Brain tumor classification with tumor segmentations and a dual path residual convolutional neural network from MRI and pathology images

Yunzhe Xue[1], Yanan Yang[1], Fadi G. Farhat[1], Frank Y. Shih[1], Olga Boukrina[2], A.M. Barrett[3], Jeffrey R. Binder[4], William W. Graves[5], and Usman W. Roshan[1]

[1] Department of Computer Science, New Jersey Institute of Technology, Newark NJ, USA
[2] Stroke Rehabilitation Research, Kessler Foundation, West Orange, NJ, USA
[3] Emory University and Atlanta VA Medical Center, Atlanta, GA, USA
[4] Department of Neurology, Medical College of Wisconsin, Milwaukee, WI, USA
[5] Department of Psychology, Rutgers University – Newark, Newark, NJ, USA

Abstract. Brain tumor classification plays an important role in brain cancer diagnosis and treatment. Pathologists typically have to work through numerous pathology images that can be in the order of hundreds or thousands which takes time and is prone to manual error. Here we investigate automating this task given pathology images as well as 3D MRI volumes without lesion maps. We use data provided by the CPM-RadPath 2019 MICCAI challenge. We first evaluate accuracy on the validation dataset with MRI and pathology images separately. We predict the 3D tumor mask with our custom developed tumor segmentation model that we used for the BraTS 2019 challenge. We show that the predicted tumor segmentations give a higher validation accuracy of 77.1% vs. 69.8% with MRI images when trained by a 3D residual convolutional neural network. For pathology images we train a 2D residual network and obtain a 66.2% validation accuracy. In both cases we find high training accuracies above 95% which suggests overfitting. We propose a dual path residual convolutional neural network model that trains simultaneously from both MRI and pathology images and we use a simple method to prevent overfitting. One path of our network is fully 3D and considers 3D tumor segmentations as input while the other path considers pathology images. To prevent overfitting we stop training after 90% training accuracy at the epoch number where our network loss increases in the following one. With this approach we achieve a validation accuracy of 84.9% showing that indeed combining the two image sources yields a better overall accuracy.

Keywords: Convolutional neural networks · residual networks · pathology · brain MRI

1 Introduction

Brain cancer tumors fall into different categories as given by the World Health Organization [1–3]. The correct prediction of tumor type plays a key role in diagnosis and treatment. However, pathologists typically have to browse numerous images to determine the tumor type which requires considerable training, is time intensive, and is prone to manual errors. The automated classification of tumor type can greatly speed up physician diagnosis and lead to better care and treatment.

The CPM-RadPath 2019 MICCAI challenge is to automatically predict three tumor types given below.

- Lower grade astrocytoma, IDH-mutant (Grade II or III)
- Oligodendroglioma, IDH-mutant, 1p/19q codeleted (Grade II or III)
- Glioblastoma and Diffuse astrocytic glioma with molecular features of glioblastoma, IDH-wildtype (Grade IV).

The contest provides MRI and pathology images from 221 patients as training data and 35 as validation. For each patient we have 3D MRI images in four modalities: native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR). All brain scans were obtained with different clinical protocols and from various scanners from different institutions. The images were all co-registered to the same anatomical template, interpolated to the same resolution (1 mm^3) and skull-stripped.

We are also given varying number of pathology images for each patient. These are digitized whole slide tissue images captured from Hematoxylin and Eosin (H&E) stained tissue specimens. The tissue specimens were scanned at 20x or 40x magnifications. In Figure 1 we see a cropped pathology image with a Grade IV tumor (class G) from this dataset.

Inspired by the success of convolutional neural networks in image recognition tasks, we present a dual-path residual convolutional neural network solution to this problem. We find that using predicted tumor segmentations of each MRI image leads to higher overall validation accuracy than if we used the original MRI images. We also see that our model achieves above 95% training accuracy which suggests overfitting. With careful training we achieve a validation accuracy of 84.9% with both datasets which is higher than the accuracy with predicted tumor segmentations or pathology images alone.

2 Methods

2.1 Custom designed U-Network for predicting tumor segmentations

In Figure 2 we show our custom designed U-Network to predict tumor segmentations from MRI images [4]. We trained our network on data from the Brain Tumor Segmentation (BraTS) 2019 MICCAI challenge [5, 6]. We see our network takes images in four modalities and predicts segmentations of three regions of the tumor.

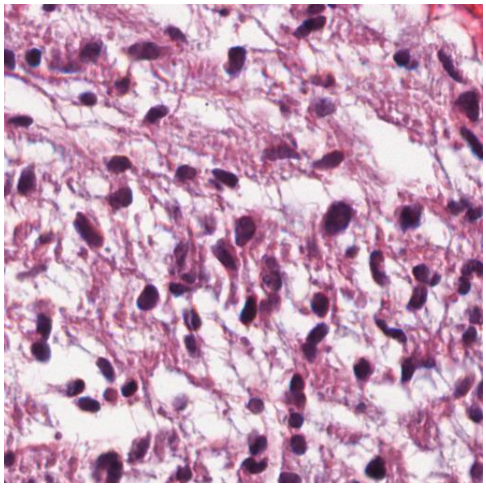


Fig. 1. A typical cropped pathology image taken from the CPM-RadPath dataset with a Grade IV tumor (class G)

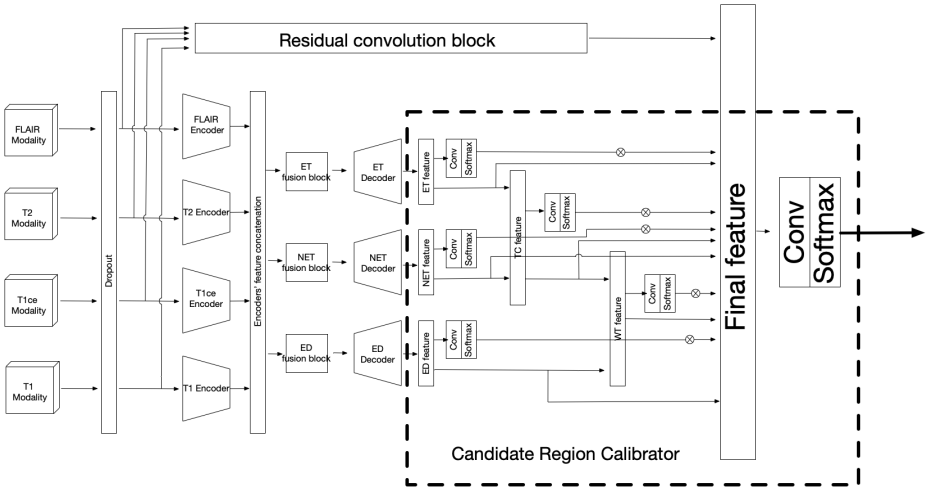
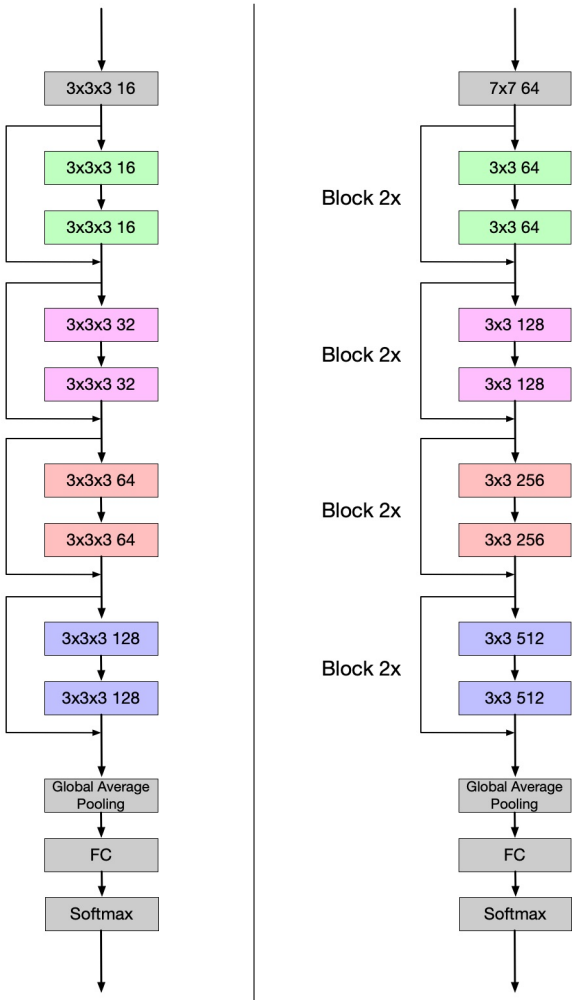


Fig. 2. Our custom designed multi-modal tumor segmentation network

2.2 Dual-path residual convolutional neural network

The ResNet18 architecture [7] uses residual connections between layers to prevent gradient vanishing problems and is a highly successful approach. In Figure 3(a) and 3(b) we show the ResNet18 convolutional neural network archi-

tectures that we use separately on MRI and pathology images respectively. We combine them in a dual-path model as shown in Figure 4.



(a) ResNet18 for 3D brain MRI and for tumor segmentations

(b) ResNet18 for pathology images

Fig. 3. Our ResNet18 networks for 3D tumor and pathology images. In each block is shown the size and number of convolutional kernels all with stride 1 except for the first convolutional block that has stride 2.

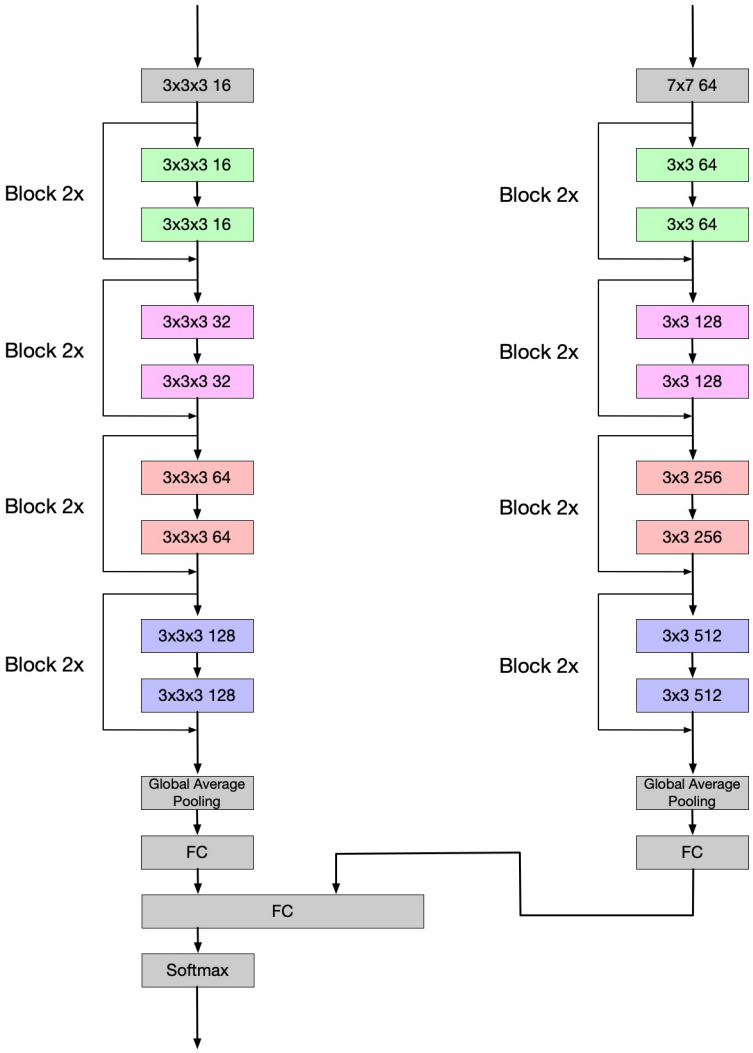


Fig. 4. Our combined network model for both tumor segmentations and pathology images. In each block is shown the size and number of convolutional kernels all with stride 1 except for the first convolutional block that has stride 2.

2.3 Model training and parameters

We use the standard cross-entropy loss function [8] to predict the three tumor classes. We implement our network using the Pytorch library [9].

3D ResNet18 training We train our network for 60 epochs, learning rate of 0.01, stochastic gradient descent with Nesterov, a batch size of 8, and no weight decay.

2D ResNet18 training We train our network for 100 epochs, learning rate of 0.01, stochastic gradient descent with Nesterov, a batch size of 128, and no weight decay.

Combined model training Our combined model takes in both tumor segmentations and pathology images as input for each patient. For each tumor segmentation we randomly pick 8 pathology images of the patient that go into the same batch during training. If a patient has less than 8 pathology images (which occur in some cases) we simply select randomly with replacement. At the end of the 2D part of our combined model is an average operation that averages the features of the 8 images into one layer that is then concatenated into the 3D part (see Figure 4),

We train our network for 50 epochs, learning rate of 0.01, stochastic gradient descent with Nesterov, a batch size of 8, and no weight decay.

Early stopping to prevent overfitting To prevent overfitting we train our model until it reaches a 90% training accuracy. After that we will stop at the epoch if loss increases in the following one.

2.4 Data preprocessing and augmentation

3D ResNet18 data preprocessing We normalize the data by subtracting the mean and dividing by standard deviation to give 0 mean and unit variance. We crop and pad original images from dimensions $240 \times 240 \times 155$ to $160 \times 192 \times 160$.

2D ResNet18 preprocessing We randomly crop each image from dimensions 512×512 to 224×224. We also study a center crop variant. We perform random horizontal flip on images during both the training and inference processes.

Combined model preprocessing Here we preprocess the MRI images and pathology ones using the same methods described above in the individual networks.

3 Results

We use our custom designed 3D network [4] for the Brain Tumor Segmentation (BraTS) 2019 MICCAI challenge [5, 6] to predict tumor segmentations for each MRI image. In Figure 5 we show tumor segmentations by our BraTS model for each of the three different axial planes of a given slice of an MRI image. We

see that the predicted tumor is highly accurate when compared with the true tumor segmentation across all four image modalities. We conjecture that the tumor position and size play a bigger role in determining the tumor type than the entire MRI image. Thus we consider these as inputs to our models vs. the original MRI images.

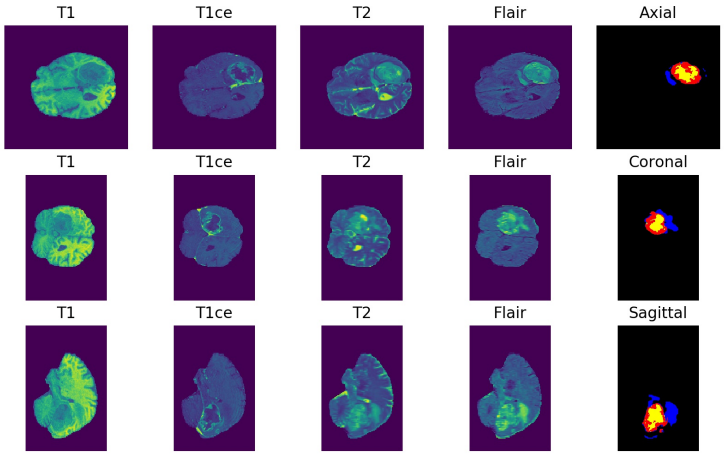
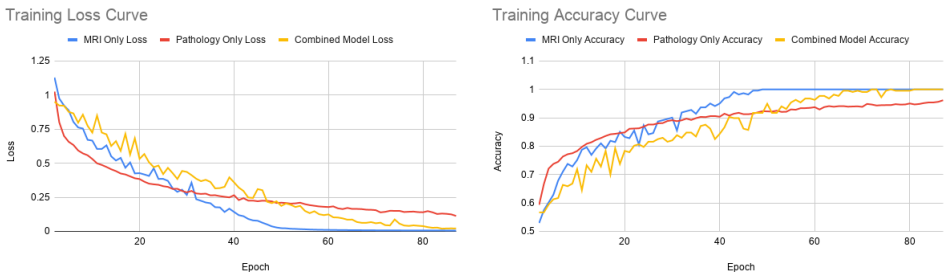


Fig. 5. Tumor segmentations given by our BraTS model in all three axial planes for a given slice across four image modalities. We use the predicted tumor segmentations (that we see are highly accurate in this example) as input to our model to classify the tumor type.

We first examine our model training lose and accuracy. In Figure 6 we see the training loss and accuracy of our models on the predicted tumor segmentations, pathology images, and combined images model. In all three cases we see a high training accuracy suggesting that we may be overfitting. In order to avoid this we perform early stopping as described above.

We now proceed to the validation accuracies with different training datasets in Table 1. First we see that indeed the tumor segmentations give a higher validation accuracy of 77.1% than using MRI images alone which give 69.8%. We also see that validation accuracy on pathology images alone is lower than that of MRI and tumor images. In the case of random crops on pathology images it varies between 66.2% and 69.2%.

Combining the MRI images with pathology images under random crops gives us 78.7% validation accuracy whereas combining with tumor segmentations gives us 81.6%. Finally combing MRI images with pathology under center crop also



(a) Training loss (b) Training accuracy

Fig. 6. Our individual and combined model training loss and accuracy.

gives 78.7% while combining tumor segmentations with pathology images under center crop gives us the best validation accuracy of 84.9%.

Brain MRI images	69.8
Predicted tumor segmentations	77.1
Pathology (center crop)	66.2
Pathology (random crop)	66.2-69.2
Combined MRI + pathology (random crop)	78.7
Combined MRI + pathology (center crop)	78.7
Combined tumor + pathology (random crop)	81.6
Combined tumor + pathology (center crop)	84.9

Table 1. Validation accuracy from different training datasets

4 Conclusion

We show that with predicted tumor segmentations we can achieve a higher accuracy for predicting tumor category than if we used the original MRI images. We present a dual path residual convolutional neural network trained on both tumor segmentations and pathology images simultaneously. We show that the combined model achieves a higher accuracy of 84.9% than if we used the tumor or pathology images alone which achieve 77.1% and 66.2% respectively.

References

1. David N Louis, Arie Perry, Guido Reifenberger, Andreas Von Deimling, Dominique Figarella-Branger, Webster K Cavenee, Hiroko Ohgaki, Otmar D Wiestler, Paul Kleihues, and David W Ellison. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*, 131(6):803–820, 2016.

2. Paul Kleihues, David N Louis, Bernd W Scheithauer, Lucy B Rorke, Guido Reifenberger, Peter C Burger, and Webster K Cavenee. The WHO classification of tumors of the nervous system. *Journal of Neuropathology & Experimental Neurology*, 61(3):215–225, 2002.
3. Daniel J Brat, Kenneth Aldape, Howard Colman, Eric C Holland, David N Louis, Robert B Jenkins, BK Kleinschmidt-DeMasters, Arie Perry, Guido Reifenberger, Roger Stupp, et al. cIMPACT-NOW update 3: recommended diagnostic criteria for diffuse astrocytic glioma, idh-wildtype, with molecular features of glioblastoma, WHO grade iv. *Acta neuropathologica*, 136(5):805–810, 2018.
4. Yunzhe Xue, Meiyan Xie, Fadi Farhat, Olga Boukrina, A. M. Barrett, Jeffrey R. Binder, Usman W. Roshan, and William W. Graves. A multi-path decoder network for brain tumor segmentation. *Proceedings of MICCAI BraTS 2019 challenge*, 2019.
5. 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):1993–2024, 2014.
6. Spyridon Bakas, Mauricio Reyes, Andras Jakab, Stefan Bauer, Markus Rempfler, Alessandro Crimi, Russell Takeshi Shinohara, Christoph Berger, Sung Min Ha, Martin Rozycki, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:1811.02629*, 2018.
7. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
8. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
9. Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *NIPS-W*, 2017.