Literature Survey on Semantic Segmentation of Brain Tumor

Melih Berk Yılmaz, *Member, Bilkent University,* Fuat Arslan, *Member, Bilkent University,*

***Abstract*—This literature review explores the semantic segmentation of brain tumors, discussing the different imaging methods used, common evaluation metrics, how data is handled in research, and the various designs used to accurately segment tumors. It explores the significance of magnetic resonance imaging (MRI) modalities in tumor detection. The review carefully examines methods involving input data types and architectural designs, such as Convolutional Models with Graphical Models, Encoder-Decoder, Generative Adversarial, and Multi-Task Learning approaches. Throughout, it highlights the challenges and nuances within the domain of brain tumor segmentation.**

***Index Terms*—Image Segmentation, Brain tumor segmentation, MRI, Semantic Segmentation, Deep Convolutional Neural Networks.**

I. Introduction

BRAIN tumor segmentation from medical images is a critical task in clinical diagnostics, aiding in treatment planning, patient monitoring, and outcome prediction. Magnetic Resonance Imaging (MRI) stands as a pivotal modality for brain tumor examination, offering multiple imaging sequences such as FLAIR, T2, T1, and T1c, each providing distinct but complementary information about tumor location, type, and characteristics. The utilization of advanced computational techniques for semantic segmentation has evolved significantly, driven by the availability of large datasets like the BraTS Challenge Dataset and the need for accurate, efficient, and reproducible tumor delineation methodologies.

This literature review aims to comprehensively explore the methodologies and techniques employed in brain tumor segmentation from MRI data. It will discuss the variety of imaging modalities, evaluation metrics, input data modalities, and diverse architectural designs utilized in state-of-the-art semantic segmentation models for brain tumors.

II. Brain Tumor Viewing Modalities and Data

Various methods are employed to acquire medical images for diagnostic purposes, including Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), and Computed Tomography (CT). MRI, renowned for its efficacy in examining soft tissues and the nervous system, stands out as the most impactful technique. In tumor detection, widely used MRI modalities encompass Fluid-attenuated inversion recovery (FLAIR), T2-weighted images (T2), T1-weighted images (T1), and T1-weighted images with contrast enhancement (T1c). [1], [2]

FLAIR modality aids in detecting lesions, T2 modality facilitates visualizing edema and distinguishing abnormal tissues, T1 modality helps determine the presence and location of the tumor, and T1c modality assists in detecting tumor presence. Consequently, all these modalities provide complementary information for detecting the tumor, identifying its type, and determining its location [1], [2]. The visual distinctions between these modalities are depicted in Figure 1.

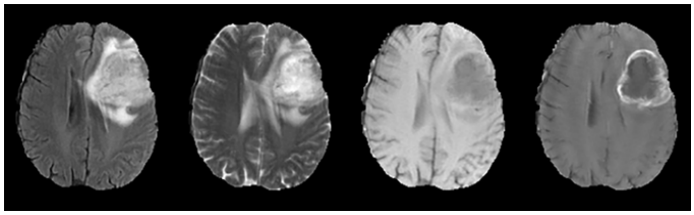

Fig. 1: Modalities of MRI: FLAIR, T2, T1, T1c (from left to right) [2]

The Brain Tumor Segmentation (BraTS) Challenge Dataset, commonly used and publicly available, consists of multi-modality images including T1, T1c, T2, and FLAIR. This dataset offers segmentation labels for three classes: enhancing tumor (ET), peritumoral edema (ED), and necrotic and non-enhancing tumor core (NCR). The provided data undergo preprocessing, wherein they are co-registered to the same anatomical template, interpolated to a uniform resolution ($1\,\mathrm{mm}^3$), and skull-stripped. Figure 2 demonstrates 2D slice image belonging to a patient with its corresponding mask. [3]

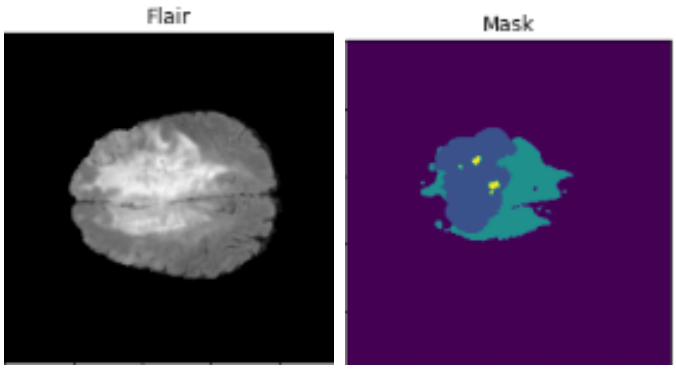

Fig. 2: Sample 2D Slice with FLAIR Modality and Corresponding Mask [2]

III. Evaluation Metrics

Evaluation metrics are a crucial aspect for both the training and testing phases in machine learning. For segmentation tasks, like the ones considered in this survey, several metrics commonly cited in the literature are:

1) Dice Loss: The Dice Loss is a prevalent approach for segmentation tasks due to its suitability for measuring the overlap between the prediction result and the ground truth. The Dice loss, defined as this overlap rate, is formulated as follows:

$$\text{LOSS}_{\text{Dice}} = 1 - \frac{2\sum_{i=1}^{n} p_i y_i}{\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} y_i^2} \tag{1}$$

In this formula, y_i represnets label pixels and p_i are prediction pixels. $1 - \text{LOSS}_{\text{Dice}}$ is the dice coefficient. This metric is particularly useful in addressing class imbalance, which is typical in pixel counts for foreground and background segmentation [4].Since brain tumor segmentation often deals with imbalanced data in MRI images, Dice Loss is highly suitable. Its effectiveness for brain image segmentation has been the focus of its originating paper and subsequent studies [5]. The Dice coefficient is also commonly used as a measure of model performance on test data, serving as both a loss function and an evaluation metric in brain tumor segmentation literature [4], [5], [6].

2) Cross Entropy: Its formulation given below:

$$\text{LOSS}_{\text{CE}} = -\sum_{i=1}^{n} y_i \log(p_i) \tag{2}$$

This metric quantifies the uncertainty of a prediction. It penalizes incorrect predictions logarithmically. It applies pixel wise. Binary Cross Entropy, a variant used for binary masks (0 and 1), is not as prevalent as Dice Loss in contemporary approaches but is sometimes combined with other loss functions. One drawback of Cross Entropy is its performance with imbalanced datasets, which is a notable concern in brain tumor segmentation [4], [7].

3) Jaccard Similarity: Also known as the Intersection over Union (IoU), Jaccard Similarity measures the intersection over the union of two datasets. It is another common metric for evaluation and is well-suited for brain tumor segmentation due to its interpretability and alignment with the task's objectives [8], [9]. The formula for Jaccard Similarity:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{3}$$

There are other metrics which are not very unique to segmentation like Sensitivity, Precision etc. which can be used as evaluation of the model.

IV. Methodology

The section specifically focuses on the anticipated input data type for the models and outlines how cutting-edge models in this domain manage MRI data preparation. It then delves into the diverse architectural designs specifically crafted for the semantic segmentation of brain tumors.

A. Input Data Modalities

Before delving into the specific architectural designs for brain tumor segmentation, it's crucial to analyze the input data modalities. Many datasets provide hundreds of consecutive 2D slices of the brain, each with four modalities for every patient. To articulate this mathematically, suppose a single patient's data consists of 200 2D slices across four modalities, resulting in a total of 800 images along with their corresponding semantic segmentation masks.

Some methodologies employ 2D image-based approaches, which involve developing segmentation models that process each entire image independently, potentially leading to the loss of spatial information specific to a patient. Within this framework, two approaches emerge: one solely utilizes FLAIR or T1c images to localize and classify tumors, while an alternative approach simultaneously considers all four modalities. The latter involves either stacking these modalities in the channel dimension or separately passing them through the model and merging their outputs to leverage the complementary information from each modality. [7]

Another strategy involves a 3D approach. Here, the 2D slices pertaining to a single patient are stacked along the third dimension, thereby creating a volumetric image for each patient. Although this method significantly reduces the volume of data by condensing hundreds of images into one sample, it retains crucial spatial information for each patient. Within this approach, certain methods either utilize only FLAIR or T1c images or concatenate all four modalities along the channel dimension. Subsequently, the volumetric data undergoes processing through a cascade of 3D convolutional blocks. [10]

B. Architecture Design

1) Convolutional Models with Graphical Models Approach: Various methods have been explored to enhance MRI brain tumor segmentation, leveraging Convolutional Neural Networks (CNNs) alongside Graphical Models. One approach involves combining Fully Convolutional Networks (FCNs) with Conditional Random Fields (CRFs) to refine pixel-wise predictions and ensure spatial consistency in segmentation results.

One study presents the incorporation of CRFs, such as CRF-RNN, within CNN architectures aids in refining pixel-wise predictions obtained from CNNs by considering spatial dependencies among pixels. These CRF-based approaches enable the preservation of spatial coherence and fine-tuning of segmentation boundaries, ensuring smoother and more accurate segmentations. By combining the power of CNNs in feature learning with the spatial consistency offered by CRFs, these hybrid models significantly improve brain tumor segmentation results, leveraging both local and contextual information in medical imaging datasets. The model architecture can be seen in Figure 3 (top). [11]

Additionally, another study introduces a comprehensive solution for MRI brain tumor segmentation, employing three concurrent Fully Convolutional Networks (FCNs) and a fully connected Conditional Random Field (CRF) as key components. The resulting FCN predictions are fused using linear regression to refine the segmentation. Following this, a fully connected CRF serves as a post-processing step, enhancing the FCN segmentation. This CRF step, leveraging state-of-the-art CRF-RNN techniques, refines boundaries by considering both pixel-level intensities and pairwise interactions. Notably, the

CRF model's use of Gaussian kernels and appearance-based potentials significantly contributes to improving segmentation accuracy. The study's contributions primarily lie in the integration of FCNs and CRFs, replacing patch-based CNNs with FCNs for pixel-wise segmentation, and enhancing MRI brain tumor segmentation accuracy through CRF-based spatial refinement. The model architecture can be seen in Figure 3 (bottom). [12]

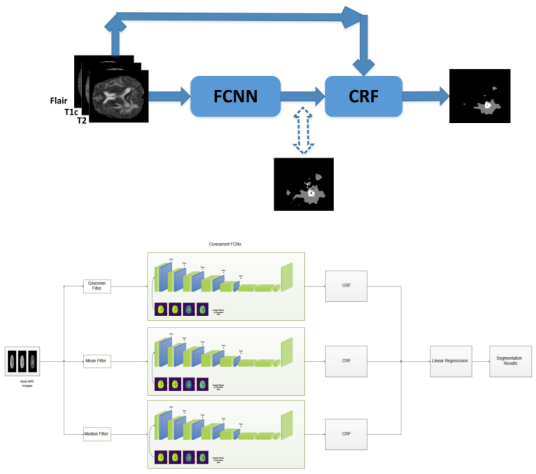

Fig. 3: CNN+CRF Model Architectures [11], [12]

2) Encoder Decoder Approach: Encoder-Decoder Networks have become a fundamental architecture for image segmentation tasks in modern deep learning models. In this architecture, the network consists of an encoder section that gradually reduces the spatial dimensions of the input image while increasing the channel dimensions through a series of convolutional blocks. This process allows for the extraction of hierarchical and abstract features from the input image.

The encoder's objective is to transform the input image into a dense, lower-dimensional representation, often referred to as a latent representation or feature map. This dense representation aims to capture essential information while filtering out irrelevant or noisy features from the image, thereby enhancing the network's ability to understand and process the image content effectively.

Once the encoder produces a compact and rich feature representation, the decoder section, typically composed of transpose convolutional blocks or upsampling layers, takes this latent representation and gradually upscales it back to the original input image size. This upsampling process reconstructs the spatial information lost during the encoding phase, allowing the network to generate a segmented output that aligns with the original input dimensions. Encoder-Decoder Networks aim to preserve intricate details and context while segmenting images.

The U-Net architecture represents a groundbreaking advancement in brain tumor segmentation, revolutionizing biomedical image segmentation methods. It is specialized convolutional for the complex task of segmenting brain tumors.It brings benefits of the encoder-decoder approach to biomedical segmentation domain. U-Net's unique structure, featuring a contracting path for context capture and a symmetric expanding path for precise localization, is exceptionally effective in medical imaging, ensuring accurate tumor boundary delineation. The encoder-decoder framework of U-Net manages brain tumor segmentation by compressing images for global context understanding and then reconstructing them at higher resolutions for detailed segmentation. This approach is particularly advantageous when training data is scarce, a frequent issue in medical imaging. U-Net's proficiency in learning from limited data sets, achieving high accuracy in segmentation, has established a new benchmark in the field and paved the way for future innovations in medical image segmentation [13].

The literature has seen significant expansions of the U-Net architecture. One notable example is the Attention U-Net, which introduces attention gates to the residual connections between the downscaling and corresponding upscaling blocks [14]. Another advanced model is the Swin UNETR, which overcomes the traditional U-Net's limitations in modeling long-range information. This limitation is primarily due to the restricted kernel size of its convolution layers. Swin UNETR incorporates Swin Transformers as the encoder within the U-Net framework, enabling the efficient capture of long-range dependencies through interactions between token embeddings. This integration significantly enhances both local and global contextual understanding [15].

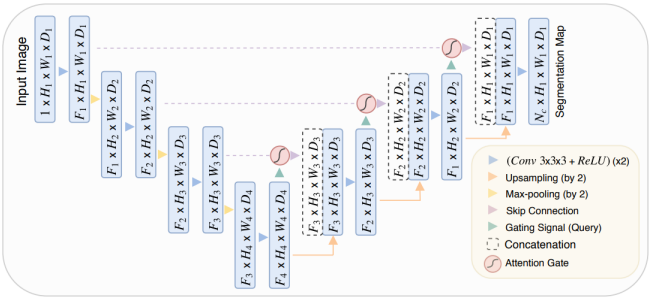

Fig. 4: Model Architecture of Attention U-Net [14]

Additionally, the U2-Net has improved the traditional U-Net architecture by introducing a two-level nested structure. It features advanced ReSidual U-block (RSU) modules, which allow for more efficient, high-resolution feature extraction in segmentation. Remarkably, this enhancement does not substantially increase computational load [16].

Processing MRI data slice by slice leads to the loss of patient-specific spatial information. In place of this approach, Fausto et al. proposed the V-Net architecture, which operates on volumetric data (as explained in section 4.A) and processes the data using a cascade of 3D convolutions. It employs an encoder-decoder architecture divided into left and right portions. The left side operates at different resolutions and uses residual functions in its convolutional stages.

It downsamples the input using convolutional kernels and doubling the number of feature maps at each stage. By employing feature maps at different resolutions, the network can capture multi-scale information present in the input data. This helps in handling objects or structures of varying sizes within

the image. Using different resolutions allows the network to process information at multiple scales simultaneously. It enables the network to extract both local and global context. [17]

Instead of traditional pooling operations, convolutional operations are utilized for downsampling, aiding in feature extraction and preserving memory during training. The right portion expands lower resolution feature maps and generates a two-channel volumetric segmentation output through soft-max voxelwise operations. [17]

Horizontal connections between the left and right portions facilitate the transfer of fine-grained details, enhancing contour predictions and accelerating model convergence. Overall, this architecture combines residual learning, convolutional operations for downsampling, and interconnections between encoder and decoder segments to achieve accurate volumetric segmentation in medical images. Figure 5 shows the architectural design of the V-Net. [17]

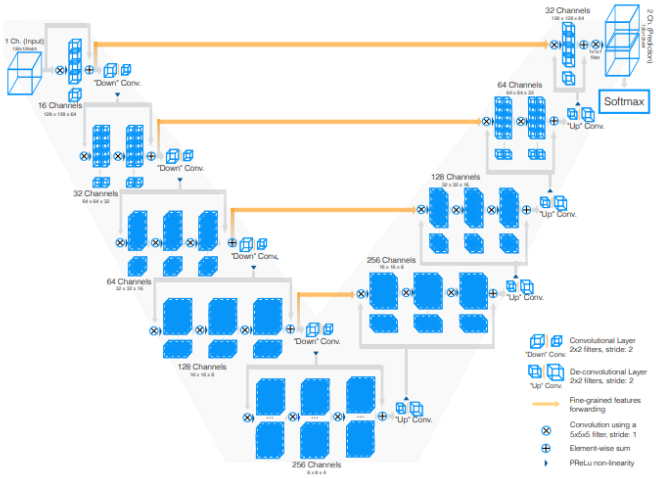

Fig. 5: Model Architecture of Attention V-Net [17]

3) Generative Adversarial Approach: The Generative Adversarial Network (GAN) [18] is a modern approach in image generation. This method has been diversified into various forms, primarily focusing on generating images from a known distribution through adversarial training. Specifically in image segmentation, the task can be reframed as generating an image that represents the segmentation, thus adapting it to an image generation task. However, unlike typical GAN applications, this requires generating an image from another image instead of a known distribution.

One solution involves encoding images into a distribution and subsequently generating images from this encoded form [19]. In this process, the original image intended for segmentation passes through an encoder architecture to achieve an efficient representation. Contrary to traditional GANs that generate from $p(z)$, this method initiates generation from a more complex distribution $p(w)$. The encoder establishes a mapping from the z-space to the w-space using the training data distribution. Following this, the adversarial loss functions as usual during the generation phase.

Another approach involves deriving a feature map as the starting point for generation, aiming to identify the $p(z|x^*)$ conditional distribution [20]. A feature extractor (typically a CNN) creates an image representation to condition the generator. As a result, the generator receives information about the desired output before generation, unlike scenarios where it starts with complete noise. This added information, combined with the noise's stochasticity, aids in generating the segmentation map. The process is then completed with traditional adversarial training.

Moreover, various GAN variants for segmentation incorporate not only adversarial training but also additional losses such as cross-entropy and perceptual loss, as mentioned in [21]. While adversarial loss ensures the generated image aligns with the distribution of segmented images, it doesn't necessarily enforce precise learning of the specific segmentation. Therefore, most variants introduce supplementary loss functions to enhance accuracy, as elaborated in [20], [22], and [21]. This GAN-based segmentation, particularly in the context of brain tumor segmentation, has shown promising results and has established its significance in the literature of brain tumor segmentation, also highlighted in these studies [20], [22], [21].

4) Multi-Task Approach: Multi-task Learning (MTL) [23] in machine learning involves jointly learning multiple related tasks, with the primary goal of enhancing the performance of each task by leveraging the learning processes of the others. This approach is particularly advantageous in addressing data scarcity issues by utilizing data from each task to train a shared model, thereby maximizing the use of existing data [24]. Additionally, MTL facilitates more effective learning of representations. For example, while focusing solely on brain tumor segmentation, a model might struggle to identify certain significant features which could be more easily learned through another task. MTL compels the model to assimilate essential information for every task, allowing individual tasks to benefit from these complexly interrelated features [25].

Another significant advantage of MTL is its role in regularization. Training concurrently on different tasks prevents the feature extractor from overfitting to the training data of a single task. The model must develop more generalizable features to accommodate the requirements of multiple tasks, resulting in a more robust final model [24], [25].

Reflecting on these benefits, the brain tumor segmentation field has also adopted MTL, integrating auxiliary tasks into the learning process. A straightforward approach involves dividing the segmentation task into sub-tasks, such as differentiating non-tumor from tumor brain areas, identifying the tumor core, and enhancing tumor detection [26]. This subdivision results in an MTL framework. The architecture proposed in [26], which can be seen in Figure 6, defines three distinct losses for each sub-task. These are trained jointly to achieve a more effective segmentation model.

One of the most intriguing and commonly employed auxiliary tasks in Multi-task Learning (MTL) is the reconstruction

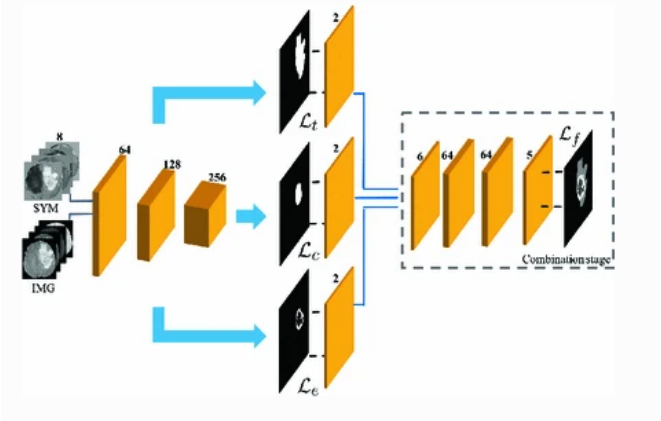

Fig. 6: Model Architecture at [26]

task, which involves regenerating the input data using representations produced by the model. A prominent example of this approach is detailed in [27], where the authors' architecture secured the first place in the BraTS 2018 competition. This architecture employs an autoencoder to reconstruct the input image from the model's intermediate representation, functioning as a form of regularization and harnessing the benefits associated with MTL.

The architecture first encodes the image into a latent representation, which is then used to generate a prior distribution. By sampling from this distribution, the model reconstructs the input image. The auxiliary task in this setup is framed as a variational autoencoder, while the primary task remains focused on segmentation [27]. This concept achieved remarkable success at the time of its introduction. The overall architecture of this approach is depicted in Figure 7.

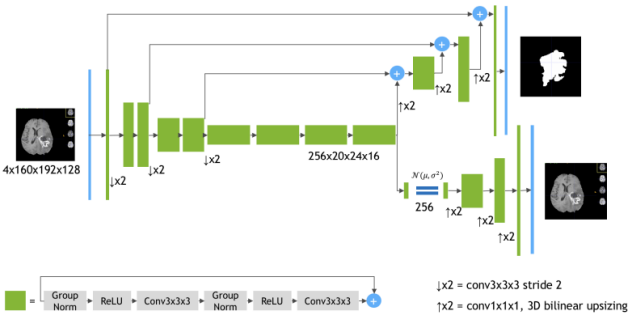

Fig. 7: Model Architecture at [27]

The concept of reconstruction using a variational autoencoder has remained a popular technique in the literature, as evidenced by subsequent studies [28], [29], [30]. However, these studies began introducing additional auxiliary tasks to further enhance the model's capabilities. In [28], the authors added the task of identifying boundaries as an auxiliary task. Meanwhile, [29] expanded the approach by incorporating a classification task alongside the reconstruction with a Variational Autoencoder (VAE). These developments reflect an ongoing effort to diversify and strengthen the utility of multi-task learning frameworks in various applications.

V. Current Challanges

The challenges in brain tumor segmentation encompass several critical aspects. Firstly, the scarcity of diverse data integration, especially the absence of additional modalities such as genomic information, limits the depth and breadth of understanding tumor patterns. Secondly, the arduous and resource-intensive process of annotating medical image data obstructs the availability of extensive annotated datasets. Strategies like data augmentation, transfer learning, and weakly supervised methods aim to alleviate this challenge.

Challenges persist in enhancing computational efficiency for large-scale medical imaging datasets and ensuring model generalizability across diverse patient demographics and tumor variations. Innovations addressing computational optimization and robustness to heterogeneous data remain crucial for advancing brain tumor semantic segmentation.

VI. Conclusion

In conclusion, the landscape of semantic segmentation for brain tumors from MRI imaging data is diverse and continually evolving. The evaluation metrics such as Dice Loss, Cross Entropy, and Jaccard Similarity serve as crucial benchmarks for assessing the accuracy and efficacy of segmentation models. The methodologies adopted for input data modalities and architectural designs showcase a spectrum of approaches, ranging from 2D and 3D modeling to the fusion of multiple MRI modalities, each with its unique advantages and considerations.

The integration of Convolutional Models with Graphical Models, Encoder-Decoder Models, Generative Adversarial Networks, and Multi-Task Learning approaches has significantly advanced the field, offering promising avenues for accurate and robust brain tumor segmentation. Yet, there are challenges, like making computations faster and ensuring that the models work well for different groups of patients. These areas need more exploration and new ideas in the field of brain tumor research.

References

[1] E. K. Lee, E. J. Lee, S. Kim, and Y. S. Lee, "Importance of Contrast-Enhanced Fluid-Attenuated inversion recovery magnetic resonance imaging in various intracranial pathologic conditions," *Korean J Radiol*, vol. 17, no. 1, pp. 127–141, Jan. 2016.

[2] J. Hu, X. Gu, and X. Gu, "Dual-pathway densenets with fully lateral connections for multimodal brain tumor segmentation," *International Journal of Imaging Systems and Technology*, vol. 31, 08 2020.

[3] B. H. Menze and Jakab, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans Med Imaging*, vol. 34, no. 10, pp. 1993–2024, Dec. 2014.

[4] V. Rajput, "Robustness of different loss functions and their impact on networks learning capability," 2021.

[5] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, *Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations*. Springer International Publishing, 2017, p. 240–248. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-67558-9$_2$8

[6] *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers*. Springer International Publishing, 2018. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-75238-9

[7] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in mri images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

[8] F. Dehghani, A. Karimian, and H. Arabi, "Joint brain tumor segmentation from multi mr sequences through a deep convolutional neural network," 2022.

[9] Q. Ru, G. Chen, and Z. Tang, "Brain tumor image segmentation method based on m-unet network," in *2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, 2021, pp. 243–246.

[10] K.-L. Tseng, Y.-L. Lin, W. Hsu, and C.-Y. Huang, "Joint sequence learning and cross-modality convolution for 3d biomedical segmentation," 2017.

[11] Z. Xiaomei, Y. Wu, G. Song, l. Zhenye, Y. Fan, and Y. Zhang, "Brain tumor segmentation using a fully convolutional neural network with conditional random fields," 04 2016, pp. 75–87.

[12] G. Shen, Y. Ding, T. Lan, H. Chen, and Z. Qin, "Brain tumor segmentation using concurrent fully convolutional networks and conditional random fields," 03 2018, pp. 24–30.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[14] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention u-net: Learning where to look for the pancreas," 2018.

[15] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. Roth, and D. Xu, "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images," 2022.

[16] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U2-net: Going deeper with nested u-structure for salient object detection," *Pattern Recognition*, vol. 106, p. 107404, Oct. 2020. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2020.107404

[17] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," 2016.

[18] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[19] D. Li, J. Yang, K. Kreis, A. Torralba, and S. Fidler, "Semantic segmentation with generative models: Semi-supervised learning and strong out-of-domain generalization," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[20] C. Zhang, Y. Song, S. Liu, S. Lill, C. Wang, Z. Tang, Y. You, Y. Gao, A. Klistorner, M. Barnett, and W. Cai, "Ms-gan: Gan-based semantic segmentation of multiple sclerosis lesions in brain magnetic resonance imaging," in *2018 Digital Image Computing: Techniques and Applications (DICTA)*, 2018, pp. 1–8.

[21] G. Sohaliya and K. Sharma, "Semantic segmentation using generative adversarial networks with a feature reconstruction loss," in *2021 Asian Conference on Innovation in Technology (ASIANCON)*, 2021, pp. 1–7.

[22] H. Chen, Z. Qin, Y. Ding, and T. Lan, "Brain tumor segmentation with generative adversarial nets," in *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 2019, pp. 301–305.

[23] R. Caruana, "Multitask learning," *Machine Learning*, vol. 28, no. 1, pp. 41–75, Jul 1997. [Online]. Available: https://doi.org/10.1023/A:1007379606734

[24] Y. Zhang and Q. Yang, "A survey on multi-task learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 12, pp. 5586–5609, 2022.

[25] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017.

[26] H. Shen, R. Wang, J. Zhang, and S. McKenna, "Multi-task fully convolutional network for brain tumour segmentation," in *Medical Image Understanding and Analysis*, M. Valdés Hernández and V. González-Castro, Eds. Cham: Springer International Publishing, 2017, pp. 239–248.

[27] A. Myronenko, "3d mri brain tumor segmentation using autoencoder regularization," 2018.

[28] J. Iwasawa, Y. Hirano, and Y. Sugawara, "Label-efficient multi-task segmentation using contrastive learning," 2020.

[29] L. Weninger, Q. Liu, and D. Merhof, "Multi-task learning for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi and S. Bakas, Eds. Springer International Publishing, 2020, pp. 327–337.

[30] Y. Liu, F. Mu, Y. Shi, and X. Chen, "Sf-net: A multi-task model for brain tumor segmentation in multimodal mri via image fusion," *IEEE Signal Processing Letters*, vol. 29, pp. 1799–1803, 2022.