# Chapter 3

## Automatic Segmentation of Brain Tumour in Magnetic Resonance Images using an Enhanced Deep Learning Approach

***Highlights of the Chapter***

- *The present chapter proposes an automatic method for segmentation of brain tumour in MR images.*
- *Deep learning method based on cross-channel normalization along with residual connections.*

***Contribution of the chapter***

This chapter presents a deep learning method to quantify the tumour region in brain Magnetic Resonance images as the accurate diagnosis of brain tumour region is necessary for the treatment of patients. The irregular and confusing boundaries of tumours regions make it a challenging task to accurately figure out such regions. Another challenge with the segmentation task is of preserving the boundary details of the segmented tumour regions. The proposed network focuses on delineating the irregular brain tumour region as the best feature maps are learnt by the network, which is used for decoding; thus, it preserves the accurate boundary and pixel details. The proposed method incorporates internal residual connections in encoder and decoder to transfer the feature maps directly to the successive layers to avoid the loss of information contained in the images. The use of cross channel normalization (CCN) and parametric rectified linear unit (PRELU) gives a more balanced network output. The qualitative image statistics of our proposed network are closest to the expected results. The trained network produced remarkable results when tested on images of other datasets. Further, external clinical

validation was performed by comparison of the algorithmic segmented images with those generated by a manual segmentation done by an experienced radiologist. We have termed our network as CCN-PR-Seg-net.

### 3.1 Introduction

The undesired growth of cells causes a lump in the brain which is termed as brain tumour because the cells divide and grow in uncontrolled manner. The tumour in brain can be cancerous or non-cancerous which are termed as malignant and benign respectively. When any of these type of tumour grow, it causes pressure inside the skull of brain as the skull of the brain is rigid and it encloses whole brain. This pressure might cause the damage in brain and can also cause threat to life. The categorization of brain tumours is done as primary or secondary. The tumour which originates in the brain is known as primary tumour while the tumour which originates because of cancer cells which spread to brain from other organs of body such as lung or breast etc.

Brain tumour is a severe healthcare issue now a days and can lead to the death of patient. Magnetic resonance imaging (MRI) has wide application in the diagnosis of the patient affected with brain tumour but due to lack of effective means to process large data volume produced by the image acquisition system, the full exploitation of such imaging techniques is not possible. Magnetic Resonance Imaging is non-invasive in vivo imaging technique in which the target tissues are excited by radio frequency signals to produce the internal images of their structure under the influence of strong magnetic field [83]. Distinct MRI modalities are produced by changing the excitation and repetition times during the image acquisition process. Tissue contrast images of different types are produced by different MRI modalities which provides valuable information of structure for the purpose of accurate segmentation of tumour and correct diagnosis [84].

The challenging issue in such cases is to precisely depict the region of interest (ROI) i.e. the tumorous part from the healthy tissues of the brain so that any therapy could be applied to healthy tissues while damaging the cancerous tissues. The accurate segmentation of tumour from normal part of brain is of utmost importance as an inaccurate segmentation may lead to spurious diagnosis which can affect the process of treatment. The size, shape and location of tumour in the tumour bearing brain MRI data differs from one patient to another. The tumour boundaries are unclear having irregular discontinuities. Thus it poses a great challenge to accurately figure out the boundaries of the tumour and minimising the errors. There are many segmentation techniques evolved over a period of time varying in use and accuracy of the algorithm[85] . The algorithms used for segmentation can be categorized as manual, semi-automatic and automatic algorithms. The segmentation methods incorporating manual procedures requires a radiologist with anatomical and physiological knowledge gained through training and experience to investigate the multimodal information presented by MR images. These procedures are radiologist dependent and results are subject to greater inter and intra rater variability. In semi-automatic methods the user interaction is needed for the purpose of initialization, intervention and evaluation. The process of initialization is carried out by defining an approximate area which consists of tumour. This can further be modified using the feedbacks received about the region and then can steer towards the final result. Finally, the user can evaluate the result and can amend it in case not satisfied with the results. Tumour cut method proposed by Hamamci et al. is an example of semi-automatic algorithm in which user draws a maximum diameter first and then a cellular automata based seeded tumour segmentation method run twice [86].

The automatic algorithm fully eliminates the need of user intervention as artificial intelligence is used in these methods to supervise the segmentation process [87] [88]. Many deep learning based network architectures for segmentation are presented over the years such as CNN [89],

VGGNet [90], Alexnet [66] and Googlenet [65]. These methods are capable of extracting the feature maps on their own and hence no handcrafted features are required in the form of manual inputs. A network using the sliding window setup to predict the class label of each pixel by providing a local patch was developed by Dan C. Ciresan [91]. The Fully connected networks (FCN) has wide application used the field of biological image segmentation and provides fine results as compared with conventional methods [92] [93]. A more elegant architecture based on FCN was presented by Olaf Ronneberger et al. for the purpose of biomedical image segmentation [30]. Seg-Net is another extraordinary deep learning architecture used for pixel-wise segmentation of images of different types [94]. To enhance the quality of segmentation we introduced an new method based on encoder and decoder approach to segment the biomedical images of different modalities for the purpose of getting accurate diagnosis. Our proposed network can yield better precise segmentation results while running on very few images. Moreover our proposed model is able to delineate the uneven boundaries of the brain tumour more precisely and accurately. In addition, the proposed model is trained once and tested on other datasets gives finer results thus avoiding retraining of the network.

### 3.2 Methodology

To segment MR images more efficiently we proposed some modifications to the seg-net network and the results obtained were remarkable. Figure 3.1 shows the architecture of our proposed network model consisting of an encoder and decoder network [95]. We have introduced cross channel normalization with modified parameters. Moreover, short residual connections [96] are introduced internally in both encoder and decoder part at each set of convolution and cross channel normalization layer [97]. Parametric RELU (PRELU) [98] with value of $\alpha = 0.05$ is used instead to normal RELU. In our proposed model every encoder layer corresponds to a decoder layer and thus the number of layers in encoder and decoder networks are same [99]. Initially, the purpose of convolution operation is to obtain the feature maps.

Then, the normalization is applied, in our case we have used cross channel normalization with tuned parameters. The normalized data is then fed to a parametric RELU layer with the value of $\alpha = 0.05$. Following the Parametric RELU layer, a max-pooling layer is used with a window of size $2\times2$ and stride 2, having non overlapping window. The purpose of using the max pooling layer to retain the important information and discard all unwanted information as in case of seg-net. Moreover, max pooling allows to get translational invariance over small spatial shifts.

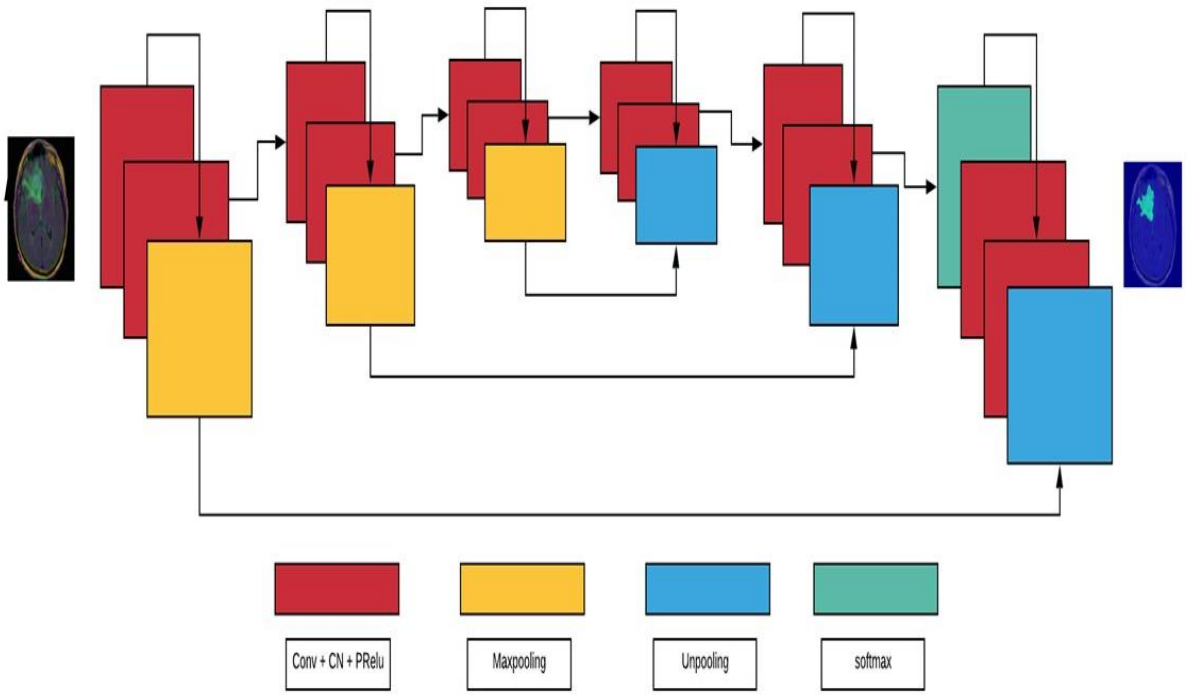


***Figure 3.1: The architecture of CCN-PR-Seg-net segmentation network***

The convolution operation converts an image of high resolution to a low resolution image by performing the down sampling. The receptive field is increased in this process i.e. the information of 'what' is present in the image but it losses the information of 'where' it is present. The information of boundary of the object starts getting lost at every convolution operation. This is not at all desirable when the task of image segmentation is to be performed where the boundary delineation is of utmost importance as in the case of segmentation of brain

tumour, the exact shape of the tumour matters a lot for the diagnostic purpose to get the correct treatment. Thus the knowledge of boundary details is most important in case of segmentation. For this purpose, the value of the location of the maximum feature value i.e. the indices of max-pooling is stored and while decoding, the decoder up-samples the input feature map provided to it using the stored max-pooling indices. For every encoder a corresponding decoder is present in the network having the same order i.e., the first encoder will correspond to the first decoder. After successive sets of decoding, the data is presented to a softmax layer prior to the classification layer. The softmax layer applies [100], softmax function to the input data which turns the raw score i.e. logits in to probability which sums to 1.Softmax function can be interpreted as multi class sigmoid i.e. they can be used for the determination of probability at once of multiple classes [101]. The activation function of softmax is by equation 3.1:

$$Y_r(x) = \frac{\exp(a_r(x))}{\sum_{j=1}^{k} \exp\left(a_j(x)\right)} \quad (3.1)$$

Where r denotes the class $0 \leq Y_r \leq 1$ and $\sum_{j=1}^{k} Y_j = 1$. Y and x are the outputs and the inputs of the layer respectively.

The softmax performs pixelwise classification and the output of the softmax layer is an image of 'n' channels where the number of classes are represented by 'n' in the image. Finally, a categorical label is provide to each pixel in the image by the pixel classification layer. The cross entropy loss function is used for optimising the parameters of the network. In a minibatch, the loss is summed up over all the pixels [102]. For loss calculation following equation 3.2 is used:

$$loss = -\frac{1}{N}\sum_{i=1}^{M} T_i \log(X_i) \quad (3.2)$$

Where $X_i$ is the response of network, $T_i$ is the target value, M is the total number of responses in the image and N is the total number of response in X.

In biomedical images, it is desired to find the exact boundaries of the tumour for the diagnostic purposes so determination of proper region of tumour plays a significant part in the segmentation of images having tumour. For this purpose, we have proposed our model that automatically segments the brain tumour and the produces the results with higher similarity to the ground truth data as compared with other methods. We trained our proposed model using cross channel normalization. In cross channel normalization, $a_{x,y}^{i}$denotes the activity of a neuron computed by applying kernel i at position (x, y) and then applying the RELU, the response-normalized activity $b_{x,y}^{i}$ [97] is given by the equation 3.3:

$$b_i = \frac{a_i}{\left(k + \nu \sum_{j=\max\left(0,i-\frac{n}{2}\right)}^{\min\left(N-1,i+\frac{n}{2}\right)} \left(a_{x,y}^{j}\right)^2\right)^{\mu}} \tag{3.3}$$

where the sum runs over n "adjacent" kernel maps at the same spatial position, and N is the total number of kernels in the layer. k, ν, μ are the hyperparameters of normalization and their values are determined using validation set. *K* is used to avoid any singularities (division by zero), $\nu$ is normalization constant, while $\mu$ is contrast constant. We used the default values of the hyperparameters of the layer except normalization constant $\nu$. The value of hyperparameter $\nu$ was selected experimentally by tuning it to a range of values.

The cross channel normalization was applied with hyperparameters taken as: k = 2, ν = 0.1, μ = 0.75. In the proposed network, parametric RELU (α = 0.05) is used in conjunction with cross channel normalization. The parametric RELU (α = 0.05) prevents the problem of dying RELU and produces some output for negative inputs thus it also removes the issues of vanishing gradient. The value of α = 0.05 was selected to have a considerable range for negative input values so that the network should not saturate for negative inputs.

### 3.3 Clinical external validation of the segmented data and acquisition of dataset

The proposed network for segmentation got trained on the dataset obtained from 'The cancer imaging archive'. For validating the effectiveness of the proposed network, the trained network

was tested on the dataset acquired under different environment. The results obtained will illustrate the comprehensive capability of the network. Moreover, manual segmentation of the dataset under evaluation was performed subsequent to the careful visual inspection of the images by a senior radiologist having experience of more than ten years in clinical reporting of brain tumour. This was done for the purpose of external validation from the clinical application point of view. FLAIR images were used for the purpose of the identification of the viable tumoral component was done using McDonald's criteria. The outline of the most solid components was included within the region of segmentation with the exclusion of any normal brain parenchyma projecting within a 2-D slice from any aspect of 3-D tumoral anatomy. The variable intensity components areas of necrosis cystic change and calcification were however included with in the region of segmentation. The segmented image hence attained was compared to the various automatically segmented images, including that obtained from our proposed network. The evaluation of segmentation performance was done by comparing the proposed method with the other two well-established methods.

### 3.4 Dataset Acquisition

The dataset used in this paper is obtained from 'The Cancer Imaging Archive'. This dataset is sponsored by 'National Cancer Institute' and the images corresponds to the TCGA lower grade glioma collection with at least fluid-attenuated inversion recovery (FLAIR) sequence [103]. The dataset consists of the data of 110 patients with all sequence available. The dataset of registered images along with the corresponding masks is available at Kaggle website. The second dataset used is acquired from figshare website [104]. In this dataset slices of brain contrast-enhanced MRI (CE-MRI) with a large slice gap are acquired. The dataset of brain T1-weighted CE-MRI images consists of 3064 slices from 233 patients, including 708 meningiomas, 1426 gliomas, and 930 pituitary tumours, which are publicly available.

### 3.5 Results:

#### 3.5.1 Results obtained when network was trained and tested on the Kaggle data

##### 3.5.1.1 Results obtained from 50 epochs of training

Figure 3.2 shows the results obtained when the network was trained on 50 epochs. The value of evaluation parameters obtained are summarised in Table 3-1.The segmented results of the datasets were compared using the segmentation evaluation metrics such as Global accuracy, mean accuracy, mean intersection over union(mIoU), mean BF score. The analysis of the results obtained depicts the superiority of our proposed model over the existing architecture. The mIoU of our proposed model is highest (nearly 2.5% higher than the second best) among the networks being compared which shows that the overlapping area of the image obtained by our network is the larger with the ground truth image as compared with the other existing methods. The BF score for predicting the boundary accuracy obtains maximum value in our proposed network. Thus our model defines the segmented boundaries more accurately than the networks being compared.

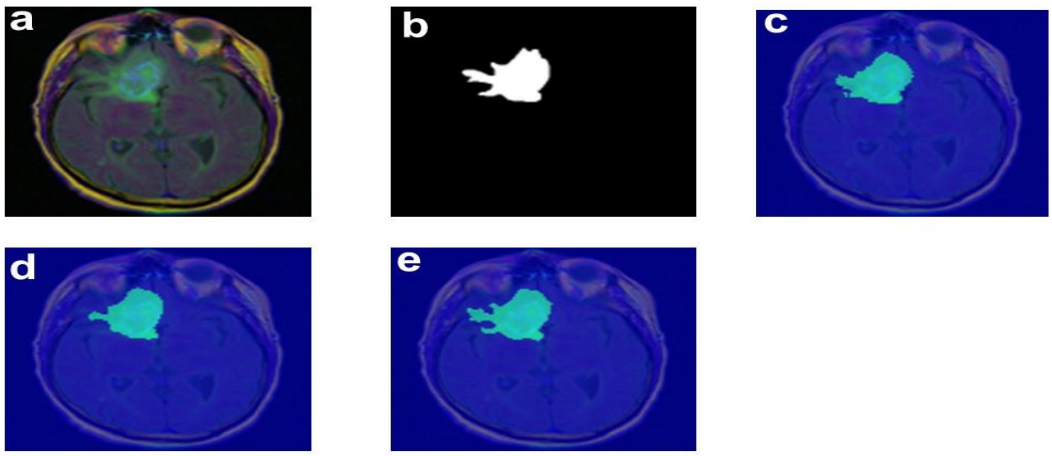


***Figure 3.2: Segmentation results when network was trained on Kaggle dataset and tested on same dataset (a) original image, (b) segmentation mask, (c) seg-net output, (d) u-net output, (e) CCN-PR-seg-net output for 50 epochs of training.***

*Table 3-1: Parameter comparison of the networks for 50 epochs of training*

| Networks | Global Accuracy | Mean Accuracy | mIoU | Mean BF Score |
|---|---|---|---|---|
| Seg-net | 0.99768 | 0.97165 | 0.9527 | 0.91358 |
| U-Net | 0.99655 | 0.95736 | 0.93101 | 0.89001 |
| CCN-PR-Seg-net | 0.9986 | 0.99082 | 0.97652 | 0.95272 |

### 3.5.1.2 Results obtained from 100 epochs of training

The results obtained when the network was trained on 100 epochs are shown in Figure 3.3.The values of evaluation metrics obtained by the networks being compared for the segmentation results are summarized in Table 3-2.It is evident that after 100 epochs of training the results obtained by the different networks tends to get more finer. On comparing the results we infer that the mIoU of our proposed network is maximum and the value of BF score also attains the highest value among all the networks as shown in figure. 3.3. Thus again we can say that for more number of repetitions our proposed network yields much better segmentation results and the network delineates the more finer boundaries in the segmented results as depicted by the value of mean BF score and the obtained images are closer to the ground truth data.

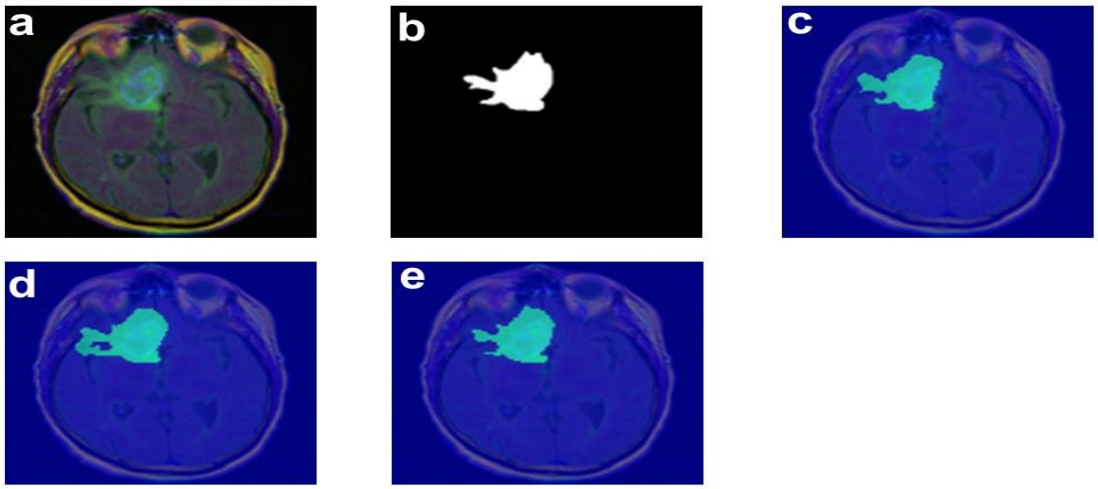


***Figure 3.3: Segmentation results when network was trained on Kaggle dataset and tested on same dataset (a) original image, (b) segmentation mask, (c) seg-net output, (d) u-net output, (e) CCN-PR-seg-net output for 100 epochs of training.***

*Table 3-2: Parameter comparison of the networks for 100 epochs of training*

| Networks | Global Accuracy | Mean Accuracy | mIoU | Mean BF Score |
|---|---|---|---|---|
| Seg-net | 0.99772 | 0.9751 | 0.95387 | 0.917 |
| U-Net | 0.99818 | 0.97104 | 0.96222 | 0.92133 |
| CCN-PR-Seg-net | 0.99916 | 0.9914 | 0.98861 | 0.9752 |

### 3.5.2 Results obtained when network was trained on Kaggle data and tested fig share dataset

Figure 3.4 and figure 3.5 shows the results obtained when the network was trained on Kaggle dataset and tested on fig share dataset for 50 and 100 epochs of training. The network being trained on different dataset yielded remarkably good results when tested on the other dataset. On observing figure 3.4 and figure 3.5, it is evident that the shape of the tumour boundary segmented by the proposed network is closest to the expected results when compared with the other networks under consideration. Table 3-3 and Table 3-4 compares the segmentation results obtained by the networks. The proposed network tends to produce finer results when the number of epochs is increased. It can be noticed from figure 3.4 and figure 3.5 that the segmented results produced from 100 epochs of training are having more regular and fine boundary detail as compared to the results obtained from 50 epochs of training.

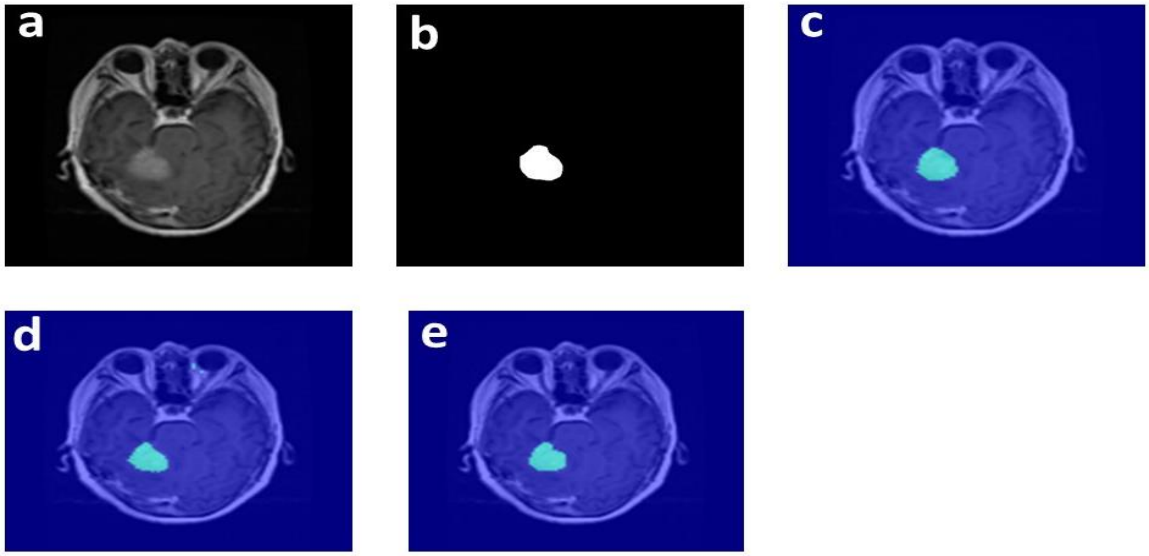


*Figure 3.4: Segmentation results when the network was trained on Kaggle dataset and tested on figshare dataset (a) original image, (b) segmentation mask, (c) seg-net output, (d) u-net output, (e) CCN-PR-seg-net output for 50 epochs of training.*

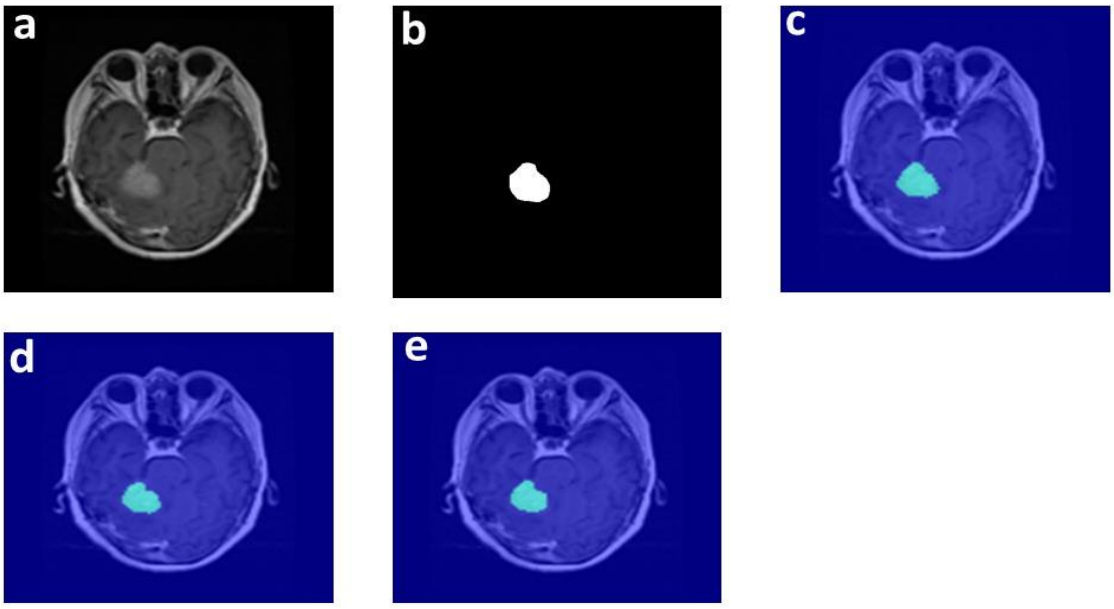


*Figure 3.5: Segmentation results when network was trained on Kaggle dataset and tested on figshare dataset (a) original image, (b) segmentation mask, (c) seg-net output, (d) u-net output, (e) CCN-PR-seg-net output for 100 epochs of training.*

***Table 3-3:Parameter comparison of the networks for 50 epochs of training when network is trained on kaggle dataset and tested on figshare dataset***

| **Networks** | **Global Accuracy** | **Mean Accuracy** | **mIoU** | **Mean BF Score** |
|---|---|---|---|---|
| Seg-net | 0.99664 | 0.96193 | 0.93318 | 0.88271 |
| U-Net | 0.99679 | 0.9373 | 0.94255 | 0.89674 |
| CCN-PR-Seg-net | 0.99821 | 0.9865 | 0.96387 | 0.92397 |

***Table 3-4:Parameter comparison of the networks for 100 epochs of training when network is trained on kaggle dataset and tested on figshare dataset.***

| **Networks** | **Global Accuracy** | **Mean Accuracy** | **mIoU** | **Mean BF Score** |
|---|---|---|---|---|
| Seg-net | 0.99186 | 0.97304 | 0.94299 | 0.92062 |
| U-Net | 0.99446 | 0.98287 | 0.96255 | 0.92827 |
| CCN-PR-Seg-net | 0.99882 | 0.98961 | 0.98681 | 0.9567 |

### 3.6 Discussion

In the field of biomedical imaging, even minute details have importance as the incorrect analysis may lead to the blunder in diagnosis. So, for segmenting the brain tumour we present a deep learning approach which incorporates residual connections along with parametric RELU with the value of $\alpha = 0.05$.We used cross channel normalization in our proposed model for normalizing the image data. Our proposed model yields the better results as compared with the other networks in this paper. The value of mean intersection over union (mIou) indicates how well a segmented result overlaps with the ground truth data [105]. Higher value of mIoU is desirable for the segmentation to be considered as good. As shown in Table 3-1 and 3-2, we obtained above 2.5% higher value of mIoU in the results of our proposed network which indicates the better segmentation results as compared with the networks in this paper when the network was trained and tested on the Kaggle dataset (dataset being partitioned in to train and test dataset). The BF contour matching score which is the measure of the computational

precision in evaluating the segmented boundary tends to yield the highest value for our network among all the networks being compared [106] [107]. For 50 epochs of training the margin of BF score is about 4% higher as compared with the second best method. As the epochs are increased, the BF score tends to get better and it attains a higher level of above 5% when compared with the immediate second best method. For testing the comprehensive capability of the network, the network trained on Kaggle dataset was fed with the figshare dataset. The network being trained on different dataset and tested for other dataset produced exceptionally good results. Figure 3.4 and figure 3.5 shows the segmented tumour regions. Table 3-3 and Table 3-4 shows the comparison of segmentation parameters of all the networks under study. The value of mIoU and BF score is about 2% and 3% higher respectively for the proposed network. It indicates that the boundary details segmented by the proposed network are preserved more accurately than the other networks under study. Moreover, the area of overlap with expected ground truth data is much higher in case of network proposed. This indicates the better capability of our network to segment the tumour region more precisely without retraining the network on other dataset. So, it can be seen as the versatile potential of the network for segmentation. Training a large dataset is time consuming process and if the network is able to produce good results after training once on a particular dataset and tested on different datasets then it can be considered as a figure of merit for that network. The combination of mean intersection over union (mIoU) with the Berkley contour matching score (BF score) gives the best results which are more aligned with the human qualitative visual perception.

The results for more number of iterations have higher similarity to the ground truth data i.e. the results have better segmented boundaries and have more area of overlap with the ground truth images. Thus the increase in number of iterations gives the more smoothed results which can be judged as having more details, with the human visual raking of segmentation results. The global and mean accuracies of the network proposed by us are yielding the higher values as

compared with the results of other networks. In biomedical imaging even minute details have higher relevance as in case of tumour the identification of correct boundaries and shape of tumour is of great importance. As shown by figure 3.2, figure 3.3,figure 3.4 and figure 3.5 the boundary delineated by our proposed model is very much close to the ground truth images and the evaluation parameters shows the correctness of segmentation as the correctly classified pixels by our network are larger as compared to other networks. The use of residual connections enriches the network with the capability of best performance along with the fast convergence of the network. The layers which are nearer to the centre of the network does not get effective updates because of the problem of vanishing gradient.

The comparison of performance on the basis of epoch counts gives the understanding that running the network for higher number of epochs the performance tends to get better as indicated by Table 3-1,Table 3-2,Table 3-3 and Table 3-4. On comparing the various architecture being studied in this paper for 100 epochs, we get to know that our network has the highest accuracy along with the maximum BF score. The BF scores shows increment of above 5% over the results obtained by the second best method. When the trained network was tested on different dataset then also the values of the BF score obtained were remarkably good. An increment of above 3% was achieved in the BF score when the already trained network was supplied with different dataset. The literature review suggests that the BF score is the judgment of the segmented boundary deail. A high BF score states that the boundary is well segmented in the result obtained [106] [107]. The BF score measures how close the predicted boundary of an object matches the ground truth boundary.

The use of convolution and max-pooling operation tends to learn the feature of the object being segmented i.e. this combination focuses on learning the detailed characteristic of the segmentation object. In this process of learning the details of the boundary of the object being segmented is lost. Thus we can say, that the network starts losing the spatial resolution of the

feature maps which is important as under the absence of such details the correct contour of the object under observation cannot be identified in correct manner which will lead to the blunder in the diagnosis process. Thus those max-pooling indices which possess the boundary details are learnt by the network and same are used at the time of decoding to get the exact boundary details. Figure 3.2 and figure 3.3 shows the fine boundaries segmented by our proposed network. Parametric RELU (PRELU) has non zero value for the negative values of the inputs i.e. it has some slope in the negative region. Parametric RELU removes the problem of 'dying RELU' in the network as it gives some output for the negative values of the input too. It learns more faster as compared to the normal RELU as it is more balanced. The use of residual connections provides the flexibility to the network to add more number of layers for better feature acquisition without degrading the performance of the network. It was here by noted that the segmentation achieved by CCN-PR-seg-net was clinically more pertinent in terms of accurate inclusion of pathological tissue and exclusion of normal brain tissue projecting with in the areas of complex tumoral anatomy.

### 3.7 Conclusion

A deep convolutional neural network is presented in this work which is capable of segmenting the brain tumour with more precise boundary and pixel details as desired for correct diagnosis in the domain of biomedical image analysis by the radiology experts. Our proposed network not only segmented the brain tumour but also preserved the minute contour curvatures and boundary details. The results produced by the network when the already trained network is fed with different dataset are exceedingly good. The network proposed is more flexible with the number of layers and more balanced as compared to other networks because of incorporation of residual connections and parametric RELU. In conclusion the remarkable increment achieved in the evaluation metrics mIoU and mean BF score by the proposed network proved the better segmentation capability of the network over other network architectures.