Brain Tumor Segmentation from MRI Images Using U-Net Architecture

Imesh Udara Ekanayake
Dept. Computer Engineering
University of Peradeniya
Peradeniya, 20400, Sri Lanka
imeshuek@eng.pdn.ac.lk

***Abstract*—Brain tumor is the 22nd most common cancer worldwide with 1.8% of the total number of new cancers worldwide which ranks number 12 for mortality. This work describes a framework for automatic brain tumor segmentation from Magnetic Resonance images (MRI). The detection of edema is done simultaneously with tumor segmentation, as the knowledge of the extent of edema is important for diagnosis, planning, and treatment. Whereas many other tumor segmentation methods rely on the intensity enhancement produced by the gadolinium contrast agent in the weighted image, the method proposed here does not require contrast enhanced image channels. A 3-D Convolution Neural Network with 2D Conv-LSTM is proposed to segment Magnetic Resonance images with brain tumor into background and three hierarchical regions: whole tumor,tumor core and enhancing tumor core. The architecture is designed to decompose the multi-class segmentation problem into a sequence of three binary segmentation problems according to the sub-region. Since most of the area is consist of the background and only a fraction of area consist of the tumor, an F1 score has defended for each pixel to handle the imbalances of the pixels. Since there are 155 layers of 240x240 images with 4 classes for each MRI instance, a part of the image has been selected which consist minimum of 5% of pixels with brain tumor. After detecting the regions we have taken confidence benchmark to sharpen the margins of the tumors.**

***Index Terms*—brain tumor, 3D convolution, machine learning, U-net, 2D conv-LSTM, neural network, imege segmentation**

I. Introduction

Automatic brain tumor segmentation from MR images is a difficult task that involves various disciplines covering pathology, MRI physics, radiologist's perception, and image analysis based on intensity and shape [1], [2]. There are many issues and challenges associated with brain tumor segmentation. Brain tumors may be of any size, may have a variety of shapes, may appear at any location, and may appear in different image intensities. Some tumors also deform other structures and appear together with edema that changes intensity properties of the nearby region. For many human experts, manual segmentation is a difficult and time consuming task, which makes an automated brain tumor segmentation method desirable.

Brain MRI segmentation is an essential task in many clinical applications because it influences the outcome of the entire analysis. This is because different processing steps rely on accurate segmentation of anatomical regions. For example, MRI segmentation is commonly used for measuring and visualizing different brain structures, for delineating lesions, for analysing brain development, and for image-guided interventions and surgical planning. This diversity of image processing applications has led to development of various segmentation techniques of different accuracy and degree of complexity [2], [3].

The typical use of convolution networks is on classification tasks, where the output to an image is a single class label. However, in many visual tasks, especially in biomedical image processing, the desired output should include localization, i.e., a class label is supposed to be assigned to each pixel.

To to train the segmentation algorithm "A large annotated medical image dataset for the development and evaluation of segmentation algorithms" has been used, which consist of 10 MRI and X-ray image data-sets. Where when visualizing the all different classes in one image it looks like (check Fig. 1)

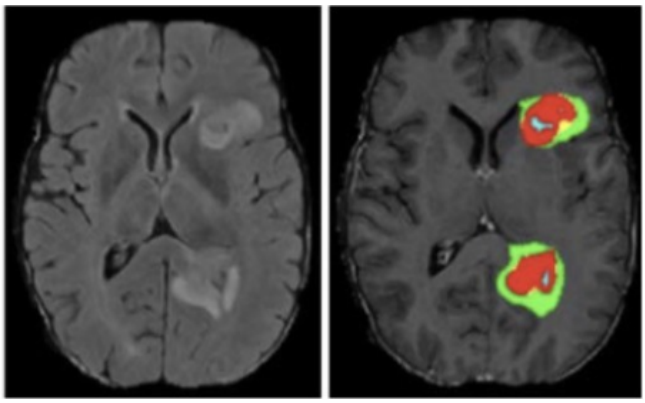

Fig. 1. Visualization of all classes in one image.

When Consider the dataset it has 4 different layers as T1, T2, T1 enhanced and Fluid Attenuated Inversion Recovery (FLAIR). where a sample layer from each category is shows in the (Fig. 2). When applying segmentation algorithms one of the major concerns is the dataset size here we had only 480 instances which is like the bare minimum number of instances for a image classification or a segmentation, There the U-net architecture and sub-volume concept has considered.

II. Related work

Tumor is an uncontrolled growth of cancer cells in any part of the body. Tumors are of different types and have

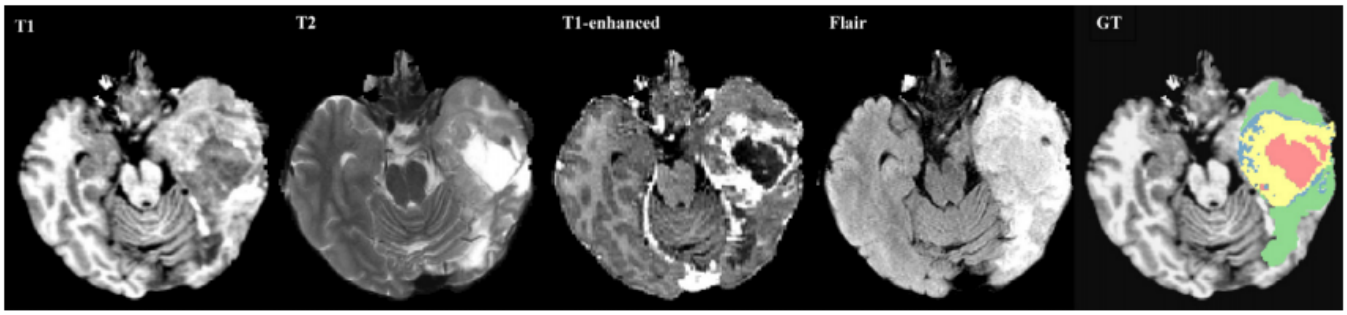

Fig. 2. Four Channels in one layer of data.

different characteristics and different treatments [4].Brain tumor segmentation aims to separate the different tumor tissues such as active cells, necrotic core ,and edema from normal brain tissues of White Matter (WM), Gray Matter (GM), and Cerebrospinal Fluid (CSF). [5]Many research studies have been carried out on MRI-based brain tumor segmentation and, at the same time those could gather the attention in recent years due to non-invasive imaging and good soft tissue contrast of Magnetic Resonance Imaging (MRI) images.

Guotai Et al.(2018) proposed a cascade of fully convolutional neural networks to segment multimodal Magnetic Resonance (MR) images with brain tumor into the background and three hierarchical regions: whole tumor, tumor core and enhancing tumor core [6].In there the whole tumor is portioned in the initial step and the bouncing box of the outcome is utilized for the tumor center division in the subsequent advance.The improving tumor center is then fragmented dependent on the bouncing box of the tumor core division result.

Automatic brain tumor segmentation from MR images is a difficult task that involves various disciplines covering pathology, MRI physics, radiologist's perception, and image analysis based on intensity and shape. [**?**]Marcel Et al.(2004)depicts a system for automatic brain tumor segmentation from MR images. [**?**]There ,the detection of edema is done at the same time with tumor segmentation.Whereas many other tumor segmentation methods believe the intensity enhancement produced by the gadolinium contrast agent within the T1-weighted image, the tactic proposed here doesn't require contrast enhanced image channels.The only required input for the segmentation procedure is that the T2 MR image channel, but it can make use of any additional non-enhanced image channels for improved tissue segmentation.

Marcel Et al.(2003) performed the segmentation of a registered set of resonance images using an expectation-maximization scheme [3]. The segmentation is guided by a spatial probabilistic atlas that contains expert prior knowledge about brain structures. This atlas is modified with the subject-specific brain tumour prior that's computed based on contrast enhancement.The results obtained from the automated segmentation program are compared with results from manual and semi-automated methods. The automated method yields results that have [3] surface distances at roughly 1–4 mm compared with the manual results.

Jason J.Et al.(2007) presented a replacement method for automatic segmentation of heterogeneous image data that takes a step toward bridging the gap between bottom-up affinity-based segmentation methods and top-down generative model based approaches. The most contribution of the paper was a Bayesian formulation for incorporating soft model assignments into the calculation of affinities, which are conventionally model free. [7]There they have integrated the resulting model-aware affinities into the multilevel segmentation by weighted aggregation algorithm, and apply the technique to the task of detecting and segmenting brain tumor and edema in multichannel magnetic resonance (MR) volumes. [**?**], [8], [9]

Atiq Islam Et al.proposed A stochastic model for characterizing tumor texture in brain MR images [7].The efficacy of the model is demonstrated in patient-independent brain tumor texture features extraction and tumor segmentation in magnetic resonance images (MRIs).Due to complex appearance in MRI, brain tumor texture is formulated using a multiresolution-fractal model known as multi-fractional Brownian motion (mBm). Detailed mathematical derivation for mBm model and corresponding novel algorithm to extract spatially-varying multi-fractal features are proposed. A multifractal feature-based brain tumor segmentation method is developed next. To evaluate efficacy, tumor segmentation performance using proposed multi-fractal features is compared with that using Gabor-like multi-scale texton features.

Due to the rapid development of medical image modalities, new application-specific segmentation problems are emerging and new methods are continuously explored and introduced. Selection of the most appropriate technique for a given application is a difficult task. In many cases, a combination of several techniques may be necessary to obtain the segmentation goal. Very often integration of multimodal information (acquired from different modalities or over time) can help to segment structures that otherwise could not be detected on single images. [10]

III. Methodology

The proposed methodology consists of 3 key steps: Data preprocessing, models training and model evaluation (Fig. 3).

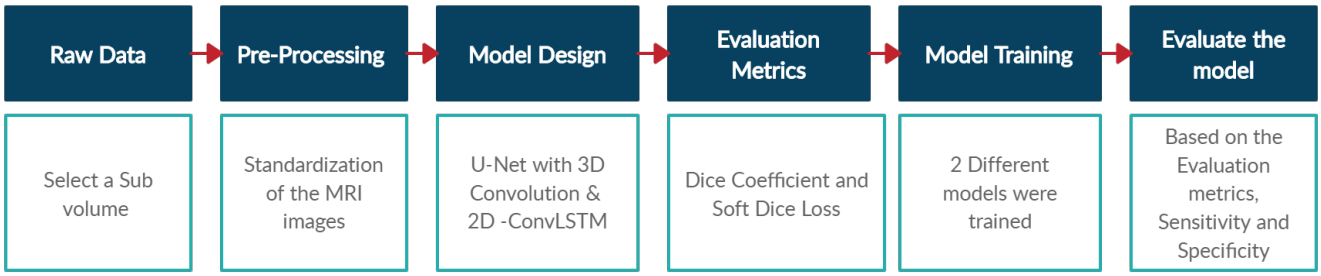

Fig. 3. Proposed workflow.

A. *Data Preprocessing : Sub-volume Selection*

When considering the whole volume of the image it is a [240x240x155]x4 space, where this volume costs a single image to be nearly 35 to 40 Megabytes and training the model with a single instance of size of that would require a huge amount of RAM. Moreover, there are only 480 MRI instances which makes a really low amount of data points to train a good image segmentation model, where repeating with the same dataset would over-fit the model easily. There I decided to sample the small blocks of images with minimum of 5% of the image consist of the tumor. which gave an output of 160x160x16 training data and 3x160x160x16 labels.

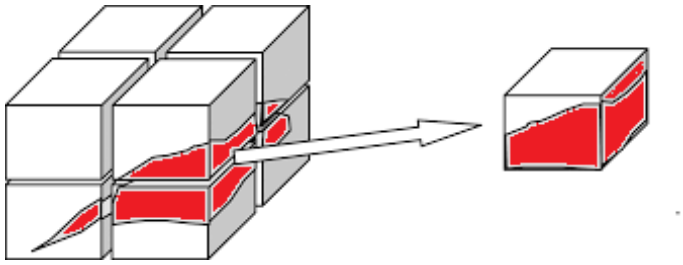

Fig. 4. Sub-volume Selection.

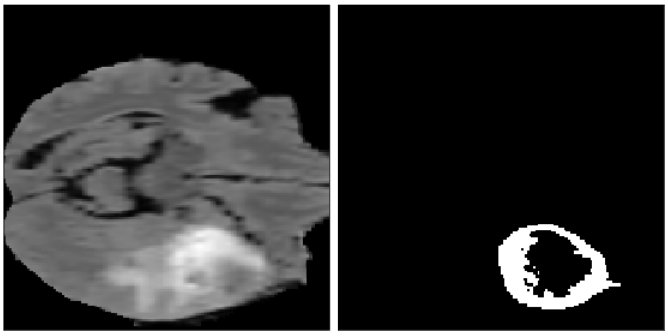

Fig. 5. sample of a sub-volume layer.

B. *Data Preprocessing : Standardizing the Image*

The images pixel values has a variation of 0-255. In medical image processing, preprocessing of an image is very important so that the extracted image does not have any impurities, and it is accomplished to be better for the forthcoming process such as segmentation, feature extraction, etc. Only the correct segmentation of the tumor will yield the accurate result.There for the image pixels have been normalized to have a standard deviation to be 1.

C. *Model Designing : Convolution Algorithm*

Model selection and designing is a major part of this workflow, where the segmentation is all depending on that. There we could find 2 main operations, which are in two different extends of performance requirements. One is 3D convolution, which is simile to 2D convolution and uses a cube to convolute the selected volume.

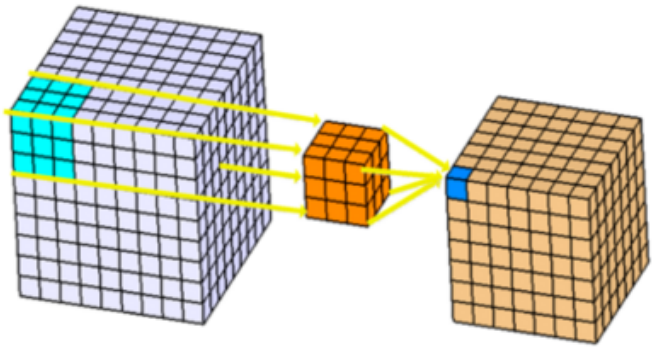

Fig. 6. 3D convolution.

where it keeps all the values of the kernels and has to store a separate image for each kernel. which exhaustively requires the memory and a considerable amount of computational power.

in the other hand, 2D convolution long short term memory model (2D Conv-LSTM). Which is an RNN model where each output is a 2D convolution which keeping the memory of last predefined number of images. Here, the most important part is that it requires more computational power than 3D convolution and way less memory to store the model while training and storing. Since we are not capable of providing any of these extensively, we had to find a middle ground with both algorithms to build the model.

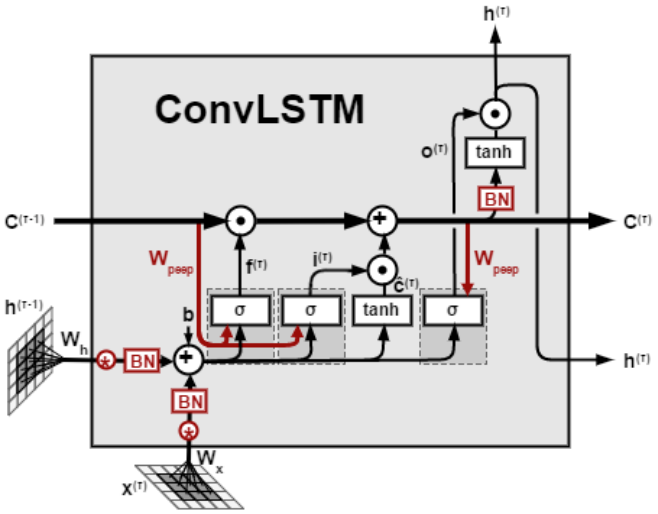

Fig. 7. 2D convolution LSTM.

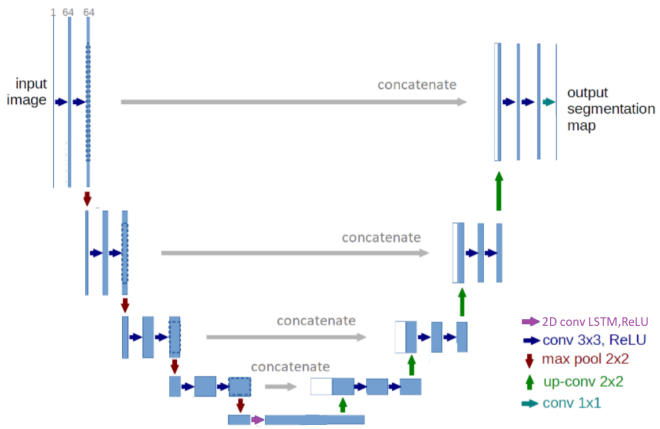

Fig. 8. U-Net architecture.

D. Model Designing : Neural Network Architecture

One of the main concerns in this work flow is number of instances there to train the model and 3D data, initially we planed to do transfer learning by training the data from image-net and fine-tune the model for the brain tumor detection. Therefor the U-Net [11], [12] architecture has been selected where it needs only a small amount of data compared to the rest of the architectures for image segmentation.

When building the model we have used 6 main layers, which are Activation, Conv3D, MaxPooling3D, Deconvolution3D, UpSampling3D and ConvLSTM2D.

Initially 3D convolution with 3x3x3 kernel has been used and Relu has been used as the activation layer, where by adding a padding the image has been kept as the same size after each convolution. In each convolution stage by starting at 32 multiples of 2 kernels were used (1st convolution 32, 2nd convolution 64 kernels and so on). After 2 convolutions a 3D max pooling has been done to reduce the image size. like wise 4 max pooling steps have been taken after 8 convolutions.

There after a 2D convolution LSTM has been done with the size of 3x3 kernel and maximum of 10 pipeline memory iterations. Which made it reduced the work of more 3D convolution layers.

Next, at the 9th layer (20x20x768) images has to be get back to the original size with the segmentation, were we did it in 4 steps. In each step previous layer images were up sampled and concatenated with the parallel max pooling layer. where it sharpen the tumor boundaries. There also the ReLU has been selected as the activation layer for UP sampling. in between two up sampling and concatenation layers 3D convolution has done to maintain the local consistency of pixels.

E. Cost Function : Dice Similarity Coefficient

Aside from the architecture, one of the most important elements of any deep learning method is the choice of the loss function.A natural choice that you may be familiar with is the cross-entropy loss function, However, this loss function is not ideal for segmentation tasks due to heavy class imbalance (there are typically not many positive regions).A much more common loss for segmentation tasks is the Dice similarity coefficient, which is a measure of how well two contours overlap. The Dice index ranges from 0 (complete mismatch), To 1 (perfect match). In general, for two sets ”A” and ”B”, the Dice similarity coefficient is defined as:

$$\mathrm{DSC}(A,B) = \frac{2 \times |A \cap B|}{|A| + |B|}.$$

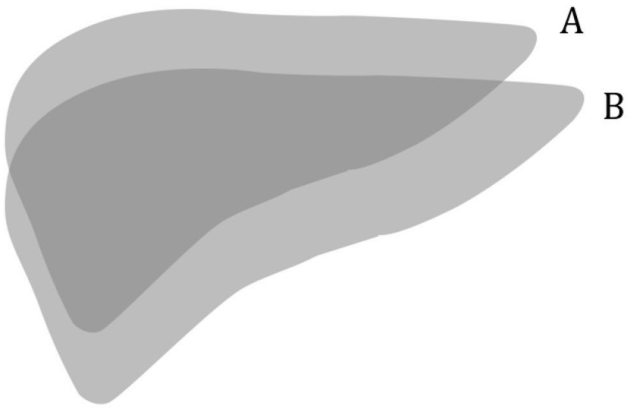

Fig. 9. Dice Similarity Coefficient.

Here we can interpret ”A” and ”B” as sets of voxels, ”A” being the predicted tumor region and ”B” being the ground truth. The model will map each voxel to 0 or 1, 0 means it is a background voxel and 1 means it is part of the segmented region.In the dice coefficient, the variables in the formula are:

x : the input image

f(x) : the model output (prediction)

y : the label (actual ground truth)

The dice coefficient ”DSC” is:

$$\mathrm{DSC}(f,x,y) = \frac{2 \times \sum_{i,j} f(x)_{ij} \times y_{ij} + \epsilon}{\sum_{i,j} f(x)_{ij} + \sum_{i,j} y_{ij} + \epsilon}$$

- ϵ is a small number that is added to avoid division by zero

F. Cost Function : Dice Coefficient for Multiple classes

From a single class case, we can think about how to approach the multi class context segmentation for each of the 3 classes of abnormality (edema, enhancing tumor, non-enhancing tumor).This will give 3 different dice coefficients (one for each abnormality class). To combine these,we can just take the average and calculate the overall dice coefficient is:

$$DC(f,x,y) = \frac{1}{3}\left(DC_1(f,x,y) + DC_2(f,x,y) + DC_3(f,x,y)\right)$$

- DC_1, DC_2 and DC_3 are edema, enhancing tumor, and non-enhancing tumor dice coefficients.
For any number of classes, the equation becomes:

$$DC(f,x,y) = \frac{1}{N}\sum_{c=1}^{C}\left(DC_c(f,x,y)\right)$$

In this case, with three categories, $C = 3$

Implementing the mean dice coefficient is not different from singe-class implementation.

G. Cost Function : Soft Dice loss

Similarly as measure the coefficient the loss should also be calculate, even though Dice Coefficient makes intuitive sense, it is not the best for training.

It takes in discrete values (zeros and ones) and the model outputs probabilities that each pixel is, say, a tumor or not, and also can use to back-propagate through those outputs.

Therefore, we need an analogue of the Dice loss which takes real valued input. This is where the Soft Dice loss comes in. The formula is:

$$\mathcal{L}_{Dice}(p,q) = 1 - \frac{2 \times \sum_{i,j} p_{ij} q_{ij} + \epsilon}{\left(\sum_{i,j} p_{ij}^2\right) + \left(\sum_{i,j} q_{ij}^2\right) + \epsilon}$$

p is predictions, q is the ground truth, In practice each q_i will either be 0 or 1, ϵ is a small number that is added to avoid division by zero

We can also check that if p_i and q_i are each 0 or 1, then the soft Dice loss is just one minus the dice coefficient.

Similarly when consider than multi-class, For any number of categories of diseases, the expression becomes:

$$\mathcal{L}_{Dice}(p,q) = 1 - \frac{1}{N}\sum_{c=1}^{C} \frac{2 \times \sum_{i,j} p_{cij} q_{cij} + \epsilon}{\left(\sum_{i,j} p_{cij}^2\right) + \left(\sum_{i,j} q_{cij}^2\right) + \epsilon}$$

H. Thresholding

When segmented the data, it is hard to clearly mark the margins therefor it blurs out from the core of tumor, to handle that problem we only consider the prediction values above a certain threshold, which clearly breaks the margins.

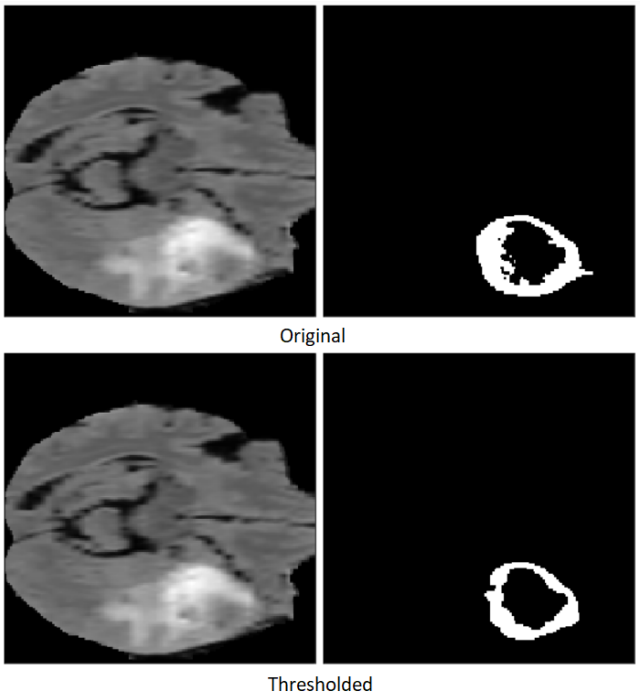

Fig. 10. Threshold and original image .

IV. Results

Check the Fig. 11.

V. Discussion

Starting from the process this workflow has focused on the segmentation of the MR Images. Selecting a sub-volume with minimum of 5% pixels of brain tumor made a huge impact on training and number of iterations it required to train, moreover, since this neural network requires higher amount of memory, it makes hard to have more images in a single batch of training. Therefor, it was limited to maximum on 4 images for a batch to train in a system with 16GB RAM, but it was way faster than 2D Conv-LSTM.

When mixing both algorithms, it made the system learn faster than going towards one end. Combined system usually takes about 21mins for single epoch. In addition, concatenating and up sampling has made system predict sharp and clear margins of the final image.

Selecting a cost function was one of the crucial parts of the workflow, initially we have selected the prediction error, and mean square error. Since the area of the tumor is small even though it doesn't predict anything it gives a good accuracy therefor we decided to do F1 score for each pixel and we applied the Soft Dice loss.

Finally, after checking the accuracies and F1 scores it has not been improved over time then we have introduced the thresholding and it rapidly increases the values of accuracies.

VI. Conclusion

In conclusion this work flow has made the process faster to learn and get in to a good accuracy witha high recall.

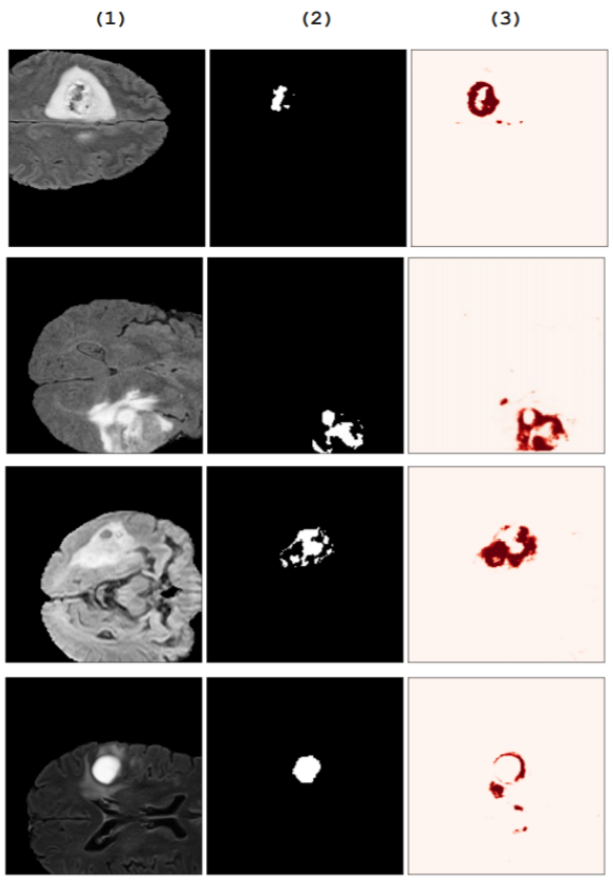

Fig. 11. Predictions.

REFERENCES

[1] M. Styner, J. Lee, B. Chin, M. Chin, O. Commowick, and H. Tran, "3d segmentation in the clinic: A grand challenge ii: Ms lesion segmentation," 2008.

[2] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in *International MICCAI brainlesion workshop*. Springer, 2017, pp. 178–190.

[3] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Medical image analysis*, vol. 8, no. 3, pp. 275–283, 2004.

[4] M. P. Gupta, M. M. Shringirishi *et al.*, "Implementation of brain tumor segmentation in brain mr images using k-means clustering and fuzzy c-means algorithm," *International Journal of Computers & Technology*, vol. 5, no. 1, pp. 54–59, 2013.

[5] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of mri-based brain tumor segmentation methods," *Tsinghua science and technology*, vol. 19, no. 6, pp. 578–595, 2014.

[6] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in *International MICCAI brainlesion workshop*. Springer, 2017, pp. 178–190.

[7] I. Despotović, B. Goossens, and W. Philips, "Mri segmentation of the human brain: challenges, methods, and applications," *Computational and mathematical methods in medicine*, vol. 2015, 2015.

[8] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of mri-based brain tumor segmentation methods," *Tsinghua science and technology*, vol. 19, no. 6, pp. 578–595, 2014.

[9] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in mri images," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

[10] A. Islam, S. M. Reza, and K. M. Iftekharuddin, "Multifractal texture estimation for detection and segmentation of brain tumors," *IEEE transactions on biomedical engineering*, vol. 60, no. 11, pp. 3204–3215, 2013.

[11] W. Chen, B. Liu, S. Peng, J. Sun, and X. Qiao, "S3d-unet: separable 3d u-net for brain tumor segmentation," in *International MICCAI Brainlesion Workshop*. Springer, 2018, pp. 358–368.

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.