International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

Unet 3+ For Brain Tumor Segmentation : A Study

Asna P K[1]

[1]PG Student, Department of CSIS, Rajagiri Scool Of Engineering & Technology, Kerala, India

---

Abstract - *Deep learning based image segmentation has varying applications in this modern era. Automatically segmenting the Brain Tumor is an important one among them. Manual segmentation of brain tumor is a very difficult task in the field of medical image processing. Moreover it is a time consuming process and chances for getting false results is high. Hence an accurate method for automatic brain tumor segmentation is necessary for the clicians to obtain a correct result and hence for the proper analysis , monitoring of the deseases. Unet is a CNN based architecture which is widely used for image segmentation process. Unet has its different modifications such as unet++, unet3+ . Unet 3+ having full-scale skip connections is capable of extracting more features than other unet varients. Unet 3+ model provides a better accuracy and other metrics values compared to basic unet architecture in brain tumor segmentation.*

***Key Words*: Unet, Unet++, Unet3+**

1.INTRODUCTION

The quick increase in deep learning algorithm development suggests that deep neural networks have excellent potential for use in computer-aided automated or semi-automatic procedures for clinical fact processing. Convolutional neural networks quickly improved, leading to models that could match or even outperform human level performance in several applications, including microscopic image segmentation. Deep learning knowledge-based semantic segmentation has recently attracted a lot of attention. Medical picture segmentation frequently use UNet, a deep learning network with an encoder-decoder architecture. One of the crucial elements for accurate segmentation is the fusion of many scale variables. UNet++ was updated to resemble Unet by creating a structure with dense and stacked skip connections.

However, it doesn't look at many records from broad scales, and there may still be a tonne of room for improvement. The full-scale skip connections advantage is used by the UNet 3+ paradigm. The full-scale skip connections use feature maps at various scales that provide low-stage information with high-degree semantics. In addition to boosting accuracy, the UNet 3+ can decrease network parameters to increase processing performance. The changing shape and appearance of brain tumours in multi-modal magnetic resonance imaging makes accurate segmentation of these lesions a challenging medical image analysis task. Such types of brain tumours must be manually segmented, which takes a great lot of clinical skill and time and increases the possibility of human error. Additionally, the manual method lacks consistency and reproducibility, which has a detrimental impact on the outcomes and may, in the long run, lead to inaccurate analysis and treatment. A well-designed U-Net based architecture may be able to segment brain tumours effectively. So the automatic segmentation of brain tumours can employ the Unet 3+ model. Brain tumor segmentation is an essential step in medical image segmentation.Automatic brain tumor segmentation will gives an accurate result and will helps the clicians for proper diagnosis.Moreover it saves the time and will reduces the chances for getting errors.

Unet is a CNN based architecture having encoder and decoder path which is widely used for image segmentation process. Unet has its different modifications such as unet++, unet3+ . UNet++ does not sufficiently examine the information from complete scales when using layered and dense skip connections. UNet 3+ utilises full-scale skip connections to combine high-level semantics with low-level semantics from feature maps at various scales. The suggested Unet3+ technique produces significantly better segmentation results than existing models and aids in accuracy improvements. An encoder-decoder structure called Unet 3+ has connections between each decoder layer as well as between the encoder and decoder levels. These full-scale skip connections enable them to extract more characteristics and produce more precisely segmented results.

2. RELATED WORKS

In order to make better use of the existing annotated examples, Ronneberger, Olaf, Philipp Fischer, and Thomas Brox suggested a project that presents a network and training technique that heavily relies on data augmentation. The design consists of a decoding path for accurate localisation and an encoding path for capturing context. Network speed is quick. On a modern GPU, segmenting an image may take less than a second. This approach is built upon the so-called "fully convolutional network," a more elegant architecture. This architecture is updated and extended so that it can function with a small number of training photos and produce segmentations that are more accurate. The fundamental concept is to use upsampling operators to replace pooling operators. Thus, these layers raise the output's quality. [1].

The dental panoramic image segmentation using unet architecture was discussed by S. Sivagami, P. Chitra, G. S. R. Kailash, and S. R. Muralidharan. In order to accurately segment dental panoramic x-ray images, this method suggests a UNet architecture that makes use of convolutional neural networks. This allows the dentist to identify any impacted teeth, locate the precise location for placing dental implants, and ascertain the direction of the teeth's structure. This method recommends a UNet architecture that uses convolutional neural networks to accurately segment dental panoramic x-ray images. Medical experts can identify and diagnose disease more precisely with the help of dental radiography pictures. Images obtained by radiographic procedures like X-rays, CT scans, and MRIs. X-ray pictures are frequently complex in nature. It is challenging to accurately distinguish between the various dental portions when there is noise. This makes segmentation a very difficult task. Dental image segmentation helps the dentist locate the best spot for dental implants, spot impacted teeth, and determine the direction of the teeth's structure. A more recent method is the segmentation of medically utilised images using the UNet architectural model. [2].

Attention Unet++ approach for liver CT image segmentation was discussed by C. Li et al. In order to allow the features recovered at various levels to be combined with a task-related selection, this research suggested a layered attention-aware segmentation network called Attention UNet++ that integrates attention mechanisms between nested convolutional blocks. Here, the hierarchical Attention UNet++ attention-aware segmentation network was suggested. Liver cancer is one of the cancers with the highest death rates. It is imperative to develop an automatic liver segmentation model since manual segmentation takes a lot of time and is prone to mistakes, making it challenging for medical professionals to identify and treat liver lesions. The recommended method includes a deep supervised encoder-decoder architecture and an improved dense skip connection. The attention method introduced by Attention UNet++ between nested convolutional blocks enables the combination of a task-related selection with the features extracted at different levels. Furthermore, the addition of deep supervision increases the prediction speed of the pruned network at the cost of a modest performance impact. The segmentation of the liver by UNet++ was quite successful. [3].

The article "UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation" was written by H. Huang et al. UNet 3+ utilises deep supervisions and full-scale skip connections in this project. Full-scale skip connections combine high-level feature maps with low-level features at various scales. From the feature maps, the deep supervision develops hierarchical representations. Organs that occur at various scales can benefit from the suggested methodology. The suggested UNet 3+ can decrease network parameters to increase computation efficiency in addition to improving accuracy. Furthermore, it develops a classification-guided module and a hybrid loss function to improve the organ boundary and lessen over-segmentation in a non-organ image, producing better segmentation outcomes. [4].

J The authors of the research, "Optimized U-Net for Brain Tumor Segmentation," are Micha Futrega, Alexandre Milesi, Michal Marcinkiewicz, and Pablo Ribalta. This study suggested an improved U-Net architecture for segmenting brain tumours. Run a thorough ablation research to test: deep supervision loss, Focal loss, decoder attention, drop block, and residual connections in order to determine the best model architecture and learning schedule. Additionally, the ideal UNet encoder depth, convolutional channel count, and post-processing method were looked for. Automatic brain tumour segmentation is one of the most difficult issues in medical image processing. A more accurate, dependable, and uniform method for disease identification, treatment planning, and monitoring would be made possible by developing a computational model that is capable of performing better than a trained human. Deep learning-based semantic segmentation has recently attracted increasing attention. Medical picture segmentation frequently use UNet, a deep learning network with an encoder-decoder architecture. [5].

It is essential for the diagnosis, follow-up, and therapy planning of the condition to automatically segment brain tumours using 3D MRIs. Manual delineation techniques are expensive, time-consuming, labor-intensive, and subject to human mistake. Myronenko.A present a semantic segmentation network based on encoder-decoders for the segmentation of tumour subregions from 3D MRIs.Due to a small training dataset size, a variational auto-encoder branch is added to reconstruct the input image itself in order to regularise the shared decoder and impose additional restrictions on its layers. The segmentation strategy is based on an encoder-decoder CNN architecture, with an asymmetrically larger encoder to extract image features and a smaller decoder to reconstruct the segmentation mask . An additional branch was introduced to the encoder endpoint in a manner reminiscent of autoencoder architecture in order to reconstruct the original image. [6]

A unique two-stage cascaded U-Net is presented to segment brain tumour substructures from coarse to fine. The network is trained end-to-end using the Multimodal Brain Tumor Segmentation Challenge training dataset.Experimental findings on the testing set show that, for the enhancing tumour, whole tumour, and tumour core, respectively, With more than 70 teams competing, the strategy took first position in the segmentation task of the BraTS 2019 challenge. Multi-modal magnetic resonance images are sent to the first stage U-Net, which approximates a segmentation map. The second stage U-net is fed the raw images and the coarse segmentation map combined. A segmentation map with more network parameters may be provided in the second stage. A two-stage cascaded network is trained from beginning to end.[7]

The field of deep learning is encouraging advances in biomedical imaging, which is a key element of medical care and a driver of scientific advancement. While many applications use semantic segmentation techniques to enable picture analysis and quantification, designing the corresponding customised solutions is difficult and heavily dependent on the characteristics of the dataset and the available technology. Isensee, F., Ja¨ger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H created the deep learning-based segmentation technique nnU-Net, which adapts automatically to any new task, preprocessing, network architecture, training, and post-processing. A set of fixed parameters, interdependent rules, and empirical decisions are used to model the process's major design decisions. nnU-Net outperforms the majority of current techniques without the need for manual intervention, including highly specialised solutions on 23 open datasets used in international biomedical segmentation competitions.[8]

Convolutional neural networks (CNNs) have recently been used to solve issues in the fields of medical image analysis and computer vision. Despite their widespread use, most methods can only process 2D images, but the majority of medical data utilised in clinical practise is made up of 3D volumes. Fausto Milletari, Nassir Navab, Seyed-Ahmad Ahmadi proposed a volumetric, fully convolutional neural network-based method for 3D picture segmentation. CNN learns to predict segmentation for the entire volume at once after being trained end-to-end on MRI data showing prostate.Provide a novel goal function based on the Dice coefficient that we optimise during training. In this method, can address circumstances in which the ratio of foreground to background voxels is significantly out of balance.Add more data using random non-linear transformations and histogram matching to make up for the small number of annotated volumes that are available for training.[9]

The U-Net3+ segmentation network proposed in this study by Chuanbo Qin, Yujie Wu, Wenbin Liao, Junying Zeng1, Shufen Liang, and Xiaozhi Zhang is improved. It is based on stage residual. The encoder section of the network uses an encoder based on the stage residual structure to overcome the vanishing gradient problem brought on by increasing network depth and improves the encoder's feature extraction capabilities, which are essential for full feature fusion when the network is up-sampled. In order to remove the impact of batch size on the network, filter response normalisation (FRN) layer was substituted for batch normalisation (BN) layer. The IResUnet3+ three-dimensional (3D) model is built using the improved U-Net3+ two-dimensional (2D) model with stage residual.[10]

3. UNET WORKING

A symmetric U-shape defines the U-Net architecture, which is composed of two components, the encoder and the decoder. The input volume is converted into a lower-dimensional space by the contracting path (encoder), which is the first component. The modular structure of the encoder is made up of convolution blocks that repeat. Two smaller blocks of transformations make up each block. The input feature map's spatial dimensions are first cut in half by using a convolutional layer with kernels of 3x3x3 and a stride of 2x2x2, followed by instance normalisation and Leaky ReLU activation with a negative slope if 0.01 is used. With the exception of the convolutional layer having a stride of 1x1x1, the subsequent feature map is modified using nearly the same set of procedures.

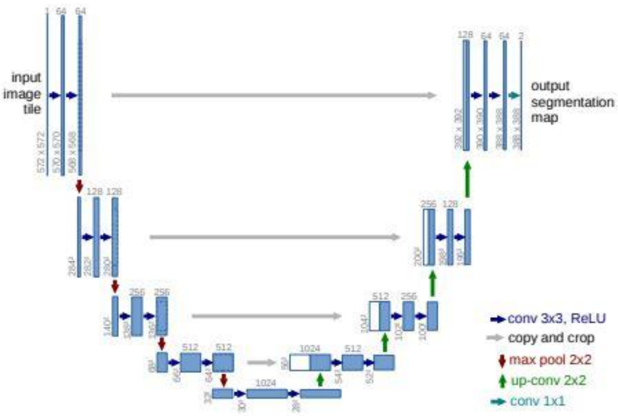

Fig 1.1: Architecture Of Unet

The decoder portion begins after the feature map's spatial dimensions are changed to measure 2x2x2. Although the decoder also has a modular design, its objective is to decrease the encoder feature map in order to improve the spatial dimensions. The decoder's block is composed of three smaller blocks. The first one doubles the spatial dimensions of the feature map by transposed convolution with kernels and strides of 2x2x2. A convolutional layer with kernels 3x3x3 and stride 1x1x1, instance normalisation, and Leaky ReLU activation with a negative slope if 0.01 are then applied to the upsampled feature map and encoder feature map from the comparable spatial level.

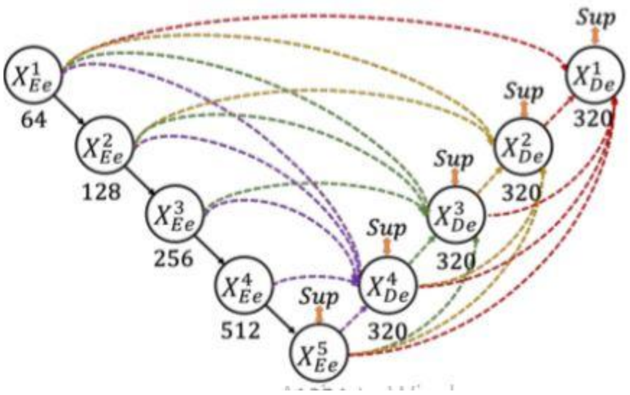

Fig1.2: Architecture Of Unet 3+

Encoder consists of five blocks.Each blocks are connected each other and pooling operations are applied.Also convolutional blocks are applied with corresponding filters.The fifth block is the bottle neck layer.Decoder part also consists of five blocks.Each layers are concatenated with the previous layers along with encoding layers.Batch normalization and relu is applied to each layers except the last layer. The interconnection between the encoder and decoder as well as the intraconnection between the decoder sub-networks are both converted by the suggested full-scale skip connections. Both UNet with plain connections and UNet++ with nested and dense connections fall short of clearly learning the place and border of an organ by not exploring enough data from full scales. Each decoder layer in UNet 3+ integrates smaller- and same-scale feature maps from the encoder and larger-scale feature maps from the decoder, which capture fine-grained details and coarse-grained semantics in complete sizes, to address the flaw in UNet and UNet++.

The feature map from the same-scale encoder layers is received directly in the decoder, just like the UNet. As opposed to the UNet, a network of inter-encoder-decoder skip connections transmits low-level detailed information from the smaller-scale encoder layer using non-overlapping max pooling, while a network of intra-decoder skip connections sends high-level semantic information from the larger-scale decoder layer using bilinear interpolation. We need to further unify the number of channels and cut down on the unnecessary data now that we have five feature maps of the same quality. We had the idea that the convolution with 64 filters of size 3 3 might be a good option. to smoothly combine deep semantic knowledge with shallow exquisite information.Apply a feature aggregation process that consists of 320 filters of size 3 3, batch normalisation, and a ReLU activation function to the concatenated feature map from five scales.

CONCLUSION

U-Net 3+ model experimentation is done to choose the best model architecture. A good dice score is provided by unet3+ architecture, which increases segmentation accuracy. Complete avoidance of oversegmentation in complex backgrounds is a UNet 3+ skill. Comparing this strategy to all others, it is superior. Comparing Unet3+ to Unet's basic design, Unet3+ improved +'s version, which has intra and inter connections between encoder and decoder, is the best method for segmenting brain tumours.In order to make the most of feature maps at full scales for precise segmentation and effective network architecture with fewer parameters, UNet 3+ is utilised, which is a full-scale linked UNet. In settings with little data, it is a very effective segmentation method that performs admirably in a variety of biomedical segmentation applications. It is anticipated that UNet 3+ will outperform all prior state-of-the-art techniques and produce accurate segmentation results.

REFERENCES

[1] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional networks forbiomedical image segmentation. arXiv 2015." arXiv preprint arXiv:1505.04597 (2019).

[2]S. Sivagami, P. Chitra, G. S. R. Kailash and S. R. Muralidharan, "UNet Architecture Based Dental Panoramic Image Segmentation," 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 2020, pp. 187-191.

[3] Li et al., "Attention Unet++: A Nested Attention-Aware U-Net for Liver CT Image Segmentation," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 345-349.

[4]H. Huang et al., "UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 1055-1059

[5]Michał Futrega, Alexandre Milesi, Michal Marcinkiewicz, Pablo Ribalta "Optimized UNet for Brain Tumor Segmentation" arXiv:2110.03352 [eess.IV],2021]

[6] Myronenko, A.: 3D MRI Brain Tumor Segmentation Using Autoencoder Regularization. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. pp. 311–320. Springer International Publishing, Cham (2019)

[7] Jiang, Z., Ding, C., Liu, M., Tao, D.: Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In: Crimi, A., Bakas, S. (eds.) Brainlesion:

Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. pp. 231–241. Springer International Publishing, Cham (2020)

[8] Isensee, F., Ja¨ger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. pp. 1–9. Nature Methods (2020)

[9] Fausto Milletari, Nassir Navab, Seyed-Ahmad Ahmadi: V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. International Conference on 3D Vision (3DV) (2016)

[10] Chuanbo Qin1, Yujie Wu1, Wenbin Liao1,2, Junying Zeng1*, Shufen Liang1 and Xiaozhi Zhang: Improved U-Net3+ with stage residual for brain tumor segmentation. Qin et al. BMC Medical Imaging (2022)