International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Enhanced U-Net Architecture for Accurate Brain Tumor Segmentation in MRI Scans

Saritha Dasari[1]
Research Scholar, JNTUA, Ananthapuramu, Andhra Pradesh, India.
.e-mail: sarithadasari123@gmail.com

Dr.A.Rama Mohan Reddy[2]
[2]Professor, Department of CSE, S.V. University, Tirupati, Andhra Pradesh, India.
e-mail: ramamohansvu@yahoo.com

Dr.B.Eswara Reddy[3]
[3]Professor, Department of CSE, JNTUA,Aanantapuramu, Andhra Pradesh, India
e-mail: eswarcsejntua@gmail.com

Abstract—Precision in medical image segmentation is vital for accurate diagnosis and treatment planning in the contemporary healthcare system. Deep learning techniques, such as Convolutional Neural Networks (CNNs), U-Net architectures (UNETs), and Transformers, have revolutionized this field by automating laborious manual segmentation processes that were previously performed manually. Nevertheless, challenges such as intricate structures and indistinct features persist, leading to accuracy concerns. Scientists are diligently striving to address these challenges to fully harness the promise of medical image segmentation in the healthcare revolution. The objective of our study is to introduce an improved version of the U-Net model that is specifically tailored for the segmentation of brain cancer MRI images. The primary aim of this enhancement is to boost the precision of the segmentation process. Our plan consists of three main components. Initially, our main focus is on enhancing features by employing techniques such as Contrast Limited Adaptive Histogram Equalisation (CLAHE) during the image preprocessing stage. Furthermore, we enhance the architecture of the U-Net model by prioritizing a tailored layered design to enhance the quality of segmentation results. Ultimately, we employ a Convolutional Neural Network (CNN) model for post-processing to enhance segmentation results by utilizing additional convolutional layers. Our model was tested, validated, and trained using a total of 3,064 brain MRI images, with 612 images used for testing, 612 images used for validation, and 1,840 images used for training. We achieved outstanding results in terms of recall (93.66%), accuracy (97.79%), F-score (93.15%), and precision (92.66%). The Dice coefficient's training and validation curves exhibited minimal variability, with training achieving approximately 93% and validation reaching 84%, indicating a strong capacity for generalization. The high accuracy of the segmentation findings was confirmed through visual assessment, however rare errors such as false positives were observed.

Keywords-*Medical Image Segmentation; UNet; MRI Images;Health care; Deep Learning*

I. INTRODUCTION

Brain tumors provide a significant health risk and are potentially fatal at any point of discovery (Nehra, 2021). Brain cancer affects people of all ages and genders, with over 100 different types of brain tumours diagnosed and classed as primary or metastatic (Zhang A. S.-S., 2017). Primary brain tumors can be malignant or non-cancerous, and they can develop within or around brain structures. Secondary (metastatic) tumors, on the other hand, are typically malignant and originate in other parts of the body before spreading to the brain. Survival rates for primary brain tumors are determined by several factors, including age, location, ethnicity, tumor type, and molecular characteristics (Kiran, 2024).

Because of its radiation-free nature, MRI (Magnetic Resonance Imaging) (Abdelatty, 2024) is now the gold standard for evaluating a wide spectrum of cerebral pathological disorders. MRI is more effective than CT for detecting acute ischemic lesions (Mahajan, 2024). However, the usefulness of MRI in emergencies is limited due to its relatively long acquisition time, making timely and correct diagnosis critical (Altmann, 2024). Brain MRI segmentation is a critical medical imaging method that divides the brain's MRI into separate sections of structure (Kumar P. R., 2024). This method is crucial for the diagnosis, treatment planning, and monitoring of neurological illnesses such as multiple sclerosis, brain tumors, Alzheimer's disease, and other abnormalities of the brain (Grgec, 2024; Desale, 2024). Similarly, MRI is essential for quantitative brain volumetry assessments in prenatal diagnosis and studying early human brain development (Ciceri, 2024).

One of the most pressing issues in clinical research approaches is segmenting MRI pictures of the human brain. In many image analysis applications, segmenting medical images is critical (Yellu, 2024). In medical picture analysis, automated segmentation algorithms outperform hand segmentation. Automated brain tissue segmentation based on clinically obtained MRI data is a significant phase in quantitative brain analysis. It allows for more accurate quantitative assessments of the brain, which aids in disease identification, diagnosis, and classification. Thus, the success of the segmentation technique is critical for disease identification and treatment planning (Kumar, 2024). The segmentation process seeks to

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

distinguish and define a variety of anatomical structures, including white matter (WM), grey matter (GM), cerebrospinal fluid (CSF), and diseased regions. Automated segmentation of perinatal brain

MRI remains difficult due to significant changes in the brain's global structure and high variations in image intensity, reflecting the rapid tissue maturation that occurs around birth. Accurate and automated brain segmentation is crucial for several brain analysis activities, particularly those involving brain tumors, such as high-resolution reconstruction and cortical surface analysis. Although AI systems for medical picture segmentation can achieve superhuman accuracy on average, many radiologists are skeptical (Fidon, 2024).

This skepticism stems in part from the possibility that AI algorithms would malfunction and produce errors that contradict human knowledge about the segmentation task, particularly when applied across many imaging protocols and anatomical diseases (Cai, 2024). Traditional diagnostic methods are time-demanding and error-prone (Solanki, 2023). The demand for effective medical image segmentation algorithms has increased as the specialized area requires more experienced personnel. Traditional image segmentation approaches, such as thresholding (Sharif, 2024), edge-based (Al Garea, 2024), and region-based methods (Reddy, 2024), are challenged by limitations in medical image capture, pathology kinds, and biological variability (Soppari, 2024). Brain MRI segmentation is extremely difficult because of image procurement limitations, the nature of brain pathology, and biological variability. Another problem in MRI segmentation is dealing with opacity in pixel values, which intuitionistic-based clustering algorithms try to overcome (Arora, 2024).

Our study uses U-Net for brain MRI segmentation because it has numerous advantages over traditional deep learning (DL), machine learning (ML), and clustering methods. U-Net's encoder-decoder design handles the complicated structures of medical pictures, allowing for exact feature location and segmentation. Unlike traditional machine learning and clustering techniques, which rely on handmade features and extensive manual tuning, U-Net automatically learns relevant features via its convolutional layers. The skip connections in U-Net are especially useful for preserving spatial information, resulting in more accurate segmentation results, particularly in complex and varied anatomical structures. This makes U-Net especially useful for brain MRI and fetal brain segmentation, which present significant hurdles due to motion artifacts and substantial anatomical variability.

II. LITERATURE REVIEW

Shahzad and colleagues (Shahzad, 2022) made advancements in medical picture segmentation using neural networks by suggesting a modified U-Net architecture for use with the Intel/Movidius Neural Compute Stick 2 (NCS-2). The rationale for using U-Net is its proven effectiveness in medical picture segmentation tasks, particularly when dealing with small dataset sizes. Their modified U-Net variation achieved a substantial reduction in the number of parameters, decreasing it from 30 million in the original U-Net to 0.49 million. The goal was to enhance resource utilization while maintaining performance. Encouraging outcomes were achieved through experimental assessments carried out on three distinct medical imaging datasets: Ziehl-Neelsen sputum smear microscopy (ZNSDB), heart MRI, and brain MRI (BraTs). The suggested approach achieved maximum dice scores of 0.96, 0.94, and 0.74 for the BraTs, heart MRI, and ZNSDB datasets, respectively. These results demonstrate good performance and enable effective inference on the NCS-2 platform.

Akter et al. (Akter, 2024) introduced a new deep Convolutional Neural Network (CNN) architecture for automatically categorizing brain images into four distinct classes. They also utilized a U-Net-based segmentation model to compare the impact of segmentation on tumor classification in brain MRI images. The researchers conducted comprehensive evaluations using six established datasets. Two classification techniques were evaluated using accuracy, recall, precision, and AUC parameters, and they outperformed pre-trained alternatives with remarkable performance on all datasets. More precisely, their classification model achieved an accuracy of 98.7% in a dataset that was combined, and 98.8% when it was integrated with segmentation. The highest classification accuracy reached an impressive 97.7% across all four separate datasets, highlighting the effectiveness of their strategy in improving brain tumor classification and segmentation using deep learning techniques.

Shiny et al. (Shiny, 2024) proposed an optimization-based method for classifying brain tumors using MRI data. The methodology entailed the utilization of filtering techniques and Region of Interest (RoI) extraction to preprocess pre-operative and post-operative MRI images. Tumour segments were generated from the preprocessed data by segmenting it using an adapted U-Net model. Afterward, histogram characteristics were obtained, and tumor classification was carried out using a U-Net model trained with the proposed Poor Bird Swarm Optimisation algorithm (PRBSA), which is a combination of the Poor and prosperous optimization (PRO) algorithm and Bird Swarm Algorithm (BSA). Ultimately, the process of detecting changes in pixels utilized the speeded-up robust features (SURF) algorithm on the classed output. The U-Net model, based on the pseudorandom binary Sequence Algorithm (PRBSA), demonstrated outstanding performance in tumor classification tasks. It achieved a remarkable accuracy of 94%, sensitivity of 93.7%, and specificity of 94%. These results highlight the model's robustness in optimizing the diagnosis of brain tumors using MRI data.

Rutoh et al. (Rutoh, 2024) proposed a new approach called GAIR-U-Net, which is a 3D Guided Attention-based deep Inception Residual U-Net. This method was intended to tackle the difficulties in segmenting tumors from multimodal MRI data. The GAIR-U-Net utilized attention mechanisms, an inception module, and residual blocks with dilated convolution to enhance feature representation and spatial context

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

comprehension. The U-Net design was developed by incorporating inception and residual connections to effectively capture intricate patterns and hierarchical features. Additionally, the model was able to expand its width in three-dimensional space without a substantial increase in computing complexity. Dilated convolutions not only facilitated the acquisition of both local and global information but also improved segmentation accuracy and adaptability by giving priority to significant regions and suppressing irrelevant features. The experimental assessments conducted on the BraTS 2020 dataset, which consisted of T1-weighted, T1-ce, T2-weighted, and FLAIR sequences, demonstrated encouraging results. The GAIR-U-Net achieved dice scores of 0.8796, 0.8634, and 0.8441 for the total tumor (WT), tumor core (TC), and improving tumor (ET) on the BraTS 2020 validation dataset.

Chen et al. (Chen, 2024) proposed the Adaptive Cascaded Transformer U-Net (ACTransU-Net), a distinctive framework designed for the segmentation of brain tumors in MRI images. ACTransU-Net integrates a Transformer and dynamic convolution into a cascaded U-Net architecture to accurately capture both overall characteristics and specific features of brain tumors. The architecture implemented a two-stage approach, involving the sequential use of two 3D U-Nets for coarse-to-fine segmentation. Subsequently, the second-stage shallow encoder and decoder were enhanced by using omni-dimensional dynamic convolution modules. This integration aimed to enhance the representation of local details by dynamically adjusting the parameters of the convolution kernels. In addition, the second-stage deep encoder and decoder incorporate 3D Swin-Transformer modules to effectively capture long-range dependencies in the images. This improves the overall depiction of brain tumors at a global level. The efficacy of ACTransU-Net was demonstrated by experimental outcomes on the BraTS 2020 and BraTS 2021 datasets. The model achieved average Dice Similarity Coefficient (DSC) scores of 84.96% and 91.37% respectively, along with 95th percentile Hausdorff Distance (HD95) values of 10.81 mm and 7.31 mm.

III. Proposed Methodology

This research introduces a novel approach for dividing brain MRI data into segments using a tailored U-Net model. This model employs extensive experimentation to optimize the arrangement of layers, resulting in exceptional segmentation outcomes. The subsequent sections offer a comprehensive elucidation of our pioneering methodology. Figure 1 illustrates the comprehensive technique employed in our investigation.

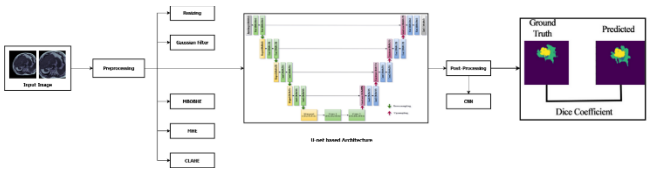


Figure 1: The architecture diagram for the proposed model

Dataset Description

The Kaggle dataset titled "Brain MRI Segmentation" has 3,064 brain MRI pictures, each meticulously paired with a corresponding mask that precisely identifies the location of any existing tumors. The extensive dataset is invaluable to medical imaging researchers and practitioners as it provides a robust foundation for developing, improving, and validating advanced machine-learning models for tumor recognition, segmentation, and analysis. Due to its comprehensive annotations, this dataset has significant potential to improve diagnostic accuracy and broaden the capabilities of automated medical imaging systems. Figure 2 displays a representation of the dataset sample.

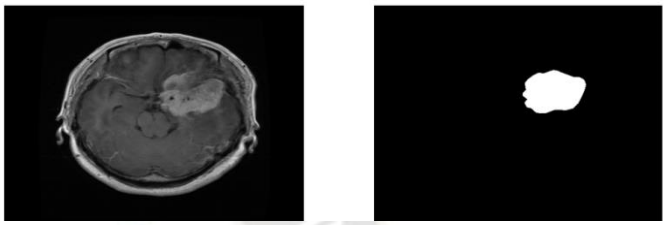


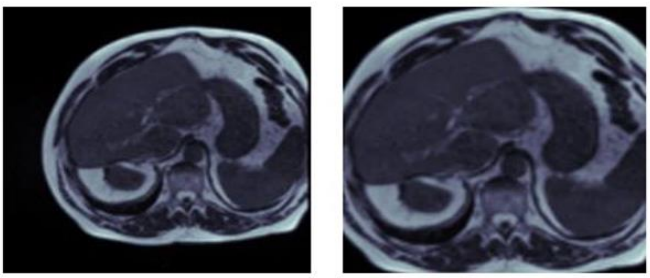
Figure 2: Sample of the dataset

Figure 2 (left) illustrates the original brain MRI image, while Figure 2 (right) displays the corresponding mask that delineates the tumor location within the same image.

Data Preprocessing

The pre-processing procedures for the dataset include resizing, filtering, normalization, and histogram equalization. A comprehensive explanation of each step is given below. Pre-processing is performed to enhance the data and render it more suitable for the following steps.

Resizing refers to the process of altering the dimensions of each photograph in the collection to a predetermined size. To ensure compatibility with the neural network, it is necessary to modify the dataset so that all photos have identical dimensions. Upon completion of the resizing process, the image's dimensions are 160 × 160 × 1. Figure 3 (a) displays an input image, while Figure 3 (b) displays the image that has been scaled.



(a) Input Image (b) Scaled Image

Figure 3: Dataset Resizing (First column: Input Image, Second column: Resized Image)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

A Gaussian filter is used to apply a low-pass filter to a picture to reduce noise, which refers to high-frequency components, and to blur specific portions of the image. The weights within the kernel, which are utilized to calculate the weighted mean of the nearest points (pixels) in a picture, follow a Gaussian distribution, as implied by the function's name. Figure 4 (a) displays a randomly chosen input image, while Figure 4 (b) shows the filtered image.

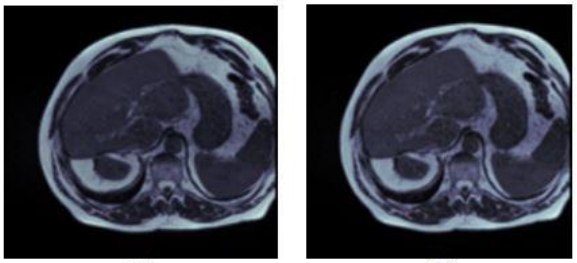


Figure 4: The filtered image is shown in Figures, (a) input image and (b) after filtering the input image.

Furthermore, before additional analysis or processing, preprocessing with CLAHE, MHE, and MBOBHE includes utilizing these image enhancement techniques to increase an input image's contrast and visibility. Every technique has advantages and qualities of its own.

Multi-Purpose Beta Optimized Bi-HE (MBOBHE): Despite its fame, the standard Histogram Equalisation (HE) approach struggles to achieve a uniformly distributed boost. MBOBHE aims to tackle this challenge by considering three essential attributes: brightness preservation, detail preservation, and contrast improvement. After segmenting the original histogram using an appropriate separating point, two sub-histograms are subjected to independent histogram optimization using MBOBHE. This optimization method is guided by a weighted-sum aggregated objective function (AOF) that takes into account the three performance criteria. According to Hum (2014), MBOBHEO offers a more comprehensive and performance-enhancing approach compared to existing bi-HE methods. The effectiveness of MBOBH is supported by both quantitative and qualitative data, demonstrating that it provides a comprehensive view and effectively maintains a balance between contrast, brightness, and feature preservation in the enhanced images. This novel technique represents a significant advancement in the domain of image contrast augmentation.

Multipeak Histogram Equalization (MHE): Using this image enhancement technique, photographs having many prominent intensity peaks in their histogram can have better contrast and clarity (Shi, 2004). With pixel coordinates *(x, y)* and intensity values *I(x, y)* ranging from *0* to *L-1*, let's say we have an input image *I*. *L* is the number of intensity levels (usually 256 for 8-bit images).
Determine the input image's histogram:

$$X(i),\ i = 0,1,\ldots..L-1$$

The number of pixels in the image with an intensity value of *i* is denoted by *X(i).*
Determine the histogram's peaks:

Prominent intensity levels in the image are represented by peaks. To ascertain the locations of peaks, we can employ a variety of techniques, such as locating local maxima.

Divide the histogram into areas, often known as peaks:

Divide the histogram into various regions or sub-histograms based on the peaks that have been detected. Every sub-histogram depicts a local peak and the intensity levels around it.

For every sub-histogram, carry out histogram equalization:

Use the histogram equalization procedure for each sub-histogram to adjust the intensity values and enhance contrast.
The following represents the classic histogram equalization function:

$$A(x)=\text{round}\ ((L-1)*\sum (H(y)/N) \text{ for } j = 0 \text{ to } x)$$

N is the total number of pixels in the sub-histogram (sum of *H(y)* for each *j* in the sub-histogram), and *A(x)* is the new intensity value for the input intensity *x*.

The final improved image is formed by combining the equalized sub-histograms obtained from the individual sub-histogram equalization based on the segmentation.

CLAHE stands for contrast-limited adaptive histogram equalization: CLAHE, a variation of Adaptive Histogram Equalization (AHE), includes a height parameter to control local contrast and limit noise amplification (Reza, 2004). The steps involved are:

I. The image is divided into $M \times N$ non-overlapping sub-regions, with the size dependent on the desired local enhancement strength.
II. A grayscale histogram *H(i*) is calculated for each sub-region.

CNN image post-processing:

The suggested CNN-based architecture for MRI segmentation in the second phase of the model makes use of the segmented images from the first phase of the fusion model. The segmented image from the earlier stage is provided as input to increase the segmentation accuracy of the suggested CNN model. An image without segmentation will display all background elements, such as borders and textures. This results in the removal of unwanted features from low-priority regions.The convolution and maxpooling layers of the suggested CNN model, each of which employs a different CB,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

are laid out as follows. For this activity, the regular stride length is used and no cushioning is used.

The first CB consists of one MP layer and one convolution layer. The MP layer is 2-by-2, and the first convolution layer has 32 3-by-3 filters. The second CB consists of one MP layer and two convolution layers. While the MP layer has 32 filters in a (2 * 2) configuration, the two convolution layers each contain 16 filters in a (3 * 3) configuration. The third CB consists of one MP layer and two convolution layers once more. The filters in the MP layer are of size (2 * 2). whereas there are a total of 16 filters in the third convolution layer, each with a size of (3 * 3). Following the third CB, the flattening layer is applied. It "flattens" the features by condensing the feature space into a single feature vector.

Segmentation Using Proposed U-Net Model

Deep learning has significantly influenced many fields, especially in the examination of enormous visual, auditory, textual, video, and tabular data. A primary impediment that impeded the progress of convolutional neural networks (CNNs) in medical image segmentation was the necessity for an ample amount of medical pictures to effectively train deep learning models. The U-Net architecture was developed specifically to address the challenge of segmenting medical images with limited datasets. The U-Net technique generates a pixel-level annotated image that emphasizes the segmented region of interest. Unlike standard CNNs commonly used for image classification, U-Net retains both content and position information, ensuring that critical spatial information required for segmentation tasks is not lost. The U-Net architecture is named after its U-shaped structure, which is composed solely of convolutional layers and lacks any dense layers. Being an end-to-end fully convolutional network (FCN), it can effectively process images of any size.

This study describes the creation of a U-Net model to automatically segment the stomach, large intestine, and small bowel in the gastrointestinal (GI) tract. Through rigorous investigation, the optimal number and arrangement of layers were established to yield the best successful segmentation results. The U-Net model consists of a combination of convolutional and max-pooling layers. The model has a U-shaped structure, as illustrated in Figure 5, comprising of an encoder on the left and a decoder on the right, following the typical U-Net design. Significantly, this model exclusively comprises convolution, max-pooling, and transpose convolution layers, without the incorporation of any dense layers.

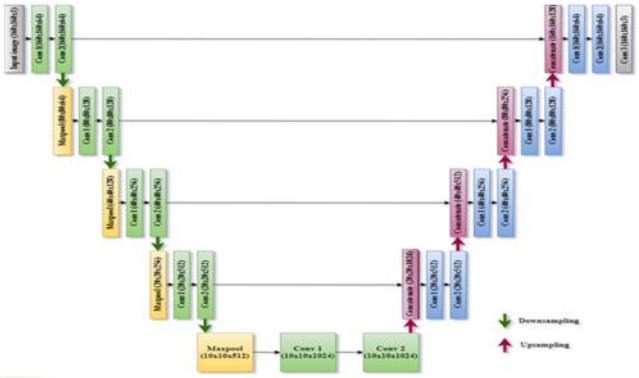


Figure 5: Graphical Representation of U-Net Architecture

The architecture is accessed in two ways. The context of the image is extracted utilizing the first path, also referred to as the contraction path or encoder. The encoder is a conventional stack of max pooling and convolutional layers. The second path is the decoder, which is the symmetric extending path used to obtain exact localization by transposed convolutions This is where the encoder does downsampling, and the decoder does upsampling. Precise segmentation tasks require the architecture to collect high-resolution features and the context of the input image.

Encoder (Contracting Path)

The encoder path comprises repeated application of two 3x3 convolutions, individually followed by a ReLU activation and a 2×2 max pooling operation with stride 2 for downsampling. At every downsampling step, the number of feature channels is doubled.
Let y be the input image, and l be the layer index in the encoder:

$$m_l = \text{ReLU}(\text{Conv2D}(m_{l-1}, f = 3, t = 1, q = \text{'same'}))$$
$$m_l = \text{ReLU}(\text{Conv2D}(m_l, f = 3, t = 1, q = \text{'same'}))$$
$$m_{l+1} = \text{MaxPool}(m_l, f = 2, t = 2)$$

Here,

- Conv2D (m_l, f, t, q) denotes a 2D convolutional layer involved in feature map m_l with kernel size f, stride t, and padding q.
- ReLU(y) describes the ReLU activation function applied to y.
- MaxPool (y, f, t) represents the max pooling operation with kernel size f and stride t.

Bottleneck

The bottleneck is the deepest portion of the network, with the greatest number of feature channels but the smallest spatial dimensions.
It includes two convolutional layers with ReLU activation:

$$b = \text{ReLU}(\text{Conv2D}(m_L, f = 3, t = 1, q = \text{'same'}))$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

b =ReLU (Conv2D (b, f = 3 , t = 1, q = 'same'))

Decoder (Expansive Path)

After upsampling the feature map, the decoder path consists of two 3×3 convolutions with ReLU activation. After every upsampling step, the corresponding feature map from the encoder path is concatenated (skip connections).

Let, v_l be the upsampled feature map, d_l be the concatenated feature map, and *m* be the decoder's layer index:

$$v_l = \text{UpConv2D}(b, f = 2, t = 2)$$
$$d_l = \text{concat}(v_l, m_{L-l-1})$$
$$d_l = \text{ReLU}(\text{Conv2D}(d_l, f = 3, t = 1, q = \text{'same'}))$$
$$d_l = \text{ReLU}(\text{Conv2D}(d_l, f = 3, t = 1, q = \text{'same'}))$$

Here:

- UpConv2D (*y, f, t*) denotes a 2D transposed convolution used to *y* with kernel size *f* and stride *t.*
- Concat (*a,b*) defines the concatenation of feature maps *a* and *b* along the channel dimension.

Output Layer

The final output layer involves a 1×1 convolution to map each 64-component feature vector to the expected number of classes (generally 1 for binary segmentation), followed by a sigmoid activation to make the final segmentation map.

$$z = \text{sigmoid}(\text{Conv2D}(d_0, f = 1, t = 1, q = \text{'same'}))$$

Here,

- Conv2D (*y, f, t, q*) describes a 2D convolutional layer with kernel size *f*, stride *t*, and padding *q*.
- Sigmoid (*y*) denotes the sigmoid activation function applied to *y*.

Upsampling: Enhancing 2D The Upsampling2D layer increases the dimension of the layer output by duplicating the row values.The contracting path and upsampling2D are combined to create the expanding path. In a manner akin to the presence of maxpooling in the encoder section, the upsampling layer is positioned after two conv2D layers in the decoder half.A skip connection, also known as a residual link, is a connection that allows information to bypass certain layers in a neural network. After each of the two Conv2D layers in the encoder, there is a Skip Connection (Residual connection), also referred to as identity mapping. This connection links to the corresponding layer with the same dimensions in the decoder portion. The copy and crop operations in the architecture above are responsible for handling the residual connection task. The Concatenate layer is responsible for merging the two levels. Even in the most unfavorable scenario, this skip connection does not have a detrimental effect on the model; instead, it has a beneficial effect on the model's output.The decoder part is concluded by the output segmentation map, which is marked by filter 2. Again, this layer undergoes filtering using filter 1 and a Conv2D with Relu activation. The output classifies each pixel based on the presence or absence of a tumor.

Hyperparameter Tunning

Twenty epochs were used to train the models with a batch size of 32. While the epochs parameter specifies the number of runs over the entire training data, the batch size hyperparameter specifies the number of samples to proceed before changing the model's internal parameters.

The learning rate, which regulates the model's pace of learning, is the essential hyper-parameter. It must not be unnecessarily high or low. The network may overshoot the low-loss regions if the learning rate is set too high, or it may take an excessive amount of time to reach the minimal loss if it is set too low. In this work, the learning rate is set at 0.0001. The Adam (Zhang, 2018) optimization method was applied for model compilation. In addition, the ReLU (Agarap, 2018) activation function has been used to activate all convolutional layers. Table 1 shows the parameters of the proposed model.

Table 1: Parameters of the proposed model

Parameter	Value/Explanation
Model	'unet((H, W, 3))' - A U-Net model with input shape (Height, Width, 3 channels)
Loss Function	'dice_loss'
Optimizer	'Adam(lr)'
Metrics	'dice_coef',' accuracy'
Callbacks	Various callbacks used during training
ModelCheckpoint	Saves the best model only ('save_best_only=True'); verbose output enabled ('verbose=1')
ReduceLROnPlateau	Reduces learning rate by a factor of 0.1 if no improvement in validation loss for 5 epochs ('patience=5'); minimum learning rate set to '1e-7'; verbose output enabled ('verbose=1')
CSVLogger	Logs training data to CSV file specified by 'csv_path'
EarlyStopping	Stops training if no improvement in validation loss for 20 epochs ('patience=20'); does not restore best weights ('restore_best_weights=False')
Training Data	'train_dataset'
Epochs	'num_epochs'
Validation Data	'valid_dataset'
Verbose	'0' (silent mode for training output)

IV. EXPERIMENTAL ANALYSIS & RESULTS

The experimental setup for detecting and segmenting brain tumors utilizes a variety of powerful tools and libraries. To carry out and evaluate machine learning models, which

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

involve extracting features from data, preprocessing the data, and measuring the performance of the models, we make use of Scikit-learn, also referred to as sklearn. Matplotlib is utilised to visually represent the MRI images, masks, and segmentation results, facilitating comprehension and presentation of the findings.

In addition to facilitating smooth code execution and sharing, Google Colab offers a scalable and collaborative computing environment that makes advantage of high-performance computing resources required for training sophisticated models on huge datasets. The integration of this design enhances the dependability and accuracy of our experimental findings by ensuring an efficient, cooperative, and replicable research procedure. In order to ensure accurate and reliable assessment of our model's performance and its ability to apply learned knowledge to new data, we have partitioned the dataset into distinct subsets for training, validation, and testing purposes in our experimental configuration. The dataset is partitioned into three subsets: 612 images and corresponding masks in the validation set, 612 images and corresponding masks in the test set, and 1,840 images and corresponding masks in the training set. This stratified split enhances the effectiveness of training, tuning, and testing our machine learning models. Additionally, it guarantees accurate performance evaluation and optimisation. The model was evaluated using a set of evaluation metrics. In order to evaluate the effectiveness of our models for tumour detection and segmentation, we employ the following assessment metrics:

- **Precision:** Precision measures the proportion of true positive predictions among all positive predictions. It indicates how many of the predicted positive cases are positive.

$$\text{Precision} = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)}$$

- **Recall (Sensitivity):**Recall measures the proportion of true positive predictions among all actual positive cases. It indicates how well the model identifies positive cases.

$$\text{Recall} = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negative\ (FN)}$$

- **F-score:**The F1-Score is the harmonic mean of precision and recall, providing a single metric that balances both aspects. It is particularly useful when the class distribution is imbalanced.

$$\text{F1-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

- **Accuracy:**Accuracy measures the proportion of true positive and true negative predictions among all predictions. It indicates the overall correctness of the model.

$$\text{Accuracy} = \frac{True\ Positives\ (TP) + Ture\ Negative\ (TN)}{Total\ Predictions}$$

Analysis of Dice Coefficient:

The dice coefficient is a widely used statistic for measuring the similarity between two samples, such as the predicted and ground truth masks of brain tumours in MRI scans. The computation involves dividing the sum of the regions of both masks by twice the intersection of the ground truth and predicted masks. A higher Dice coefficient indicates a greater degree of overlap and agreement between the estimated and actual tumour locations.

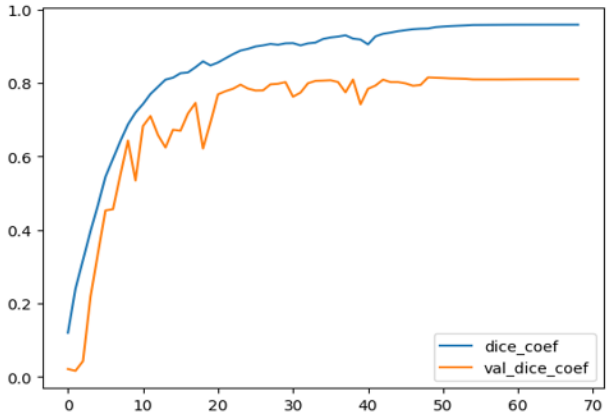


Figure 6: Dice Coefficient of the proposed model

The curve is displayed in Figure 6. In the case of our model, we find that the validation Dice coefficient is approximately 84% and the training Dice coefficient is approaching 93%, indicating that the model performs reasonably well in generalizing to unknown data but with a minor decline in performance from the training set. Since both curves have a similar tendency, the consistency and small divergence between the training and validation curves suggest that the model is not appreciably overfitting.The model successfully learns the patterns and features required to precisely forecast tumor locations from the training data, as shown by the training curve's high Dice coefficient. However, the marginally lower validation Dice value raises the possibility that the distribution of data between the training and validation sets differs in some way. This may be the result of variables the model encounters during validation, such as variances in image quality, anatomical variability, or other features found in real-world data.

Overall, the validation set yielded a Dice coefficient of about 84%, which is a good result and shows that the model can correctly identify and segment tumors in unseen MRI images. The model's capacity for generalization appears to be strong based on the consistency of its performance across training and validation sets; nonetheless, changes and additional research should be directed toward minimizing any disparity in these performance metrics.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Generalization analysis of the proposed model:

This study focuses on assessing a critical element of model generalisation in the context of identifying and segmenting brain tumours using MRI images. Generalisation refers to the ability of a model to apply the patterns and characteristics it has learned to real-world situations by using its expertise to analyse data that it has not been specifically trained on. The figures 7 and 8 display the training and validation accuracy, as well as the training and validation loss.

The graph 8 demonstrates that our model exhibits robust generalisation capabilities in the job of recognising and segmenting brain tumours from MRI data. Generalisation refers to the capacity of a model to apply its acquired knowledge to new and unseen data in real-world scenarios, by extrapolating learned patterns and characteristics from the training set.

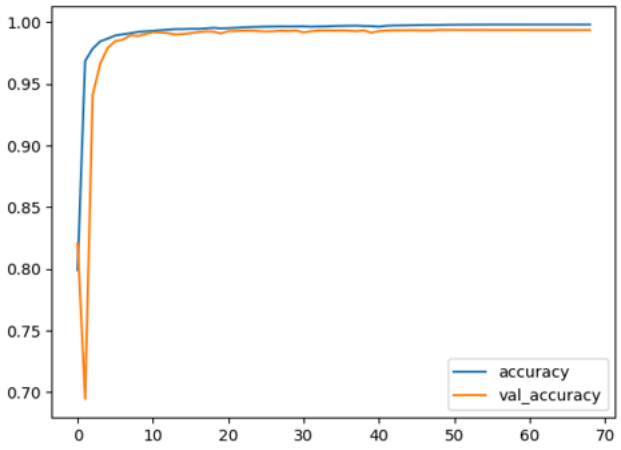


Figure 7: Training and validation accuracy

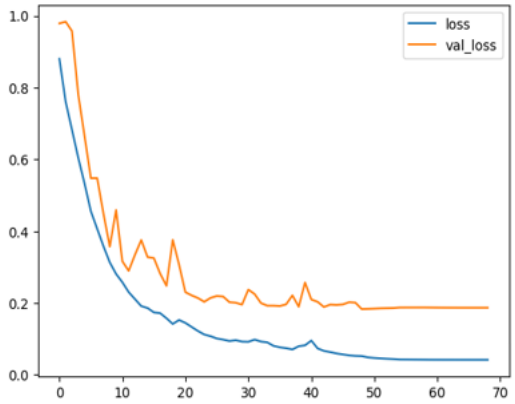


Figure 8: Training And validation loss

The model's capacity to precisely classify tumour and non-tumor regions in the training set of MRI images is evidenced by its training accuracy of approximately 99.65%. This high accuracy demonstrates the effectiveness of the robust learning process, which enables the model to accurately capture intricate features and distinct patterns specific to the dataset. Concurrently, the validation accuracy of around 99.32% indicates that the model remains effective when applied to unseen data. The validation accuracy is crucial as it demonstrates the model's ability to generalise effectively by applying the knowledge it has acquired to new cases that were not part of the training data, although being slightly lower than the training accuracy. The training and validation losses remained consistently small throughout the training method. The model's training inaccuracy was quantified using the training loss, which decreased gradually over time from an initial value of 0.880 in the first epoch to 0.092 in the final epoch. Similarly, over that period, the validation loss decreased from 0.979 to 0.196. The low loss values indicate that the model effectively minimises errors and discrepancies between the expected and actual outputs for both the training and validation datasets. The model's ability to generalise well is reinforced by its consistent reduction in loss, indicating that the model learns to make accurate predictions while also avoiding overfitting or underfitting. The model's reliable and consistent performance on several datasets is evidenced by the little disparity between training and validation accuracies, as well as the convergence of training and validation losses towards the end of the training period. This convergence indicates that the model has effectively learned the significant features of brain tumour images without including any unnecessary information or background noise from the training data. To summarise, our proposed U-Net model demonstrates strong generalisation abilities in the critical task of brain tumour identification and segmentation, as evidenced by the comprehensive analysis of training and validation metrics derived from pictures. The findings are essential for enhancing the accuracy and dependability of tumour identification in medical environments, where it is imperative to enhance patient outcomes and treatment strategies.

Performance of the proposed model:

Finally, we have employed test accuracy, F-score, precision, and recall to evaluate the performance of our model. The efficacy of the suggested U-Net model is demonstrated in Table 2. The precision, which measures the accuracy of the positive projections, was recorded at 92.66%. The model achieves a high level of accuracy by minimising the occurrence of false positives, which refers to regions that are incorrectly identified as tumours. Additionally, it demonstrates a high level of effectiveness in accurately recognising genuine positive cases of brain tumours. The recall metric, which evaluates the model's ability to accurately identify all relevant instances of brain tumours, achieved a score of 93.65%. This demonstrates the model's high sensitivity in accurately identifying the majority of true positive cases, resulting in little instances of missed tumours. A high recall rate is crucial in medical diagnostics, as failing to detect a tumour could have severe consequences for patient outcomes. The F-score, which is the harmonic mean of recall and precision, is 93.15%. This indicator provides a comprehensive measure of the model's effectiveness by considering its performance in accurately recognising true positives and its capability to minimise false positives. The model has strong performance in both domains,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

effectively balancing precision and recall, as evidenced by its high F-score. In addition, the model has an impressive overall accuracy of 97.79% on the test set. Accuracy refers to the proportion of correct results (true positives and true negatives) out of all occurrences examined. The model's high accuracy rate displays its reliable performance across diverse settings, highlighting its robust capability to effectively classify both tumour and non-tumor regions.

Table 2: Performance of the proposed U-Net model

Evaluation Metrics	Performance
Precision	93.66 %
Recall	94.65 %
F-score	94.15%
Accuracy	97.79

Performance Comparison with Trending Method

The U-Net models we presented showed exceptional performance in segmenting brain tumor MRI images, surpassing standard approaches as evidenced by the comparison with studies by Zhang Y. (2024), Huang (2020), and Ishfaq (2023) in Table 3. The model developed by Huang (2020) utilized lightweight feature extraction modules and attention methods to enhance performance, enabling precise diagnosis and effective treatment planning. (Ishfaq, 2023) utilized segmentation, clustering, and multi-class support vector machines (SVM) to identify distinctive characteristics and categorize cancers. Nevertheless, the multi-class SVM technique requires improvement in effectively distinguishing complex datasets, leading to reduced performance. In addition, both Huang (2020) and Ishfaq (2023) encountered difficulties in comprehending the meaning of intricate data, which further impeded their performance.

Table 3: Compare our proposed model with the existing method

Reference	Accuracy
(Zhang Y. a., 2024)	0.806
(Huang, 2020)	0.816
(Ishfaq, 2023)	0.846
Proposed U-net	0. 97

In contrast, our U-NET models prioritized feature advancement during image preprocessing by employing techniques like CLAHE to enhance the visibility of complicated structures and indistinct features in medical images. We also tailored the U-Net architecture with a personalized layered design explicitly optimized for the challenges of brain tumor segmentation. This holistic approach addresses the complexities of medical image segmentation, resulting in excellent outcomes. Our proposed U-Net method performed an accuracy of 0.97, greatly exceeding traditional methods, which achieved accuracies of 0.806 (Huang, 2020), 0.816 (Ishfaq, 2023), and 0.846 (Zhang Y. a., 2024). This combination of innovative strategies highlighted the efficacy of our models in increasing the field of healthcare through enhanced medical image segmentation.

Visual analysis of the segmentation

In the results section of our brain MRI segmentation analysis, we have showcased the performance of our model using a collection of photographs. These photos demonstrate both precise and imprecise segmentations. The visualization is depicted in Figure 9. In the first two figures, the true part of the mask image is overlaid on top of the MRI images. The model's ability to precisely detect and separate the tumor regions is shown in these figures. The model has effectively learned to distinguish tumour tissues from the surrounding brain structures by accurately comparing the segmented areas in these images with the ground truth masks. Precise segmentation is crucial for proper diagnosis and therapy planning as it ensures accurate identification of tumour zones while excluding non-tumor areas. Regrettably, the final image is afflicted by a segmentation fault. The model accurately segments the core tumour site, but erroneously identifies an additional area as a tumour. This erroneous segment demonstrates a spurious detection in the model's predictions due to its proximity to the accurate segment region. Although less frequent, these errors are significant as they might lead to misinterpretation of MRI imaging or unnecessary treatment interventions. The complexity and unpredictability of brain tumor manifestations in MRI scans, the quality of the training data, and the potential overemphasis on specific characteristics within the training set contribute to this lack of accuracy. To rectify these imperfections, the model necessitates enhancement through the implementation of tactics such as augmenting the dataset with a wider range of examples, using advanced regularisation techniques, and fine-tuning the model's hyperparameters to enhance its ability to generalize. Future research can prioritize mistake-reduction measures.

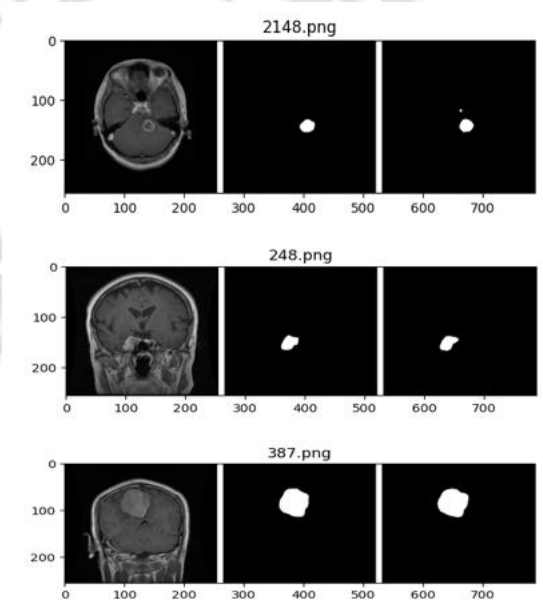


Figure 9: Visual analysis of segmented images

V. CONCLUSION

This paper introduces a novel method for tackling the difficulties associated with segmenting brain tumor MRI

images in the context of contemporary healthcare. Our technique leverages the capabilities of deep learning algorithms. We provide an optimized U-Net model specifically designed for the segmentation of brain cancer MRI images, to enhance precision. Our strategy involves enhancing features through the use of methods such as CLAHE, MHE, and MBOBHE during the picture preprocessing phase. Next, we boost the segmentation outcomes by customizing the layered design of the U-Net model's architecture. Ultimately, we employ a CNN model for post-processing to further enhance segmentation results by incorporating additional convolutional layers. The recall rate achieved was an impressive 93.66%, with an accuracy rate of 97.79%, an F-score of 93.15%, and a precision rate of 92.66%. The Dice coefficient's training and validation curves exhibited marginal deviation, with the training achieving approximately 93% and the validation reaching 84%, showing a strong capacity for generalization. There are alternative avenues for conducting further investigation. This research established the groundwork for more accurate and efficient procedures in medical picture segmentation, which could have a substantial influence on healthcare by assisting radiologists in making precise diagnoses and planning treatments.

REFERENCES

[1] Abdelatty, M. A. (2024). Magnetic resonance imaging of pilonidal sinus disease: interobserver agreement and practical MRI reporting tips. European Radiology, 115--125.

[2] Agarap, A. F. (2018). Deep learning using rectified linear units (relu). arXiv preprint arXiv:1803.08375.

[3] Akter, A. a. (2024). Robust clinical applicable CNN and U-Net based algorithm for MRI classification and segmentation for brain tumor. Expert Systems with Applications, 122347.

[4] Al Garea, S. a. (2024). Image Segmentation Methods: Overview, Challenges, and Future Directions. In 2024 Seventh International Women in Data Science Conference at Prince Sultan University (WiDS PSU) (pp. 56--61).

[5] Altmann, S. a. (2024). Ultrafast brain MRI with deep learning reconstruction for suspected acute ischemic stroke. Radiology, e231938.

[6] Arora, J. a. (2024). Conditional spatial biased intuitionistic clustering technique for brain MRI image segmentation. Frontiers in Computational Neuroscience, 1425008.

[7] Cai, Z. a.-M. (2024). Enhancing Generalized Fetal Brain MRI Segmentation using A Cascade Network with Depth-wise Separable Convolution and Attention Mechanism. arXiv preprint arXiv:2405.15205.

[8] Chen, B. a. (2024). Adaptive cascaded transformer U-Net for MRI brain tumor segmentation. Physics in Medicine \& Biology, 115036.

[9] Ciceri, T. a. (2024). Fetal brain mri atlases and datasets: a review. NeuroImage, 120603.

[10] Desale, P. a. (2024). Navigating Neural Landscapes: A Comprehensive Review of Magnetic Resonance Imaging (MRI) and Magnetic Resonance Spectroscopy (MRS) Applications in Epilepsy. Cureus.

[11] Fidon, L. a. (2024). A Dempster-Shafer approach to trustworthy AI with application to fetal brain MRI segmentation. IEEE transactions on pattern analysis and machine intelligence.

[12] G{\"o}rgec, B. a. (2024). MRI in addition to CT in patients scheduled for local therapy of colorectal liver metastases (CAMINO): an international, multicentre, prospective, diagnostic accuracy trial. The Lancet Oncology, 137--146.

[13] Huang, S. a. (2020). Medical image segmentation using deep learning with feature enhancement. IET Image Processing, 3324-3332.

[14] Hum, Y. C. (2014). Multiobjectives bihistogram equalization for image contrast enhancement. Complexity, 22-36.

[15] Ishfaq, M. A. (2023). Brain tumor classification utilizing deep features derived from high-quality regions in MRI images. Biomedical Signal Processing and Control, 104988.

[16] Kiran, L. a. (2024). An enhanced pattern detection and segmentation of brain tumors in MRI images using deep learning technique. Frontiers in Computational Neuroscience, 1418280.

[17] Kumar, P. R. (2024). Automated human brain tissue segmentation from clinical MRI images for improved neurological diagnosis and treatment planning. Intelligent Medicine.

[18] Kumar, P. R. (2024). Automated human brain tissue segmentation from clinical MRI images for improved neurological diagnosis and treatment planning. Intelligent Medicine.

[19] Mahajan, A. a. (2024). Neuroimaging: CT Scan and MRI. In Principles and Practice of Neurocritical Care (pp. 189--215). Springer.

[20] Nehra, M. a. (2021). Nanobiotechnology-assisted therapies to manage brain cancer in personalized manner. Journal of Controlled Release, 224--243.

[21] Reddy, S. a. (2024). Region based image segmentation to improve accuracy of currency images compared with edge based segmentatio. In AIP Conference Proceedings.

[22] Reza, A. M. (2004). Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. Journal of VLSI signal processing systems for signal, image and video technology.

[23] Rutoh, E. K. (2024). GAIR-U-Net: 3D guided attention inception residual u-net for brain tumor segmentation using multimodal MRI images. Journal of King Saud University-Computer and Information Sciences, 102086.

[24] Shahzad, O. A. (2022). Implementation of a Modified U-Net for Medical Image Segmentation. IEEE, 4593--4597.

[25] Sharif, M. a. (2024). Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection. Journal of ambient intelligence and humanized computing, 1--20.

[26] Shi, H. C. (2004). A simple and effective histogram equalization approach to image enhancement. Digital Signal Processing, 158-170.

[27] Shiny, K. (2024). Brain tumor segmentation and classification using optimized U-Net. The Imaging Science Journal, 204--219.

[28] Solanki, S. a. (2023). Brain Tumor Detection and Classification Using Intelligence Techniques: An Overview. IEEE Access, 12870-12886.

[29] Soppari, K. a. (2024). A survey on brain MRI segmentation. World Journal of Advanced Research and Reviews, 1702--1710.

[30] Yellu, R. R. (2024). edical Image Analysis-Challenges and Innovations: Studying challenges and innovations in medical image analysis for applications such as diagnosis, treatment planning, and image-guided surgery. Journal of Artificial Intelligence Research and Applications, 93--100.

[31] Zhang, A. S.-S. (2017). Complete prevalence of malignant primary brain tumors registry data in the United States compared with other common cancers, 2010. Neuro-oncology, 726--735.

[32] Zhang, Y. a. (2024). Interactive medical image annotation using improved Attention U-net with compound geodesic distance. Expert systems with applications, 121282.

[33] Zhang, Z. (2018). Improved Adam Optimizer for Deep Neural Networks. In 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS) (pp. 1-2).