

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

# Brain Tumor Segmentation from MRI Images using Hybrid Convolutional Neural Networks

Dinthisrang Daimary, Mayur Bhargab Bora, Khwairakpam Amitab\*, Debdatta Kandar

*Department of Information Technology, North-Eastern Hill University, Shillong- 793022, India*

---

## Abstract

Brain tumor segmentation is a process of identifying the cancerous brain tissues and labeling them automatically based on the tumor types. Manual segmentation of tumor from brain MRI is time-consuming and error-prone. There is a need for fast and accurate brain tumor segmentation technique. Convolutional Neural Networks (CNNs) have recently shown outstanding performance in computer vision for image segmentation and classification tasks. U-Net, SegNet and ResNet18 are the most popular CNN for image segmentation. The U-Net architecture uses skip connection that captures the fine and soars information but requires higher computational time for training. SegNet is computationally efficient. The ResNet18 also uses skip connection and has a layer which adds inputs from multiple neural network layers to get more accurate results. The proposed U-SegNet and Seg-UNet is a hybridization of the novel architecture, SegNet, and U-Net. The main difference between them is the depth, Seg-UNet uses five convolution blocks compared to U-SegNet, which has three convolution blocks and both the models has a skip connection inspired from U-Net after the first convolutional layer by using a depth concatenation layer. And proposed Res-SegNet is also a hybridization of SegNet and ResNet18. It is inspired by ResNet18 and uses an element-wise addition layer as a skip connection. The SegNet3, SegNet5, U-Net, Seg-UNet, U-SegNet, and Res-SegNet are implemented to compare the performance based on their accuracy. For experimentation, BraTS dataset are used for training and testing the models. From the simulation results, it is observed that the hybrid architecture has higher accuracy.

© 2020 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).

**Keywords:** Deep learning; Hybridization of CNN; MRI; Res-SegNet; Segmentation of tumor; Seg-UNet; U-SegNet;

---

## 1. Introduction

**Glioma is a brain tumor, the tumors may also spread to the nearby region of the brain.** The glioma can be fatal and may lead to death if not treated at the early stage. According to the World Health Organization (WHO) [1], the

---

\* Corresponding author. Tel.: +91-364-272-3628;

E-mail address: [kamitab@nehu.ac.in](mailto:kamitab@nehu.ac.in)

mortality rate is higher in adult compared to children. As per the data published in cancer.net, it is estimated that about 23,820 adults and 5270 children will be diagnosed with brain tumors in the United States alone, this year.

The treatment of the brain tumor depends on the patient age, tumor type, and its location. The tumors may grow and spread to the nearby healthy tissue, causing difficulty in diagnosis and treatments. Therefore, accurate segmentation of brain tumor from the surrounding tissues is important to detect the tumors at the early stage to increase the chances of survival for the patients. In this work, the glioma is classified into peritumoral edema, necrotic & non enhancing and enhancing tumor [2], which are important factors in the treatment of the patient.

Magnetic Resonance Imaging (MRI) is commonly used imaging techniques for capturing images of brain tumors. By configuring the MRI scanner different modalities can be captured, such as T1-weighted (T1), T1-Post contrast-enhanced (T1ce), T2-weighted (T2), and T2-weighted fluid-attenuated inversion recovery (Flair). T1 is good for segmentation of tumor from healthy brain tissue. T1ce has higher visibility of the tumor boundaries. In T2 the edema (fluid) around the brain tumor is visible. And the Flair is suitable for identifying edema region from cerebrospinal fluid [3]. The MRI images can be viewed in three dimensions (Sagittal, Axial, and Coronal), which helps the medical experts in examining the tumors [4]. Enhancing tumor shows hyper-intensity mostly in T1-weighted. Non-enhancing and necrotic tumor both are core tumor, it looks hypo-intense in T1-weighted. The peritumoral edema are form in meningiomas, it is spread from the nearby tumor.

Accurate segmentation of brain tumors from MRI is a complicated and difficult task because of the complex structure and appearance of the tumors, the borders of the tumor are often fuzzy, and the tumors may spread into the nearby region of the brain, which result in difficulty to differentiate the affected tissue from the other surrounding healthy tissue. Therefore, manually identifying the boundary (delineation) of tumors in MRI images is time-consuming, and error-prone. Automatic brain tumor segmentation using MRI images would solve these problems, which will provide a fast and reliable diagnosis, by identifying the type and the exact location of the tumors. Early treatment of tumors may cure the patient.

Recently, the deep neural network is gaining popularity among the researchers and have shown outstanding performance with very high accuracy in image segmentation. CNN is a type of deep neural network, which can learn and extract features from the images. Many researchers have used CNN for automatic brain tumor segmentation in MRI images. The objective of this paper is to explore the architecture of the popular CNN for segmentation (SegNets, U-Net and Resnet18) and find the advantages of each model, and to develop hybrid architectures by inheriting the advantages of the popular CNN models. It is expected that the hybrid architecture will give a more accurate result.

The SegNet architecture can be classified into SegNet3 (consisting of three convolution blocks) and SegNet5 (consisting of five convolution blocks) [5]. The SegNet3, has two convolutional layers with  $3 \times 3$  filter, a stride of  $[1 \times 1]$  and padding of  $[1 \ 1 \ 1]$  in each of the three convolution blocks for the feature extraction from input by sliding the filter kernel and performing the convolution operation. Batch normalization layers are used after each convolutional layer for normalizing the channels of the extracted features and ReLU layers are used for converting the negative input to zero without changing its dimensions. The SegNet5 is created by using Vgg16 [6]. The Vgg16 is a CNN model which is also popular in the image classification and segmentation. It has 13 convolutional layers and batch normalization layers, ReLU layers followed by max-pooling layers and three fully connected layers, but for performing image segmentation the fully connected layers are removed and decoders (convolutional layer) corresponding to the encoders (existing convolutional layer) are included in the architecture and it is called as SegNet5 in this paper. U-Net has become one of the most popular techniques for medical image segmentation [7]. It is capable to capture fine and soar pieces of information from the encoder to the decoder using skip connection, but it requires higher computational time compared to SegNet. The skip connection passes the whole captured features to the corresponding upsampling convolution blocks in the decoder [8].

This paper focus on the hybridization of the CNN architecture, the hybrid Seg-UNet, U-SegNet [9], and Res-SegNet are presented. Seg-UNet is a fusion of a novel architecture SegNet5 and U-Net. This architecture has a property of downsampling and upsampling, mimicked from the SegNet5 and concatenation of features map size from multiple layers inspired by U-Net. The U-SegNet is identical to Seg-UNet, the main difference is the convolution block (depth), five convolution blocks are used in Seg-UNet thus it has deeper layers while the U-SegNet composed of only three convolution blocks. Res-SegNet is a hybridization of ResNet18 and SegNet5. This model has a property of downsampling and upsampling identical to SegNet5 and addition of inputs from multiples layers inspired by Resnet18. The

steps involved in the segmentation of brain tumors from the MRI images are shown in Figure 1. In preprocessing, the 2D images from the 3D MRI images are extracted. And fed to the CNN models which gives the segmented image.

The rest of the paper is organized as follows. Section 2 discusses the related existing works, Section 3 presents the proposed hybrid CNN architectures for segmenting brain tumors from MRI images, Section 4 presents the training and implementation details, Section 5 discusses the experimental result and the conclusion and future work are discussed in the final section.

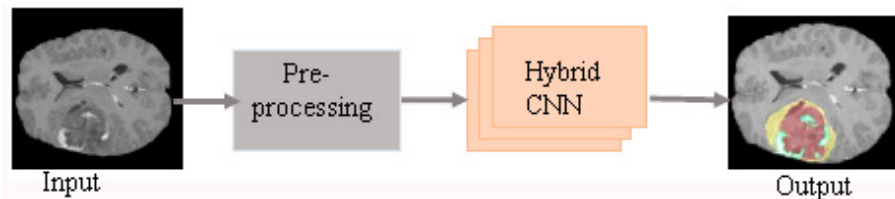


Fig. 1. Block diagram of the proposed hybrid CNN.

## 2. Related Work

There is no perfect segmentation technique for segmenting brain tumor, so continuously numerous advanced approaches have been introduced from time-to-time for automatic segmentation of the brain tumor. The contrasts of intensities, shapes, location, boundaries of the brain tissue vary from person to person, and it is the major challenges for the automatic segmentation. The state-of-the-art result of deep learning in image segmentation and classification is a signatory to overcome these problems. In this section, some of the deep learning based approaches will be discuss which are used in automatic segmentation of brain tumor in MRI using BraTS dataset.

The deep neural network architecture uses small convolution filters such as  $3 \times 3$  to maintain the longer depth of CNN, and learn more features from inputs in each learnable layer of the network. One of this kind was proposed by Pereira (2016) [11], it has three convolution blocks (11 layers of depth) which consist of 6 convolutional layers with  $3 \times 3$  filters, two max-pooling layers and followed by three fully-connected layers. Before training the network authors have performed preprocessing to equalize intensity across all the images by normalizing and also performed filtering of noise by computing the standard deviation and the mean intensity value across all training images. The proposed model obtained an accuracy of 88% for the whole tumor, 83% for core tumor and 77% for an active tumor in the BraTS dataset.

Urban, G. proposed a novel three-dimensional convolutional neural network for automatic brain tumor segmentation in multi-modality MRI images (2014)[12]. It demands high computational time, but the 3D visualization makes radiologists easy to understand the development of the tumor. The  $3 \times 3 \times 3$  convolution filters are used to reduce the size of the feature map along with the batch normalization layers, ReLU, and 3D max-pooling layers. The 3D inputs are stacked into a 4D volume, the four dimensions represent the height, width, channels of the images and the number of modalities. This architecture has achieved an accuracy of 87%, 77%, and 73% for the whole tumor, core tumor, and active tumor region respectively in the BraTS dataset.

Salma X sun (2019) [13] presented a SegNet for automated brain tumor segmentation by training all the modalities separately and combining the output of SegNet at post-processing. Firstly, the inputs are preprocessed to remove unwanted artifacts by normalizing and bias field correction, which increases the performance in segmentation. SegNet is used to train the four different MRI modalities separately. The architecture consists of a pair of the encoder (down-sampling) and decoder (upsampling). 13 convolutional layers with  $3 \times 3$  filters, batch normalization layers, ReLU layers and follow by max-pooling layers with  $2 \times 2$  filter are used in the encoder. The decoder also has 13 convolutional layers to match with the corresponding encoder. High dimensionality features, extracted in the decoder is fed to softmax layers to classify each class pixel separately. The segmentation technique achieved accuracy of 85% for the whole tumor, 81% for core tumor, and 79% for enhancing tumor.

Patch-based CNN approach utilizes the inherent properties of CNN for pattern recognition and also performs highly accurate segmentation in MRI. Hussain S have proposed a cascaded two pathway CNN Model (2017) [14], it extracts

the large patch size of  $37 \times 37$  and smaller patch size of  $19 \times 19$  at the same time. The architecture has a lot of learnable parameters, which may result in overfitting during the training, to avoid overfitting used maxout and dropout layers in the architecture. The model consists of 6 convolutional layers with a different filter size that enable the CNN model to learn the feature of different size and uses ReLU. The network was trained end to end by cascading the output of the first CNN and the input of the second CNN. The 3D slicer toolkit was applied on MRI images for bias field correction to similarized the artifact in preprocessing, which contributed to better performance in segmentation. The model achieved an accuracy of 80%, 67% and 85% for complete, core and enhancing tumor respectively in BraTS dataset.

### 3. Proposed Approaches

The objective of this work is to combine the popular deep CNN models for the automatic segmentation of tumors in the brain MRI images, by inheriting the advantages of each model. It is expected that the hybrid models will give a higher segmentation accuracy. The following subsection presents the proposed hybrid architecture.

#### 3.1. The Seg-UNet Architecture

The Seg-UNet is a hybridization of the novel SegNet5 and U-Net architecture, which are widely used for segmentation of the image. The SegNet5 consists of a pair of an encoder (for downsampling path) and decoder (for upsampling path). The encoder is similar to vgg16 (13 convolutional layers) and the decoder is similar to the inverse of the vgg16 [6], but the SegNet5 has one more convolutional layer in each convolution block. It is included to match the feature map size of the third dimension. SegNet5 is modified by adding a skip connection in the first convolution blocks inspired by U-Net [7]. The skip connection recaptured the features which are already captured in the corresponding encoder, for the reconstruction of the images from the loss of local information and to consolidate finer contextual information at the upsampling layer of the decoder. While downsampling some of the information is lost in each max-pooling layer. The skip connection is implemented by using the depth concatenation layer. The architecture of the proposed Seg-UNet is presented in Figure 2.

The Seg-UNet takes input of size  $240 \times 240 \times 3$ , 64 convolution filters of size  $3 \times 3$  are used to convolve across the input with a stride [1 1] and padding [1 1 1], to extract feature map of size  $240 \times 240 \times 64$ . Followed by the batch normalization layer to normalized the third dimension and ReLU to convert the negative value to zeros. The first convolution block also has one more convolutional layer, normalization layer and ReLU layer with the same parameters. Followed by a max-pooling layer of  $2 \times 2$  window and stride of 2 to downsample the feature map size to  $120 \times 120 \times 64$ . In the second convolution block, convolutional layers of the same filter size are used on the output of the previous layer and reduce the third dimension to 128, while the first two dimensions remain unchanged  $120 \times 120 \times 128$  and the max-pooling layer is used to reduce the first two dimensions to  $60 \times 60 \times 128$ . In third convolution blocks, features becomes more complex  $60 \times 60$ , furthermore, it is downsampled to  $30 \times 30 \times 256$  by max pooling. In the fourth convolution blocks, the feature map size is reduced to half and then downsample to  $15 \times 15 \times 512$ . The fifth convolution block also applies the same filter size and generates the feature map of size  $15 \times 15 \times 512$  and downsampled to  $7 \times 7 \times 512$ . It observed that the last dimension (depth) cannot be increased anymore, as the feature becomes very small and difficult to differentiate among them. Therefore, going deeper may not be useful and it will lead to unnecessarily higher computational time.

An upsampling path or decoder scales up to the feature map dimension of lower resolution to a higher resolution to get back the original input resolution, this will preserve correct boundary delineation of the tumors. The decoder has a symmetrical structure of the encoder path, except the max-pooling operation is replaced by an un-pooling operation. The un-pooling layer takes outputs of the previous layer as input and the indices of the corresponding max-pooling encoder. The output of the final decoder, which has huge dimensional feature representation, is fed into a softmax layer, which classifies each pixel separately. Subsequently, the output of the softmax layer is a 4 channel image, where 4 channels represent the number of desired classes; necrotic & none enhancing, peritumoral edema, enhancing tumor, and others (everything else).

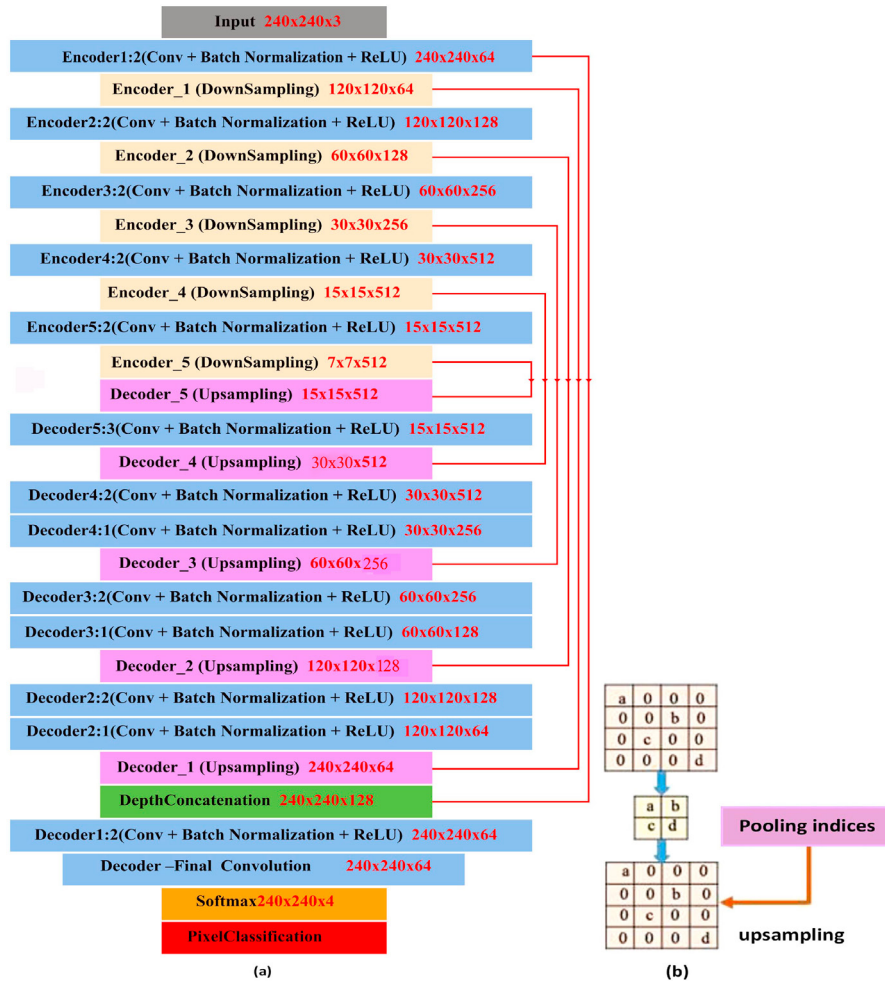


Fig. 2. Architecture of Seg-UNet, is a hybrid CNN re-designed using high tune information captured by a skip (depth concatenation) connection in the novel architecture SegNet5 and U-Net.

### 3.2. The Res-SegNet Architecture

The Res-SegNet is a combination of ResNet18 and SegNet5. A skip connection is added in SegNet5, inspired by ResNet18 which is also widely used for image segmentation. The skip connection of the ResNet18 is used for re-capturing the information at the upsampling path from already captured information in the corresponding encoder by element-wise addition. This model has the advantages of relatively less training time and high accuracy. The skip connection in the Resnet18 uses an additional layer which adds inputs from multiple layer element-wise.

The five convolution blocks of SegNet5 are used for the initialization of weight and each convolution block of the encoder composed of 2 convolutional layers, batch normalization layers, ReLU layers, and max-pooling layer. Each convolution block of the decoder composed of an un-pooling layer and followed by 3 convolutional layers, batch normalization layers, and ReLU layers. Max-pooling reduces the size of the features and retains the contextual information.  $3 \times 3$  convolution filter and stride of size [1 1] with padding [1 1 1 1] are used in all the convolutional layers. The convolutional layers acquire the weights of  $3 \times 3 \times 3 \times 64$  and  $3 \times 3 \times 64 \times 64$  at first convolution block and continuously learns weight till the fifth convolution block where learnable parameters acquired the weight of  $3 \times 3 \times 512 \times 512$  at last layer of downsampling. The max-pooling layer decreases the spatial dimensions of the feature



which are extracted in each convolutional layer by convolving with the  $3 \times 3$  filter on the input. The output of the encoder needs to be upsampled, to obtain a reasonably high-resolution representation of the features.

The upsampling path increased the first two dimensions of the feature size in each un-pooling layer using pooling indices. The last convolutional layers of each convolution block decrease the third dimension. All convolutional layers learn weight according to their filter size, the number of filters and the number of the channel of the previous layer. Upsampling is required for preserving the boundary depiction information, because sometimes the tumors may be very small which may be lost during downsampling. The use of a skip connection in SegNet5 can avoid this problem and improved the performance of segmentation. The architecture of the proposed Res-SegNet is presented in Figure 3.

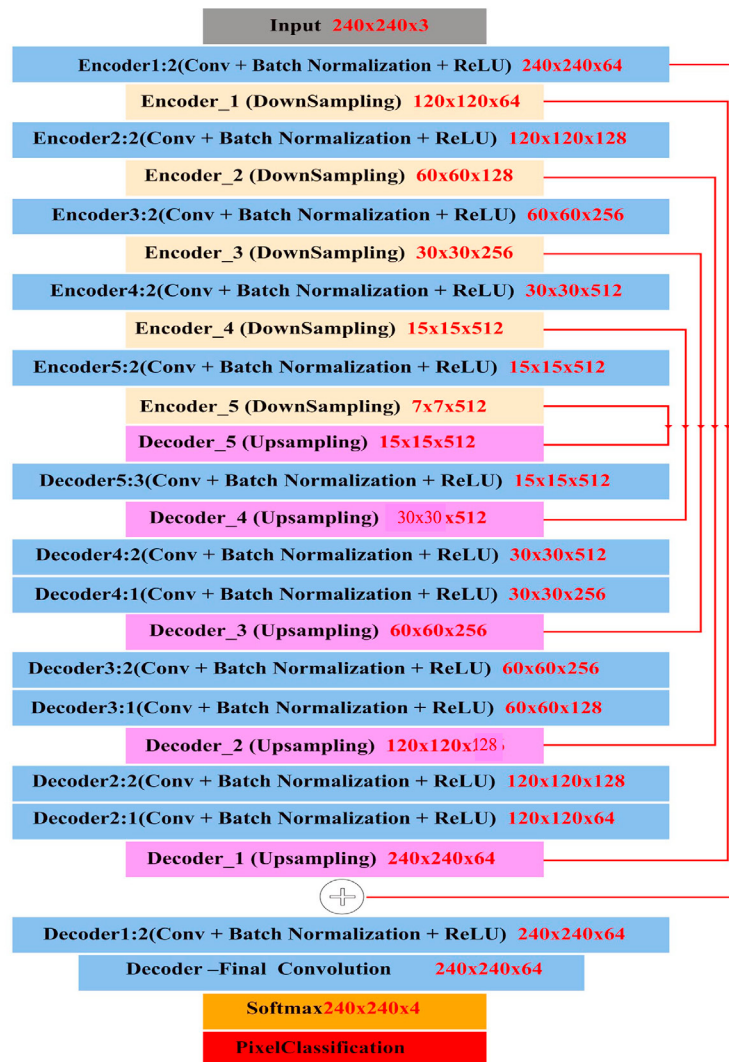


Fig. 3. Architecture of Res-SegNet, is a hybrid CNN re-designed using high tune information captured by a skip (addition layer) connection in the novel architecture SegNet5 and Res-Net18.

### 3.3. The U-SegNet Architecture

The U-SegNet was proposed by Pulkit Kumar [9], it consists of 3 convolution block, is a combination of SegNet3 and U-Net. The SegNet3 is hybridized by using a skip connection at the first convolution block, inspired by U-Net architecture. The authors used a  $1 \times 1$  convolution filter in the last convolutional layer to find out the small feature present in the images. The segmentation method obtained a mean accuracy of 89.74% in the BraTS dataset. The network was designed to segment the MRI images into white matter, gray matter and cerebrospinal fluid. The original U-SegNet uses  $40 \times 40 \times 3$ , in this paper the input layer is modified by changing it to  $240 \times 240 \times 3$  to make it uniform with the other proposed architecture for evaluation and the last convolutional layer is also modified by using  $3 \times 3$  kernel in order to reduce the feature size instead of  $1 \times 1$  in the original model. The modified U-SegNet is implemented to segment the MRI images into 4 classes.

## 4. Training and Implementation details

The BraTS dataset used for training and testing the CNN models is collected from the CBICA portal, it contains both the MRI images and the corresponding segmented output images, the output or the ground truth labeled images are revised by clinically expert neuro-radiologists. In the dataset, the multimodal glioblastoma MRI scan is divided into two categories; high graded glioma and low-grade glioma. The BraTS datasets are 3D volumetric nifty formats with dimensions of  $240 \times 240 \times 154$  [15], the first two dimensions represent the height and width of the images, and the third dimension represents the number of channels or slice. The datasets contain images obtained by four different MRI scan modalities; T1, T1ce, T2 and Flair. A sample of MRI image obtained by different MRI modality is shown in Figure 4.

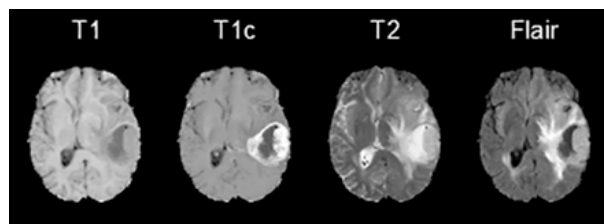


Fig. 4. Axial view of T1, T1ce, T2 and Flair.

After analyzing the dataset it is observed that the axial planes have more visibility of tumor boundaries and diffusion to other regions of the brain which will contribute to better segmentation of the tumor. Therefore, the 3D MRI scan is extracted into 154 numbers of the 2D slice along the axial plane by taking 16 bits depth. The images are pre-processed by normalizing the intensity value. The dataset consists of 775 T1ce images, extracted slice by slice from five HGG patients. 60 % of the dataset is used for training and the rest 40 % is used for testing. The training images are rotated at different angles using the data augmentation which helps in more generalization and improves the accuracy.

The CNN models were trained on a single CPU. The same training parameters used for all the models to comparatively evaluate their performance. The stochastic gradient descent with momentum algorithm [16] is used for training all the models. The training parameters; momentum, initial learning rate, L2Regularization, maximum epochs, mini-batch size are set to 0.9, 0.0001, 0.0001, 80 and 16 respectively. The momentum is the contribution of the gradient from the previous iteration to the current iteration. The low initial learning rate will give an optimal solution but require longer training time, whereas the higher initial learning rate will give sub-optimal results. To avoid overfitting L2Regularization is used. Larger epoch minimizes the error and will give better segmentation results. A mini-batch is a subset of the training set, mini-batch with 16 is suitable for training in a single CPU, increasing the mini-batch size may give better results, but requires higher memory.

Besides the proposed models the two different kinds of SegNet and U-Net are also implemented. The main difference between SegNet3 and SegNet5 is the depth and the feature map size. The weights of convolutional layers in SegNet5 are initialized using VGG-16. This is deep enough to capture high-level features in the last layer of the

encoder, results in dense feature maps. The encoder upsamples the feature map using pooling indices. While the SegNet3 has a thinner feature map. U-Net uses the  $3 \times 3$  convolution filter in all the convolutional layers, except the last convolutional layers which use a  $1 \times 1$  filter.

## 5. Results and Discussion

The proposed hybrid architectures (Seg-UNet, Res-SegNet, and U-SegNet) are implemented and analyzed the segmentation capability in terms of accuracy by comparing with the popular CNN models for image segmentation namely SegNet3, SegNet5, and U-Net. Figure 5 shows a sample of ground truth images and the corresponding predicted output by the CNN models. The green, red, yellow and gray colors represent the enhancing tumor, necrotic & none enhancing tumor, peritumoral edema and anything else respectively. All models accept 172,800 neurons as the input in the input layers and pass them through the several hidden layers. These neurons are classified into four classes in the output layer.

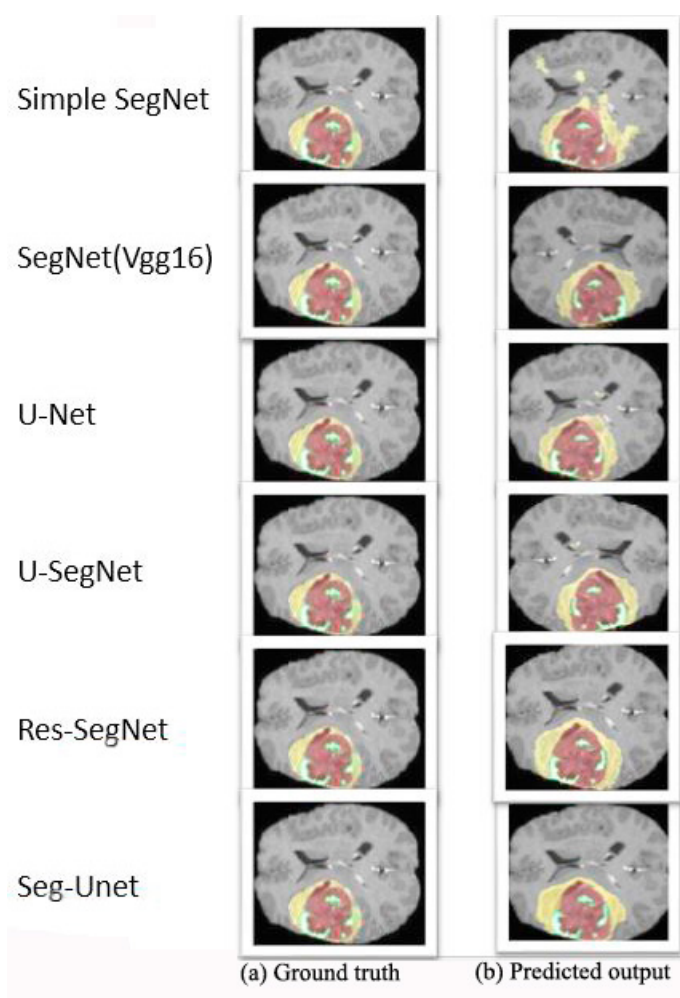


Fig. 5. Ground Truth vs Predicted Image

To evaluate how well the model performs in segmentation tasks, five parameters are considered, these are global accuracy, mean accuracy, mean Intersection Over Union (IOU), weighted IOU and mean BF-scores. Global accuracy represents the ratio of the highest correctly classified pixels of one class regardless of all of the classes, to the total



number of pixels, while mean accuracy indicates the averages percentage of correctly identified pixels for each class. IOU also known as the Jaccard similarity coefficient, the mean IOU is the averaged IOU of each class. IOU is defined in Equation 1.

$$IOU\ score = \frac{TP}{TP + FP + FN} \quad (1)$$

Where, TP, FP, and FN are the true positives, false positives and false negatives respectively. The weighted IOU indicates the number of pixels of each class that is weighted in the disproportion pixel class to reduce the bigger class overlapping on the smaller class. The mean BF-scores indicate the alignment of the predicted boundary and boundary in the ground truth. The performance of all the models are presented chronologically in the Table 1. From

Table 1. Segmentation performance of the CNN models.

Model name	Global accuracy	Mean accuracy	Mean IOU	Weighted IOU	Mean BF-score
SegNet3	0.97628	0.89327	0.53646	0.95859	0.77267
SegNet5	0.98194	0.91787	0.60213	0.98567	0.64461
U-Net	0.98085	0.90425	0.59213	0.97567	0.63499
U-SegNet	0.9824	0.91689	0.64791	0.98221	0.8451
Res-SegNet	0.98854	0.93352	0.68914	0.98293	0.82147
Seg-UNet	0.99117	0.93124	0.73409	0.986357	0.85078

the simulation result presented in Table 1, it is observed that the hybrid architectures perform better than the popular CNN models in all the performance measures. U-SegNet performed better than SegNet3 and U-Net of depth 3, which are used in the hybridization of U-SegNet. Res-SegNet performed comparably better than SegNet5 and U-Net of depth 3 that are used in the hybridization of Res-SegNet. The Seg-UNet achieved better performance in segmentation than SegNet5 and U-Net of depth 3, which are used to hybridize the CNN using a skip connection. It is observed that even though Res-SegNet has a higher mean accuracy than Seg-UNet, it has lower global accuracy, mean IOU, weighted IOU and mean BF-score compared to Seg-UNet. The Seg-UNet overall performs better than all the other models and gives good segmented output with well boundary alignment of each class.

The training parameters are the key factor for studying the computational time of CNN. Therefore setting up all the training parameters same for all the models and using the same dataset is important. The SegNet3 takes less time to train because it has fewer layers compared to others, and completed the training in 530 min. The SegNet5 took 1339 min and 48 sec to complete the training, it takes more time compared to SegNet3 as the architecture has more layers, but has better accuracy. The U-Net took 730 min and 14 sec to train, the U-Net and SegNet3 have similar depth, but the U-Net requires more time compared to SegNet3 due to the use of skip connection, but it gives a better segmentation result. The hybrid architectures U-SegNet, Res-SegNet, and Seg-UNet took 741 min and 26 sec, 1690 min and 28 sec and 2694 min and 48 sec, respectively. The hybrid architectures require more time to train as it has the number of layers and training parameters, but it gives a better segmentation result. Once the network is trained, the network can be used for segmentation, segmentation of an image by using the trained model takes just a few seconds. Whereas, manual segmentation of tumors by clinical experts may take hours. The proposed image segmentation techniques are accurate, fast and can be implemented at a low cost. This will help the doctors in the diagnosis of brain tumor fast and accurate, which may save the lives of many patients.

## 6. Conclusion and future work

This paper presents three hybrid CNN models, namely U-SegNet, Res-SegNet, and Seg-UNet designed for reliable automatic segmentation of the brain tumor from the MRI images with high accuracy. The proposed models inherit the properties of SegNet, U-Net, and ResNet, which are the most popular CNN models for semantic segmentation. Sometimes brain tumors of small size are lost during downsampling which causes inappropriate segmentation. The hybrid models can overcome such problem by adding skip connection in the SegNet, which is inherited from U-Net and Res-Net18. The CNN models are trained and validated by using the BraTS dataset. The segmentation accuracy

is evaluated by considering global accuracy, mean accuracy, mean IOU, weighted IOU and mean BF-score. From the experimentation result, it is found that the proposed hybrid architectures achieved more accurate output compared to the other existing CNN models. The U-SegNet, Res-SegNet and Seg-UNet achieved mean accuracy of 91.6%, 93.3% and 93.1% respectively. The hybrid architectures are associated with the higher number of layers and trainable parameters, therefore more time is required for training, but once the network is trained. The system can automatically segment the brain tumors from the MRI images in few seconds. In the future, the proposed hybrid models will be improved by using different sizes of filters and all the modalities of the MRI images will be considered in segmentation of the tumors, the segmentation result will be further enhanced by increasing the mini-batch size from 16 to 64 and max-epoch from 80 to 120.

## References

- [1] Louis, David N and Perry, Arie and Reifenberger, Guido and Von Deimling, Andreas and Figarella-Branger, Dominique and Cavenee, Webster K and Ohgaki, Hiroko and Wiestler, Otmar D and Kleihues, Paul and Ellison, David W. (2016) "The 2016 World Health Organization classification of tumors of the central nervous system: a summary." *Acta neuropathologica* **131** (6): 803–820.
- [2] Menze, Bjoern H., et al. (2014) "The multimodal brain tumor image segmentation benchmark (BRATS)." *IEEE transactions on medical imaging* **34** (10): 1993–2024.
- [3] Bauer, Stefan and Wiest, Roland and Nolte, Lutz-P and Reyes, Mauricio. (2013) "A survey of MRI-based medical image analysis for brain tumor studies." *Physics in Medicine & Biology* **58** (13): R97.
- [4] Juan-Albarracín, Javier and Fuster-García, Elies and Manjon, Jose V and Robles, Montserrat and Aparici, F and Martí-Bonmatí, L and García-Gómez, Juan M. (2015) "Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification." *PLoS One* **10** (5): e0125143.
- [5] Badrinarayanan, Vijay and Kendall, Alex and Cipolla, Roberto. (2017) "SegNet: A deep convolutional encoder-decoder architecture for image segmentation." *IEEE transactions on pattern analysis and machine intelligence* **39** (12): 2481–2495.
- [6] Simonyan, Karen and Zisserman, Andrew. (2014) "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409*: 1556
- [7] Ronneberger, Olaf and Fischer, Philipp and Brox, Thomas. (2015) "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention* **234**: 241.
- [8] Drozdal, Michal and Vorontsov, Eugene and Chartrand, Gabriel and Kadoury, Samuel and Pal, Chris. (2016) "The importance of skip connections in biomedical image segmentation." *Deep Learning and Data Labeling for Medical Applications* **179**: 187.
- [9] Kumar, Pulkit and Nagar, Pravin and Arora, Chetan and Gupta, Anubha. (2018) "U-SegNet: Fully Convolutional Neural Network Based Automated Brain Tissue Segmentation Tool." *2018 25th IEEE International Conference on Image Processing (ICIP)* **3503**: 3507.
- [10] Işın, Ali and Direkçioğlu, Cem and Şah, Melike. (2016) "Review of MRI-based brain tumor image segmentation using deep learning methods." *Procedia Computer Science* **102**: 317–324.
- [11] Pereira, Sérgio and Pinto, Adriano and Alves, Victor and Silva, Carlos A. (2016) "Brain tumor segmentation using convolutional neural networks in MRI images." *IEEE transactions on medical imaging* **35**(5): 1240–1251.
- [12] Urban, Gregor and Bendszus, M and Hamprecht, F and Kleesiek, J (2014), "Multi-modal brain tumor segmentation using deep convolutional neural networks", *MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings, winning contribution* **31**: 35.
- [13] Alqazzaz, Salma and Sun, Xianfang and Yang, Xin and Nokes. (2019) "Automated brain tumor segmentation on multi-modal MR image using SegNet." *Len, Computational Visual Media* **1**: 11.
- [14] Hussain, Saddam and Anwar, Syed Muhammad and Majid, Muhammad. (2017) "Brain tumor segmentation using cascaded deep convolutional neural network." *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* **1998**: 2001.
- [15] Menze, Bjoern H and Jakab, Andras and Bauer, Stefan and Kalpathy-Cramer, Jayashree and Farahani, Keyvan and Kirby, Justin and Burren, Yuliya and Porz, Nicole and Slotboom, Johannes and Wiest, Roland and others. (2014) "The multimodal brain tumor image segmentation benchmark (BRATS)." *IEEE transactions on medical imaging* **34**(10): 1993–2024.
- [16] Bottou, Léon. (2010) "Large-scale machine learning with stochastic gradient descent." *Proceedings of COMPSTAT'2010* **177**: 186.