Using 3D U-Net for Brain Tumour Segmentation from Magnetic Resonance Images

[1] Muhammed Uhudhan ATEŞ, [2] Recep Tahir GÜNLÜ, [*3] Ekin EKİNCİ, [4] Zeynep GARİP

[1] Department of Computer Engineering, Sakarya University of Applied Sciences, Türkiye, b200109024@subu.edu.tr

[2] Department of Computer Engineering, Sakarya University of Applied Sciences, Türkiye, b200109030@subu.edu.tr

[*3] Corresponding Author, Department of Computer Engineering, Sakarya University of Applied Sciences, Türkiye, ekinekinci@subu.edu.tr

[4] Department of Computer Engineering, Sakarya University of Applied Sciences, Türkiye, zbatik@subu.edu.tr

Abstract

Brain tumors within the skull can lead to serious health issues. The rapid and accurate detection and segmentation of tumor regions allow patients to receive appropriate treatment at an early stage, increasing their chances of recovery and survival. Various medical imaging methods, such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), digital pathology, and Computed Tomography (CT), are used for the detection of brain tumors. Nowadays, with advancing technology and hardware, concepts like artificial intelligence (AI) and deep learning (DL) are becoming increasingly popular. Many AI methods are also being utilized in studies on brain tumor segmentation. This paper proposes a 3D U-Net DL model for brain tumor segmentation. The training and testing processes are carried out on the Brain Tumor Segmentation (BraTS) 2020 dataset, which is widely used in the literature. As a result, an Intersection over Union (IoU) score of 0.81, a Dice score of 0.87, and a pixel accuracy of 0.99 are achieved. The proposed model has the potential to assist experts in diagnosing the disease and developing appropriate treatment plans, thanks to its ability to segment brain tumors quickly and with high accuracy.

Keywords: Magnetic resonance images; Brain tumour; Deep learning; Semantic segmentation; 3D U-Net

1. INTRODUCTION

Cancer is the second most common cause of death in many countries, whether developed or underdeveloped [1]. Brain tumors are among the highest-risk types of cancer, especially in young patients [2]. Between 85% and 90% of primary central nervous system (CNS) cancers are of this kind. In 2020, it was projected that 308,102 individuals globally received a primary brain or spinal cord tumor diagnosis [3]. Early diagnosis and treatment of brain tumors greatly increase the chances of recovery.

Medical imaging is crucial for diagnosing diseases and planning their treatment. Nowadays, medical imaging equipment is rapidly developing, and a range of techniques like magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and digital pathology are used. These techniques can provide incredibly detailed images of tissues and organs. In brain tumor patients, structures in the central nervous system and brain are generally examined using MRI, CT, and PET imaging technologies. However, evaluating and interpreting these images is a time-consuming and error-prone process.

Semantic segmentation is an image processing technique essential for the diagnosis and treatment of brain tumors.

This method assigns each pixel in brain images obtained through techniques such as MRI or CT to a specific class. For example, it can distinguish between different structures such as brain tissue, tumor lesions, and normal tissue [4]. In this way, the semantic segmentation method helps medical imaging experts quickly and accurately identify and classify brain tumors. This provides valuable information to healthcare professionals in important decision-making processes such as surgical planning, treatment management, and patient follow-up. It is also an important tool for early diagnosis and formulating personalized treatment strategies. Consequently, the significance of semantic segmentation in the realm of brain tumor diagnosis and treatment is garnering growing interest in both medical imaging research and applications [4].

In recent years, with the proliferation of deep learning (DL), there has been a surge of studies in the literature on brain tumor segmentation. In the study by Zhou et al., an efficient 3D residual neural network (ERV-Net) was proposed for brain tumor segmentation, featuring reduced computational complexity and lower GPU memory usage [5]. ERV-Net used 3D ShuffleNetV2, a computationally efficient network, primarily to reduce GPU memory usage and improve the efficiency of ERV-Net. The challenge of multimodal brain

tumor segmentation was tested on the BraTS 2018 dataset, achieving a Dice score of 91.21%.

The method developed in the study by Ye et al. represented a significant advancement in brain tumor segmentation through radiomic analysis [6]. Manual segmentation by radiologists is laborious and subjective, limiting the number of cases studied and hindering the reproducibility of clinical studies. In this context, a model incorporating a 3D Center-Cut Dense Block and a parallel network architecture was proposed for automatic tumor segmentation. This structure aimed to achieve detailed segmentation of medical images by decentralizing high-level control to lower layers of neural networks. Their architecture achieved notable success, obtaining an 88.4% Dice score and 83.8% sensitivity on the BraTS 2015 dataset, and an 88.7% Dice score and 84.3% sensitivity on the BraTS 2017 dataset.

The state-of-the-art model by Zhou et al. addressed challenges such as spatial information loss and inadequate multi-scale lesion processing faced by traditional deep convolutional neural networks in brain tumor segmentation [7]. This architecture enhanced the model's discrimination ability, particularly in segmenting tumors of varying sizes. Furthermore, the inclusion of a 3D fully connected Conditional Random Field enhanced results by incorporating both appearance and spatial consistency into the model's outputs. The model demonstrated consistent performance across different time periods and data conditions, achieving success rates of 86%, 82%, and 86% on the BraTS 2013, 2015, and 2018 datasets, respectively.

In the work by Zhang et al., Triple Intersecting U-Nets (TIU-Nets) were proposed for the segmentation of brain gliomas [8]. This model consists of two main components: the binary class segmentation U-Net (BU-Net) and the multi-class segmentation U-Net (MU-Net). The MU-Net utilizes multi-resolution features generated by the BU-Net. According to experimental results, the suggested 2D/3D TIU-Nets outperformed other cutting-edge techniques such as FCN, U-Net, SegNet, IVD-Net, FCDenseNet, DMFNet, CRDN, and DeepMedic in terms of segmentation accuracy. Additionally, the method achieved success rates of 82% on the BraTS 2015 dataset and 99.5% on the BrainWeb dataset. The proposed method improved segmentation accuracy and robustness by integrating advanced deep learning models [9]. Li et al. proposed a sophisticated approach for brain tumor segmentation that utilizes a combination of cascaded 3D U-Net and 3D U-Net++ architectures and got 89% Dice score.

Henry et al. proposed a 3D U-Net architecture enhanced with self-ensembling and deep supervision strategies to improve segmentation accuracy and robustness using the BraTS 2020 dataset. The model achieved a high Dice score of 90% [10]. In another study, the authors presented a novel approach to brain tumor segmentation using a 3D Generative Adversarial Network (GAN). The proposed method, called Vox2Vox, leverages the power of GANs to enhance the accuracy and robustness of segmentation in 3D medical images from the BraTS 2020 dataset. The model achieved Dice scores of 90% [11]. Lin et al. utilized a 3D U-Net architecture with deep supervision and context aggregation to enhance segmentation performance for the BraTS 2020 dataset and

obtained a Dice score of 92.3% [12]. Sasank and Venkateswarlu devised a novel approach to brain tumor segmentation that incorporates tumor growth prediction into the segmentation process [13]. The method aimed to enhance the accuracy of brain tumor segmentation by using a Full Resolution Convolutional Network (FrCN), which is capable of processing high-resolution images while maintaining spatial details. Based on experimental results, the method achieved accuracies of 97% for BraTS 2020, 95.56% for BraTS 2019, and 95.23% for BraTS 2018.

In this study, we aim to develop a fully automatic segmentation pipeline for MR images of brain tumors obtained from the BraTS 2020 dataset. To achieve this goal, we use the state-of-the-art 3D U-Net architecture. An experimental study is conducted on the proposed model, and its performance results are compared with those of other studies in the literature. The experimental results showed that the 3D U-Net achieved a pixel accuracy of 0.99, a Dice score of 0.87, and an IoU score of 0.81. In light of these results, this study provides evidence of suitable techniques for automatic brain tumor segmentation.

The rest of the paper is organized as follows: In the next section, an overview of the dataset and a formal definition of the segmentation algorithms are presented in detail. In the third section, the proposed architecture, performance metrics, and experimental results are discussed. Finally, section four concludes the paper.

2. MATERIALS AND METHODS

This section provides information about experimental dataset, background of segmentation and segmentation algorithms.
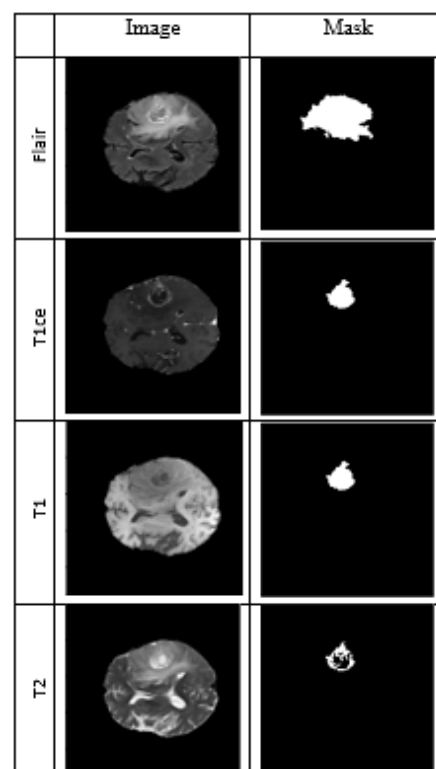

Figure 1. Samples from BraTS 2020 dataset

2.1. Dataset

The BraTS dataset, widely utilized in medical image segmentation studies, is the largest publicly available resource for brain tumor data [14]. Initially introduced in 2012 for the BraTS competition, the dataset has been regularly updated since. It consists of brain images captured using MRI devices, annotated and segmented by specialist doctors and radiologists. The dataset includes several versions, such as BraTS 2012, BraTS 2015, BraTS 2017, BraTS 2018, and BraTS 2020. Researchers have employed this dataset in various studies, including brain tumor detection, computer-aided diagnosis systems, and automatic medical image segmentation. In this study, the BraTS 2020 dataset obtained via the Kaggle platform is used [15].

The BraTS 2020 dataset used includes 369 MR images for training and 125 MR images for validation. Each brain image includes four distinct MRI modalities (FLAIR, T1ce, T1, T2) along with their corresponding masks, as illustrated in Figure 1.

These images in the dataset provide information about brain tissue. FLAIR images identify areas of edema, while T1ce images enhance the tumor boundary, revealing bright signals from substances accumulated in active cell areas [16]. T1 images depict healthy tissues, and T2 images are effective in highlighting bright signals indicating areas of edema. Each MRI modality consists of 155 slices, forming a 3D data structure for each brain when combined. This allows for a comprehensive examination and analysis of each modality and mask in 3D.

2.2. Segmentation

Image segmentation has been a fundamental problem in fields like computer vision and image processing for many years. This technique, used to understand and interpret images, involves dividing images into multiple segments and objects. Image segmentation plays a central role in various fields, including disease detection through medical images, tissue volume measurement, autonomous vehicles, video surveillance, and augmented reality [17].

Examining the network structure designs of medical image segmentation approaches reveals five main classes: Convolutional Neural Networks (CNNs), Autoencoder Networks (AEs), Fully Convolutional Networks (FCNs), Generative Adversarial Networks (GANs), and Region-Based Networks (R-FCNs) [4].

Our 3D U-Net model in this study is built on the U-Net architecture, placing it within the FCN class in terms of network structure.

2.2.1. Fully Convolutional Networks

FCNs are deep learning models commonly used for pixel-level classification tasks such as image segmentation. They are termed "Fully Convolutional" because each layer includes convolution processes. These networks evolved from Convolutional Neural Network architectures [18]. Figure 2 illustrates the FCN architecture; unlike traditional CNNs, FCNs maintain the size of the input image at the output. In other words, the output size of FCNs matches the input size. To achieve this, up-sampling is performed on the output image [19].

FCNs are widely recognized in medical image segmentation studies for consistently improving and achieving successful results. The U-Net model, frequently employed in this field, represents an advanced form of the FCN model.
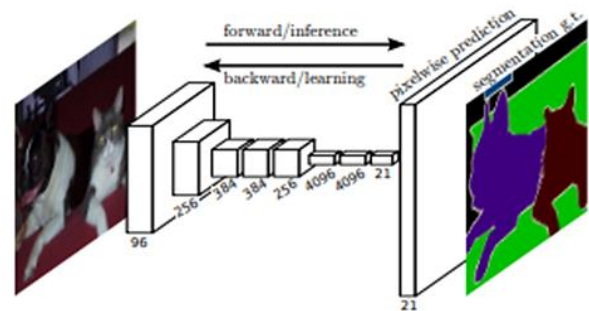

Figure 2. FCN architecture [19]

2.2.2. U-Net

The U-Net architecture, devised by Ronneberger et al. in 2015 for biomedical image segmentation, comprises 23 convolutional layers [20].

The model consists of two main components: the encoder and the decoder, as depicted in Figure 3. The encoder extracts features from medical images using convolutional layers, activation functions (ReLU), and pooling layers. Conversely, the decoder expands these features using upsampling layers. This expansion process enables precise localization of features within the image, thereby ensuring accurate and detailed segmentation [20].

2.2.3. 3D U-Net

This paper presents the 3D U-Net model, a powerful deep learning architecture created specifically for medical image segmentation [21]. Unlike traditional 2D convolutional neural networks, 3D U-Net processes data in three dimensions, making it particularly suitable for 3D medical imaging applications like MRI or CT scans. The model enhances the ability to identify complex structures in medical data by simultaneously capturing spatial features across height, width, and depth.

The encoder structure of the 3D U-Net model is composed of four blocks, similar to the original U-Net architecture. Each block includes convolution, dropout, and max-pooling 3D layers. In the decoder structure, each block begins with upsampling, followed by convolution and dropout layers. This ensures that the input brain MRI image is processed without altering its size at the output.
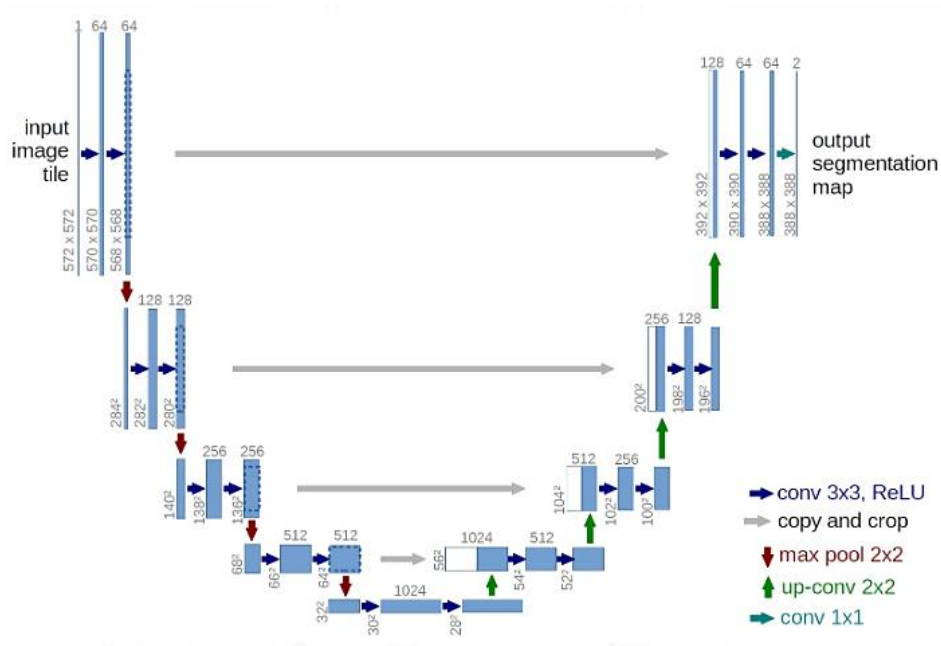
Figure 3. U-Net architecture [15]

3. EXPERIMENTAL STUDY

This section explains the proposed architecture with details. And evaluates the performance of 3D U-Net model via figures and tables.

3.1. Proposed Architecture

The proposed model comprises a total of 23 convolutional layers, as depicted in Figure 4. The encoder structure consists of four blocks. In the first block, the convolutional process utilizes 16 filters of size 3 x 3 x 3. The feature maps resulting from each convolutional process are then passed through the ReLU activation function. To reduce overfitting, a dropout layer with a rate of 0.1 is added, and a MaxPooling3D layer is applied to reduce the spatial dimensions by half.

The second block has a similar structure to the first block but uses 32 filters for the convolutional process. In the third and fourth blocks, 64 filters with a 0.2 dropout rate and 128 filters with a 0.2 dropout rate are used, respectively. The sub-block, the deepest layer of the 3D U-Net architecture, utilizes 256 filters and a dropout ratio of 0.3. Max-pooling is not applied here because this block transitions to the decoder path.

In the decoder path, each block starts with upsampling, accomplished using the Conv3DTranspose layer. In the first upsampling block, convolution is performed with 128 filters and a dropout ratio of 0.2, and the feature maps from the corresponding block in the encoder path are concatenated. The second upsampling block follows a similar structure with 64 filters and a 0.2 dropout rate. The third and fourth upsampling blocks use 32 filters with a 0.1 dropout rate and 16 filters with a 0.1 dropout rate, respectively. Finally, a 1x1x1 convolutional process is applied for each class in the output layer, and the final segmentation output is produced using the softmax activation function. These stages enable the model to successfully learn fine details and complex structures in medical data, optimizing segmentation performance.
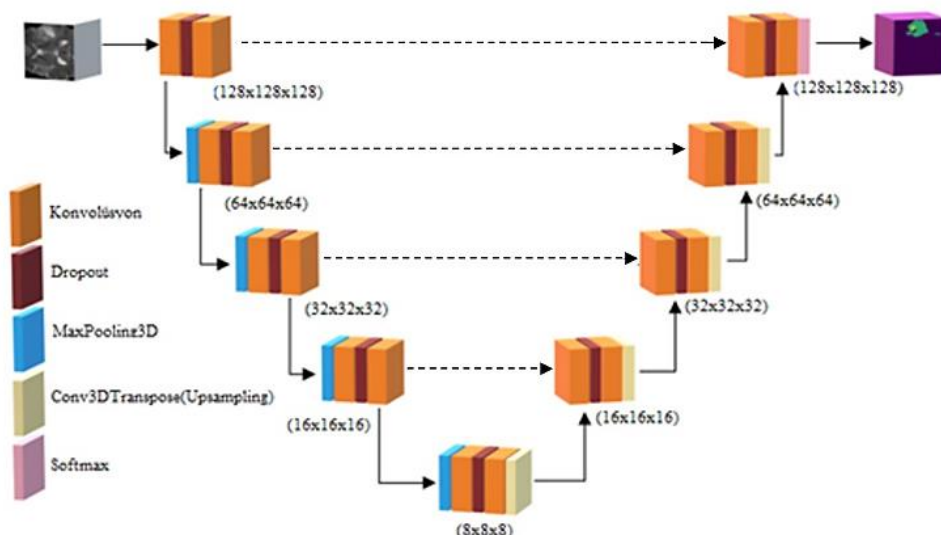

Figure 4. Proposed 3D U-Net architecture

3.2. Performance Metrics

The training and testing processes of the brain MRIs from the BraTS 2020 dataset were completed using the proposed 3D U-Net model. To evaluate the segmentation performance of the model, the commonly used metrics of Pixel Accuracy (1), Dice score (2), and IoU score (3) are employed based on the literature [16]. These metrics are calculated using True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

TP represents the number of pixels or instances in both the ground truth and predicted segmentation that are accurately classified as part of the target class. FP is the number of pixels or instances that are incorrectly labelled as part of the target class in the segmentation prediction. TN refers to the pixels or instances that are correctly identified as not belonging to the target class in both the predicted segmentation and the ground truth segmentation. FN represents the number of pixels or instances in the ground truth that are part of the target class but were not recognized as such in the predicted segmentation.

Pixel accuracy refers to the ratio of correctly classified pixels to the total pixels in an image. This metric is often used in pixel-based classification problems such as image segmentation and is crucial for evaluating the overall performance of the model.

$$Pixel\ Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

The Dice score is one of the most frequently used evaluation metrics in published medical image research.

$$Dice\ Score = \frac{2TP}{2TP+FP+FN} \qquad (2)$$

The formula above measures the agreement between the predicted and ground truth segmentations, with a Dice score of 1 indicating perfect agreement and a score of 0 indicating no overlap.

The IoU score is a performance measure used in classification and segmentation problems, often employed in object detection and image segmentation. It measures how accurately a model's predictions match real objects. The IoU score is calculated using the following equation:

$$IoU = \frac{TP}{TP+FP+FN} \qquad (3)$$

The IoU score calculates the ratio of the intersection area between the predicted objects and the real objects to their union area. It indicates the degree of overlap between the two regions representing the two clusters.

3.3. Experimental Results

The experimental research on the 3D U-Net model, developed for segmenting brain MRI images from the BraTS 2020 dataset, was conducted using the Python programming language. The model's training and testing were performed on a machine equipped with an Nvidia GeForce RTX 4060

GPU, utilizing Jupyter Notebook. TensorFlow and Keras libraries were employed for model development. During training, the batch size was optimized to 2, with 100 epochs and a learning rate set to 0.0001. The Adam optimizer was selected as the learning algorithm.

In the experiment, the pixel accuracy and loss values obtained during the training can be observed in Figures 4 and 5, respectively. Figure 4 illustrates that the proposed model achieved significant performance results in terms of pixel accuracy during both the training and validation stages.

Figure 4. Train and validation set pixel accuracy

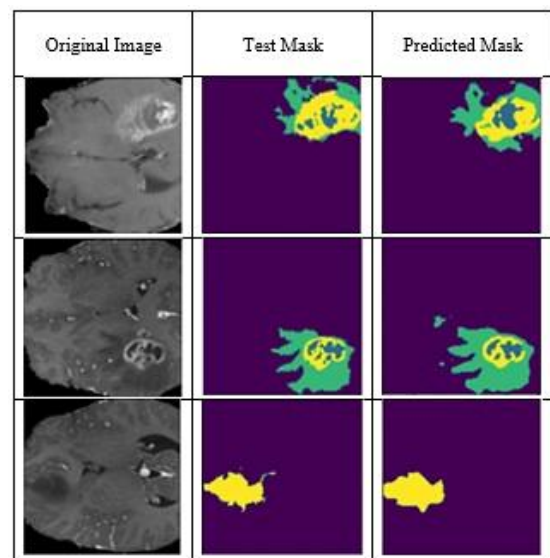**Figure 5.** Train and validation set loss

Figure 6. Visualized segmentation examples of MRI images

The test results of the proposed 3D U-Net model were evaluated using pixel accuracy, Dice score, and IoU score metrics. The model achieved a pixel accuracy of 0.99 during testing. The average Dice score and IoU score were 0.87 and 0.81, respectively.

Figure 6 displays visual results of the proposed model's brain tumour segmentation. The images illustrate the original brain MRI, the actual segmentation result, and the model's predicted segmentation output.

The model proposed in this study is compared with segmentation studies conducted on the BraTS datasets in Table 1, considering Dice score and pixel accuracy values from relevant studies. Reviewing the literature, it is evident that the proposed 3D U-Net model has achieved significant success.

Table 1. Comparing performances of studies using the BraTS dataset

Reference	Dataset	Pixel Accuracy	Dice Score
[4]	BraTS 2018	0.91	0.91
[5]	BraTS 2017	N/A	0.88
[6]	BraTS 2018	N/A	0.76
[7]	BraTS 2015	N/A	0.82
[10]	BraTS 2020	N/A	0.90
[11]	BraTS 2020	N/A	0.90
[12]	BraTS 2020	N/A	0.92
[13]	BraTS 2020	N/A	0.97
Proposed	**BraTS 2020**	**0.99**	**0.87**

4. CONCLUSION

A 3D U-Net segmentation model for the early diagnosis and detection of brain cancers was proposed in this study. The BraTS 2020 dataset was used to train and evaluate the suggested model. According to the experimental results, the model achieved an IoU score of 0.81, a Dice score of 0.87, and a pixel accuracy of 0.99.

Comparison with other studies in the literature demonstrated that the proposed model achieved significant performance results. It is anticipated that this model will be highly beneficial for radiologists, neurologists, and neurosurgeons working with brain MRI images, enabling rapid and accurate diagnoses. Therefore, the 3D U-Net segmentation model represents a promising approach for facilitating fast and accurate decision-making.

Author contributions: Concept – M.U.A., R.T.G, E.E., Z.G.; Data Collection & Processing – M.U.A., R.T.G; Literature Search – M.U.A., R.T.G; Writing – M.U.A., R.T.G, E.E., Z.G.

REFERENCES

[1] R. Bütüner and M. H. Calp, "Predicting of Melanoma Skin Cancer Using Machine Learning Methods," Gazi Journal of Engineering Sciences, vol. 10, no. 1, pp. 141–154, 2024.

[2] A. Ergin, R. Özdilek, and N. Dutucu, "The Distribution and Types of Women's Cancers Seen Between 2012 And 2017: A University Hospital Example," Journal of Women's Health Nursing Jowhen, vol. 5, no. 1, pp. 1–21, 2019.

[3] The American Cancer Society medical and editorial content team, About Brain and Spinal Cord Tumours in Adults. American Cancer Society, 2020. Accessed on: May. 10, 2024. [Online]. Available: https://www.cancer.org/content/dam/CRC/PDF/Public/8567.00.pdf

[4] T. Şentürk and F. Latifoğlu, "Deep Learning Based Methods for Biomedical Image Segmentation: A Review," Dicle University Journal of the Institute of Natural and Applied Sciences, vol. 12, no. 1, pp. 161–187, 2023.

[5] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen and X. Gao, "ERV-Net: An efficient 3D residual neural network for brain tumour segmentation," Expert Systems with Applications, vol. 170, p. 114566, 2021.

[6] F. Ye, Y. Zheng, H. Ye, X. Han, Y. Li, J. Wang and J. Pu, "Parallel pathway dense neural network with weighted fusion structure for brain tumour segmentation," Neurocomputing, vol. 425, pp. 1–11, 2021.

[7] Z. Zhou, Z. He and Y. Jia, "AFPNet: A 3D fully convolutional neural network with atrous-convolution feature pyramid for brain tumour segmentation via MRI images," Neurocomputing, vol. 402, pp. 235–244, 2020.

[8] J. Zhang, J. Zeng, P. Qin and L. Zhao, "Brain tumour segmentation of multi-modality MR images via triple intersecting U-Nets," Neurocomputing, vol. 421, pp. 195–209, 2021.

[9] P. Li, W. Wu, L. Liu, F. M. Serry, J. Wang, and H. Han, "Automatic brain tumor segmentation from Multiparametric MRI based on cascaded 3D U-Net and 3D U-Net++," Biomedical Signal Processing and Control, vol. 78, 103979, 2022.

[10] T. Henry, A. Carré, M. Lerousseau, T. Estienne, C. Robert, N. Paragios, and E. Deutsch, "Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-net neural networks: a BraTS 2020 challenge solution," In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes, Held in Conjunction with MICCAI 2020, pp. 327–339, 2020.

[11] M. D. Cirillo, D. Abramian, and A. Eklund, "Vox2Vox: 3D-GAN for brain tumour segmentation," In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, pp. 274–284, 2020.

[12] M. Lin, S. Momin, Y. Lei, H. Wang, W. J. Curran, T. Liu, and X. Yang, "Fully automated segmentation of brain tumor from multiparametric MRI using 3D context deep supervised U-Net," Medical Physics, vol. 48, no. 8, pp. 4365–4374, 2021.

[13] V. V. V. Sasank, and S. Venkateswarlu, "An automatic tumour growth prediction based segmentation using full resolution convolutional network for brain tumour," Biomedical Signal Processing and Control, vol. 71, 103090, 2022.

[14] D. LaBella, et al., "A multi-institutional meningioma MRI dataset for automated multi-sequence image segmentation," Scientific Data, vol. 11, no. 496, pp. 1–8, 2024.

[15] Brain Tumour Segmentation 2020 Dataset. Accessed on: May. 2, 2024. [Online]. https://www.kaggle.com/awsaf49/brats20-dataset-training-validation.

[16] E. Gökçe, M. F. Demiral, A. H. Isık, and M. Bilen, "Brain Tumour Segmentation with Convolutional Neural Networks," El-Cezerî Journal of Science and Engineering, vol. 9, no. 4, pp. 1518–1528, 2022.

[17] P. Khurana, A. Sharma, S. N. Singh and P. K. Singh, "A survey on object recognition and segmentation techniques," in 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 3822–3826, 2016.

[18] B. Kayhan, "Deep learning based multi organ segmentation in computed tomography images," Master thesis, Konya Technical Univ., Konya, Türkiye, 2022.

[19] J. Long, E. Shelhamer, T. Darrell, "Fully convolutional networks for semantic segmentation," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440, 2015.

[20] O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in Medical image computing and computer-assisted intervention–MICCAI 2015, pp. 234–241, 2015.

[21] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, pp. 424–432, 2016.