2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Brain Tumors Classification for MR images based on Attention Guided Deep Learning Model

Yuhao Zhang[1], Shuhang Wang[2], Haoxiang Wu[3], Kejia Hu[4,5,*], Shufan Ji[6,*]

***Abstract*— Magnetic Resonance Imaging (MRI) technology has been widely applied to generate high-resolution images for brain tumor diagnosis. However, manual image reading is very time and labor consuming. Instead, automatic tumor detection based on deep learning models has emerged recently. Although existing models could well detect brain tumors from MR images, they seldom distinguished primary intracranial tumors from secondary ones. Therefore, in this paper, we propose an attention guided deep Convolution Neural Network (CNN) model for brain tumor diagnosis. Experimental results show that our model could effectively detect tumors from brain MR images with 99.18% average accuracy, and distinguish the primary and secondary intracranial tumors with 83.38% average accuracy, both under ten-fold cross-validation. Our model, outperforming existing works, is competitive to medical experts on brain tumor diagnosis.**

I. INTRODUCTION

According to World Cancer Report provided by World Health Organization (WHO), there are approximately 297,000 new cases of brain tumor worldwide in 2018. Brain tumors are usually classified into two types according to their origins: the primary intracranial tumor originates in the brain (e.g. meningioma, glioma, and pituitary tumor); while the secondary intracranial tumor originates from malignant tumors elsewhere in the body spreading to the brain (e.g. metastatic sites from colorectal, lung, and breast). Magnetic Resonance Imaging (MRI) technology, which can generate high-resolution MR images of brain structures, is widely adopted for clinical brain tumor diagnosis. However, manual MR image reading is very time and labor consuming. Instead, automatic tumor detection based on deep learning models has emerged recently.

Nowadays, CNN deep learning models [1]-[7] are widely applied in computer vision classification. In 2017, Xie et al. combined residual block and inception structure to propose ResNeXt [3], while Hu et al. delivered SENet [5] based on channel attention module. In 2020, Zhang et al. proposed ResNeSt [7] based on the split attention module. Those CNN models provided an effective tool for automatic brain tumor detection [8]. For example, Ge et al. proposed a novel multi-scale CNN model for gliomas grading with 89.47% accuracy [9]. However, the CNN models usually depend on large amounts of labelled images for training. In clinical applications, it is very difficult to get adequate training images from patients, well labelled with tumor locations and types by medical experts.

Fortunately, transfer learning technique [10] was proposed to solve the problem of inadequate training data for CNN models. Transfer learning is a popular approach in deep learning, where a model trained for a task is reused as the starting point for another model on a second task. In transfer learning, we first train a base model on the base dataset (with adequate training data) and task, and then we repurpose the learned features, or transfer them, to a second target model to be trained on the target dataset (with inadequate training data) and task. This process will tend to work if the features are general to both base and target tasks, instead of specific to the base task. Accordingly, in 2019, Swati et al. used a pre-trained CNN model and proposed a block-wise fine-tuning strategy to classify glioma, meningioma, and pituitary tumor, with an accuracy of 94.82% [11].

Although existing models could well detect brain tumors from MR images, they seldom distinguished primary intracranial tumors from secondary ones. Therefore, in this paper, we propose a novel attention guided CNN model named ResNeSAt, for automatic brain tumor detection and tumor type classification on brain MR images. We adopt an advanced CNN model ResNeSt [7] trained by images from ImageNet [12] as the base model, and then employ transfer learning technique to train the target model on brain tumor images (labelled with tumor locations and types) for our target task. Moreover, the target model is further improved by adding spatial attention strategy for tumor location detection, forming our proposed attention guided CNN model ResNeSAt. After fine-tuning, our model could effectively detect tumors from brain MR images with 99.18% average accuracy, and distinguish the primary and secondary intracranial tumors with 83.38% average accuracy, both under ten-fold cross-validation. Experimental results show that, our model not only outperforms existing works but also provides tumor type classification, which is competitive to medical experts on brain tumor diagnosis.

II. MATERIAL

In this study, we use the brain MR images provided by Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine. All experiments were performed in compliance

[1]School of Sino-French Engineer, Beihang University, Beijing, 100191, China
[2]Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, 02138, USA
[3]School of Medicine, Shanghai Jiao Tong University, 200025, China.
[4]Center for Functional Neurosurgery, Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, 200025, China.
[5]Department of Neurosurgery, Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, 200025, China.
[6]School of Computer Science and Engineering, Beihang University, Beijing, 100191, China
*Co-corresponding author; E-mail address: dockejiahu@gmail.com, jishufan@buaa.edu.cn

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

TABLE I

DETAILS OF MR IMAGES.

Tumor Presence	Tumor Type	Tumor Sub-type	No. of Patients	No.of MR Images
Yes	Primary intracranial tumors	Cavernous hemangioma	3	75
		Glioma	5	316
		Meningioma	5	78
		Acoustic neuroma	2	38
	Secondary intracranial tumors	Colorectal cancer brain metastasis	7	226
		Lung cancer brain metastasis	29	818
		Breast cancer brain metastasis	8	285
No	-	-	11	1834
Total	**-**	**-**	**70**	**3670**

with the Declaration of Helsinki. Written informed consent was acquired from each patient or next of kin. Patients underwent MR scans on a 3T scanner using a 16-channel head coil in a supine position. In order to facilitate the therapeutic planning and treatment of GKRS, a gamma knife rigid head frame (Elekta, Stockholm, Sweden) matched with the head coil, was fixed on patient's head. The T1 MPRAGE sequence was performed on all of the patients 5 minutes after administration of 0.1 mmol/kg body weight of contrast agent (Gadoxetic Acid Disodium Injection, Primovist, Bayer Vital GmbH, Leverkusen, Germany). Finally, we collected 3670 T1 MPRAGE sequence brain MR images from 70 patients (59 with tumors and 11 without tumors). Among patients with tumors, 15 patients were labelled with primary intracranial tumors (including cavernous hemangioma, glioma, meningioma, and acoustic neuroma); while 44 patients were labelled with secondary intracranial tumors (including colorectal cancer brain metastasis, lung cancer brain metastasis, and breast cancer brain metastasis). Table 1 and Figure 1 represent the details and examples of those MR images, respectively.

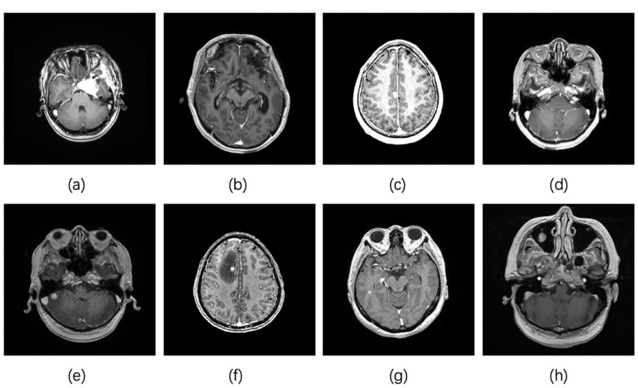

Fig. 1. Examples of MR images. (a) Cavernous hemangioma, (b) Glioma, (c) Meningioma, (d) Acoustic neuroma, (e) Colorectal cancer brain metastasis, (f) Lung cancer brain metastasis, (g) Breast cancer brain metastasis, (h) No tumor.

III. METHODOLOGY

We propose a novel attention guided CNN model ResNeSAt, for automatic brain tumor detection and tumor type classification on brain MR images. The framework of our method is shown in Figure 2.

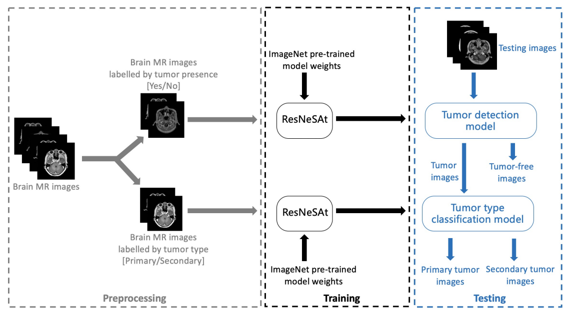

Fig. 2. Framework of the proposed method.

A. MR Image Preprocessing

The brain MR images were preprocessed to facilitate model training. First, all images were unified to size 256×256 pixels by bilinear interpolation; and then 50% of the images were randomly flipped horizontally, changing pixel positions but keeping image features, so as to enhance the model generalization; at last, all pixel values ranged [0,255] were normalized into the range [-1,1] to avoid imbalanced pixel intensity distribution.

B. Model Training

After preprocessing, the MR brain images were randomly divided into training set and testing set. The training set was applied to train two models: images labelled with tumor presence (Yes/No) were employed for tumor detection model; while images labelled with tumor type (Primary/Secondary) were employed for tumor type classification. As the amount of MR images was inadequate for CNN model training, we took the transfer learning technique by adopting the base model ResNeSt [7] pre-trained by ImageNet [12] on pytorch. While training our target model on brain MR images, we improved ResNeSt by inserting the Spatial Attention module (SA) [13] into each bottleneck layer to emphasize the tumor area, as shown in Figure 3. $\otimes$ and $\oplus$ represent element-wise multiplication and element-wise summation, respectively. This improvement makes our ResNeSAt perform better on tumor detection than ResNeSt [7]. The detailed network layers of our ResNeSAt is presented in Table 2.

TABLE II

DETAILED NETWORK LAYERS OF RESNESAT.

ResNeSAt	Name	Details	Repeat
	Input		
InitConv	InitConv 1	Conv3, BN, ReLU	2
	InitConv 2	Conv3, BN, ReLU, MaxPool	1
Layer 1	Bottleneck 1	Conv1, BN, SplAtConv, Conv1, BN, ReLU, SA, AvgPool, Conv1, BN	1
	Bottleneck 2	Conv1, BN, SplAtConv, Conv1, BN, ReLU, SA	2
Layer 2	Bottleneck 3	Conv1, BN, AvgPool, SplAtConv, Conv1, BN, ReLU, SA, AvgPool, Conv1, BN	1
	Bottleneck 4	Conv1, BN, SplAtConv, Conv1, BN, ReLU, SA	3
Layer 3	Bottleneck 5	Conv1, BN, AvgPool, SplAtConv, Conv1, BN, ReLU, SA, AvgPool, Conv1, BN	1
	Bottleneck 6	Conv1, BN, SplAtConv, Conv1, BN, ReLU, SA	5
Layer 4	Bottleneck 7	Conv1, BN, AvgPool, SplAtConv, Conv1, BN, ReLU, SA, AvgPool, Conv1, BN	1
	Bottleneck 8	Conv1, BN, SplAtConv, Conv1, BN, ReLU, SA	2
	Classifier	GlobalAvgPool, FC	1

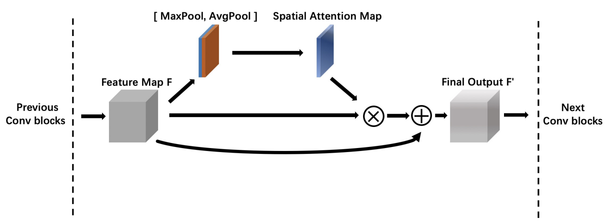

Fig. 3. SA integrated with bottleneck in ResNeSAt.

C. Testing and Evaluation

After model training and fine-tuning, the MR images in testing set were taken to evaluate the performance of models on tumor detection and tumor type classification. As shown in Figure 2, the testing images were first sent into tumor detection model to classify whether there exists any tumor or not; and then images with tumors were sent into tumor type classification model to distinguish their tumor types. The performance of above two models were evaluated on recall, specificity, precision, F1-score, and accuracy, defined in the following formulas by four indices: Truly Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN).

$$Recall = \frac{TP}{TP + FN} \tag{1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1 - Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{4}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Recall and Specificity describe how good the classifier is at classifying the positive and negative condition, respectively; Precision is the positive predictive rate; F1-score measures the classification performance in term of recall and precision; and Accuracy is the global measure of classification performance.

IV. EXPERIMENTAL RESULTS

In this section, we will compare our proposed model ResNeSAt with state-of-the-arts CNN deep learning models [1]-[7] on brain tumor detection and tumor type classification. All models were implemented by pytorch, on Huawei G5500 series servers with 1 NVIDIA V100 GPU card.

In each round of iterative training, model parameters are automatically adjusted based on the prediction results of the previous round. The maximum number of training iterations is set to 100 epochs. We take SGD optimizer to update the weights of the model network with a batch size of 16. And the learning rate α follows the cosine decays defined as:

$$\alpha = \frac{1}{2} \times (1 + cos(\frac{e \times \pi}{N_e})) \times \alpha_0 \tag{6}$$

e is an epoch counter, N_e is a total number of epochs, and α_0 is the initial learning rate. α_0 is set to 1e-3 and 1e-4 for the tumor detection model and tumor type classification model, respectively.

Firstly, we compared state-of-the-arts CNN deep learning models [1]-[7] on brain tumor detection and tumor type classification. 80% of the brain MR images were randomly selected as the training set; and the rest 20% were left to evaluate their performance on accuracy. Note that images from the same patient were never divided into the different two sets.

The performance of seven state-of-the-arts CNN models is shown in Figure 4. It is clear that the ResNeSt model (red line) performs best on both tasks with highest accuracy. Therefore, we would like to compare our model ResNeSAt against ResNeSt in more detailed aspects.

To remove the performance variation caused by random training image selection, we evaluated the performance of ResNeSt and ResNeSAt based on ten fold cross-validation. Table III and Table IV show the average recall, specificity, precision, F1-score and accuracy of ResNeSt vs. ResNeSAt. In table III, we find that ResNeSAt outperforms ResNeSt in all aspects on brain tumor detection. Table IV shows that, ResNeSAt performs better than ResNeSt on tumor type classification in terms of recall, F1-score, and accuracy. In addition, the parameters in ResNeSt and ResNeSAt are 97.45 MB and 97.46 MB, respectively. That is, ResNeSAt effectively improves the performance of ResNeSt on brain tumor detection and tumor type classification, with few increases on model size.

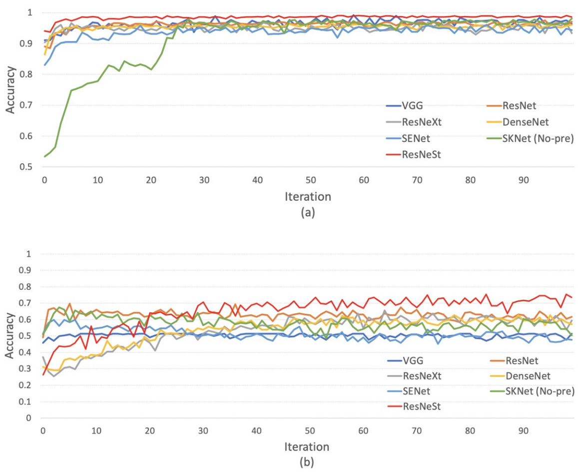

Fig. 4. Comparison on state-of-the-arts CNN Models. (a) Tumor presence detection, (b) Tumor type classification.

TABLE III
COMPARISON ON BRAIN TUMOR DETECTION.

	Recall	Specificity	Precision	F1-score	Accuracy
ResNeSt	0.9733	0.9772	0.9803	0.9768	0.9751
ResNeSAt	0.9864	0.9981	0.9983	0.9923	0.9918

TABLE IV
COMPARISON ON TUMOR TYPE CLASSIFICATION.

	Recall	Specificity	Precision	F1-score	Accuracy
ResNeSt	0.7927	0.8684	0.8718	0.8304	0.8283
ResNeSAt	0.8212	0.8480	0.8591	0.8397	0.8338

V. CONCLUSIONS & FUTURE WORK

In this paper, we propose a novel attention guided CNN model ResNeSAt for automatic tumor detection and tumor type classification on brain MR images. Compared with existing works, we are the first to distinguish the primary intracranial tumors from secondary ones. In addition, our ResNeSAt model outperforms state-of-the-arts CNN models on brain tumor detection and tumor type classification with few increases on model size. Our ResNeSAt model could effectively detect tumors with 99.18% average accuracy, and classify tumor types with 83.38% average accuracy, which would be a fast and helpful tool for clinical diagnosis. Future work will be done on tumor area segmentation and tumor sub-type classification with increased amounts of clinical data.

ACKNOWLEDGMENT

This research is supported by the National Key R&D Program of China (Grant No.2019YFB2102100); the National Natural Science Foundation of China (Grant No.82161148011, No.92046010); Shanghai Pujiang Program (KH, 19PJ1407500); and Medical and Engineering Cross Research Fund from Shanghai Jiao Tong University (KH, YG2019QNA31). We would like to thank all patients or next of kin for providing tumor image data.

REFERENCES

[1] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[2] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[3] Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1492-1500.

[4] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.

[5] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.

[6] Li X, Wang W, Hu X, et al. Selective kernel networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2019: 510-519.

[7] Zhang H, Wu C, Zhang Z, et al. Resnest: Split-attention networks[J]. arXiv preprint arXiv:2004.08955, 2020.

[8] Muhammad K, Khan S, Del Ser J, et al. Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020.

[9] Ge C, Qu Q, Gu I Y H, et al. 3D multi-scale convolutional networks for glioma grading using MR images[C]//2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, 2018: 141-145.

[10] Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks?[C]//Advances in neural information processing systems. 2014: 3320-3328.

[11] Swati Z N K, Zhao Q, Kabir M, et al. Brain tumor classification for MR images using transfer learning and fine-tuning[J]. Computerized Medical Imaging and Graphics, 2019, 75: 34-46.

[12] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[13] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.