPM University, Pune and Pimpri Chinchwad College of Engineering Pune. We are grateful for the access to state-of-the-art laboratories, research materials, and technological resources offered by the Institutions.

REFERENCES

[1] S. Kumar, S. Choudhary, A. Jain, K. Singh, A. Ahmadian, and M. Y. Bajuri, “Brain tumor classification using deep neural network and transfer learning,” *Brain Topography,* vol. 36, no. 3, pp. 305-318, Jan. 2023, doi: 10.1007/s10548-023-00953-0.

[2] A. Raza *et al*., "A hybrid deep learning-based approach for brain tumor classification," *Electronics*, vol. 11, no. 7, Jun. 2022, doi: 10.3390/electronics11071146.
[3] M. Rasool *et al*., "A hybrid deep learning model for brain tumour classification," *Entropy*, vol. 24, no. 6, Jun. 2022, doi: 10.3390/e24060799.
[4] H. Mehnatkesh, S. M. J. Jalali, A. Khosravi, and S. Nahavandi, "An intelligent driven deep residual learning framework for brain tumor classification using MRI images," *Expert Systems with Applications*, vol. 213, p. 119087, Oct. 2023, doi:10.1016/j.eswa.2022.119087.
[5] H. ZainEldin *et al*, "Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization," *Bioengineering*, vol. 10, no. 1, p. 18, Dec. 2022, doi: 10.3390/bioengineering10010018.
[6] M. A. Talukder *et al*., "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning," *Expert systems with applications*, vol. 230, p.120534, 2023, doi: 10.1016/j.eswa.2023.120534
[7] E. U. Haq, H. Jianjun, K. Li, H. U. Haq, and T. Zhang, "An MRI-based deep learning approach for efficient classification of brain tumors," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 6697–6718, Oct. 2023, doi: 10.1007/s12652-021-03535-9.
[8] S. Shanthi, S. Saradha, J. A. Smitha, N. Prasath, and H. Anandakumar, "An efficient automatic brain tumor classification using optimized hybrid deep neural network," *International Journal of Intelligent Networks*, vol. 3, pp. 188-196, Jan. 2022, doi: 10.1016/j.ijin.2022.11.003.
[9] T. Muezzinoglu *et al*., "Patchresnet: multiple patch division–based deep feature fusion framework for brain tumor classification using MRI Images," *Journal Digital Imaging*, vol. 36, no. 3, pp. 973-987, Jun. 2023, doi: 10.1007/s10278-023-00789-x.
[10] R. Vankdothu, M. A. Hameed, and H. Fatima, "A brain tumor identification and classification using deep learning based on CNN-LSTM method," *Computers and Electrical Engineering*, vol. 101, Jul. 2022, doi: 10.1016/j.compeleceng.2022.107960.
[11] G. Satyanarayana, P. A. Naidu, V. S. Desanamukula, and B. C. Rao, "A mass correlation based deep learning approach using deep convolutional neural network to classify the brain tumor," *Biomedical Signal Processing and Control*, vol. 81, Mar. 2023, doi: 10.1016/j.bspc.2022.104395.
[12] D. Rammurthy and P. K. Mahesh, "Whale harris hawks optimization based deep learning classifier for brain tumor detection using MRI images," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3259-3272, Jun. 2022, doi: 10.1016/j.jksuci.2020.08.006.
[13] P. Agrawal, N. Katal, and N. Hooda, "Segmentation and classification of brain tumor using 3D-Unet deep neural networks," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 199-210, Jun. 2022, doi: 10.1016/j.ijcce.2022.11.001.
[14] F. Demir, Y. Akbulut, B. Taşçı, and K. Demir, "Improving brain tumor classification performance with an effective approach based on new deep learning model named 3ACL from 3D MRI data," *Biomedical Signal Processing and Control*, vol. 81, Mar. 2022, doi: 10.1016/j.bspc.2022.104424.
[15] H. Kibriya, M. Masood, M. Nawaz, and T. Nazir, "Multiclass classification of brain tumors using a novel CNN architecture," *Multimedia Tools and Applications*, vol. 81, no. 21, pp. 29847-29863, Sep. 2022, doi: 10.1007/s11042-022-12977-y.
[16] N. S. Shaik and T. K. Cherukuri, "Multi-level attention network: application to brain tumor classification," *Signal, Image Video Process*, vol. 16, no. 3, pp. 817-824, Apr. 2022, doi: 10.1007/s11760-021-02022-0.
[17] T. Rahman and M. S. Islam, "MRI brain tumor detection and classification using parallel deep convolutional neural networks," *Measurement: Sensors*, vol. 26, p. 100694, Apr. 2023, doi: 10.1016/j.measen.2023.100694.
[18] A. S. Ladkat *et al*., "Deep neural network-based novel mathematical model for 3D brain tumor segmentation," *Computational Intelligence and Neuroscience*, Aug. 2022, doi: 10.1155/2022/4271711.
[19] S. Rao and K. Karunakara, "Efficient detection and classification of brain tumor using kernel based SVM for MRI," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 7393-7417, Feb. 2022, doi: 10.1007/s11042-021-11821-z.
[20] A. Chattopadhyay and M. Maitra, "MRI-based brain tumour image detection using CNN based deep learning method," *Neuroscience Informatics*, vol. 2, no. 4, Dec. 2022, doi: 10.1016/j.neuri.2022.100060.
[21] Figshare Dataset link: https://figshare.com/articles/dataset/brain_tumor_dataset/1512427, (Accessed Dec. 28, 2024).
[22] Brats 2019 dataset link: https://www.med.upenn.edu/cbica/brats2019/data.html, (Accessed Dec. 28, 2024).
[23] Brats 2020 dataset link: https://github.com/akhanss/BraTS-2020, (Accessed Dec. 28, 2024).
[24] H. Avcı and J. Karakaya, "A novel medical image enhancement algorithm for breast cancer detection on mammography images using machine learning," *Diagnostics*, vol. 13, no. 3, p. 348, Jan. 2023, doi: 10.3390/diagnostics13030348.
[25] J. Walsh, A. Othmani, M. Jain, and S. Dev, "Using U-Net network for efficient brain tumor segmentation in MRI images," *Healthcare Analytics*, vol. 2, Nov. 2023, doi: 10.1016/j.health.2022.100098.

BIOGRAPHIES OF AUTHORS

Preeti Sadanand Topannavar has completed her Engineering Degree from Amravati University, India in Electronics and Telecommunication in 2002. She has received a master's degree in Electronics and Telecommunication engineering from Sinhgad College of Engineering Pune 2010 respectively. She is awarded a Ph.D. degree from Savitribai Phule Pune University in 2024. She has total teaching experience of 22 years and currently working at JSPM University Pune India. She has 2 granted patents and more than 25 papers in Journals and international conferences. Her current research interests include biomedical image processing, human-computer interactions, and machine learning. She has published various research papers in different international and national journals and conferences. She can be contacted at email: topannavarp@gmail.com.

Varsha Sachin Bendre received a bachelors in electronics and telecommunication engineering from Amaravati University and an M.E. degree from Rajarshi Shahu College of Engineering, Pune in 2000 and 2010, respectively. She completed her Ph.D. in low power VLSI from the Rajarshi Shahu College of Engineering, Pune, affiliated with Savitribai Phule Pune University, Pune, India, in January 2020. She has a total teaching experience of 19 years and is currently working as an Associate Professor at Pimpri Chinchwad College of Engineering, Pune, India. Her research interests include nanotechnology, VLSI design, microelectronics, low-power analog circuits, and signal processing. She has published several research papers in various SCI/Scopus-listed journals and more than 30 research papers in peer-reviewed international conferences. She can be contacted at email: dineshyadav800@gmail.com.

Deepti Khurge is certified by DAAD (German Academic Exchange) DIES ProGRANT proposal writing for research grants by University of Cologne, Focusing on practical-oriented training for new researchers in India 2023. She has 4 granted patents, 4 copy rights, more than 30 papers in Journals and international conference. Her area of interest in VLSI design. She completed innovation ambassador training conducted by MoE's Innovation Cell and AICTE during the period from 30th June-30th July 2021 in online mode. She successfully completed accelerating project-based learning using internet of things and thing speak, 19th 2021 Alborg UNESCO center, IUCEE foundation along with regional research Symposium on. She can be contacted at email: dipti.khurge@pccoepune.org.