DEPARTMENT OF COMPUTER SCIENCE AND
ENGINEERING
INDIAN INSTITUTE OF INFORMATION
TECHNOLOGY, DESIGN AND
MANUFACTURING, KANCHEEPURAM
CHENNAI - 600127

Synopsis Of

Deep Learning Approaches for Brain Tumor Segmentation Using Magnetic Resonance Images

A Thesis

To be submitted by

SUBIN SAHAYAM M

For the award of the degree

Of

DOCTOR OF PHILOSOPHY

1 Abstract

Brain tumors are abnormal growths of cells that develop within the brain tissue. They can originate from different cell types within the brain and can be benign (non-cancerous) or malignant (cancerous). Brain tumors can interfere with normal brain function and can cause a variety of symptoms depending on their location, invasiveness, growth rate, and size. These tumors can be identified using imaging techniques such as magnetic resonance imaging (MRI), computed tomography (CT), and positron-emitted tomography (PET). They can further be classified based on their histological characteristics, molecular profiles, and location within the brain. Brain tumors are a significant health concern with profound implications for patient diagnosis, treatment, and overall prognosis. Early detection, accurate diagnosis, and appropriate treatment are essential in managing brain tumors and improving patient outcomes.

In recent years, advanced medical imaging techniques, such as magnetic resonance imaging (MRI), have played a pivotal role in the non-invasive assessment of brain tumors. These imaging modalities provide detailed anatomical and functional information, enabling clinicians to visualize and analyze the extent, location, and characteristics of brain tumors. MRI is a powerful medical imaging technique that uses a combination of magnetic fields and radio waves to generate detailed and high-resolution images of the internal structures of the body, including the brain. It provides a non-invasive method for visualizing the soft tissues, organs, and structures within the brain, allowing healthcare professionals to evaluate and diagnose various neurological conditions and abnormalities. MRI scans offer better contrast between different types of brain tissue than other imaging techniques, enabling the detection of tumors, lesions, blood vessel abnormalities, and other structural anomalies. Based on the mode of operation, MR scanners can generate T1-weighted, T2-weighted, Fluid attenuated inversion recovery (FLAIR), and contrast-enhanced T1-weighted (T1CE) images focusing on different information like tissues, fat, contrast agents, and water. MR imaging modalities are crucial in guiding treatment planning, monitoring disease progression, and assessing treatment response in patients with brain disorders.

Manual identification of brain tumors in MR images is laborious, time-consuming, and human error-prone. Automatic segmentation of brain tumors from MR images aims to bridge the gap. The integration of advanced image processing algorithms and machine learning techniques has shown great promise in enhancing the accuracy and efficiency of brain tumor detection, segmentation, and classification. In recent years, deep learning models have been the go-to approach for segmenting brain tumors. U-Net and its' variants for semantic segmentation of medical images have achieved good results in the literature. However, they tend to over-segment tumor regions and may not accurately segment the tumor edges and small regions. It has a significant impact on deploying such models for practical use. The thesis provides an overview of the current state-of-the-art techniques and methodologies in segmenting brain tumors. It explores the challenges associated with brain tumor segmentation and discusses three deep learning-based solutions that aim to bridge the gap in the segmentation of brain tumors. All the solutions have been studied on the BraTS2020 benchmark dataset, and the results are tabulated using the dice and Hausdorff95 metrics.

The initial work proposes a hybrid U-Net which adds residual, multi-resolution,

dual attention, and deep supervision blocks to the baseline U-Net model. The U-Net has several drawbacks, namely, merging of low-level features with high-level features in the skip connection, over-segmentation, lack of global context, and learning lots of similar feature maps resulting in wastage of memory. The goal of introducing residual blocks is to extract features efficiently to reduce the semantic gap between low-level features from the decoder and high-level features from skip connections. The multiple resolution blocks have been added to extract features and analyze tumors of varying scales. The dual attention mechanism has been incorporated to highlight tumor representations and reduce over-segmentation. Finally, deep supervision blocks have been added to utilize features from various decoder layers to obtain the target segmentation. The proposed model has been trained and evaluated on the BraTS2020 training and validation datasets. On the validation data, the proposed model has achieved a dice score of 0.691, 0.876, and 0.644 for enhancing tumor (ET), whole tumor (WT), and tumor core (TC), respectively, and a Hausdorff95 score of 34.43, 14.58, and 32.93, respectively. Compared to the baseline U-Net, the proposed model has outperformed WT and ET volumes in the dice score metric except for the CT volume. In the Hausdorff95 metric, the proposed method outperforms other U-Net models in the ET volume. The model achieved the performance with 505,559 parameters significantly smaller than the parameters required by the other models. At the end of the work, there is a significant need to improve the model in the Hausdorff metric.

The next work proposes a data-oriented approach to improve the edge detection capabilities of semantic segmentation models. The intermediate edges are extracted from the ground truth using a 3D derivative-like filter. The edges are reconstructed from the ground truth to obtain the tumor edges. The generated edge map is considered an additional ground truth. Utilizing both ground truths, several U-Net and its' variant architectures are trained with and without tumor edges as a target. Compared to the baseline U-Net and its variants, the models that learned edges along with the tumor regions performed well in the enhancing and core tumor regions in both training and validation datasets. The improved performance of edge-trained models trained on baseline models like U-Net and V-Net achieved performance similar to baseline state-of-the-art models like Swin U-Net and hybrid MR-U-Net. The edge-target trained models can generate edge maps useful for treatment planning. Additionally, for further explainability of the results, the activation map generated by the hybrid MR-U-Net has been studied. On the validation data, the hybrid edge MR-U-Net model has achieved a dice score of 0.682, 0.847, and 0.693 for enhancing tumor (ET), whole tumor (WT), and tumor core (TC), respectively, and a Hausdorff95 score of 40.70, 19.76, and 20.88, respectively. Compared to the baseline Hybrid MR-U-Net, the edge Hybrid MR-U-Net model has outperformed in TC in the Hausdorff95 distance and dice metric but has degraded performance in WT and ET volumes. One reason for the poor performance is the high penalty for false positives without a tumor region. Also, the dataset is multi-institutional, and there could be under-represented MR images resulting in poor results in the validation dataset.

To address under-represented data and false positives, the final work focuses on learning hard samples and giving more weight to false positive cases. Hard samples are images with a higher loss value at the end of an epoch. The loss for each sample is stored in a list and sorted in non-increasing order at the end of the epoch. In the

next epoch, δ percent of the whole dataset is selected for training instead of the whole dataset. The idea is to train high-loss samples multiple times to learn under-represented samples. In the proposed work, δ is set at 20%. In general, false positives are heavily penalized without a tumor region. So, two losses, namely, a weighted focal loss and a false positive (FP) loss, have been proposed to better learn false positive cases in the dataset. On the validation data, the Hybrid MR-U-Net Weighted Focal model has achieved a dice score of 0.674, 0.874, and 0.710 for enhancing tumor (ET), whole tumor (WT), and tumor core (TC), respectively, and a Hausdorff95 score of 41.62, 8.81, and 12.81, respectively. Similarly, the Hybrid MR-U-Net FP-Loss model has achieved a dice score of 0.674, 0.870, and 0.676 for enhancing tumor (ET), whole tumor (WT), and tumor core (TC), respectively, and a Hausdorff95 score of 38.66, 8.43, and 17.43, respectively. Compared to the popular U-Net models and several works in the literature, both the models have outperformed in WT and TC regions in the Hausdorff95 distance showing an improved detection of edges. However, the false positive rates in ET remain high, and there is a scope for improvement in the dice score. Also, the number of times a sample image is trained has been suggested as a methodology to identify outliers. It could act as an indicator to identify under-represented samples in the training data, potentially helping data collection. It could also help in identifying gaps in the current segmentation methods.

2 Objectives

The objectives of the research are as follows,

(a) To propose algorithms for automatically segmenting brain tumors using magnetic resonance imaging.

(b) To improve the U-Net model for better semantic segmentation performance for brain tumor segmentation.

(c) To extract tumor edges and study the performance of deep learning models provided with and without tumor edges for brain tumor segmentation.

(d) To develop a dynamic batch training algorithm to identify hard samples, study the performance of deep learning models and compare with the traditional training algorithm.

(e) To utilize dynamic batch training algorithm to identify potential outliers for brain tumor segmentation.

3 Existing Gaps Which Were Bridged

The existing gaps which were bridged in the work are as follows,

(a) **Improved Hybrid U-Net:** Implemented a U-Net that includes multi-resolution blocks, dual attention blocks, and deep supervision blocks. The model performs better with $\frac{1}{4}$th the number of parameters of the baseline U-Net.

(b) **Explored Data-oriented Approach:** Extract edges from the ground truth and learn edges and tumor regions as targets for brain tumor segmentation using MR images. By providing the model with explicit edge information as learning targets to enable more precise and accurate segmentation of brain tumors.

(c) **Confident Tumor Detection:** Learning of tumor edges along with the tumor segments resulted in higher activation of the tumor regions. The activation maps capture it.

(d) **Dynamic Batch Training to Learn Hard Samples:** The training of hard samples focused on false positives and high lossy samples showed an improved performance measured by the Hausdorff95 metric.

(e) **Dynamic Batch Training to Identify Outliers:** Identify outliers in the training data using the number of times a data point is trained.

4 Most Important Contributions

Figure 1 illustrates a sample 2D axial input MR Images and the corresponding ground truth. The input images (a-c), and the different MR modalities, namely, FLAIR, T1CE, and T2 images, are visualized. The ground truth consists of peritumoral edema (ED) marked in light grey given by an intensity value of 2, enhancing tumor (ET) represented as a white region with 4 as an intensity value, and the non-enhancing tumor (NET) and necrotic core region (NCR) as dark grey with an intensity value of 1. The models are evaluated as enhancing tumor (ET), tumor core (TC), and whole tumor (WT). The tumor core corresponds to NET/NCR and the enhancing tumor region. The whole tumor corresponds to all three tumor regions put together. The goal of the research is to learn the ED, NET/NCR, and ET regions. All the experiments have been carried out on the BraTS2020 dataset. The three most important contributions are given below,

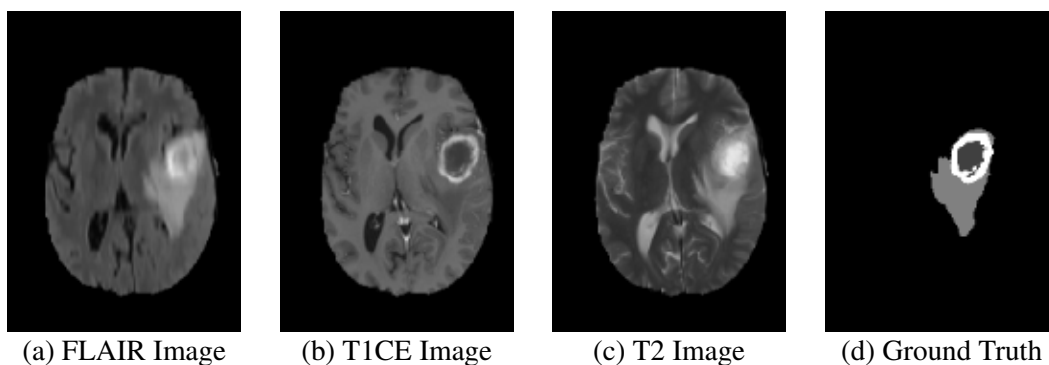| (a) FLAIR Image | (b) T1CE Image | (c) T2 Image | (d) Ground Truth |

Figure 1: Shows a sample 2D axial input MR Images (a-c) and the corresponding ground truth (d). In the ground truth, the white region corresponds to the enhancing tumor (ET), the dark grey region corresponds to the necrosis and non-enhancing tumor (NCR/NET), and the light grey area represents the edema region (ED) (Sahayam *et al.* (2022)).

4.1 Hybrid Multi-Resolution U-Net For Brain Tumor Segmentation

Figure 2 shows the architecture of the hybrid MR-U-Net. The drawbacks of the U-Net model include merging low-level features with high-level features at the skip connection, difficulty detecting small-scale tumors, and over-segmentation. The hybrid U-Net model utilizes multiple resolutions blocks (MRB) Ibtehaz and Rahman (2020), residual blocks (RB) He *et al.* (2016), dual attention blocks (DAB) Hariyani *et al.* (2020) and a deep supervision block (DSB) Lee *et al.* (2015) to enhance the segmentation ability of baseline U-Net model.
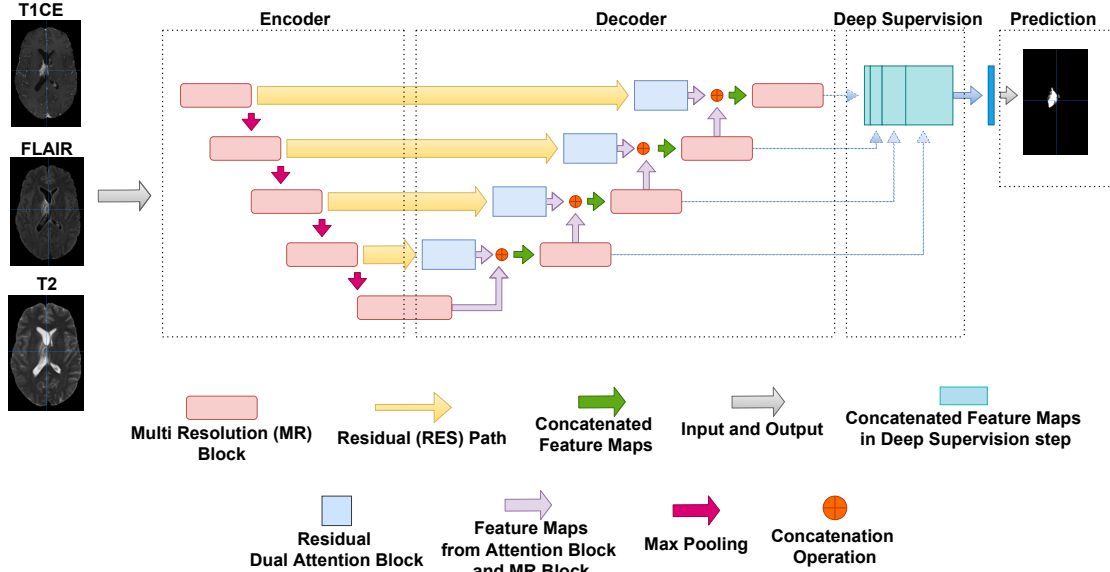
Figure 2: Proposed MultiResolution U-net with Residual Dual Attention and Deep Supervision Sahayam *et al.* (2022)

Table 1 compares different models on the validation dataset for brain tumor segmentation. The models are evaluated based on their Dice Score and Hausdorff95 metrics, with higher values for Dice Score and lower values for Hausdorff95 indicating better performance. The proposed Hybrid MR-U-Net achieves the overall best performance with a Dice Score of 0.691 for ET and 0.876 for WT. The Hybrid MR-U-Net also achieves the best Hausdorff95 score for ET, which is 34.43. The proposed model has 505,559 parameters, and it is $\frac{1}{4}$th number of parameters of baseline U-Net to achieve better performance.

4.2 Learning Tumor Edges and Segments for Brain Tumor Segmentation Using Deep Learning

The workflow of the proposed methodology is shown in Figure 3. The workflow can be broken down into Z-score normalization, edge extraction, one-hot representation, and training deep learning models. The work utilizes FLAIR, T2, and T1CE 3D MRI modalities and the ground truth for segmentation. The edges are extracted and reconstructed using a Laplacian-like 3D filter. Several U-Net models are trained with edge

Table 1: Comparison of the mean of different models on the validation dataset. The values in **bold** are the overall best.

Model	Dice Score ↑			Hausdorff95 ↓			Parameters
	ET	WT	TC	ET	WT	TC	
U-Net Ronneberger *et al.* (2015)	0.688	0.875	0.649	36.15	15.15	26.90	2,145,860
V-Net Milletari *et al.* (2016)	0.664	0.873	0.673	51.19	15.94	24.69	3,799,236
Attention U-Net Oktay *et al.* (2018)	0.646	0.863	0.679	56.37	16.81	27.09	2,167,703
U-Net 3 Plus Huang *et al.* (2020)	0.663	0.868	0.665	48.21	14.50	26.47	1,994,860
Swin U-Net Cao *et al.* (2023)	0.641	0.860	**0.682**	43.57	**11.13**	**16.95**	715,532
Hybrid MR-U-Net	**0.691**	**0.876**	0.644	**34.43**	14.58	32.93	505,559

and tumor regions as targets.
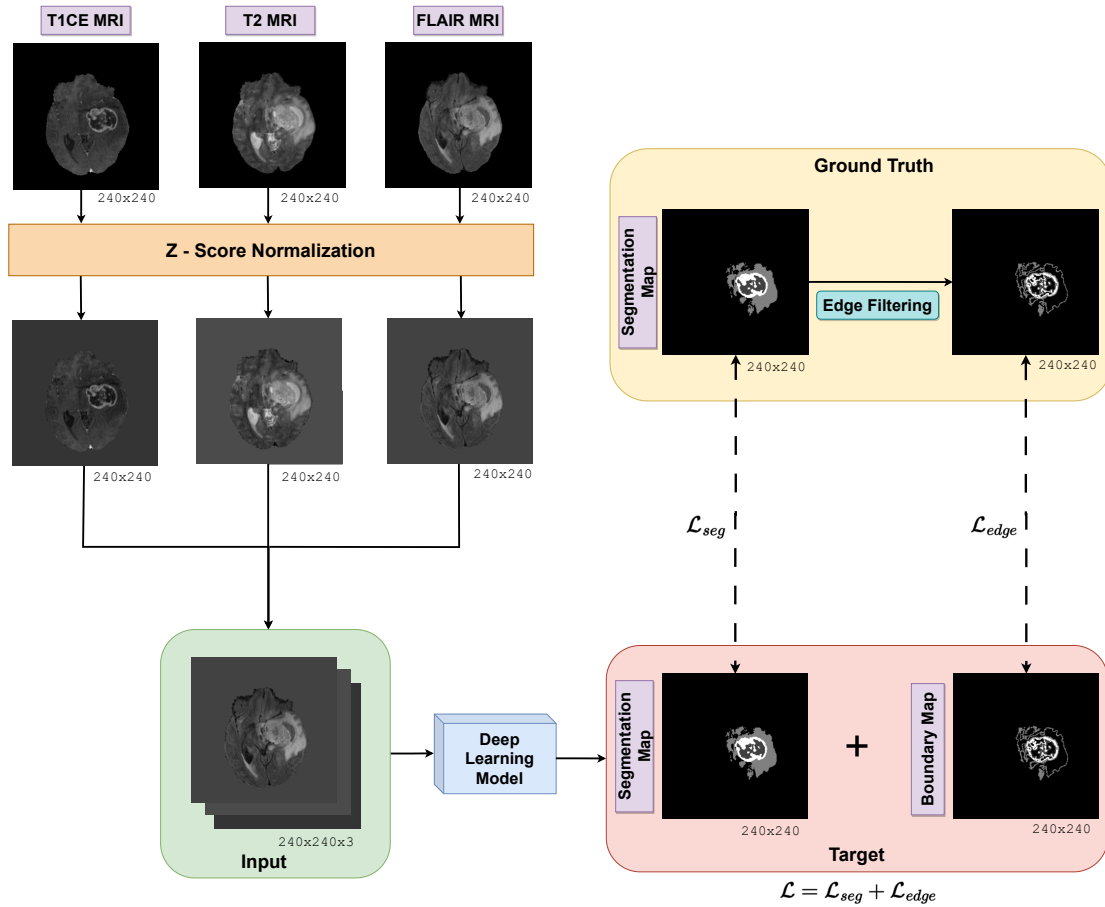
Figure 3: Shows the flow diagram of the proposed methodology to learn tumor edges and segments together.

Table 2 shows the mean validation results obtained for popular U-Net models and the proposed Hybrid MR-U-Net models. It can be observed the earlier models like U-Net, V-Net, and Attention U-Net show improvements when training with edges as targets in the enhancing and core tumor dice and Hausdorff95 score. All the models except the swin transformer performed well in dice and Hausdorff95 for tumor core trained with edges as targets along with the tumor regions.

Table 2: Comparison of the mean of different models on the validation dataset.[1]

Model	Dice Score ↑			Hausdorff95 ↓		
	ET	WT	TC	ET	WT	TC
U-Net Ronneberger *et al.* (2015)	0.688	**0.875**	0.649	**36.15**	15.15	26.90
U-Net Edge	**0.692** ▲ 0.58%	0.867 ▼ 0.91%	**0.671** ▲ 3.38%	38.69 ▲ 7.02%	**13.41** ▼ 11.48%	**22.83** ▼ 15.13%
V-Net Milletari *et al.* (2016)	0.664	**0.873**	0.673	51.19	15.94	24.69
V-Net Edge	**0.678** ▲ 2.11%	0.869 ▼ 0.46%	**0.675** ▲ 0.30%	**42.39** ▼ 17.19%	**13.66** ▼ 14.30%	**21.14** ▼ 14.37%
Attention U-Net Oktay *et al.* (2018)	0.646	**0.863**	**0.679**	56.37	**16.81**	27.09
Attention U-Net Edge	**0.660** ▲ 2.17%	0.820 ▼ 4.98%	0.656 ▼ 3.38%	**49.53** ▼ 12.13%	35.18 ▲ 109.28%	**17.36** ▼ 56.04%
U-Net 3 Plus Huang *et al.* (2020)	**0.663**	**0.868**	0.665	**48.21**	**14.50**	26.47
U-Net 3 Plus Edge	**0.663** ▼ 0.00%	0.862 ▼ 0.69%	**0.693** ▲ 4.21%	49.44 ▲ 2.55%	14.81 ▲ 2.13%	**21.13** ▼ 20.17%
Swin U-Net Cao *et al.* (2023)	**0.641**	**0.860**	**0.682**	**43.57**	**11.13**	**16.95**
Swin U-Net Edge	0.628 ▼ 2.02%	0.858 ▼ 0.23%	0.675 ▼ 1.02%	48.48 ▲ 11.26%	13.81 ▲ 24.07%	21.31 ▲ 25.72%
Hybrid MR-U-Net Sahayam *et al.* (2022)	**0.691**	**0.876**	0.644	**34.43**	**14.58**	32.93
Hybrid MR-U-Net Edge	0.681 ▼ 1.44%	0.847 ▼ 3.31%	**0.693** ▲ 7.60%	40.70 ▲ 18.21%	19.76 ▲ 35.52%	**20.88** ▼ 36.59%

The swin transformer did not improve in any metrics when learning along with target edges. Additionally, it can be noted that the swin transformer is the only model that uses linear artificial neural networks to learn features instead of convolutional layers. Also, the whole tumor degrades in performance in dice score and Hausdorff95 distance for all the models except U-Net and V-Net in the Hausdorff95 metric.

One possible explanation is that the recent models might have blocks that take care of learning edges, like the swin transformer blocks and self-attention blocks in the swin transformer and the dual attention blocks in hybrid MR-U-Net. It could also be that the edge target models might require more parameters than their earlier counterparts, like V-Net. Coincidentally, V-Net showed the most improvement in it has the highest number of parameters. Also, learning edges along with tumors shows that even earlier models like U-Net and V-Net could perform close to recent models like swin U-Net and hybrid MR-U-Net.

For better explainability of the models trained on tumor regions with and without tumor edges, the activation maps are generated by obtaining the predictions from the last layer of each of the models. For example, a sample 2D axial one-hot slice and activation map for hybrid MR-U-Net trained on tumor regions with and without edges is shown in Figure 4. The models trained along with the edges appear to have sharper edge activation than models trained without edges as targets. From Figure 4 (i) to (l), it can be observed that the model trained with tumor regions along with edges had higher activation (closer to 1.0) for tumor areas compared to the model trained with tumor regions only as in Figure 4 (e) to (h). Also, the latter has low activation, marked by purple color, for tumor regions that it shouldn't predict. For example, the activation of NCT/NET from Figure 4 (g) shouldn't show activation for edema and enhancing tumor regions. However, the model still has a low activation marked by the purple color. A similar activation map is observed for all other models trained for tumor regions with and without edges. The edges as targets might act as data-level attention to learn better tumor and edge segments. Additionally, the higher activation for models trained with tumor and edge regions could be the reason for improving the Hausdorff95 score.

---

[1]The values in **bold** represent the best performance of the existing and proposed models. The values in **blue** are the overall best. ▲ and ▼ % denotes improvement, and ▼ and ▲ % denotes deterioration of the proposed edge method to the normal method.
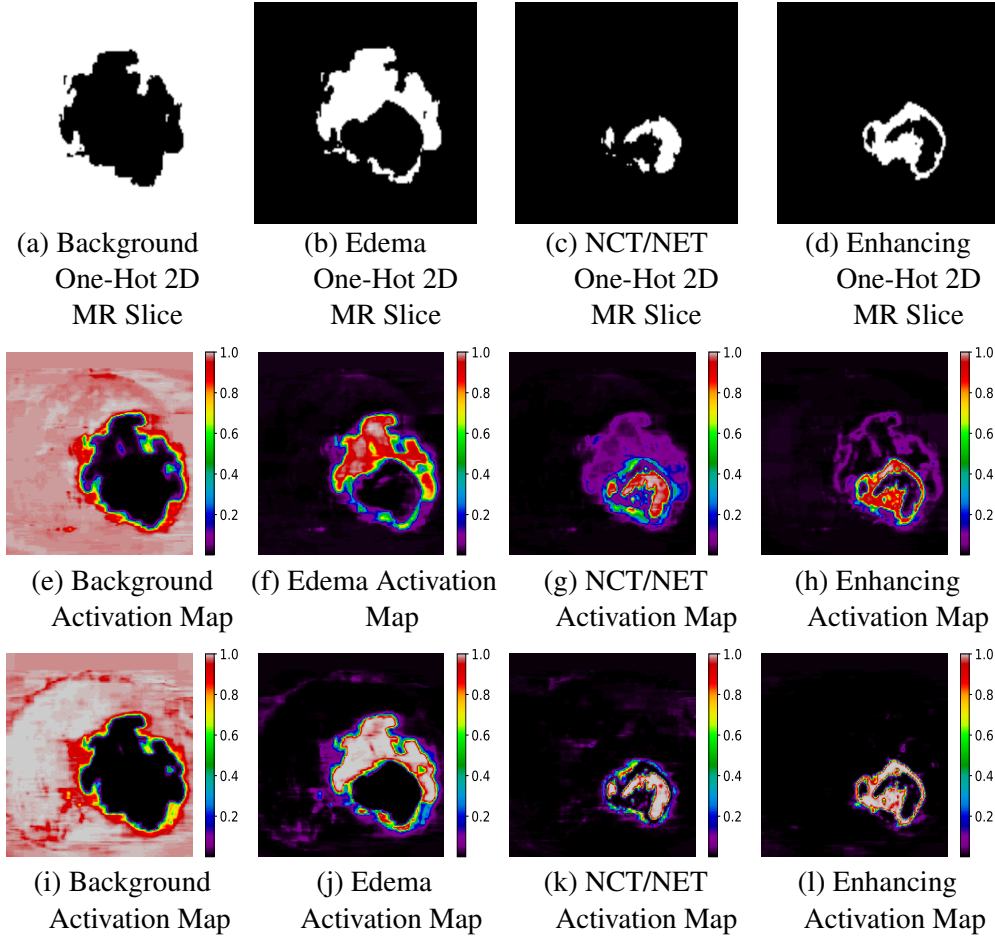
Figure 4: (a) to (d) Shows the ground truth one-hot vector for background, edema, NC-T/NET, and enhancing tumor. (e) to (h) shows the respective activation maps generated by the hybrid MR-U-Net trained with only tumor regions. (i) to (l) show the respective activation maps generated by the hybrid MR-U-Net trained with the tumor and edge regions.

4.3 Detection of Under-represented Samples Using Dynamic Batch Training for Brain Tumor Segmentation

Figure 5 depicts the workflow of dynamic batching for learning from hard samples. Hard samples are data points with a higher loss value at the end of an epoch. The loss for each sample is stored in a list and sorted in non-increasing order at the end of the epoch. In the next epoch, δ percent of the whole dataset is selected for training. It ends up training harder samples more than easier samples. δ is set at 20%. Two losses, namely, a weighted focal loss and a false positive (FP) loss have been proposed to better learn false positive cases in the dataset.

In the overall comparison Table 3, the models proposed in work have been compared with popular U-Net models and other semantic segmentation models proposed in the literature. It can be observed that the proposed hybrid model trained using a dynamic batch with proposed hybrid focal losses achieved 8.81 and 12.81 in the whole tumor and tumor core regions.
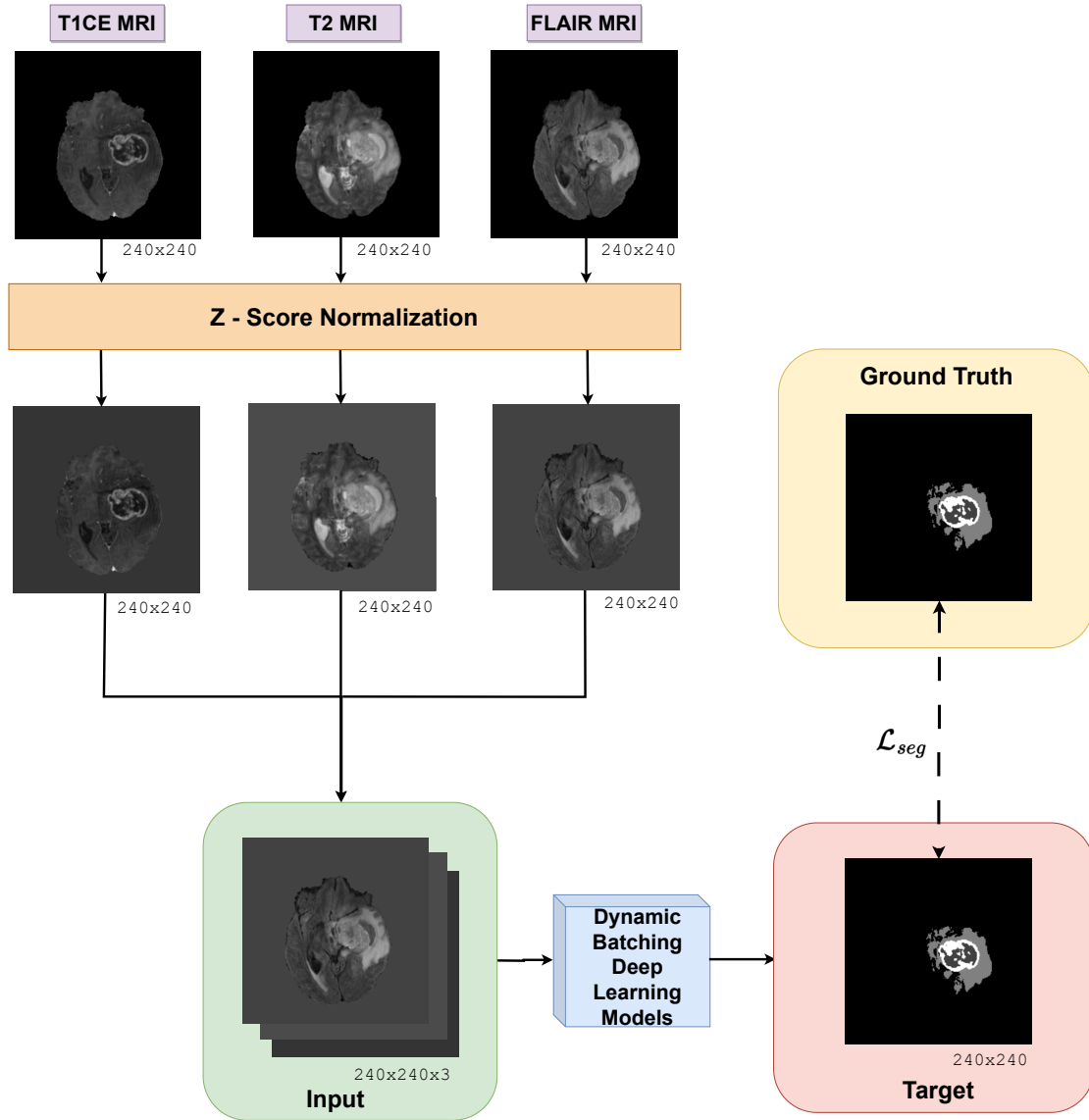
Figure 5: Proposed Workflow of Dynamic Batching for Learning Hard Samples

Figure 6 shows a scatterplot between patient ID and the number of times a patient's MR input image has been trained on the BraTS2020 dataset using traditional and dynamic batch training methods. The provided information can help identify underrepresented hard samples. It could give a direction to study data points in which the model fails to train. The regular hybrid MR-U-Net will see all the data points equivalent to the number of epochs. Since the number of epochs for the hybrid MR-U-Net is 50, each sample is trained 50 times denoted by the green scatter points. The traditional training method does not provide any additional information regarding the difficulty of the data points learned. From the other three models trained using dynamic batch training, it can be observed that the model sees a sample at least 25 times and up to a maximum of 150 times. The patients that get trained 25 times are likely easier samples to segment than those trained to around 150 times. So, the scatter plot given in Figure 6 has an excellent potential to identify, find, and take corrective measures to handle outliers.

Table 3: Shows the comparison study of various models in the literature with the BraTS2020 validation dataset.

| Model	Dice Score ↑			Hausdorff95 ↓		
	ET	WT	TC	ET	WT	TC
U-Net Ronneberger *et al.* (2015)	0.688	0.875	0.649	36.15	15.15	26.90
V-Net Milletari *et al.* (2016)	0.664	0.873	0.673	51.19	15.94	24.69
3D U-Net Çiçek *et al.* (2016)	**0.729**	0.851	**0.779**	31.69	12.50	18.75
Attention U-Net Oktay *et al.* (2018)	0.646	0.863	0.679	56.37	16.81	27.09
U-Net 3 Plus Huang *et al.* (2020)	0.663	0.868	0.665	48.21	14.50	26.47
MTAU Awasthi *et al.* (2021)	0.570	0.730	0.610	38.87	20.81	24.22
Probabilistic U-Net Savadikar *et al.* (2021)	0.689	0.819	0.717	36.89	41.52	26.28
DRU-Net Colman *et al.* (2021)	0.670	**0.880**	0.670	47.62	12.11	15.74
MVP U-NetZhao *et al.* (2021)	0.670	0.860	0.620	47.33	12.58	50.14
3D U-Net Ballestar and Vilaplana (2021)	0.720	0.830	0.770	37.42	12.34	13.11
Swin U-Net Cao *et al.* (2023)	0.641	0.860	0.682	43.57	11.13	16.95
Hybrid MR-U-Net Sahayam *et al.* (2022)	0.691	0.876	0.644	**34.43**	14.58	32.93
Hybrid MR-U-Net Edge	0.681	0.847	0.693	40.70	19.76	20.88
Hybrid MR U-Net MFP Focal Loss	0.674	0.870	0.676	38.66	**8.43**	17.43
Hybrid MR U-Net Hybrid Focal Loss	0.674	0.874	0.710	41.62	8.81	**12.81**
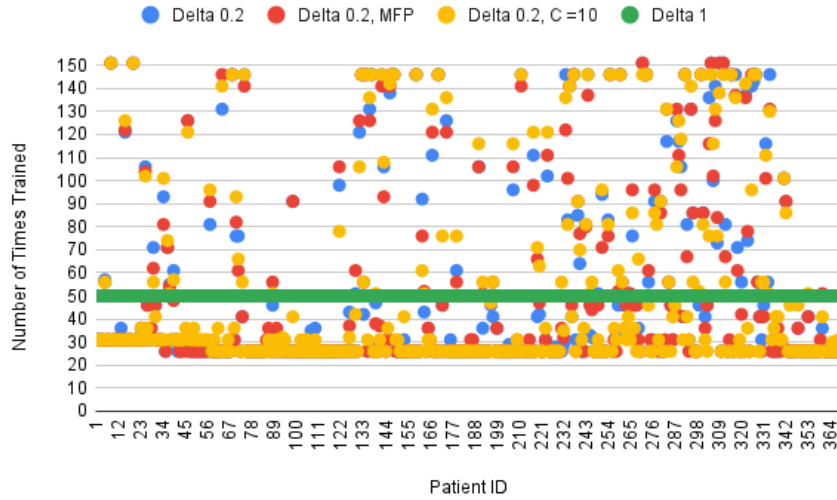
Figure 6: Shows a scatterplot with patient ID vs. the number of times a patient's MR image has been trained.

5 Conclusions

In conclusion, the works reviewed in the thesis have proposed modifications to the popular U-Net model for more accurate segmentation of brain tumors in MR images. The first work proposed a model incorporating residual, multi-resolution, dual attention, and deep supervision blocks to improve feature extraction and segmentation accuracy. The proposed model achieved better results than the baseline U-Net regarding the Hausdorff95 and dice metrics for the whole tumor and enhancing tumor regions. However, there is room for improvement in the segmentation of tumor core regions and general improvement requirements in the Hausdorff95 score.

The second work addressed the issue of accurately segmenting tumor edges, which are crucial for accurate diagnosis, surgical precision, and treatment planning. The proposed edge-target training strategy uses Laplacian-like 3D filter to extract tumor edges from the ground truth and trains the model to segment both the tumor regions and edges. The proposed model achieved competitive results compared to popular U-Net models and published literature in the BraTS2020 benchmark dataset. The edge maps generated by the model can be helpful for treatment planning. An activation map has been generated to explain better the results obtained by the models with and without edge information as target images.

The third work proposes a dynamic batch training method to improve the performance of brain tumor segmentation has been proposed. The proposed method employs two loss functions, namely, hybrid focal loss and means false positive focal loss, to address the problem of class imbalance and false positives. The experimental results showed that the proposed method outperformed several models in the literature regarding Hausdorff distance and dice score metrics. A scatter plot has been created between patient ID and the number of patients to demonstrate that the proposed dynamic batch training can be used to identify under-represented and hard samples. The proposed dynamic batch training method can be a valuable tool in clinical settings, and identifying hard samples can help researchers fine-tune the model, collect more data, and ultimately lead to improved overall performance on the task at hand.

Overall, the segmentation of tumors from MR images is a challenging task. There is still potential to develop end-to-end deep learning models for the segmentation task before applications can be deployed for real-world segmentation problems. Until then, deep learning segmentation models with post-processing and additional data for training can yield good performance. Future works could still focus on handling the false positives in the absence of a tumor region and improving dice scores in the validation dataset.

6 Organization of the Thesis

The proposed outline of the thesis is as follows:

(a) Chapter 1: Introduction

(b) Chapter 2: Related Work

(c) Chapter 3: Preliminaries: Popular U-Net Models

(d) Chapter 4: Model-oriented Approach: Hybrid Multi-Resolution U-Net for Brain Tumor Segmentation

(e) Chapter 5: Data-oriented Approach: Learning Tumor Edges and Segments for Brain Tumor Segmentation Using Deep Learning

(f) Chapter 6: Detection of Under-represented Samples Using Dynamic Batch Training for Brain Tumor Segmentation

(g) Chapter 7: Conclusions and Future Work

7 List of Publications
I. REFEREED JOURNALS BASED ON THE THESIS

1. Subin Sahayam, Rahul Nenavath, Umarani Jayaraman and Surya Prakash, Brain tumor segmentation using a hybrid multi-resolution U-Net with residual dual attention and deep supervision on MR images, *Biomedical Signal Processing*, 78, p.103939, (2022).

2. Subin Sahayam, and Umarani Jayaraman, Learning Tumor Edges and Segments for Brain Tumor Segmentation Using Deep Learning, *Computers in Biology and Medicine* (**Submitted**).

3. Subin Sahayam, John Zakkam and Umarani Jayaraman, Detection of Under-Represented Samples Using Dynamic Batch Training for Brain Tumor Segmentation, *Medical Image Analysis* (**Under Preparation**).

II. REFEREED JOURNALS (Others)

1. Anil Kumar Adepu, Subin Sahayam, Umarani Jayaraman and Rashmika Arramraju, Melanoma classification from dermatoscopy images using knowledge distillation for highly imbalanced data, *Computers in Biology and Medicine*, 154, p.106571, (2023).

III. PRESENTATIONS/PUBLICATIONS IN CONFERENCES (Others)

1. Subin Sahayam, Nanda H Krishna and Umarani Jayaraman, Brain Tumour Segmentation on MRI Images by Voxel Classification Using Neural Networks and Patient Survival Prediction, *In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes2019, Held in Conjunction with MICCAI2019*, pp. 284-294 Springer International Publishing, (2019).

2. Subin Sahayam, Umarani Jayaraman and Bhaskar Teja Multi-class Glioma Classification from MRI Images Using 3D Convolutional Neural Networks, *In Computer Vision and Image Processing: 5th International Conference, CVIP 2020*, Part I 5 pp. 327-337 Cham: Springer International Publishing, (2020).

3. Subin Sahayam, Abirami and Umarani Jayaraman, A Novel Modified U-shaped 3-D Capsule Network (MUDCap3) for Stroke Lesion Segmentation from Brain MRI, *In 2020 IEEE 4th Conference on Information & Communication Technology (CICT)*, pp. 1-6 IEEE, (2020).

4. Subin Sahayam, Silambarasan and Umarani Jayaraman, Detection of Cataract from Fundus Images Using Deep Transfer Learning, *In Computer Vision and Image Processing: 6th International Conference, CVIP 2021*, Part I pp. 175-186 Cham: Springer International Publishing, (2021).

5. Subin Sahayam, Tutturu Lakshmi Manasa and Umarani Jayaraman, Ensemble Models for Multi-class Classification of Diabetic Retinopathy, *In Pattern Recognition and Machine Intelligence: Ninth International Conference, PReMI 2021, Kolkata*, Springer International Publishing, (2021).

6. Anil Kumar Adepu, Subin Sahayam, Rashmika Arramraju and Umarani Jayaraman A Study on an Ensemble Model for Automatic Classification of Melanoma from Dermoscopy Images *In Computer Vision and Image Processing: 7th International Conference, CVIP 2022*, Cham: Springer International Publishing, (2022).

References

1. **Awasthi, N.**, **R. Pardasani**, and **S. Gupta** (2021). Multi-threshold attention u-net (mtau) based model for multimodal brain tumor segmentation in mri scans. *In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6*. Springer.

2. **Ballestar, L. M.** and **V. Vilaplana** (2021). Mri brain tumor segmentation and uncertainty estimation using 3d-unet architectures. *In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part I 6*. Springer.

3. **Cao, H.**, **Y. Wang**, **J. Chen**, **D. Jiang**, **X. Zhang**, **Q. Tian**, and **M. Wang** (2023). Swin-unet: Unet-like pure transformer for medical image segmentation. *In Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part III*. Springer, `doi:https://doi.org/10.1007/978-3-031-25066-8_9`.

4. **Çiçek, Ö.**, **A. Abdulkadir**, **S. S. Lienkamp**, **T. Brox**, and **O. Ronneberger** (2016). 3d u-net: learning dense volumetric segmentation from sparse annotation. *In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*. Springer.

5. **Colman, J.**, **L. Zhang**, **W. Duan**, and **X. Ye** (2021). Dr-unet104 for multimodal mri brain tumor segmentation. *In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6*. Springer.

6. **Hariyani, Y. S.**, **H. Eom**, and **C. Park** (2020). Da-capnet: Dual attention deep learning based on u-net for nailfold capillary segmentation. *IEEE Access*, **8**, 10543–10553.

7. **He, K.**, **X. Zhang**, **S. Ren**, and **J. Sun** (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition.*

8. **Huang, H.**, **L. Lin**, **R. Tong**, **H. Hu**, **Q. Zhang**, **Y. Iwamoto**, **X. Han**, **Y.-W. Chen**, and **J. Wu** (2020). Unet 3+: A full-scale connected unet for medical image segmentation. *In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, `doi:https://doi.org/10.48550/arXiv.2004.08790`.

9. **Ibtehaz, N.** and **M. S. Rahman** (2020). Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Networks*, **121**, 74–87.

10. **Lee, C.-Y.**, **S. Xie**, **P. Gallagher**, **Z. Zhang**, and **Z. Tu** (2015). Deeply-supervised nets. *In Artificial intelligence and statistics*. PMLR.

11. **Milletari, F.**, **N. Navab**, and **S.-A. Ahmadi** (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. *In 2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, `doi:10.1109/3DV.2016.79`.

12. **Oktay, O.**, **J. Schlemper**, **L. L. Folgoc**, **M. Lee**, **M. Heinrich**, **K. Misawa**, **K. Mori**, **S. McDonagh**, **N. Y. Hammerla**, **B. Kainz**, *et al.* (2018). Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, `doi:https://doi.org/10.48550/arXiv.1804.03999`.

13. **Ronneberger, O.**, **P. Fischer**, and **T. Brox** (2015). U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention*. Springer.

14. **Sahayam, S.**, **R. Nenavath**, **U. Jayaraman**, and **S. Prakash** (2022). Brain tumor segmentation using a hybrid multi resolution u-net with residual dual attention and deep supervision on mr images. *Biomedical Signal Processing and Control*, **78**, 103939, `doi:https://doi.org/10.1016/j.bspc.2022.103939`.

15. **Savadikar, C.**, **R. Kulhalli**, and **B. Garware** (2021). Brain tumour segmentation using probabilistic u-net. *In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6*. Springer.

16. **Zhao, C.**, **Z. Zhao**, **Q. Zeng**, and **Y. Feng** (2021). Mvp u-net: multi-view pointwise u-net for brain tumor segmentation. *In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6*. Springer.