

A Novel Approach For Brain Tumor Classification Using Convolutional Neural Network

F R Shiny Malar¹, S Bavithra², S Padmalal³

^{1, 2, 3} Dept of CSE

^{1, 2, 3} Stella marys College Of Engineering

Abstract- A novel CNN-based method for segmentation of brain tumors in MRI images. Pre-processing stage consisting of bias field correction, intensity and patch normalization. After that, during training, the number of training patches is artificially augmented by rotating the training patches, and using samples of HGG to augment the number of rare LGG classes. The CNN is built over convolutional layers with small 3X3 kernels to allow deeper architectures. Augmenting training data set by rotating the patches as well as by sampling from classes of HGG that were underrepresented in LGG. Found that data augmentation was also quite effective, although not thoroughly explored in Deep Learning methods for brain tumor segmentation. Also, we investigated the potential of deep architectures through small kernels by comparing our deep CNN with shallow architectures with larger filters. We found that shallow architectures presented a lower performance, even when using a larger number of feature maps. Finally, verified that the activation function LReLU was more important than ReLU in effectively training our CNN.

Keywords- Segmentation, normalization, training patches, Mini Augmenting training, Deep Learning.

I. INTRODUCTION

Brain cancer can be counted among the most deadly and intractable diseases. Tumors may be embedded in regions of the brain that are critical to orchestrating the body's vital functions, while they shed cells to invade other parts of the brain, forming more tumors too small to detect using conventional imaging techniques. Brain cancer's location and ability to spread quickly makes treatment with surgery or radiation like fighting an enemy hiding out among minefields and caves. In recent years, the occurrence of brain tumors has been on the rise. Unfortunately, many of these tumors will be detected too late, after symptoms appear. The past few years had witnessed a rapid and multi directional increase in the applications of image processing. In today's digital era, capturing, storing and analysis of medical image had been digitized.

The challenge stands tall especially in regions with abnormal color and shape which needs to be identified by radiologists for future studies. The key task in designing such image processing and computer vision applications is the accurate segmentation of medical images. Image segmentation is the process of partitioning different regions of the image based on different criteria. Surgical planning, post-surgical assessment, abnormality detection, and much other medical application require medical image segmentation.

Brain image segmentation from MRI images is complicated and challenging but its precise and exact segmentation is necessary for tumors detection and their classification, edema, hemorrhage detection and necrotic tissues. For early detection of abnormalities in brain parts, MRI imaging is the most efficient imaging technique. The morphological operations are basically applied on some assumptions about the size and shape of the tumor and in the end the tumor is mapped onto the original gray scale image with 255 intensity to make visible the tumor in the image.

For the treatment of patients with brain tumors, imaging of the brain is often indicated at different stages and usually has a significant role in each of them. Several stages of management may be considered:

- Detection or confirmation that a structural abnormality is present,
- Localization and assessment of the extent of any abnormality,
- Characterization of the abnormality,
- Assessment of the nature of a tumor,
- Facilitation of additional diagnosis procedures, and planning for surgery or other types of therapy,
- Intra-operative control of rejection progress, Monitoring of response to therapy.

A variety of imaging techniques are used to study brain tumors, including computed tomography (CT), magnetic resonance (MR) imaging, and single photon emission computed tomography (SPECT) imaging, positron emission tomography (PET) scanning, and cerebral angiography. At this

moment, CT and MR imaging are the most widely used techniques, because of their widespread availability and their ability to produce high resolution images of normal anatomic structures and pathological tissues. CT is the fastest modality, making it the preferred examination for imaging critically ill or medically unstable patients. SPECT and PET imaging serve smaller roles, although their ability to provide information on tissue biology and physiology can be greatly helpful. PET scanning is also used to evaluate tumor grade.

II. LITERATURE SURVEY

C. Senthilkumar (2018), focuses on how a brain image is being segmented to diagnose the brain tumor by using spatial fuzzy clustering algorithm. The segmentation process may cause error while diagonalizing MR images due to the artifacts and noises exist in it. This may leads to misclassify the normal tissue as abnormal tissue. The proposed method is to segment normal tissues such as White Matter, Gray Matter, Cerebrospinal fluid and abnormal tissue like tumor part from MR images automatically. It comprises of pre-processing step, using wrapping based curvelet transform to remove noise and modified spatial fuzzy C-Means algorithm segments normal tissues by considering spatial information. The neighboring pixels are highly correlated and construct initial membership matrix randomly. The process is also intended to segment the tumor cells as well as the removal of background noises for smoothening the region, results in presenting segmented tissues and parameter evaluation to produce the algorithm efficiency.

Fannia Pacheco, (2016) proposes a new unsupervised algorithm for feature selection based on attribute clustering and rough set theory. Rough set theory is used to compute similarities between features through the relative dependency. The clustering approach combines classification based on distance with clustering based on prototype to group similar features, without requiring the number of clusters as an input. Additionally, the algorithm has an evolving property that allows the dynamic adjustment of the cluster structure during the clustering process, even when a new set of attributes feeds the algorithm. That gives to the algorithm an incremental learning property, avoiding a retraining process. These properties define the main contribution and significance of the proposed algorithm. Two fault diagnosis problems of fault severity classification in gears and bearings are studied to test the algorithm. Classification results show that the proposed algorithm is able to select adequate features as accurate as other feature selection and reduction approaches.

Deepika Koundal, (2015) introduces an automated delineation method that integrates spatial information with

neutrosophic clustering and level-sets for accurate and effective segmentation of thyroid nodules in ultrasound images. The proposed delineation method named as Spatial Neutrosophic Distance Regularized Level Set (SNDRLS) is based on Neutrosophic L-Means (NLM) clustering which incorporates spatial information for Level Set evolution. The SNDRLS takes rough estimation of region of interest (ROI) as input provided by Spatial NLM (SNLM) clustering for precise delineation of one or more nodules. The performance of the proposed method is compared with level set, NLM clustering, Active Contour Without Edges (ACWE), Fuzzy C-Means (FCM) clustering and Neutrosophic based Watershed segmentation methods using the same image dataset. To validate the SNDRLS method, the manual demarcations from three expert radiologists are employed as ground truth. The SNDRLS yields the closest boundaries to the ground truth compared to other methods as revealed by six assessment measures (true positive rate is $95.45 \pm 3.5\%$, false positive rate is $7.32 \pm 5.3\%$ and overlap is $93.15 \pm 5.2\%$, mean absolute distance is 1.8 ± 1.4 pixels, Hausdorff distance is 0.7 ± 0.4 pixels and Dice metric is $94.25 \pm 4.6\%$). The experimental results show that the SNDRLS is able to delineate multiple nodules in thyroid ultrasound images accurately and effectively. The proposed method achieves the automated nodule boundary even for low-contrast, blurred, and noisy thyroid ultrasound images without any human intervention. Additionally, the SNDRLS has the ability to determine the controlling parameters adaptively from SNLM clustering.

Yong Xia List (2015), introduce the variational Bayes inference to brain MRI image segmentation, and thus propose a novel segmentation algorithm based on learning a cohort of local variational Gaussian mixture (LVGM) models. By assuming all Gaussian parameters to be random variables, the LVGM model has more flexibility than GMM in characterizing the complexity of brain voxel distributions. To alleviate the impact of bias field, we train each LVGM model on a sampled small data volume and linearly combine the trained models to classify each brain voxel. Also construct a co-registered probabilistic brain atlas for each MRI image to incorporate the prior knowledge about brain anatomy into the segmentation process. The proposed LVGM learning algorithm has been evaluated against five state-of-the-art brain MRI image segmentation methods on both synthetic and clinical data. Our results suggest that the LVGM algorithm can segment brain MRI images more effectively and provide more precise distribution of major brain tissues.

Anupama Namburu (2017) propose, a novel clustering algorithm, namely generalized rough intuitionistic fuzzy c-means (GRIFCM) for brain magnetic resonance (MR) image segmentation avoiding the dependency with the fuzzy

membership function. In this algorithm, each pixel is categorized into three rough regions based on the thresholds obtained by the image data by minimizing the noise. These regions are used to create IFS. The distance measure based on IFS eliminates the influence of noise and INU present in the image producing accurate brain tissue segmentation. The proposed algorithm is evaluated through simulation and compared it with existing k-means (KM), fuzzy c-means (FCM), Rough fuzzy c-means (RFCM), Generalized rough fuzzy c-means (GRFCM), soft rough fuzzy c-means (SRFCM) and rough intuitionistic fuzzy c-means (RIFCM) algorithms. Experimental results prove the superiority of the proposed algorithm over the considered algorithms in all analyzed scenarios.

Dishant Mittal, (2016) Over the years data clustering algorithms have been used for image segmentation. Due to the presence of uncertainty in real life datasets, several uncertainty based data clustering algorithms have been developed. The C-means clustering algorithms form one such family of algorithms. Starting with the Fuzzy C-means (FCM) a sub family of this family comprises of Rough C-means (RCM), Intuitionistic Fuzzy C-means (IFCM) and their hybrids like Rough Fuzzy C-means (RFCM) and Rough Intuitionistic Fuzzy C-means (RIFCM). In the basic sub family of this family of algorithms, the Euclidean distance was being used to measure the similarity of data. However, the sub family of algorithms obtained replacing the Euclidean distance by Kernel based similarities produced better results. Especially, these algorithms were useful in handling viably cluster data points which are linearly inseparable in original input space. During this period it was inferred by Krishnapuram and Keller that the membership constraints in some rudimentary uncertainty based clustering techniques like Fuzzy C-means imparts them a probabilistic nature, hence they suggested its possibilistic version. In fact all the other member algorithms from basic subfamily have been extended to incorporate this new notion. Currently, the use of image data is growing vigorously and constantly, accounting to huge figures leading to big data. Moreover, since image segmentation happens to be one of the most time consuming processes, industries are in the need of algorithms which can solve this problem at a rapid pace and with high accuracy. In this paper, we propose to combine the notions of kernel and possibilistic approach together in a distributed environment provided by Apache Hadoop. Integrate this combined notion with map-reduce paradigm of Hadoop and put forth three novel algorithms.

Pan Shao (2016) presents a novel approach for unsupervised change detection in multi-temporal remotely sensed images. This method addresses the problem of the

analysis of the difference image by proposing a novel and robust semi-supervised fuzzy C-means (RSFCM) clustering algorithm. The advantage of the RSFCM is to further introduce the pseudolabels from the difference image compared with the existing change detection methods; these methods, mainly use difference intensity levels and spatial context. First, the patterns with a high probability of belonging to the changed or unchanged class are identified by selectively thresholding the difference image histogram. Second, the pseudolabels of these nearly certain pixel-patterns are jointly exploited with the intensity levels and spatial information in the properly defined RSFCM classifier in order to discriminate the changed pixels from the unchanged pixels. Specifically, labeling knowledge is used to guide the RSFCM clustering process to enhance the change information and obtain a more accurate membership; information on spatial context helps to lower the effect of noise and outliers by modifying the membership. RSFCM can detect more changes and provide noise immunity by the synergistic exploitation of pseudolabels and spatial context. The two main contributions of this study are as follows: it proposes the idea of combining the three information types from the difference image, namely, intensity levels, labels, and spatial context; and it develops the novel RSFCM algorithm for image segmentation and forms the proposed change detection framework. The proposed method is effective and efficient for change detection as confirmed by six experimental results of this study.

Indrajit Saha, (2015) propose a hybrid technique, called Rough Possibilistic Type-2 Fuzzy C-Means clustering with the integration of Random Forest. In the proposed method, possibilistic approach handles the noisy data better, whereas the other various uncertainties and inherent vagueness are taken care by type-2 fuzzy set and rough set theories. After clustering, it produces rough and crisp points. Thereafter, such crisp points are used to train the Random Forest classifier in order to classify the rough points for yielding better clustering solution. The performance of the proposed method has been demonstrated in comparison with several other recently proposed methods for MR brain image segmentation. Finally, superiority of the results produced by the proposed hybrid method has also been validated through statistical significance test.

Dwarikanath Mahapatra (2016) proposes an accuracy of the final segmentation determines to a large extent the accuracy of automated segmentation and disease detection algorithms. Medical image segmentation requires consensus ground truth segmentations to be derived from multiple expert annotations. A novel approach is proposed that obtains consensus segmentations from experts using graph cuts (GC) and semi supervised learning (SSL). Popular approaches use

iterative Expectation Maximization (EM) to estimate the final annotation and quantify annotator's performance. Such techniques pose the risk of getting trapped in local minima. Propose a self consistency (SC) score to quantify annotator consistency using low level image features. SSL is used to predict missing annotations by considering global features and local image consistency. The SC score also serves as the penalty cost in a second order Markov random field (MRF) cost function optimized using graph cuts to derive the final consensus label.

Anupama Namburu, (2016) Fuzzy sets, rough sets are efficient tools to handle uncertainty and vagueness in the medical images and are widely used for medical image segmentation. Soft sets are a new mathematical approach to uncertainty and vagueness. In this paper, a hybrid segmentation algorithm based on soft sets namely soft fuzzy rough c-means is proposed to extract the white matter, gray matter and the cerebro spinal fluid from MR brain image with bias field correction. In this algorithm, soft fuzzy rough approximations are applied to obtain the rough regions of image. These approximations are free from defining thresholds, weight parameters and are less complex compared to the existing rough set based algorithms. Soft sets use similarity coefficients to find the similarity of the clusters formed in present and previous step. The proposed algorithm does not involve any negative region; hence all the pixels participate in clustering avoiding clustering mistakes. Also, the histogram based centroids choose the centroids close to the ground truth that in turn effect the definition of approximations, standardizing the clusters. The proposed algorithm evaluated through simulation, compared it with existing k-means, rough k-means, fuzzy c-means and other hybrid algorithms. The soft fuzzy rough c-means algorithm outperforms the considered algorithms in all analyzed scenarios even in extracting the tumor from the brain tissue.

III. PROPOSED SYSTEM

Magnetic Resonance Imaging (MRI) is a widely used imaging technique to assess these tumors, but the large amount of data produced by MRI prevents manual segmentation in a reasonable time, limiting the use of precise quantitative measurements in the clinical practice. So, automatic and reliable segmentation methods are required; however, the large spatial and structural variability among brain tumors make automatic segmentation a challenging problem. Propose an automatic segmentation method based on Convolutional Neural Networks (CNN), exploring small kernels. The use of small kernels allows designing a deeper architecture, besides having a positive effect against over-fitting, given the fewer number of weights in the network.

Also investigated the use of intensity normalization as a pre-processing step, which though not common in CNN-based segmentation methods, proved together with data augmentation to be very effective for brain tumor segmentation in MRI images.

PRE-PROCESSING

MRI images are altered by the bias field distortion. This makes the intensity of the same tissues to vary across the image. To correct it, applied the N4ITK method. However, this is not enough to ensure that the intensity distribution of a tissue type is in a similar intensity scale across different subjects for the same MRI sequence, which is an explicit or implicit assumption in most segmentation methods. In fact, it can vary even if the image of the same patient is acquired in the same scanner in different time points, or in the presence of a pathology. So, to make the contrast and intensity ranges more similar across patients and acquisitions, apply the intensity normalization method proposed by Nyul on each sequence.

After normalizing the MRI images, compute the mean intensity value and standard deviation across all training patches extracted for each sequence. Then, normalize the patches on each sequence to have zero mean and unit variance.

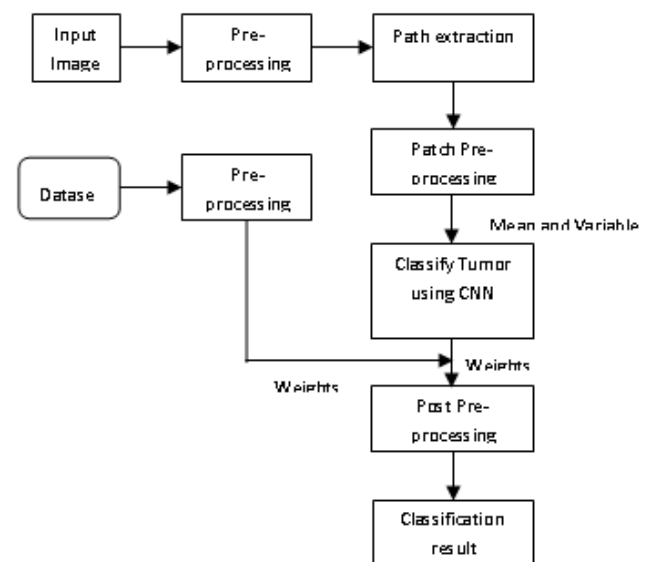


Fig.3.1 Proposed Block Diagram

CONVOLUTIONAL NEURAL NETWORK

CNN were used to achieve some breakthrough results and win well-known contests. The application of convolutional layers consists in convolving a signal or an

image with kernels to obtain feature maps. So, a unit in a feature map is connected to the previous layer through the weights of the kernels. The weights of the kernels are adapted during the training phase by back propagation, in order to enhance certain characteristics of the input. Since the kernels are shared among all units of the same feature maps, convolutional layers have fewer weights to train than dense FC layers, making CNN easier to train and less prone to over-fitting. Moreover, since the same kernel is convolved overall the image, the same feature is detected independently of the location – translation invariance. By using kernels, information of the neighborhood is taken into account, which is a useful source of context information. Usually, a nonlinear activation function is applied on the output of each neural unit. If stack several convolutional layers, the extracted features become more abstract with the increasing depth. The first layers enhance features such as edges, which are aggregated in the following layers as motifs, parts, or objects. The following concepts are important in the context of CNN

Initialization

It is important to achieve convergence. Use the Xavier initialization. With this, the activations and the gradients are maintained in controlled levels, otherwise back-propagated gradients could vanish or explode.

Activation Function

It is responsible for non-linearly transforming the data. Rectifier linear units defined as

$$f(x) = \max(0, x)$$

Were found to achieve better results than the more classical sigmoid, or hyperbolic tangent functions, and speed up training. However, imposing a constant 0 can impair the gradient flowing and consequent adjustment of the weights. Cope with these limitations using a variant called leaky rectifier linear unit (LReLU) that introduces a small slope on the negative part of the function.

Pooling

It combines spatially nearby features in the feature maps. This combination of possibly redundant features makes the representation more compact and invariant to small image changes, such as in significant details; it also decreases the computational load of the next stages. To join features it is more common to use max-pooling or average-pooling.

Regularization

It is used to reduce over-fitting. Use Dropout in the FC layers. In each training step, it removes nodes from the network with probability p . In this way, it forces all nodes of the FC layers to learn better representations of the data, preventing nodes from co-adapting to each other. At test time, all nodes are used. Dropout can be seen as an ensemble of different networks and a form of bagging, since each network is trained with a portion of the training data.

Data Augmentation

It can be used to increase the size of training sets and reduce over-fitting. Since the class of the patch is obtained by the central voxel, restricted the data augmentation to rotating operations. Some authors also consider image translations, but for segmentation this could result in attributing a wrong class to the patch. So, increased our data set during training by generating new patches through the rotation of the original patch. In our proposal, used angles multiple of 90, although another alternative will be evaluated.

Architecture

Aim at a reliable segmentation method however, brain tumors present large variability in intra-tumoral structures, which makes the segmentation a challenging problem. To reduce such complexity, designed a CNN and tuned the intensity normalization transformation for each tumor grade – LGG and HGG.

The architecture used for HGG is deeper than the one for LGG, because going deeper did not improve results in the latter. To go deeper, one must include more layers with weights, which may increase over-fitting, given the smaller training set of LGG. This is supported by the need of setting Dropout with $p = 0:5$ in LGG, while it is $p = 0:1$ in HGG, since the database used for evaluation contained more HGG than LGG cases. Additionally, the appearance and patterns are different in HGG and LGG. Since are doing segmentation, need a precise sense of location. Pooling can be positive to achieve invariance and to eliminate irrelevant details however; it can also have a negative effect by eliminating important details. Apply overlapping pooling with 3 x3 receptive fields and 2 x2 strides to keep more information of location.

In the convolutional layers the feature maps are padded before convolution, so that the resulting feature maps could maintain the same dimensions. In the case of HGG there are 2,118,213 weights to train, while in LGG it lowers to 1,933,701 weights because it has two less convolutional layers.

Training

To train the CNN the loss function must be minimized, but it is highly non-linear. Use Stochastic Gradient Descent as an optimization algorithm, which takes steps proportionally to the negative of the gradient in the direction of local minima. Nevertheless, in regions of low curvature it can be slow. So, also use Nesterov's Accelerated Momentum to accelerate the algorithm in those regions. The momentum μ is kept constant, while the learning rate ϵ was linearly decreased, after each epoch. Consider an epoch as a complete pass over all the training samples.

Post-processing

Some small clusters may be erroneously classified as tumor. To deal with that, impose volumetric constraints by removing clusters in the segmentation obtained by the CNN that are smaller than a predefined threshold.

DEEPER ARCHITECTURES/SMALL KERNELS

Using cascaded layers with small 3 x3 kernels has the advantage of maintaining the same effective receptive field of bigger kernels, while reducing the number of weights, and allowing more non-linear transformations on the data. To evaluate the real impact of this technique on brain tumor segmentation changed the cascaded convolutional layers before each max pooling of the proposed architecture by one layer with larger kernels with the equivalent effective receptive field. So, in HGG changed the groups of layers 1, 2, 3 and 5, 6, 7 by one convolutional layer with 7 x7 kernels each, while in the LGG changed the groups of layers 1 and 2, and 4 and 5 by one layer with 5 x5 kernels each. Using these architectures, experimented two variants for both grades: 1) maintained the 64 feature maps in the first convolutional layer and 128 in the second; 2) increased the capacity of the CNN by using wider layers, namely, 128 feature maps in the first convolutional layer and 256 in the second. Present the results obtained in the Leader board and Challenge data, while for variant 1 it was 2:1%. In the majority of metrics, the proposed method obtained higher scores than both variants with bigger kernels, with some of them with statistical significance, while the variants achieved better scores in PPV.

In the box plots, both variants seem to have larger dispersion and more outliers. In the segmentations, although the segmentations by the variants appear with good quality, the proposed method can capture more details, and variant 2 classified some non enhanced tumor inside the enhancing ring, which does not happen in the manual segmentation in HGG in

LGG the architecture with bigger kernels also identified an excess of non-enhancing tumor.

PATCH EXTRACTION PLANE

The use of 2D patches in a MRI image requires that define a plane perpendicular to an axis to extract patches. So, following the procedure defined in the previous subsection, investigated the use of patches extracted in a plane perpendicular to the Axial, Coronal, and Sagittal axis. The results in both the Leader board and Challenge data sets. As can be observed, extracting patches in the plane perpendicular to the Axial axis presented the best overall performance with a mean gain of 2:33% relative to the Coronal plane and 4:00% relative to the Sagittal plane. The Axial plane presented better DSC and PPV scores for both data sets than the Sagittal plane, but worst sensitivity for the Challenge data set and for the complete region in the Leader board data set.

Considering, this can be explained by an over-segmentation of the tumor, which is corroborated by the lower PPV score. A similar pattern is found for the Coronal plane, which was better in the enhanced region for the PPV score and in the complete region for the Sensitivity score. The better performance obtained using patches extracted in the axial plane can be explained by some acquisitions having lower spatial resolution in the Coronal and Sagittal planes, which can be considered a limitation of the BRATS databases. Finally, as an overall analysis, note some general trends across all experiments. Considering the box plots, verify a lower dispersion for the complete region, presenting also a higher mean value for the same region. This lower dispersion is less expressive in the Leader board than in the Challenge data set, which may be explained by the worst performance of the algorithms on LGG subjects in this data.

Gliomas are the brain tumors with the highest mortality rate and prevalence. These neoplasms can be graded into Low Grade Gliomas (LGG) and High Grade Gliomas (HGG), with the former being less aggressive and infiltrative than the latter. Even under treatment, patients do not survive on average more than 14 months after diagnosis. Current treatments include surgery, chemotherapy, radiotherapy, or a combination of them. MRI is especially useful to assess gliomas in clinical practice, since it is possible to acquire MRI sequences providing complementary information.

The accurate segmentation of gliomas and its intratumoral structures is important not only for treatment planning, but also for follow-up evaluations. However, manual segmentation is time-consuming and subjected to inter- and intra-rater errors difficult to characterize. Thus, physicians

usually use rough measures for evaluation. For these reasons, accurate semiautomatic or automatic methods are required. However, it is a challenging task, since the shape, structure, and location of these abnormalities are highly variable. Additionally, the tumor mass effect changes the arrangement of the surrounding normal tissues. Also, MRI images may present some problems, such as intensity in homogeneity, or different intensity ranges among the same sequences and acquisition scanners.

In brain tumor segmentation, find several methods that explicitly develop a parametric or non-parametric probabilistic model for the underlying data. These models usually include a likelihood function corresponding to the observations and a prior model. Being abnormalities, tumors can be segmented as outliers of normal tissue, subjected to shape and connectivity constrains. Other approaches rely on probabilistic atlases. In the case of brain tumors, the atlas must be estimated at segmentation time, because of the variable shape and location of the neoplasms.

Tumor growth models can be used as estimates of its mass effect, being useful to improve the atlases. The neighborhood of the voxels provides useful information for achieving smoother segmentations through Markov Random Fields (MRF). Zhao et al. also used a MRF to segment brain tumors after a first over segmentation of the image into super voxels, with a histogram-based estimation of the likelihood function. As observed by Menze et al. generative models generalize well in unseen data, but it may be difficult to explicitly translate prior knowledge into an appropriate probabilistic model. Another class of methods learns a distribution directly from the data. Although a training stage can be a disadvantage these methods can learn brain tumor patterns that do not follow a specific model. This kind of approaches commonly considers voxels as independent and identically distributed, although context information may be introduced through the features. Because of this, some isolated voxels or small clusters may be mistakenly classified with the wrong class, sometimes in physiological and anatomically unlikely locations. To overcome this problem, some authors include information of the neighborhood by embedding the probabilistic predictions of the classifier into a Conditional Random Field. Classifiers such as Support Vector Machines and, more recently, Random Forests (RF) were successfully applied in brain tumor segmentation.

CNNs have been used to win several object recognition and biological image segmentation challenges. Since a CNN operates over patches using kernels, it has the advantages of taking context into account and being used with raw data. In the field of brain tumor segmentation, recent

proposals also investigate the use of CNNs. Used a shallow CNN with two convolutional layers separated by max-pooling with stride 3, followed by one fully-connected (FC) layer and a soft max layer. Urban et al. evaluated the use of 3D filters, although the majority of authors opted for 2D filters. 3D filters can take advantage of the 3D nature of the images, but it increases the computational load. Some proposals evaluated two-pathway networks to allow one of the branches to receive bigger patches than the other, thus having a larger context view over the image.

Then, a cellular automaton smooths the segmentation, before a multiclass CNN discriminates the sub-regions of tumor. Extracted patches in each plane of each voxel and trained a CNN in each MRI sequence; the outputs of the last FC layer with softmax of each CNN are concatenated and used to train a RF classifier. Divided the brain tumor regions segmentation tasks into binary sub-tasks and proposed structured predictions using a CNN as learning method. Patches of labels are clustered into a dictionary of label patches, and the CNN must predict the membership of the input to each of the clusters.

IV. RESULT AND DISCUSSION

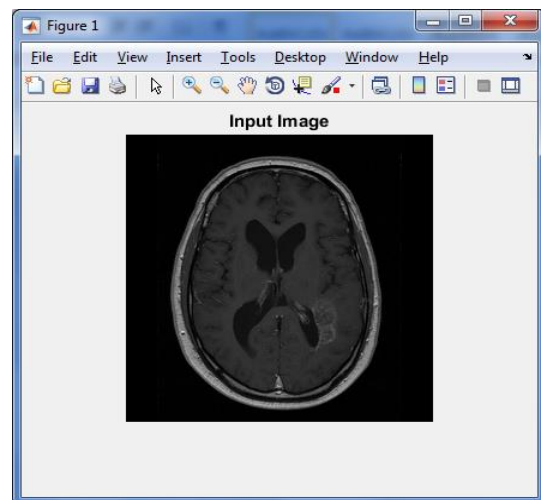


Figure 4.1 Input Image

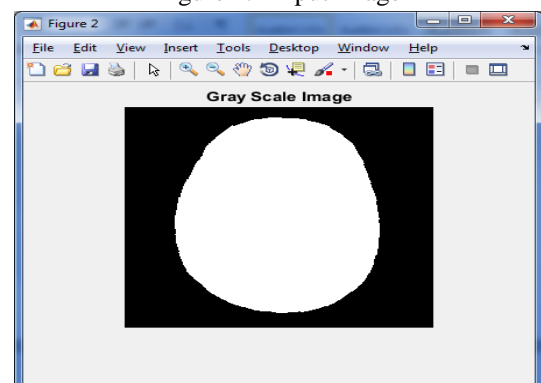


Figure 4.2 Gray Scale Image

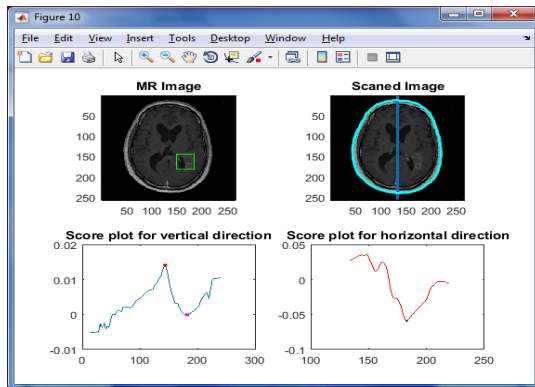


Fig 4.3 LIPC-based feature extraction

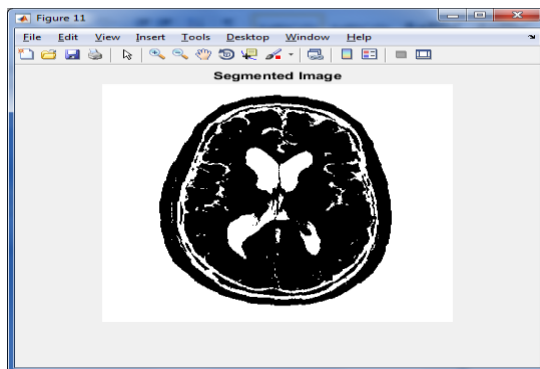


Fig 4.4 Segmented Image

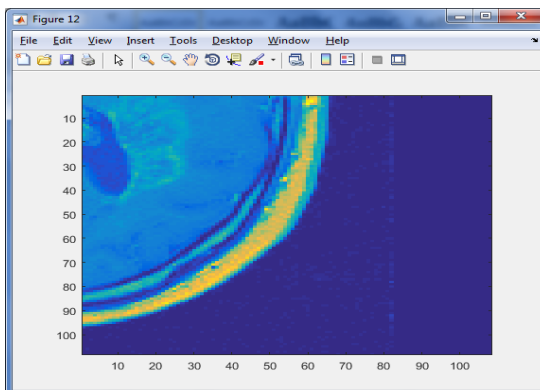


Fig 4.5 Segmented Tumor Part

V. CONCLUSION

Thus propose a novel CNN-based method for segmentation of brain tumors in MRI images. Pre-processing stage consisting of bias field correction, intensity and patch normalization. After that, during training, the number of training patches is artificially augmented by rotating the training patches, and using samples of HGG to augment the number of rare LGG classes. The CNN is built over convolutional layers with small 3X3 kernels to allow deeper architectures. Augmenting training data set by rotating the patches as well as by sampling from classes of HGG that were underrepresented in LGG. Found that data augmentation was

also quite effective, although not thoroughly explored in Deep Learning methods for brain tumor segmentation. Also, we investigated the potential of deep architectures through small kernels by comparing our deep CNN with shallow architectures with larger filters. We found that shallow architectures presented a lower performance, even when using a larger number of feature maps. Finally, verified that the activation function LReLU was more important than ReLU in effectively training our CNN.

In future work propose a label-fusion aided deep learning approach for automatically segmenting isointense infant brain images into white matter, gray matter and cerebrospinal fluid using T1- and T2-weighted magnetic resonance images. A key idea of our approach is to apply the fully convolutional neural network (FCNN) to individual brain regions determined by a traditional registration-based segmentation method instead of training a single model for the whole brain.

REFERENCES

- [1] Y. Zhang, S. Ye, and W. Ding, "Based on rough set and fuzzy clustering of MRI brain segmentation," *Int. J. Biomath.*, vol. 10, no. 2, p. 1750026, 2017.
- [2] Namburu, S. S. Samayamantula, and S. R. Edara, "Generalised rough intuitionistic fuzzy c-means for magnetic resonance brain image segmentation," *IET Image Process.*, vol. 11, no. 9, pp. 777-785, 2017.
- [3] K. Tripathy and D. Mittal, "Hadoop based uncertain possibilistic kernelized c-means algorithms for image segmentation and a comparative analysis," *Appl. Soft Comput.*, vol. 46, pp. 886923, Sep. 2016.
- [4] J. P. Sarker, I. Saha, and U. Maulik, "Rough possibilistic type-2 fuzzy C-means clustering for MR brain image segmentation," *Appl. Soft Comput.*, vol. 46, pp. 527536, Sep. 2016.
- [5] A. Namburu, S. K. Samay, and S. R. Edara, "Soft fuzzy rough set-based MR brain image segmentation," *Appl. Soft Comput.*, vol. 54, pp. 456466, May 2017.
- [6] P. Shao, W. Shi, P. He, M. Hao, and X. Zhang, "Novel approach to unsupervised change detection based on a robust semi-supervised FCM clustering algorithm," *Remote Sens.*, vol. 8, no. 3, p. 264, 2016.
- [7] S. U. Kumar and H. H. Inbarani, "PSO-based feature selection and neighborhood rough set-based classification for BCI multiclass motor imagery task," *Neural Comput. Appl.*, vol. 28, no. 11, pp. 32393258, 2017.
- [8] F. Pacheco, M. Cerrada, R. V. Sánchez, D. Cabrera, C. Li, and J. V. de Oliveira, "Attribute clustering using rough set theory for feature selection in fault severity classification of

- rotating machinery”, *Expert Syst. Appl.*, vol. 71, pp. 6986, Apr. 2017.
- [9] Y. Zhang, N. Ye, R. Wang, and R. Malekian, “A method for traffic congestion clustering judgment based on grey relational analysis,” *ISPRS Int. J. GeoInf.*, vol. 5, no. 5, p. 71, 2016.
- [10] Bai, D. Dhavale, and J. Sarkis, “Complex investment decisions using rough set and fuzzy c-means: An example of investment in green supply chains,” *Eur. J. Oper. Res.*, vol. 248, no. 2, pp. 507-521, 2016.
- [11] J. Vijaya and E. Sivasankar, “Computing efficient features using rough set theory combined with ensemble classification techniques to improve the customer churn prediction in telecommunication sector,” *Computing*, vol. 100, pp. 839860, Aug. 2018.
- [12] Koundal, S. Gupta, and S. Singh, “Automated delineation of thyroid nodules in ultrasound images using spatial neutrosophic clustering and level set,” *Appl. Soft Comput.*, vol. 40, pp. 8697, Mar. 2016.
- [13] G. Nápoles, E. Papageorgiou, R. Bello, K. Vanhoof, “On the convergence of sigmoid fuzzy cognitive maps,” *Inf. Sci.*, vols. 349350, pp. 154171, Jul. 2016.
- [14] C.-T. Chen, E. C.-K. Tsao, and W. C. Lin, “Medical image segmentation by a constraint satisfaction neural network,” *IEEE Trans. Nucl. Sci.*, vol. 38, no. 2, pp. 678-686, Apr. 2016.
- [15] Y. Xue, T. Xu, H. Zhang, R. Long, and X. Huang, “SegAN: Adversarial network with multi-scale loss for medical image segmentation,” *Neuroinformatics*, vol. 16, pp. 383392, May 2018.