DOI: 10.15849/IJASCA.250730.05

Int. J. Advance Soft Compu. Appl, Vol. 17, No. 2, July 2025
Print ISSN: 2710-1274, Online ISSN: 2074-8523

IRU-Net: Brain Tumor Detection in MRI Images Using the End-to-End Inception Residual UNet Model

Eman Abdulaziz Ghani Aldhaher[1*], Amir Lakizadeh[2]

[1] Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran, Department of Computer Science, Faculty of Education for Girls, University of Kufa, Iraq
*Corresponding author
e-mail: emana.aldhaher@uokufa.edu.iq
[2] Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran
e-mail: lakizadeh@qom.ac.ir

Abstract

Nowadays, brain tumors are prevalent conditions affecting millions of people worldwide. Detecting the brain tumor accurately in its early phases by brain MRI images is important for diagnosing the disease and avoiding mortality. The rise of deep learning techniques over the past few years has led to effective tools for the diagnosis of medical illnesses, predominantly for the segmentation of brain tumors. Unet is considered one of the most effective end-to-end convolutional neural networks for biomedical image segmentation. This article introduces an Inception pre-trained network on ImageNet with residual connections based on the UNet architecture (called IRU-Net) for brain tumor segmentation. IRU-Net enhances the ability to integrate contextual information by using the GoogLeNet network and adding a residual connection as an encoder part. Additionally, the Inception module that incorporates residual connections is used as a decoder part, enabling the model to capture multi-scale features and ensure efficient gradient flow. The experiments conducted on two different size datasets obtained from Kaggle, namely Brain Tumor Segmentation (BraTS2020) and LGG MRI segmentation (LGG) which is smaller than BraTS2020. The presented results show the influence of pre-training on the network's performance using datasets of varying sizes. The experimental results demonstrate that the proposed IRU-Net architecture outperforms both the traditional UNet and other models that are based on pre-trained UNet..

Keywords: *Brain tumors, MRI, Deep learning, segmentation, Pre-trained, UNet, Inception, Residual connections.*

Received 21 April 2025; Accepted 5 June 2025

1 Introduction

The brain is a crucial and intricate component of the person's body, controlling other organs and all processes that regulate bodily functions [1,2,3]. As a result, an unusual brain abnormality can put an individual's health at risk or, in some cases, can be life-threatening. Among the various types of abnormalities, brain tumors represent the most severe diseases, which can be either benign (noncancerous) or malignant (cancerous) [4,5,6,7,8]. Early and correct brain tumor detection is quite important to help plan the treatment and monitor patients' conditions to improve the survival rate [9,10]. Various medical imaging techniques [11] are used to help diagnose brain tumors by providing important information about their location, size, shape, and kind. Brain MR images (MRI) are considered a standard and precise technique for the early detection and classification of brain tumors [12]. MRI creates several 2D and 3D images of the interior organs of the body in a non-invasive and painless manner [13]. Detecting brain tumors involves a doctor or radiologist analyzing MRI images to identify lesion regions and make diagnostic decisions. Because of the diverse nature of tumors, MRI images sometimes lack distinct features, making accurate decision-making challenging. As a result, integrating AI technologies, particularly deep learning (DL), has become essential for analyzing medical images for diagnostic purposes using computer-aided diagnosis (CAD) systems. [1,2,4,14].

Recently, deep learning (DL) has profoundly impacted numerous medical applications across diverse domains [12,16]. The rapid increase of graphics processing units (GPUs) and the accessibility of medical imaging datasets for training have enabled deep learning to develop advanced methods for processing medical images [16]. For segmentation tasks, DL systems utilizing convolutional neural networks (CNNs) have proven successful, outperforming traditional neural network techniques. Many DL models have been employed during the image segmentation process [17]. Among these models [18-21], UNet [22] has garnered significant attention for its tremendous performance in medical image segmentation since its introduction in 2015 [22,23]. The symmetrical design of UNet, which consists of an encoder and a decoder paths linked by a skip connection, improves its performance. These concatenation-based skip connections can actively preserve tiny features in images and save high-resolution data both throughout the encoding and decoding processes. [17].

Although the unique UNet architecture improves segmentation accuracy in medical imaging tasks, the model has several significant limitations. Some of these include its reliance on enormous datasets that have been annotated, sensitivity to imbalanced class distributions, and its limited ability to capture multi-scale contextual factors. Furthermore, the high costs of training and memory requirements may limit its performance when applied to large datasets [26,12]. Transfer learning (TL) has emerged as a technique enabling deep learning algorithms to deal with the issue of limited training data by utilizing the weights of a model that has already been trained on large-scale datasets (e.g., ImageNet) for various applications and research, particularly in the medical domain [27,28]. The three benefits of TL are that it requires less training time, enhances neural network performance, and operates effectively with minimal data [29]. In a UNet design, the encoder part can be substituted with any pre-trained model. [25].

This study presents an enhanced model based on the UNet architecture, called IRU-Net, that uses a pre-trained GoogLeNet [30] (also known as Inception) model with residual connections [25] for brain tumor detection. Unlike earlier architectures such as AlexNet [31] and VGG-16 [32], GoogLeNet featured a more complex and profound architecture but used a fewer number of parameters without waiving the network's performance [33]. The use of residual connections helps avoid the vanishing gradient problem in

backpropagation and makes it easier to train a deeper network. The aim of the IRU-Net structure is to combine the advantages of both GoogLeNet and UNet in order to increase accuracy, especially in more complex medical imaging tasks like brain tumor detection. Following are the major contributions from this study:

- A new approach (IRU-Net) to detecting brain tumors using end-to-end deep learning is suggested.
- IRU-Net adopted the foundational structure of the UNet architecture. A transfer learning approach was employed, utilizing a pre-trained Inception model with a residual connection as the encoding path. In the corresponding decoding path, the Inception module with residual connections is utilized.
- The effectiveness of the proposed IRU-Net was tested using some common segmentation metrics, such as the Dice Similarity Coefficient, the Jaccard Similarity Index, accuracy, sensitivity and specificity.
- IRU-Net and comparison methods were trained and tested using different datasets in two modes: from scratch and with pre-trained weights. The results indicate that the suggested network performs better in both modes on all datasets.
- Two datasets of different sizes, a larger one (BraTS2020) and a smaller one (LGG), were used. The influence of pre-training shows more with the LGG dataset since its training set is smaller than the BraTS2020 dataset. The research indicates that pre-training enhances model performance when utilizing smaller datasets.

The paper is divided into multiple sections as detailed below: The second section provides a review of related work, whereas the third section outlines the proposed methodology. Section 4 presents the experiment results. The paper's conclusions and future work are finally presented in section 5.

2 Related Work

A brain tumor is challenging to identify manually because of its irregular shapes and ambiguous boundaries [1]. As a result, deep learning and image processing play crucial for the early detection of brain tumors. For the automatic early segmentation and classification of tumor regions, various intelligent techniques have been developed. [17]. Among these techniques, CNN, ensemble learning and pre-trained deep learning models are the most prevalent. Below is a brief overview of some of the most well-known and latest techniques. Ujalambkar et al. [34] introduced a segmentation model for deep learning that employs the U-Net architecture in conjunction with EfficientNet-B7 to improve the detection of lower-grade gliomas (LGGs). The model addresses the limitations of conventional segmentation techniques, including manual inaccuracies, inefficiencies, and prolonged processing times, by employing various levels of feature extraction and discerning distinctions in MRI scans. Model Dice scores averaged 0,92 using a dataset of 110 LGG patients obtained from TCIA and TCGA, highlighting its expert proficiency. Shu-You Lin and Chun-Ling Lin [35] presented an enhanced model by combining pre-trained EfficientNetV2 model as an encoder together with U-Net. Utilizing the Brats 2019 dataset supplied by the (MICCAI), the proposed methodology was assessed for its effectiveness. Experimental results demonstrated that using the proposed architecture enhances the segmentation model's performance, achieving a Dice score of 0.9133, accuracy of 0.9977, and a loss of 0.0866. Rehan Razat et al. [36] introduced dResU-Net a segmentation model based on a 3D deep

residual U-Net for brain tumors segmentation from multi-modality MRI data. The encoder contains residual blocks to deal with the vanishing gradient problems and help to maintain low-level features, while the decoder maintains the U-Net structure for up-sampling. The model attained DSC of 0.8357 for the tumor core, 0.8004 for the enhancing tumor and 0.8660 for whole tumor during cross-validation within the BraTS 2020 dataset. Ali TM et al. [37] suggested a U-Net-based model that uses a pre-trained VGG19 network with the convolutional unit as the encoder part followed by the decoder part uses the encoder's output with an attention mechanism. Testing was done using the BRATS 2020 dataset, achieving DSC values for TC were 0.86, 0.83 for the ET and 0.90 for the WT regions. Isensee et al. [38] employed the nnU-Net architecture for the segmentation of brain tumours. Although nnU-Net's original model did well, the authors improved it by adding region-based training, advanced data augmentation, customised post-processing, and other pipeline modifications. To determine the best configuration, they also re-implemented the BraTS rating system. With respective HD95 values of 8.498, 17.337, and 17.805, their model's Dice scores of 88.95 (total tumor), 85.06 (tumor core), and 82.03 (enhancing tumor). A new deep learning system based on a modified U-Net architecture was developed by Zeldin et al. [39]. The encoder part was used to extract features using different pre-trained networks like DenseNet, ResNet, and NASNet. Then, the decoder part was utilized to obtain the semantic probability map. The BRATS'19 dataset was used to validate the suggested technique. The DSC for the U-Net, VGGNet, ResNet, and DenseNet encoders were 0.809, 0.837, 0.811, and 0.839, respectively. Pei et al. [40] described a framework for brain tumor segmentation using a context-aware deep neural network (CANet). A context encoding module is incorporated into the U-Net's encoder-decoder architecture to compute scaling factors for all tumor classes, facilitating the acquisition of a global context representation. The BraTS 2019 and 2020 datasets were utilized for evaluating the efficacy of the proposed model using DSC. The model reached DSC values of 0.821 for ET, 0.895 for WT, and 0.835 for TC. Silva et al. (41) proposed a fully connected, multicascaded, deep neural network of brain tumor segmentation. A basic convolutional block, an aggregation block, and a convolutional block make up the three deep layer aggregation neural networks that are proposed. The BRATS'20 datasets were employed to evaluate the proposed method, with Hausdorff distance and DSC serving as evaluation metrics. The DSC values for whole tumors were 0.88, 0.82 for enhancing tumors, and 0.79 for core tumors, respectively. distance. The DSC values recorded were 0.88 for whole tumors, 0.82 for enhancing tumors, and 0.79 for core tumors, respectively. Colman et al. (42) introduced a two-dimensional deep residual U-Net design, DR-Unet104, consisting of 104 convolutional layers for brain tumor segmentation from multimodal MRI. The model incorporated bottleneck residual blocks in the encoder and applied dropout regularization after each convolutional block to improve generalization. Evaluated on the BraTS 2020 dataset, it achieved Dice scores of 0.8862 (whole tumor), 0.6756 (enhancing tumor), and 0.6721 (tumor core) on validation data and 0.8673, 0.7514, and 0.7983, respectively, on test data. Despite being a 2D model, DR-Unet104 demonstrated competitive performance while maintaining lower computational complexity compared to 3D models. Despite being a 2D model, DR-Unet104 demonstrated competitive

performance while maintaining lower computational complexity compared to 3D models. Myronenko et al. (43) presented an automated 3D semantic segmentation of brain tumors utilizing an encoder–decoder architecture based on MRI. The encoder consists of a pre-trained ResNet network, followed by the corresponding decoder. The suggested technique was tested with the BRATS 2018 dataset, resulting in DSC values of 0.82 for enhanced tumors, 0.91 for complete tumors, and 0.86 for core tumors, respectively.

The majority of researchers concentrate on the results of segmentation rather than its effectiveness. As a result, for every machine learning task, it is crucial to obtain the fewest features that are still useful. In solving this issue, we use a feature extraction method based on residual connections and inception units, which is light but strong, as it captures meaningful information from the whole MRI image. Similarly, to lessen the task's algorithmic and computational complexity, we will employ a pre-trained network rather than creating a neural network from scratch. This method allows us to optimize segmentation results without affecting the other task's requirements.

3 Methods

This section will explain the suggested model and the fundamental techniques employed in its design and implementation details.

3.1 Research Methods

3.1.1 UNet

A UNet is a deep learning architecture typically used for image segmentation, particularly in medical imaging segmentation tasks [44,45]. It is called UNet because its configuration resembles the letter “U”, and the architecture itself consists of two paths: the encoder and the decoder linked together by a skip connection [25,46]. The down-sampling encoder is composed of multiple convolutional layers that gradually reduce the dimensions of the input image while capturing its low-level features [24,47]. In a corresponding up-sampling decoder, it captures semantic features, and spatial information is recovered by using a series of transposed convolutional layers to reproduce segmented output [22,47]. A key feature of UNet is its use of skipped connections, which significantly enhances the accuracy of the segmentation map [48]. Skip connections facilitate the integration of low-level features with semantic features, enabling the acquisition of more informative features and the preservation of crucial details throughout the segmentation process [16, 4]. The original UNet architecture is shown in Figure.1.

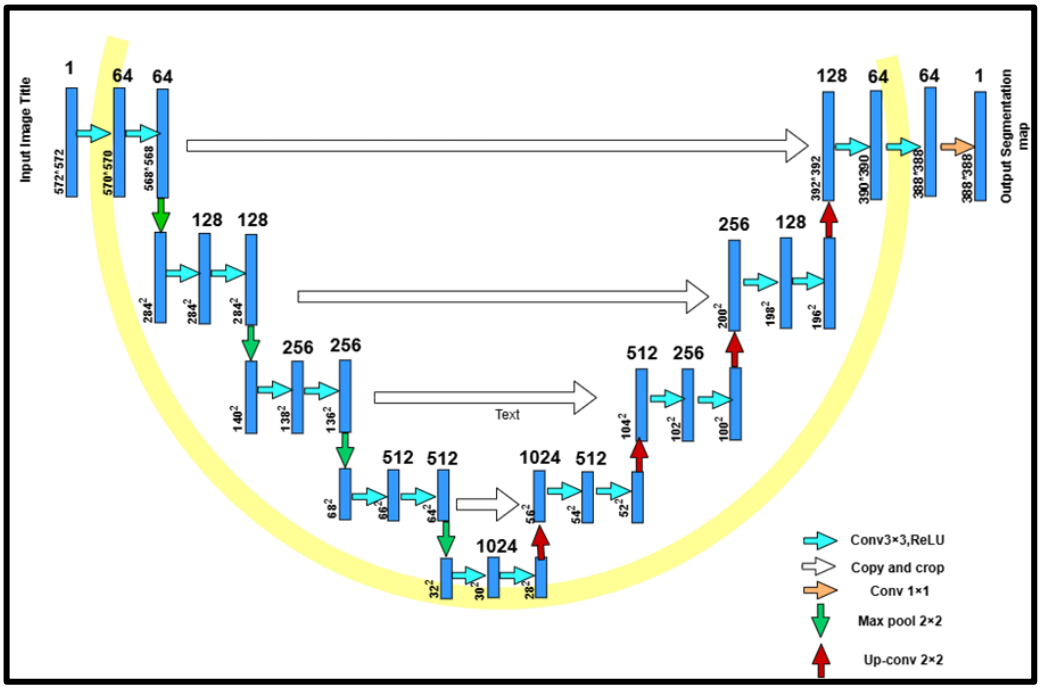

Figure. 1 UNet CNN architecture [22].

3.2.1. GoogLeNet:

Inception, or GoogLeNet, is a pre-trained network architecture proposed by Szegedy et al. [30] in 2014; its design is shown in Figure.2. Networks utilizing the Inception design are faster compared to those employing non-Inception architectures. The model is composed of 22 layers, incorporating nine Inception modules [30,49.50]. This advanced Inception module utilizes trainable filters with different kernel sizes, ranging from (1×1) to (5×5), to execute parallel convolution operations [17, 25]. Figure.3 displays the structure of the Inception module. This design enables the extraction of richer image features at multiple levels of detail. GoogLeNet features a deeper and more complex architecture while utilizing significantly fewer parameters without compromising performance. This efficiency is achieved by replacing the fully connected layers at the top of the network with a global average pooling layer and incorporating auxiliary classifiers to enhance gradient flow in deeper layers. These design elements not only improve convergence during training but also help mitigate the vanishing gradient problem, which is common in deeper networks [30,33,49,50,51].

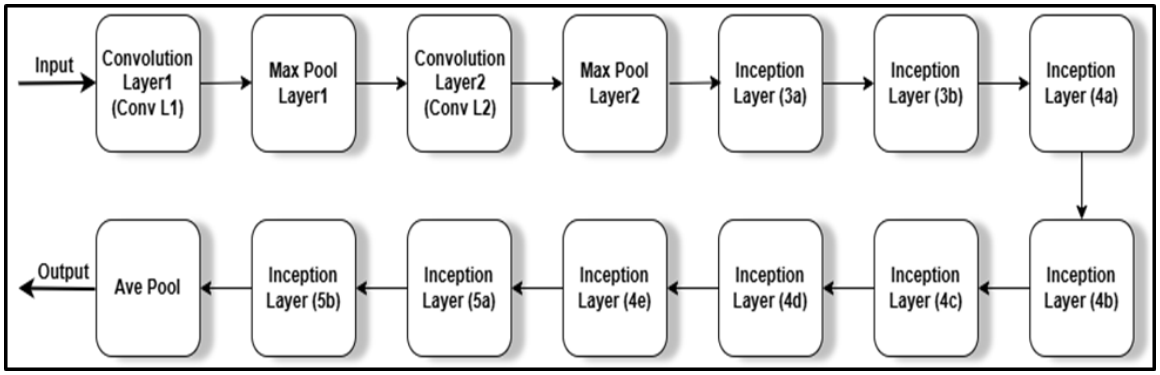

Figure.2 The structure diagram of GoogLeNet

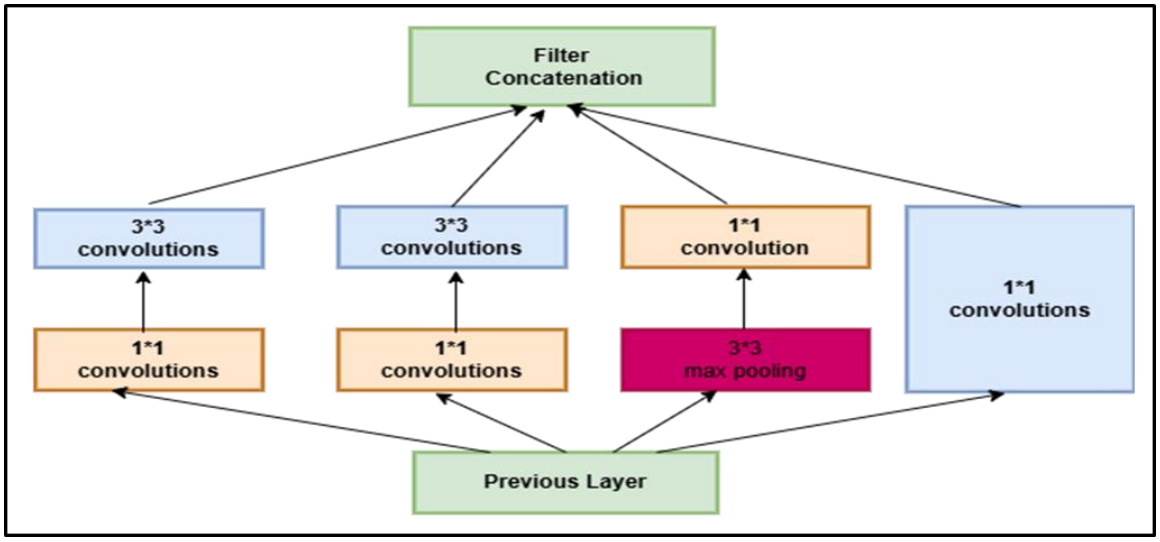

Figure.3 The components of the inception's module

3.2.2. Residual Connections

Residual connections [51] were introduced to enhance network learning efficiency and mitigate the vanishing gradient problem during backpropagation, particularly in deeper architectures with numerous layers. The following is the formula for a residual block:

$$H(x) = F(x) + x \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

In a residual block, the output of each layer is forwarded to the next layer while simultaneously being directly combined with the input through a shortcut connection. Figure.4 illustrates the residual block, which facilitates the design of deeper networks while reducing the risk overfitting.

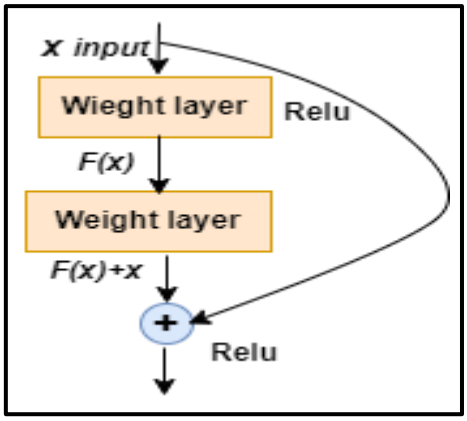

Figure.4 Original residual block

3.2. The Proposed Method

IRU-Net is a deep learning model that is proposed for the purpose of brain tumor segmentation. It is based on the basic architecture of UNet. There are two primary pathways that make up the structure of the IRU-Net. The encoder path is situated on the left, and the decoder path is situated on the right. Each of these pathways is composed of six layers. A skip connection is used to establish a connection between each encoder layer

and the decoder layer that corresponds to it. Figure.5 illustrates the overall structure of the proposed IRU-Net.

The standard GoogLeNet is used as the encoder in the UNet architecture's encoding path. Employing a pre-trained encoder allows the model to converge more quickly and achieve better performance than a model without pre-training. The encoder weights are initialized from the pre-trained GoogLeNet on the ImageNet dataset and then fine-tuned using a brain tumor dataset. Each encoder layer consisting of two Inception layers, convolution1x1 (Conv1x1), rectified linear unit (ReLU) activation function and 2x2 spatial max pooling (MP). In each layer, the residual connection is passed from base convolution to enhance feature propagation and prevent information loss. The output layer is concatenated and transmitted directly to the appropriate decoder layer via skip connections to preserve spatial features and increase segmentation performance. The max pooling operation halves the output feature map resolution before passing it to the next encoding layer.

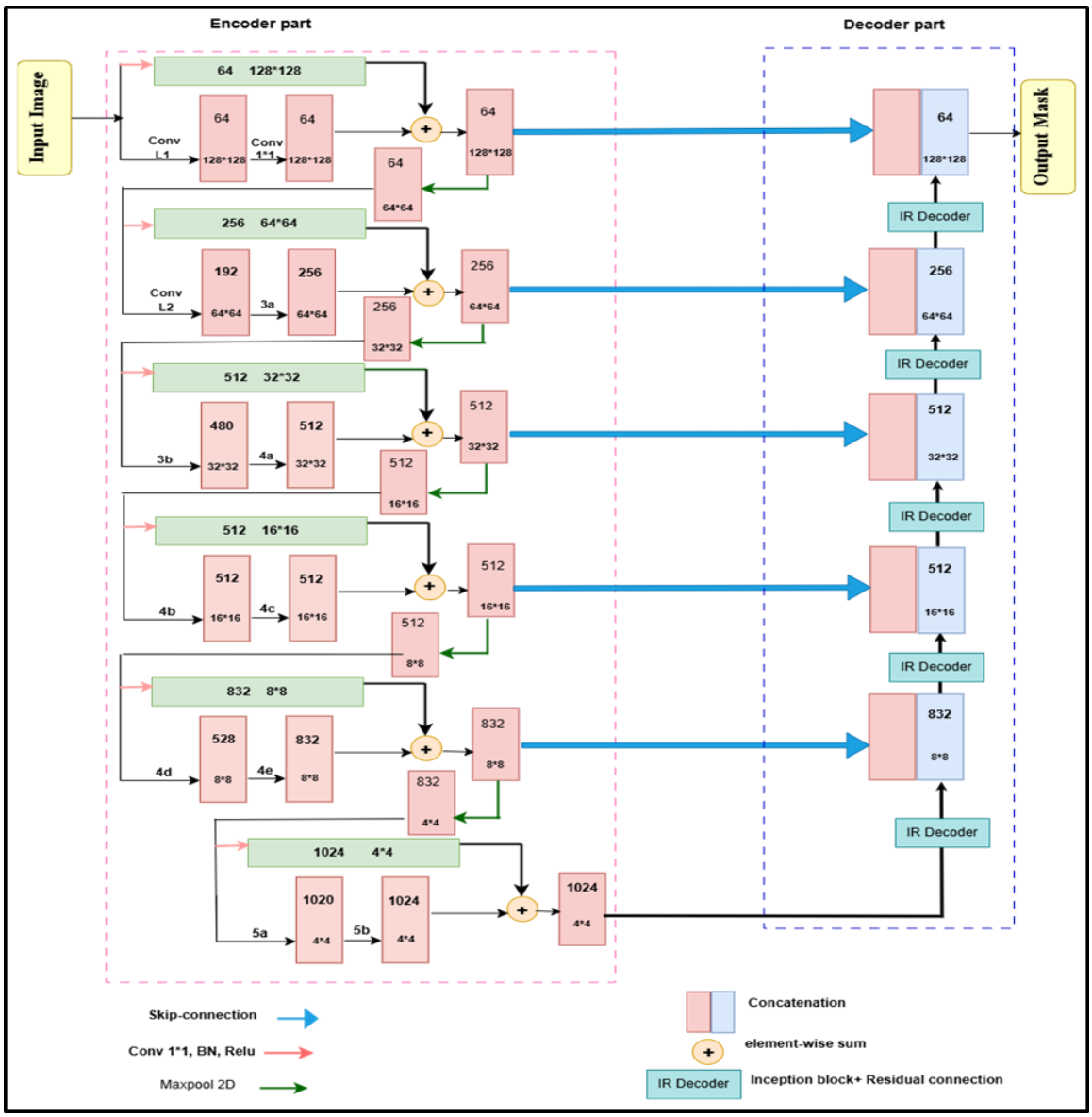

Figure.5 The proposed brain tumor detection focused on the combination of UNet architecture and GoogLeNet model.

The feature map that were extracted by the encoder component is up-sampled for the decoding path using an inception module and the residual connection as a decoder portion. The inception module, which relies on learnable filters with sizes ranging from (1x1) to (5x5), is applied after the input has first been passed to convolution 3 x 3 (Conv 3x3) in each inception residual decoder (IR Decoder) block to extract the key features. The parallel convolution in different kernel size aids in capturing features with varying levels of detail, enhancing the network's learning capacity. Last but not least, to provide a deeper network design without performance deterioration, a residual connection is routed from the base convolution. The IR Decoder block is using convolution 1x1 (Conv1x1) for equalizing the number of the encoder feature maps. Each decoder block's output is sent into transposed convolutions (TConv), which use skip connections to concatenate it with the matching encoder layer, doubling the size of the feature map. The feature map's size gets restored to the input image's original size at the final layer of the decoding route. The structure of IR Decoder is shown in Figure.6.

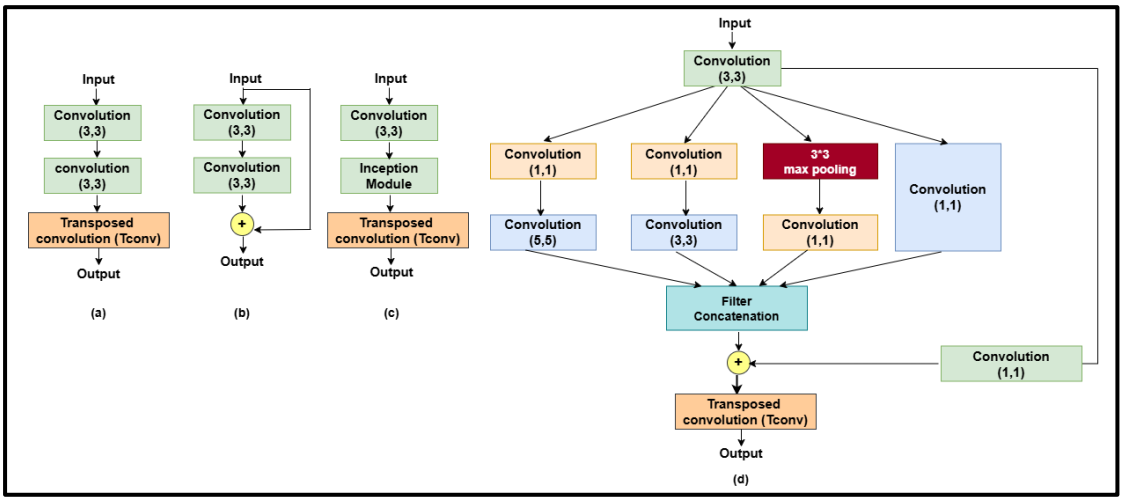

Figure.6 IRU-Net Decoder Structure. (a) UNet decoder. (b) Residual UNet decoder. (c) Inception UNet decoder. (d) Proposed IR Decoder

3.2 Implementation Details

The experimental work executed with PyTorch framework [52] within Google's Collaboratory environment, a cloud-based platform designed for machine learning and deep learning tasks. The model was trained on the BraTS2020 and Brain LGG training datasets using a batch size of 16 for 50 epochs. The resolution of the input images scaled down to 128 × 128 pixels for the BraTS2020 dataset and 256 × 256 pixels for the LGG dataset. The Adam optimizer was employed with its default settings [53]. The learning rate was started at 0.001 and then reduced if the DCS metric stopped to improve over the course of seven epochs.

4 Results and Discussion

This part discusses the impact of pre-training on the efficacy and performance of the proposed model. Furthermore, a comparison with other pre-trained deep learning models as well as other state-of-the-art models will be presented.

4.1 MRI Datasets

The proposed model's accuracy and performance are evaluated using the MRI brain tumor dataset. The two different datasets used in this study to perform a convincing evaluation were taken from the Kaggle database, which publicly provides brain MRI images. The first dataset, namely Brain Tumor Segmentation (BraTS2020) [54]. It comprises 3D photos of dimensions 240 x 240 x 155 from 660 patients, sourced from various institutions employing distinct clinical methods and MRI scanners. Each patient has four types of 3D MRI images: regular T1-weighted (T1), T1-weighted after contrast (T1Gd), T2-weighted (T2), and T2 Fluid-Attenuated Inversion Recovery (T2-FLAIR) volumes. Given that the model architecture utilised in this study accepts 2D images as inputs, the 3D images from the BraTS 2020 dataset are segmented into 2D slices. Consequently, each image divided the axial plane, producing a collection of 155 two-dimensional slices matching to each original three-dimensional image [55]. The second dataset is the LGG MRI segmentation dataset (LGG) [56]. The number of MRI slices of patient brains ranges from 20 to 88, and the imaging data taken before surgery includes a fluid-attenuated inversion recovery (FLAIR) sequence. This database has roughly 3,929 brain MRI images accompanied by the associated manual FLAIR segmentation results. The ground truth for all images was manually generated following a consistent annotation protocol and was validated by an expert neuroradiologist. Table.1 describes the splitting of the training and testing size for datasets used in our experiments.

Table.1 Datasets information

DATASET	Training size	Testing size
Brain Tumor Segmentation (BraTS2020)	14,784	6,336
LGG MRI segmentation (LGG)	3,179	693

4.2 Evaluation Metrics

The proposed model's effectiveness is evaluated using several well-known metrics for image segmentation, including Dice Score (DSC), Jaccard Similarity (JSC), Sensitivity (SEN), Specificity (SPE) and Accuracy (ACU). Dice Score and Jaccard Similarity are the essential metrics in brain tumor segmentation. They are appropriate for brain tumor datasets due to their imbalance issues [57]. Dice Score (F1-score) quantifies the agreement or spatial overlap between the predicted segmentation mask and the ground truth mask. Its formula is as follows in Equation (2) as follows:-

$$\mathrm{DC}S = \frac{2*\mathrm{Tp}}{2*\mathrm{Tp}+\mathrm{Fp}+\mathrm{Fn}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Jaccard Similarity quantifies the extent of overlap between the predicted mask and the ground truth. Both DSC and JSC are used to evaluate the similarity between the predicted and actual masks. The formula for computing JSC is presented as follows in Equation (3): -

$$JSC = \frac{Tp}{Tp+Fp+Fn} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

ACU, SEN, and SPE are used as additional indicators for model evaluation. These metrics criteria are calculated using the formulas presented in equations (4) to (6), respectively:

$$ACU = \frac{Tp+Tn}{Tp+Tn+Fp+Fn} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \quad (4)$$

$$SEN = \frac{Tp}{Tp+Fn} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

$$SPE = \frac{Tn}{Fp+Tn} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

True positives (Tp), false positives (Fp), true negatives (Tn), and false negatives (Fn) are the four primary outcomes that indicate the pixel's number in the predicted image in comparison to the ground-truth image.

4.3. Ablation Study

To find out how various changes affect the suggested network performance, ablation research is conducted. As shown in Table.2, the suggested IR-Unet contains three changes; as a result, it trained and tested MRI datasets using these modifications, both from scratch and with pre-training, using the same number of epochs and hyper-parameter for every experiment. Since the DCS is the most widely used metric to reflect segmentation performance, it is used to evaluate the testing set outcomes for the datasets included in this research for each experiment.

The first experiment, which we have dubbed IRU-Net1 for convenience, employed the pre-trained Inception network as the encoder part for the UNet architecture in both the training and testing datasets. The model, which we refer to as IRU-Net2, was created and evaluated in the second experiment by including the residual connection to improve the outcomes. After employing the inception block with a residual connection as a decoder component for the UNet architecture, the last experiment, which exemplifies the suggested model, tests the IRU-Net. Incorporating residual connections enhances the network's learning ability and improves performance by addressing the vanishing gradient problem, which often occurs in deeper networks.

Table. 2 Ablation analysis of the proposed network was conducted on two datasets, comparing performance with and without pre-training on ImageNet.

MRI DATASET	Model	DCS	
		Pre-trained	Scratch
BraTS2020	IRU-Net1	0.872	0.866
	IRU-Net2	0.874	0.871
	IRU-Net	0.891	0.878
LGG	IRU-Net1	0.901	0.854
	IRU-Net2	0.919	0.868
	IRU-Net	0.922	0.883

Table.2 presents quantitative results indicating that the proposed network achieved better scores on pre-trained datasets. Furthermore, the dice score curves in Figure.7 illustrate the effect of pre-training on model performance. The curves show the model's training and testing behavior with and without pre-training on each dataset. from the curves, it is noticeable that the model with pre-training converged faster. This makes sense, as the

low-level features that are obtained from pre-training on the ImageNet dataset (one of the largest and most popular datasets) are effective in the learning process.

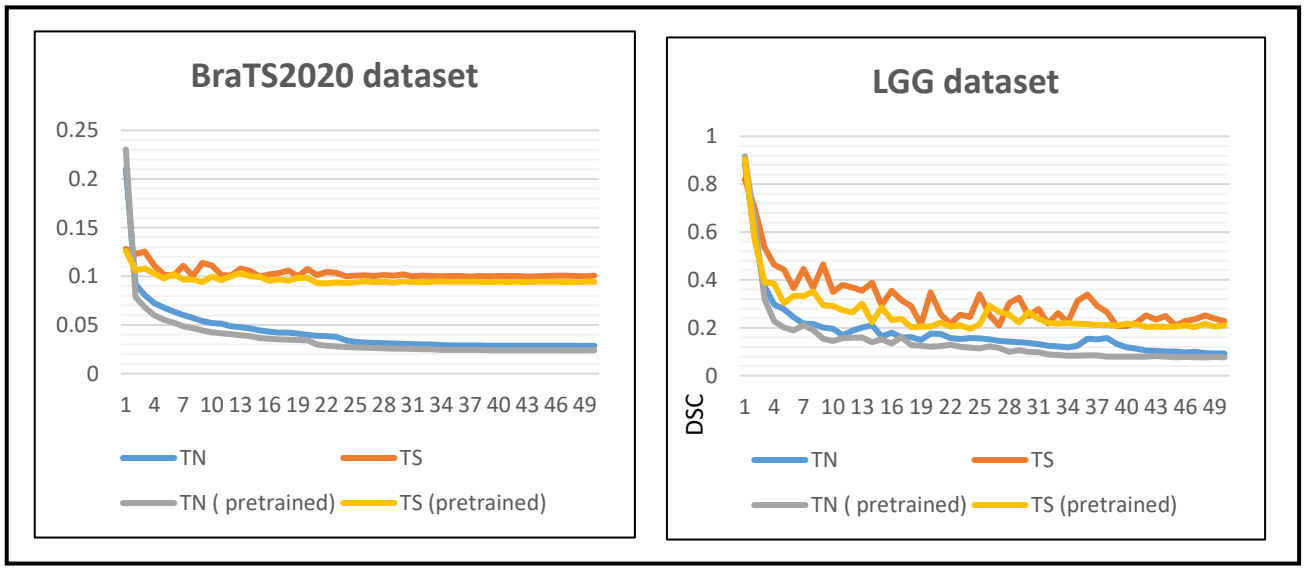

Figure.7 The DCS curves for MRI datasets (BraTS2020 and LGG) train on the proposed (IRU-Net) model. The training set is denoted by TN, the testing set by TS, and the number of epochs by EPS.

4.4. Comparison with Other Pre-trained Deep Learning Methods

Other pre-trained models that follow the U-shaped structure were compared to the suggested IRU-Net using the BraTS2020 and LGG datasets with pre-training and from scratch. Table.3 presents key metrics, including the number of parameters and computing complexity as shown by floating point operations (FLOPs), for both the proposed method and alternative methods.

Table. 3 Resources utilized for both the suggested approach and alternative approaches.

Model	Alex-Unet	IR-Unet (Proposed)	ResNet18-Unet	UNet	VGG11-Unet
Number of parameters	8,759,073	20,098,369	28,976,321	31,037,633	45,995,457
Number of FLOPs	629,277,984	2,023,036,416	5,958,262,784	12,081,823,744	6,946,193,408
Size in Memory	72.99 MB	284.75 MB	246.47 MB	373.59 MB	421.4 MB
Inference (s)	0.19	0.19	0.19	0.19	0.19

According to Table.3, IR-Unet is more computationally efficient than ResNet18-Unet, UNet, and VGG-Unet because it uses less FLOPs and parameters. But compared to Alex-Unet, it needs more parameters and FLOPs. On the other hand, IRU-Net uses less memory than UNet and VGG-Unet but more than Alex-Unet and ResNet18-Unet combined. In terms of inference time, the suggested model's time on a single test image is close to other models. Thus, IRU-Net uses reasonable amounts of computational resources. As shown in Table.4, the suggested approach and alternative deep learning techniques were assessed using the datasets in Table.1. Five performance metric measures (DC, JSC, ACU, SEN, and SPE) were used, both with and without ImageNet pre-training.

Table. 4 Results obtained from proposed model and other deep learning models. Bold indicates high-value performance metrics.

DATA -SET	Model	DCS		JSC		ACU		SEN		SPE	
		Pre-trained	Scratch	Pre-trained	Scratch	Pre-trained	Scratch	Pre-trained	Scratch	Pre-trained	Scratch
BraTS 2020	UNet.	0.866		0.789		0.984		0.864		0.993	
	ResNet18-Unet. [44]	0.871	0.864	0.795	0.784	0.984	0.983	0.795	0.857	0.992	0.993
	Alex-Unet. [16]	0.852	0.850	0.768	0.765	0.982	0.981	0.844	0.843	0.992	0.992
	VGG-Unet. [39]	0.862	0.856	0.796	0.781	0.984	0.984	**0.876**	0.878	0.992	0.992
	IRU-Net. **(Proposed)**	**0.891**	0.878	**0.796**	0.786	**0.984**	0.983	0.868	0.866	**0.993**	0.992
LGG	UNet.	0.823		0.745		0.940		0.820		0.949	
	ResNet18-Unet.	0.706	0.676	0.683	0.676	0.997	0.997	0.686	0.676	0.899	0.896
	Alex-Unet.	0.882	0.861	0.852	0.833	0.992	0.991	0.903	0.895	0.999	0.998
	VGG-Unet.	0.917	0.852	0.889	0.830	**0.998**	0.996	**0.934**	0.870	0.998	0.998
	IRU-Net. **(Proposed)**	**0.922**	0.852	**0.893**	0.853	0.997	0.997	0.918	0.910	**0.998**	0.998

The results presented in Table.4 for the BraTS2020 dataset confirm that the IRU-Net model with pre-training and from scratch offers better values for all metrics, with the exception of a small ratio in SEN when compared to VGG-Unet. The IRU-Net model achieved Dice Score, Jaccard Similarity, Accuracy and Specificity of 0.891, 0.796, 0.984 and 0.993, respectively. The results on the LGG dataset, as shown in Table.4, confirm that the proposed model both with pre-training and from scratch, offers better values of all metrics except for a small ratio in SEN and ACU compared to VGG-Unet with pre-training. The IRU-Net model achieved Dice Score, Jaccard Similarity and Accuracy of 0.922, 0.893, 0.984 and 0.998, respectively. The IRU-Net outperformed other deep learning methods recently built for brain tumor segmentation, Using the BraTS2020 and LGG datasets, as illustrated in Table.5

Pre-training is an effortless adjustment that enhances performance and allows the networks to converge more quickly. The pre-trained weights positively impacted our model and demonstrated improved outputs in terms of convergence speed and accuracy. The suggested approach and other deep learning methods that used pre-training perform better than those built from scratch, as shown in Table.5 of the experimental results for the datasets, which require more data and computing power to achieve similar results. The influence of pre-training is more evident on the LGG dataset because it has a smaller training set when compared to other datasets. In summary, based on its DCS and JSC, the proposed pre-trained model outperformed the other pre-trained models. According to the results, the IR-Unet model is effective in detecting brain tumor images and could be a useful tool for medical practitioners to diagnose brain tumors

Table. 5 DSC values of the proposed method and other state-of-the-art methods.

Model	DSC	Dataset
EfficientNet-B7-Unet [34]	0.920	LGG
dResU-Net [36]	0.834	BraTS2020
VGG19-Unet [37]	0.863	BraTS2020
nnU-Net [38]	0.853	BraTS2020
IRU-Net (Proposed)	**0.922**	LGG
IRU-Net (Proposed)	**0.891**	BraTS2020

.

5 Conclusion

This paper introduces an end-to-end deep learning model called IRU-Net for the segmentation of brain tumors in MRI images. IRU-Net is an improved UNet model, replacing its encoder with a pre-trained GoogLeNet network with residual connections. The combination enables capturing different levels of features to fuse them from different layers prior to sending them to the decoder. Furthermore, the inception block with residual connections was used in the path of the decoder to enhance multi-scale feature representation and improve information flow; thus, it improves the results. A public datasets BraTS2020 and LGG with five evaluation metrics were employed to evaluate the performance of the proposed pre-trained model against other pre-trained deep learning models. The results indicated that all models performed better with pre-training compared to training from scratch. Moreover, the proposed model outperformed all other comparable models. The IRU-Net model achieved Dice Score, Jaccard Similarity, Accuracy and Specificity 0.922, 0.893, 0.984 and 0.998, respectively in BraTS2020 and of 0.922, 0. of 893, 0.984 and 0.998, respectively in LGG dataset. Future work will focus on improving the accuracy and efficiency of the proposed model by integrating additional training data and utilizing fine-tuning techniques with pre-trained weights rather than initiating training from scratch. Subsequent investigations could assess the model's robustness by analyzing its performance on images influenced by noise or enhanced by using super-resolution methods to replicate diverse clinical imaging conditions. Furthermore, extending the model to accommodate additional imaging modalities, including CT and PET, along with the incorporation of innovative heuristic optimization techniques, could enhance diagnostic accuracy and overall model efficacy.

References

[1] Krishnapriya S and Karuna Y (2023) Pre-trained deep learning models for brain MRI image classification. Front. Hum. Neurosci. 17:1150120. doi: 10.3389/fnhum.2023.1150120.

[2] Alsubai S, Khan HU, Alqahtani A, Sha M, Abbas S and Mohammad UG (2022) Ensemble deep learning for brain tumor detection. Front. Comput. Neurosci. 16:1005617. doi: 10.3389/fncom.2022.1005617.

[3] Ramtekkar, P.K., Pandey, A. & Pawar, M.K. Accurate detection of brain tumor using optimized feature selection based on deep learning techniques. Multimed Tools Appl 82, 44623–44653 (2023). https://doi.org/10.1007/s11042-023-15239-7.

[4] Senan EM, Jadhav ME, Rassem TH, Aljaloud AS, Mohammed BA, Al-Mekhlafi ZG. Early Diagnosis of Brain Tumour MRI Images Using Hybrid Techniques between Deep and Machine Learning. Comput Math Methods Med. 2022 May 18;2022:8330833. doi: 10.1155/2022/8330833. PMID: 35633922; PMCID: PMC9132638.

[5] Saeedi, S., Rezayi, S., Keshavarz, H. et al. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Med Inform Decis Mak 23, 16 (2023). https://doi.org/10.1186/s12911-023-02114-6.

[6] Lamrani, Driss & Cherradi, Bouchaib & el Gannour, Oussama & Bouqentar, Mohammed & Bahatti, Lhoussaine. (2022). Brain Tumor Detection using MRI Images and Convolutional Neural Network. International Journal of Advanced Computer Science and Applications. 13. 10.14569/IJACSA.2022.0130755.

[7] Seetha J, Raja S. S. Brain Tumor Classification Using Convolutional Neural Networks. Biomed Pharmacol J 2018;11(3).
[8] P.M. Siva Raja, Antony Viswasa rani, Brain tumor classification using a hybrid deep autoencoder with Bayesian fuzzy clustering-based segmentation approach, Biocybernetics and Biomedical Engineering, Volume 40, Issue 1, 2020, Pages 440-453, ISSN 0208-5216, https://doi.org/10.1016/j.bbe.2020.01.006.
[9] El-Sayed A. El-Dahshan, Heba M. Mohsen, Kenneth Revett, Abdel-Badeeh M. Salem, Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm, Expert Systems with Applications, Volume 41, Issue 11, 2014, Pages 5526-5545, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2014.01.021.
[10] Nodirov, J.; Abdusalomov, A.B.; Whangbo, T.K. Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images. Sensors 2022, 22, 6501. https://doi.org/10.3390/ s22176501.
[11] Abd-Ellah MK, Awad AI, Khalaf AAM, Hamed HFA. A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. Magn Reson Imaging. 2019 Sep;61:300-318. doi: 10.1016/j.mri.2019.05.028. Epub 2019 Jun 5. PMID: 31173851.
[12] Rai, Hari & Chatterjee, Kalyan. (2021). 2D MRI image analysis and brain tumor detection using deep learning CNN model LeU-Net. Multimedia Tools and Applications. 80. 10.1007/s11042-021-11504-9.
[13] Syed, Sibtain & Khan, Maqbool & Ahmed, Rehan & Talha, Syed. (2023). Intracranial Tumor Detection using Magnetic Resonance Imaging and Deep Learning. International Journal of Emerging Multidisciplinaries: Computer Science & Artificial Intelligence. 2. 10.54938/ijemdcsai.2023.02.1.239.
[14] Shah, Hasnain & Saeed, Faisal & Yun, Sangseok & Park, Jun-Hyun & Paul, Anand & Kang, Jae-Mo. (2022). A Robust Approach for Brain Tumor Detection in Magnetic Resonance Images Using Finetuned EfficientNet. IEEE Access. 10. 1-1. 10.1109/ACCESS.2022.3184113.
[15] Tiwari, Pallavi & Pant, Bhaskar & Elarabawy, Mahmoud & Abd-Elnaby, Mohammed & Mohd, Noor & Dhiman, Gaurav & Sharma, Subhash. (2022). CNN Based Multiclass Brain Tumor Detection Using Medical Imaging. Computational Intelligence and Neuroscience. 2022. 1-8. 10.1155/2022/1830010.
[16] Khalaf, Muna and Dhannoon, Ban N. (2022) "Skin Lesion Segmentation based on U-Shaped Network," Karbala International Journal of Modern Science: Vol. 8 : Iss. 3 , Article 20. Available at: https://doi.org/10.33640/2405-609X.3248.
[17] Wang, R., Lei, T., Cui, R., Zhang, B., Meng, H., Nandi, A.K.: Medical image segmentation using deep learning: A survey. IET Image Process. 16, 1243–1267 (2022). https://doi.org/10.1049/ipr2.12419.
[18] Li, H.; Li, A.; Wang, M. A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. Comput. Biol. Med. 2019, 108, 150–160.
[19] Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation.I EEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 2481–2495.
[20] Lizarraga-Morales, R.A.; Sanchez-Yanez, R.E.; Ayala-Ramirez, V.; Patlan-Rosales, A.J. Improving a rough set theory-based segmentation approach using adaptable threshold selection and perceptual color spaces. J. Electron. Imaging 2014, 23, 013024.
[21] Wang,Z.; Jensen, J.R.; Im, J.J.E.M. An automatic region-based image segmentation algorithm for remote sensing applications.Environ. Modell. Softw. 2010, 25, 1149–1165.

[22] Ronneberger, O., Fischer, P., Brox, T., 2015. UNet: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention. Springer, pp. 234-241..
[23] Saeed, M.U.; Ali, G.; Bin, W.; Almotiri, S.H.; AlGhamdi, M.A.; Nagra, A.A.; Masood, K.; Amin, R.u. RMU-Net: A Novel Residual Mobile U-Net Model for Brain Tumor Segmentation from MR Images. Electronics 2021, 10, 1962. https://doi.org/10.3390/ electronics101619.
[24] M. Vidyarthi et al., "Novel Architecture and Comparative Study for Deep Learning-Based Segmentation of 3D Brain MRI Images Using SegNet, V-Net, and U-Net," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 714-719, doi: 10.1109/IC3SE62002.2024.10593388.
[25] Aghayari, S., Hadavand, A., Mohamadnezhad Niazi, S., and Omidalizarandi, M.: BUILDING DETECTION FROM AERIAL IMAGERY USING INCEPTION RESNET UNET AND UNET ARCHITECTURES, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-4/W1-2022, 9–17. https://doi.org/10.5194/isprs-annals-X-4-W1-2022-9-2023, 2023.
[26] Md. Eshmam Rayed, S.M. Sajibul Islam, Sadia Islam Niha, Jamin Rahman Jim, Md Mohsin Kabir, M.F. Mridha, Deep learning for medical image segmentation: State-of-the-art advancements and challenges, Informatics in Medicine Unlocked, Volume 47,2024,101504,ISSN 2352-9148,https://doi.org/10.1016/j.imu.2024.101504.
[27] Filatov, Dmytro & Yar, Ghulam. (2022). Brain Tumor Diagnosis and Classification via Pre-Trained Convolutional Neural Networks. 10.48550/arXiv.2208.00768.
[28] Wen, Yang & Chen, Leiting & Deng, Yu & Zhou, Chuan. (2021). Rethinking pre-training on medical imaging. Journal of Visual Communication and Image Representation. 78. 103145. 10.1016/j.jvcir.2021.103145.
[29] K. Nishanth Rao, Osamah Ibrahim Khalaf, V. Krishnasree, Aruru Sai Kumar, Deema Mohammed Alsekait, S. Siva Priyanka, Ahmed Saleh Alattas, Diaa Salama AbdElminaam, An efficient brain tumor detection and classification using pre-trained convolutional neural network models, Heliyon, Volume 10, Issue 17, 2024, e36773, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2024.e36773.
[30] Szegedy, Christian & Liu, Wei & Jia, Yangqing & Sermanet, Pierre & Reed, Scott & Anguelov, Dragomir & Erhan, Dumitru & Vanhoucke, Vincent & Rabinovich, Andrew. (2015). Going deeper with convolutions. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1-9. 10.1109/CVPR.2015.7298594.
[31] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Adv Neural Inf Process Syst. 25 (2012), https://doi.org/10.1145/3383972. 3383975.
[32] Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 2015.
[33] Al-Huseiny, Muayed & Sajit, Ahmed. (2021). Transfer learning with GoogLeNet for detection of lung cancer. Indonesian Journal of Electrical Engineering and Computer Science. 22. 1078. 10.11591/ijeecs. v22.i2. pp1078-1086.
[34] Ujalambkar, D. M., Itkarkar, R. R., Patil, V. N., & Karve, S. M. (2025). Automated detection of lower-grade gliomas using deep learning with UNet and EfficientNet-B7. Journal of Information Systems Engineering and Management, 10(33s). https://www.jisem-journal.com.
[35] T., G., & K. , S. K. (2024). Brain Tumor Segmentation and Classification Using CNN Pre-Trained VGG-16 Model in MRI Images. IIUM Engineering Journal, 25(2), 196–211. https://doi.org/10.31436/iiumej.v25i2.2963

[36] Rehan Raza, Usama Ijaz Bajwa, Yasar Mehmood, Muhammad Waqas Anwar, M. Hassan Jamal,dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI, Biomedical Signal Processing and Control, Volume 79, Part 1,2023,103861,ISSN 1746-8094,https://doi.org/10.1016/j.bspc.2022.103861.
[37] Ali TM, Nawaz A, Ur Rehman A, Ahmad RZ, Javed AR, Gadekallu TR, Chen C-L and Wu C-M (2022) A Sequential Machine Learning-cumAttention Mechanism for Effective Segmentation of Brain Tumor. Front. Oncol. 12:873268. doi: 10.3389/fonc.2022.873268.
[38] Isensee, F., Jäger, P.F., Full, P.M., Vollmuth, P., Maier-Hein, K.H. (2021). nnU-Net for Brain Tumor Segmentation. In: Crimi, A., Bakas, S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2020. Lecture Notes in Computer Science(), vol 12659. Springer, Cham. https://doi.org/10.1007/978-3-030-72087-2_11.
[39] Zeineldin, R.A., Karar, M.E., Coburger, J. et al. DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. Int J CARS 15, 909–920 (2020). https://doi.org/10.1007/s11548-020-02186-z.
[40] Pei, L., Vidyaratne, L., Rahman, M.M. et al. Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. Sci Rep 10, 19726 (2020). https://doi.org/10.1038/s41598-020-74419.
[41] Silva CA, Pinto A, Pereira S, Lopes A. Multi-Stage Deep Layer Aggregation for Brain Tumor Segmentation. Int MICCAI Brainles Worksh (Spring) (2020) 179–88. doi: 10.1007/978-3-030-72087-2_16.
[42] Colman, Jordan & Zhang, Lei & Duan, Wenting & Ye, Xujiong. (2021). DR-Unet104 for Multimodal MRI Brain Tumor Segmentation. 10.1007/978-3-030-72087-2_36.
[43] Myronenko A. 3d Mri Brain Tumor Segmentation Using Autoencoder Regularization. Int MICCAI Brainles Worksh (Spring) (2018) 311–20.
[44] Pamungkas, Yuri, Evi Triandini, Wawan Yunanto, & Yamin Thwe. "Impact of Hyperparameter Tuning on ResNet-UNet Models for Enhanced Brain Tumor Segmentation in MRI Scans." International Journal of Robotics and Control Systems [Online], 5.2 (2025): 917-936. Web. 31 May. 2025. http://dx.doi.org/10.31763/ijrcs.v5i2.1802.
[45] Lee, Y.; Sim, W.; Park, J.; Lee, J. Evaluation of Hyperparameter Combinations of the U-Net Model for Land Cover Classification. Forests 2022, 13, 1813. https://doi.org/ 10.3390/f13111813.
[46] Jwaid, W. M., Al-Husseini, Z. S. M., Sabry, A. H. (2021). Development of brain tumor segmentation of magnetic reso nance imaging (MRI) using U-net deep learning. Eastern-European Journal of Enterprise Technologies, 4 (9 (112)), 23–31. doi: https://doi.org/10.15587/1729-4061.2021.238957.
[47] Khalaf, Muna and Dhannoon, Ban N. (2022) "MSRD-Unet: Multiscale Residual Dilated U-Net for MedicalImage Segmentation," Baghdad Science Journal: Vol. 19: Iss. 6, Article 48.DOI: https://doi.org/10.21123/bsj.2022.7559.
[48] Gupta, Aitik & Dhar, Joydip. (2023). A hybrid approach for improving U-Net variants in medical image segmentation. 10.48550/arXiv.2307.16462.
[49] Kuang, Zhuodiao. (2022). Transfer Learning in Brain Tumor Detection: from AlexNet to Hyb-DCNN-ResNet. Highlights in Science, Engineering and Technology. 4. 313-324. 10.54097/hset.v4i.919
[50] Akinyelu, A.A.; Zaccagna, F.; Grist, J.T.; Castelli, M.; Rundo, L. Brain Tumor Diagnosis Using Machine Learning, Convolutional Neural Networks, Capsule Neural Networks and Vision Transformers, Applied to MRI: A Survey. J. Imaging 2022, 8, 205. https://doi.org/ 10.3390/jimaging8080205.

[51] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770- 778.
[42] D. Gutman, N.C.F. Codella, E. Celebi, B. Helba, M. Mar chetti, N. Mishra, A. Halpern, Skin lesion analysis toward melanoma detection: a challenge at the international sym posium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC) 2016, Eprint ArXiv1605, 2016: pp. 3e7, 01397, http://arxiv.org/abs/1605. 01397.
[53] N. Codella, V. Rotemberg, P. Tschandl, M.E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, A. Halpern, Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC), ArXiv Prepr. ArXiv1902, 2019, 03368, http://arxiv.org/abs/1902.03368..
[54] Brain Tumor Segmentation (BraTS2020). (n.d.). Available online: https://www.kaggle.com/datasets/awsaf49/brats2020- training-data (accessed on 11 March 2023).
[55] Träff, H. (2023). Comparative Analysis of Transformer and CNN Based Models for 2D Brain Tumor Segmentation (Dissertation). Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-195095.
[56] https://kaggle.com/mateuszbuda/lgg-mri-segmentation.
[57] Wang R, Lei T, Cui R, Zhang B, Meng H, Nandi AK. Medical image segmentation using deep learning: A survey. IET Image Process. 2022; 16(5): 1243–67.