

Brain tumors are the growth of abnormal cells or a mass in a brain. Numerous kinds of brain tumors were discovered, which need accurate and early detection techniques. Currently, most diagnosis and detection methods rely on the decision of neuro-specialists and radiologists to evaluate brain images, which may be time-consuming and cause human errors. This paper proposes a robust U-Net deep learning Convolutional Neural Network (CNN) model that can classify if the subject has a tumor or not based on Brain Magnetic resonance imaging (MRI) with acceptable accuracy for medical-grade application. The study built and trained the 3D U-Net CNN including encoding/decoding relationship architecture to perform the brain tumor segmentation because it requires fewer training images and provides more precise segmentation. The algorithm consists of three parts; the first part, the downsampling part, the bottleneck part, and the optimum part. The resultant semantic maps are inserted into the decoder fraction to obtain the full-resolution probability maps. The developed U-Net architecture has been applied on the MRI scan brain tumor segmentation dataset in MICCAI BraTS 2017. The results using Matlab-based toolbox indicate that the proposed architecture has been successfully evaluated and experienced for MRI datasets of brain tumor segmentation including 336 images as training data and 125 images for validation. This work demonstrated comparative performance and successful feasibility of implementing U-Net CNN architecture in an automated framework of brain tumor segmentations in Fluid-attenuated inversion recovery (FLAIR) MR Slices. The developed U-Net CNN model succeeded in performing the brain tumor segmentation task to classify the input brain images into a tumor or not based on the MRI dataset

Keywords: *Magnetic Resonance Imaging (MRI), deep learning, Convolutional Neural Network (CNN), 3D U-Net architecture, brain tumors, segmentations*

DEVELOPMENT OF BRAIN TUMOR SEGMENTATION OF MAGNETIC RESONANCE IMAGING (MRI) USING U-NET DEEP LEARNING

Wasan M Jwaid

Assistant Teacher
Department of Banking and
Finance Administration and Economics
University of Thi-Qar
Nasiriyah, Iraq, 0096442

Zainab Shaker Matar Al-Husseini

Assistant Teacher
Department of Computer Technology Engineering
College of Information Technology
Imam Ja'afar Al-Sadiq University
Nasiriyah, Iraq

Ahmad H. Sabry

Corresponding author
Doctor of Control and Automation Engineering
Department of Institute of Sustainable Energy
Universiti Tenaga Nasional
Jalan Ikram-Uniten, Kajang,
Selangor, Malaysia, 43000
E-mail: ahs4771384@gmail.com

Received date 11.06.2021

Accepted date 28.07.2021

Published date 31.08.2021

How to Cite: Jwaid, W. M., Al-Husseini, Z. S. M., Sabry, A. H. (2021). Development of brain tumor segmentation of magnetic resonance imaging (MRI) using U-net deep learning. *Eastern-European Journal of Enterprise Technologies*, 4 (9 (112)), 23–31.

doi: <https://doi.org/10.15587/1729-4061.2021.238957>

1. Introduction

The method of distinguishing tumor borders from healthy cells is still a challenging mission in the medical habit. Fluid-attenuated inversion recovery (FLAIR) and Magnetic resonance imaging (MRI) modalities can provide physicians with excellent information concerning tumor penetration [1]. Brain tumors are the growth of abnormal cells or a mass in a brain. Several studies discussed the application of Convolutional Neural Network (CNN) model as a deep learning architecture for MRI-based brain tumor segmentation by using magnetic resonance FLAIR images [2], multimodal MRI scans [3–5], and automatic semantic segmentation [6]. The study [7] presented 3D convolutional neural networks for tumor segmentation using long-range 2D context, which was then updated to more accurately classify and detect brain cancer cells in MRI and computerized tomography (CT) images using nano-contrast agents [8], and using dense residual refine networks for automatic brain tumor segmentation in [9, 10]. Numerous kinds

of brain tumors were discovered, which need accurate and early detection techniques. Currently, most diagnosis and detection methods rely on the decision of neuro-specialists and radiologists to evaluate brain images, which may be time-consuming and cause human errors.

U-Net Convolutional Neural Network (CNN) model is typically used for image segmentation. U-Net CNN model is a customized network, which has been introduced originally in [11]. It is called U-Net because it looks like U and the architecture itself contains two paths. Among numerous types of CNNs, U-Net architectures are the major fully convolutional network models for semantic segmentation in medical tasks. Current works show that U-Net networks can be significantly deeper to improve the performance of segmentation tasks. A network became deeper and may lead to redundant computation or gradient vanishing during training through adding extra layers directly.

As a result, segmentation, detection, and extraction of contaminated tumor regions from magnetic resonance (MR) images is a major concern, but it is a time-consuming and labo-

rious task conducted by radiologists or clinical experts, whose accuracy is solely dependent on their experience. As a result, the employment of computer-assisted technology becomes increasingly important in order to overcome these constraints.

2. Literature review and problem statement

The study [2] developed CNN-based fully automated detection and segmentation as an application of CNN on gliomas, which is the most aggressive and common type of brain tumors due to their rapid progression and infiltrative nature. Although the study uses U-Net including encoding/decoding relationship architecture and demonstrated that using U-Net has a comparative performance and is feasible compared to applying another deep learning CNN architectures, this work didn't show the network loss and accuracy through training results. The paper [12] introduced brain tumor segmentations and grading of lower-grade glioma (LGG) for MRI imaging. This paper also discussed the grading and segmentation models using the same pipeline of FLAIR, T1-precontrast, and T1-postcontrast for 110 patients of LGG. The paper classified LGG by calculating the accuracy, sensitivity. However, this work also didn't show the network loss and accuracy through training results, and didn't provide more details about the model. A new study [13] introduced a new CNN architecture called DIU-Net model to integrate the modules; convolutional densely connecting by the Inception-Res within a U-Net architecture. Their experiments of combining the inception with dense connections modules with the U-Net were applied on images of the lung segmentation of CT Data from the benchmark Kaggle dataset, blood vessel segmentations from retina images, and the MRI scan brain tumor segmentations dataset in MICCAI-BraTS 2017. The limitation was that the dense-inception part was raising the growth rates that might lead to numerous parameters, making the architecture slower to train and further difficult.

A way to overcome these difficulties can be a brain tumor segmentation of MRI imaging using a hybrid CNN in [14]. Although this work relayed on the measured accuracy to compare the performances of each of the SegNet3, SegNet5, U-SegNet, Res-SegNet, Seg-UNet, and U-Net, the study doesn't show the network loss and accuracy through training results, and doesn't provide more details about the model. Multi-perspective scaling CNN is proposed by [15] for high-resolution MRI brain segmentations. The simulation results of this work showed that the anticipated MPS-CNN and U-Net have better trade-off as compared with other algorithms like FCN, SegNet, Deep V3, CNN, and Deep FCN. A recent study [16] presented a Multi-modal cascaded U-Net architecture based on Fully Convolutional Deep Network (FCN) and Deep Convolutional Neural Network (D-CNN) implemented to classify brain glioblastoma tumors into High-Grade (HG) and Low-Grade (LG). Multistage-based multi-modal image classification of brain tumors was also discussed by [17] using CNN to segregate and identify the tumor classes by applying a patch-wise classification method. The study results are 0.9636 of accuracy and 0.9214 of sensitivity.

All this suggests that it is advisable to conduct a study on developing a brain tumor segmentation of MRI using U-Net

deep learning CNN to classify the input brain images into a tumor or not based on the MRI dataset.

3. The aim and the objectives of the study

The study aims to build and train the 3D U-Net CNN including encoding/decoding relationship architecture to perform the brain tumor segmentation.

To achieve this aim, the following objectives are accomplished:

- to evaluate the developed 3D U-Net CNN by showing acceptable network accuracy and loss through the training;
- to classify if the subject has a tumor or not based on MRI of brain tumors segmentation task;
- to evaluate the proposed U-Net CNN architecture by comparing it with a well-known CNN model.

4. Materials and methods

The main purpose of this project is to build a robust Convolutional Neural Network (CNN) model that can classify if the subject has a tumor or not based on Brain Magnetic resonance imaging (MRI) scan images with acceptable accuracy for medical-grade application. Deep learning is able to actually learn features, extract features, and make classifications. The predictions are based on that features. The workflow of the developed segmentation for Brain Tumor MRI is shown in Fig. 1.

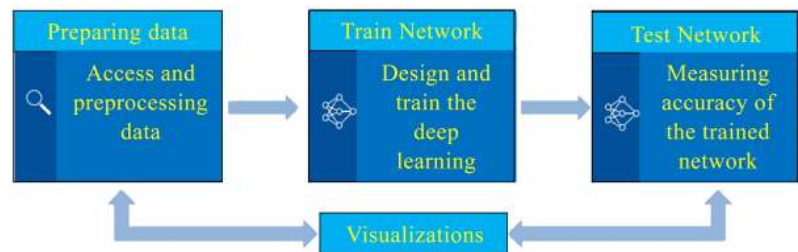


Fig. 1. Segmentation workflow for Brain Tumor MRIs

We will follow these steps and start with loading and cleaning the data. The MRI data set is publicly available and can be downloaded for free. It is essential to perform some pre-processing to crop and normalize the 3D MRI images. These images are also referred to as volumes. U-Net CNN model is shown in Fig. 2.

All blue boxes correspond to multi-channel feature maps. On top of the box, there is an indication of the number of channels, while at the lower-left edge of the boxes, the x-y-size is provided. The white box represents copied feature map, while the arrows show different actions.

The left side of U-Net is called contraction path or encoder path, while that on the right is the expansion, expanding, or decoder path. The concatenation of in between feature maps is the reason to get localized information, which makes a semantic segmentation possible using the U-Net model.

In the developed U-Net architecture, there are three parts; the first part, on the left-hand side, is the downsampling part from B1 to B6. The second part is the bottleneck represented by B7, and the third part is the optimum part from B8 to B13. To clearly explain the operation of the architecture, we divide it into two parts as shown in Fig. 3.

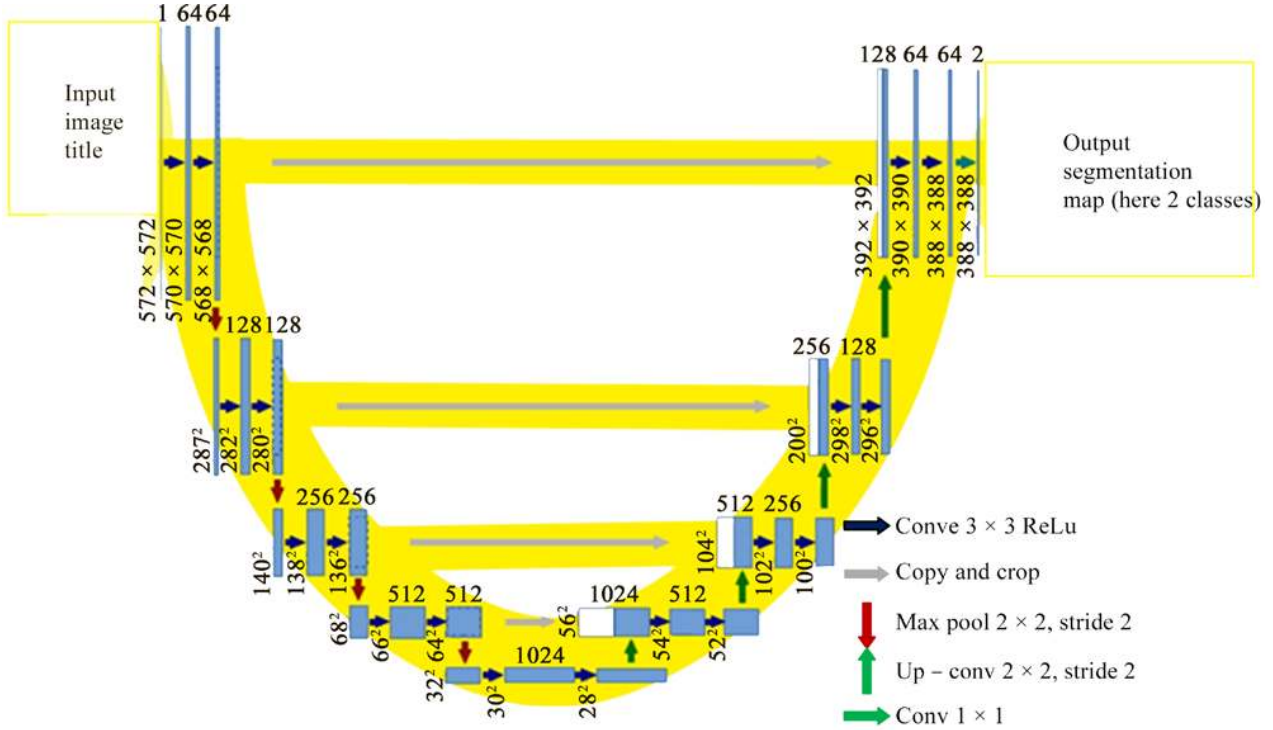


Fig. 2. The architecture of U-Net CNN

B1 represents the input image with size 256×256 , and there are three channels for the input image of red, green and blue as an RGB image. There are different arrows in different colors. The green arrow indicates a convolutional layer with a kernel size of 3×3 , and it comes with a ReLu function as an activation function. After passing it to the first convolutional layers, we are getting eight feature maps. The second convolutional layer has another eight feature maps passing then to a max pooling layer, which is represented by the red arrow. The kernel size for the max pooling layer is 2×2 , which will extract the more important features from the previous feature maps. After passing it to the max pooling layer, the image size will become from 256^2 to 128^2 . Then, the data pass to two 3×3 convolutional layers. At box B2, 16 feature maps basically do the similar things from B1 to B6. At B7, the feature maps are taken from B6 and pass it to the max pooling layer. After passing the feature maps to the max pooling layer, an image size of 4×4 is obtained, and the number of filters is 256 in this case. This will pass to 2 convolutional layers at B7, which is called bottleneck, which has one max pooling layer and up convolutional layer. After passing the feature maps to the up convolutional layer, 64 feature maps are obtained with an image size of 8×8 , and the grey arrow here means copy and crop, where the feature maps are copied from B6 and crop it to the feature maps at B8. The copy and crop here is getting more features from the previous boxes and attach it to the current box. After the copy and crop, more features are obtained for the official maps so the input for the boxes has more features. Then, the input at the gray box passes to two convolutional layers, and then go to up convolutional layers at B9. The image size here is 16×16 , and a copy and crop is performed on the feature maps from B5 and concatenate it to B9. Then the input is passed to two convolutional layers. For box eight, we are doing a similar thing, from B8 to B13. At B13, we are concatenating the feature maps from B1 as

well, and then we will pass it to two convolutional layers. The last layer for B13 is a 1×1 convolutional layer. Here we apply a sigmoid activation function, which will confine the pixel value in the range of 0 and 1. Therefore, the values on the last feature map will be confined in the range of zero and one. Connected components are employed to reduce false positives. Only one forward pass produces the segmentation masks for the whole brain slice.

The final output function is calculated by a soft-max pixel-wise at the final feature map together with the loss function of the cross entropy. The function of the soft-max is given by [18]:

$$p_k(x) = \frac{\exp(a_k(x))}{\sum_{k=1}^{k=K} \exp(a_k(x))}, \quad (1)$$

where $p_k(x)$ represents the maximum function approximation nearly equal to 1 when k has the maximum activation of $a_k(x)$, which represents the activation function channel k at the position of the pixel ($x \in \Omega$) and ($\Omega \subset \mathbb{Z}^2$). K denotes the classes' number. $p_k(x) \approx 0$ for all other k . The cross entropy at all positions is given by:

$$En = \sum_{x \in \Omega} \omega(x) \log(p_l(x)). \quad (2)$$

where $l: \Omega \rightarrow \{1, \dots, K\}$, which is the true label of each pixel, while $\omega: \Omega \rightarrow \mathbb{R}$, which is a weight map presented to provide a priority to some pixels during training.

A voxel is the 3D equivalent of a pixel. The 3D datasets are typically very large so a random patch extraction is used to avoid running out of memory as demonstrated in Fig. 4.

We intend to extract random patches from ground truth slices and the corresponding label information for training and validation as demonstrated in Fig. 5.

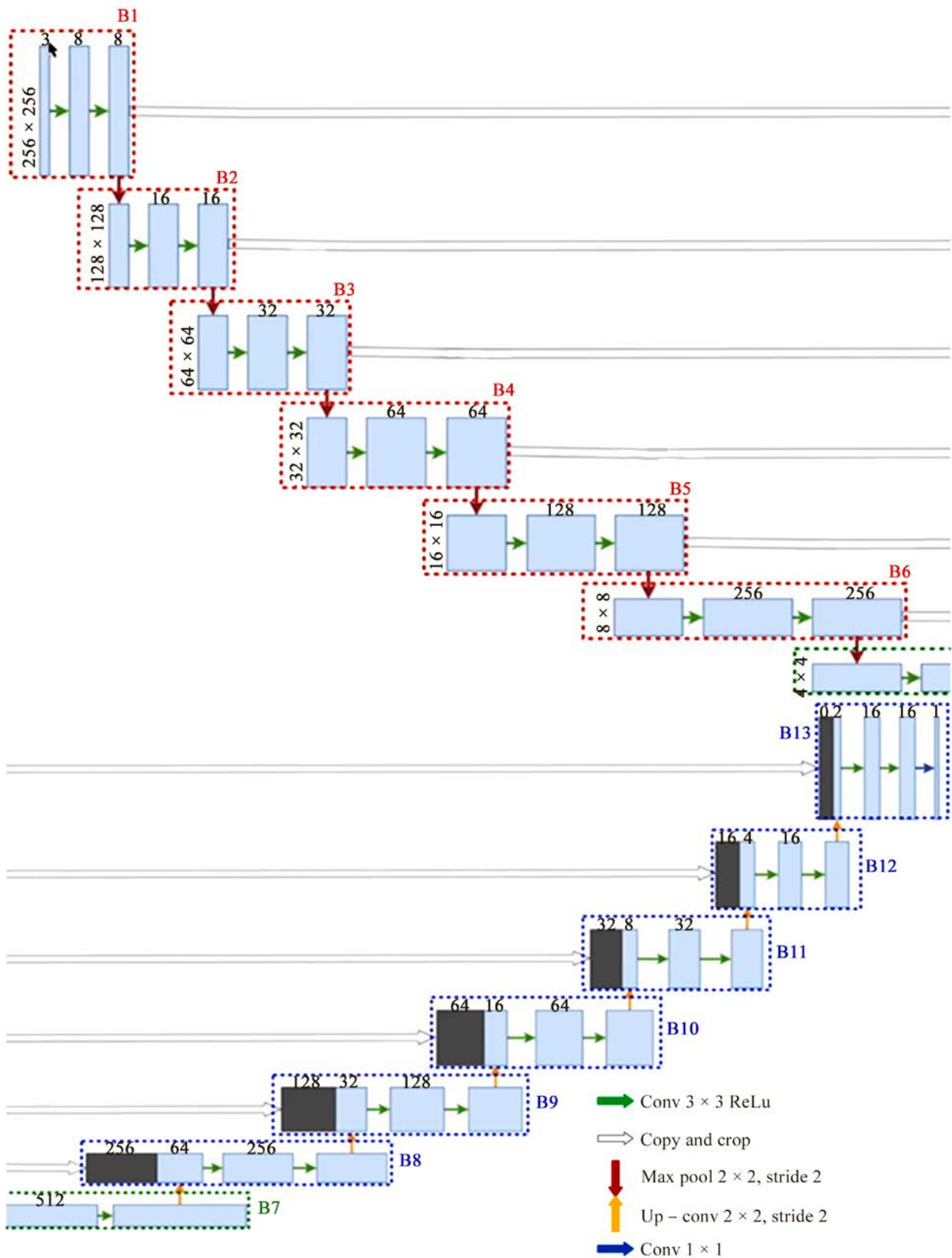


Fig. 3. The developed U-Net CNN architecture

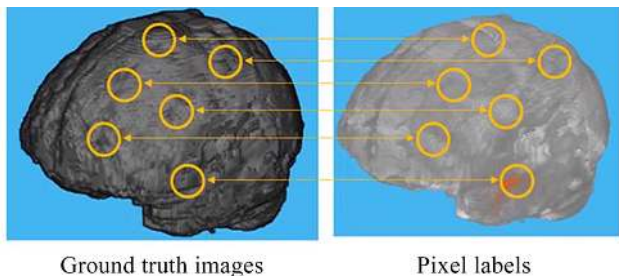


Fig. 4. Random patch extraction datastore

Next is to build and train the 3D U-Net CNN architecture to perform the brain tumor segmentation because it requires fewer training images and provides more precise segmentation. We first create the layers, and then concatenate them. Training the network on the entire dataset takes a long time. Therefore, the pre-trained network is considered. We can see the network accuracy and loss through the training results and we perform segmentation on the test images to check the performance of our network.

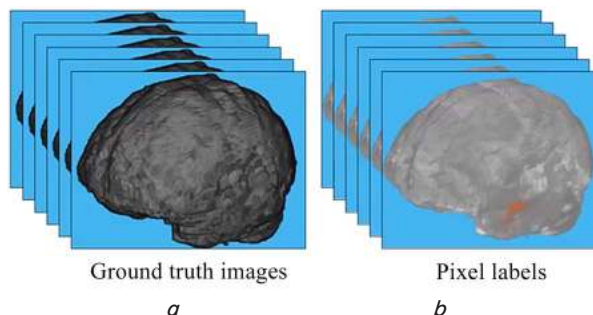


Fig. 5. Extraction of random patches from ground truth slices and the corresponding label data:
a – ground truth images; *b* – pixel labels

5. Results of the proposed approach

5.1. Results of network accuracy and loss through the training

The training network can be shown in Fig. 6, while the measurements of network accuracy and loss through the training are shown in Fig. 7.

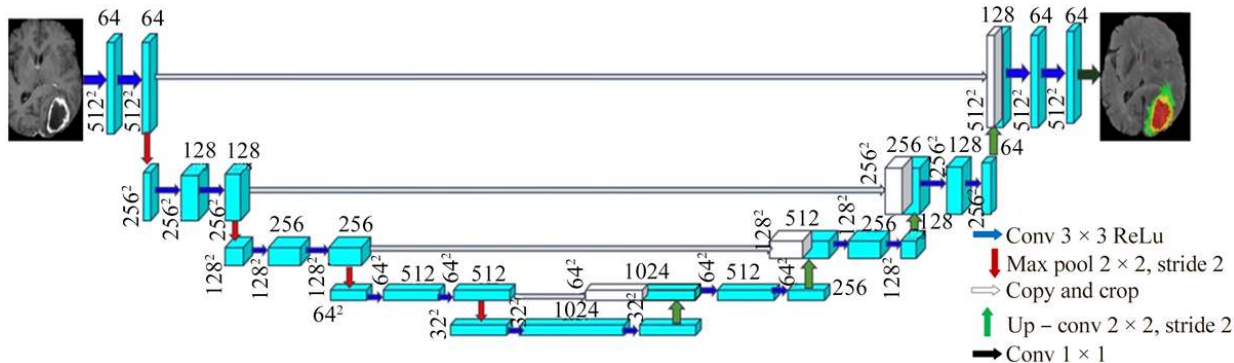


Fig. 6. The training network

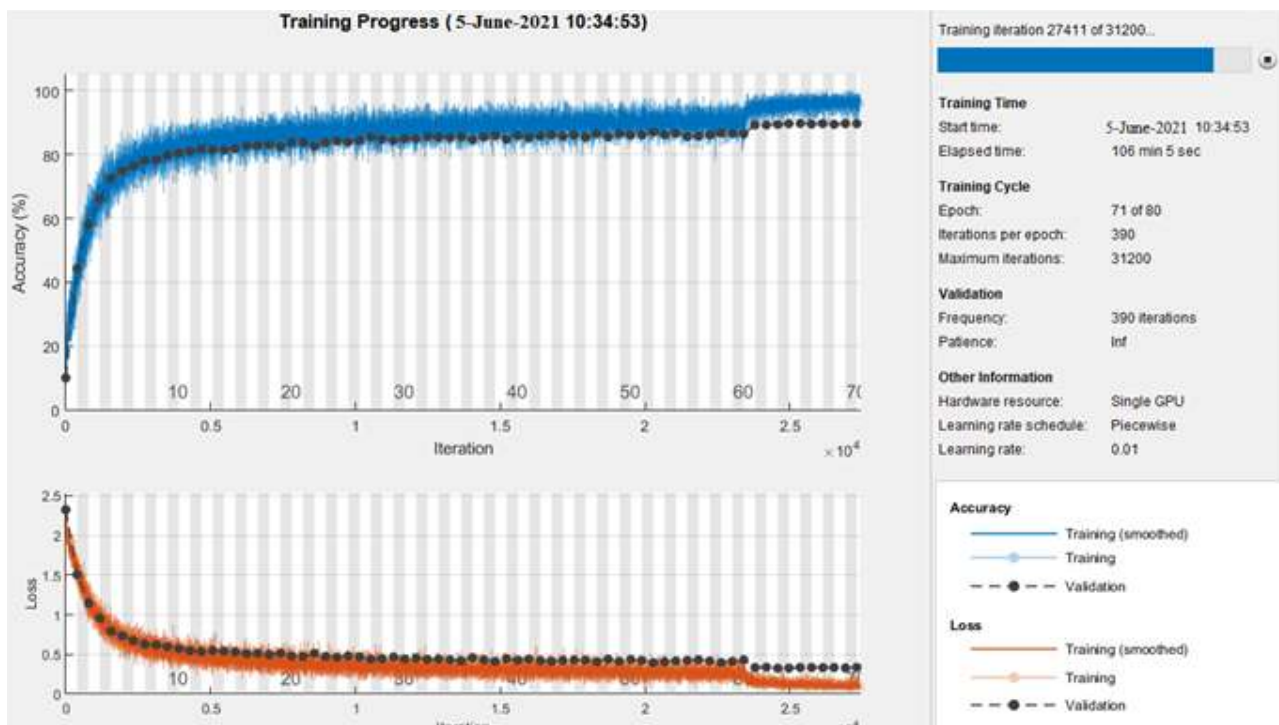


Fig. 7. The network accuracy and loss through the training results

This result shows the application of the U-Net CNN to classify if the subject has a tumor or not based on brain MRI with acceptable accuracy for brain tumor segmentation tasks.

5. 2. Results of brain tumor segmentation

The dynamic representation of the predicted segmentations as compared to the ground truth segmentations is shown in Fig. 8.

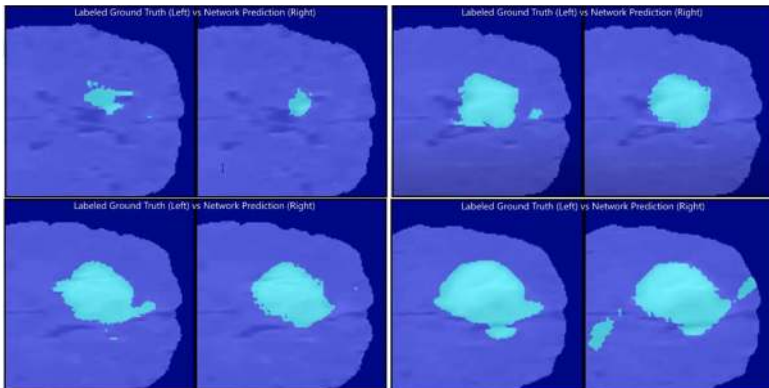


Fig. 8. Dynamic representation of the predicted segmentations as compared to the ground truth segmentations

Since we used the Matlab-based toolbox, it is possible to utilize the 3D volumetric application to measure brain tumor defect as demonstrated in Fig. 9.

Fig. 10 shows samples of ground truth images with their corresponding predicted outcome from the U-Net CNN model from three different views. The yellow, green, red, and gray colors indicate the enhancing brain tumors and non-enhancing tumor.

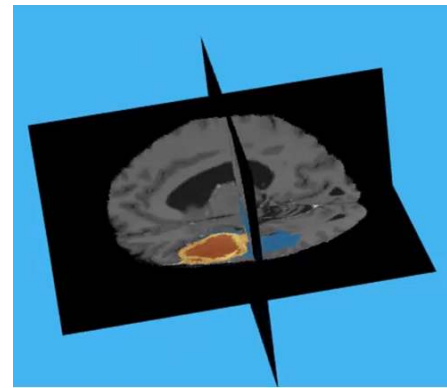


Fig. 9. 3D volumetric application to measure brain tumor defect

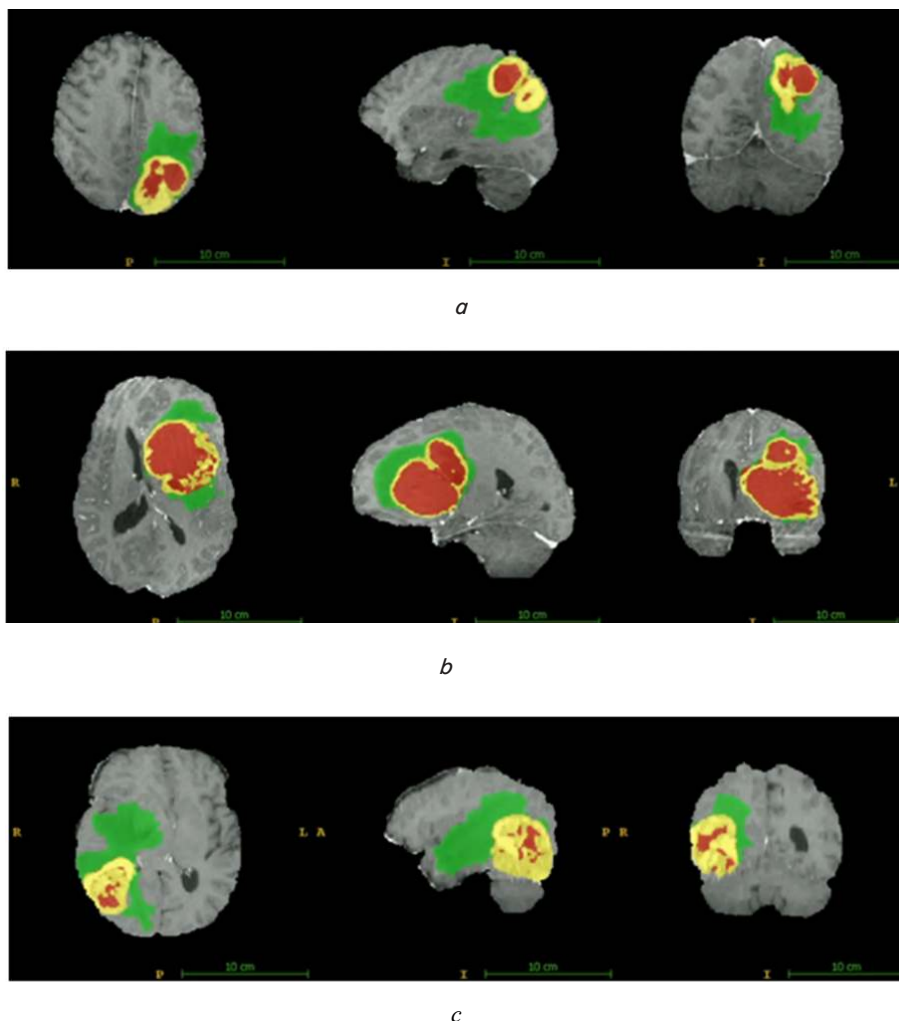


Fig. 10. Ground truth images with their corresponding predicted outcome from U-Net CNN model: a – top-views; b – side-view; c – bottom-view

We've constructed a number of brain voxels transparent to visualize tumors, as depicted in Fig. 11.

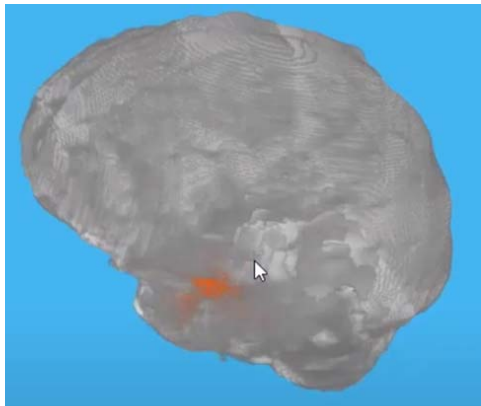


Fig. 11. Brain voxels transparent

Fig. 11 - demonstrates the label wall show-function to preview a label training image as tumors are located in brain tissues.

5.3. Comparing the proposed U-Net CNN architecture to the SegNet5 CNN model

The proposed U-Net architecture was implemented, and the accuracy of segmentation was assessed by comparing it to a prominent CNN model for image segmentation, SegNet5. Fig. 12 depicts a set of ground truth photos along with the models' anticipated output. The enhancing tumor, necrotic and non-enhancing tumor, peritumoral edema, and anything else are represented by the green, red, yellow, and gray colors, respectively.

Five metrics are used to assess how well the model performs in segmentation tasks: global accuracy, mean accuracy, mean Intersection-Over-Union (IOU), weighted IOU, and mean BF-scores. The ratio of the highest correctly categorized pixels of one class to the entire number of pixels is referred to as global accuracy, whereas mean accuracy refers to the average percentage of correctly identified pixels for each class. The mean IOU is the averaged IOU of each class, commonly known as the Jaccard similarity coefficient. IOU can be defined by:

$$IOU = \frac{Tp}{Tp + Fp + Fn}$$

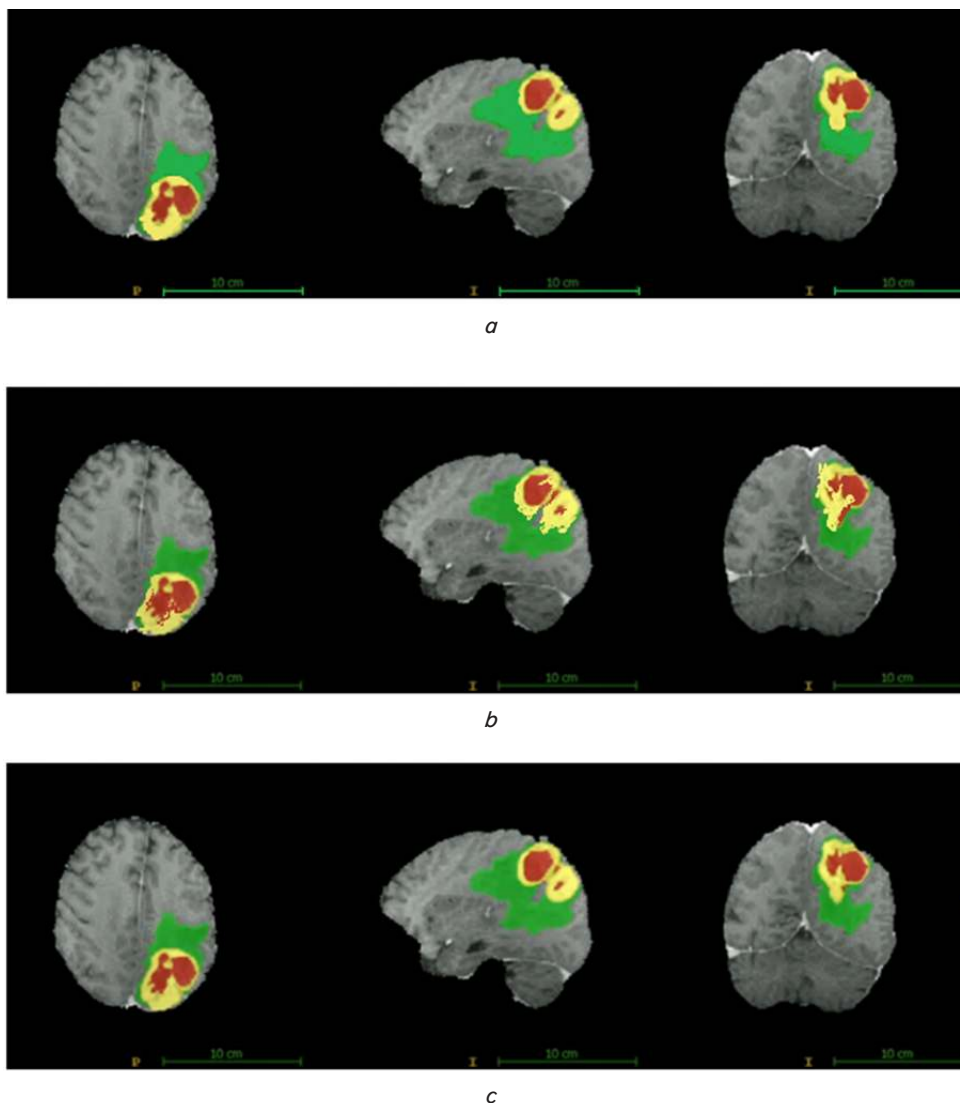


Fig. 12. Comparison showing the brain tumor segmentation on the model: *a* – ground truth; *b* – SegNet5; *c* – proposed U-Net CNN

True positives, false positives, and false negatives are denoted by the letters Tp , Fp and Fn , respectively. The weighted IOU defines how many pixels of each class are weighted in the disproportion pixel class to prevent the larger class from overlapping the smaller class. Table 1 shows the segmentation performance parameters of the proposed and SegNet5 CNN models.

Table 1

Performance evaluation for the proposed and SegNet5 CNN models.

Model name	Global accuracy	Mean accuracy	Mean IOU	Weighted IOU	Training time
SegNet5	0.9625	0.9048	0.5587	0.9397	530 min
U-Net	0.9898	0.915	0.69313	0.9846	644 min

Therefore, the developed U-Net CNN architecture is introduced by modifying the upper/lower sampling module to replace the upper/layer pooling layer of the current U-Net architecture.

6. Discussion of experimental results of the proposed approach

In order to explain the results of this work, the network in Fig. 6 demonstrates the construction of the network, where it accepts inputs of dimension 64×64 and predicts the class associated to all the 64×64 pixels in a single forward pass. The network comprises an encoder and decoder. The encoder comprises convolutional layers and max pooling layer. The decoder comprises transposed convolution and convolutional layers. Skip connection is used in the network to combine low-level high-resolution features and high-level low-resolution features.

For Fig. 7, the upper graph demonstrates the accuracy during the learning, which was found to be 0.96, while the cross-entropy loss was close to 0.132 after 71 epochs training (27411 iterations), which can be seen on the right side of the figure. The elapsed time here was 106 minutes as the hardware resource is a single CPU and not GPU. The dice similarity coefficient is used to measure the accuracy of the network, which is 85 % for the segmentation of brain tumors. Using deep learning for segmentation also allows for further classification. This work can be extended to identifying different types of demos.

The figures (Fig. 8–11) show that the developed U-Net CNN is feasible, and the yellow, green, red, and gray colors indicate the enhancing brain tumors and non-enhancing tumors effectively.

Overall, the U-Net outperforms the other models (SegNet5) and produces well-segmented output with accurate grading in each class. The training parameters are the most important consideration when calculating the computing time of a CNN. As a result, it is critical to set up all of the training parameters consistently across all models and to use the same dataset. Because it has fewer layers than the intended U-Net, the SegNet5 requires less time to train and took 530 minutes to finish.

The limitations of this work are that the developed U-Net CNN architecture takes more time to train because of using skip connections, as well as the number of layers and training settings. It is also solely applied to medical brain tumor images and also the need for adjusting the network element features, which reflects its disadvantages.

These limitations can be overcome by developing a parameter optimization technique combining the presented U-Net CNN architecture to tune the structure elements automatically.

7. Conclusions

1. The training accuracy and loss of the developed 3D U-Net CNN network to achieve the segmentation was acceptable, which are evaluated by showing 0.96 network accuracy and 0.132 losses.

2. The segmentation task of the MRI-based tumor subject was performed successfully after 71 epochs training (27411 iterations).

3. Because of using skip connections, as well as the number of layers and training settings, the suggested U-Net architecture takes more time to train (644 minutes), but it produces better segmentation results. After the network has been trained, it can be utilized for image segmentation. Image segmentation using the trained model takes only a few seconds. Manual tumor segmentation by clinical professionals, on the other hand, can take hours. The proposed image segmentation approaches are accurate, quick, and low-cost to deploy. This will aid doctors in making a quick and precise diagnosis of a brain tumor, perhaps saving the lives of many people.

References

- Zwanenburg, J. J. M., Hendrikse, J., Visser, F., Takahara, T., Luijten, P. R. (2010). Fluid attenuated inversion recovery (FLAIR) MRI at 7.0 Tesla: comparison with 1.5 and 3.0 Tesla. *European Radiology*, 20 (4), 915–922. doi: <https://doi.org/10.1007/s00330-009-1620-2>
- Zeineldin, R. A., Karar, M. E., Coburger, J., Wirtz, C. R., Burgert, O. (2020). DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. *International Journal of Computer Assisted Radiology and Surgery*, 15 (6), 909–920. doi: <https://doi.org/10.1007/s11548-020-02186-z>
- Sun, L., Zhang, S., Chen, H., Luo, L. (2019). Brain Tumor Segmentation and Survival Prediction Using Multimodal MRI Scans With Deep Learning. *Frontiers in Neuroscience*, 13. doi: <https://doi.org/10.3389/fnins.2019.00810>
- Kulkarni, S. M., Sundari, G. (2020). A Framework for Brain Tumor Segmentation and Classification using Deep Learning Algorithm. *International Journal of Advanced Computer Science and Applications*, 11 (8). doi: <https://doi.org/10.14569/ijacsa.2020.0110848>
- Al-qazzaz Salma, Sun, X., Yang, H., Yang, Y., Xu, R., Nokes, L., Yang, X. (2020). Image classification-based brain tumour tissue segmentation. *Multimedia Tools and Applications*, 80 (1), 993–1008. doi: <https://doi.org/10.1007/s11042-020-09661-4>
- Cui, S., Mao, L., Jiang, J., Liu, C., Xiong, S. (2018). Automatic Semantic Segmentation of Brain Gliomas from MRI Images Using a Deep Cascaded Neural Network. *Journal of Healthcare Engineering*, 2018, 1–14. doi: <https://doi.org/10.1155/2018/4940593>

7. Mlynarski, P., Delingette, H., Criminisi, A., Ayache, N. (2019). 3D convolutional neural networks for tumor segmentation using long-range 2D context. *Computerized Medical Imaging and Graphics*, 73, 60–72. doi: <https://doi.org/10.1016/j.compmedimag.2019.02.001>
8. Ruba, T., Tamilselvi, R., Beham, M. P., Aparna, N. (2020). Accurate Classification and Detection of Brain Cancer Cells in MRI and CT Images using Nano Contrast Agents. *Biomedical and Pharmacology Journal*, 13 (03), 1227–1237. doi: <https://doi.org/10.13005/bpj/1991>
9. Sun, J., Chen, W., Peng, S., Liu, B. (2019). DRRNet: Dense Residual Refine Networks for Automatic Brain Tumor Segmentation. *Journal of Medical Systems*, 43 (7). doi: <https://doi.org/10.1007/s10916-019-1358-6>
10. Sajid, S., Hussain, S., Sarwar, A. (2019). Brain Tumor Detection and Segmentation in MR Images Using Deep Learning. *Arabian Journal for Science and Engineering*, 44 (11), 9249–9261. doi: <https://doi.org/10.1007/s13369-019-03967-8>
11. Farahani, A., Mohseni, H. (2020). Medical image segmentation using customized U-Net with adaptive activation functions. *Neural Computing and Applications*, 33 (11), 6307–6323. doi: <https://doi.org/10.1007/s00521-020-05396-3>
12. Naser, M. A., Deen, M. J. (2020). Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Computers in Biology and Medicine*, 121, 103758. doi: <https://doi.org/10.1016/j.compbiomed.2020.103758>
13. Abdelhafiz, D., Bi, J., Ammar, R., Yang, C., Nabavi, S. (2020). Convolutional neural network for automated mass segmentation in mammography. *BMC Bioinformatics*, 21 (S1). doi: <https://doi.org/10.1186/s12859-020-3521-y>
14. Rao, S., Lingappa, B. (2019). Image Analysis for MRI Based Brain Tumour Detection Using Hybrid Segmentation and Deep Learning Classification Technique. *International Journal of Intelligent Engineering and Systems*, 12 (5), 53–62. doi: <https://doi.org/10.22266/ijies2019.1031.06>
15. Yuvaraj, D., Noori, S. F., Swaminathan, S. (2021). Multi-perspective scaling convolutional neural networks for high-resolution MRI brain image segmentation. *Materials Today: Proceedings*. doi: <https://doi.org/10.1016/j.matpr.2020.12.199>
16. Mzoughi, H., Njeh, I., Slima, M. B., Ben Hamida, A., Mhiri, C., Mahfoudh, K. B. (2020). Towards a computer aided diagnosis (CAD) for brain MRI glioblastomas tumor exploration based on a deep convolutional neuronal networks (D-CNN) architectures. *Multimedia Tools and Applications*, 80 (1), 899–919. doi: <https://doi.org/10.1007/s11042-020-09786-6>
17. Rajasree, R., Columbus, C. C., Shilaja, C. (2020). Multiscale-based multimodal image classification of brain tumor using deep learning method. *Neural Computing and Applications*, 33 (11), 5543–5553. doi: <https://doi.org/10.1007/s00521-020-05332-5>
18. Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 234–241. doi: https://doi.org/10.1007/978-3-319-24574-4_28