Brain Tumour Segmentation and Classification using Convolutional Neural Network in MRI images

Jijith M P[1], Sadhik M S[2], Prof. Linda Sara Mathew[3]

[1,2,3] Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, India

ABSTRACT

In brain tumors, gliomas are the most common and aggressive, leading to a very short life expectancy in their highest grade. Thus, treatment planning is a key stage to improve the quality of life of oncological patients. 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. Here we propose an automatic segmentation method based on Convolutional Neural Networks (CNN), exploring small 33 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. We 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. We also try to find out the area of the tumor effected potion in the input image. There are mainly three stages includes. The first stage is pre-processing, second stage is classification via deep neural network and the final stage is the post-processing .

Keywords: Convolutional Neural Networks (CNN), Magnetic Resonance Imaging (MRI)

I. INTRODUCTION

Data mining is the process of extracting hidden knowledge, useful trends and patterns from large databases which is used by organizations for decision making purpose. Data mining deals with the difficulty of extracting patterns from the information by paying suspicious attention to computing, communication and human- computer interface issues. There are various data mining techniques available like clustering, classification, prediction, outlier analysis. Convolutional neural networks are widely used in pattern- and image-recognition problems as they have a number of

advantages compared to other techniques. This white paper covers the basics of CNNs including a description of the various layers used. Using traffic sign recognition as an example, we discuss the challenges of the general problem and introduce algorithms and implementation software developed by Cadence that can trade off computational burden and energy for a modest degradation in sign recognition rates. We outline the challenges of using CNNs in embedded systems and introduce the key characteristics of the Cadence Tensilica Vision digital signal processor for Imaging and Computer Vision and software that make it so suitable

1

for CNN applications across many imaging and related recognition tasks.

Clusters are very useful to extract interesting patterns from large data. But as the application grows it generates large amount of data. Pattern representation refers to the number of classes, the number of available patterns, and the number, type, and scale of the features available to the clustering algorithm. The growing need for clustering algorithms is required to handle huge size of databases that is common nowadays. Clustering has become an increasingly important task in modern applications such as marketing, bio informatics, spatial analysis, molecular biology as well. In all these applications generates a large amount of data. On the other hand data are originally collected at different sites. This leads us to the requirement of clustering large data sets. Typical clustering algorithms cluster a data set stored in a single site. Moreover, most of the data collected in many problems seem to have some inherent properties that lend themselves to natural groupings. Clustering algorithms are used extensively not only to organize and categorize data, but are also useful for data compression and model construction

Emerging data mining applications place many requirements on clustering techniques, motivating need for developing algorithms to handle large data sets. Such as 1) Effective treatment of high dimensionality and type of attributes algorithm can handle, where an object typically has dozens of attributes and the domain for each attribute can be large. Many dimensions or combinations of dimensions can have noise or values that are uniformly distributed. Therefore, distance functions that use all the dimensions of the data may be ineffective. 2) Interpretability of results: It is particularly important to have simple representations because most visualization techniques do not work well in high dimensional spaces. 3) Scalability and usability to large datasets: The clustering technique should be fast and scale with the number of dimensions and the size of input. It should be insensitive to the order in which the data records are presented. 4) Handling outliers and ability to find clusters of irregular shape. 5) Data order dependency: Finally, it should not presume some canonical form for data distribution. These are the main requirements considered while clustering large data sets and is driven by the need of applying algorithms for clustering datasets separated with noise efficiently.
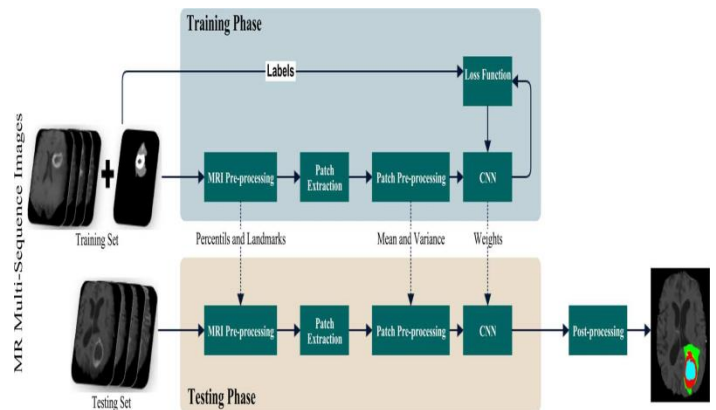
Fig 1: Stages in MRI image Processing

II. LITERATURE SURVEY

Literature survey is the process of referring some base papers that are related to our topic and then choose the best methods to implement our system.

In 2014, Ed-EdilyMohd Azhari[1] proposed gives the details about cells each type of cell has special functions. Most cells in the body grow and then divide in an orderly way to form new cells as they are needed to keep the body healthy and work properly. When cells lose the ability to control their growth, they divide too often and without any order. The extra cells form a mass of tissue called a tumor. Brain tumors are created by abnormal and uncontrolled cell division in brain itself. Generally, if the growth becomes more than 50%, then the patient may not be able to recover. Hence detection of brain tumor at its early stage with its accurate diagnosis is very important. Identification of tumor involves tests like CT and MRI. MRI plays vital role in identifying location, size and type of brain tumor.

An image may be defined as a two-dimensional function f(x, y), where x & y are spatial coordinates, & the amplitude of at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. Digital image is composed of a finite number of elements, each of which has a particular location & value. The elements are called pixels. For the extraction of

useful features from the complex brain structure Magnetic resonance imaging (MRI) is reliable. MRI is very important in order to improve the diagnosis and treatment of brain tumor, by detecting tumor at its early stage. Segmentation of medical images is first important step in their analysis, the segmentation gives organ detection and variation of growth of tissues as a output in medical images. Some segmentation approaches are Global image threshold using Otsu's method, Region Growing, Edge Based Segmentation, K-means Clustering, Fuzzy C-means Clustering.

Clustering the process of collection of objects which are similar between them and are dissimilar objects belonging to other clusters. . Region growing is a technique of segmentation in which pixels with similar intensities are grouped in order to find the regions directly. This group of pixels belonging to the region of focus is known as seeds. K-mean is example of exclusive clustering algorithm. In overlapping clustering, one data (pixel) is belonging two or more clusters.

Watershed Segmentation: It is one of the best methods to group pixels of an image on the basis of their intensities. Pixels falling under similar intensities are grouped together. It is a good segmentation technique for dividing an image to separate a tumor from the image Watershed is a mathematical morphological operating tool. Watershed is normally used for checking output rather than using as an input segmentation technique because it usually suffers from over segmentation and under segmentation. Morphological image processing (or morphology) describes a range of image processing techniques that deal with the shape (or morphology) of features in an image and morphological operations are typically applied to remove imperfections introduced during segmentation, and so typically operate on bi-level images i.e. binary images.

In 2014, Kailash Sinha[2] describes The structure and function of the brain and researchers using MRI imaging techniques Brain tumor extraction and its analysis are challenging tasks in medical image processing because brain image and its structure is complicated that can be analyzed only by expert radiologists. This paper presents a comparative study of three segmentation methods implemented for tumor detection. The methods include k-means clustering with watershed segmentation algorithm, optimized k-means clustering with genetic algorithm and optimized c- means clustering with genetic algorithm. Traditional k-means algorithm is sensitive to the initial cluster centers. Genetic c-means and k-means clustering techniques are used to detect tumor in MRI of brain images .At the end of process the tumor is extracted from the MR image and its exact position and the shape are determined .Over segmentation and sensitivity to false edges are other difficulties in ordinary k-means method. Determination of exact location and area of brain tumor using k-means method becomes very difficult and hence use of genetic algorithm is suggested. Fittest is searched by the algorithm and hence used in optimization tasks, The implementation of genetic algorithm begins with an initial population of chromosomes which are randomly selected. Chromosome is a long thread of DNA.

The c-means clustering method has been implemented and its performance can be improved by using optimization with the use of genetic algorithm. The combined method results an improvement in segmentation efficiency and higher area of affected region extraction and detection. MRI images were segmented using k-means clustering and Watershed algorithm. The method is implemented using process of two stages. The first stage of the process uses k-means clustering and primary segmentation results are produced for the brain MRI images. Second stage of the process is applied as watershed segmentation algorithm to improve the results of

the primary segmentation; and the results obtained are final results.

Datta et al (2011) introduced colour-based segmentation using k-means clustering for brain tumor detection. The developed algorithm shows better result than Canny based edge detection. Nandha et al (2010) designed intelligent system to diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy c-means along with intelligent optimization tools, such as Genetic Algorithm (GA), and Particle Swarm Optimization (PSO).

Jobin et al (2012) proposed a method which integrated the k-means clustering algorithm with the marker controlled watershed segmentation algorithm.

Yang et al (2010) presented a new image segmentation algorithm W-SPK (combining watershed and K-means clustering method based on simulated annealing particle swarm optimization) to overcome the shortcomings of watershed and realize fast and accurate image segmentation [9]. Sasikala et al (2006) presented an automatic segmentation of malignant tumor in magnetic resonance images (MRI's) of brain using optimal texture features.

III. PROPOSED SYSTEM

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, we 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 [37]. 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, we apply the intensity normalization method proposed by Nyúlon each sequence. In this intensity normalization method, a set of intensity landmarks are learned for each sequence from the training set. And are chosen for each MRI sequence as described represents the intensity at the percentile. After training, the intensity normalization is accomplished by linearly transforming the original intensities between two landmarks into the corresponding learned landmarks. In this way, the histogram of each sequence is more similar across subjects. After normalizing the MRI images, we compute the mean intensity value and standard deviation across all training patches extracted for each sequence. Then, we normalize the patches on each sequence to have zero mean and unit variance

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 easir to train and less prone to over fitting. Moreover, since the same kernel is convolved over all the image, the same feature is detected independently of the locating—translation invariance. By using kernels, information of the neighborhood is taken into account, which is a useful source of context information. Usually, a non-linear activation function is applied on the output of each neural unit. If we 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:

1) Initialization: It is important to achieve convergence. The activations and the gradients are maintained in controlled levels; otherwise back-propagated gradients could vanish or explode.

2) Activation Function: It is responsible for non-linearly transforming the data. Rectifier linear units (ReLU), defined as 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 we cope with these limitations using a variant called leaky rectifier linear unit that introduces a small slope on the negative part of the function. This function is defined as where is the leakiness parameter. In the last FC layer, we use softmax.

3) 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.

4) Regularization: It is used to reduce over fitting. We use Dropout in the FC layers. In each training step, it removes nodes from the network with probability 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.

5) Data Augmentation: It can be used to increase the size of training sets and reduce overfitting. Since the class of the patch is obtained by the central voxel, we 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, we increased our data set during training by generating new patches through the rotation of the original patch. In our proposal, we used angles multiple of 90 although another alternative will be evaluated.

6) Loss Function: It is the function to be minimized during training. We used the Categorical Cross-entropy, where represents the probabilistic predictions (after the softmax) and is the target. In the next subsections, we discuss the architecture and training of our CNN.

7) Architecture: We 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, we designed a CNN and tuned the intensity normalization transformation for each tumor grade LGG and HGG.

This is supported by the need of setting Dropout with in LGG, while it is in HGG, since the database used for evaluation contained more HGG then LGG cases. Additionally, the appearance and patterns are different in HGG and LGG. Since we are doing segmentation, we 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. We apply overlapping pooling with 33 receptive fields and 2 2 stride 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 all layers with weights, with the exception of the last that uses softmax. Dropout was used only in the FC layers.

8) Training: To train the CNN the loss function must be minimized, but it is highly non-linear. We 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, we

also use Nesterov's Accelerated Momentum to accelerate the algorithm in those regions.

Here we use post preprocessing technique, in this method Some small clusters may be erroneously classified as tumor. To deal with that, we impose volumetric constrains by removing clusters in the segmentation obtained by the CNN that are smaller than a predefined threshold.

IV. CONCLUSION

The proposed brain tumor detection and localization framework detects and localizes brain tumor in MR imaging. In summary, we propose a novel CNN-based method for segmentation of brain tumors in MRI images. We start by a 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. In designing our method, we address the heterogeneity caused by multi-site multi-scanner acquisitions of MRI images using intensity normalization. We show that this is important in achieving a good segmentation. Brain tumors are highly variable in their spatial localization and structural composition, so we have investigated the use of data augmentation to cope with such variability.

V. REFERENCES

[1] Ed-EdilyMohd. Azhari1, Muhd. MudzakkirMohd. Hatta1, Zaw ZawHtike1 and Shoon Lei Win2, \Brain Tumor Detection And Localization In Magnetic Resonance Imaging", International Journal of Information Technology Convergence and Services (IJITCS) Vol.4, No.1,February 2014.

[2] Kailash Sinha1, G.R. Sinha., \ E_cient Segmentation Methods for Tumor Detection in MRI Images", IEEE Student's Conference on Electrical, Electronics and Computer Science, 2014.

[3] Pratibha Sharma ,Manoj Diwakar ,sangam Choudhary, \ Application of Edge Detection for Brain Tumor Detection", International Journa lof Computer Applications (0975 { 8887) Volume 58, November2012.

[4] Riries Rulaningtyas1 and Khusnul Ain2, \Edge Detection for Brain Tumor Pattern Recognition".

[5] S. Datta, M. Chakraborty, \Brain Tumor Detection from Pre-ProcessedMR Images using Segmentation Techniques",Special Issue on 2nd National Conference-Computing, Communication and Sensor Network(CCSN) Published by Foundation of Computer Science, NewYork, USA. vol.2, pp.1-5, 2011.

[6] Gopal, N.N., Karnan, M., "Diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy C Meansalong with intelligent optimization Techniques", IEEE International Conference on Computational Intelligence and Computing Research(ICCIC),Vol.2, No.3, pp.1-4, 2010.

[7] Christ, M. J., Parvathi, R. M. S., \Magnetic Resonance Brain Image Segmentation", International Journal of VLSI Design and Communication Systems, Vol.3,No.4, pp.121-133,2012.

[8] Wenli Yang, Zhiyuan Zeng, Sizhe Zhang, \Application of Combining Watershed and Fast Clustering Method in Image Segmentation",Computer Modeling and Simulation. ICCMS Second International Conference on, Vol.3,No.,pp.170-174,22-24,Jan.2010.

[9] P.Tamije;V. Palanisamy; T. Purusothaman: "Performance Analysis of Clustering Algorithms in Brain Tumor Detection of MR Images" European Journal of Scientific Research, ISSN 1450-216X Vol.62 No.3 (2011), pp. 321-330.

[10] Ratan, Rajeev, Sanjay Sharma, and S. K. Sharma. "Brain tumor detection based on multi-parameter MRI image analysis." International Journal on Graphics, Vision and Image Processing vol 9.no.3, pp.9-17,2009.

[11] S.K.Bandyopadhyay and D.Saha, "Brain region extraction volume calculation," UNIASCIT, vol. 1, no. 1, pp. 44-48, 2011.

[12] Gopal, N.N.; Karnan, M., "Diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy C Means along with intelligent optimization techniques," IEEE International Conference on Computational Intelligence and Computing Research (ICCIC),vol.2, no.3, pp.1-4, 2010.

[13] Amanpreet Kaur; Gagan Jindal "Tumor Detection Using Genetic Algorithm" International Journal on Computer Science and Technology, vol. 4, no.1,pp. 423-427 2013.