International Journal of Machine Learning and Image Processing (IJMLIP) ISSN online

Vol. 1, Issue 2, July 2023 2583-908X

U-Net for Segmenting Brain Tumors in 3D MRI: Application of CNN in medical imaging

S.B. Mohan, Department of Electronics and Communication Engineering, S.A. Engineering College,

Chennai, Tamil Nadu, India. Email: drsbmohan@gmail.com

Abstract:

Brain tumors are very dangerous and can kill patients if they are not found in time. Finding them is very important for their survival. The most accurate way to find tumors is with Magnetic Resonance Imaging (MRI), which can clearly show that they are there on the video. But it can also lead to less accurate reviews when an expert looks at the pictures by hand. This is mostly because they are tired, don't know what they're doing, or there isn't enough information in the picture. This can happen if the tumor isn't big enough to show up on the pictures or if it has overlapping parts of the brain that make it hard for the specialist to find the right one because they are mistaken for healthy brain areas. In order to improve the accuracy of diagnoses, this study will suggest a segmentation method to help doctors find brain tumors. On the MICCAI BraTS'20 standard dataset, this method can accurately separate and group brain tumors with 98.81% pixel-level and 98.93% classification accuracy. A comparison with past studies shows that the offered way works the best.

Keywords: Brain Tumors, MRI, Deep Learning,U-Net

1.Introduction

Brain tumors are growths that happen when brain cells multiply in a way that isn't normal and the brain's control systems stop working. Brain tumors affect about 700,000 people around the world, and 86,000 new cases were found in 2019. Finding brain tumors early is very important for diagnosing cancer because it helps doctors choose the best treatment plan to save the patient's life.

On the other hand, it's hard to find and separate tumors using images and human analysis because tumors can have both low-grade and high-grade features. This also takes a lot of time and doesn't always give accurate results [2]. A computer-aided (CAD) fully automatic diagnostic method that doesn't involve any harm could help doctors make earlier diagnoses, analyze data, and plan treatments. It could also lower the death rate by making tumor identification more consistent, faster, and more accurate [3]. Because of this, several CAD-based picture segmentation methods are being looked into right now for this purpose [4]. Because CNN has grown, computer vision and different deep learning-based brain tumor segmentation methods have also come a long way [5].

We are focusing on tumor segmentation in this study, which is thought to be one of the hardest things to do with mixed MRI scans. This study used a set of files called BraTS 2020 from the Brain Tumor Segmentation Challenge 2020. For each patient, the collection includes four different types of MRI scans, along with a carefully divided area of interest (ROI). These are the important things that this study adds to the field:

- A convolutional network, also known as the U-net model, is made to split images quickly and accurately.
- A new way of building the model architecture is shown instead of using fixed hyper-parameters. In this way, the model is improved through an ablation study in which different hyper-parameters are changed in a number of experiments.
- During the pre-processing step, the middle slice of the 3D MRI is automatically removed, and a levelling method is used, which has an effect on the total performance.
- Three of the four patterns are used to train and test the model individually to see which one gives the best results.
- The model can achieve ideal performance across a range of hyper-parameters, which further proves that the design is stable and consistent in its performance.

2. Related Works

In the past few years, many new and different ways of automatically separating brain tumors have been shown. Pereira et al. [6] showed how to use a CNN model to handle the segmentation process. Our method is meant to get the best results with the least amount of work. We did not use 3D data to train the model; instead, we used a single slice of a 3D MRI.Other authors [7] shown a computer system that, with the use of MRI pictures, can differentiate between a normal brain and an abnormal brain. Their system did a good job of finding HGGs and LGGs, with scores of 99% for accuracy, 98.03% for precision, and 100% for sensitivity. Noori et al. [8] made a 2D U-Net, a low-parameter network that works in two different ways. They used the BraTS 2018 validation data set and got die scores of 89.5 for the whole tumor (WT), 81.3 for the enhancing tumor (ET), and 82.3% for the tumor core (TC). With the help of a 3D masked U-net multi-scale, Xu et al. [9] made it easier to tell tumors apart. In order to get features from this network, you may feed it multi-scale building photos and then add a 3D atrous spatial pyramid pooling layer (ASPP). On the BraTS 2018 proof set, they achieved a total score of 80.94% for ET, 90.34% for WT, and 83.19% for TC. This technique received ratings of 76.90% for ET, 87.11 percent for WT, and 77.92 percent for TC on the BraTS 2018 test set. There were two research [10] that tested a U-Net model using the BraTS dataset and found that it performed well. Both of these studies employed the BraTS dataset.When you look at the data from the first and third sets side by side, you can see that they were 66.96% and 62.22% accurate at finding HGG cases. For LGG, the second set was 63.15% accurate and the third set was 62.28% accurate. They came up with a model that is a mix of a U-Net and a VGG16 [11].

The algorithm was successful in locating the region containing the tumor. Within the context of this work [12], the segmentation algorithms Gamma, U-Net, SegNet3, Res-SegNet, and U-SegNet were all used using BraTS files. It was observed that Res-SegNet, U-SegNet, and Seg-UNet were each right 93.3% of the time, 91.6% of the time, and 93.1% of the time, respectively. Using feature-based fusion approaches may assist figure out how to better separate tumor lesions into three regions, as stated by Pei et al. [13]. Although the strategy was successful in improving the true separation of WT and ET, the improvement was not very significant (DSCWT = 0.85 0.055, DSCET = 0.837 0.074). The use of a 3D deep learning system known as context deep-supervised U-Net is suggested as a method for separating brain cancers from 3D MRI images in a study conducted by Lin et al. [14]. It received a DSC of 0.92 for WT, 0.89 for TC, and 0.846 for ET when using the approach that was proposed. In their paper, Punn et al. [15] discussed a 3D model called the U-net that may be used in conjunction with 3D MRI to classify brain tumors into one of three categories: WT, TC, and ET. A DSC value of 0.92 on WT, 0.91 on CT, and 0.84 on ET showed that the suggested deep learning model worked the best. Getting the data to a normal state came first in both of these tests. They came up with the idea of a 3D U-Net model that could automatically tell brain tumors apart from other parts of 3D MRI data [16].

DATASET

The BraTS2020 dataset, which we got from Kaggle [17], is used to train and test our system. There are a total of 473 3D pictures of people in the collection. Four types of MRI scans are done on each patient: FLAIR, T1ce, T1, T2, and the related ROI (seg). The experts named the ground facts that were given. Each 3D volume has 155 2D slices, or pictures, of brain MRIs taken in different parts of the brain

3. Research Methodology

3.1. Data processing

Brats 2021 was the set that was used. First, the MRI pictures were made the same size, which was 160 x 192 x 192 pixels. This size was found through testing, and the comparison was based on Dice Score, which is talked about more in the data. In Figure 1, you can see some MRI scans from two different files. Using trimming and scaling, the pictures are made to the same size. For images that are smaller than the specified size, padding is used. For images that are bigger than the specified size, cropping is used. Since there aren't many data sets, augmentations are very important. To make things even better, different types of augmentations are used at random. It is used with the following additions:

- Flip the picture
- Change the brightness
- Changes that are flexible
- Changes in strength

It's important to note that the augmentations used are chosen at random from this set. This makes the model resistant to overfitting. By picking the augmentations at random based on a

level, the following can be done. The data was split into 5 folds using K-Folds. Fold 1 was used to test how well the model worked, and the other folds were used to train it. Table 1 shows how the data used was spread out. Fold 1 was picked because it is stable in terms of metrics. Many people noticed that the results they got on Fold 1 went up in a pretty smooth way. There are a good number of both busy and clean data examples in this collection.

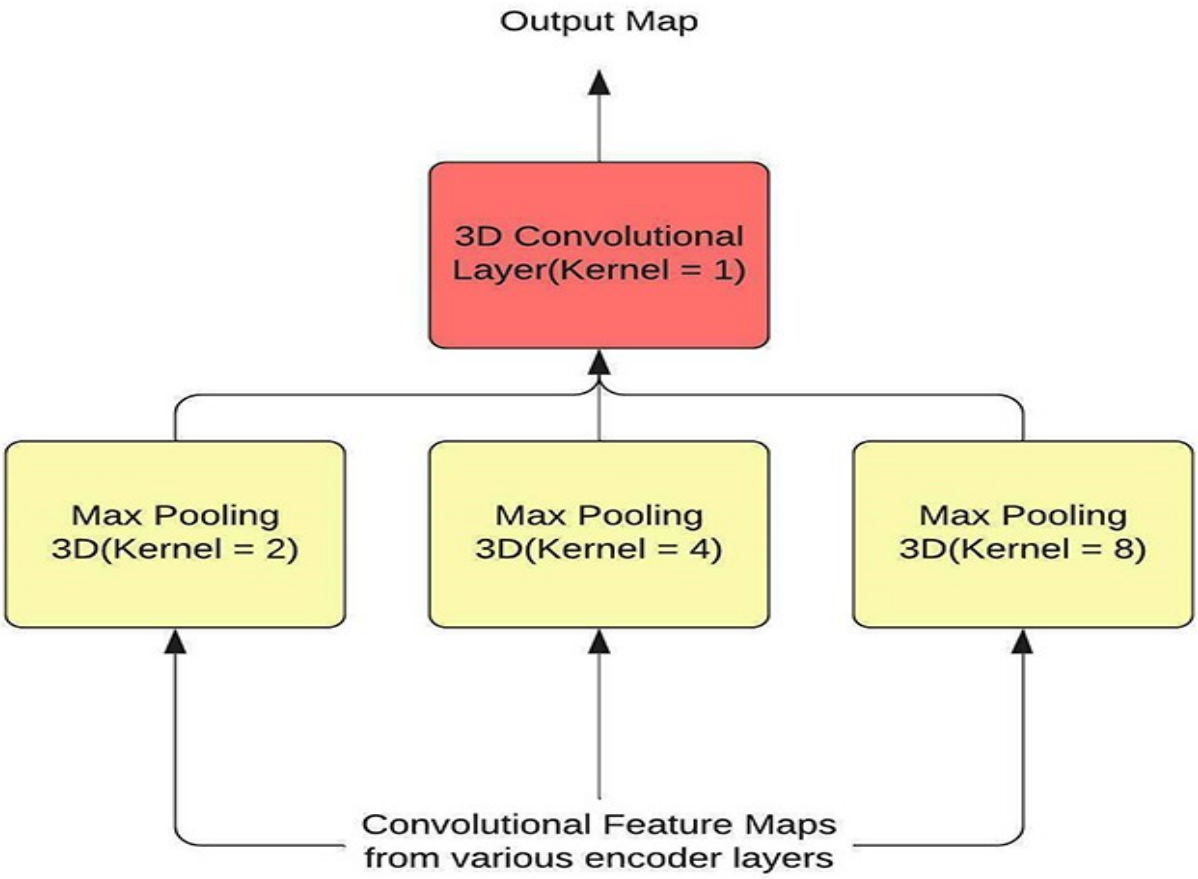
Table 1. To separate the data for brats in 2021.

Data Split	
Split Name	Number Of Samples
Train Split	1000
Validation Split	250

3.2. Three-dimensional U-Net model with residual spatial pyramid pooling

SPP is where the idea for Atrous Spatial Pyramid Pooling came from. It builds on the idea of SPP with parallel Atrous Convolutional layers.

Figure 2. The schematic representation of SPP's architecture.



Attention is the process by which people concentrate on doing certain things. When we read, we understand what's going on by looking at the words around us in a line. This information-gathering system operates at the attention layer, and its primary responsibility is to perform that function. Deep learning has made significant use of the attention layer, which has contributed to the development of cutting-edge outcomes. The model is brought into focus by the combination of the outputs from two different encoder settings. The outcome of the two encoder layers is then passed on to two distinct 3D convolutional layers, which perform the operations listed below. After adding the outputs of the two layers together, the relu variable is then utilized to turn it on. The process begins with an active output, which is then sent via a 3D convolutional layer. After that, it is readjusted, and finally switched on. When you combine the result of these layers with activation, you can keep the context without losing any of the dimensions.

This is done to keep the idea of depth. Since the SPP layer gets its information from different encoder layers, the output's size changes a lot when sharing is used.

In this study, tests are done on three model designs shown in Figure 3:

• No SPP: The SPP blocks would not be included, which would result in an improvement to the context of the model.

• 1 SPP: The top SPP block would be eliminated from the design, leaving just the bottom SPP block in its place. This would be the sole block that remained. Therefore, it is a sound strategy to enhance the model by making use of both the context and the attributes.

• 2 SPP: We used both of the available SPP blocks. This kind has a greater number of characteristics than the other two categories combined.

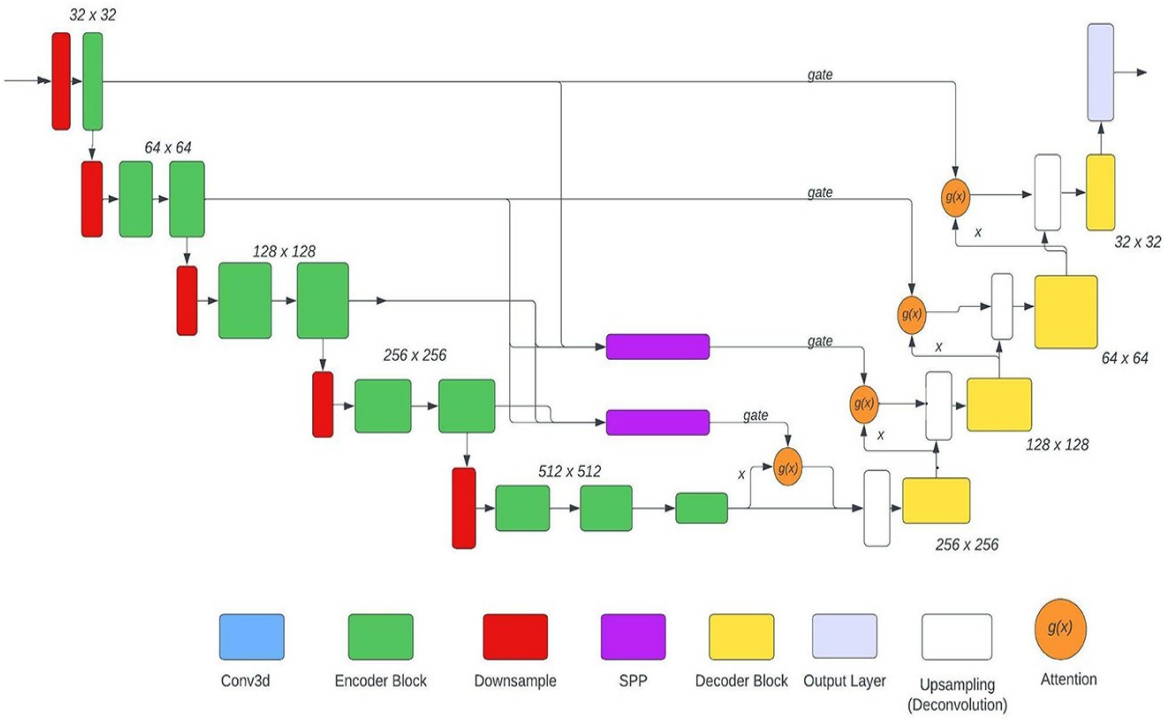


Figure 3. The architecture that was utilized for U-Net

3.2.1. The process of training

The picture size format is channel * length * width. Some of the numbers we tried are 160 × 160 × 160, 128 × 160 × 160, and 160 × 192 × 192. Equation (1) shows how this works. gi stands for the model forecast. The main reason this function was chosen was to keep the loss function's attention off of fields in the background.

$$Dice\ Loss = \frac{2^* \sum_i^N p_i g_i}{p_i^2 g_i^2} \quad \text{---------- (1)}$$

We found the problem of gradient growth through our experiments, which is why group normalization was used. We couldn't use a large batch size for training, so batch normalization didn't work for us. Because of this, batch size 1 was chosen. For the training of the models, an Nvidia Tesla V100 GPU was used.

Table 2. Training hyperparameters

Hyperparameters used	
Hyperparameter	Value
Image size (channels * length * width)	160 × 192 × 192
Epochs	60
Learning rate	2.50E-04
Weight decay	1.00E-07
Scheduler	Cosine annealing LR
Criterion	Dice loss
Optimizer	Adam
Normalization	Group norm

3.2.2. Training procedure

Cross-validation was used to test the model, and it was tested on two different criteria:

The Dice Score, which can be seen in Equation 2, is the F1-Score given for each picture pixel. The marked dots are the ground truth in this case. The dice score is a good measure because it punishes false positives: This map is used in the denominator instead of the numerator if it has a lot of fake results.

$$Dice\ Score = \frac{2 * Region\ of\ Overlap}{Region\ of\ Union}$$

--------(2)

• **Hausdorff Distance:**This value (31), which indicates the degree to which each point in a model set resembles a point in an image set, also indicates the degree to which the reverse is true. So, this closeness can be used to figure out how similar two things are that are on top of each other.It's important to note that Hausdorff Distance doesn't care about how big the background is. The Hausdorff distance works with the Dice measure to find the extensive distance between the two shapes' ends. Since it punishes outliers harshly, a forecast may show almost voxel-perfect agreement. Even though it is messier than the Dice index, this number is still a good way to figure out how important segmentation is in the clinical setting.

4. Results and Discussions

This model is compared against two models taken from the Brats 2021 dataset, as well as one model taken from the Brats 2020 dataset, and one model taken from the Brats 2019 dataset.

• Out of all the other designs that were proposed, the model consisting of a single SPP block performed the best. As a result of this, we can observe that the improvement of features and context work hand in hand with one another.

• The Enhancing Tumor class results in all three models having the smallest Hausdorff Distance.

Table 3. Results accomplished by maintaining the status quo.

Model name	Metrics							
	WT, whole tumor; TC, tumor core; ET, enhancing tumor (sub regions of tumor affected brain)							
	Hausdorff distance				Dice			
	WT	TC	ET	Average	WT	TC	ET	Average
No SPP	13.070	11.010	10.210	11.430	0.908	0.877	0.838	0.870
1 SPP	9.430	7.780	6.300	7.840	0.899	0.899	0.850	0.883
2 SPP	16.060	5.650	5.270	8.990	0.904	0.880	0.845	0.876
Hatamizadeh et al. (32)	4.739	15.309	16.326	12.120	0.927	0.876	0.853	0.890
Jia and Shu (22)	3.000	2.236	1.414	2.220	0.926	0.935	0.887	0.920
Qamar et al. (19)	–	–	–	–	0.875	0.837	0.795	0.840
Jiang et al. (12)	4.610	4.130	2.650	3.800	0.888	0.837	0.833	0.850

The model was trained using a 160x160x160 picture. To determine the effect of a lower picture size. Table 4 shows it.

Table 4.Results using 160-pixel-per-side images.

Model name	Metrics							
	WT, whole tumor; TC, tumor core; ET, enhancing tumor (sub regions of tumor affected brain)							
	Hausdorff distance				Dice			
	WT	TC	ET	Average	WT	TC	ET	Average
No SPP	34.1	7.97	7.13	16.4	0.895	0.872	0.837	0.868
1 SPP	18.6	6.13	4.88	9.87	0.887	0.879	0.842	0.869
2 SPP	20.12	7.42	6.22	11.25	0.886	0.876	0.843	0.868

Table 4 allows these kind of inference. The average roll score was 0.01 lower for models that were trained on images of different sizes. There was, however, a big difference in the Hausdorff distances.The models all follow the same pattern: the dice scores stop going up after 60 epochs, no matter what size the training picture is. Secondly, models trained on 160 × 160 × 160 do not do as well as models trained on 160 × 192 × 192. In addition, it can be seen that models with SPP take longer for the loss to converge. Even though SPP doesn't add any new factors that can be trained, it still stops the model from converge. A trend-based approach can be used to fine-tune the model at very low learning rates in order to make it work better.

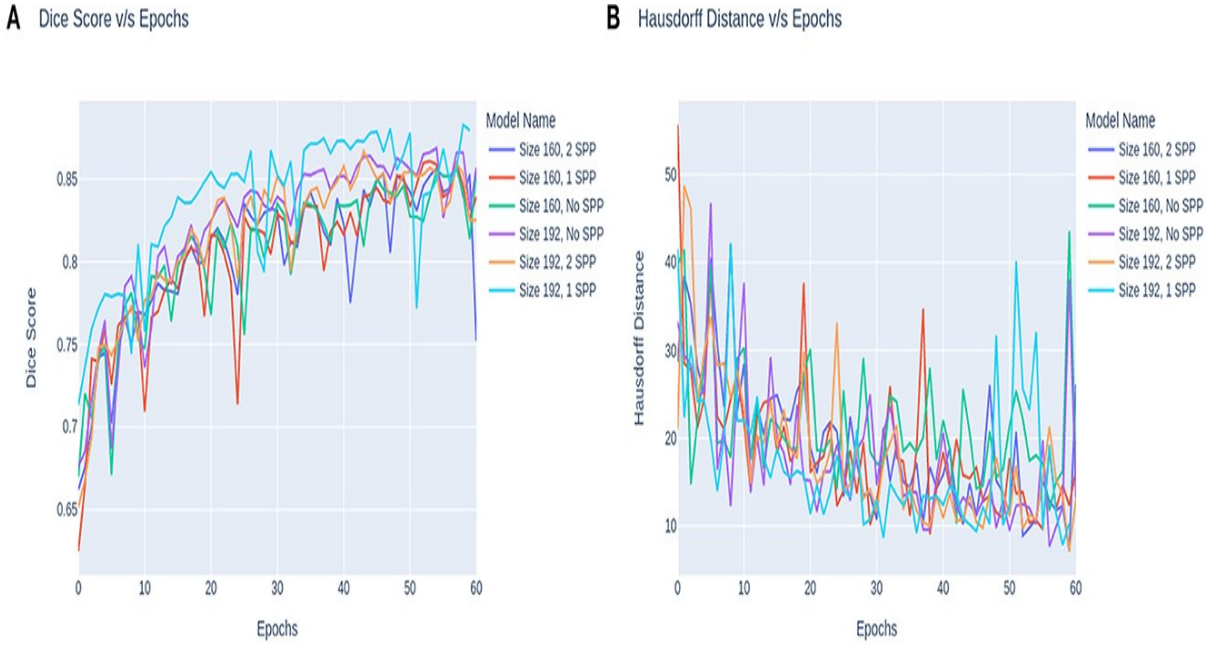


Figure 4. **Dice score against. epochs. Hausdorff scores against. epochs.**

If you look at the forecasts, you can see that the computer can accurately guess all kinds of situations. The purpose of the image is to show how well the model can predict enhancing tumors, which isn't seen much in the data. Also, the model can easily find whole tumor forms that aren't normal, which shows that it can be used for real-life detection.

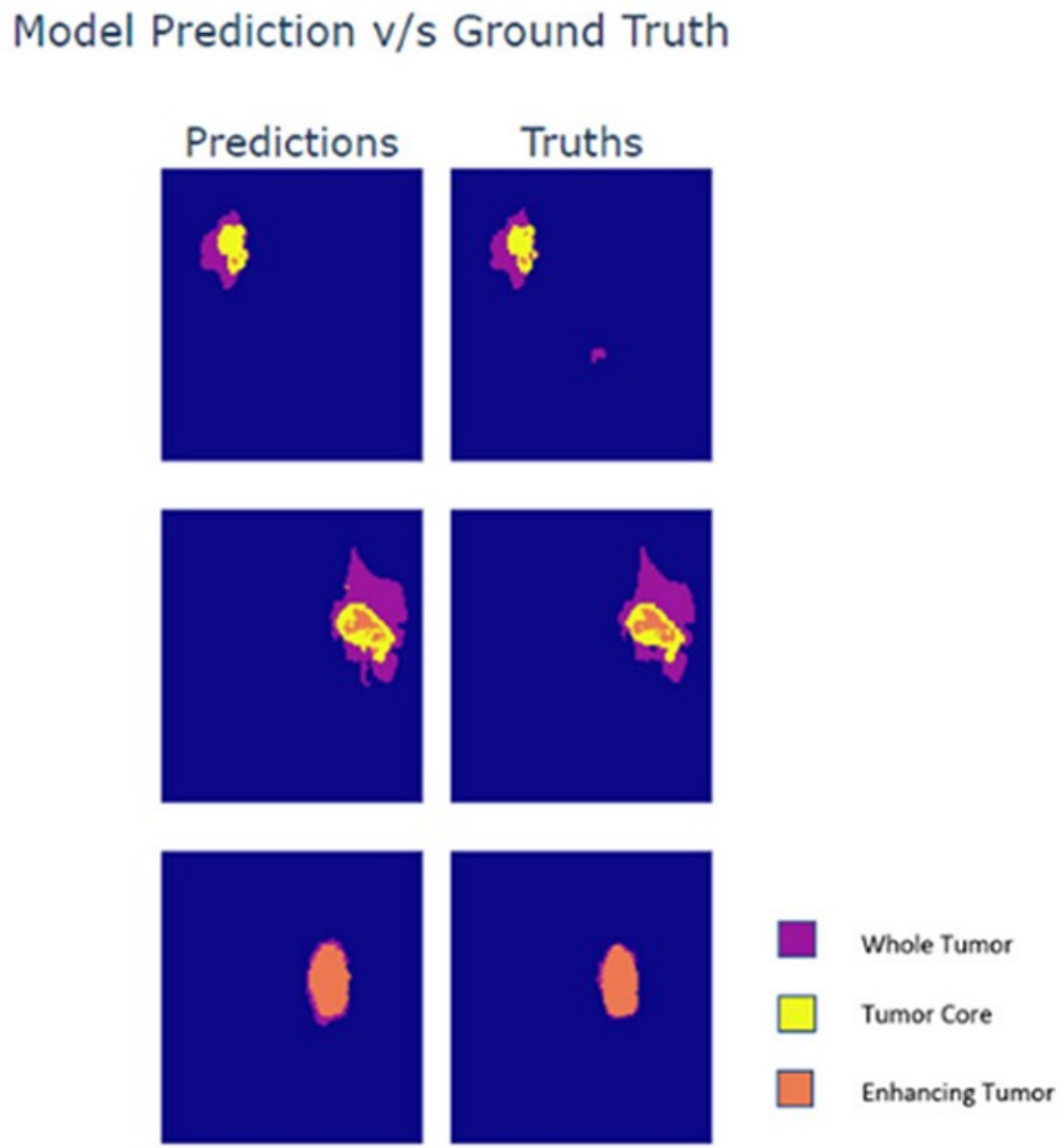


Figure 5. Comparison of prediction and ground truth segmentation masks.

Conclusion

We propose U-Net with SPP and Attention Residual Connections in this study. By providing high- and low-resolution data to the Unet decoders, this model connection notion improves model knowledge and context. Model with attention, Model with attention and 1 SPP, and Model with attention and 2 SPP are the results of applying the suggested method to the NvNet model at different rates in order to build these three versions of the model. The model that just has one SPP but requires greater attention performs the best and produces results that are on par with those produced by models with substantial transformer residue attachments. Improving information and context will boost performance significantly. This strategy works well for edge computers that must balance speed and cost. This technology might be

employed in mobile health care stations with low processing capacity that need to diagnose quickly. Sometimes the speed gain isn't as noticeable as with hefty techniques. Because components contribute trainable elements, patterns are picked up more. The technique has only been modified for one model. We hope to support more 3D-Unet designs in the future.

References

1. Rehman A, Khan MA, Saba T, Mehmood Z, Tariq U, Ayesha N. Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture. Microsc Res Tech. 2021;84(1):133–49. https://doi.org/10.1002/jemt.23597.

2. Irmak E. Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework. Iran J Sci Technol. 2021;45(3):1015–36. https://doi.org/10.1007/s40998-021-00426-9.

3. Sharif MI, Khan MA, Alhussein M, Aurangzeb K, Raza M. A decision support system for multimodal brain tumor classification using deep learning. Complex Intell Syst. 2022;8(4):3007–20. https://doi.org/10.1007/s40747-021-00321-0.

4. H. N. T. K. Kaldera, S. R. Gunasekara, and M. B. DIssanayake, "Brain tumor Classification and Segmentation using Faster R-CNN," *2019 Adv. Sci. Eng. Technol. Int. Conf. ASET 2019*, pp. 1–6, 2019, doi: https://doi.org/10.1109/ICASET.2019.8714263.

5. Crimi A, et al. Glioma, multiple sclerosis, stroke and traumatic brain injuries, vol. 2. Cham: Springer International Publishing; 2018.

6. Pereira S, Pinto A, Alves V, Silva CA. Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging. 2016;35(5):1240–51. https://doi.org/10.1109/TMI.2016.2538465.

7. Han'guk Chŏngbo Kwahakhoe, IEEE Computer Society, and Institute of Electrical and Electronics Engineers, "The 32nd International Conference on Information Networking (ICOIN 2018) : January 10 (Wed.)-12 (Fri.), 2018, Holiday Inn Chiang Mai, Chiang Mai, Thailand," pp. 473–475, 2018.

8. Noori M, Bahri A, Mohammadi K. Attention-guided version of 2D UNet for automatic brain tumor segmentation. In: 2019 9th international conference on computer and knowledge engineering. Piscataway: IEEE; 2019. p. 269–75. https://doi.org/10.1109/ICCKE48569.2019.8964956.

9. Le XH, Ho HV, Lee G, Jung S. Application of long short-term memory (LSTM) neural network for flood forecasting. Water (Switzerland). 2019;11(7):1387. https://doi.org/10.3390/w11071387.

10. Choi SG, Sohn CB. Detection of HGG and LGG brain tumors using u-net. Medico-Legal Updat. 2019;19(1):560–5. https://doi.org/10.5958/0974-1283.2019.00132.4.

11. Pravitasari AA, et al. "UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation. Telkomnika (Telecommun Comput Electron Control). 2020;18(3):1310–8. https://doi.org/10.12928/TELKOMNIKA.v18i3.14753.

12. Daimary D, Bora MB, Amitab K, Kandar D. Brain tumor segmentation from MRI images using hybrid convolutional neural networks. Proc Comput Sci. 2020;167(2019):2419–28. https://doi.org/10.1016/j.procs.2020.03.295.

13. Pei L, Bakas S, Vossough A, Reza SMS, Davatzikos C, Iftekharuddin KM. Longitudinal brain tumor segmentation prediction in MRI using feature and label fusion. Biomed Signal Process Control. 2020;55:101648. https://doi.org/10.1016/j.bspc.2019.101648.

14. Lin M, et al. Fully automated segmentation of brain tumor from multiparametric MRI using 3D context deep supervised U-Net. Med Phys. 2021;48(8):4365–74. https://doi.org/10.1002/mp.15032.

15. Punn NS, Agarwal S. Multi-modality encoded fusion with 3D inception U-net and decoder model for brain tumor segmentation. Multimed Tools Appl. 2021;80(20):30305–20. https://doi.org/10.1007/s11042-020-09271-0.

16. Ullah F, et al. Brain MR image enhancement for tumor segmentation using 3d u-net. Sensors. 2021;21(22):1–14. https://doi.org/10.3390/s21227528.

17. Brain Tumor Segmentation 2020 Dataset. [Online]. https://www.kaggle.com/awsaf49/brats20-dataset-training-validation.

18. Montaha S, Azam S, Rafid AKMRH, Hasan MZ, Karim A, Islam A. TimeDistributed-CNN-LSTM: a hybrid approach combining CNN and LSTM to classify brain tumor on 3D MRI scans performing ablation study. IEEE Access. 2022;10:60039–59. https://doi.org/10.1109/ACCESS.2022.3179577.

19. Khalil HA, Darwish S, Ibrahim YM, Hassan OF. 3D-MRI brain tumor detection model using modified version of level set segmentation based on dragonfly algorithm. Symmetry (Basel). 2020;12(8):1256. https://doi.org/10.3390/SYM12081256.

20. Gadosey PK, et al. SD-UNET: stripping down U-net for segmentation of biomedical images on platforms with low computational budgets. Diagnostics. 2020;10(2):1–18. https://doi.org/10.3390/diagnostics10020110.

21. Weng W, Zhu X. INet: convolutional networks for biomedical image segmentation. IEEE Access. 2021;9:16591–603. https://doi.org/10.1109/ACCESS.2021.3053408.

22. S. Banerjee, S. Mitra, F. Masulli, and S. Rovetta, “Deep Radiomics for Brain Tumor Detection and Classification from Multi-Sequence MRI,” pp. 1–15, 2019, [Online]. http://arxiv.org/abs/1903.09240.

23. Rehman MU, Cho S, Kim JH, Chong KT. Bu-net: brain tumor segmentation using modified u-net architecture. Electron. 2020;9(12):1–12. https://doi.org/10.3390/electronics9122203.

24. Asim M, Rashid A, Ahmad T. Scour modeling using deep neural networks based on hyperparameter optimization. ICT Express. 2021;8:357–62. https://doi.org/10.1016/j.icte.2021.09.012.

25. Ding P, Li J, Wang L, Wen M, Guan Y. HYBRID-CNN: an efficient scheme for abnormal flow detection in the SDN-based smart grid. Secur Commun Netw. 2020;2020:1–20. https://doi.org/10.1155/2020/8850550.

26. Kandel I, Castelli M. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. ICT Express. 2020;6(4):312–5. https://doi.org/10.1016/j.icte.2020.04.010.

27. Khan MA, Sharif MI, Raza M, Anjum A, Saba T, Shad SA. Skin lesion segmentation and classification: a unified framework of deep neural network features fusion and selection. Expert Syst. 2022;39(7):1–21. https://doi.org/10.1111/exsy.12497.

28. Khan MA, et al. Lungs cancer classification from CT images: an integrated design of contrast based classical features fusion and selection. Pattern Recognit Lett. 2020;129:77–85. https://doi.org/10.1016/j.patrec.2019.11.014.

29. Ghosh S, Bandyopadhyay A, Sahay S, Ghosh R, Kundu I, Santosh KC. Colorectal histology tumor detection using ensemble deep neural network. Eng Appl Artif Intell. 2021;100:104202. https://doi.org/10.1016/j.engappai.2021.104202.

30. Montaha S, et al. MNet-10: a robust shallow convolutional neural network model performing ablation study on medical images assessing the effectiveness of applying optimal data augmentation technique. Front Med. 2022. https://doi.org/10.3389/fmed.2022.924979.

31. Naser MA, Deen MJ. Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. Comput Biol Med. 2020;121:103758. https://doi.org/10.1016/j.compbiomed.2020.103758.

32.Vijay Sanchit, Guhan Thejineaswar, Srinivasan Kathiravan, Vincent P. M. Durai Raj, Chang Chuan-Yu, MRI brain tumor segmentation using residual Spatial Pyramid Pooling-powered 3D U-Net, Frontiers in Public Health, volume 11, 2023