

# Efficient Brain Tumor Segmentation with Multiscale Two-Pathway-Group Conventional Neural Networks

Muhammad Imran Razzak, Muhammad Imran, Guandong Xu

**Abstract**—Manual segmentation of the brain tumors for cancer diagnosis from MRI images is a difficult, tedious and time-consuming task. The accuracy and the robustness of brain tumor segmentation, therefore, are crucial for the diagnosis, treatment planning, and treatment outcome evaluation. Mostly, the automatic brain tumor segmentation methods use hand designed features. Similarly, traditional methods of Deep learning such as Convolutional Neural Networks require a large amount of annotated data to learn from, which is often difficult to obtain in the medical domain. Here we describe a new model Two-Pathway-Group CNN architecture for brain tumor segmentation, which exploits local features and global contextual features simultaneously. This model enforces equivariance in the Two-Pathway CNN model to reduce instabilities and overfitting parameter sharing. Finally, we embed the cascade architecture into Two-Pathway-Group CNN in which the output of a basic CNN is treated as an additional source and concatenated at the last layer. Validation of the model on BRATS2013 and BRATS2015 datasets revealed that embedding of a group CNN into a two pathway architecture improved the overall performance over the currently published state-of-the-art while computational complexity remains attractive.

**Keywords**—Brain tumor, Group CNN, CNN, Deep neural network, group convolutional neural networks, cascade CNN, Two-Pathway CNN, 2PG-CNN

## I. INTRODUCTION

CANCEROUS brain tumors present themselves as unnatural, uncontrolled growth and division of cells in the brain. While brain tumors are not very common, they are one of the most lethal cancers. In the US alone for example, approximately 23,000 new cases of brain cancer were diagnosed in 2015. It is an abnormality in the brain tissues that damage the nervous system severely, which result patient death. are the most common brain tumors that are infiltrative in nature, and occur near white matter fibers. They may spread to any part of the brain making it difficult to detect. High-grade gliomas are considered one of the most aggressive tumors with a median survival of 15 months. Gliomas can be measured by MRI using multiple sequences such as T2-weighted uid attenuated inversion recovery (Flair), T1-weighted (T1), T1-weighted contrast-enhanced (T1c), and T2-weighted (T2) using existing automated brain tumor segmentation techniques.

Healthy brains consist of three types of tissues: gray matter, white matter and cerebrospinal fluid. Detection and segmentation of cancerous cells using MRI not only helps to detect the presence of tumors and their location, but it also enables the identification of tumor size, necrotic tissue, tumorous tissue (vascularized or not) and edema (swelling near the tumor). Brain tumors are heterogeneous in shape and appearance (gliomas may have the same appearance as gliosis and stroke etc), which makes segmentation challenging for radiologists. Furthermore, they may appear at any location in the brain: depending on the origin of the brain tumor, they can be classified as either primary tumors or metastatic brain tumors. The edges of brain tumors are often ambiguous and fuzzy, and are hard to distinguish from healthy tissues. Therefore a more sensitive alternative to MRI is needed to improve the detection of tumors and to increase the survival rate of people with brain tumors.

Machine-aided image segmentation, and its subsequent quantitative assessment of cancerous tissue, provides valuable information for the early diagnosis and characterization of neuropathologies, which then informs appropriate treatment strategies. Quantitative analysis of affected cells reveals clues about the disease progression, its characteristics, and effects on the particular anatomical structure. Furthermore, it very important for early diagnosis that helps early prevention by planning the treatment strategies. Cancerous cells are normally quantified by means of the number of lesions, their volume, and biomarkers that have been shown to be related to cognitive deficits. As a result, the quantitative analysis of effected regions requires accurate lesion segmentation, which is a challenging task because of the variations in size, shape, location and frequency of cancerous lesions. Arguably, the most accurate brain tumor segmentation results are achieved manually by an expert; however, this is an expensive, time-consuming, tedious, and impractical task, that is prone to errors and is affected by inter-observer variability. Hence, clinicians often use qualitative or visual inspection only, or at best, crude measures like approximating the volume and numbers of tumors.

Manual segmentation of brain tumors from large MRI images is a difficult and time-consuming task. Existing brain tumor segmentation approaches can be broadly categorized into either generative models or discriminative models [27]. Generative models require prior information and segmentation of brain tumors, whereas discriminative models depend on a set of features and classifiers. The most commonly adopted classifiers are support vector machines (SVMs), random forests, neural networks and genetic algorithms. In contrast, automatic brain tumor segmentation methods use hand-designed features

M. I.Razzak : (email: imran.razzak@ieee.org) is with the Advanced Analytics Institute, University of Technology, Sydney, Australia.

Muhammad Imran (email: dr.m.imran@ieee.org) is with King Saud University, Riyadh, Saudi Arabia.

Guandong Xu : (email: guandong.xu@uts.edu.au) is associate professor at Advanced Analytics Institute, University of Technology, Sydney, Australia.

Manuscript received April 19, 2005; revised January 11, 2007.

and a variety of image features (e.g. shape, area, perimeter, circularity etc.), intensity (e.g. mean, variance, standard deviations) and texture (e.g. contrast, entropy, correlation etc.).

Recently, deep learning, and in particular the Convolutional Neural Network (CNN) has become the methodology of choice for medical image analysis, following its tremendous success in routine computer vision applications [31], [30]. With regards to tumor detection, candidate generation and false positive reduction, using deep learning based methods, unambiguously outperformed traditional machine learning approaches [16], [41], [28], [45]. This achievement was acknowledged in 2015 where the DeepMedic software for brain lesion segmentation, which was based on a 3D-CNN coupled with a 3D fully connected CRF, won the ISLES 2015 challenge [20]. Additional deep learning based brain tumor segmentation methods were presented in the 2013, 2015 and 2017 challenges. Different deep learning models were adopted, including FCNN [33], [7], [24], [16], 3DCNN [6], [14], FCNN with CRF [44], 3D U-Net [2], [3] and Autoencoders [1], [37].

Results of the 2015, 2016 and 2017 ISLES brain tumor challenge showed that, among the deep learning based tumor segmentation methods, CNN based approaches achieved better performance as compared to other methods. CNNs, however, require a substantial amount of labeled data for training, which is typically scarce within the medical imaging community as a result of patient confidentiality and the time needed to obtain high-quality annotations. Furthermore, CNNs do not exploit symmetries which result in the creation of multiple rotated and reflected copies of filters. The problem is further compounded by the fact that many CAD systems will need to be developed to accommodate different imaging modalities, scanner types, settings, resolutions, and patient populations. Therefore, data efficiency is a major hurdle for the scalable development of CAD systems for tumor detection.

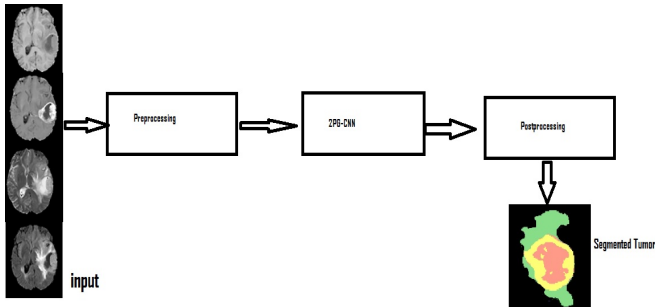


Fig. 1. Block diagram of the proposed Brain tumor segmentation

CNN is somehow data efficient as compared to fully connected networks due to its translation weight sharing properties in the convolutional layers. Layers in CNN are translation equivariant i.e. when the network input is shifted, internal representations are also shifted, thus transnational weight sharing is effective in each layer. Brain tumors maintain their identity not just under translation but also rotation and reflection. Thus, embedding these properties by using group convolutions, weight sharing, and equivariance into CNN for brain tumor detection is a viable option. Furthermore, contextual

information has played a very useful role in computer vision and image segmentation tasks. Currently, the use of conditional random fields is a widely used approach to model context information; however, it relies on fixed topologies and offers limited flexibility. To incorporate the spatial correspondence of labels, the output of one network could be concatenated with another network.

We developed a Two-Pathway-Group Convolutional Neural Network (2PG-CNN) to address current hurdles in brain tumor segmentation by embedding a group convolutional neural network, that exploits transnational, rotational and reflection properties in itself, and that exploits local and global features through a Two-Pathway CNN. The **key contributions** of this paper are:

- Development of a Two-Pathway-Group CNN that substantially increases the expressive capacity of a network through rotational and transnational invariance properties and which embeds local and global features by utilizing two-pathway architecture without increasing the number of parameters.
- A novel deep learning approach that utilizes symmetries together with local and global features for brain tumor segmentation that takes approximately 3 minutes and 18 seconds which is almost faster, or comparable with, most current state-of-the-art systems.
- A fully automatic brain segmentation method with state-of-the-art results on BraTS-2013 and BraTS-2015 datasets.
- An equivariance in the Two-Pathway CNN model to reduce instabilities and overfitted parameter sharing.
- An extension of the Two-Pathway-Group CNN to include novel cascaded architecture by performing concatenation prior to its output layer.

We produced new insights into the higher degree of weight sharing in the Two-Pathway-Group CNN compared with the Two-Pathway CNN, and have increased the capacity of CNN by modeling symmetry into it without increasing the parameter. We observed that the network is able to learn the anatomical features of a brain in order to identify and segment brain tumors. A block diagram of our new 2PG-CNN approach to brain tumor segmentation is given in Fig. 1. In what follows, we provide the related work in section 2, followed by the architecture of the proposed approach. In section 4, we provided the comprehensive evaluation.

## II. RELATED WORK

Accurate brain tumor segmentation is critical for early diagnosis and patient survival. Manual segmentation (human-based) is typically undertaken for the diagnosis of brain tumors by radiologists, and it is a difficult, tedious and time-consuming task. Automated (machine-based) brain tumor segmentation offers an objective, reproducible, and scalable alternative to manual segmentation, and is a popular research focus in the medical imaging community. Brain tumor segmentation techniques can be roughly classified into two categories: generative and discriminative model-based approaches. Generative model-based approaches rely on differentiating healthy

from cancerous tissues based on appearance and therefore rely on knowledge of the anatomical structure of the organ that could be computed by aligning the image with several images of healthy tissues. Contour-based generative methods for tumor segmentation rely on alignment-based features or left-right brain symmetry features. The main advantage of these methods is that they do not rely on labeled data. In contrast, discriminative approaches exploit prior knowledge of the brains anatomy. Segmentation is based on low-level image features such as raw pixels and Gabor filterbanks. Typical discriminative methods include conventional machine learning approaches such as random forest and support vector machines; however, these methods rely on handcrafted features consisting of high discriminative power.

Deep learning based approaches for medical object detection and classification are emerging as viable alternatives for applications in medical imaging. In brain tumors, these approaches out perform classical machine learning techniques in terms of false positive reduction. They are able to learn discriminative features automatically by outperforming hand-crafted feature sets [32]. Among the deep learning based approaches for brain tumor segmentation, methods based on CNNs have provided performance as compared to others. In other words, CNN has become the methodology of choice and have been extensively applied on variety of medical image task and showed promising results [5], [8], [9], [12], [18], [36], [39], [45]. Particularly, both 3D-CNNs and 2D-CNNs are being used to develop a brain tumor segmentation system. However, deep learning based methods are data hungry i.e. more the data the better the performance. Whereas it is not easy to obtain the labeled data in medical fields due to its sensitivity and cost of labeling the data as annotation by single expert is not enough due to human error and requires to have consensus annotations by multiple expert observers.

Both 2D-CNNs and 3D-CNNs models are being adopted to develop brain tumor segmentation methods [3], [8], [26]. Although, 3D-CNNs takes full advantage of 3D information of the MRI volume-metric data, however, size of networks is much bigger than 2D-CNN which results in high computational complexity. Thus, 2D-CNNs has been widely preferred for brain tumor segmentation. To overcome the computational challenge for 3D processing and to incorporate 3D contextual information, several researchers focus on using the 2D CNNs on three orthogonal 2D patches. Kamnitsas et al. presented an effective dense training approach by processing the adjacent patches into one pass [20]. They have applied dual pathway architecture to incorporate both local and larger contextual information. At the final stage, a 3D fully connected conditional random field was used for soft segmentation, which effectively removed false positives. Similarly, Havaei et. al. simultaneously exploited local contextual features, together with global contextual features by using two different CNN paths, concatenated at the last layer [16]. To overcome the issue of data imbalance, they used a 2-phase training procedure: patches from the true distribution followed by an imbalanced dataset. The network was extended by adding another network, and concatenated it at the last layer. In other work, Kamnitsas et. al. presented an efficient and effective dense training

scheme to join the processing of adjacent image patches into one pass through the network [20]. The dual pathway was applied in order to model the local and global features.

In most of the pattern recognition tasks, image maintains its identity under translation, rotation as well as reflection. Generalizing CNNs to deal with such kind of data is shown great performance by using group convolutions and weight sharing. Recently, modeling symmetry into CNN is actively being studied [4], [10], [13], [22], [29], [42]. Recently, Cohen and Welling presented Group equivariant CNN that reduces sample complexity by exploiting symmetries such as translation, rotation, reflection [11]. Group equivariant CNNs use G-convolutions than regular convolution layers for weight sharing which helps to increase the expressive capacity of the network without increasing the number of parameters. Bekkers et.al. presented a framework for rotation and translation covariant CNN using group convolutions that are invariant to shape and appearance variability for medical image analysis [4]. Results on histopathology, retinal imaging, and electron microscopy images showed achieved better performance without the need for data augmentation by rotation. Winkels and Cohen presented 3D roto-translation group CNN based approach for pulmonary nodule detection instead of using more conventional translational convolutions [41]. Results showed that CNN by leveraging the symmetrical behavior reduce not only reduce the false positive but also improve the performance of nodules detection. In another work, Veeling et.al. applied rotation equivariant CNNs for pathological image analysis [38]. The method shows strong generalization under limited dataset size and are more robust under adversarial perturbations in the rotation, translation, and local geometric distortions.

### III. TWO-PATH-GROUP CNN ARCHITECTURE

In this section, we presented Two-Path-Group convolutional neural network (2PG-CNN) in detail focusing on the algorithm rather than the underlying mathematical theory and using visual aids where this is helpful. We start the brain tumor segmentation with a novel representation learning on features scheme by embedding the symmetry into two pathway CNN. Since glioblastomas are infiltrative tumors, their borders are often fuzzy in appearance that make them quite hard to distinguish them from healthy tissues thus additional attention is needed to segment the border pixels. Recently, multi-scale features representation has been successfully applied for brain tumor segmentation that fully capture both local and global contextual information [25], [16], [17]. Thus, to overcome this fuzzy appearance of tumor segmentation, we have also adopted two pathway CNN to model the local and global contextual information in CNN.

As discussed earlier, the performance of CNNs is significantly influenced by the feature matrix. Furthermore, it does not accommodate local features as well as global contextual features simultaneously. To overcome these challenges, we present 2PG-CNN that simultaneously accommodate the global and local features as well as embed additional transformations like rotations and reflections in itself by applying not only translation but also rotational and reflection to the

filters which result in an increase in the degree of weight sharing. More specifically, starting with a canonical filter with learnable parameters, one produces a number of transformed copies, which are then convolved (translationally) with the input feature maps to produce a set of output feature maps. Enforcing equivariance in the Two-Pathway CNNs model reduces instabilities and overfitting parameter sharing.

Since the BraTS image dataset has lack of resolution in third dimension, thus we considered the two dimensional approach by performing the tumor segmentation slice by slice (2D axial image) where each pixel in each slice is associated with different image modalities (T1, T2, T1C and FLAIR). Thus, input of two dimension 2PG-CNN is the  $x \times x$  patch that extract the group equivariant features from its preceding layers into the hierarchy to which it is concerned. In the following subsections, we first describe the group CNN and modelling of rotational, reflection and transnational invariant properties in to Two-Pathway CNN.

#### A. Group Equivariant CNN

Practically, the convolutional neural network is transnational equivariant and can be expressed as  $[[L_t f] * \Phi](x) = [L_t[f * \Phi]](x)$  however, it is not rotational equivariant i.e.  $[[L_r f] * \Phi](x) = L_r[f * [L_{r^{-1}} \Phi]](x)$ ; where  $L$  denotes the operator that translate the features map  $f : \mathbb{R}^K$  by  $t \in \mathbb{Z}^2$ . As a result, conventional CNN requires a substantial amount of labeled data for training, however, it is quite tedious, expensive and prone to error to label complex data especially in medical. The problem is further compounded by the fact that in all likelihood, many CAD systems will have to be developed for different imaging modalities, scanner types, settings, resolutions, and patient populations. All of this suggests that data efficiency is a major hurdle to the scalable development of CAD systems.

For a traditional CNN filter (translational), the filter is translated across the features, as a result, each cell of feature map associated with translation. However, symmetry is preserved by each layer that could be used rather than shifting the image and feeding it. To overcome this issue, we have used a group convolutional neural network [11] for brain tumor segmentation. Mathematically, group convolution for a group  $G$  and input space  $X$  is defined as  $[f * \Phi](g) = \sum_{h \in X} \sum_K f_k(h) \Phi(g^{-1}h)$ ; where  $f_k$  and  $\Phi_k$  are the signals defined on  $X$ ,  $k$  is the input channel and  $g$  is the transformation in  $G$  [11]. Now, we can say that the  $G$ -convolution is equivariant to  $G$  transformations i.e.  $[[L_u f] * \Phi](g) = [L_u[f * \Phi]](g)$ . In conclusion, group equivariant CNN is based on the symmetry of an object is a transformation that leaves the object invariant. It embed additional transformations like rotations and reflections in itself by applying not only translation but also rotational and reflection to the filters which result in an increase in the degree of weight sharing. More specifically, starting with a canonical filter with learn-able parameters, one produces a number of transformed copies, which are then convolved (translationally) with the input feature maps to produce a set of output feature maps. To embed these symmetries, we used the recent CNN that is rotation equivariant [11], [43], [41], [22].

Generally, the symmetry group of an object is the set of transformations that map the object back onto itself without changing its structure (leaves the object invariant) i.e. rectangle or square can be rotated any degrees without changing its appearance. In other words, rotation of an object does not effect its shape. For example, flipping of an image  $Z$ , we get  $-Z$  such that  $-Z = -x, -y \in Z = Z$ . We have used p4m that consist of all compositions of translations mirror reflections, and rotations (0, 90, 180, 270) about any center of rotation in the grid each learnable filter produces a number of orientation channels, each of which detects the same feature in a different orientation. The group p4m consists of all compositions of translations, mirror reflections, and rotations by 90 degrees about any center of rotation in the grid. Like p4, we can parameterize this group by integers:

$$g(m, r, u, v) = \begin{bmatrix} (1)^m \cos(\frac{r\pi}{2}) & -(-1)^m \sin(\frac{r\pi}{2}) & u \\ \sin(r\pi/2) & \cos(r\pi/2) & v \\ 0 & 0 & 1 \end{bmatrix}$$

where  $m \in [0, 1]$ ,  $0 \leq r < 4$  and  $(u, v) \in \mathbb{R}^2$ .

To embed the symmetric properties into Two-Pathway CNN with a feature map, the filter is translated across the feature map, and a dot product is computed at each position, as a result, each cell of output features map consist of translation, rotation and reflection properties in itself. Output filter has several obvious properties, such as closure, associativity for transformations, identity map, and inverses. Further detail of GCNN can be found at [11], [38], [42].

#### B. Two-Pathway-Group CNN

Two-Pathway CNNs exploit both local features of an image, as well as the larger context simultaneously. The proposed patch classification model used here is shown in Fig. 2 for p4 (the p4m variant is a trivial extension). The Two-Pathway-Group CNN architecture was based on the pathway convolutional network, which consisted of dense blocks with layers that use the stack of all previous layers as input, alternated with transition blocks consisting of an 1x1 convolutional layer and 2x2-strided average pooling [16]. Two-Pathway-Group convolutional neural networks architecture consist of two parallel streams of CNN. Stream-I, convolutional neural networks with smaller receptive fields (7x7 or 5x5) and stream-II that consist of larger receptive fields (13x13, 15x15 or 17x17). In the following throughout discussion, we refer to two convolutional neural networks streams as the local CNN pathway and the global CNN pathway, respectively. Two-Pathway-Group convolutional neural networks is able to predict the label of pixel by using visual detail of the pixel through local CNN pathway and larger context of the pixel through global CNN pathway. The motivation of using two-pathway CNN is to utilized the local as well as context information of the image to reach precise segmentation. Fig. 2 and Fig. 3, shows the two pathway architecture that embed symmetrical properties in itself. As in Two-pathway group CNN, we have two different streams of CNNs (local pathway CNN and global pathway CNN), we have performed concatenation of both CNN at

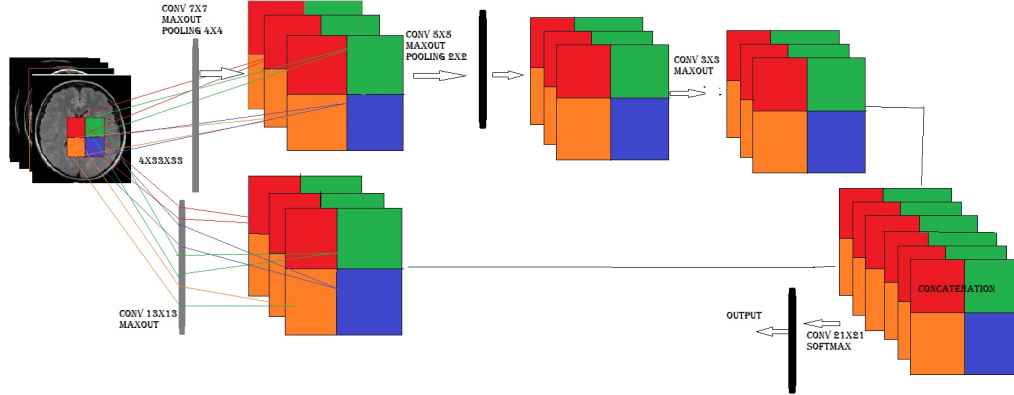


Fig. 2. Two-Pathway-Group CNN architecture (2PG-CNN) showing that the input patch is processed by two different group CNNs. The four blocks in a feature map of both CNNs show a p4 group features map that inherits group CNN properties. The upper CNN represents a local feature map and the lower CNN shows a global feature map.

second-last layer. The concatenated features set is finely fed to the output layer.

The next step is modeling the Group Equivariant symmetry in the two pathway CNN i.e. the feature maps of both CNNs are thought of as functions on p4m group i.e. p4m group means that feature channels come in groups of 8 ( 4 for rotations and 4 for reflections). We have modeled the symmetrical features in both CNNs individually. In the first layer, these were produced using the  $(Z^2 \rightarrow G)$ -convolution. Group-pooling layer is used to ensure that the output is either invariant.

$$[f * \varphi](g) = \sum_{y \in Z^2} \sum_{k=1}^K f_k(y) \varphi_k(g^{-1}y)$$

where  $g = (r, t)$  is the roto-reflection-translation.

In next layer, both feature maps of earlier layer and filters are functions on G and combined using convolution.

$$[f * \varphi](g) = \sum_{h \in G} \sum_{k=1}^K f_k(h) \varphi_k(g^{-1}h)$$

Two-Pathway Group CNN show strong generalization under limited dataset size and are more robust under adversarial perturbations in rotation, translation and local geometric distortion. Thus, each learnable filter produces a number of orientation channels, each of which detects the same feature in a different orientation. We will refer to the set of orientation channels associated with one feature / filter as one feature map.

1) *Cascaded 2PG-CNN* : One main disadvantage of the convolutional neural networks described is that conventional CNNs predict segmentation label separately from each other that are of the same semantic class. As CNN predict the segmentation labels separately from each other and CRF is being used to overcome this issue by performing the mean-field message passing inference to segment completely. However, it results in biasness as a label at a given position is effectively

influenced by the model's beliefs about what the label is in the vicinity of that position. In literature, joint models for segmentation are being used however these methods are typically more computationally expensive than a simple feedforward pass through a CNN. To overcome this issue, we simply concatenate the output layer of the first CNN with any of the layers in the second CNN. Moreover, we use the same two-pathway structure for both CNNs. This effectively corresponds to a cascade of two CNNs.

In this work, we investigated the cascaded architectures that concatenate the first CNNs output at end of other CNNs i.e. concatenation is performed right before the output layer as shown in Fig. 3. In detail, the first CNNs output is actually the first iteration of mean-field whereas the output of second CNN is the second iteration.

#### IV. EXPERIMENTAL SETTING

The brain tumor segmentation is performed on the multimodal MRI volumes image dataset where each voxel is associated with several image modalities (such as T1, T1c, T2 and FLAIR). Segmentation of brain tumor is a data imbalanced problem due huge variations in different types of tumors such as healthy voxels comprise 98% of total voxels. 2% of them are pathological voxels whereas only 0.18%, 1.1%, 0.12% and 0.38% belongs to necrosis, edema, non-enhanced and enhanced tumor respectively. In medical imaging, even small image features may be very important from clinical perspective as some of the pathologies are rare but life threatening, thus approach should be sensitive to imbalance class. CNNs performance is significantly effected by the strategy of training sample being used. As dataset is biased, thus, trained model could be biased to the classes with more training samples which result in over segmentation. In conclusion, using path from the true distribution for initial training would result in imbalance training of the model as in this case model will be overwhelmed by healthy patches.

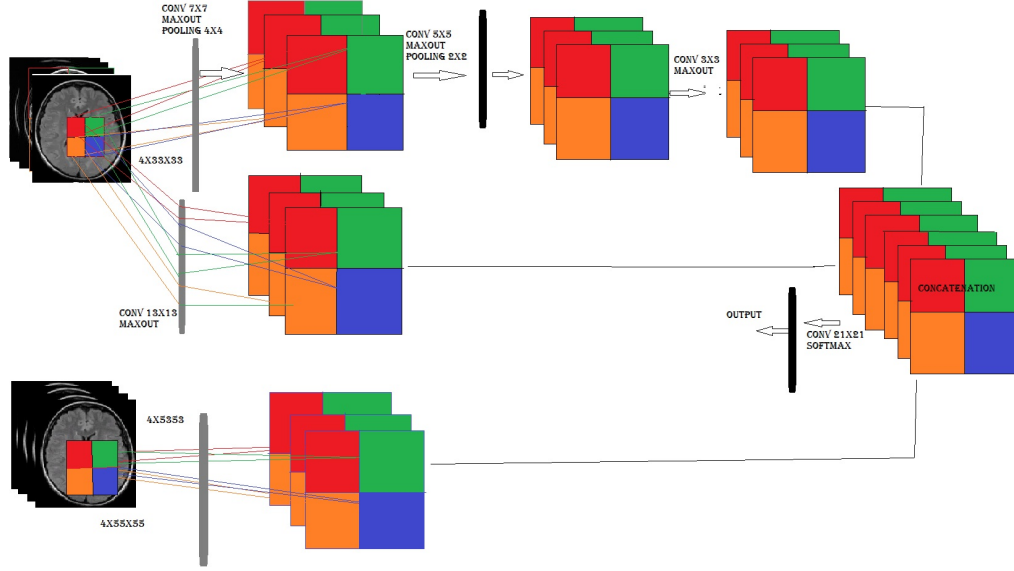


Fig. 3. Cascaded Two-Pathway Group CNN architecture based on pre-output concatenation.

A commonly approach adopted to overcome the imbalance issue of data is the training on image patches that are equally sampled from each tumor type. This, however, add the biasness in the classifier towards class having more data and it may result in over-segmentation. To overcome this issue, initially, we consider the patches. In conclusion, using path from the true distribution for initial training would result in imbalance training of the model as in this case model will be overwhelmed by healthy patches. To overcome this issue, initially, we consider the patches that are equal probable. Once the initial training on done on balance dataset, we moved to unbalanced nature of the data and train the output layer only. First training phase results the the networks having the balanced properties whereas the second result the output probabilities to be calibrated correctly. As we dataset is small, 2PG-CNN showed strong generalization as it is robust under adversarial perturbations in rotation and translation as well as geometric distortions.

The data in the BraTS have lack resolution in the third dimension thus, in this work, we have considered segmenting the tumor form image slice by slice from the axial view. We processes each 2D axial slice sequentially where each pixel in slice is associated with different image modalities namely such as T1, T2, T1C and FLAIR. The input is NN 2D patch with several modalities that predicts the class of a pixel. Final layer consists of a group-pooling layer followed by a sigmoid activation, resulting out is probability of tumor output.

#### A. Dataset

We evaluated our network using BraTS 2013 and 2015 training sets, leaderboard datasets and challenge datasets [40], [15]. The complex characteristics of the medical imaging

datasets imposed several challenges as the image acquisition was quite expensive and sensitive as compared to real life data. Thus, medical images were collected under controlled conditions to enable more predictable data distributions. In many modalities, images were calibrated before capturing, such that the spatial relationships and image intensities mapped directly to physical quantities and were inherently normalized across subjects.

The training sets contained images from 285 individuals (220 HGG and 54 LGG) captured under controlled environments, whereas the testing datasets consisted of 110 cases. The BraTS 2016 dataset also shared the same training dataset with BraTS 2015. The BraTS 2015 validation and testing set contained images from 46 and 146 patients with brain tumors of unknown grade, respectively. For each brain there existed four modalities (T1, T1C, T2 and FLAIR) for the same patient, which coregistered and aligned onto the same anatomical template space. The training brains came with manual segmented, ground-truth results given by experts for which five segmentation labels were provided: non-tumor, necrosis, edema, non-enhancing tumor and enhancing tumor. All brain images were skull-stripped to an isotropic  $1mm^3$  resolution.

#### B. Data Preprocessing

Since CNNs are able to learn useful features from the beginning, data should be clean and standard. However, intensity values being non standardized in MRI data as data may be from different MRI scanners with varying protocols that causes false positives in predicted outputs i.e. images obtained from different modalities may be affected by artefacts that causes false intensity levels. Data normalization is therefore required



especially in the case of deep learning where input data is considered as color channels i.e. raw data is directly input to the neural network. Thus, we employed preprocessing, by normalizing the value ranges, to assure that data integrity matched not only between patients but between the modalities in order to avoid initial biases of the network.

We applied the same preprocessing as described by Menze et.al. (winner of the 2013 BRATS challenge [27]) and Gibson et al. [15]. We follow the simple work flow and removed the 1% highest and lowest intensities followed by N4ITK bias correction on T1 and T1C modalities [35]. We have applied non-uniform intensity normalization algorithm [35], which is widely used for artefact removal and intensity normalization.

Finally, each modality is normalized by subtracting the mean and dividing by the standard deviation of the brain region and clipping to [5, 5] and scaling to [0, 1] to remove outliers and segmenting the interested region respectively.

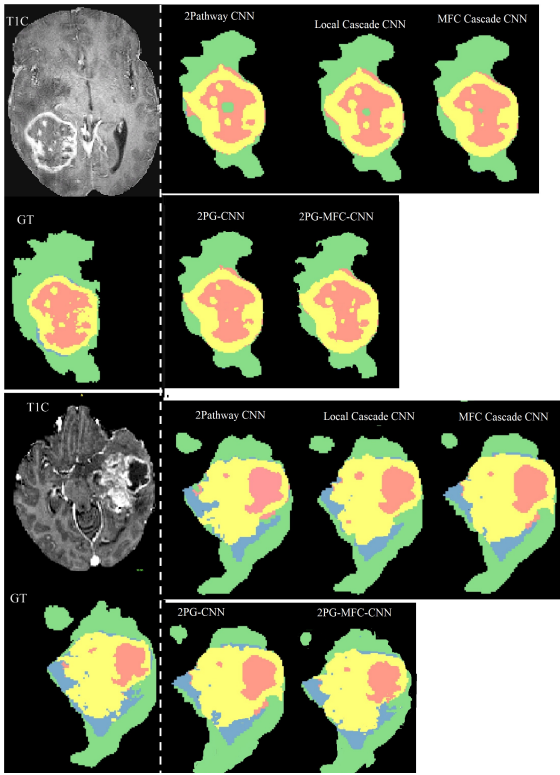


Fig. 4. Brain Tumor segmentation visual comparison Two-Pathway CNN ([16]) vs Group Two pathway CNN (proposed) on BraTS 2013 dataset (color code: green:edema, blue: non enhanced tumor, pink:necrosis)

### C. Parameter Setting

The proposed tumor segmentation model is shown in Fig. 2 and Fig. 3. We have provided visual demonstration of P4 in Fig. 2 as p4m variant is a trivial extension. Our implementation is based on the PyTorch library. It is an open source machine learning library for Python, based on Torch and supports the use of GPUs, which can greatly accelerate the execution of deep learning algorithms. The architecture is based on the

traditional convolutional neural network. Models are optimized using Adam [21]. We have replaced the traditional layer of CNN with group equivariance convolution layers [11]. Training parameter of both Two-Pathway-Group CNN and Cascade Two-Pathway-Group CNN can be seen in Fig. 2 and 3. For max pooling, we used a stride of 1 to keep per-pixel accuracy during full image prediction. The final layer consists of a group-pooling layer followed by a sigmoid activation and group-pooling layer followed by a concatenation layer respectively. The final layer consist of  $5 \times 1 \times 1$ .

Stochastic gradient descent (SGD) in update the parameter for each training example  $x_i$  and label  $y_i$ . To maximize the label probability in training dataset, we used SGD by selecting labels  $Y_{ij}$  repeatedly at random subset of patches within each brain and computing the average negative log probability  $-\log p(Y/X) = \sum_{ij} -\log p(Y_{ij}|X)$  for each label image. Batch gradient descent performs redundant computations for large datasets, however, SGD performed one update at a time and frequent updates with a high variance that cause the objective function to fluctuate heavily. We performed update based on small subset patches to avoid the process for whole brain image. As training the network with simple gradient descent can take quite a long time, however, momentum can accelerates the gradient descent learning. We have used the momentum strategy as described by Krizhevsky et.al. in [23].

By interpreting the output of 2PG-CNN as a model for the distribution of over segmentation labels, a natural training criterion was to minimize the negative log-probability, equivalently, or maximize the probability of all labels in the training dataset. To do this, we followed a stochastic gradient descent approach by repeatedly selecting labels at a random subset of patches. We set the learning rate to 0.005 with the decay 0.1. We gradually increase the momentum coefficient form 0.5 to 0.9 during training. The learning rate is decreased by a factor at every epoch. To learn the useful features, we used dropout [34] by multiplying each hidden layer by 0 with certain probability.

The optimization improves the performance and it results in tumor probability output on the plane Z. Training took 3 minute and 14 second per epoch.

### D. Training

We developed two networks: 2PG-CNN and Cascade 2PG-CNN. In 2PG-CNN, we trained two group CNNs with different receptive field sizes to model local and global properties. The architecture was made of two streams: a local-receptive field size of  $7 \times 7$  to model the local properties, and global receptive field of size  $13 \times 13$  to model the global properties. We have selected this size based on earlier experiments [16]. Finally two networks are concatenated followed by the soft max. To train cascaded 2PG-CNN, In 2PG-CNN, we have used the same network (2PG-CNN) and concatenation is performed at last layer. In this network, there are three networks, Cascade with receptive field size  $53 \times 53$ , local  $7 \times 7$  and global  $13 \times 13$ . Unlike CNN and 2PG-CNN, cascaded 2PG-CNN model the dependencies between adjacent labels in segmentation.

As the data was imbalanced, we considered patches to overcome this issue. A path from the true distribution for

TABLE I. PERFORMANCE EVALUATION OF 2PG-CNN AND CASCADE 2PG-CNN

Method	Dice			Precision			Sensitivity		
	Complete	Core	Enhance	Complete	Core	Enhance	Complete	Core	Enhance
TwoPathCNN [16] [BRATS'13]	88	79	73	89	79	<b>68</b>	87	<b>79</b>	<b>80</b>
DeepMedic [19][BRATS'15]	83.6	67	62.0	82.3	84.6	64	88.5	61.6	65.6
DeepMedic+CRF [BRATS'15] [20]	84.7	67	62.9	85.0	84.8	63.4	87.6	60.7	66.2
Ensemble [20] [BRATS'15]	84.5	66.7	63.3	83.3	86.1	63.2	<b>88.9</b>	59.9	67.3
Ensemble+CRF [20] [BRATS'15]	84.9	66.7	63.4	85.3	86.1	63.4	87.7	60	67.4
FCNN + 3D CRF [45] [BRATS'15]	84	73	62	89	76	63	82	76	67
<b>2PG-CNN</b> [BRATS'13]	88.3	<b>81.2</b>	73.1	86.43	87.1	67.6	87.4	64.2	66.45
<b>Cascade 2PGCNN</b> [BRATS'13]	<b>88.9</b>	81.1	<b>73.7</b>	<b>86.91</b>	<b>87.2</b>	73.9	87.71	66.21	69.42
<b>2PG-CNN</b> [BRATS'15]	86.65	76.4	74.7	87.2	87.8	64.6	87.8	79.2	79.25
<b>Cascade 2PGCNN</b> [BRATS'15]	<b>89.2</b>	<b>79.18</b>	<b>75.1</b>	<b>88.22</b>	<b>87.9</b>	68.41	88.32	<b>80.66</b>	<b>80.21</b>

initial training resulted in the imbalanced training of the model because of the presence of an overwhelming number of healthy patches. To overcome this issue, we considered the patches that were equally probable. Once the initial training was done on the balanced dataset, we moved to an unbalanced dataset and trained the output layer only. In the first training phase the networks had balanced properties and in the second training phase, the output probabilities calibrated correctly. As our dataset was small, 2PG-CNN showed a strong generalization. Furthermore, it was robust under adversarial perturbations in rotation and translation, as well as in geometric distortions.

#### E. Evaluation Parameters

For the quantitative evaluation of the segmentation performance on the test set, we have used the three metrics namely: Sensitivity, Specificity, specificity, and the dice similarity coefficient because As the tumor structure is was grouped into three different tumor regions: such as (complete, core and enhancing tumor). Thus, we have computed the all three metrics for all regions as shown in Fig.4. For each tumor region, the dice score (identical to F-measure) was computed by overlapping the predicted output image with the ground truth image. Evaluation metrics, dice similarity, sensitivity and specificity were computed as

$$Dice(P, G) = \frac{|P_1 \cap G_1|}{(|P_1| + |G_1|)/2}$$

$$Sensitivity(P, G) = \frac{|P_1 \cap G_1|}{|G_1|}$$

$$Specificity(P, G) = \frac{|P_o \cap G_o|}{|G_o|}$$

Where  $P$  is the machine based segmentation region and  $G$  is the manually labeled ground truth. Here  $P_1$  and  $P_o$  are the positive and negative predicted tumor regions respectively. Similarly, for  $G_1$  and  $G_o$ .

#### F. Results

The experiments were carried out on real patient data obtained from the 2015 brain tumor segmentation challenge (BraTS2015), as part of the MICCAI conference [15], [40]. As mentioned earlier, Two-Pathway Group CNN has properties of modeling local and global features as well as symmetry in itself. The architecture discussed in earlier sections has two path

way, one group CNN is used to model the local whereas the other global CNN model the context. The results reported in table-I shows that joint embedding of of group CNN into Two-Pathway CNN benefits the segmentation performance. The results indicate that the proposed 2PG-CNN and Cascade 2PG-CNN performed consistently better than competing methods. Compared to [16], [20] results, we see that the superiority of both 2PG-CNN and cascade 2PG-CNN architecture. It is due to the increased parameter sharing by the p4m-equivariance as well as local and global contextual features. Furthermore, cascade CNN embedding with Group CNN helps to reduces the redundancy of detecting the same tumor in in different orientations. Table 1 presents the quantitative results of the proposed approach. Results showed that Cascade 2PG-CNN provided slightly better results compared to 2PG-CNN. Fig. 4 showed the segmentation result comparison of the Two-Pathway- Group CNN and the cascade Cascade Two-Pathway-Group CNN. We have noticed that the architecture shown in Fig. 3 provided better segmentation with smooth boundaries that might be due to the concatenation at layer before the softmax output layer that resulted in learning of a center pixel similar to its surrounding.

#### V. CONCLUSION

In summary, this work presented a fully automatic brain tumor segmentation approach (Two-Pathway-Group CNN) that embedded symmetrical properties, as well as local and global contextual features, in itself. In contrast to previous multiscale features, 2PG-CNN benefited from symmetrical embedding (i.e. additional transformations like rotations and reflections in itself) by applying not only translation but also rotational and reflection to the filters, which resulted in an increase in the degree of weight sharing. We extended 2PG-CNN to predict the segmentation label using Cascade 2PG-CNN. Evaluation results showed that G -CNN outperformed the CNN. We concluded that modeling symmetrical equivariance into Two-Pathway CNN improved the model reliability with similar numbers of parameters. Furthermore, experiments showed that 2PG-CNN and Cascade 2PG-CNN were more data efficient than conventional CNNs.

#### ACKNOWLEDGEMENT

This work is partially supported by Australian Research Council Linkage Project under LP170100891.



## REFERENCES

- [1] V Alex and G Krishnamurthi. Brain tumor segmentation from multi modal mr images using stacked denoising autoencoders. *proc. of BRATS-MICCAI*, 2016.
- [2] Chagas V. S. Escudero G. G Oliveira D. D. C. Pereira S. M Santos H. M. Scussel A. A Amorim, P. H. A. 3d u-nets for brain tumor segmentation in miccai 2017 brats challenge. In *2017 International MICCAI BraTS Challenge*, pages 9–14. Springer, 2017.
- [3] Andrew Beers, Ken Chang, James Brown, Emmett Sartor, CP Mammen, Elizabeth Gerstner, Bruce Rosen, and Jayashree Kalpathy-Cramer. Sequential 3d u-nets for biologically-informed brain tumor segmentation. *arXiv preprint arXiv:1709.02967*, 2017.
- [4] Erik J Bekkers, Maxime W Lafarge, Mitko Veta, Koen AJ Eppenhof, and Josien PW Pluim. Roto-translation covariant convolutional networks for medical image analysis. *arXiv preprint arXiv:1804.03393*, 2018.
- [5] Radiation Oncology Branch. Brain tumor segmentation in mri scans using deeply-supervised neural networks. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers*, volume 10670, page 320. Springer, 2018.
- [6] Adrià Casamitjana, Santi Puch, Asier Aduriz, Elisa Sayrol, and Verónica Vilaplana. 3d convolutional networks for brain tumor segmentation. *Proceedings of the MICCAI Challenge on Multimodal Brain Tumor Image Segmentation (BRATS)*, pages 65–68, 2016.
- [7] Peter D Chang. Fully convolutional neural networks with hyperlocal features for brain tumor segmentation. In *Proceedings MICCAI-BRATS Workshop*, pages 4–9, 2016.
- [8] Lele Chen, Yue Wu, Adora M DSouza, Anas Z Abidin, Axel Wismüller, and Chenliang Xu. Mri tumor segmentation with densely connected 3d cnn. In *Medical Imaging 2018: Image Processing*, volume 10574, page 105741F. International Society for Optics and Photonics, 2018.
- [9] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2018.
- [10] Benjamin Chidester, Minh N Do, and Jian Ma. Rotation equivariance and invariance in convolutional neural networks. *arXiv preprint arXiv:1805.12301*, 2018.
- [11] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999, 2016.
- [12] Nicholas C Cullen and Brian B Avants. Convolutional neural networks for rapid and simultaneous brain extraction and tissue segmentation. In *Brain Morphometry*, pages 13–34. Springer, 2018.
- [13] Sander Dieleman, Jeffrey De Fauw, and Koray Kavukcuoglu. Exploiting cyclic symmetry in convolutional neural networks. *arXiv preprint arXiv:1602.02660*, 2016.
- [14] Pavel Dvorak and Bjoern Menze. Structured prediction with convolutional neural networks for multimodal brain tumor segmentation. *Proceeding of the Multimodal Brain Tumor Image Segmentation Challenge*, pages 13–24, 2015.
- [15] Eli Gibson, Wenqi Li, Carole Sudre, Lucas Fidon, Dzhoshkun I Shakir, Guotai Wang, Zach Eaton-Rosen, Robert Gray, Tom Doel, Yipeng Hu, et al. Niftynet: a deep-learning platform for medical imaging. *Computer methods and programs in biomedicine*, 158:113–122, 2018.
- [16] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31, 2017.
- [17] Saddam Hussain, Syed Muhammad Anwar, and Muhammad Majid. Segmentation of glioma tumors in brain using deep convolutional neural network. *Neurocomputing*, 282:248–261, 2018.
- [18] Sajid Iqbal, M Usman Ghani, Tanzila Saba, and Amjad Rehman. Brain tumor segmentation in multi-spectral mri using convolutional neural networks (cnn). *Microscopy research and technique*, 81(4):419–427, 2018.
- [19] Konstantinos Kamnitsas, Enzo Ferrante, Sarah Parisot, Christian Ledig, Aditya V Nori, Antonio Criminisi, Daniel Rueckert, and Ben Glocker. Deepmedic for brain tumor segmentation. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 138–149. Springer, 2016.
- [20] Konstantinos Kamnitsas, Christian Ledig, Virginia FJ Newcombe, Joanna P Simpson, Andrew D Kane, David K Menon, Daniel Rueckert, and Ben Glocker. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical image analysis*, 36:61–78, 2017.
- [21] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [22] Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. *arXiv preprint arXiv:1802.03690*, 2018.
- [23] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [24] TKTK Lun and W Hsu. Brain tumor segmentation using deep convolutional neural network. *Proceedings of BRATS-MICCAI*, 2016.
- [25] Chao Ma, Gongning Luo, and Kuanquan Wang. Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of mr images. *IEEE Transactions on Medical Imaging*, 2018.
- [26] Midas Meijs and Rashindra Manniesing. Artery and vein segmentation of the cerebral vasculature in 4d ct using a 3d fully convolutional neural network. In *Medical Imaging 2018: Computer-Aided Diagnosis*, volume 10575, page 105751Q. International Society for Optics and Photonics, 2018.
- [27] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993, 2015.
- [28] Sérgio Pereira, Adriano Pinto, Victor Alves, and Carlos A Silva. Brain tumor segmentation using convolutional neural networks in mri images. *IEEE transactions on medical imaging*, 35(5):1240–1251, 2016.
- [29] Siamak Ravanbakhsh, Jeff Schneider, and Barnabas Poczos. Equivariance through parameter-sharing. *arXiv preprint arXiv:1702.08389*, 2017.
- [30] Muhammad Imran Razzak and Bandar Alhaqabani. Automatic detection of malarial parasite using microscopic blood images. *Journal of Medical Imaging and Health Informatics*, 5(3):591–598, 2015.
- [31] Muhammad Imran Razzak and Saeeda Naz. Microscopic blood smear segmentation and classification using deep contour aware cnn and extreme machine learning. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, pages 801–807. IEEE, 2017.
- [32] Holger R Roth, Hirohisa Oda, Xiangrong Zhou, Natsuki Shimizu, Ying Yang, Yuichiro Hayashi, Masahiro Oda, Michitaka Fujiwara, Kazunari Misawa, and Kensaku Mori. An application of cascaded 3d fully convolutional networks for medical image segmentation. *Computerized Medical Imaging and Graphics*, 66:90–99, 2018.
- [33] Rupsa Saha, Ashish Phophalia, and Suman K Mitra. Brain tumor segmentation from multimodal mr images using rough sets. In *International Conference on Computer Vision, Graphics, and Image processing*, pages 133–144. Springer, 2016.
- [34] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

- [35] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320, 2010.
- [36] Gregor Urban, M Bendszus, F Hamprecht, and J Kleesiek. Multi-modal brain tumor segmentation using deep convolutional neural networks. *MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings, winning contribution*, pages 31–35, 2014.
- [37] Kiran Vaidhya, Subramaniam Thirunavukkarasu, Varghese Alex, and Ganapathy Krishnamurthi. Multi-modal brain tumor segmentation using stacked denoising autoencoders. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 181–194. Springer, 2015.
- [38] Bastiaan S Veeling, Jasper Linmans, Jim Winkens, Taco Cohen, and Max Welling. Rotation equivariant cnns for digital pathology. *arXiv preprint arXiv:1806.03962*, 2018.
- [39] Christian Wachinger, Martin Reuter, and Tassilo Klein. Deepnat: Deep convolutional neural network for segmenting neuroanatomy. *NeuroImage*, 170:434–445, 2018.
- [40] Guotai Wang, Wenqi Li, Sébastien Ourselin, and Tom Vercauteren. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In *International MICCAI Brainlesion Workshop*, pages 178–190. Springer, 2017.
- [41] Marysia Winkels and Taco S Cohen. 3d g-cnns for pulmonary nodule detection. *arXiv preprint arXiv:1804.04656*, 2018.
- [42] Jim Winkens, Jasper Linmans, Bastiaan S Veeling, Taco S Cohen, and Max Welling. Improved semantic segmentation for histopathology using rotation equivariant convolutional networks. 2018.
- [43] Daniel E Worrall, Stephan J Garbin, Daniyar Turmukhambetov, and Gabriel J Brostow. Harmonic networks: Deep translation and rotation equivariance. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, volume 2, 2017.
- [44] Xiaomei Zhao, Yihong Wu, Guidong Song, Zhenye Li, Yong Fan, and Yazhuo Zhang. Brain tumor segmentation using a fully convolutional neural network with conditional random fields. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 75–87. Springer, 2016.
- [45] Xiaomei Zhao, Yihong Wu, Guidong Song, Zhenye Li, Yazhuo Zhang, and Yong Fan. A deep learning model integrating fcnn and crfs for brain tumor segmentation. *Medical image analysis*, 43:98–111, 2018.