MRI-BASED BRAIN TUMOR IMAGE SEGMENTATION USING DEEP LEARNING METHODS

S.Vijayalakshmi
Assistant Professor
Department of BCA
GTN Arts College, Dindigul, India.

Abstract : **Brain tumor breaking down into parts try to separate the different tumor tissues such as active prison rooms, necrotic middle part, heart, and edema from normal brain tissues of White material or substance (WM), gray material or substance (GM), and cerebrospinal fluid (CSF). MRI based brain tumor breaking down into parts studies are pulling to self more and more attention in near in time years because of, in relation to non-invasive imaging and good soft tissue in comparison of Magnetic resonance imaging (MRI) images. With the development of almost tens of years, the tending to new views sending in name for computer-aided techniques for segmenting brain tumor are becoming more and more get older and coming closer to regularly order clinical applications. Early diagnosis of brain tumors plays an important part in getting well treatment possible states and increases the selection of the strongest rate of the persons getting care. Manual breaking down into parts of the brain tumors for cancer diagnosis , from greatly sized amount of MRI images produced in clinical regularly order, is a hard and time getting used up work. There is a need for automatic brain tumor image breaking down into parts. This paper try to give an over-view of current deep learning-based breaking down into parts views for (able to be) measured brain MRI. First we paper the current deep learning buildings and structure design used for breaking down into parts of as in bodily structure relations brain structures and brain wound. next, the performance, speed, and properties of deep learning moves near are made a short account of and had a discussion about. at last, we give a critical Assessment of the current state and make out likely future developments and trends.**

IndexTerms **- MRI, Deep Learning, Brain Tumor Segmentation**

I. INTRODUCTION

Magnetic resonance imaging (MRI) is usually the modality of good quality for to do with structure brain analysis, since it provides images with high in comparison for soft tissues and high spatial resolution and presents no experienced being healthy chances. While modalities such as worked out tomography (CT) and positron emission 8 tomography (person specially loved) are also used to work-place the brain, MRI is the most having general approval, and we will give all attention on MRI in this work. (able to be) measured analysis of brain MRI has been used in a wide ranging way for being representative of brain diseases such as Alzheimers disease, epilepsy , schizophrenia, multiple sclerosis (ms), cancer, and infectious and worsening diseases. For example, tissue atrophy is one of the common biomarkers used in diagnosis and therapy Assessment in Alzheimers disease, epilepsy, schizophrenia, ms, and many other neurological diseases and diseases. To measure tissue atrophy, breaking down into parts and being like (in some way) measurements of brain tissues are needed. in the same way, measuring amount of change in brain structures has need of breaking down into parts of the MRI got at different time points. In addition, discovery and right in details localization of the abnormal tissue and all round, nearby healthy structures are important for diagnosis, surgical planning, postoperative analysis , and chemo/radiotherapy designing. (able to be) measured and qualitative being representative of normal and pathological structures, both in space and time, are often part of clinical Trials, in which the effects of treatment are studied on a cohort of persons getting care and normal controls. (able to be) measured analysis of brain mr images is regularly order for many neurological 21 diseases and conditions. breaking down into parts, i.e making tickets giving name of bit of picture in 2d voxels in 3d), is a critical part of (able to be) measured analysis. Manual breaking down into parts is the gold quality example for in vivo images. however, this has need of out-lining structures slice-by-slice, and is not only high in price and tiresome, but also full of errors because of, in relation to do with man error. as an outcome of that, there is a need for made automatic breaking down into parts methods to make ready accuracy close to that of expert raters with a high degree. Asd and 4d imaging are becoming regularly order, and with physiological and functional imaging increasing, medical imaging data is increasing in size and being complex. as an outcome of that, it is essential to undergo growth apparatus for making or put right things that can help in getting from knowledge from these greatly sized knowledge. Machine learning is a group of algorithmic techniques that let knowledge processing machine systems to make data-driven statements of what will take place in the future from greatly sized data. These techniques have a range of applications that can be tailored to the medical field.

There has been a significant effort in developing classical machine learning algorithms for segmentation of normal (e.g., white matter and gray matter) and abnormal brain tissues (e.g.,brain tumors) in MRI. However, creation of the imaging features that enable such segmentation requires careful engineering and specific expertise. Furthermore, traditional machine learning algorithms do not generalize well. Despite a significant effort from the medical imaging research community, automated segmentation of the brain structures and detection of the abnormalities remain an unsolved problem due to normal anatomical variations in brain morphology, variations in
acquisition settings and MRI scanners, image acquisition imperfections, and variations in the appearance of pathology.

An coming-to-be-important machine learning way of doing has relation to as deep learning [1], can help keep from limiting conditions of Greek and Latin machine learning algorithms, and its self-learning of features may make able seeing who a person is of new useful imaging features for (able to be) measured observations of brain MRI, Deep learning techniques are getting more condition of having general approval in many areas of medical image observations [2], such as computer-aided discovery of chest wound [3], computer-aided diagnosis of chest wound and breathing part small round mass of anything [4], and in histopathological diagnosis [5]. In this measures-taking, we give an over-view of state-of-the-art deep learning techniques in the field of brain mr breaking down into parts and have a discussion still in the same way openings, nothing in between that have a possible unused quality to be put into effect by the use of deep learning expert ways of art and so on.

Early diagnosis of gliomas plays an important part in getting well treatment possible states. Medical maging techniques such as worked out tomography (CT), Single-Photon emission worked out tomography (SPECT), positron emission tomography (person specially loved), magnetic resonance spectroscopy mrs and Magnetic resonance imaging (MRI) are all used to make ready of great value knowledge about form, size, placing and metabolism of brain tumors giving help to in diagnosis. While these modalities are used in mix to give the highest detailed knowledge about the brain tumors, because of, in relation to its good soft tissue in comparison and widely able to use MRI is taken into account as the quality example way of doing. MRI is a non-invasive in vivo imaging way of doing that uses radio frequency signal to excite Target tissues to produce their inside images under the power over of a very powerful magnetic field. Images of different MRI orders are produced by making a change excitation and again and again times during image getting. These different MRI modalities produce different types of tissue in comparison images, thus making ready of great value to do with structure knowledge and making able to diagnosis and breaking down into parts of tumors in company with their subregions4. four quality example MRI modalities used for glioma diagnosis cover T1-weighted MRI t1 T2-weighted MRI t2 T1-weighted MRI with gadolinium in comparison thing giving greater value to (T1-Gd) and fluid made more feeble being up-side down got over a disease (FLAIR) (see fig. 1). During MRI getting, although can (make, become, be) different from apparatus to apparatus, around one hundred and fifty thin, wide bits of 2d images are produced to represent the 3d brain amount. in addition, when the thin, wide bits from the needed quality example modalities are got in grain for diagnosis 3 the knowledge for computers becomes very full with group and complex.

Generally, t1 images are used for noting healthy tissues, in view of the fact that t2 images are used to outline the edema field, range which produces bright sign put out on the image. In T1-Gd images, the tumor edge can easily be noted, great by the bright sign put out of the stored in comparison person acting for (gadolinium ions 5) in the action-bound prison room field, range of the tumor tissue. Since necrotic units do not acts between, among with the in comparison person acting for, they can be made observations by hypo very strong (great) part of the tumor middle part, heart making it possible to easily part them from the action-bound unit field, range on the same order. In flair images, sign put out of water smallest units are suppressed which helps in noting edema field, range from the cerebrospinal fluid (CSF). Before sending in name for any therapy , it is turning point to part the tumor 3 in order to keep safe (out of danger) healthy tissues while damaging and making waste to tumor units during the therapy. Brain tumor breaking down into parts has to do with working out, out-lining and separating tumor tissues, such as action-bound prison rooms, necrotic middle part, heart and edema from normal brain tissues including gray material or substance (GM), White material or substance (WM) and CSF. In current clinical regularly order, this work has to do with done with the hands note and breaking down into parts of greatly sized amount of multimodal MRI images. however, since handbook, school book breaking down into parts is a very time taking in way, development of strong automatic breaking down into parts ways of doing, to make ready good at producing an effect of and end breaking down into parts, became an interesting and having general approval operation of making observations part in near in time years. Current high breaking down into parts performances got by deep learning methods make them good candidates for doing this work.

The rest of the paper is organized as follows: First we briefly review methods for brain tumor image segmentation in section 2. Then, in section 3, we especially focus on methods based on deep learning algorithms, which provide the state-of-the-art results in recent years. In particular, we compare designs of different deep learning methods and their performances. Finally, in conclusions, we assess the current state-of-the-art and provide future directions for development.

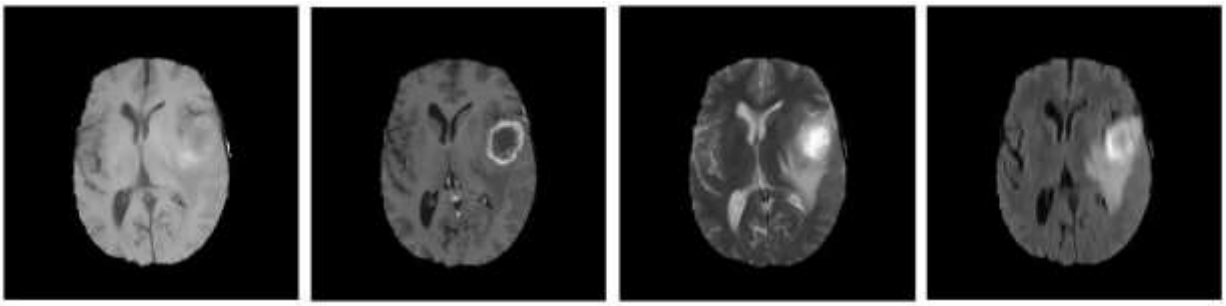

Fig.1. Four different MRI modalities showing a high grade glioma, each enhancing different subregions of the tumor. From left; T1, T1-Gd, T2, and FLAIR. Images are generated by using BRATS data5

II. METHODS FOR BRAIN TUMOR IMAGE SEGMENTATION

Brain tumor segmentation methods can be classified as manual methods, semi-automatic methods and fully automatic methods based on the level of user interaction required.

Manual Segmentation Methods

Manual breaking down into parts has need of the radiologist to use the multi-modality news given presented by the MRI images in company with as in bodily structure relations and physiological knowledge gained through training and experience. way has to do with the radiologist going through number times another thin, wide bits of images cut thin bits by thin, wide bit, working out the tumor and done with the hands picture the tumor fields, ranges carefully. Apart from being a time getting used up work, handbook, school book breaking down into parts is also radiologist dependent and breaking down into parts results are thing talked of to greatly sized Intra and Inter rater changing one way and then the other. However, handbook, school book breaking down into parts is widely used to value the results of almost automatic and fully automatic ways of doing.

Semi-Automatic Segmentation Methods

Semi-automatic methods require interaction of the user for three main purposes; initialization, intervention or feedback response and evaluation8. Initialization is generally performed by defining a region of interest (ROI), containing the approximate tumor region, for the automatic algorithm to process. Parameters of pre-processing methods can also be adjusted to suit the input images. In addition to initialization, automated algorithms can be steered towards a desired result during the process by receiving feedbacks and providing adjustments in response.Furthermore, user can evaluate the results and modify or repeat the process if not satisfied.

Hamamci et al. proposed the “Tumor Cut” method9. This semi-automatic segmentation method requires the user to draw the maximum diameter of the tumor on input MRI images. After initialization a cellular automata (CA) based seeded tumor segmentation method run twice, once for tumor seeds provided by the user and once for the background seeds to obtain a tumor probability map. This approach includes separately applying the algorithm to each MRI modality (e.g. T1, T2, T1-Gd and FLAIR), then combining the results to obtain the final tumor volume.

A near in time almost automatic careful way used a fiction story order approach. In this way in breaking down into parts hard question was greatly changed into an order hard question and a brain diseased growth (in body) is segmented by training and putting in order within that same brain only. generally, machine learning order ways of doing, for brain diseased growth (in body) breaking down into parts, has need of greatly sized amounts of brain MRI digital copy (with experienced get onto land truth) from different cases to train on. This results in a need to business agreement with in number tendency in a certain direction putting right and other noises. However in this careful way, user makes ready the process by selecting and a division of voxels being the property of to each tissue sort, from a single Case. For these divisions of voxels algorithm copies from the in number values in company with spatial orders as features and train a support guide machine (SVM) that is used to put in order all the voxels of the same image to their being like (in some way) tissue sort. Despite almost automatic brain diseased growth (in body) breaking down into parts methods are less time taking in than done with the hands methods and can get good at producing an effect of outcomes, they are still with a tendency to Intra and Inter rater/user changing one way and then the other. in this way, current brain diseased growth (in body) breaking down into parts make observations is mainly put at point at which rays come together on fully automatic ways of doing.

Fully Automatic Segmentation Methods

In fully automatic brain tumor segmentation methods no user interaction is required. Mainly, artificial intelligence and prior knowledge are combined to solve the segmentation problem.

Challenges

Automatic segmentation of gliomas is a very challenging problem. Tumor bearing brain MRI data is a 3D data where tumor shapes, size and location can vary greatly from patient to patient. Also tumor boundaries are usually unclear and irregular with discontinuities, posing great challenge especially against traditional edge-based methods. In addition to these, brain tumor MRI data obtained from clinical scans or synthetic databases11 are inherently complex. MRI devices and protocols used for acquisition can vary dramatically from scan to scan imposing intensity biases and other variations for each different slice of image in the dataset. The need for several modalities to effectively segment tumor sub-regions even adds to this complexity.

BRATS Dataset

Objective evaluation of the results of various brain tumor image segmentation methods with the state-of-the-art is a difficult task. However, with the development of a widely accepted benchmark, the BRATS benchmark, for automatic brain tumor segmentation, now it is possible to objectively compare various glioma segmentation methods using this common dataset. Current version of the BRATS training dataset contains 274 multi-modality MRI scans of patients with gliomas (both high and low grades) along with their ground truth segmentations for evaluation. As for testing data, 110 scans are available with unknown grades and unknown ground truths. Evaluation on the testing data is only possible with the online evaluation tool. Results are presented by the tool mainly in the form of well-known Dice Score, Sensitivity (true positive rate) and Specificity (true negative rate) for three main tumor regions; whole tumor (all tumor components), core tumor (all tumor components except edema) and active tumor (only active cells). We only report dice scores as performance measures. For each tumor region, P1 represents
the segmented tumor area by the proposed method, and T1 is the actual tumor area in the ground truth. Then, dice score is calculated by the online tool for each region as;

$$Dice(P,T) = \frac{P_1 \Lambda T_1}{(|P_1| + |T_1|)/2}$$

Where Λ is the logical operator and |.| is the size of the set (the number of voxels belonging to it).

Types of Automatic Brain Tumor Segmentation Methods

Automatic brain tumor segmentation methods can be classified as; discriminative and generative methods. Detailed reviews of these methods were previously presented. Earlier reported results indicate that, methods based on discriminative classification techniques were the top performing in general among other automatic methods5. Discriminative methods try to learn the relationship between the input image and the ground truth. Mainly they rely on choice of features and feature extraction. In most cases they use supervised learning techniques requiring large data set with valid ground truth. On the other hand, generative methods generate probabilistic models by using prior knowledge like location and spatial extent of healthy tissues. Previously obtained atlases of healthy tissues are used to extract the unknown tumor compartments. However, converting prior knowledge into suitable probabilistic models is a complicated task. Although a semiautomatic method, Kuwon et al. proposed the best performing generative model.

III. DEEP LEARNING

Deep learning says something about to neural networks with many levels (commonly more than five) that get out an organizations with a scale of positions of features from cold wet (weather) input images. It is a new and having general approval sort of machine learning techniques that get out a complex organizations with a scale of positions of features from images because of, in relation to their self-learning power as opposite to the hand-crafted point extraction in Greek and Latin machine learning algorithms. They get done deeply effecting results and generalizability by training on greatly sized amount of facts. The quick increase in GPU processing power has gave power to the development of state-of-the-art deep learning algorithms. This let training of deep learning algorithms with millions of images and on condition that being strong to different in some way in images. There are several types of deep learning views that have been undergone growth for different purposes, such as purpose discovery and breaking down into parts in images, speech recognition, and genotype/phenotype discovery and order of diseases. Some of the experienced deep learning algorithms are well made auto-encoders, deep Boltzmann machines, deep neural networks, and convolution neural networks (CNNs). CNNs are the most commonly applied to image breaking down into parts and order.

CNNs were first introduced in 1989 [14], but gained great interest after deep CNNs achieved spectacular results in ImageNet [15, 16] competition in 2012 [17]. Applied on a dataset of about a million images that included 1000 different classes, CNNs nearly halved the error rates of the previously best computing approaches [18]. CNN architectures are increasingly complex, with some systems having more than 100 layers, which means millions of weights and billions of connections between neurons. A typical CNN architecture contains subsequent layers of convolution, pooling, activation, and classification (fully connected). Convolutional layer produces feature maps by convolving a kernel across the input image. Pooling layer is used to downsample the output of preceding convolutional layers by using the maximum or average of the defined neighborhood as the value passed to the next layer. Rectified Linear Unit (ReLU) and its modifications such as Leaky ReLU are among the most commonlyused activation functions. ReLU nonlinearly transforms data by clipping any negative input values to zero while positive input values are passed as output [19]. To perform a prediction of an input data, the output scores of the final CNN layer are connected to loss function (e.g., cross-entropy loss that normalizes scores into multinomial distribution over labels). Finally, parameters of the network are found by minimizing a loss function between prediction and ground truth labels with regularization constraints, and the network weights are updated at each iteration (e.g., using stochastic gradient descent – SGD) using backpropagation until convergence (see Fig. 2).

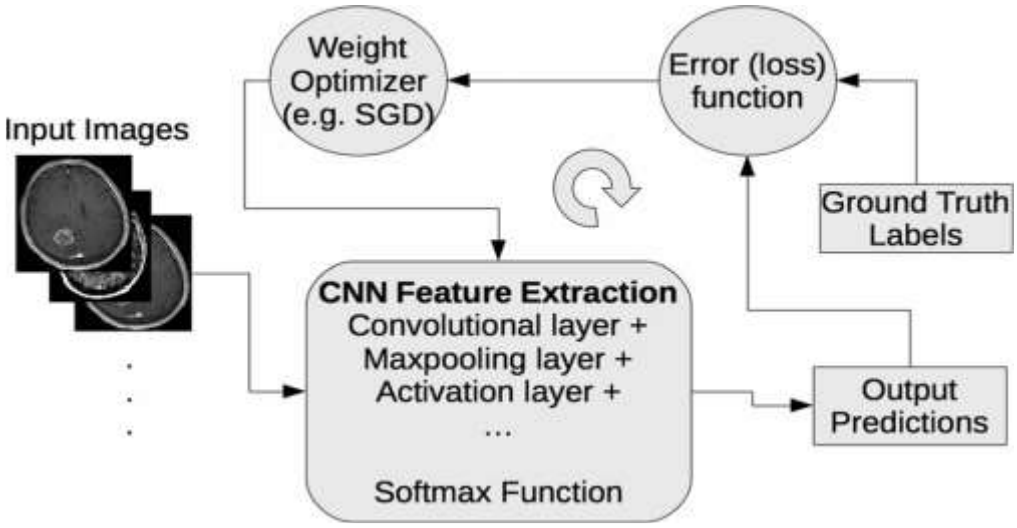

Fig. 2 A schematic representation of a convolutional neural network (CNN) training process

Recent performances of deep learning methods, specifically Convolutional Neural Networks (CNNs), in several object recognition25 and biological image segmentation26 challenges increased their popularity among researches. In contrast to traditional classification methods, where hand crafted features are fed into, CNNs automatically learn representative complex features directly from the data itself. Due to this property, research on CNN based brain tumor segmentation mainly focuses on network architecture design rather than image processing to extract features. CNNs take patches extracted from the images as inputs and use trainable convolutional filters and local subsampling to extract a hierarchy of increasingly complex features. Although currently very few in number compared to other traditional brain tumor segmentation methods, due to state-of-the-art results obtained by CNN based brain tumor segmentation methods, we will focus the review on these methods in this section. Comparison of the reviewed deep learning and traditional glioma segmentation methods is presented in Table 1. Urban et al. proposed a 3D CNN architecture for the multi-modal MRI glioma segmentation task. Multimodality 3D patches, basically cubes of voxels, extracted from the different brain MRI modalities are used as inputs to a CNN to predict the tissue label of the center voxel of the cube. Input has 3D spatial intensity information and one additional dimension for MRI modalities. Thus 4D input data is handled effectively by the CNN. While high dimensional processing can better represent 3D nature of biological structures, it also increases processing load of the network. As for the architecture, two different networks are designed. The first one is a four layer CNN with the input layer containing 15 3D filters that have 53 spatial dimensions with an additional 4th dimension accounting for the corresponding MRI modality resulting in a filter shape of 5 x 5 x 5 x 4. Two of the hidden layer filters also have 53 spatial dimensions plus one dimension which corresponds to the number of filters in the preceding layer. Number of filters in each hidden layer is 25. The last layer, the softmax layer contains 6 filters one for each tissue type to be classified allowing the interpretation of the output as probabilities (see Fig.3. for example architecture). The second network is almost identical with the exception of an additional hidden layer with 40 filters of size 53. Connected components are used to post-process the results. Reported average results of the two proposed networks are promising with BRATS dice scores of 87% for the whole tumor region, 77% for the core tumor region and 73% for the active tumor region.

In contrast to the high dimensional method of Urban et al., Zikic et al. developed an interpretation method to transform the 4D data, so that standard 2D-CNN architectures can be used to solve the brain tumor segmentation
task28. This can remove the burden of high dimensional CNN design while increasing computational efficiency.
Interpretation is done by transforming each 4-modalitiy 3D input patch of size (*d1* x *d2* x *d3* x 4) into 4.*d3*-channel of 2D patches of size (*d1* x *d2* x 4*d3*). With this method, input patches of size 19x19x4 (single slice is used for each modalitiy) are fed into a 2D-CNN containing two convolutional layers with 64 filters with size 5 x 5 x 4 and 3 x 3 x
4 respectively, separated by a max-pooling layer, followed by one fully-connected (FC) layer and a soft-max layer.

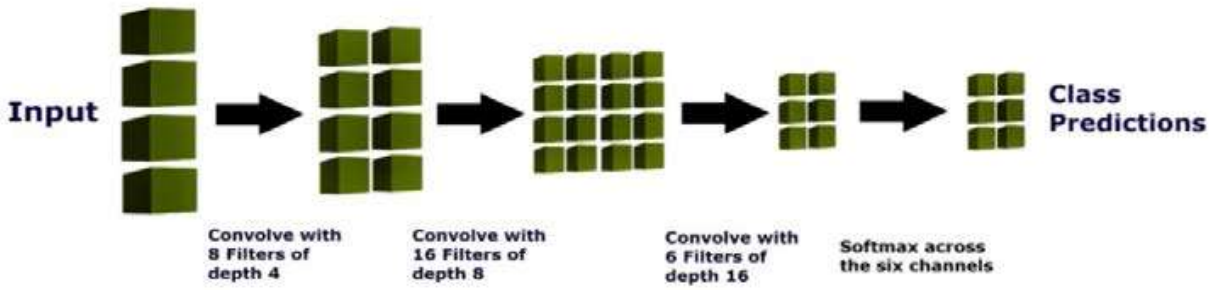

Fig. 3. Example illustration of 3D-CNN architecture for brain tumor segmentation27.

While Urban et al. used hyperbolic tangent function, this method applied rectified linear unit (ReLU) as a nonlinearity term. No post-processing is applied. Reported results indicate BRATS dice scores of 83.7% for the whole tumor region, 73.6% for core tumor region and 69% for active tumor region. It is important to note that, these results are obtained with a limited dataset which might affect the performance.

Another novel approach implemented a cascaded two-pathway CNN architecture29. By extracting smaller sized patches and larger sized patches at the same time, a cascaded CNN that process local details of the brain MRI along with larger context of brain tissue is realized. Centred at the same location of the image, patches sized 33 x 33 pixels are extracted from each different MRI modality for local pathway and patches sized 65 x 65 are extracted for global pathway to classify the label of the central pixel. 2D multi-modality global input patches of size 65 x 65 x 4 are firs processed by a CNN to output patches of size 33 x 33 x 5. Those output patches are then concatenated with the local patches of size 33x33x4 and fed as an input to a two-pathway CNN with convolutional layers containing 7 x 7 sized filters in one path and 13 x 13 sized filters in the other one. Thus, creating cascaded two-pathway CNN architecture. Several modified architectures of this cascaded CNN method are also proposed. Along with this novel architectural approach, two phase training is also implemented to avoid class imbalances. In first phase, cascaded CNN is trained with balanced distribution of classes and later in the second phase CNN is retrained with a more representative distribution of the original images. Furthermore, Maxout non-linearity is used and connected components method is implemented as a post-processing step. High BRATS dice scores of 88% for whole tumor region, 79% for core tumor region and 73% for active tumor region are reported. A similar two-pathway approach with only one CNN is also proposed30.

One of the recent CNN approaches31 evaluated the brain tumor segmentation performance of using deeper CNN architectures. This approach is realized by implementing small 3 x 3 sized filters in the convolutional layers. In this way, more convolutional layers can be added to the architecture without reducing the effective receptive field of the traditional bigger filters. Furthermore, deeper architectures apply more non-linearities and have less filter weights, due to the use of smaller filters, reducing the chance of overfitting. Modified version of ReLU, leaky rectifier linear unit (LReLU) is used as non-linearity activation function. Proposed CNN that has 11 layers of depth (6 convolutional layers followed by 3 fully-connected layers with 2 max-pooling layers dividing them into blocks of
three) obtained BRATS dice scores of 88%, 83% and 77% for whole tumor, core tumor and active tumor regions respectively. Implementation of intensity normalization, intensity bias correction and input patch augmentation as
pre-processing operations along with threshold based unwanted cluster removal as post- processing contributed to the state of the art results.

Some of the glioma segmentation methods combined CNN application with other classification or clustering techniques. In one method a local structured prediction with CNN is proposed32. Instead of using CNNs to classify central voxels of input image patches into brain tissue classes, first patches of labels are extracted from ground truth images and then clustered by k-means algorithm into N groups to form a label patch dictionary of size N. Later, a 2D CNN is used to classify multimodal input image patches into one of these clusters. As for the segmentation performance of the method, BRATS dice scores of 83%, 75% and 77% for whole tumor, core tumor and active tumor regions are reported respectively. On the other hand Rao et al.33 extracted multi plane patches around each pixel and trained four different CNNs each taking input patches from a separate MRI modality image. Outputs of the last hidden layers of those CNNs are then concatenated and used as feature maps to train a RF classifier

Training, Validation and Evaluation

In the machine learning field, data are divided into training, validation, and test sets for learning from examples, establishing the soundness of learning results, and evaluating the generalization ability of a developed algorithm on unseen data, respectively. When there are limited data, cross validation methods (e.g., one-leave out, fivefold, or tenfold validations) are preferred. In a k-fold cross-validation, the data are randomly partitioned into k equal sized parts. One of the k parts is retained as the validation data for testing the algorithm, and the remaining k – 1 parts are used as training data. Training is typically done with a supervised approach which requires ground truth for the task. Ground truth is usually obtainedwith manual delineations of brain lesions or structures by experts for segmentation tasks. Even though this is the gold standard for the learning and evaluation, it is a tedious and laborious task and contains subjectivity. In their work, Mazzara et al. [11] reported intra-expert variabilities of 20 ± 15% and interexperts

variabilities of 28 ± 12% for manual segmentations of brain tumor images. To alleviate this variability, multiple expert segmentations are combined in an optimal way by using label fusion algorithms such as STAPLE [12, 13]. For classification tasks of brain lesions, the ground truth is obtained with biopsy and pathological tests.

To evaluate performance of a newly developed deep learning approach on a task, it is essential to compare its performance against available state of the art methods. In general, most of the algorithms are evaluated on different sets of data and reported different similarity metrics. This makes it hard to compare the performance of different algorithms against each other. Over the last decade, the brain imaging community has become more aware of this and created publicly available datasets with ground truth for evaluating the performance of algorithms against each other in an unbiased way. One of the first such datasets was released in the framework of an MS lesion segmentation challenge, which was held in conjunction with MICCAI 2008. The dataset is maintained as an online challenge dataset (https:// www.nitrc.org/projects/msseg), meaning the training data is released with the ground truth to the public, while the test dataset is released without the ground truth and thus can be evaluated only by the organizers. The latter helps avoid overfitting of the methods and makes comparison more objective. Following the same paradigm, many other datasets have been released since then. Some of the other well-known publicly available datasets for brain MRI are Brain Tumor Segmentation (BRATS), Ischemic Stroke Lesion Segmentation (ISLES), Mild Traumatic Brain Injury Outcome Prediction (mTOP), Multiple Sclerosis Segmentation (MSSEG), Neonatal Brain Segmentation (NeoBrainS12), and MR Brain Image Segmentation (MRBrainS).

IV IMPLEMENTATION DETAILS

Our implementation is based on the Pylearn2 library [20]. Pylearn2 is an open-source machine learning library specializingin deep learning algorithms. It also supports the use of GPUs, which can greatly accelerate the execution of deep learning algorithms. Since CNN's are able to learn useful features from scratch, we applied only minimal pre-processing. We employed the same pre-processing as Tustison et al., the winner of the 2013 BRATS challenge. The pre-processing follows three steps. First, the 1% highest and lowest intensities are removed. Then, we apply an N4ITK bias correction to T1 and T1C modalities. The data is then normalized within each input channel by subtracting the channel's mean and dividing by the channel's standard deviation. As for post-processing, a simple method based on connected components was implemented to remove flat blobs which might appear in the predictions due to bright corners of the brains close to the skull. The hyper-parameters of the di erent architectures (kernel and max pooling size for each layer and the number of layers) can be seen in Figure 3. Hyper-parameters were tuned using grid search and cross-validation on a validation set. The chosen hyper-parameters were the ones for which the model performed best on the validation set. For max pooling, we always use a stride of 1. This is to keep per-pixel accuracy during full image prediction. We observed in practice that max pooling in the global path does not improve accuracy. We also found that adding additional layers to the architectures or increasing the capacity of the model by adding additional feature maps to the convolutional blocks do not provide any meaningful performance improvement.

Biases are initialized to zero except for the softmax layer for which we initialized them to the log of the label frequencies. The kernels are randomly initialized from U (□0:005; 0:005). Training takes about 3 minutes per epoch for the TwoPathCNN model on an NVIDIA Titan black card.

At test time, we run our code on a GPU in order to exploit its computational speed. Moreover, the convolutional nature of the output layer allows us to further accelerate computations at test time. This is done by feeding as input a full image and not individual patches. Therefore, convolutions at all layers can be extended to obtain all label probabilities p(Yi jjX) for the entire image. With this implementation, we are able to produce a segmentation in 25 seconds per brain on the Titan black card with the TwoPathCNN model. This turns out to be 45 times faster than when we extracted a patch at each pixel and processed them individually for the entire brain. Predictions for the MFCascadeCNN model, the LocalCas- cadeCNN model, and InputCascadeCNN model take on average 1.5 minutes, 1.7 minutes and 3 minutes respectively.

V. EXPERIMENTS AND RESULTS

The experiments were carried out on real patient data obtained from the brain tumor segmentation challenge (BRATS), as part of the MICCAI conference. The BRATS dataset is comprised of 3 sub-datasets. The training dataset, which contains 30 patient subjects all with pixel accurate ground truth (20 high grade and 10 low grade tumors); the test dataset which contains 10 (all high grade tumors) and the leader board dataset which contains 25 patient subjects (21 high grade and 4 low grade tumors). There is no ground truth provided for the test and leaderboard datasets. All brains in the dataset have the same orientation. For each brain there exists 4 modalities, namely T1, T1C, T2 and Flair which are coregistered. The training brains come with ground truth for which 5 segmentation labels are provided, namely non-tumor, necrosis, edema, non-enhancing tumor and enhancing tumor.

In total, the model iterates over about 2.2 million examples of tumorous patches (this consists of all the 4 sub-tumor classes) and goes through 3.2 million of the healthy patches. we work with 2D slices due to the fact that the MRI volumes in the dataset do not posses an isotropic resolution and the spacing in the third dimension is not consistent across the data. We explored the use of 3D information (by treating the third dimension as extra input channels or by having an architecture which takes orthogonal slices from each view and makes the prediction on the intersecting center pixel), but that didn't improve performance and made our method very slow. Quantitative evaluation of the models performance on the test set is achieved by uploading the segmentation results to the online BRATS evaluation system . The online system provides the quantitative results as follows: The tumor structures are grouped in 3 different tumor regions. This is mainly due to practical clinical applications. As described by Menze et al. , tumor regions are defined as:

a) The complete tumor region (including all four tumor structures).
b) The core tumor region (including all tumor structures exept "edema").
c) The enhancing tumor region (including the "enhanced tumor" structure).

For each tumor region, Dice (identical to F measure), Sensitivity and Specificity are computed as follows :

$$Dice(P,T) = \frac{|P_1 \Lambda T_1|}{(|P_1| + |T_1|)/2}$$

$$Sensitivity(P,T) = \frac{|P_1 \Lambda T_1|}{|T_1|}$$

$$Sensitivity(P,T) = \frac{|P_0 \Lambda T_0|}{|T_0|}$$

where P represents the model predictions and T represents the ground truth labels. We also note as T1 and T0 the subset of voxels predicted as positives and negatives for the tumor region in question. Similarly for P1 and P0. The online evaluation system also provides a ranking for every method submitted for evaluation. This includes methods from the BRATS challenge published in as well as anonymized unpublished methods for which no reference is available.

CONCLUSION:

Despite the significant impact of deep learning techniques in quantitative brain MRI, it is still challenging to have a generic method that will be robust to all variations in brain MR images from different institutions and MRI scanners. The performance of the deep learning methods depends highly on several key steps such as preprocessing, initialization, and post processing. Moreover, current deep learning architectures are based on supervised learning and require generation of manual ground truth labels, which is tedious work on a large-scale data. Therefore, deep learning models that are highly robust to variations in brain MRI or have unsupervised learning capability with less requirement on ground truth labels are needed. In addition, data augmentation approaches that realistically mimic variations in brain MRI data could alleviate the need of large amount of data. Transfer learning could be used to share well-performing deep learning models, which are trained on normal and pathological brain MRI data, among brain imaging research community and improve the generalization ability of these models across datasets with less effort than learning from scratch.

REFFERENCES

1. Y. LeCun, Y. Bengio, and G. Hinton, BDeep learning,^ Nature, vol. 521, no. 7553, pp. 436–444, 2015.
2. Lin D, Vasilakos AV, Tang Y, Yao Y: Neural networks for computer-aided diagnosis in medicine: A review. Neurocomputing 216:700–708, 2016
3. Kooi T et al.: Large scale deep learning for computer aided detection of mammographic lesions. Med. Image Anal. 35:303–312, 2017
4. Cheng J-Z et al.: Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans. Sci. Rep. 6:24454, 2016
5. Litjens G et al.: Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci. Rep. 6:26286, 2016
6. Menze B, et al. The Multimodal brain tumor image segmentation benchmark (brats). IEEE Trans Med Imaging 2015; 34(10):1993-2024.
7. Gordillo N, Montseny E, Sobrevilla P. State of the art survey on MRI brain tumor segmentation. Magn Reson Imaging 2013; 31(8):1426–38.
8. Bauer S, Wiest R, Nolte L, Reyes M. A survey of MRI-based medical image analysis for brain tumor studies. Phys Med Biol.2013;58:97-129.
9. Liu J, Wang J, Wu F, Liu T, Pan Y. A survey of MRI-based brain tumor segmentation methods. Tsinghua Science and Technology 2014;19(6):578-595.
10. Angelini E D, Clatz O, Mandonnet E,Konukoglu E,Capelle L, Duffau H. Glioma dynamics and computational models: a review of segmentation, registration, and in silico growth algorithms and their clinical applications. Curr. Med. Imaging 2007; 3: 262–76.
11. D. Kwon et al. Combining generative models for multifocal glioma segmentation and registration. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014. Springer, 2014:763–770.
12. Nowak R D, Wavelet-based rician noise removal for magnetic resonance imaging. IEEE Trans Image Processing 1999; 8(10):1408-1419.
13. Zhuang AH, Valentino DJ, Toga AW. Skull stripping magnetic resonance brain images using a model based level set. NeuroImage 2006;32(1):79-92.
14. Y. LeCun et al., BBackpropagation applied to handwritten zip code recognition,^ Neural Comput., vol. 1, no. 4, pp. 541–551, 1989.

15. Deng J, et al.: BImageNet: A large-scale hierarchical image database, in 2009 I.E. Conference on Computer Vision and Pattern Recognition, 2009.
16. O. Russakovsky et al., BImageNet large scale visual recognition challenge, Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015.
17. Krizhevsky A, Sutskever I, Hinton GE: ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ Eds. Advances in neural information processing systems 25. USA: Curran Associates, Inc., 2012, pp. 1097–1105
18. He K, Zhang X, Ren S, Sun J: BDelving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, in 2015 I.E. International Conference on Computer Vision (ICCV), 2015.
19. Goodfellow, I.J., Warde-Farley, D., Lamblin, P., Dumoulin, V., Mirza, M., Pascanu, R., Bergstra, J., Bastien, F., Bengio, Y., 2013a. Pylearn2: a machine learning research library. arXiv preprint arXiv:1308.4214 .
20. Kwon, D., Akbari, H., Da, X., Gaonkar, B., Davatzikos, C., 2014. Multimodal brain tumor image segmentation using glistr, in: in proc of BRATS Challenge - MICCAI.