

Journal of Integrated SCIENCE & TECHNOLOGY

Review

Brain tumor segmentation in multi-modal MRI: A comparative study

Ghanshyam Raghuwanshi,¹ Devanarayanan Vinu¹, Punit Gupta², Mayank Kumar Goyal^{3*}, Supriya Khaitan⁴, Madhusudhan HS⁵

¹Computer and Communication Engineering, Manipal University Jaipur. India. ²Pandit Deendayal Energy University, Gandhinagar, Gujarat. India. ³Department of Computer Science & Engineering, Sharda School of Computing Science & Engineering, Sharda University, Greater Noida. India. ⁴Department of Engineering and Technology, Bharati Vidyapeeth (Deemed to be University), Navi Mumbai. India. ⁵Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka. India.



ABSTRACT

A brain tumor is considered as an abnormal growth or lesion inside of the brain or near it. These tumors can be cancerous; malignant, known as Glioma due to abnormal growth of glial cells, or it can be noncancerous or benign as well. As such, the early detection and



treatment of brain tumors is a crucial task today. Various technologies exist and are continuously being developed to capture high quality images of brain tissue, allowing experienced medical professionals to diagnose tumors in its early stage, so that they can take action to provide the necessary treatment based on their findings. With emerging interests and technological advancements in Artificial Intelligence and Machine Learning models, the task of manually classifying and segmenting medical images can become less burdensome. Automated diagnosis and classification of brain tumors using deep learning models can provide a way to overcome the problems of manual segmentation. This work aims to compare various studies involving brain tumor segmentation using deep learning methods on various aspects like accuracy the data used and hyper parameters of those architectures for task of brain tumor segmentation. The work compares existing model proposed using U-Net ,Link-Net,PSP-Net and FPN. Where the the performance stydy is compared using Accuracy, Dice Coefficient, Sensitivity, Precision and Specificity. The result shows PSP-Net achieves the highest Dice coefficient, indicating superior segmentation accuracy at the expense of computational intensity.

Keywords: Brain, Tumor, Segmentation, Deep Learning, U-Net, Link-Net, PSP-Net, FPN

INTRODUCTION

A brain tumor can be defined as abnormal growth of cells inside the brain or skull. The cells grow uncontrollably for reasons we still do not fully understand. Brain tumors are of different types, and can vary based on the place of growth, growth rate, and the type of cells that are growing uncontrollably.¹ The three important categories of brain tumors are benign, malignant, and metastatic tumors. Within

^{*}Corresponding Author: Dr. Mayank Kumar Goyal, Department of Computer Science & Engineering, Sharda School of Computing Science & Engineering, Sharda University, Greater Noida . Tel: +91 91367 80090; Email: mayank.rkg.it@gmail.com

Cite as: J. Integr. Sci. Technol., 2025, 13(7), 1153. URN:NBN:sciencein.jist.2025.v13.1153 DOI:10.62110/sciencein.jist.2025.v13.1153



©Authors CC4-NC-ND, ScienceIN https://pubs.thesciencein.org/jist

Journal of Integrated Science and Technology

these categories there are various subtypes as well. A malignant tumor is a tumor comprising of cancerous cells that typically grows from the glial tissue of the brain. The cancerous tumors that grow from the glial tissue are also known as gliomas. These tumors grow at variable rates and can spread to other regions of the body, so swift detection and treatment are necessary.^{2,3} A benign tumor is a noncancerous tumor. The most common type of benign brain tumors are meningiomas, which are tumors that form in the meninges, which is a layer of tissue protecting the brain and also the spinal cord. Another category of brain tumors are metastatic, which are tumors that form from the spreading of cancerous growths from other parts of the body.

Whether a brain tumor is cancerous or not, it can still cause life threatening problems. The skull or cranium is solid and does not move to accommodate the tumor, causing other regions of the brain to be hampered or under pressure, which can affect brain from functioning properly. So, it is highly beneficial if the tumors are detected early, and treatment is provided at the right time. Numerous medical imaging techniques, including magnetic resonance imaging (MRI), computed tomography (CT), simple photon emission computed tomography (SPECT), ultrasound, positron emission tomography (PET), and X-ray, are available to diagnose brain tumors. Because it provides a more contrasting image than other medical imaging methods, magnetic resonance imaging (MRI) is the imaging technique most frequently used to detect tumors. A medical professional such as a radiologist or a neurologist would use these images to detect brain tumors and then segment or outline the tumor so that it is easily visible. But there are parts of a tumor, such as the necrotic region containing dead cells or the fluid region surrounding the tumor, known as edema, which needs to be clearly differentiated from the actual core of the tumor. This can be strenuous work for medical professionals. In a 3-Dimensional MRI, the number of sections of the brain being scanned goes up to 240 slides, and having multiple viewing angles of the brain also adds to the number of images to be segmented. Here, automating the segmentation of brain MRIs can provide significant help to the radiologist or neurologist. Over the years, deep learning techniques such as Convolutional Neural Networks (CNNs), has shown great promise in automating the segmentation and classification tasks. The advanced models being created today use neural networks to extract important features from medical images, thereby creating an accurate identification and categorization of brain tumors.



Figure 1. An MRI scan of brain with a tumor from the BraTS 2023 dataset.

The motivation behind this project is the important need for more accurate, reliable, and efficient methods for diagnosing brain tumors. Brain tumor detection in medical imaging provides significant challenge, requiring accurate identification and segmentation of the tumor so that we can give the patient the required treatment at the right time before it is too late. Manual segmentation of brain tumors from MRI scans is subjective, time consuming, and can lead to human error. This highlights the pressing need for an automatic method of segmenting brain MRIs.

LITERATURE REVIEW

During the research phase of this project leading up to the review of semantic segmentation models, I read and reviewed papers and articles from peers and different sources to gain an understand of the general approach and techniques of segmenting brain tumor images. Across most sources, semantic segmentation was achieved using many different methods.⁴ Of them, 5 common techniques are going to be reviewed in this study, namely U-Net, Link-Net, SegCaps-Net, AttU-Net, and SegFormer.

U-NET ARCHITECTURE:

The U-Net architecture is a model that consists of a contracting or downsampling path known as encoder, and an expanding or upsampling path known as decoder. This path gives the model a 'U' shape, which is why the model is named so. The encoder-decoder architecture is a common structure in image segmentation models. Each encoder consists of two convolutional layers and a max pooling layer to decrease the spatial dimensions of the image.



Figure 2. Architecture of U-Net.

The value from max pooling layer is sent to two places, first, for further downsampling, and second, to merge with the upsampling layer on the other side. The decoder consists of 2 convolutional layers and a transposed convolution layer that increases the spatial dimensions of the image. The deepest layer in the network is the bottleneck which has the smallest spatial dimensions and the most feature channels. The following is a review of 5 studies conducted on brain tumor segmentation using this model.

H. Dong et al.⁵ introduced a U-Net-like architecture with an encoding and decoding branch with zero padding along the encoding path to ensure that all output sizes in the convolutional layers of the encoding and decoding path are matching. Their approach reached comparable results to us in the total tumor volumes and core tumor volumes with the advantage over by having a Dice Similarity Coefficient of 0.86 for High Grade and Low-Grade Gliomas. T. A. Tuan et al.⁶ proposed a technique where first round of processing was applied on Bitplane slicing for arranging pixels of significant bit pattern and then further image samples were generated from the database. The U-Net CNN network is then used. with multi-kernels employed for segmentation to figure out the most suitable k: states. Form validation in 2018 BRATS challenge data, the method achieved 82%, 68% and 70% of Dice in validation and 51%, 48%, 77% in testing for Tumor Core, Enhancing Tumor and Whole Tumor respectively. A. Kermi et al.⁷ presented a U-Net-like network for 2D images based on Deep Convolutional Neural Networks,

optimised with Stochastic Gradient Descent for loss minimisation. On average, our method achieved an enhancing tumor, whole tumor and core tumor dice score of 0.783, 0.868 and 0.805 respectively which are competitive of top models. W. Chen et al.⁸ questioned the restric-tions of 2D and 3D convolutions, and brought in a Separable 3D U-Net for brain tumor segmentation. Their work designed by a Stereoscopic 3D (S3D) model and proposed an architecture, i.e., the S3D-UNet, for brain tumor segmentation. On the testing set, they achieved similar performance (Dice score: 0.68946 for enhancing tumor, 0.83893 for whole tumor, and 0.78347 for tumor core). Another investigation of Isensee et al.9 introduced a network based on a variation of the U-Net architecture, specifically designed to be used in conjunction with 3D input data, with blocks of voxels 128x128x128 in size, instead of 2D patches of pixels. Another advantage is that this U-Net based framework fuses input information with multiple scale, since the modified U-Net-like network could naturally fuse information at each level across scales, compared to early methods that manually integrated the inputs with a mixture of resolutions and/or pathways with different F sizes. On the BRATS 2017 test set, the proposed model achieves combined large dice score of 0.8523, 0.7762, 0.6475 for whole, core and enhancing tumor respectively. Zhu Meng et al.¹⁰ presented a novel U-Net-like model to improve the brain tumor segmentation on MRI by incorporating an end-to-end noise suppression.

The architecture, named ENS-Unet, incorporates a noise suppression into the UNet directly, it makes the network to be more noise robust. This method attempts to better delineate glioma subregions such as the Enhancing Tumor (ET), Tumor Core (TC) and Whole Tumor (WT). The ENS-Unet was tested on the BRATS 2018 and produced excellent Dice scores of 89% for the Whole Tumor, 80% for Tumor Core, and 74% for the Enhancing Tumor. These results indicate that the model performs better than conventional U-Net implementations, and has clear potential for clinical applications in the diagnosis and treatment planning of brain tumor. Pravitasari et al.¹¹, they introduced a hybrid U-Net architecture with VGG-16 and a transfer learning to simplify the network by decreasing the layers and parameters. The U-VGG16 U-Net was obtained by modifying the VGG16 by appending a set of multiple upsampling and convolution layers to the bottom of the VGG16 so as to mirror its U-Net architecture. This iteration was repeated until the global model structure became symmetrical, making a U. The model achieved an accuracy of 96%, and CCR was 95.69%. P. Gadosey et al.¹² presented a Stripping-Down UNet (SD-UNet), a light-weight novel U-Net version to perform efficient biomedical image segmentation over resource-constrained computing ecosystem by exploiting the depthwise separable convolutions, and weight-standardization and group-normalization for drastically minimizing the model magnitude and complexity, while keeping the competitive performance on the ISBI neuronal structure segmentation and MSD BRATs brain tumor segmentation benchmarks. It obtained a mean Dice Score of 82.75% on Brats dataset. U. Baid et al.¹³ proposed the use of thw Deep Learning Radiomics Algorithm for Gliomas (DRAG) Model, which is used to both segment and predict survival of glioma brain tumors using MRI data with a 3D patch-based U-Net architecture tested on the

G. Raghuwanshi et. al.

BraTS challenge 2018. On the validation set the model achieved promising Dice score values as follows -whole tumor (0.88), core (0.83), enhancing (0.75). Regarding survival prediction, it achieved an accuracy of 57.1% on the validation data, ranking the third in the task of OS prediction of the BraTS 2018 challenge. Q. Jin et al.¹⁴ proposed a residual attention U-Net (RAU-Net) architecture for liver tumor segmentation, which was motivated by fusing the advantages of residual attention learning and U-Net. This method consists of three main stages targeted to extract, in a sequential order, the liver and tumor regions. We first use a 2D RA-UNet (RA-UNet-I) that is cascaded to obtain a crude liver boundary box at a lower computational complexity. Then 3D RA-UNet (RA-UNet-II) follows to finetune such learned contours for getting a complete precise liver VOI. The architecture was capable of obtaining a Dice score of 83% for the on 3DIRCADb liver dataset.

Outcome: It is seen that applying U-Net architecture on the brain tumor semantic segmentation task is fruitful, as the U-Net is a network that was designed for this task. The architecture is able to give decent dice scores and various studies were able to use the network using different approaches in segmenting the brain MRI scans.

LINK-NET ARCHITECTURE

The Link-Net architecture is another architecture that utilizes convolutional layers designed for image segmentation tasks. It follows the same encoder-decoder design used in most architectures for the task of segmentation. The main idea behind the network is to create a balance between high accuracy and efficient computation, making it suitable for real-time applications and scenarios with limited computational resources.



Figure 3. Architecture of Link-Net.

As seen from Figure 3, the architecture of Link-Net and U-Net is quite similar, with the only difference being that Link-Net uses simple addition to keep the lost features while U-Net uses concatenation. This greatly reduces the complexity and intensiveness of the model when processing images.

Gayathri Ramasamy et al.¹⁵ used a modified multi-modal Link-Net model with a Squeeze and Excitation ResNet152 backbone to develop an effective deep learning model for semantic segmentation. A multi-modal MRI dataset comprising T1weighted, FLAIR-weighted, and T2-weighted MRI images of the human brain is used by their model. With a 99.2% accuracy rate in tumor segmentation, it shows strong performance. In another study, due to differences in tumor shape and size depending on image angles, Zahra Sobhaninia et al.¹⁶ used individual LinkNet networks for three different types of MR images (sagittal, coronal, and axial). When they used a single LinkNet network for all images, they did not achieve the same Dice scores for two angles (0.78 for coronal and 0.79 for sagittal). T. Ruba et al.¹⁷ suggested an architecture which first uses a light deep neural network to segment brain images, then integrates the semantic segmentation results with a CNN made up of LinkNet Layers. It achieved an accuracy of 0.9867. Adel Sulaiman et al.18 presented a technique that concentrates on the encoder portion's image downsampling. An intelligent LinkNet-34 model based on semantic segmentation using the EfficientNetB7 encoder was developed for this technique. Then, three separate optimizers (RMSProp, Adamax, and Adam) were used to optimize the model. With the aid of the Adamax optimizer, a very high dice coefficient of 0.915 was attained. Zahra Sobhaninia et al.¹⁶ in their further study proposed an architecture that utilized a Cascaded Dual-Scale LinkNet model that addresses the issues of tumor location and brightness level similarity with brain tissue. The Cascaded model was able to achieve a Dice coefficient of 0.8003 on evaluation criteria. P. Sameer et al.¹⁹ compared the performance of various transfer learning models such as Inception, ResNet, and ResNext when embedded with the LinkNet Decoder. The proposed models were able to produce high Dice scores, one of which achieved a notable score of 94%. P. Rajesh et al.²⁰ compared the performance of LinkNet with a ResNet50 encoder against UNet with EfficientNetB7 encoder. The LinkNet model outperformed the U-Net model, achieving a Dice score of 98.76%. T. M. Geethanjali et al.²¹ proposed a LinkNet architecture with a ResNet18 backbone for the task of Semantic Segmentation in Kidney tumors. The model achieved a remarkable dice score of 93%. C. Akyel et al.²² proposed a LinkNet with EfficientNetB7 encoder to segment lesions in images of Skin Cancer. The model achieved a high Dice score of 96.75%. An article by M. Widiansyahet al.²³ proposed a model which used a MobileNet encoder and a LinkNet decoder for segmentation of skin cancer. It achieved a considerable segmentation score, with intersection over union score as 71.5%.

Outcome: The LinkNet architecture is an impressive architecture that can provide highly accurate results that can even outperform other networks, while having a smaller amount of data to learn or train on. Since it is also a lighter network, being able to be applied in scenarios where the computational power is not enough to maintain other networks, is a huge advantage.

PSP-Net

The Pyramid Scene Parsing Network (PSPNet) is a deep learning architecture designed for semantic segmentation, introduced by Hengshuang Zhao et al. in 2017.²⁴ It extends traditional convolutional neural networks by incorporating a pyramid pooling module, which captures contextual information at multiple scales.





The PSPNet divides feature maps into four different pyramid scales, applying pooling operations of varying kernel sizes to capture global and local context. Subsequently, the original feature map is concatenated with the pooled features, improving the network's comprehension of intricate scenes. By integrating contextual information across different image regions, PSPNet is able to achieve state-of-the-art performance in a variety of semantic segmentation benchmarks thanks to this multi-scale approach.

Zhao et al.²⁵ published an article on using the PSPNet for the task of segmentation of brain tumors from MRI images. The authors leveraged the multi-scale feature extraction capabilities of PSPNet to effectively segment brain tumors. While the article did not provide quantitative results, it demonstrated the potential of the PSPNet model for this application. D. Seeli et al.²⁶ discussed a brain disease detection approach using the PSPNet model for multiclass segmentation. The authors claimed that the method was useful for evaluating and optimizing brain tumor segmentation models. Zhu et al.²⁷ proposed a coronary angiography image segmentation method based on the PSPNet network. Their technique reached an accuracy of 95.7%, outperforming traditional algorithms by 26.75% and the U-Net model by 4.59%. By employing transfer learning with the PSPNet model, they boosted the segmentation accuracy from 92.6% to 93.6%, increased sensitivity from 84.6% to 86.5%, and enhanced specificity from 92.1% to 94.9%. The authors concluded that the PSPNet network not only minimizes the need for manual intervention in diagnoses but also reduces reliance on medical personnel and improves overall efficiency in disease diagnosis.

Weng et al.²⁸ the automatic brain tumor segmentation approach using three models of convolutional neural network, SegNet, U-Net and PSP-Net, were compared. The proposed models were evaluated on the BraTS 2018 database. The PSP-Net model was reported to have a Dice score of 0.89 in brain tumor segmentation by the authors. They emphasized the capability of the pyramid pooling module to collect global context information, which helped the PSP-Net outperform state-of-the-art methods in the segmentation task. In summary, our results indicate the PSP-Net architecture offers an effective solution to accurate and robust brain tumor segmentation on 3D MRI scans, thus a promising model for CADx and treatment planning purposes. S. Suhara et al.²⁹ used a FCN and PSP-Net fusion approach for brain tumor segmentation. The proposed architecture was evaluated against SegNet, FCN and PSP-Net. The model ensemble reached a Dice score of 87%. Samudrala et al.³⁰ presented a model scheduled on the combination of DenseNet and PSPNet for improved segmentation of breast cancer images. The goal of this work was to exploit the connectivity patterns of DenseNet, which are known to enable deeper and dense feature propagation and to use a PSPNet-inspired pyramid pooling module, to alleviate the limitations incurred by the limited context in DenseNet and other densely connected architectures. The Dice coefficient score of the model was 0.897, suggesting the model provided an excellent segmentation for breast cancer images. Sindhura et al.³¹, a light-weight model that fuses with efficient computing operations PSPNet to improve performance with reducing the complexity is proposed. This strategy can find a tradeoff between accuracy and efficiency which may be ad-vantageous in clinical, real-time situations. The achieved Dice score (a measure to evaluate segmentation's correctness) is 0.82, which demonstrates the performance of proposed method to correctly segment the liver and tumor regions in CT images.

Li-Yin Ye et al.³² proposed a deep learning PSP-Net + VGG16 architecture on Prostate tumor MRIs. The three-dimensional prostate MRI scans are transformed into two-dimensional image slices, which are then processed using the PSP-Net neural network for training. To analyze regions of interest and differentiate between prostate cancer and normal prostate tissue, the VGG16 network is employed. This method has demonstrated high classification accuracy and recognition rates, achieving around 87%. P. Wen et al.³³ proposed an Advanced PSP-Net architecture that integrates an adaptive PSP-Network with the traditional PSP-Net architecture to enhance the segmentation performance by utilizing multiscale feature aggregation and context information. It was tested on renal ultrasound images and indicated significant improvements in segmentation accuracy, achieving a Dice score of 0.875, demonstrating the method's effectiveness in accurately delineating renal structures in ultrasound images. S. Wang et al.³⁴ proposed a DPAM-PSPNet model, where a Dual Path Attention Mechanism (DPAM) captures global information through one channel and focuses on nodal margins on another channel, before being used in conjunction with PSP-Net. The model was tested on Thyroid nodule images for segmentation and achieved a high dice score of 92.13

Outcome: The PSPNet architecture shows its versatility and superior performance in various medical imaging tasks, particularly in medical image segmentation tasks. The high Dice Similarity Coefficient received from brain tumor segmentation in Weng et al.'s research illustrate PSPNet's robust multi-scale feature extraction capabilities, which enhance segmentation accuracy and efficiency across diverse medical imaging applications.

FPN

The Feature Pyramid Network (FPN) is a type of neural network architecture designed to enhance object detection and segmentation by effectively utilizing features at different scales. Think of it like a multi-story building where each floor provides a different view of the surroundings. In an FPN, the network takes an image and processes it through several layers, each capturing different levels of detail - from fine details like textures to broader, more abstract features.



Figure 5. Architecture of FPN.

These multi-scale features are then combined, creating a rich, pyramid-like representation of the image. This helps the network recognize and locate objects of various sizes more accurately, improving performance in tasks like detecting small objects in cluttered scenes or segmenting different parts of an image.

Bai et al.³⁵ proposed a fully convolutional network (FCN) based on the Feature Pyramid Network (FPN) architecture for automated segmentation of the left ventricle in cardiac MRI images. They reported that their FPN-based model achieved a Dice similarity coefficient (DSC) of 0.92, outperforming the standard U-Net architecture which had a DSC of 0.89. The authors attributed the improved performance to the FPN's ability to effectively capture multi-scale features, which is crucial for accurate delineation of the left ventricle boundaries. Syazwany et al.³⁶ proposes a novel multimodality fusion network called MM-BiFPN for accurate segmentation of brain tumors from MRI scans. The key innovation is the use of a Bi-directional Feature Pyramid Network (Bi-FPN) to effectively fuse features from different MRI modalities (T1, T1ce, T2, FLAIR) at multiple scales. The authors report that their MM-BiFPN model achieved a Dice similarity coefficient (DSC) of 0.91 for whole tumor segmentation, 0.87 for tumor core, and 0.84 for enhancing tumor on the BraTS 2020 dataset. This outperformed other state-of-the-art methods like UNet++, nnU-Net, and TransBTS. The authors attribute the strong performance to the Bi-FPN's ability to capture cross-scale and cross-modality dependencies, as well as the effective fusion of multi-modal MRI features.

Our findings present the evidence of the promising performance for the accurate and robust segmentation of brain tumor, a critical step in treatment planning and evaluation. S. Ghosh et al.³⁷ report that the FPN models perform well for brain tumor segmentation, and it further preserved the tumor boundaries and small tumor regions with a good performance. Wang et al.³⁸ developed a RFPNet and added a residual link embedded between the bottomup and top-down components to enhance the medical image segmentation model. We also tested our proposed model on multiple medical image segmentation tasks, e.g., skin lesion segmentation, polyp segmentation, organ segmentation. RFPNet also achieved a mean Dice coefficient of 0.8421 on the ISIC 2018 skin lesion segmentation challenge, which surpassed the performance of the original FPN by 2.5%. When polyp is segmented on the ETIS-Larib Polyp DB dataset, RFPNet achieved 82.01% Dice score, which is 3.2% higher than FPN. On the Synapse multi-organ segmentation test set, an average Dice score of 0.8012 was achieved by the RFPNet, which outperformed FPN by 2.8%. The strong performance is attributed by the authors to the reorganized pyramid structure and the adaptive feature fusion, which can capture multi-scale context and preserve the details well in RFPNet. Experiments show the superiority of the adopted modifications to the FPN structure on accurate medical image segmentation. Z. Xu et al.³⁹ presented a novel enhanced backbone model, named EFPN (Enhanced Feature Pyramid Network), to detect medical images. The multi-scale feature fusion problems are solved by a top-down pyramid, scale attention modules and feature fusion attention modules in EFPN. Tested on the datasets PenD (Xray) and BraTs (MRI), the presented EFPN-based models yielded significant performance boosts.

Y. Su et al.40 introduced a Feature Augmented Pyramid Network (FAPN), designed for the task of polyp segmentation. The structure of the model consists of three modules: Cross-Embedding Module (CEM), Predictive Calibration Module (PCM), and Hierarchical Feature Fusion Module (HFFM). The CEM is comprised by a two-stages fusion method, which first performs an interaction embedding of multi-level feature, and then performs a second fusion to refine the fused feature representation. Following this, the PCM calibrates the fused feature representations using its stage-wise predicted probability maps, fine-tuned with supervision. This procedure emphasizes areas of interest while reducing other unrelated noises. Finally, HFFM adopts top-down sequential feature fusion based multi-scale feature extraction, which enhances the ability of polyp segmentation. The proposed model showed a promising Dice score of 90.2% on the Kvasir-SEG public dataset, demonstrating its effectiveness and outperforming other methods. Chiu-Han Hsiao et al.41 utilized EfficientNetB7 and FPN as encoder and segmentor for sematic segmentation of Kidney Tumor. The proposed model has attained a good performance; it reached a Dice score up to 96%. W. Huang et al.⁴² introduced the FPN along with a level-aware attention mechanism to improve the condition of capturing rich and abundant multi-scale features. This levelaware attention can adaptively enhances desirable features in different pyramid layers, which brings higher accuracy to segment meningiomas that have various sizes and appearances. The model obtained a high Dice score of 0.876, which demonstrates the model is capable to accurately segment meningiomas as compared with other existing methods. V. Allapakam et al.43 fused CT scans and PET scans of Gastrointestinal stromal tumors based on DenseNet and VGG model and used FPN to realize effective tumor detection. The FPN using EfficientNetB0 was capable of segmenting the fusion images with an IoU of 85.21%.

Outcome: Overall, the literature review highlights the potential of the FPN architecture for medical image segmentation, particularly in the domain of brain tumor detection and cardiac structure analysis. The model's ability to capture multiscale features and its state-of-the-art performance make it a promising approach for computer-aided diagnosis and image analysis⁴⁴ in the healthcare domain.

Table 1:	Literature	Review	analy	ysis
----------	------------	--------	-------	------

Author	Model	Analysis report
H Dong et	II-Net	Reached a Score of 0.86 for both High Grade and
al	0-1101	Low-Grade Gliomas
T A Tuan	U Not	Achieved Dice scores of 82% 68% and 70% on
et al	0-1101	validation data and 77% 48% and 51% on testing
ot ui.		data for the three modalities of the tumor scans.
A Kermi	U-Net	Achieved a mean Dice score of 0.783, 0.868, and
et al.	0 1101	0.805 for the three modalities of the tumor scans.
W Chen et	U-Net	Achieved Dice scores of 0.68946 0.83893 and
al.	0 100	0.78347 for the three modalities of the tumor scans.
Isensee et	U-Net	Achieved Dice scores of 0.858 for whole, 0.775 for
al.		core, and 0.647 for enhancing tumor on the test set
		for BRATS 2017.
Zhu Meng	U-Net	Achieved Dice scores of 89%, 80%, and 74% for
et al.		the Whole Tumor, Tumor Core, and Enhancing
		Tumor, respectively.
Pravitasari	U-Net	Achieved an accuracy of 96% and a Correct
et al.		Classification Ratio (CCR) of 95.69%.
Р.	U-Net	Achieved an average Dice Score of 82.75% on the
Gadosey		BraTS dataset.
et al.		
Weng et	PSP-	Achieved a Dice score of 0.89 for brain tumor
al.	Net	segmentation on the BraTS 2018 dataset,
		highlighting the effectiveness of the pyramid
~ ~ .		pooling module.
S. Suhara	PSP-	Combined FCN and PSPNet for brain tumor
et al.	Net	segmentation, achieving a high Dice score of 87%.
Samudrala	PSP-	Integrated DenseNet and PSPNet for breast cancer
et al.	Net	image segmentation, achieving a Dice coefficient
C'a dhaana	DCD	Score of 0.897.
Sindnura	PSP- Not	tumor segmentation in CT images, achieving a
et al.	INCL	Dice score of 0.82
Li-Vin Ve	DSD-	Developed a PSPNet \pm VGG16 architecture for
et al	Net	prostate tumor MRIs achieving high classification
et un	1.00	accuracy and recognition rates around 87%.
P. Wen et	PSP-	Proposed an Advanced PSPNet for renal
al.	Net	ultrasound image segmentation, achieving a Dice
		score of 0.875.
S. Wang et	PSP-	Developed a DPAM-PSPNet model for thyroid
al.	Net	nodule segmentation, achieving a high Dice score
		of
		92.13%.
Bai et al.	FPN	Proposed an FPN-based model for automated
		segmentation of the left ventricle in cardiac MRI
C	EDM	Images, achieving a DSC of 0.92.
Syazwany	FPN	representation achieving DSCs of 0.91 for whole
et al.		tumor 0.87 for tumor core and 0.84 for enhancing
		tumor.
Sourodip	FPN	Suggested FPN models achieve strong
Ghosh et		performance for brain tumor segmentation,
al.		preserving tumor boundaries and small tumor
		regions.
Wang et	FPN	Proposed RFPNet with a residual connection,
al.		achieving Dice scores of 0.8421 for skin lesion,
		0.8201 for polyp, and 0.8012 for organ
		segmentation.
Zhenghua	FPN	Proposed EFPN, achieving substantial
Xu et al.		performance improvements on PenD (X-ray) and
		Brais (MRI) datasets.
Haimei Li	FPN	Proposed a 3D Improved FPN for gastric tumor
et al.		segmentation, achieving a Dice score of 65.5%.

Yanzhou Su et al.	FPN	Proposed FAPN for polyp segmentation, achieving a Dice score of 90.2% on Kvasir-SEG dataset.
Chiu-Han Hsiao et al.	FPN	Used EfficientNetB7 and FPN for kidney tumor segmentation, achieving a Dice score of 96%.
W. Huang et al.	FPN	Integrated FPN with a level-aware attention mechanism for meningioma segmentation, achieving a Dice score of 0.876.
Venu Allapakam et al.	FPN	Used FPN with EfficientNetB0 for segmentation of fused CT and PET scans of gastrointestinal stromal tumors, achieving an IoU score of 85.21%.

This study summarizes the performance of three popular architectures – U-Net, PSP-Net (Pyramid Scene Parsing Network) and FPN (Feature Pyramid Network) on the BraTS (Brain Tumor Segmentation) data sets as showsn in table 1. U-Net based architectures also achieve strong results with 68% to 89% union of compare (Dice) scores of different sub-regions (whole, core, enhancing) of the tumor, while the PSP-Net gives competitive results (Dice of 82–92%) by using the pyramid pooling for the multi-scale context aggregation.30 FPN-based models push segmentation accuracy further, with Dice scores ranging from 84 to 96% and are even more successful at preserving the boundary of tumor and small regions through hierarchical feature fusion. Hybrid methods (eg FPN coupled with EfficientNet or residual connections) demonstrate state-ofthe-art performance (eg, 96% Dice for kidney tumors). Compared to U-Net, which is also still a baseline benchmark under different degrees of task specificity, the PSP-Net and FPN demonstrate improvements towards complex sub-regions of the tumor.

METHODOLOGY

This section provides a detailed explanation of the approach used to create the model for our needs. We will be creating a model for each segmentation architecture and comparing the results of the model on the BraTS 2024 dataset.¹⁵

The BraTS (Brain Tumor Segmentation) 2024 dataset is a multimodal magnetic resonance imaging (MRI) dataset aimed at fostering research in automated brain tumor segmentation, radiogenomic classification, and as well as overall and progression-free survival prediction. This dataset contains high quality expert annotated MRi (T1, T1Gd, T2 and FLAIR sequences) images of glimoa patients with labeled tumor subregions which include enhancing tumor, edema and necrotic core. Apart from segmentation challenges, BraTS 2024 leverages clinical and molecular data, such as MGMT promoter methylation status - a fundamental biomarker for treatment response in glioblastoma - and survival outcome of patients. The dataset has a potential to be a baseline for developing and validating machine and deep learning models in computational neuro-oncology. The dataset is made publicly available for research purpose, encouraging researchers to develop novel methods for addressing this problem and allowing research communities of medical imaging and artificial intelligence collaboration. The dataset is a pre processed dataset and ready to use data.

HARDWARE AND SOFTWARE REQUIREMENTS:

Hardware Requirements:

The model was developed and trained on Kaggle's cloud-based environment, which provides the following resources:

- CPU: Dual-core Intel Xeon Processors
- GPU: NVIDIA Tesla P100
- RAM: 28GB Storage: 100GB SSD

Software Requirements:

The software environment was based on Kaggle's standard setup with the following specifics:

– Operating System: Linux-based environment – Programming Language: Python 3.8 – Libraries and Frameworks:

- * NumPy
- * Pandas
- * Scikit-learn
- * TensorFlow
- * Keras
- * Matplotlib
- * OpenCV

DATA EXPLORATOIN:

We utilize the BraTS 2023 dataset, which includes patient MRI scans across four different modalities: Native Tumor, Tumor Core Enhanced, Tumor Whole, and Flair. The segmentation mask in this dataset categorizes areas into four classes: "Not Tumor" (purple), "Necrotic and Non-Enhancing Tumor" (green), "Fluid or Peritumoral Edema" (yellow), and "Enhancing Tumor" (blue).







Figure 7. Segmentation Mask with each colour representing different parts of the tumor.

There are three different views for the images as well, that are axial, coronal, and sagittal views.



Figure 8. Three Different Views.

DATA PREPROCESSING AND GENERATION:

Our goal is to train the models to predict how the segmentation masks are going to look like for a test image. So, we need to provide it with the correct and most valuable information we have to achieve this, and also to have more efficiency in computing the images, it is important to disregard information that may seem not as helpful.

From the four modalities of the brain tumor MRI scans, we decide to leave the native tumor as it holds the least valuable information compared to the other three modalities.



Figure 9. Different Slices of the same tumor, creating a 3D representation.

Also note that each dimension or view consists of a series of twodimensional images, called slices, all containing the same number of pixels. These slices are stacked together to form the 3D representation. But as we can see in Fig. 9 most of the slides towards the left and right ends are blank or have no valuable information. So we keep the number of slides for each modality as 75 (i.e. between slide 60 and slide 135) since most of these slides have the most information regarding the tumor.

Preparing the data:

We have to split the data into training, validation and testing sets. The training set is used to teach the model. During this phase, the model learns by adjusting its parameters to minimize the difference between its predictions and the actual segmentations. The validation set helps fine-tune the model's hyperparameters, which are preset before training and influence the model's behavior. By comparing different hyperparameter configurations, we can choose the best one for our model. Finally, the test set evaluates the model's performance on new, unseen data to assess how well it has learned. For this analysis, we've used a training set of 251 samples, a validation set of 74 samples, and a testing set of 45 samples to compare the models' performance.

Creating the data generator:

To teach a neural network how to identify and segment objects in images, you need to provide it with both the original images and the correct segmentations. This way, the neural network can learn to spot tumor patterns and make precise predictions about what's in a patient's scan.

So for each sample, we will:

Get the paths for its different types of scans: T1CE, FLAIR, and T2W, as these two offer complementary details about the brain's anatomy and tissue contrast.

Retrieve the path of the Ground truth, also known as the original segmentation.

Load the three modalities & the segmentation mask.

Create an X array that includes all the important slices (60-135) from these three scan types.

Create a Y array that holds all the chosen slices of the segmentation.

Set all the values of 4 in the mask array to 3 to address the three missing cases mentioned earlier.



Figure 10. Original Array of Segmentation Mask.



Figure 11. One-Hot encoding of Non Tumor Region.

To segment regions into different classes (0 to 3), we need to use One-Hot Encoding to transform our categorical data into a format suitable for our neural network. This is crucial because neural networks operate on numerical data. Without One-Hot Encoding, using the class labels directly would imply an incorrect hierarchy, suggesting that higher class numbers are superior to lower ones (e.g., class 1 being less than class 4). One-Hot Encoding helps us avoid this by converting class labels into binary vectors, making it clear that no class is superior to another. This simplifies the network's understanding and processing of the data.

DEFINING THE MODELS:

For defining the model, the segmentation models python library⁴⁵ is used, since it works well with Keras and Tensorflow. All of the four models will be created using the same parameters so that we can compare the performance of each model. The backbone used will be resnet 34, and the activation function will be softmax. The only difference would be seen in PSP-Net, since it takes an image size of dimensions (384,384) by default. The rest of the models will take an image size of (128,128).

U-NET:

The U-Net model creation process involves several key steps, each contributing to its ability to perform image segmentation effectively:

a. Initialize Submodules: The necessary Keras submodules (backend, layers, models, utils) are imported and initialized.

b. Backbone: A pre-trained backbone model (in our case, ResNet 34) is selected and loaded. This model, without its top classification layers, serves as the encoder, extracting high-level features from the input image. The backbone can be initialized with pre-trained weights (e.g., from ImageNet) to leverage prelearned features.

c. Skip Connections: Layers from the backbone are chosen as skip connections. These intermediate feature maps from the encoder are saved and later concatenated with the decoder layers. Skip connections help the model retain spatial information that might be lost during downsampling in the encoder.

d. Center Block: If the backbone ends with a MaxPooling layer, additional convolutional layers (center block) are added to further process the features before the upsampling begins.

e. Decoder Blocks: These blocks progressively upsample the features back to the original image size. Two types of decoder blocks can be used:

Upsampling Block: Uses UpSampling2D followed by Conv2D and optional batch normalization.

Transpose Convolution Block: Uses Conv2DTranspose followed by Conv2D and optional batch normalization.Each decoder block takes the output of the previous block and the corresponding skip connection from the encoder, concatenates them, and processes the combined features.

f. Final Convolution and Activation: The number of channels is lowered to the required number of output classes using a final Conv2D layer. The final segmentation map is generated by applying an activation function, such as the sigmoid for binary segmentation.

This setup allows the U-Net model to effectively learn the mapping from input images to their segmented outputs, making it highly useful for various image segmentation tasks.

LINK-NET:

The creation of Link-Net model is very similar to the U-Net model.

a. Backbone: Utilizes a pre-trained model to extract high-level features from input images.

b. Skip Connections: Intermediate feature maps from the backbone are preserved and later fused with decoder outputs to maintain spatial information.

c. Decoder Blocks: Employed to upsample feature maps back to the original image resolution. Two types are used:

Upsampling Block: Utilizes UpSampling2D and Convolution layers.

Transpose Convolution Block: Employs Conv2DTranspose for upsampling.

d. Final Layers: A final convolutional layer reduces feature maps to the desired number of output classes. Activation functions are applied.

PSP-Net:

The PSPNet model is designed for semantic segmentation of images using a powerful backbone for feature extraction and a Pyramid Pooling Module (PSP) for context aggregation. Key steps include:

a. Backbone: Uses a pre-trained model to extract detailed features from input images.

b. PSP Module: Constructs spatial pyramid pooling blocks at different scales (1x1, 2x2, 3x3, 6x6) to capture contextual information effectively.

c. Feature Aggregation: Integrates outputs from the backbone and PSP blocks to enhance feature representation.

d. Final Layers: Applies a convolutional layer followed by upsampling to generate the segmentation map. Activation functions (e.g., softmax for multiclass segmentation) determine pixel-wise class probabilities.

This structure makes PSPNet highly effective for tasks requiring accurate pixel-level segmentation, benefiting from both local and global context awareness.

FPN:

The FPN (Feature Pyramid Network) model is crafted for semantic segmentation by leveraging a backbone network for feature extraction and a feature pyramid for multi-scale processing. Key steps include:

a. Backbone: Uses a pre-trained model to extract hierarchical features from input images.

b. FPN Blocks: Constructs feature pyramid blocks at different stages, integrating high-resolution and low-resolution features using upsampling and skip connections.

c. Segmentation Heads: Adds segmentation heads to each FPN block, refining features with convolutional layers.

d. Aggregation: Aggregates outputs from all pyramid levels, either by summing or concatenating them, followed by additional convolution and upsampling.

e. Final Layers: Applies a final convolutional layer and activation function (e.g., softmax for multiclass segmentation) to produce the segmentation map.

This design makes FPN effective for capturing and utilizing features at multiple scales, enhancing the accuracy of image segmentation tasks.

COMPILING THE MODELS:

There are other factors that can affect how the model performs as well, such as what optimizer is being used or what the loss function is. The optimizer used will be the Adam optimizer provided by the Keras library. The Adam optimizer is an adaptive learning rate optimization algorithm created especially for deep neural network training

The loss function used will be categorical cross entropy, which, for a single sample, is given by,

 $L = -XXyi, c \log(\hat{y}i, c)$ i=1 c=1 Where:

OTHER METRICS:

The following metrics will be used to monitor the model's performance:

Accuracy: This metric shows the overall percentage of pixels that are correctly classified, whether they are positive or negative.

Dice Coefficient: This metric evaluates how much the predicted segmentation matches the ground truth. It ranges from 0, indicating no overlap, to 1, indicating a perfect overlap. Its equation can be written as,

$$\mathsf{Dice} = \frac{2|X \cap Y|}{|X| + |Y|}$$

Where: X and Y are two sets.

 $|X\cap Y|$ is the number of elements in the intersection of sets X and Y . -|X| and |Y| are the number of elements in sets X and Y, respectively.

Sensitivity: Also known as recall or true positive rate. It measures the percentage of actual positive pixels in the ground truth that were accurately predicted as positive.

Its formula can be written as,

Precision (positive predictive value): It calculates the percentage of positively predicted pixels that are positively pixelated. Its formula can be written as

$$Precision = \underline{\qquad} TP$$
$$TP + FP$$

Specificity: Often called the true negative rate, it shows the percentage of actual negative pixels in the ground truth that were correctly predicted as negative.

RESULTS AND DISCUSSIONS:

The four models were run for 20 epochs against the training dataset and evaluation using the validation dataset was also done along with it. The accuracy and Dice Coefficient after training is mentioned in the table bellow.

 Table 2: Performance metrics for different models

Mod el	Accura cy	Dice Coefficie nt	Sensitivi ty	Precisio n	Specifici ty
U- Net	0.9948	0.6055	0.9928	0.9952	0.9984
Link- Net	0.9939	0.588	0.992	0.9946	0.9982
PSP- Net	0.9965	0.7487	0.9955	0.996	0.9986
FPN	0.9908	0.5051	0.9889	0.9918	0.9972

From the bloated Accuracy values we can see that it is not a good metric to use for medical image segmentation as class 0 or the background class can be over-represented in our real and predicted images. So even though the blank pixels are being represented well, it does not mean that the segmentation mask is accurate. From the dice coefficient values in our training, we can see that Link-Net has a very close score to U-Net. The computational time taken for Link-Net was significantly lesser than U-Net, but despite that it was able to achieve a very close score. Which shows that Link-Net would be more suitable in situations where computing power is low or dataset is not as rich. The high dice coefficient of PSP-Net shows that it has outperformed its outperformed

Its competition, but with the cost of being computationally intensive. The model took way more time in training, so it would not be optimal in places where computational power is low. But it would do well in places where the accuracy of the answer is important, meaning the margin of error should be low.For predicting tumor segmentations, the models gave similar results.

Table 3	3 :	Testing	Resul	lts
---------	------------	---------	-------	-----

Model	Mean Dice Coef.
U-Net	0.42088
Link-Net	0.38227
PSP-Net	0.48382
FPN	0.34547

The PSP-Net has come out at the top again with the highest mean dice coefficient due to it being computationally intensive. Link-Net is not as close to U-Net now, but again, it does depend on the dataset. FPN however is the lower performing of the four models. But it could be improved by creating hybrid models, since the computational intensity of the model is low.

Performance Divergence: The comparative results from table 2 and table 3 show distinct performance profiles for the four models in this specific study's experimental setup.

PSP-Net achieved the highest Dice Coefficient in both the training/validation (0.7487) and testing (0.48382) phases. The sources explicitly state this indicates superior segmentation accuracy.

Link-Net achieved Dice scores lower than PSP-Net and U-Net in testing (0.38227), but had a score close to U-Net during training (0.588 vs 0.6055).

U-Net showed a Dice Coefficient of 0.6055 in training/validation and 0.42088 in testing, described as able to give "decent dice scores".

FPN had the lowest Dice Coefficient in both training/validation (0.5051) and testing (0.34547), but the sources note its computational intensity is low and it could be improved by creating hybrid models.

Trade-offs Highlighted: PSP-Net's superior accuracy came at the expense of computational intensity, taking significantly more time in training and being deemed not optimal where computational power is low. Link-Net, while not reaching PSP-Net's peak accuracy in this study's test results, is highlighted for its ability to balance high accuracy and efficient computation, making it suitable for real-time applications or resource-constrained scenarios.

They attribute PSP-Net's strong segmentation performance, in part, to its pyramid pooling module's ability to capture global context information. Link-Net is highlighted for its ability to balance high accuracy and efficient computation, making it suitable for real-time applications and scenarios with limited computational resources. U-Net is described as a network designed for semantic segmentation tasks and is able to give decent dice scores. Therefore, while the sources do argue for PSP-Net's superiority in terms of segmentation accuracy based on the Dice Coefficient results from the experiments.

CONCLUSION AND FUTURE WORK

To sum up, this study shows that several neural network architectures, including U-Net, Link-Net, PSP-Net, and FPN, are effective at segmenting brain tumors from MRI images. PSP-Net achieves the highest Dice coefficient, indicating superior segmentation accuracy at the expense of computational intensity. All models demonstrate different strengths. Link-Net is appropriate for situations with constrained resources because it strikes a compromise between computational efficiency and performance. Even though they are not as good as PSP-Net and U-Net, they still offer important insights into the trade-offs between segmentation accuracy and computational demands.

Future work should aim to create hybrid architectures that merge the efficiency of Link-Net and FPN with the high performance of PSP-Net to balance accuracy and computational demands. Enhancing the training dataset with diverse samples and applying advanced optimization techniques can further boost model performance. Additionally, exploring real-time segmentation, transfer learning, and fine-tuning pretrained models could significantly enhance their practical utility in time-sensitive medical scenarios.

Integrating segmentation models into clinical workflows and gathering feedback from medical professionals will help refine these models for real-world use. Cross-validation and robustness testing across different data subsets will ensure generalizability and prevent overfitting. Expanding the evaluation to include metrics like IoU, Precision, and Recall, and exploring various imaging modalities, will provide a comprehensive assessment of model performance and applicability across different medical imaging contexts.

REFERENCES

- A.M. Ansari, N.K. Sharma, K. Tripathi, et al. Overcoming multifaceted challenges in cancer treatment: Targeting signal transduction pathways and tumor microenvironment for enhanced therapeutic efficacy. *J. Appl. Pharm. Sci.* 2025, 15 (5), 40–63.
- R. Files, C. Cardoso, J. Prada, F. Silva, I. Pires. Syndecan-1 as a therapeutic target in squamous cell carcinoma: Current insights and future directions. *J. Appl. Pharm. Sci.* 2025, 15 (2), 37–44.
- M. Hussein, E.-S. Ali. Unlocking precision oncology: the role of neoantigen-based cancer vaccines. *Biomed. Ther. Lett.* 2025, 12 (1), 1152.
- V. V Bag, M.B. Patil, S. Nagnath Kendre. Frequent CNN based ensembling for MRI classification for Abnormal Brain Growth detection. *J. Integr. Sci. Technol.* 2024, 12 (4), 785.
- H. Dong, G. Yang, F. Liu, Y. Mo, Y. Guo. Automatic brain tumor detection and segmentation using U-net based fully convolutional networks. In *Communications in Computer and Information Science*; 2017; Vol. 723, pp 506–517.
- T.A. Tuan, N.D. Duc, H.H. Huynh. U-Net convolutional neural network with multi-kernel for brain tumor segmentation. In *Proceedings of the 2019 International Conference on System Science and Engineering (ICSSE)*, *IEEE*; pp 263–268.
- A. Kermi, I. Mahmoudi, M.T. Khadir. Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 2019; Vol. 11384 LNCS, pp 37–48.
- W. Chen, B. Liu, S. Peng, J. Sun, X. Qiao. S3D-UNET: Separable 3D U-Net for brain tumor segmentation. In *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); 2019; Vol. 11384 LNCS, pp 358–368.
- F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, K.H. Maier-Hein. Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge. In *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); 2018; Vol. 10670 LNCS, pp 287–297.
- Z. Meng, Z. Fan, Z. Zhao, F. Su. ENS-Unet: End-to-End Noise Suppression U-Net for Brain Tumor Segmentation. In *Conference proceedings: ...* Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference; 2018; Vol. 2018, pp 5886–5889.
- A.E. Pravitasari, I.K.E. Purnama, S.A. Wicaksono. VGG-16 based U-Net with Transfer Learning for Brain Tumor Segmentation. *IEEE Access* 8, 209356–209366.
- P.K. Gadosey, Y. Li, E.A. Agyekum, et al. SD-UNET: Stripping down Unet for segmentation of biomedical images on platforms with low computational budgets. *Diagnostics* 2020, 10 (2), 99.
- 13. U. Baid, S. Talbar, S. Rane. DRAG: Deep Learning Radiomics Algorithm for Gliomas. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; pp 351–364.
- 14. Q. Jin, Z. Meng, C. Sun. RA-UNet: Residual Attention U-Net for Brain Tumor Segmentation. *J. Healthc. Eng.* 1–12.
- G. Ramasamy, T. Singh, X. Yuan. Multi-Modal Semantic Segmentation Model using Encoder Based Link-Net Architecture for BraTS 2020 Challenge. *Proceedia Comput. Sci.* 2022, 218, 732–740.
- Z. Sobhaninia, S. Rezaei, N. Karimi, A. Emami, S. Samavi. Brain tumor segmentation by cascaded deep neural networks using multiple image scales. In 2020 28th Iranian Conference on Electrical Engineering, ICEE 2020; Springer, Cham, 2020; pp 77–87.
- T. Ruba, R. Tamilselvi, M. Parisa Beham, M. Gayathri. Segmentation of a Brain Tumour using Modified LinkNet Architecture from MRI Images. J. Innov. Image Process. 2023, 5 (2), 161–180.
- A. Sulaiman, V. Anand, S. Gupta, et al. An intelligent LinkNet-34 model with EfficientNetB7 encoder for semantic segmentation of brain tumor. *Sci. Rep.* 2024, 14 (1), 1345.
- P. Sameer, V.A.R. Kaushik, R. Indrakanti, S. Ch Manoj Kumar. Brain Tumor Segmentation using LinkNet. In *Proceedings - 2nd International*

Conference on Next Generation Intelligent Systems, ICNGIS 2022; Kottayam, India, 2022; pp 1–5,.

- A.T. Nair, M.L. Anitha, M.N.A. Kumar. Segmentation of Retinal Images Using Improved Segmentation Network, MesU-Net. *Int. J. online Biomed. Eng.* 2023, 19 (15), 77–91.
- T.M. Geethanjali, Minavathi, M.S. Dinesh. Semantic Segmentation of Kidney Tumors Using Variants of U-Net Architecture. *Int. J. online Biomed. Eng.* 2022, 18 (10), 143–153.
- C. Akyel, N. Arıcı. LinkNet-B7: Noise Removal and Lesion Segmentation in Images of Skin Cancer. *Mathematics* 2022, 10 (5), 736.
- M. Widiansyah, S. Rasyid, P. Wisnu, A. Wibowo. Image segmentation of skin cancer using MobileNet as an encoder and linknet as a decoder. *Journal of Physics: Conference Series*. 2021, pp 1–23.
- H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia. Pyramid scene parsing network. *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017* **2017**, 2017-Janua, 6230–6239.
- M. Zhao, J. Xin, Z. Wang, X. Wang, Z. Wang. Interpretable Model Based on Pyramid Scene Parsing Features for Brain Tumor MRI Image Segmentation. *Comput. Math. Methods Med.* 2022, 2022.
- D.J.J. Seeli, K.K. Thanammal. Brain Disease Detection using PSPNet for Multiclass Segmentation. In *Proceedings of the 2nd International Conference on Applied Artificial Intelligence and Computing, ICAAIC* 2023; Salem, India, 2023; pp 1030–1034.
- X. Zhu, Z. Cheng, S. Wang, X. Chen, G. Lu. Coronary angiography image segmentation based on PSPNet. *Comput. Methods Programs Biomed.* 2021, 200, 0169–2607,.
- Y.T. Weng, H.W. Chan, T.Y. Huang. Automatic segmentation of brain tumor from 3D MR images using SegNet, U-Net, and PSP-Net. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 2020, 11993 LNCS, 226–233.
- S.F. Suhara*, D.M.S. Mary. Fully Connected Pyramid Pooling Network (FCPPN) – A Method For Brain Tumor Segmentation. *Int. J. Eng. Adv. Technol.* 2019, 9 (1), 7036–7041.
- S. Samudrala, C.K. Mohan. Semantic segmentation of breast cancer images using DenseNet with proposed PSPNet. *Multimed. Tools Appl.* 2024, 83 (15), 46037–46063.
- C. Sindhura, H.V.R. Vengama, S. Gorthi. Efficient-PSP-Net: A Lightweight Model for the Segmentation of Liver and Tumor in CT Scans. In ACM International Conference Proceeding Series; 2022; pp 1–6.
- L.Y. Ye, X.Y. Miao, W.S. Cai, W.J. Xu. Medical image diagnosis of prostate tumor based on PSP-Net+VGG16 deep learning network. *Comput. Methods Programs Biomed.* 2022, 221, 106770.

- P. Wen, Y. Guan, J. Li, et al. A-PSPNet: A novel segmentation method of renal ultrasound image. In *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*; Melbourne, Australia, 2021; pp 40–45.
- 34. S. Wang, Z. Li, L. Liao, et al. DPAM-PSPNet: ultrasonic image segmentation of thyroid nodule based on dual-path attention mechanism. *Phys. Med. Biol.* 2023, 68 (16), 165002.
- W. Bai, M. Sinclair, G. Tarroni, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J. Cardiovasc. Magn. Reson.* 2018, 20 (1), 65.
- N.S. Syazwany, J.H. Nam, S.C. Lee. MM-BiFPN: Multi-Modality Fusion Network with Bi-FPN for MRI Brain Tumor Segmentation. In *IEEE* Access; 2021; Vol. 9, pp 160708–160720.
- S. Ghosh, K.C. Santosh. Tumor segmentation in brain MRI: U-Nets versus feature pyramid network. In *Proceedings - IEEE Symposium on Computer-Based Medical Systems*; Aveiro, Portugal, 2021; Vol. 2021-June, pp 31–36.
- Z. Wang, J. Zhu, S. Fu, S. Mao, Y. Ye. RFPNet: Reorganizing feature pyramid networks for medical image segmentation. *Computers in Biology* and Medicine. 2023.
- Z. Xu, X. Zhang, H. Zhang, et al. EFPN: Effective medical image detection using feature pyramid fusion enhancement. *Comput. Biol. Med.* 2023, 163, 107149.
- Y. Su, J. Cheng, M. Yi, H. Liu. FAPN: Feature Augmented Pyramid Network for polyp segmentation. *Biomed. Signal Process. Control* 2022, 78, 103903.
- C.H. Hsiao, P.C. Lin, L.A. Chung, et al. A deep learning-based precision and automatic kidney segmentation system using efficient feature pyramid networks in computed tomography images. *Comput. Methods Programs Biomed.* 2022, 221, 106854.
- W. Huang, X. Shu, Z. Wang, et al. Feature Pyramid Network With Level-Aware Attention for Meningioma Segmentation. *IEEE Trans. Emerg. Top. Comput. Intell.* 2022, 6 (5), 1201–1210.
- V. Allapakam, Y. Karuna. A hybrid feature pyramid network and Efficient Net-B0-based GIST detection and segmentation from fused CT-PET image. *Soft Comput.* 2023, 27 (16), 11877–11893.
- 44. K. Inthanon, R. Sririwichitchai, N.M. Davies. Rapid screening of adipogenic responses to EGCG in non-induced L929 adipocytes using image-based analysis compared to the conventional 3T3-L1 method. J. Appl. Pharm. Sci. 2025, 15 (5), 138–148.
- A.P. Apte, A. Iyer, M. Thor, et al. Library of deep-learning image segmentation and outcomes model-implementations. *Phys. Medica* 2020, 73, 190–196.