Journal of Applied Data Sciences
Vol. 6, No. 2, May 2025, pp. 1277-1290

Automated Brain Tumor Analysis with Multimodal Fusion and Augmented Intelligence

R. Karthick Manoj[1,*], Aasha Nandhini S [2], M.Batumalay[3,4]

[1]Department of Electrical and Electronics Engineering, AMET Deemed to be University, Kanathur, India

[2]Department of Electronics and Communication Engineering, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, India

[3]Faculty of Information Technology, INTI International University, Malaysia

[4]Centre for Data Science and Sustainable Technologies, INTI International University, Nilai, N. Sembilan, Malaysia

(Received: December 22, 2024; Revised: February 26, 2025; Accepted: March 24, 2025; Available online: April 15, 2025)

Abstract

Brain tumor segmentation and classification are critical tasks in medical imaging, having a major impact on spotting and treating brain tumors. In the medical field, augmented intelligence has garnered a lot of attention lately since it emphasizes how human knowledge and artificial intelligence can be combined to enhance efficiency and decision-making in applications like brain tumor identification. This research concentrates on developing a novel approach utilizing Attention U-Net and Multimodal Transformers to assist doctors with precise tumor segmentation and classification while maintaining their critical clinical judgment. Attention U-Net is used to segment brain tumor because it efficiently collects detailed spatial data while focusing on key locations compared with traditional U-Net models. Multimodal Transformers provide reliable as well as effective feature extraction when utilized for early fusion to merge data from many modalities, such as T1, T2, and FLAIR This work utilizes CycleGAN-based data augmentation to supplement limited training data, thus improving the variety and quality of the dataset. The fused multimodal features are then utilized for the segmentation of the tumor and further classified as benign and malignant using hybrid transformer. The performance of the proposed system is assessed using standard metrics like accuracy for classification and Dice Similarity Coefficient and Intersection Over Union for segmentation. The proposed approach demonstrates high effectiveness in both segmentation and classification tasks, achieving 98 % accuracy showcasing its potential as a process innovation for clinical applications.

Keywords: Brain Tumor Segmentation, Multimodal Imaging, Early Fusion, U-Net Architecture, Transformer, Accuracy, Process Innovation

1. Introduction

In recent years, AI has an impact on helping doctors classify brain tumor [1], [2]. To analyze medical images, doctors need to segment and classify brain tumor. Medical imaging techniques including Positron Emission Tomography scans, Computed Tomography, and magnetic resonance imaging are used by physicians to detect brain cancers. MRI stands out as the go-to method to diagnose brain tumor. It shows clear pictures of the brain's inside revealing both healthy and unhealthy parts [3]. Doctors prefer multimodal imaging over single modal imaging to segment brain tumor. This approach gives them a full picture of the tumor. Multimodal imaging has an impact on accurate segmentation, cuts down on time, and gives extra details to get the shape and disease information of brain tumor [4]. There are four types of MRIs: fluid attenuated inversion recovery, T1-weighted (T1), T2-weighted (T2), and contrast-enhanced T1-weighted (T1ce). TI displays the composition and configuration of various brain tissues.

This makes it easier to spot tumor, cysts, and other unusual growths. T1-weighted with contrast enhanced is similar to T1 but with highlighted abnormalities. T2 shows the fluid content in different types of tissue and FLAIR suppresses the fluid content and enhanced the tumors that are not clearly visible in T1 or T2 images. Tumor segmentation from medical imaging by hand is a difficult process that calls for professional assistance. With advancements in medical imaging, automated classification of brain tumor has become a challenging task, hence DL models are used for efficient

*Corresponding author: R. Karthick Manoj (aasha.nandhu@gmail.com)
DOI: https://doi.org/10.47738/jads.v6i2.719
This is an open access article under the CC-BY license (https://creativecommons.org/licenses/by/4.0/).
© Authors retain all copyrights

and automated classification of brain tumor [5]. In literatures many have used ML models, extreme learning models and learning models are used for brain tumor classification [6]. To enhance the classification accuracy, it is always important to efficiently segment the tumor region and then make the system to learn the features corresponding to the tumor. Brain scan tumor segmentation by hand is a laborious and time-consuming procedure. Furthermore, low-quality, difficult-to-interpret images can arise from abnormalities generated during the imaging process. The presence of irregular lesions, varying shapes, and indistinct borders further complicates the manual detection of brain tumor. Hence this research focuses on developing an automated brain tumor segmentation and classification framework which can efficiently and accurately segment the tumor and classify it with high accuracy compared with existing Artificial Intelligence (AI) based models. In this work, attention U-Net is used for medical image segmentation to separate tumor from multiple images. The attention U-Net for multimodal sensing along with CycleGAN and transformer-based classification makes it a unique approach. Unlike basic augmentation techniques, CycleGAN maintains the structural integrity of tumors and surrounding tissues while converting images from one modality to another. This is particularly important in medical imaging, where precision in shape, intensity, and spatial relationships plays a crucial role.

The rest of the document is structured as follows: The most recent techniques for classifying and segmenting brain tumor in multimodal images are reviewed in Section 2. The suggested Attention U-Net based brain tumor segmentation (AU-BTS) framework is explained in depth in Section 3, with an emphasis on how it improves segmentation efficiency. The validation procedure and findings are presented in Section 4, and the paper's conclusion and possible future research areas are covered in Section 5.

2. Literature Review

In this section, several works related to the research have been reviewed, and their advantages and limitations have been discussed. This highlights opportunities for additional development and offers a thorough grasp of the field's existing situation.

It is essential to use DL-based methods for the diagnosis of the illness, which give doctors the means to identify likely Alzheimer's earlier. The gradient was calculated using several mistakes below the curves. The proposed feature selection methods may be able to find a minimal level feature set for Alzheimer's disease diagnosis from an initial larger input feature set, as this study [7] shows using standard datasets. In terms of accuracy and losses compared to training and accuracy and losses compared to validation, these algorithms provide superior system performance values. These outcomes may demonstrate how well the model fits the job. In both single- and multi-modality research, this study examines preprocessing techniques, characteristics, and biomarkers [8]. A number of DL models were evaluated for their ability to identify AD. Issues like sparse datasets and training protocols continue to exist. Discriminative feature representations are required in order to distinguish AD from similar brain patterns. This paper emphasizes the value of datasets while highlighting the promise and constraints of DL in AD detection. The creation of benchmark platforms for simplified comparisons is one of the future directions. In summary, DL has the potential to detect AD accurately, but in order to overcome obstacles and improve diagnostic accuracy, models and techniques must be improved.

Automated Alzheimer's diagnosis may be possible with the use of electroencephalogram data analysis. However, because they examine each EEG channel separately, current approaches usually ignore the complex functional connections that exist throughout the brain during various tasks. To address these issues, this work proposed a unique graph signal processing-based method. This paper presents a novel graph Fourier transform characteristic for Alzheimer's illness diagnosis based on the idea of GSP [9]. Our work has been compared to the current feature-based approaches and the well-liked discrete wavelet transform feature. In order to classify the condition of Parkinson's using EEG data, this study [10] proposes a unique method for examining EEG channel-wise analysis and identifying key brain areas. The wavelet Scattering Transform precisely captures the unique temporal and spectral qualities of the EEG signals, the AlexNet CNN recognizes complicated spatial features at several levels, enabling the successful detection of intricate patterns linked to Parkinson's illness.

In this study [11], it is suggested a new illness detection paradigm. We construct a system based on artificial intelligence that retrieves different biomedical data from homogeneous and dispersed sensors. We investigate the relationships between various biomedical data in order to identify and diagnose diseases using a variety of DL architectures with ensemble learning and attention techniques. We test biomedical data in great detail. The findings demonstrate the advantages of diagnosing illnesses during the healthcare decision-making process by utilizing DL technologies in the field of AI of medical objects. For instance, the suggested methodology provides a 92% disease identification rate.This study seeks to provide optimum representation scheme that can precisely segregate patients of Parkinson's disease based on electroencephalogram data. Two hybrid DNN is presented that integrates CNN with LSTM and generates both series and parallel models for Parkinson diagnosis through EEG inputs. A comprehensive deep CNN network will be used in the connection induced by the input signals to describe the contextual dependencies among the features. The parallel model proposed yields an impressive 99.7% accuracy [12].

Timely diagnosis and appropriate treatment can substantially enhance the quality of life for individuals with Parkinsons disease. This research empowers hand-drawn images to help identify Parkinson's disease. In this article, we implemented three different machine learning algorithms [13]. The model has been shown to adapt well to a range of shapes, such as cubes, triangles, waves, and spirals. According to the findings of the experiment the accuracy of the proposed algorithm was 98.75%. In this study, the logistic regression ML method would be employed to identify brain abnormalities autonomously. Brain MRIs images are collected for training and testing and various researchers have proposed multiple methods for brain tumor detection yet they are outdated and not dynamic [14] It is used to overcome the limitations of previous models, therefore, requiring a ML system. Real-time MRI images are used to assess various age groups and orientations. Logistic regression and threshold segmentation have been used to process the illness categorization.

For the sole purpose of medical identification, medical image segmentation is essential for locating important regions or features within pictures. Since segmentation of images using DL techniques is becoming common practice for image analysis in medicine and other domains, its importance has risen over the years. It is difficult to segment the contents of brain images which is a prerequisite in the analysis and management of a number of disorders of the brain. The challenge of this project is to characterize a DL framework called U-Net which was specifically designed for image segmentation in brain MRI studies [15]. Many brain structures such as the cerebral cortex and also the subcortical areas have successfully been segmented with the use of U-Net and its applications in these areas seem to be very promising. We survey recent advances in U-Net applications for MRI skull stripping as well as the use of DL methods for segmentation.

At present, one of the most dangerous health conditions is the brain tumor. Tumor impacts the brain through either destruction of the healthy brain tissue or increase the pressure within the cranium. Because of these reasons, a rapid multiplication of a tumor cell can become deadly. As a result, it becomes absolutely necessary to identify the presence of a brain tumor in its earlier stages which may help to prevent the patient from several adverse effects. The task of medical image processing contributes significantly in assisting the population in recognizing different diseases. The expertise and abilities of the diagnosing physician have a significant impact on how brain tumor is classified. In fact, SVM classifier has gained popularity in general and military medicine because it is well targeted and has shown effective results in modern brain image classification suggesting it as the best ML method in this case [16]. This warrants the use of the SVM classifier in this paper to assist in classifying tumor in the brain. The treatment of brain tumor, which are among the most prevalent and fatal diseases, is the main topic of this study. To locate the tumor location in the brain, MR pictures were provided as MRI sequence images for an input image for a brain tumor lesion detection and segmentation network. The ultimate aim is to find out whether a brain tumor is present or not in an apparently normal brain.

The study described in the abstract highlights the impressive progress made in the detection and recognition of neurological conditions, such as Parkinson's, Alzheimer's, and brain cancers, using DL and AI technology [17], [18]. In order to improve disease identification’s precision and effectiveness, several techniques were used, including different neural network models and data fusion methods. Such techniques allow for reliable and automatic detection which is important for primary prevention and treatment. Even though good results have been achieved with regards to classification accuracy and sensitivity, obstacles characterize the application of these methods, which include limited number of samples available, intricate architectures of the human brain, and poor feature representation. In addition, the combination of DL models with medical data, the benefit of using multiple datasets in enhancing the detection and

ISSN 2723-6471

classification of diseases. In this regard, it is imperative that such models are improved continuously, including existing problems such as the quality of the datasets, and more complete systems for competitive assessment are developed. Together with advances in data preprocessing and hybrid model creation, these developments hold the potential to greatly improve medical diagnostic capacities and improve patient outcomes by detecting neurological conditions more quickly and accurately. From the discussions of the related work, it is observed that U-Net based segmentation improves the accuracy for ML/DL based medical imaging applications. U-Net based segmentation labels in a pixel-by-pixel manner which is important for crucial applications [19], [20].

3. Proposed Attention U-Net based Brain tumor Segmentation (AU-BTS) framework

The designed Attention U-Net based Brain tumor Segmentation framework is shown in figure 1. An efficient approach for the segmentation and classification of tumor from the multimodal dataset is depicted in the block diagram. Cranial and tumor information is captured in the Multi-modal input images. Image normalization is the following stage, which guarantees that the brightness levels of the images from every direction are consistent for smooth processing. So, once the image came off the pixel calibration stage, CycleGAN data augmentation is performed to augment the data for further analysis and rising accuracy. Equally important, Image registration module is included to ensure that all multimodal images are transformed to a predetermined spatial coordinate system and the underlying anatomic structures are well overlaid in corresponding locations in all the sequences. Once the images are aligned, Attention U-Net comes into play for the purpose of tumor Segmentation, whereby through the use of attention, the model is able to emphasize on significant parts of the tumor and fine-grained details while ignoring the non-essential areas compared to tradition U-Net model which processes all features equally. Following the segmentation process, Morphological Segmentation is performed in order to slightly modify the tumor mask by further editing the tumor borders through dilation and erosion. After creating a tumor mask, the Transformer-based Classification step is done so that the tumor features that are segmented can be used to determine whether a tumor is malignant or benign. A tumor boundary mask and information on the type of tumor are the result of the complete process, which can help with detection and treatment. This pipeline incorporates state-of-the-art techniques, including CycleGAN image enhancement, Attention U-Net for effective segmentation, and Transformer models for classification, which together allow efficient and accurate detection and analysis of brain tumor.

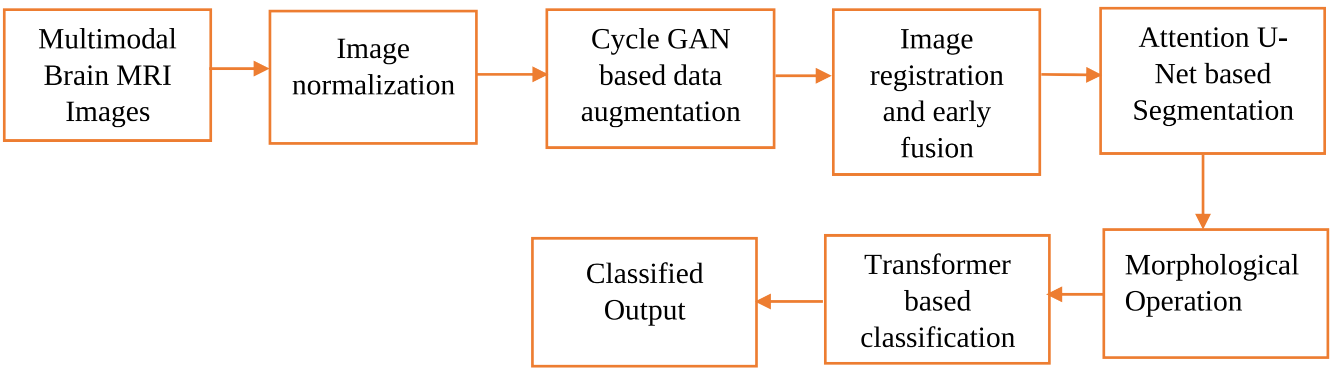


Figure 1. Proposed Multimodal Sensing based Brain tumor Segmentation and Classification Framework

3.1. Multimodal Image Input

The starting point lies in the gathering of brain MRI scans of the different types such as T1, T2, FLAIR and contrast-enhanced scans. Each of these imaging techniques provides distinct elucidations on the brain and tumorous structures. T1-weighted images offer anatomical studies in detail, T2-weighted images are useful in detecting lesions, FLAIR images show lesions in proximity of CSF, and enhanced MRI shows the altered regions including tumor. The use of these different modalities essentially enables a more precise and comprehensive imaging of the tumor which in turn assists to enhance segmentation and classification results.

3.2. Image Normalization

Image normalization is an important process which guarantees the image intensity values to be consistent among different MRI modalities by deducting the image's information average and dividing by the standard deviation, Z-score normalization is carried out to each individual image. Since different modalities of MRI have different intensity distributions (T1 vs FLAIR vs T2), hence this helps in normalizing these intensity distributions. This ensures that the model treats each modality equally, without one modality overpowering others due to high intensity values. This leads to normalization making segmentation and classification models more robust and less sensitive.

3.3. CycleGAN (Image Synthesis)

Using CycleGAN for tumor synthesis in the BraTS 2019 dataset can augment data for segmentation and classification tasks by generating realistic tumor-affected or tumor-free images. The dataset is divided into two domains: tumor-affected MRIs (Domain A) and tumor-free MRIs (Domain B). By learning to produce realistic and high-quality images from other modalities, CycleGAN helps to enhance the multimodal dataset, leading to better feature extraction and more accurate segmentation. CycleGAN is more efficient than traditional data augmentation techniques for tasks like tumor synthesis because it generates domain-specific, diverse, and realistic images while handling unpaired data. Unlike conventional methods that apply random transformations, CycleGAN learns mappings between domains, ensuring that synthetic data preserves critical medical features such as tumor shape and intensity.

3.4. Image Registration

The MRI scans from various sequences (T1, T2, FLAIR) might not be perfectly aligned due to various factors, including but not limited to patient movement or slight differences in the timings of the scans. Registration ensures that all corresponding anatomical features across these modalities are properly aligned. This is essential for performing the modality integration correctly and spatially aligning the tumor and adjacent tissues within and across images. Proper registration sets the stage for precise segmentation by ensuring all modalities contribute aligned data. In this work cross-modality registration is used which involves aligning images from different modalities (e.g., T1 and FLAIR). Firstly, affine registration is done which allows translation, rotation, scaling, and shearing. Affine registration is a linear transformation method that aligns the overall shape and orientation of the images resulting in an image similar to brain image. This step will bring the images closer together by matching global features. After affine registration, apply non-rigid registration which is a nonlinear transformation that adjusts local variations in an image. In this work B-Spline transform which is a smooth grid-based deformation is used that refines the alignment, especially in areas affected by local deformations, such as tumor regions.

3.5. Early fusion

The multimodal fusion step is essential in this study for combining the complementary information from the different MRI modalities, namely T1, T1ce, T2, and FLAIR. Each modality captures different aspects and fusion of these modalities helps in creating a more comprehensive representation of the brain and its associated tumors. Once the images are aligned, early fusion is performed by stacking the registered modalities (e.g., T1, T2, FLAIR, T1ce) as separate channels or input features for a ML model This early fusion of multimodal information allows the model to learn richer features and improve segmentation accuracy, as each modality provides unique information that aids in differentiating between tumor and non-tumor regions. To fuse the images concatenation is done which involves combining the multi-modal images (e.g., T1ce and FLAIR) along the channel axis (depth axis). This allows the model to learn unique properties from each modality by maintaining each modality as a separate channel.

3.6. Attention U-Net

As shown in figure 2, once the images are fused using early fusion technique, image segmentation is carried out. An enhanced model that focuses on image segmentation issues is the Attention U-Net. It adds attention gates to the basic U-Net structure which enables the network to concentrate on the tumor, while ignoring other background information which is not relevant. This is significant in the framework of brain tumors segmentation because brain tumors are distinctly shaped and do not have clear boundaries. The attention mechanism thus enables the model to underscore such distinctive features while the effect of noise or other artifacts is greatly reduced. As a consequence, the system is more proficient and precise in performing tumor segmentation procedures from the given multiple MRI scans. With the addition of attention gate in the skip connections, Attention U-Net is also known as a modified version of the original U-Net design. These attention gates facilitate the model to place emphasis on the relevant input images features containing the tumor by ignoring non-important ones of healthy brain tissue. This form of segmentation is especially

 ISSN 2723-6471

useful when dealing with the segmentation of a brain tumor with patterns of irregular shape, poorly defined border and other parts surrounding the tumor.

Figure 2. Task of Brain Tumor Segmentation Using Attention U-Net Model

The model was designed with several key features. It utilizes a standard U-Net architecture, which includes an encoder for feature extraction and a decoder for generating the segmentation mask. The skip connections between the encoder and decoder are enhanced with attention gates. These gates help the model focus more accurately on specific regions of the tumor while minimizing the influence of irrelevant areas, resulting in more precise tumor boundaries.

As input, the model uses a fused multi-channel image that integrates four MRI modalities: T1, T1c, T2, and FLAIR. This fusion allows the network to leverage the complementary information provided by each modality, thereby enhancing the overall segmentation performance.

3.7. Morphological Segmentation

The Attention U-Net model's output is further refined through the application of morphological segmentation. The morphological operations involved are dilation, erosion, opening, and closing that improve the tumor mask, which has been segmented by removing small noise, filling gaps, and smoothing tumor boundaries. The aim is to obtain a more continuous and accurate tumor mask that closely reflects the actual boundaries of the tumor, which is essential for the accurate quantification of the tumor and further clinical evaluation. Morphological processes including dilatation, erosion, and closure have been used in this work. Firstly, dilation expands the tumor mask to ensure any missed tumor areas are included. Erosion follows to eliminate small noise and irrelevant structures, cleaning the mask from artifacts. Closing is finally used to fill small holes or holes within the tumor mask; the result should be an unbroken, complete region. This approach enhances accuracy and quality of the result for further classification analysis as well.

3.8. Transformer-based Classification

Once the tumor area is segmented, the model of classification uses transformer to categorize the tumor as malignant or benign. Two steps are involved in the hybrid CNN and Transformer-based model for tumor classification. The segmented tumor is passed through CNN layers for extracting fine-grained local features such as shape, texture, and intensity. The local feature maps are further fed into a transformer model, which captures the global dependencies and contextual relationships among different tumor regions and surrounding tissues by utilizing its self-attention mechanism. For instance, the transformer gathers and contextualizes these features before sending the representations via a fully connected layer to identify if a tumor is benign or malignant. In this hybrid approach, it would be possible to better integrate local feature extraction with the global understanding of context while achieving improved accuracy and robustness in tumor classification. With the use of self-attention mechanisms, this model processes the features that were taken from the tumor's segmented areas. The ability of the transformer to pay special attention to important regions handles complex tumor characteristics and subtle differences, thus improving the accuracy in classification. This step has useful information about the kind of tumor and its characterization, which is necessary in clinical decision-making for treatment and prognosis.

The final output of the pipeline consists of two primary results: the segmented tumor mask and the tumor classification. The segmented mask provides a precise delineation of the tumor regions within the brain, which is essential for surgical planning, treatment, and ongoing monitoring. The tumor classification indicates whether the tumor is benign or malignant, providing key insights into the diagnosis and prognosis. When put together, these results provide a thorough examination of the tumor, assisting medical professionals in making defensible choices regarding patient care and therapeutic approaches.

4. Performance Evaluation

The BRATS 2019 dataset was used to test the suggested AU-BTS framework, and measures like accuracy, dice similarity coefficient, intersection over Union, precision, recall, F1-Score, and Area Under the Curve were used to assess the outcomes. Table 1 show the summary of dataset used in this study.

Table 1. Key details of BRATS 2019 dataset

Feature	Details
Number of patients	259 patients
Number of Images	4 MRI modalities (T1, T1ce, T2, FLAIR)
Format	NIfTI format (Neuroimaging Informatics Technology Initiative
Total Images	1036 images (259 x 4)
Resolution	240 x 240 x 155 voxels per image
Tumor types	Low-grade, High-Grade Gliomas
Annotations	No tumor/healthy, enhancing tumor, edema, necrotic core

By offering multimodal MRI images for segmentation algorithm training and evaluation, the BRATS 2019 dataset aims to facilitate research in brain tumor segmentation, namely gliomas. Images from the T1, T1c, T2, and FLAIR sequences were gathered from 259 patients and contain annotations for tumor locations like the necrotic core, enhancing tumor, edema, and tumor core. The dataset, which includes high-resolution 3D scans of both low-grade and high-grade gliomas, is utilized to build models that precisely segment and categorize different tumor sub-regions. The implementation of the framework is done using Google Colab with TPU v2-8: 8 cores, 64 GB HBM memory and python programming language.

4.1. Performance Metrics

4.1.1. Accuracy (for Classification)

One of the most often utilized performance indicators in classification activities is accuracy as shown in equation (1). It is computed as the proportion of accurately predicted tumor types (malignant or benign) to all forecasts. Accuracy is important in the categorization of brain tumors because it influences how well the model performs overall in differentiating between benign and malignant tumors. For clinical decision-making, the more accurate the model, the more trustworthy it is.

$$\text{Accuracy} = \frac{\text{Total Number of Predictions}}{\text{True Positives+True Negatives}} \tag{1}$$

4.1.2. Dice Similarity Coefficient (DSC) (for Segmentation)

The Dice Similarity Coefficient (DSC) is a statistic used to measure the similarity between two sets, in this case, the ground truth mask and the predicted segmentation mask. DSC helps evaluate how well the model identifies tumor regions compared to the actual tumor boundaries in brain tumor segmentation using equation (2). A DSC value closer to 1 indicates better segmentation, as it signifies a high overlap between the predicted and true tumor areas.

$$\text{DSC} = \frac{2\times(\text{Intersection of Predicted and Ground Truth})}{\text{Area of Predicted+Area of Ground Truth}} \tag{2}$$

4.1.3. Intersection Over Union (IoU) (for Segmentation)

Intersection over Union is the other very popular evaluation metric in segmentation model performance as shown in equation (3). It computes how much the overlap of a predicted tumor mask with its ground truth mask occupies its union with it. In simple terms, it quantifies the degree of overlap between the predicted segmentation and the true segmentation. Higher IoU values indicate that the model is better at predicting tumor regions with fewer false positives or false negatives.

$$\text{IoU} = \frac{\text{Intersection of Predicted and Ground Truth}}{\text{Union of Predicted and Ground Truth}} \tag{3}$$

4.1.4. Precision and Recall (for Classification)

Two important criteria for assessing classification performance are precision and recall, as shown in equations (4) and (5) especially in datasets that are unbalanced and may have a dominant class. Precision quantifies the proportion of tumors that are truly malignant (or benign) compared to those that were anticipated to be such. When false positives can be expensive, such as in medical diagnosis were misclassifying a benign tumor as malignant could result in needless therapy, this statistic is essential. Recall measures how many of the actual malignant (or benign) tumors have been correctly identified by the model. In the medical field, recall is vital because failing to identify a malignant tumor (false negative) could lead to missed diagnoses and delay in treatment.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}} \tag{4}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}} \tag{5}$$

4.1.5. F1 Score (for Classification)

The F1 score offers a fair assessment of a model's performance since it is the harmonic mean of precision and recall. Because it integrates precision and recall into a single statistic, it is particularly helpful in situations where the distribution of classes is not uniform. A high F1 score means that the model does a good job at reducing false negatives and false positives which helps the clinicians by reducing the risk for unnecessary intervention or missed diagnosis.

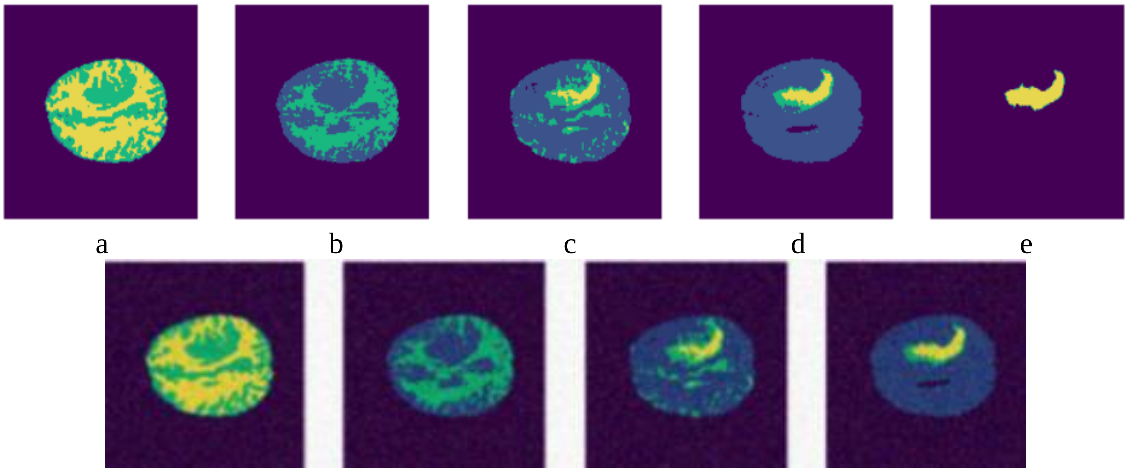
$$\text{F1 Score} = \frac{2\times(\text{Precision}\times\text{Recall})}{\text{Precision+Recall}} \tag{6}$$

4.1.6. Area Under the Curve (AUC) for ROC (Receiver Operating Characteristic)

The AUC of the ROC curve measures the model's ability to discriminate between the different classes, typically benign and malignant. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1 - specificity), and the AUC quantifies the area under this curve. A higher AUC indicates that the model has a better ability to distinguish between benign and malignant tumors.

4.2. Performance Analysis

The input images of the dataset are normalized using the z score normalization and then the cycleGAN is utilized to generate the synthetic dataset. With the help of cycleGAN the dataset consisting of 1036 images of size 240 x 240 were doubled to 2072 images of size 240 x 240. After data augmentation, cross modality registration was applied to the images to ensure the proper alignment of features from all the modalities. Early fusion is done on the registered images using the concatenation technique to stack all the modalities together. Here only T1ce and FLAIR modalities are concatenated as they give more information about the tumor compared to T2 and T1 images. The fused image is given as input to the attention U-Net model which will segment the abnormal portion from the image. Finally, a hybrid transformer comprising of CNN and transformer is employed to extract the features from the segmented image and feed those to the transformer to classify as tumor or no tumor class. Figure 3 shows the different modalities of the MRI Scan image such as T1 image, T1ce image, T2 image, FLAIR and ground truth segmented image along with CycleGAN generated images. The augmented images differ in terms of intensity, texture and noise ensuring data diversity of the dataset.



a b c d e

f g h i

Figure 3. (a) T1 image, (b) T1ce Image, (c) T2 image, (d) FLAIR image, (e) Segmented Mask, (f) CycleGAN T1, (g) CycleGAN T1ce, (h) CycleGAN T2, (i) CycleGAN FLAIR

Figure 4 and figure 5 shows the color map of the segmented output comprising of the regions such as no tumor, necrotic tumor, edema and enhancing tumor.

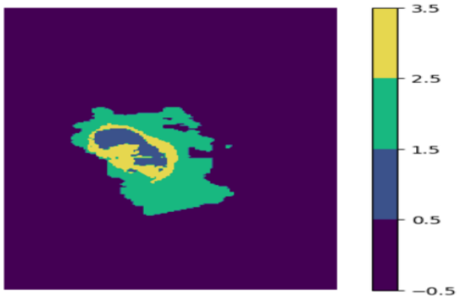
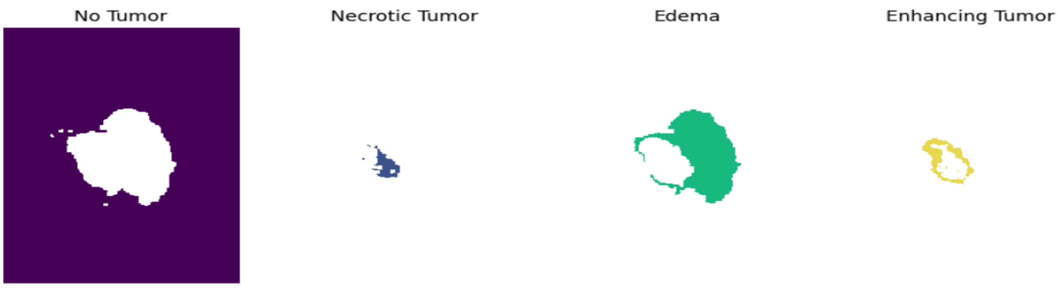


Figure 4. Color map of the segmented output

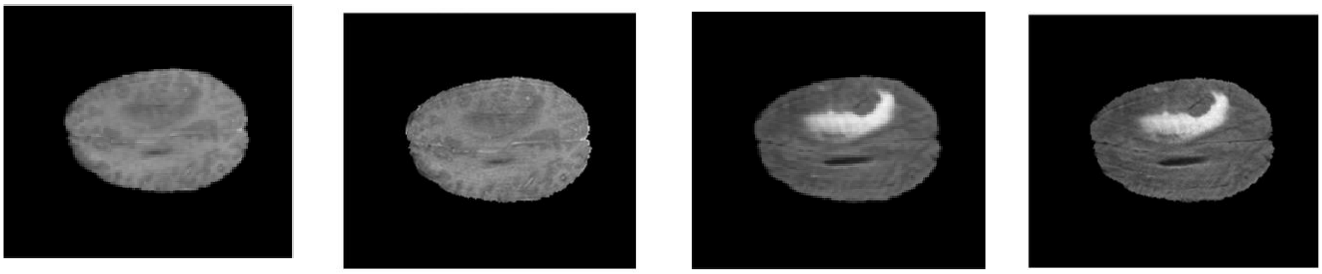


(a) (b) (c) (d)

Figure 5. (a) No tumor, (b) Necrotic tumor, (c) Edema, (d) Enhancing tumor

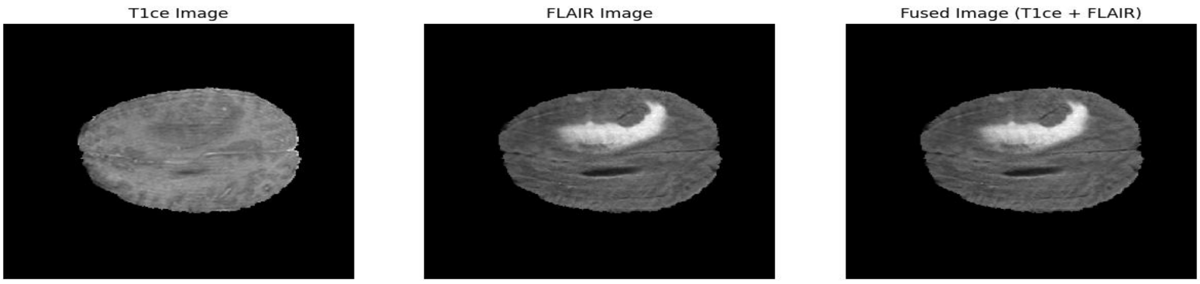
Figure 6 shows the normalized output of the T1ce and FLAIR images which will then be used for early fusion. Figure 7 depicts the fused result of the T1ce and FLAIR images. Figure 8 shows the fused input which is fed to the attention U-Net to predict the segmented output.

ISSN 2723-6471



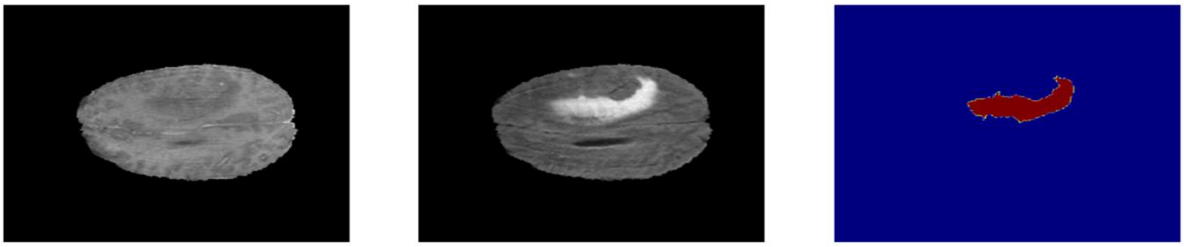
(a) (b) (c) (d)

Figure 6. (a) T1 Image, (b) T1 normalized Image, (c) T1ce Image, (d)T1ce Normalized



a b c

Figure 7. (a) T1ce, (b) FLAIR, (c) Fused Image



a b c

Figure 8. (a) T1ce, (b) FLAIR, (c) Predicted Mask

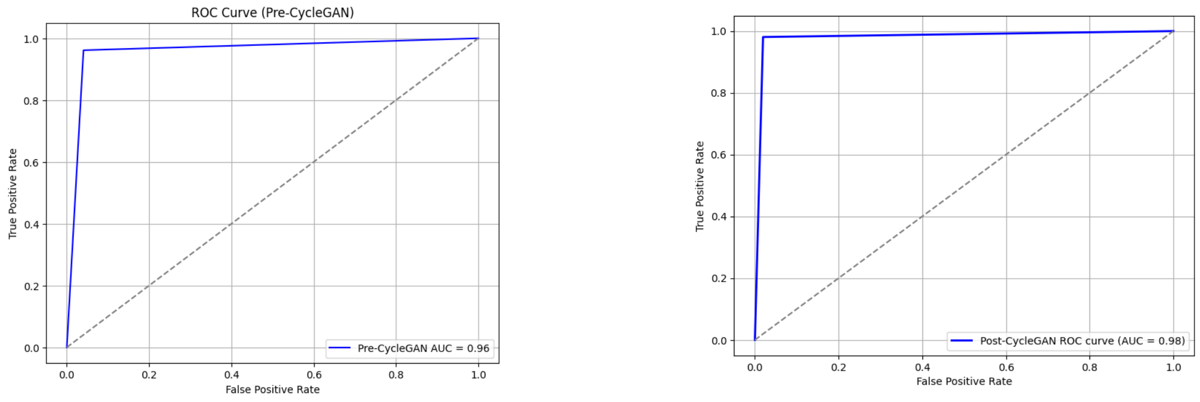
From figure 7 the shape of the fused image was found to be (240, 240, 155, 2) and the size of the Input data shape for Attention U-Net was (1, 240, 240, 155, 2). Table 2 shows the performance analysis of the proposed framework in terms of accuracy, F1 Score, IoU, Dice coefficient and AUC. The analysis was carried out with the initial dataset and then on the augmented dataset.

Table 2. Performance analysis of the proposed AU-BTS framework with existing work

Framework	Accuracy	F1_Score	IoU	Dice Coefficient (DC)	AUC
AU-BTS Without CycleGAN	96 %	0.961	0.930	0.960	0.960
AU-BTS With CycleGAN	98 %	0.979	0.974	0.986	0.980
U-Net-based model [15]	97 %	0.945	0.900	0.920	0.950
AlexNet-based model [10]	95 %	0.910	0.870	0.890	0.930

Table 2 shows that the suggested framework performs better with cycleGAN, exhibiting an accuracy of almost 98%, F1 score of 0.979, IoU of 0.974, DC of 0.986, and AUC of 0.98 compared with existing models. Therefore, a higher DSC in this experiment suggests that the Attention U-Net, when using multimodal data, is more effective at learning to segment the tumor region. The results show how efficiently the model can assist the clinicians in reducing the missed diagnosis and plan for further evaluation. The ROC curve for the suggested framework with and without data augmentation is displayed in figure 9.

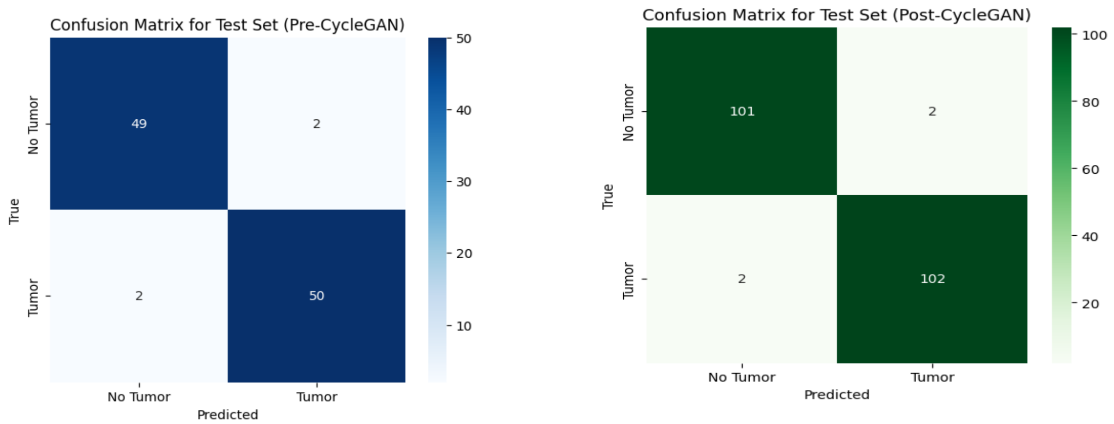
ISSN 2723-6471



a b

Figure 9. ROC Curve for the proposed AU-BTS framework with and without data augmentation

From the figure 9 it is observed that the AUC for AU-BTS with cycleGAN is 0.98 and before augmentation is 0.96 which shows the efficacy of the proposed framework with cycleGAN. Figure 10 shows the confusion matrix for AU-BTS with and without cycleGAN.



(a) (b)

Figure 10. Confusion matrix for the proposed AU-BTS framework with and without data augmentation

From figure 10 it is observed that an accuracy of 96% and 98 % indicates that the model is correctly classifying the majority of instances, with only 4% and 2 % of predictions being incorrect. This suggests the model is performing well overall. Figure 11 shows the analysis for loss vs epochs for 30 epochs.

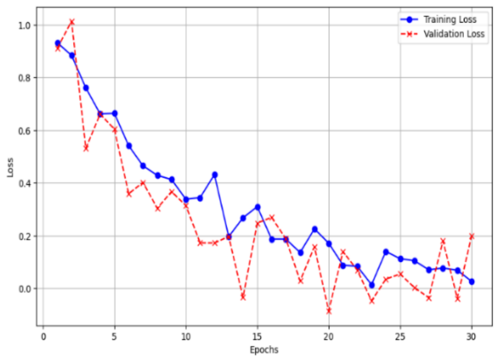


Figure 11. Loss Vs Epochs for the proposed AU-BTS model

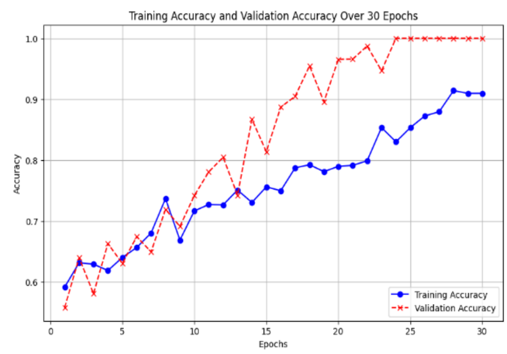


Figure 12. Accuracy Vs Epochs for the proposed AU-BTS model

The training and validation loss decreases with the increasing epochs which shows that the model is performing well by learning significant patterns from the data by adjusting its hyperparameters efficiently. Since loss is gradually decreasing it means that the model does not suffer from overfitting and poor convergence. Figure 12 shows the analysis for accuracy vs epochs for 30 epochs. The training and validation accuracy increases with the increasing epochs which shows that the model is stable. Since both training and validation accuracy increases gradually it means the model is generalizing well to unseen data. For training the time taken is 2 to 3 hours and for inferencing the model took about 1 minute which proves the efficacy of the model in real time scenarios. Table 3 shows the comparison of the epoch's vs loss, training accuracy, and validation accuracy for a model trained on the BraTS 2019 dataset before and after using CycleGAN for data augmentation.

Table 3. Loss and Accuracy analysis for the proposed AU-BTS framework with and without data augmentation

Epochs	Before Augmentation			After Augmentation		
	Loss	Training Accuracy	Validation Accuracy	Loss	Training Accuracy	Validation Accuracy
5	0.53	75%	73%	0.48	80%	78%
10	0.42	83%	81%	0.35	88%	86%
15	0.36	89%	87%	0.28	92%	91%
20	0.29	92%	91%	0.22	95%	94%
25	0.23	94%	93%	0.16	97%	96%
30	0.18	96%	96%	0.11	99%	98%

From table 3 it is observed that the model experiences slower convergence, with the loss decreasing and training accuracy increasing gradually. Validation accuracy tends to be lower at the start due to limited data. After Augmentation with CycleGAN the model benefits from the larger and more varied training dataset, resulting in faster reduction in loss and quicker improvements in both training and validation accuracy. The increased data helps the model generalize better, improving validation accuracy and reducing overfitting. Hence the CycleGAN augmentation helps the model learn more efficiently, achieving better generalization and faster convergence compared to training on the original dataset alone. The training accuracy improves faster, and validation accuracy stabilizes at a higher level due to the more diverse and abundant training data. The proposed model achieved an overall accuracy of approximately 98% after data augmentation, which is higher than or comparable to existing models, such as the U-Net-based model [15] with 97% accuracy, the deep learning-based model [11] with 92% accuracy, and the AlexNet-based model [10] with 95% accuracy.

5. Conclusion and Future work

The integration of Attention U-Net with hybrid transformer and the use of CycleGAN-based data augmentation provide a unique approach for brain tumor segmentation and classification as it enables better feature fusion across modalities and better data diversity. The Attention U-Net and hybrid Transformers are utilized to extract the global tumor segmentation efficiently from distinguished regions and perform early fusion of multiple modalities of data respectively for effective feature extraction. The CycleGAN based data augmentation mitigates the challenge of limited training dataset by increasing the diversity and quality of the dataset. Unlike traditional models, the proposed AU-BTS framework is able to achieve better accuracy as it leverages both spatial attention and cross modalities. The results indicate that the proposed framework is effective in tumor classification resulting in high accuracies achieving around 98 %, as well as segmentation resulting in higher DSC of 0.986 and IoU 0.974 values. The finding has been promising in providing clinicians with accurate diagnosis of brain tumors and their treatment approaches.

The proposed framework, although highly promising, has a few areas that could be explored further for future research. One possible direction is investigating more sophisticated data augmentation strategies such as adversarial training, style transfer augmentation, and Contrastive Learning-Based Augmentation to improve model generalization particularly for rare or complicated tumor types. In addition, 3D-CNN can be used for 3D MRI data which could leverage richer spatial information, possibly leading to improved accuracy for both segmentation and classification. Future work may also consider focusing on real-time application, where the efficiency and performance of the model would be paramount in clinical settings. Additionally, the inclusion of different imaging modalities, for instance, PET

ISSN 2723-6471

or CT scans, can enhance the system's capability in distinguishing benign from malignant tumors. Also curating a dataset that includes intermediate cases will enhance the model's ability to handle tumors with uncertain characteristics. Ultimately, further research into the explainability and interpretability of the model's decisions may foster greater confidence and adoption by clinicians, aiding larger integration into clinical workflows.

6. Declarations

6.1. Author Contributions

Conceptualization: R.K.M., A.N.S., and M.B.; Methodology: A.N.S.; Software: R.K.M.; Validation: R.K.M., A.N.S., and M.B.; Formal Analysis: R.K.M., A.N.S., and M.B.; Investigation: R.K.M.; Resources: A.N.S.; Data Curation: A.N.S.; Writing Original Draft Preparation: R.K.M., A.N.S., and M.B.; Writing Review and Editing: A.N.S., R.K.M., and M.B.; Visualization: R.K.M. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426–1438, Oct. 2013

[2] V. Harati, R. Khayati, and A. Farzan, "Fully automated tumor segmentation based on improved fuzzy connectedness algorithm in brain MR images," *Computers in Biology and Medicine*, vol. 41, no. 7, pp. 483–492, Jul. 2011

[3] M. Sharif, U. Tanvir, E. U. Munir, M. A. Khan, and M. Yasmin, "Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. Jan., pp. 1063–1082, Oct. 2018

[4] K. Pani and I. Chawla, "A hybrid approach for multi modal brain tumor segmentation using two phase transfer learning, SSL and a hybrid 3DUNET," *Computers & Electrical Engineering*, vol. 118, no. Jul., pp. 109418–109418, Jul. 2024

[5] Fethi Ghazouani, P. Véra, and S. Ruan, "Efficient brain tumor segmentation using Swin transformer and enhanced local self-attention," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, no. Feb., pp. 273–281, Oct. 2023

[6] T. Magadza and S. Viriri, "Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art," *Journal of Imaging*, vol. 7, no. 2, pp. 19-32, Jan. 2021

[7] N. Pradhan, S. Sagar, and Ajay Shankar Singh, "Analysis of MRI image data for Alzheimer disease detection using deep learning techniques," *Multimedia Tools and Applications*, vol. 83, no. 6, pp. 17729–17752, Jul. 2023

[8] M. G. Alsubaie, S. Luo, and K. Shaukat, "Alzheimer's Disease Detection Using Deep Learning on Neuroimaging: A Systematic Review," *Machine Learning and Knowledge Extraction*, vol. 6, no. 1, pp. 464–505, Mar. 2024

[9] R. Sharma and H. K. Meena, "Utilizing graph Fourier transform for automatic Alzheimer's disease detection from EEG signals," *International Journal of Information Technology*, vol. 16, no. 3, pp. 1653–1659, Jan. 2024

[10] S. Siuly, S. K. Khare, E. Kabir, M. T. Sadiq, and H. Wang, “An efficient Parkinson’s disease detection framework: Leveraging time-frequency representation and AlexNet convolutional neural network,” *Computers in Biology and Medicine*, vol. 174, no. May, pp. 1-22, May 2024

[11] Y. Djenouri, A. Belhadi, A. Yazidi, G. Srivastava, and J. C. Lin, “Artificial intelligence of medical things for disease detection using ensemble deep learning and attention mechanism,” *Expert Systems*, vol. 41, no. 6, pp. 1-13, Jun. 2022

[12] K. Li, B. Ao, X. Wu, Q. Wen, Ejaz Ul Haq, and J. Yin, “Parkinson’s disease detection and classification using EEG based on deep CNN-LSTM model,” *Biotechnology and Genetic Engineering Reviews*, vol. 40, no. 3, pp. 2577–2596, Apr. 2023

[13] A. Shrivastava, M. Chakkaravarthy, and M. Asif Shah, “A Novel Approach Using Learning Algorithm for Parkinson’s Disease Detection with Handwritten Sketches’,” *Cybernetics and Systems*, vol. 55, no. 8, pp. 2388–2404, Jan. 2023

[14] W. Li, W. Huang, and Y. Zheng, “CorrDiff: Corrective Diffusion Model for Accurate MRI Brain Tumor Segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1587–1598, Jan. 2024

[15] S. Gajula and V. Rajesh, “An MRI brain tumour detection using logistic regression-based machine learning model,” *International Journal of System Assurance Engineering and Management*, vol. 15, no. Jul., pp. 124-134, Jul. 2022

[16] Georges Dubourg et al., “Advancement of metal oxide nanomaterials on agri-food fronts,” *Science of the total environment*, vol. 928, no. Jun., pp. 172048–172048, Apr. 2024

[17] V. Kumar, P. Chugh, B. Bharti, A. Bijalwan, A. Tripathi, R. Narayan, and K. Joshi, "MRI-Based Brain Tumor Detection Using Machine Learning," in *Human Cancer Diagnosis and Detection Using Exascale Computing, K. Joshi and S. K. Gupta, Eds. 2024.*

[18] Tin Tin Ting, X. Ke, A. Aitizaz, Lee Kuok Tiung, Chong Keat Teoh, and H. Sarwar, “Machine Learning based Predictive Modelling of Cybersecurity Threats Utilising Behavioural Data,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 9, pp. 1-19, Jan. 2023

[19] B. H. Hayadi and T. Hariguna, “Predictive Analytics in Mobile Education: Evaluating Logistic Regression, Random Forest, and Gradient Boosting for Course Completion Forecasting”, *Int. J. Interact. Mob. Technol.*, vol. 19, no. 05, pp. 210–232, Mar. 2025.

[20] A. Ruangkanjanases and T. Hariguna, "Exploring the synergy of guided numeric and text analysis in e-commerce: a comprehensive investigation into univariate and multivariate distributions," *PeerJ Computer Science*, vol. 10, no. Sept., pp. 1-23, 2024.