CLASSIFICATION OF BRAIN TUMOR FROM MAGNETIC RESONANCE IMAGING USING CONVOLUTIONAL NEURAL NETWORKS

Mohammed-Amine Zyad^{1*}, Mohamed Gouskir² and Belaid Bouikhalene³

^{1,2,3}Polydisciplinary Faculty, Sultan Moulay Slimane University ¹zyad@usms.ma, ²m.gouskir@gmail.com, ³b.bouikhalene@usms.ma

Abstract— Deep learning methods gained a huge popularity in segmentation and classification of medical imaging. In this paper we propose a Convolutional Neural Network (CNN) approach which is one of the top performing methods while also being extremely computationally efficient, a balance that existing methods have struggled to achieve, we use this method as a process for segmenting brain tumor regions from magnetic resonance imaging (MRI) using CNNs. The main task for this method is using a public dataset containing 3,064 T1-weighted contrast enhanced MRI (CE-MRI) with different abnormalities from different planes. This novel method of training neural networks on this dataset has proved to be efficient than well-known methods.

Keywords- Deep Learning, brain tumor, classification, MRI

1. INTRODUCTION

Diagnosing patients for doctors and radiologists can be a time-consuming task and prone to human errors. Which can lead to mis-diagnosis or long waiting queues for patients to be diagnosed. For these reasons researchers in the field of applied health care trying to develop automated tools to help doctors in their diagnosis.

Brain magnetic resonance imaging is amongst the best imaging techniques that are used by researchers to detect brain tumors and find models of the progression of the tumor. MRI images played a big role in the automatic medical image analysis field because of the high resolution of the images which enables researcher to have a lot of information about the brain structure and pathologies within the brain regions. In fact, researchers presented different automated approaches for brain tumor segmentation and detection and also type classification. However, Support Vector Machine (SVM) and Neural Networks (NN) are widely used algorithms for the last decades because of their efficiency.

The contribution of this paper is applying Convolutional Neural Networks method and measure its accuracy to achieve automation through the classification of brain tumor types from patient's brain images to aid doctors and radiologists confidently identify the normal brain and some types of brain tumors.

2. RELATED WORK

2.1. NEURAL NETWORKS

The first apparition of Neural networks date to the early 1950's with Marvin Minsky and Dean Edmonds who tried to build the first neural network machine, able to learn, the

Received: March 20, 2019 Reviewed: April 26, 2019 Accepted: May 2, 2019



SNARC (Stochastic Neural Analog Reinforcement Calculator) [1], but this method falls out of favor until Hinton *et al.*, [2] in 2006 introduced a method of training hidden layers through unsupervised learning of restricted Boltzmann machines (RBMs) by greedily stacking these RBMs. This method showed good results in training neural networks which led it to become quickly the state of the art shown in Figure 1.

Convolutional neural networks were originally introduced in 1998 by Yann LeCun *et al.*, [3] but it gained popularity in 2012 when Krizhevsky *et al.*, [4] designed a method based on convolutional neural networks to win the ImageNet competition and outperform the state of the art model. After that the community of computer vision adopted convolutional neural network in classification of images. Since then convolutional neural networks dominated the models presented in many competitions like Galaxy Zoo Challenge and other competitions on Kaggle.

In order to prevent overfitting Ian Goodfellow *et. al.*, [5] created a neural network layer called a maxout layer, which is simply a layer where the activation function is the max of the inputs. While in in 2014 Srivastava *et. al.*, [6] introduced dropout as a simple way to prevent co-adaptation of neurons. Dropout can be thought of as a form of model averaging in which a random subnetwork is trained at every iteration and in the end the weights of the different such random networks are averaged. Since one cannot average the weights explicitly, an approximation is used. This approximation is exact for a linear network. In maxout, they do not drop the inputs to the maxout layer. Thus the identity of the input outputting the max value for a data point remains unchanged. Thus the dropout only happens in the linear part of the multilayer perceptron (MLP) but one can still approximate any function because of the maxout layer. As the dropout happens in the linear part only, they conjecture that this leads to more efficient model averaging as the averaging approximation is exact for linear networks.



Fig. 1 Architecture of a neural network showing the different layers in the network. The input layer, the hidden layers and the output layers. In addition to neurons where the calculations are done by multiplying the weights and adding biases

2.2. CONVOLUTIONAL NEURAL NETWORKS

Convolutional neural networks are very similar to ordinary neural networks, they are constituted of neurons that have learnable weights and biases by updating them in the back-propagation step. Each neuron receives some inputs, multiply it by the weights and add the biases followed by a non-linear activation function. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other. The network has a loss function on the last fully-connected layer and every intuition of neural networks still apply.

The difference between ordinary neural networks and the CNNs is that the last assumes that the inputs are images, which allows us to encode certain properties of images into the architecture. This method makes the forward propagation more efficient to implement and reduce a lot of amount of parameters in the network.

Convolutional networks have been used in medical imaging applications since 1990s in various areas like breast tissue classification and nodule detection. [7] [8] But due to the lack of datasets and the limitation of computational power by that time the research in that field were discontinued. Until recently when labeled datasets were available and more powerful graphic processing units (GPUs) have been created which enabled researchers to continue their work in the area. CNNs nowadays are widely used in brain MRI for processing data [9], detecting and segmenting lesions [10-15] and segmenting tumors [16-18], whole tissue [19-21] and sub-cortical structures [22-23].

3. METHOD

Convolutional neural networks rely on two algorithms called Forward pass (e.g., Forward propagation) and Backward pass (e.g., Backward propagation). In this section we will review those two algorithms.

3.1. FORWARD PROPAGATION

The forward pass of a fully-connected layer corresponds to one matrix multiplication followed by a bias offset and an activation function.

$$a_i^l = \mathop{a}\limits_{j=1}^n W_{ij}^l x_j + b_i$$

The variable j represents the certain input into neuron i. The activation function ReLU is then applied to layer 1 neuron is sum a_i^l to produce a new value z_i^l .

$$z_i^l = max(0, a_i^l)$$

To make our algorithm learn we need a loss function. Convolutional neural networks use

cross-entropy loss function defined as (p_j is the probability of the image being a certain type of tumor):

$$L = \frac{1}{m} \mathop{\text{a}}_{j} t_{j} log(p_{j})$$

\$m\$ here is the total number of examples in training.

In order to prevent over-fitting it is recommended to use regularization. In general, the formula for regularization term added to a loss function is:

International Journal of Advanced Science and Technology Vol.126 (2019)

$$\min_{f} \overset{n}{\underset{i=1}{\overset{n}{\overset{}}}} V(f(x_{i}), y_{i}) + / R(f)$$

With R represented as:

$$R = \frac{1}{m} \mathop{\text{a}}_{w} |w|$$

In the above formula / represents the regularization constant. It is considered as a hyper-parameter of the network and can be personalized based on the design of the CNN. Our final formula combining all the above give us the overall cost-function:

$$C = \frac{1}{m}L + R = \frac{1}{m} \mathop{a}\limits_{j}^{*} t_{j} log(p_{j}) + \frac{1}{m} \mathop{a}\limits_{w}^{*} |w|$$

3.2. BACKWARD PROPAGATION

Since convolution is a linear operation using additions and multiplications, so it's backward pass is not very complicated. In the backpropagation, we'll update the parameters from the back to start. The update consists updating the weights and the deltas.

$$\frac{\P C}{\P W^{l-1}} = \frac{\P C}{\P z^{l}} \frac{\P z^{l}}{\P W^{l-1}}$$
$$\frac{\P C}{\P b^{l-1}} = \frac{\P C}{\P z^{l}} \frac{\P z^{l}}{\P b^{l-1}}$$
$$\frac{\P C}{\P z^{l-1}} = \frac{\P C}{\P z^{l-1}} \frac{\P z^{l}}{\P z^{l-1}}$$

In the above formulas l designs the layer.

Then we update the parameters using the formula:

$$q := q - a D w_t$$

Where ∂ denotes the learning rate.

4. RESULTS AND DISCUSSION

4.1. DATASET

The dataset is taken from [24]. This brain tumor dataset containing 3064 T1-weighted contrast-enhanced images from 233 patients with three kinds of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices). Every image is originally of size 512x512. But downscaled to 256x256 for computational limits.

4.2. MODEL

In this paper we used two models as shown in Figure 2. Both of them use ReLU as activation function on every layer. Hence the difference is that the model (A) use Sigmoid as an output layer instead of Softmax layer in model (B). This difference resulted in a big improvement as we will see in the next section.

International Journal of Advanced Science and Technology Vol.126 (2019)



Fig. 2 The Two Models used to Train the Dataset. The Model (A) use a Segmoid Layer as the Output. And Model (B) use Softmax Layer as Output

4.3. RESULTS AND VALIDATION



Fig. 3 Accuracy and Loss of the Model (A)



Fig. 4 Accuracy and Loss of the Model (B)

To validate qualitatively the obtained results with ground-truth masks, we use the Dice Similarity Coefficient (DSC). This measure indicates the amount of area overlap between the automatically detected and the manually delineated brain image. This measure is calculated as follows:

$$DSC = \frac{2TP}{2TP + FP + FN}$$

where TP (True Positive) are the correct detections, FP (False Positive) are incorrect detections, and FN (False Negative) are missing detections.

5. CONCLUSION

In this paper, we proposed an automatic and supervised method for detection of brain MRI tumor, the task in this method is to automatically get accurate classification of tissues and check the presence of abnormal regions in those images using Convolutional Neural Networks algorithm to classify brain tumor types in a given MRI image. This method applied to many images proved that the use of this algorithm gives better results with little error while being computationally efficient. For future work, we consider to include color balancing step into our CNN, to boost the accuracy of our model.

REFERENCES

- [1] M. Minsky, "Neural Nets and Theories of Memory," 1963.
- [2] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," Neural computation, vol. 18, no. 7, pp. 1527–1554, 2006.

- [3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.
- [5] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio, "Maxout networks," arXiv preprint arXiv:1302.4389, 2013.
- [6] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," The Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.
- [7] S.-C. Lo, S.-L. Lou, J.-S. Lin, M. T. Freedman, M. V. Chien, and S. K. Mun, "Artificial convolution neural network techniques and applications for lung nodule detection," IEEE Transactions on Medical Imaging, vol. 14, no. 4, pp. 711–718, 1995.
- [8] B. Sahiner et al., "Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images," IEEE transactions on Medical Imaging, vol. 15, no. 5, pp. 598– 610, 1996.
- [9] J. Kleesiek et al., "Deep MRI brain extraction: a 3D convolutional neural network for skull stripping," NeuroImage, vol. 129, pp. 460–469, 2016.
- [10] Q. Dou et al., "Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1182–1195, 2016.
- [11] H. R. Roth et al., "A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations," in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2014, pp. 520–527.
- [12] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. Snead, I. A. Cree, and N. M. Rajpoot, "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1196–1206, 2016.
- [13] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1207–1216, 2016.
- [14] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, and N. Navab, "Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1313–1321, 2016.
- [15] T. Brosch, L. Y. Tang, Y. Yoo, D. K. Li, A. Traboulsee, and R. Tam, "Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1229–1239, 2016.
- [16] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1240–1251, 2016.
- [17] K. Kamnitsas et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Medical image analysis, vol. 36, pp. 61–78, 2017.
- [18] M. Havaei et al., "Brain tumor segmentation with deep neural networks," Medical image analysis, vol. 35, pp. 18–31, 2017.
- [19] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1252–1261, 2016.
- [20] W. Zhang et al., "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," NeuroImage, vol. 108, pp. 214–224, 2015.
- [21] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," NeuroImage, 2017.
- [22] J. Dolz, C. Desrosiers, and I. B. Ayed, "3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study," NeuroImage, 2017.
- [23] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: Deep convolutional neural network for segmenting neuroanatomy," NeuroImage, 2017.
- [24] J. Cheng, brain tumor dataset. figshare, 2017.

International Journal of Advanced Science and Technology Vol.126 (2019)