Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A 2.5D Attention-Enhanced Mask R-CNN with Multi-Scale Feature Fusion for Brain Tumor Segmentation from MRI Images

M GANESHKUMAR

ganeshmca1215@gmail.com

Kalasalingam Academy of Research and Education

K Karthigadevi

Kalasalingam Academy of Research and Education

A Nesarani

Koneru Lakshmaiah Education Foundation

Research Article

Keywords: Brain Tumor, Segmentation, MRI Images, BraTS2020, Mask R-CNN

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8829296/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A 2.5D Attention-Enhanced Mask R-CNN with Multi-Scale Feature Fusion for Brain Tumor Segmentation from MRI Images

Ganeshkumar M [1], Karthigadevi K[1], Nesarani A[3]

[1]Department of Computer Applications, Kalasalingam Academy of Research and Education (Deemed to be University), Krishnankoil, Tamilnadu– 626126.,

[3] Department of Computer Science and Engineering, Koneru Lakshmaiah Educational Foundation (Deemed to be University), Vadeeswaram, Andra Pradesh– 522502

ganeshmca1215@gmail.com, k.karthikrish@gmail.com,nesaraniabraham84@gmail.com

Abstract

Advancements in Image processing technology play a crucial part in evaluating the pre-operative condition of tumor morphological state and hence in biopsy guidance. Neoplasm Segmentation is significant in radiotherapy treatment, progressive monitoring of disease and for stage prediction. Due to a number of circumstances, manual diagnosis is time-consuming; therefore, automatic picture segmentation is required. Image processing frequently involves clinical data from magnetic resonance imaging (MRI). In this study, we suggested a Mask Region Convolutional Neural Network (RCNN) using an Inception U-Net architecture to carry out autonomous brain tumor segmentation. This model uses U-NET with Dilated Inception convolution as the Backbone network. Information on both a local and global scale is included in the feature map collected from the backbone network. Our model is experimented with the BraTS 2020 dataset and segmentation is carried out with the glioma portion, including tumor core, enhancing core and whole tumor, which are evaluated with various statistical parameters. Our model produces an accuracy of 96.01% for segmentation, proving the proposed approach overwhelms the comparative models with $p<0.05$.

Keywords: Brain Tumor, Segmentation, MRI Images, BraTS2020, Mask R-CNN

1. Introduction

One of the most fatal diseases brought on by abnormal brain growths is the brain tumor. Depending upon its origin and position, it affects the normal operation of human organs. Primarily, it is classified as primary and secondary tumor [1-2] based on its position. According to the survey of NBTF, the USA is estimated to have nearly 30000 cases of primary brain tumor and the UK has 42000, respectively [3].

Gliomas are one of the primary malignancies in Brain affecting the brain, affecting human health of human. Blood clots and abnormal tissues are some of the common brain abnormalities. Magnetic resonance imaging (MRI) is used to visualise the heterogeneous structure of these tumors. [4]. But these Conventional MR images don't provide significant information due to heterogeneities and variations in their structure. The limitation encountered with these conventional MR images is that they do not provide sufficient tumor physiological information. Survival rates get better on behalf of swift detection and treatment of brain tumors. Moreover, accurate segmentation is demanded for better treatment by doctors. This facilitates the need of advanced computerised strategy for precise brain tumor segmentation. Due to changes in factors like shape, position, size, and boundaries different from typical tissues, it is still a difficult task.

Deep Learning Techniques have gained importance nowadays for accurate determination of Image data analysis [5]. Most important deals with accurate segmentation of Brain tumor for efficient pre-diagnosis and treatment and post surveillance medication [6]. Multi-layered DL techniques with Convolution Neural Network (CNN) finds optimum solution for the complicated medical image segmentation tasks over the traditional methods[7-9]. Many CNN models as Google Net[10],U-Net [11], AlexNet [12], VGG[13], Dense Net[14] and Fully Convolution Network(FCN)[15] exists. Among these models, U-Net gains importance with its encoder –decoder structure.

To extract local and global contextual information regarding an image at the pixel level, lots of deep learning models have been established. One such architecture considered here is the U-Net structure, which captures high-resolution contextual information in its contracting path through down-sampling and improves output resolution through up-sampling [16]. U-Net and its variants GRA U-Net, 3D U-Net, RIC-U-Net, PsLSNet and SDRes U-Net, have been used in the segmentation task nowadays. U – Net structure aggregates the feature map from the contracting side to the expanding side using skip connections. The encoder-decoder structure is used at contracting and expanding side. The drawback of using Skip connection is about forwarding extra features in the forwarding path and the presence of large semantic deviation between the feature map in the encoding-decoding path and limited feature learning. This results in inaccurate segmentation prediction [17-18]. This problem is overcome by introducing a modified encoder-decoder with a dense dilated inception U-Net with a dense path instead of a skip connection between the encoder and decoder circuit. A bottleneck layer plus a convolution layer with 3x3 filters make up the dense path. This permits low-level feature extraction with the encoder and missing spatial information retrieval with the decoder. This dense path improves information flow,

reduces overfitting and improves overall gradient in the network. All these segmentation results with high computational cost.

Mask R-CNN, a variant of CNN, trains the model at less computation cost and includes an object mask along with bounding box regression. Mask R-CNN works as a stage prediction system, with the Region Proposal Network (RPN) acting as the first stage and each Region of Interest (ROI) possessing its own class, bounding box, and binary masks in the final stage [19]. The study designed for this paper uses a mask R CNN with a Dilated Inception module as the backbone network, followed by masked segmentation output, bounding box, and class.

The contribution of the work is as follows.

1) Used to extract multi-Scaled features using a Dilated U-Net as ac Backbone Network with end-to-end tumor segmentation.
2) Brain tumor image localisation and segmentation using Mask R CNN at lower computing cost.
3) Increase the segmentation accuracy of brain tumors.

The paper is organized as follows: Section 2 details the existing research work undertaken by different researchers and the method used. Section 3 details the working of the proposed modified Dilated U-Net Inception network with region proposal network and bounding box regression. Section 4 shows the experimental method employed in the segmentation process with detailed result analyzed with statistical parameter and the section concludes the research findings.

2. Literature Survey

The research carried out with the U-Net model and its different versions is bundled up in the section that follows.

The research on medical imaging centres on extracting features from Regions of Interest (ROI) to accelerate planning, diagnosis and disease treatment. Compared to traditional approaches, deep learning strategies achieve accurate and satisfactory results in segmentation. Researchers primarily focus on pixel-level segmentation, which are classified as patch-level and semantic segmentation. Every pixel of the image is extracted into an N X N patch during patch-level segmentation. Smaller patches just gather local information, whereas larger patches gather more spatial details. On the other hand Semantic segmentation tries to fix every pixel to a class label. It incorporates an encoder-decoder structure at its down sample and upsample side to maintain same sized image at it input and output side. Many research works using patches and semantic segmentation of image is listed in Table 1.

The work done by [20] performs segmentation on a cortical image and found that an orthogonal 2D patch performs with a DSC of 0.725. Another work by [34] uses segnet for Brain tumor segmentation for subcortical with MICCAI, IBSR, IXI dataset and achieves a better dice coefficient

Table 1. Overview of Segmentation carried out with a deep learning technique

Segmentation Type	Input Dimension	Deep Learning Technique	Image data	Dataset	Evaluation Metric	Authors
Patch	2D/3D	SegNet-CNN	Cortical	MICCAI	Error rate, Dice Score	[20]
Patch	2D	CNN	White/Grey Fluid	Clinical	Dice Score	[21]
Patch	2D	CNN	Cortical	MRI	Dice Score	[22]
Patch	2D	CNN	High Glioma/Low Glioma	BRATS	Dice Score	[23][24][25]
Patch	2D	RCNN/ANFIS	Sclerosis	MICCAI	Dice Score	[26]
Semantic	2D	mFCN	White/Grey Fluid	Clinical	Average ,Dice Score	[27]
Semantic	2D	U-Net FCN	High Glioma/Low Glioma	BRATS	Dice Score	[28]
Semantic	2D	U-Net FCN	High Glioma/Low Glioma	BRATS	Dice Score	[28]

Semantic	2D	U-Net FCN	White/Grey	MICCAI	Dice Score	[29]
Semantic	2D	FCN	Cortical tissue	ABIDE,IBSR	Dice Score	[30][31]
Semantic	2D	SC U-Net	Ischemic Lesion	ISLES	Dice Score	[32]
Semantic	2D	Inception U-Net	High Glioma/Low Glioma	BRATS	Dice Score	[33]

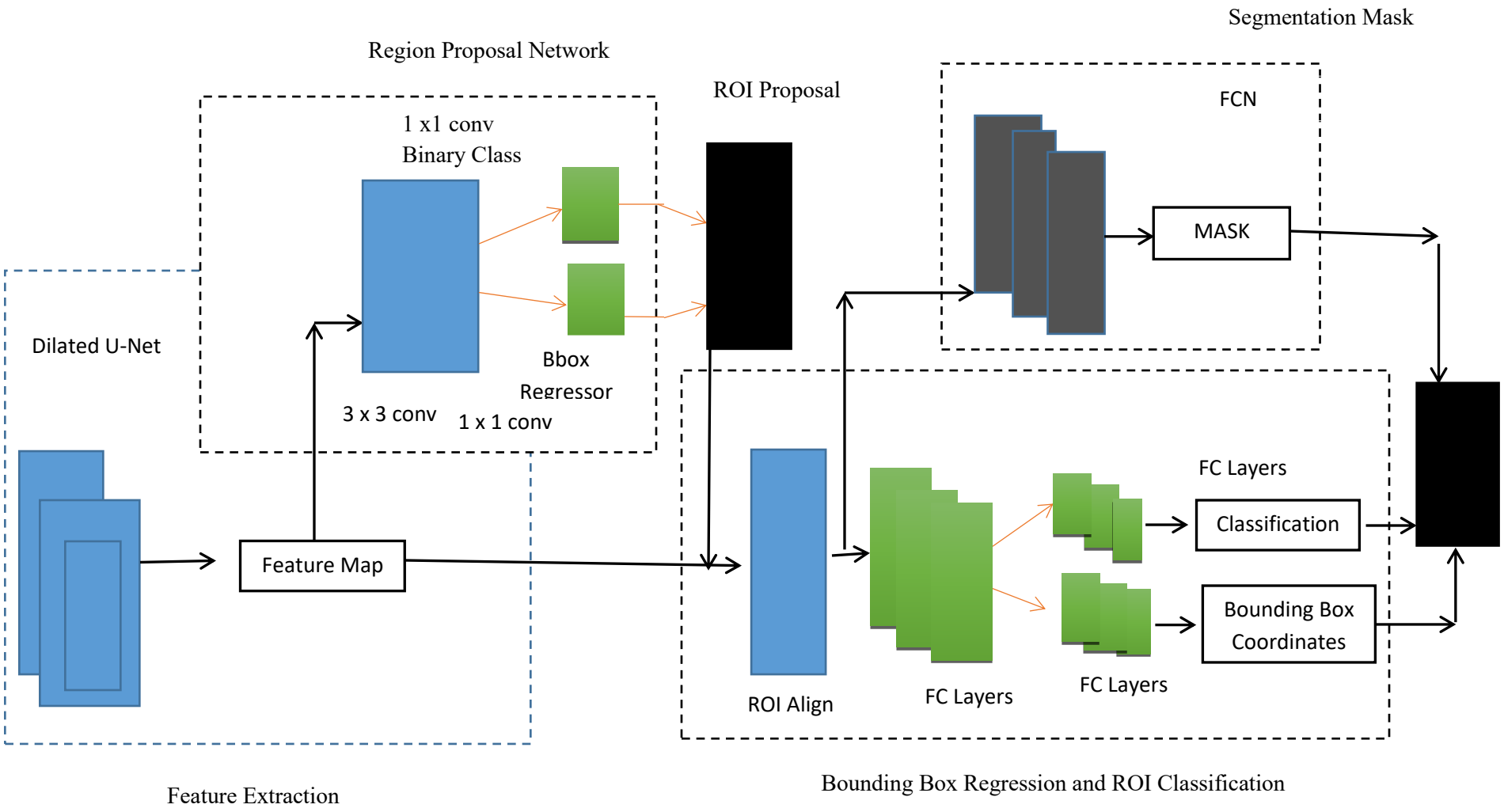

Fig 1. The Framework of the Proposed Network

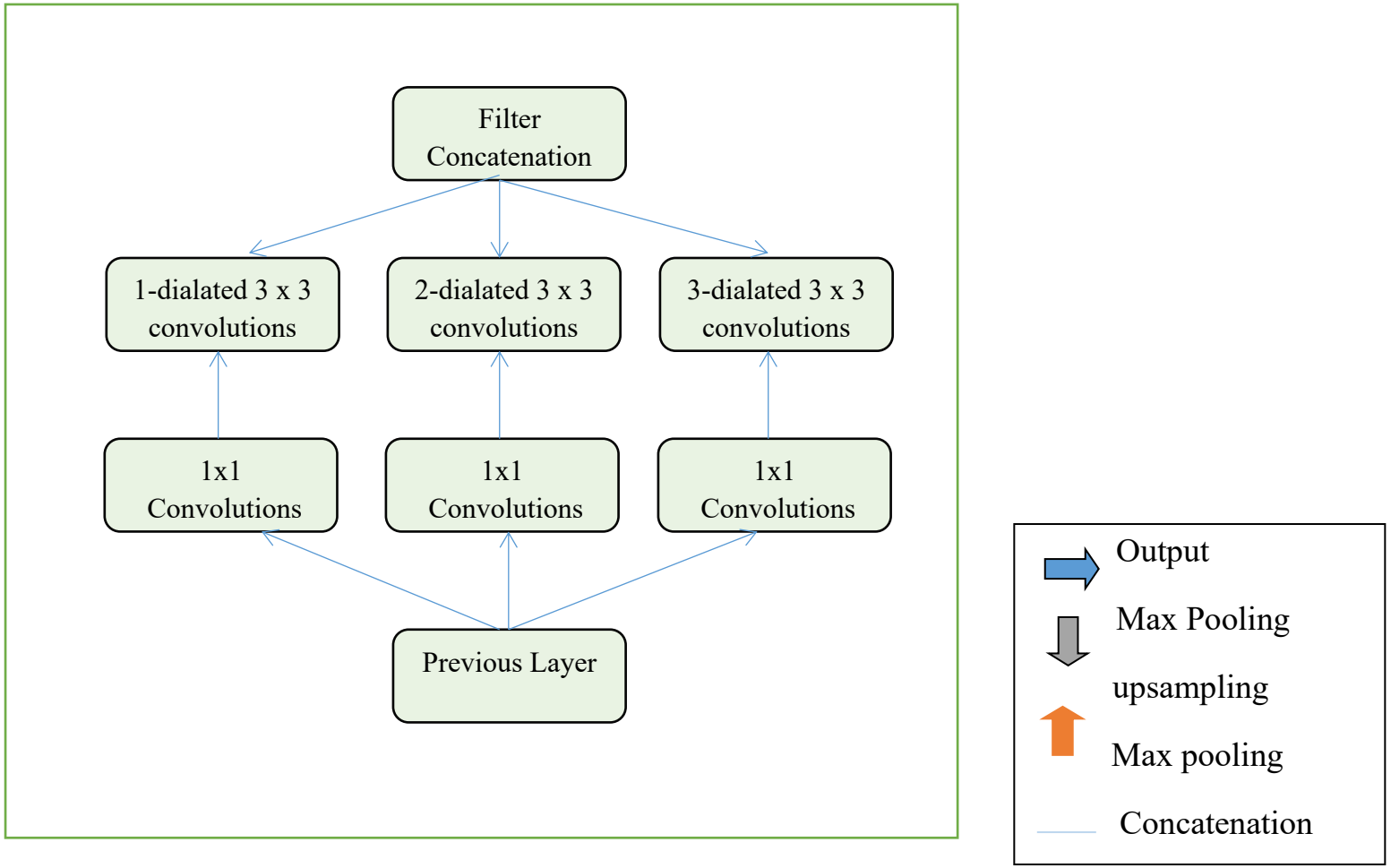

Figure 2: Dilated Inception Module showing three l-convoluted filters and 1x1 dimension reduction filters

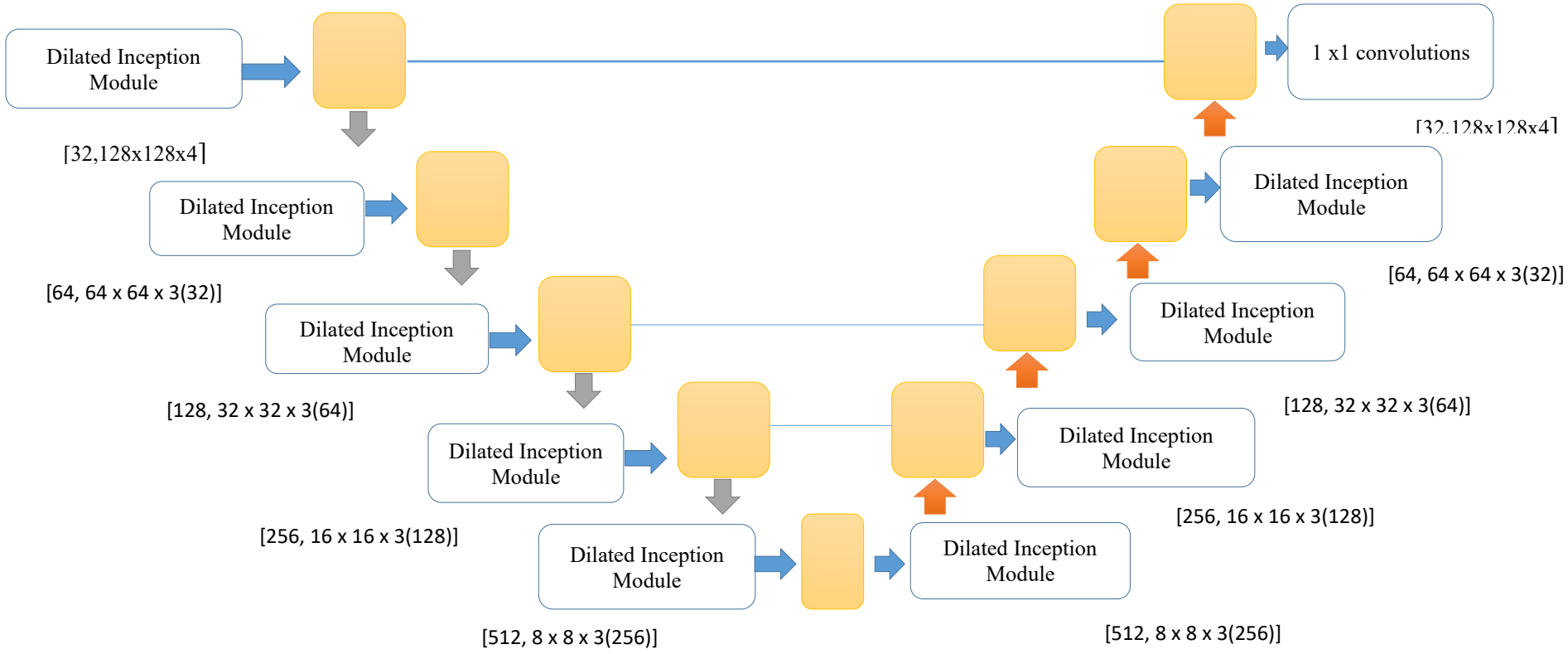

Figure 3: Structure showing Expanding and Contracting path of DIU-Net Structure

3. Proposed Methodology

3.1.1 Preprocessing: The Preprocessing is performed to achieve an enhanced image after median filtering and bias field correction.

3.1.2 Localization and Segmentation after Preprocessing. Tumor Localization and classification either into tumor or non-tumor is performed using Mask R CNN.

3.2 Structure of the proposed methodology

The section describes the architecture used in the proposed work, as shown in Figure 1. The proposed network has three layers i) Backbone Network ii) Region Proposal network iii) Bounding Regression and ROI Classification and final masked segmented output. The following section describes the functions of each layer

3.2.1 Backbone Network

The input image was processed using the backbone network, Dilated Inception U-Net, to extract multi-scale features. It outputs feature maps and fed to the Region Proposal Network to generate different-sized anchors. These anchors represent bounding boxes of different sizes corresponding to the entire image. Bounding box regression generates IOU (Intersection over Union) to predict whether the anchor has a higher IOU than the ground truth box (GT). If the anchor has a higher IOU it is classified as positive, else negative.

3.2.1.1 Dilated Inception U-Net (DIU-Net) Architecture

Long Inception Dilated convolution and inception modules are integrated into the U-Net architecture, as in Figure 2. The integrated inception and dilated network improves segmentation with $p<0.05$. The used Dilated U-net has a contracting-expanding path with a bottleneck in the middle, as in Figure 3. The use of filters doubles at the contracting side on each layer and halves on expanding path thereby the feature map is halved at contracting path and doubled at the expanding side. Upsampling is done at the expanding path and down sampling with maximum pooling is done at the contracting side. Further Features from the contracting side are concatenated correspondingly with those of the expanding side. At the final layer on the expanding side, the image's original metrics are achieved. To reduce the depth of the final feature map, a 1 X 1 convolution is carried out based on the number of segmentation classes.

3.2.1.2 Dilated Convolution

The image L considered here has size of $m * n$ with filter f of size $k*k$.. The linear convolution operation is given as below.

$$(L*f)(p) = \sum L[p+s]f[s] \quad (1)$$

The Generalization of simple convolution operation is given below.

$$(L*_l f)(p) = \sum L[p+ls]f[s] \quad (2)$$

We get simple convolution operation for l=1.when l>1,l-1 zeros are inserted between each filter elements generating $k_s * k_s$ sparse and scaled filters.

$$K_s = k+(k-1)(l-1) \quad (3)$$

$$= l(k-1)+1 \quad (4)$$

K_s is the receptive filter, and it grows linearly with l while the element count stays constant. The following equations produce the linear improvement.

$$K_s/k = k+(k-1)(l-1)/k \quad (5)$$

3.2.2 Region Proposal Network

To create an ROI (region of Interest), the backbone network's output feature map is delivered into the region proposal network. The image is scanned by a 3x3 convolution network with a sliding window that generates anchors representing bounding boxes of different sizes covering the entire image. From this Intersection over Union (IOU) is calculated for each bounding box. The higher the value, the more positive the image is compared to the ground truth image; otherwise, it is seen as negative.

3.2.3 Bounding Box Regression and ROI Classification

The Bounding Box Regression further refines the bounding boxes to fix the location and size of the bounding box to cover the tumor region. Here ROI align layer is used

to resize the downscaled feature maps and generates fixed length feature vectors for the defined candidate regions.

3.2.4 Segmentation

The positive ROI and feature map are considered as input and this network outputs binary classified segmented images.

4 Experiments

The following section discusses the input image source considered and discusses the obtained results in detail.

4.1 Testing Device and Dataset

The proposed approach is tested with a device configuration of NVIDIA TESLA P100 implemented with Python 3.8 Keras running TensorFlow background on Intel i7,2.60 GHz,16 GB RAM. The strategy consists a) Brain tumor Localization b) Brain tumor Localization. The method accepts both colour and grey coloured images with an initial fixed learning rate, which is dynamically adjusted. In Brain tumor segmentation phase DIU-Net model is used for image segmentation. The proposed approach is tested with BRATS 2020 dataset comprising MRI images .The initial image size is 128 X 128 X 4

4.2 Evaluation Metrics

The proposed methodology is evaluated with the following statistical metrics: Pixel Accuracy, Precision, Recall, Dice Similarity coefficient and Jaccard Index are mentioned below:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$\text{Dice Similarity Index(DSC)} = \frac{2TP}{2TP+FP+FN} \quad (7)$$

$$\text{Jaccard Index} = \frac{TP}{TP+FP+FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (10)$$

In equations (6)-(10), TP-True Positive represents a correctly segmented region, FP -False Positive represents an invalidly segmented region and FN- False Negative represents a missed segmented region. Further, the approach is evaluated with the Receiver Operating Curve (ROC). It is the curve that is used to compare the result of picture segmentation with the ground truth value. To plot the ROC, curve Specificity (False Positive rate) and Sensitivity (True Positive rate) are used. The other metrics evaluated here provide higher value than the other compared approaches.

4.3 Input Image

In this investigation, the BRATS 2020 dataset—which includes MRIs of 200 high-grade and 70 low-grade gliomas—is employed. Preprocessed versions of the taken BRATS dataset are available. The photos are resampled with skull-stripped, isotropic picture resolution of 1 mm3. Some additional preprocessing steps have been taken, which include 1) Image resizing to 128 X 128 pixel 2) eliminating the non-tumor portion in the GT segmentation, 3) mapping pixel values between 0 to 255 4) performing z-score normalization to each image.

Four different MRI modalities make up the patient data. pictures of the fluid-attenuated inversion recovery, T1 with gadolinium enhanced contrast, T1C, and T2 weighted (T2) images as shown in Figure 4. Three intra-tumoral features are also marked on the BRATS image: the necrotic and non-enhancing tumour core (labelled as 1), the peritumoral

edoema (labelled as 2), and the enhancing tumour (labelled as 4). The Whole tumour (WT), Tumour Core (TC), and Enhancing tumour (ET) are the three subregions of gliomas that are defined from the intra-tumoral subregions. While ET only projects an increasing tumour region and TC only specifies a tumour region lacking peritumoral edoema, WT includes all three tumour regions. The input image size to Dilated U-Net is P x Q x R pixel image, where P=Q= 128 pixels and R=4 modalities in MRI. P x Q x K pixel output image is obtained, where K points to number of segmentation classes. Each chunk of K represents a predicted segment of the jth class, where 0≤ j≤K-1.

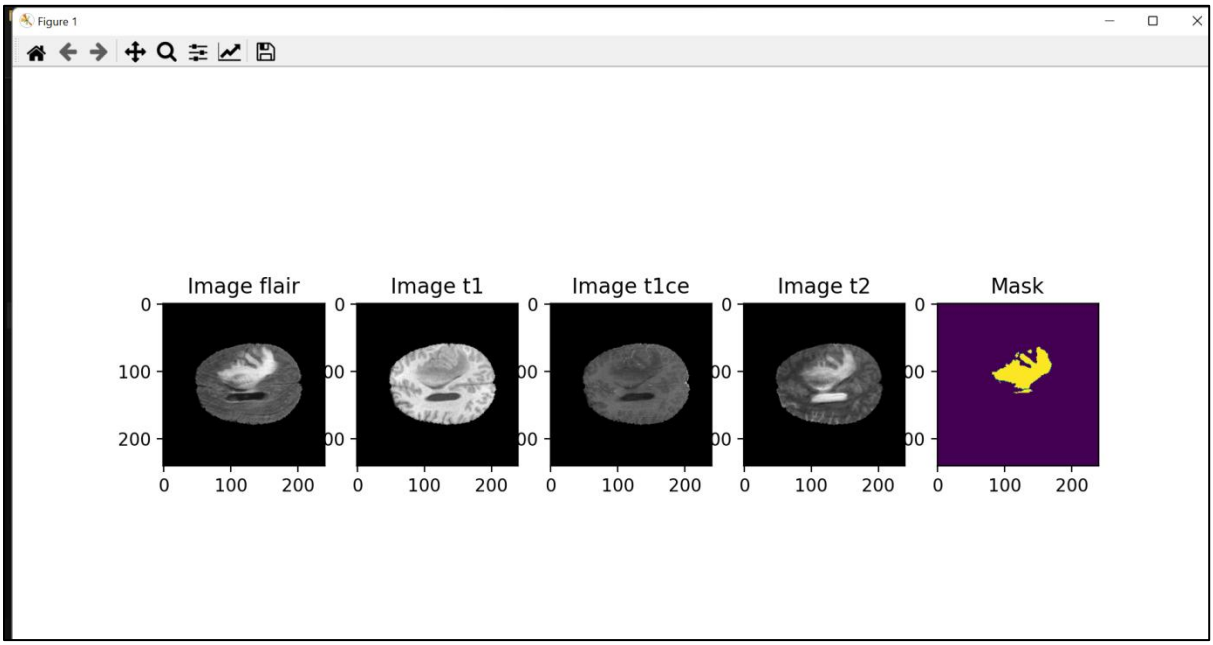

Figure 4: Nifti data showing a slice of each 3D image

Figure 5 shows BraTS dataset used in the training process of brain tumor image segmentation and Fig 6a, 6b show the segmented image trained at consecutive time periods for each of the above sliced image.



Figure 5 : BraTS Training Dataset

4.4 Evaluation of Brain Tumour Localization and Segmentation

Four-Fold cross-validation is performed randomly over BraTS2020 dataset with adherence to the standard of cross-validation. The training iterations are stopped at 100.00 for each fold with the proposed Mask R-CNN -DIU-Net structure. The average testing time has been around 20-25 sec.

4.5 Quantitative and Qualitative Results of Brain Tumour Localization and Segmentation Phase

Figure 12 represents the ROC curves for the folds 1,2,3 and 4 by the proposed methodology and it shares the margin of ground truth and original segmentation image. ROC values for four folds are 95.2%,95.13%,95.02% and 95.3%, respectively. The table represents Quantitative assessment for BraTS2020 dataset and shows the results of the combined phases of Brain Tumour Localization and Segmentation. The other metrics (DSC, JI,PRE,REC,SPE and AUC) are calculated from minimum, maximum and standard deviation values from MRI images. The proposed approach (Mask R-CNN with DIU-Net) results in 95.3% **DSC,95.32% JI, 96.05% PRE, 93.9 %REC, 96.01%**

ACC, 95.99% SPE, and 95.72% AUC. Figure 7-8 shows the qualitative Brain Tumour Segmentation of the proposed approach on MRI images against the ground truth value.

Table 2: Quantitative Results of Performance Evaluation of proposed Mask R-CNN-DIU-Net For Brain Tumor Localization and Segmentation

Fold	DSC%	JI%	PRE%
1	95.30±6.01[85.05,93.45]	95.01±4.62[85.34,93.01]	95.22±4.33[85.12,93.56]
2	95.34±5.77[85.12,94.01]	95.32±5.6[85.1,93.56]	95.018±4.6[85.33,93.22]
3	95.23 ±5.73[85.01,93.99]	95.06±4.77[85.42,93.41]	95.31±5.6[85.12,93.42]
4	95.02 ±4.945[85.59,93.56]	95.4 ±4.4[86.72,93.99]	94.12±4.7[84.69,94.12]
overall,	**95.3 ±4.67[86.59,94.02]**	**95.32 ±4.2[86.01,93.52]**	**96.05±5.25[85.002,93.8]**

Fold	REC%	ACC%	SPE%	AUC%
1	93.00±5.21[83.02,92.27]	94.89±5.03[85.01,93.01]	94.99±5.00[85.11,93.12]	95.2
2	93.42±5.67[83.01,91.98]	94.87±5.01[8.98,93.42]	94.99±4.99[85.22,93.22]	95.13
3	92.51±4.14[83.51,93.23]	95.5±5.25[85.02,93.75]	95.3 ±5.01[85.01,93.69]	95.02
4	90.56±3.43[82.25,92.01]	95.3±5.61[85.12,93.42]	95.37 ±5.2[85.13,93.47]	95.3
Overall	**93.9±6.25[83.2,92.05]**	**96.01±5.5[85.01,93.57]**	**95.99 ±5.34[85.139,93.48]**	**95.72**

4.6 **Results and Discussions**

The suggested approach is evaluated with both the BRATS and Kaggle datasets available online, which shows variations in devices, the background noise, intricate structure, observation angle, skew field effect, etc. The model's performance is assessed using criteria such as precision, recall accuracy, dice-score (DSC), and IOU. With an average

precision of 0.96 on the BRATS dataset and 0.942 on the Kaggle dataset, the suggested strategy increases the accuracy percentage in precisely localising the brain tumour.

The evaluation indicators for both datasets are shown in a boxplot that shows the performance of the proposed model. The suggested technique has an accuracy of about 96.5% over the Kaggle dataset's 94.7% accuracy.

Table 3: Setting the proposed model's parameters

Parameters	Values
Estimator	Adam
Initial Learning rate	10^{-4}
Batch Size	64
Epochs	100
API	Keras with Tensor flow
GPU	NVIDIA TESLA P100

The suggested model's parameter settings are detailed in Table 3. Each of the glioma subregions that were formed from intra tumor structures had a dice score. It is observed that the Dice score for DIU-Net shows significant improvement as 0.945 for WT, 0.96 for TC ($P<0.05$) and 0.90 for ET. There is no statistically significant variation in probability for ET. Additionally, Table 4 represents the validation dice score against epoch for all intratumoral structures, and Table 6 provides a detailed comparison of the proposed model with the other research projects. The proposed work demonstrates superior segmentation results in terms of accuracy, dice score, and IOU. Table 5 details the comparative study of the suggested method with existing RCNN approaches in terms of better accuracy and dice score.

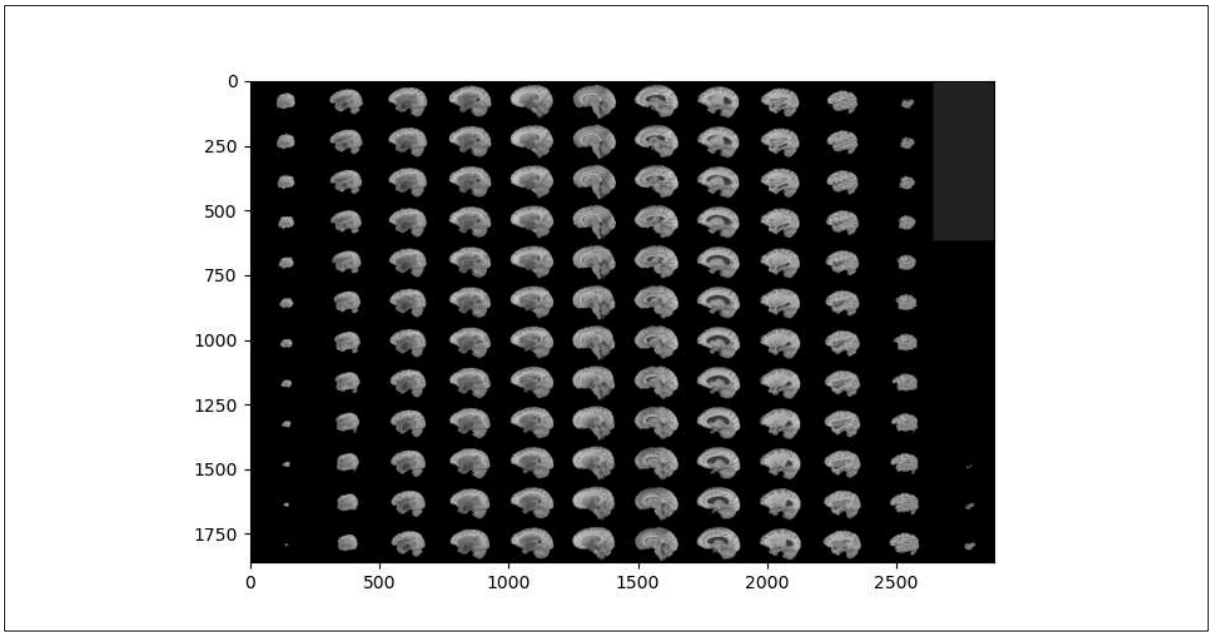

Figure 6a: Segmented Image of all class

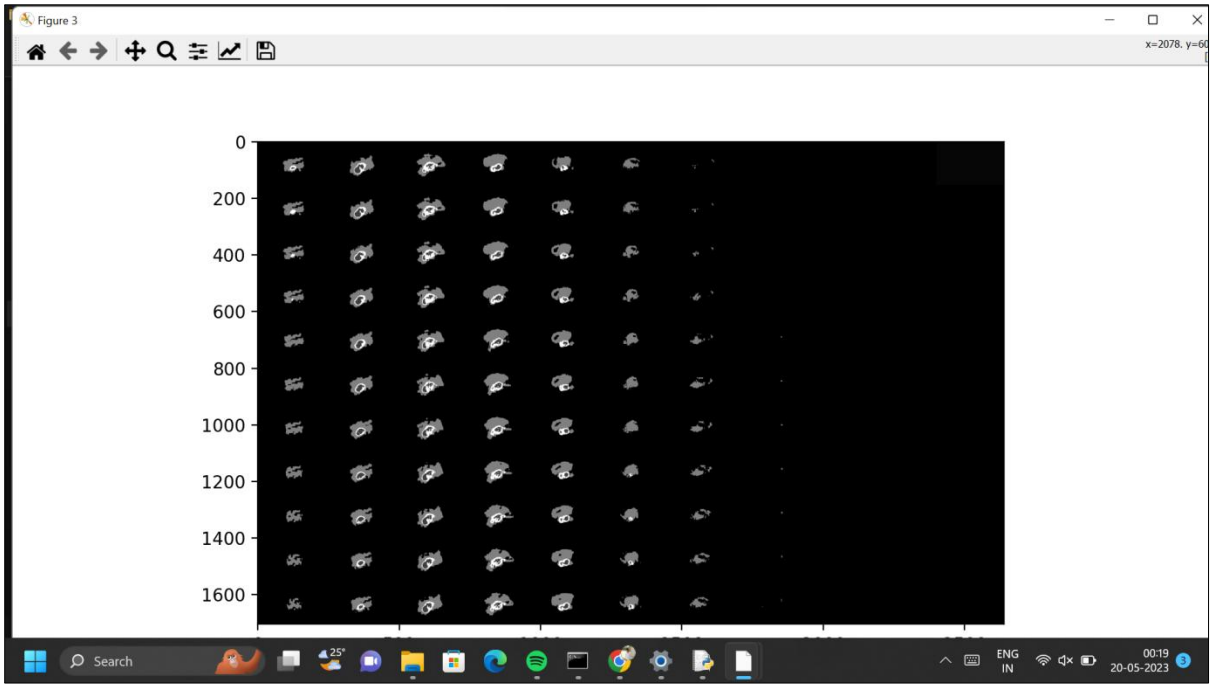

Figure 6b: Segmented Image of all class

Table 4: Comparative Chart of the suggested model with other RCNN approaches

Approaches	Accuracy	Dice Score
RCNN	0.92	0.87
Faster RCNN	0.94	0.91
Mask RCNN	0.95	0.95
Proposed model	0.96	0.953

Table 5: Comparative Chart of the suggested model with other Research Work

Research Works	Accuracy	Dice Score	IOU
[35]	86	-	-
[36]	93.4	0.93	0.76
[37]	-	0.79	-
[38]	91.4		0.94
[39]	-	0.86	-
[40]	-	0.87	0.85
[41]	-	0.88	-
[42]	-	0.86	-
[43]	-	0.88	-
Proposed model	96	0.953	0.95

Figures 7 and 8 show the prediction of all classes of tumor as Necrotic/core, Edema and Enhancing Tumor

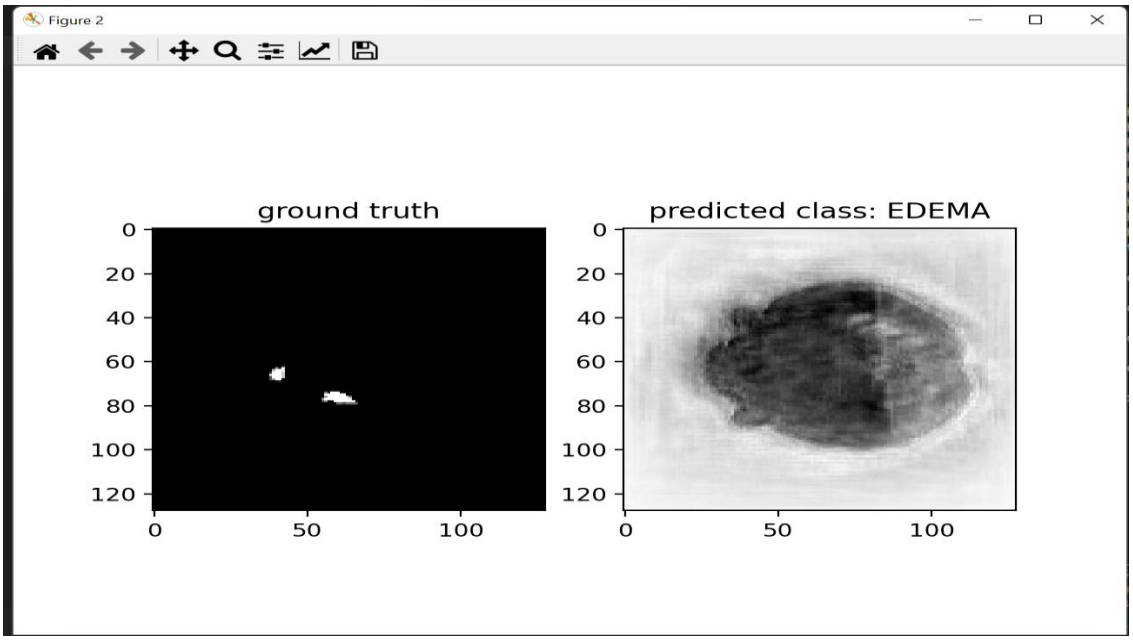

Figure 7: Prediction of Edema Tumor

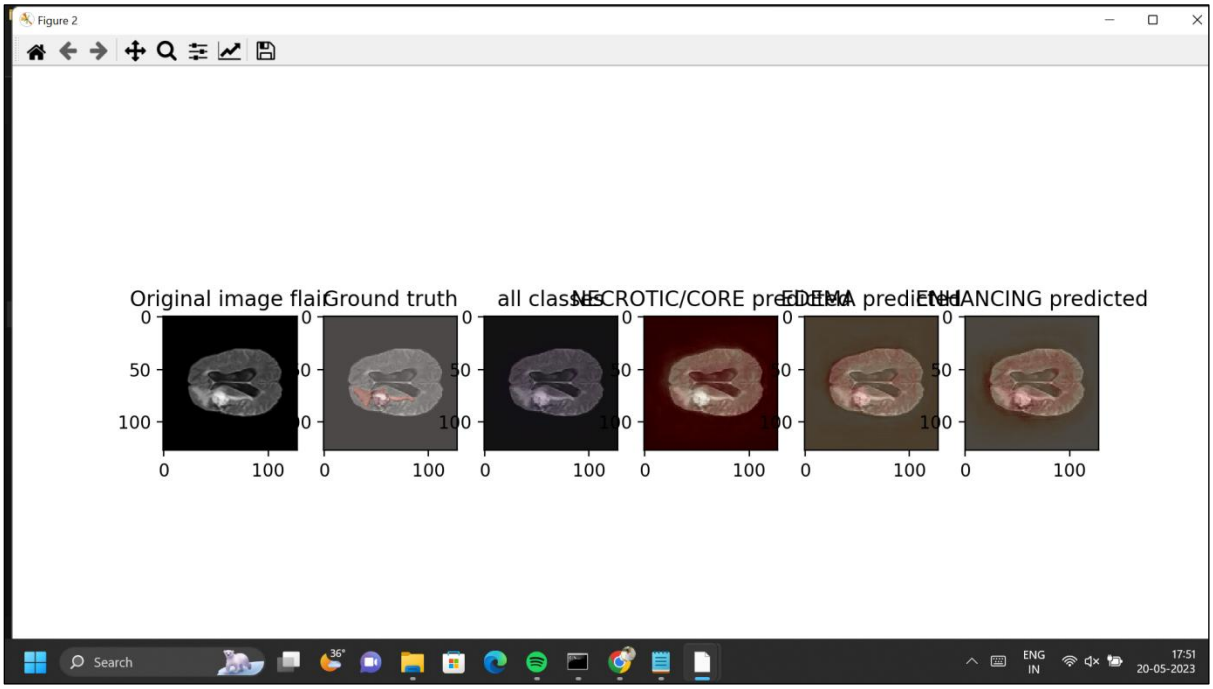

Figure 8 :Prediction of Necrotic/core, Edema and Enhancing Tumor

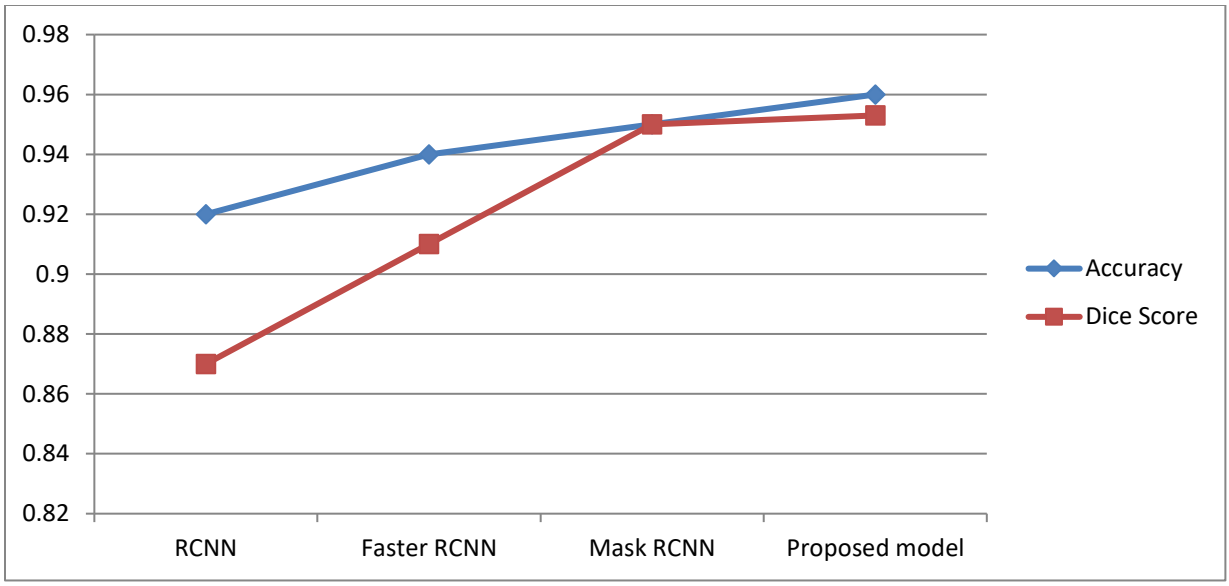

Figure 9: Comparison of the proposed model with other RCNN

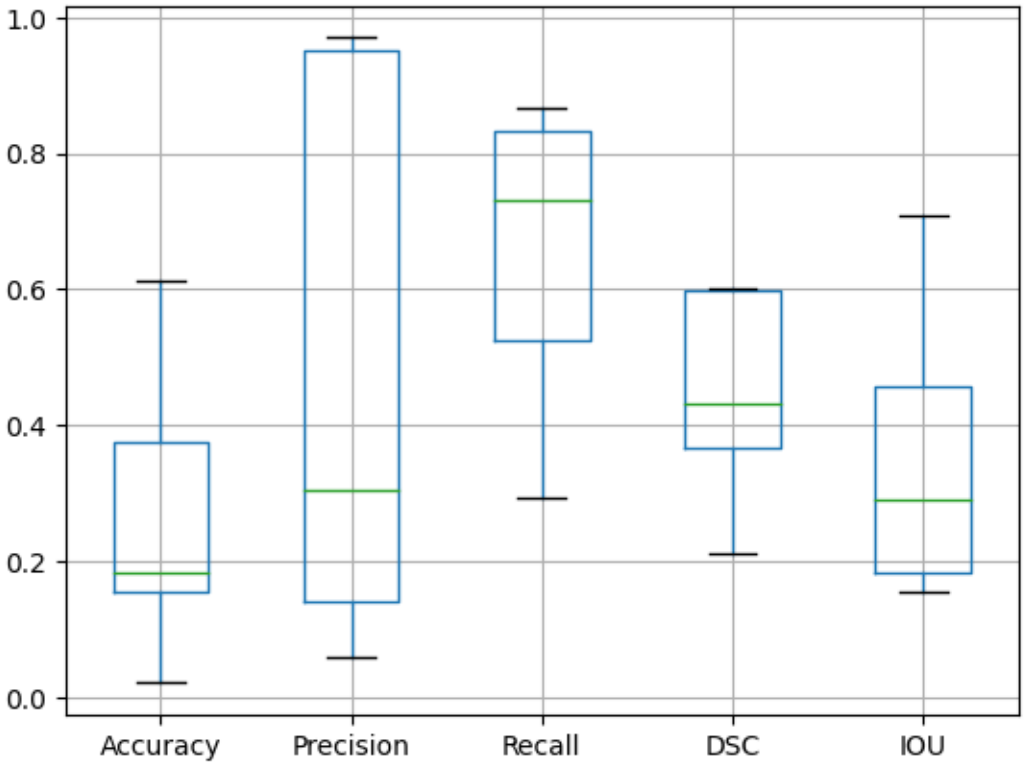

Figure 10: Statistical parameters obtained for the suggested model

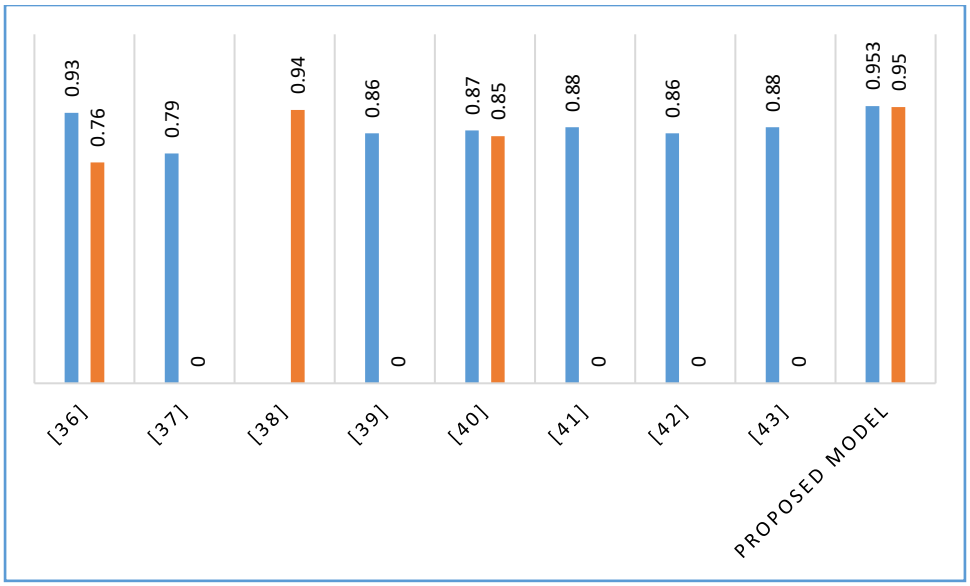

Figure 11: Comparison of the proposed model with other research work with Dice Score and IOU

The accuracy and dice score of the promoted model's performance have been compared with those of other RCNN models for Figure 9. Figure 10 shows the different evaluation metrics that were derived from the suggested model for brain tumor image segmentation. In terms of Dice Score and IOU, Figure 11 compares our models to other research models; the suggested model performs better than the other model in terms of prediction.

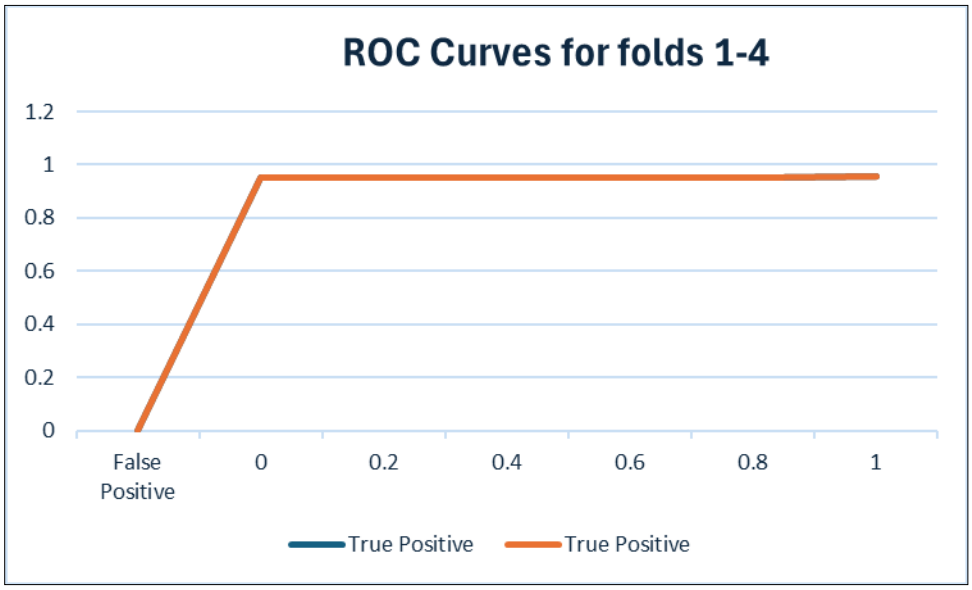

Figure 12: ROC for Brain Tumor Segmentation for folds 1,2,3 and 4

5. Conclusion

In this work, a Mask R-CNN model with DIU-NET for obtaining feature maps is designed. DIU-NET is computationally more efficient, showing a good dice score value with that of the comparable work. The accuracy percentage of delineating the tumor region is improved as segmentation is performed on a localized ROI and the use of the ROI align layer improves the segmentation output. The proposed model efficiency is evaluated with accuracy, dice score and IOU with other existing RCNN approaches and other existing research works. Comparing the suggested method to the other methods already in use, the segmentation results are better.

Conflict of Interest

Authors declared that there is no conflict of interest.

Author's Contribution

Conceptualization: Ganeshkumar M; **Methodology**: Ganeshkumar M; **Validation:** Abraham N; **Formal analysis:** Ganeshkumar M; **Writing– Original Draft**: Ganeshkumar M; Writing – **Review & Editing:** Ganeshkumar M and Abraham N; **Supervision:** Karthigadevi. K; **Final approval:** Karthigadevi. K.

Funding

Not applicable.

Acknowledgments

Not applicable.

Declaration

Competing interests: The authors declare that we have no known competing financial interests or personal relationships that could influence the work reported in this article.

6. References

1. Despotović, I., Goossens, B., & Philips, W. (2015). MRI segmentation of the human brain: challenges, methods, and applications. *Computational and mathematical methods in medicine*, *2015*.
2. Gordillo, N., Montseny, E., & Sobrevilla, P. (2013). State of the art survey on MRI brain tumor segmentation. *Magnetic resonance imaging*, *31*(8), 1426-1438.

3. Goodenberger, M. L. (2012). RB Jenkins. *Genetics of adult glioma. Cancer Genet.*, *205*(12), 613-621.
4. Lloyd, C. (2016). WorldPop Archive global gridded spatial datasets. Version Alpha 0.9.
5. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60-88.
6. Fathallah-Shaykh, H. M., DeAtkine, A., Coffee, E., Khayat, E., Bag, A. K., Han, X., & Nabors, L. B. (2019). Diagnosing growth in low-grade gliomas with and without longitudinal volume measurements: A retrospective observational study. *PLoS medicine*, *16*(5), e1002810.
7. Daimary,D., Bora,B.M.,Amitab,K., & Kandar,D.(2020). "Brain Tumor Segmentation from MRI Images using Hybrid Convolutional Neural Networks," Procedia Computer Science, 167, 2419-2428
8. Ibrahim, S. M., Ibrahim, M. S., Usman, M., Naseem, I., & Moinuddin, M. (2019). Heart segmentation from mri scans using convolutional neural network. *arXiv preprint arXiv:1911.09332*.
9. Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Shen, L., & Alzheimer's Disease Neuroimaging Initiative. (2020). A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *Neuroimage*, *208*, 116459.
10. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
11. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.
12. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.
13. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

14. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
15. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
16. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.
17. Tong, Q., Li, C., Si, W., Liao, X., Tong, Y., Yuan, Z., & Heng, P. A. (2019). RIANet: Recurrent interleaved attention network for cardiac MRI segmentation. *Computers in biology and medicine*, *109*, 290-302.
18. Zhang, Z., Zhang, X., Peng, C., Xue, X., & Sun, J. (2018). Exfuse: Enhancing feature fusion for semantic segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 269-284).
19. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
20. de Brebisson, A., & Montana, G. (2015). Deep neural networks for anatomical brain segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 20-28).
21. Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., & Shen, D. (2015). Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage*, *108*, 214-224.
22. Moeskops, P., Viergever, M. A., Mendrik, A. M., De Vries, L. S., Benders, M. J., & Išgum, I. (2016). Automatic segmentation of MR brain images with a convolutional neural network. *IEEE transactions on medical imaging*, *35*(5), 1252-1261.
23. Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE transactions on medical imaging*, *35*(5), 1240-1251.

24. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical image analysis*, *35*, 18-31.
25. Razzak, M. I., Imran, M., & Xu, G. (2018). Efficient brain tumor segmentation with multiscale two-pathway-group conventional neural networks. *IEEE journal of biomedical and health informatics*, *23*(5), 1911-1919.
26. Sajid, S., Hussain, S., & Sarwar, A. (2019). Brain tumor detection and segmentation in MR images using deep learning. *Arabian Journal for Science and Engineering*, *44*, 9249-9261.
27. Essa, E., Aldesouky, D., Hussein, S. E., & Rashad, M. Z. (2020). Neuro-fuzzy patch-wise R-CNN for multiple sclerosis segmentation. *Medical & Biological Engineering & Computing*, *58*, 2161-2175.
28. Nie, D., Wang, L., Gao, Y., & Shen, D. (2016, April). Fully convolutional networks for multi-modality isointense infant brain image segmentation. In *2016 IEEE 13Th international symposium on biomedical imaging (ISBI)* (pp. 1342-1345). IEEE.
29. Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In *Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings 21* (pp. 506-517). Springer International Publishing.
30. Li, H., Jiang, G., Zhang, J., Wang, R., Wang, Z., Zheng, W. S., & Menze, B. (2018). Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images. *NeuroImage*, *183*, 650-665.
31. Li, Q., Gao, Z., Wang, Q., Xia, J., Zhang, H., Zhang, H., & Li, S. (2018). Glioma segmentation with a unified algorithm in multimodal MRI images. *IEEE Access*, *6*, 9543-9553.
32. Dolz, J., Desrosiers, C., & Ayed, I. B. (2018). 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *NeuroImage*, *170*, 456-470.
33. Li, H., Li, A., & Wang, M. (2019). A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. *Computers in biology and medicine*, *108*, 150-160.

34. Mehta, R., Majumdar, A., & Sivaswamy, J. (2017). BrainSegNet: a Convolutional neural network architecture for automated segmentation of human brain structures. *Journal of Medical Imaging*, *4*(2), 024003-024003.
35. Qasem, S. N., Nazar, A., Qamar, A., Shamshirband, S., & Karim, A. (2019). A Learning Based Brain Tumor Detection System. *Computers, Materials & Continua*, *59*(3).
36. Naz, A. R. S., Naseem, U., Razzak, I., & Hameed, I. A. (2019). Deep autoencoder-decoder framework for semantic segmentation of brain tumor. *Aust. J. Intell. Inf. Process. Syst*, *15*(4).
37. Sobhaninia, Z., Rezaei, S., Noroozi, A., Ahmadi, M., Zarrabi, H., Karimi, N., & Samavi, S. (2018). Brain tumor segmentation using deep learning by type specific sorting of images. *arXiv preprint arXiv:1809.07786*.
38. Rayhan, F. (2018). Fr-mrinet: A deep convolutional encoder-decoder for brain tumor segmentation with relu-RGB and sliding-window. *Int. J. Comput. Appl*, *975*, 8887.
39. Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In *Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings 21* (pp. 506-517). Springer International Publishing.
40. Hasan, S. K., & Linte, C. A. (2018, October). A modified U-Net convolutional network featuring a nearest-neighbor re-sampling-based elastic-transformation for brain tissue characterization and segmentation. In *2018 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)* (pp. 1-5). IEEE.
41. Yang, C., Guo, X., Wang, T., Yang, Y., Ji, N., Li, D., & Ma, T. (2019, July). Automatic brain tumor segmentation method based on modified convolutional neural network. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 998-1001). IEEE.
42. Rodríguez Colmeiro, R. G., Verrastro, C. A., & Grosges, T. (2017, September). Multimodal brain tumor segmentation using 3D convolutional networks. In *International MICCAI Brainlesion Workshop* (pp. 226-240). Cham: Springer International Publishing.

43. Beers, A., Chang, K., Brown, J., Sartor, E., Mammen, C. P., Gerstner, E., & Kalpathy-Cramer, J. (2017). Sequential 3d u-nets for biologically-informed brain tumor segmentation. *arXiv preprint arXiv:1709.02967*.