Brain Tumor Classification Using Convolutional Neural Network

Conference Paper · May 2019

DOI: 10.1109/ICASERT.2019.8934603

3 authors, including:

Sunanda Das
Khulna University of Engineering and Technology
22 PUBLICATIONS **195** CITATIONS
SEE PROFILE

Nishat Nayla Labiba
Khulna University of Engineering and Technology
1 PUBLICATION **84** CITATIONS
SEE PROFILE

Brain Tumor Classification Using Convolutional Neural Network

Sunanda Das, O.F.M. Riaz Rahman Aranya, Nishat Nayla Labiba
Department of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
sunanda@cse.kuet.ac.bd, aranya.riaz@gmail.com, nishatnaylawork@gmail.com

***Abstract*— Medical image classification has gained tremendous attention in recent years, and Convolutional Neural Network (CNN) is the most widespread neural network model for image classification problem. CNN is designed to determine features adaptively through backpropagation by applying numerous building blocks, such as convolution layers, pooling layers, and fully connected layers. In this paper, we mainly focused on developing a CNN model for classifying brain tumors in T1-weighted contrast-enhanced MRI images. The proposed system consists of two significant steps. First, preprocess the images using different image processing techniques and then classify the preprocessed image using CNN. The experiment is conducted on a dataset of 3064 images which contain three types of brain tumor (glioma, meningioma, pituitary). We achieved a high testing accuracy of 94.39%, average precision of 93.33% and an average recall of 93% using our CNN model. The proposed system exhibited satisfying accuracy on the dataset and outperformed many of the prominent existing methods.**

Keywords*— *Brain Tumor; Convolutional Neural Network; Kernel; Histogram Equalization; Feature Maps; Adam Optimization;

I. Introduction

The brain tumor is one of the most feared diseases in medical science. According to the American Cancer Society's [1] estimation of the brain tumor that will be diagnosed in the United States in 2019 includes 23,820 malignant tumors of the brain or spinal cord (13,410 in males and 10,410 in females). These estimation does not include benign (non-cancer) tumors. According to their evaluation, there is a possibility that 17,760 people (9,910 males and 7,850 females) could die from brain and spinal cord tumors in 2019. Survival rates of a brain tumor vary according to the type of the tumor and age of the patient. The Central Brain Tumor Registry of the United States (CBTRUS) [2] shows that the survival rate can extensively vary according to age. These data were collected on the patients who were diagnosed for the brain tumor during the year 2000-2013. The report shows that the 5-year relative survival rates for Meningioma tumor are 87.0%, 77.4%, 70.5%, 53.7%, 47.3% for age 20-44, 45-54, 55-64, 65-74, 75+ respectively.

A brain tumor is the abnormal growth of cells in the brain. A brain tumor can be primary or secondary. A primary brain tumor originates in the brain itself or tissues adjacent to it, i.e. the brain-covering membranes (meninges), cranial nerves, pituitary gland or pineal gland, while a secondary brain tumor occurs when cancer cells from other organs like lung, kidney, breast, etc. spread to the brain [3]. Primary brain tumor initially arises because of the mutations of their DNA. These mutations allow the abnormal cell to grow and the normal cell dies. It can cause brain damage and sometimes it can be life-threatening.

In this work, we proposed a CNN model that is able to accurately classify the brain tumor. Hence, the treatment for the tumor can be started at an early stage.

The remaining part of the paper is organized as follows: Section II demonstrated the related works that have been done in this field. The proposed methodology with some theoretical explanations is investigated in Section III. The result is illustrated in details with some performance measures in Section IV. The conclusion of the paper is drawn in Section V.

II. Related Work

Numerous works have been done on brain tumor classification in recent years. These works are briefly outlined as follows:

Yang et al. [4] proposed the CBIR (Content-Based Image Retrieval) method to evaluate a large image dataset. This method considers the tumor region as a query and tries to find a tumor that has the same pathological properties. The system used MID (Margin Information Descriptor) as the feature. It illustrates the image content by using the surrounding tissue of a tumor. Their proposed system was able to achieve a precision of 89.3%. But the distance metric determined by their methods is globally linear, and it disappoints to produce many local projections for separate regions. Furthermore, their system is developed using manual segmentation of the brain tumor.

Huang et al. [5] suggested a CBIR system with region-specific BoVW (Bag of Visual Words) model. It uses intensity information of the lesion region and the intensity change of the surrounding of lesion region. The proposed system also summarized that normalization and whitening gave better performance in retrieving the image. Their proposed model obtained an average precision value of 91.0%. However, the brain tumors in MRI images are outlined manually which is not proficient and advantageous. When the contour of the segmentation is far from the tumor boundary, the method strives to catch the intensity variation encompassing the tumor.

Furthermore, Huang et al. [6] described a CBIR based system. It uses BoVW model to extract information features and REML (Rank Error-based Metric Learning) algorithm to decrease the gap between low-level features and high-level connotative concepts. This system outperformed many methods with an average precision of 93.1%. However, the partition learning that was used in the system to obtain discriminative features, many of the times

gives local optimal solutions. Moreover, the distance matric REML is globally linear and strives to determine multiple regional projections for various regions in the feature space.

III. Proposed Methodology

This section illustrates the detailed methodology that is used to classify brain tumors. It has the following steps: dataset collection and description of the dataset, preprocessing the images, classification of brain tumors using CNN and performance measures for the system. Each step of the proposed system is outlined as follows:

A. Dataset Collection and Description

The brain tumor dataset retrieved from [7] comprises 3064 T1-weighted contrast-enhanced images. The dataset is collected from 233 patients with three kinds of brain tumor: meningioma (708 images), glioma (1426 images), and pituitary tumor (930 images). The dataset is arranged in MATLAB data format (.mat file). Each file reserves a struct comprising various fields of an image. These fields include the label (1 for meningioma, 2 for glioma, 3 for pituitary tumor), PID (patient ID), image data, tumor border, and tumor mask.

B. Preprocessing

We used two fields, i.e. label and image data from each of the mat files. Fig. 1 illustrates the stepwise preprocessing of the work. The original images are 512×512 in size.

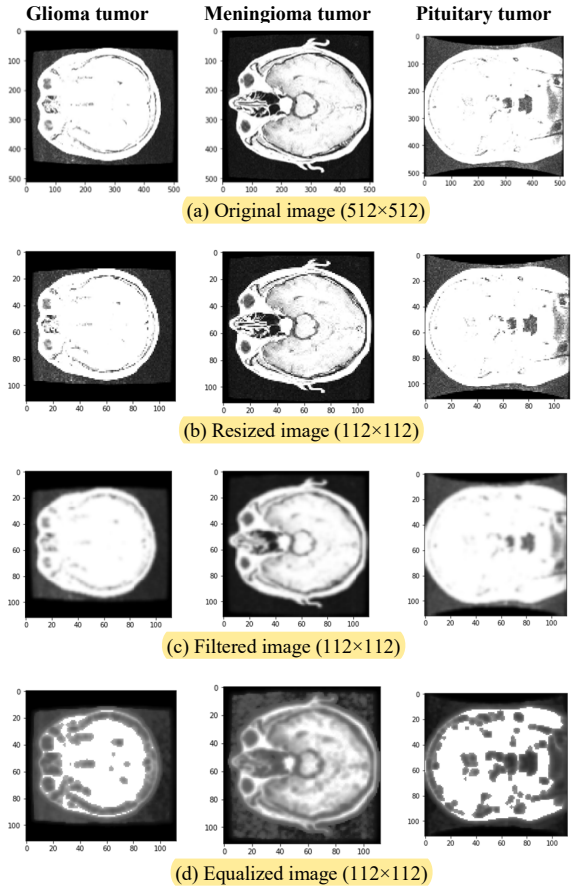


Fig. 1. Stepwise preprocessing outcome for tumor classification

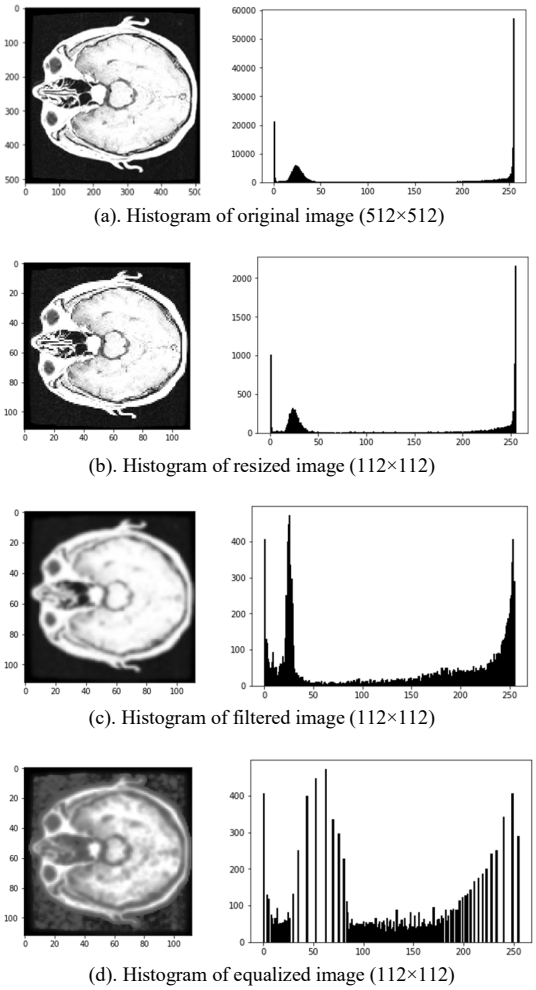


Fig. 2. Histogram of different preprocessing step image

First, we resize the images in 112×112 and then apply Gaussian filter with kernel size 5×5. Gaussian filter smooths the image. In 2-D, Gaussian function has the following form:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\pi\sigma^2}} \quad (1)$$

where x denotes the distance from the origin in the horizontal axis, y denotes the distance from the origin in the vertical axis, and σ denotes the standard deviation of the Gaussian distribution.

Then, histogram equalization is used on each of the images. Histogram graphically illustrates the intensity distribution of an image. Histogram Equalization is an image processing technique, and it enhances the contrast of images. It manages to achieve this by spreading out the most frequent intensity values of the image. Fig. 2 shows the histogram of different preprocessing step image.

The total dataset is divided into three sets. These sets include train set (2051 images), validation set (513 images) and test set (500 images). Hence, the train, validation and test set contain 66.9%, 16.7%, 16.3% of the dataset respectively. The train set is used to train the CNN model.

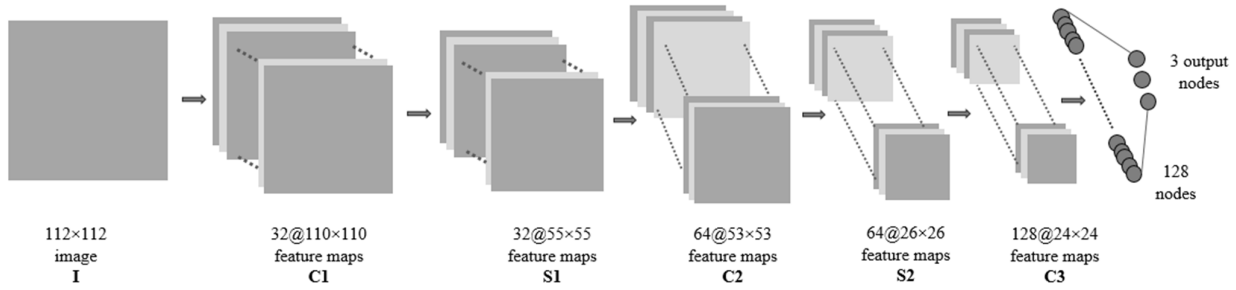


Fig. 3. CNN architecture of the model (**I, C, S** represent input image, convolution and subsampling respectively)

The validation set gives an unbiased assessment of the model while tuning the model's hyperparameters. The test set determines how well the model has been trained by evaluating the accuracy of the picked approach.

C. Tumor Classification using CNN

CNNs are unusually multi-layer neural networks. It is an effective recognition algorithm applied in pattern recognition and image processing [8]. It simulates the biological neural network through shared weight network structure. CNN is broadly practiced in the field of computer vision problem. It comprises the property of parameter sharing which decreases the number of parameters needed for the model compared to ANN (Artificial Neural Network). Moreover, the quality of the features extracted by CNN is very standard. In the CNN model we apply a series of convolution + pooling operations, followed by a fully connected layer. As we are implementing multiclass classification, the final output layer is softmax.

In our work, we used convolution on the input image using a convolution filter to generate a feature map. A 3×3 kernel is used as convolution filter. Strides = (1, 1), padding = 'valid' is used for this operation. Stride defines how much the convolution filter is moved at each step. Padding is 'valid' represents zero padding. Fig. 3 illustrates the CNN architecture that is used for this work. It includes 3 convolution layers (C1, C2, C3) with a kernel size of 5×5, 2 subsampling layers (S1, S2) with a pooling window of 2×2.

In first convolution (C1), the input image I with size 112 ×112 convolved with 32 filters with zero padding and stride = 1 which produced 110×110 sized 32 convolved features. Hence, total parameters for this operation are (3×3+1) ×32 = 320. These parameters will be learned by the algorithms in the training phase. In every convolution layer ReLU (Rectified Linear Unit) is used for activation function. It increases the non-linearity in the images. The features that are generated by the C1 layer are fed into first subsampling layer (S1). In subsampling layer max pooling is used with a window size of 2×2. Hence, 32 pooled feature maps each with size of 55×55 are produced. Pooling decreases the number of parameters which reduces the training time. Then a dropout layer is used with a dropout rate of 25%. Dropout is a regularization technique that is used to reduce overfitting [9]. Dropout is only used during training phase of the model. In dropout, some randomly selected neurons are ignored during training phase according to the dropout rate, i.e. their contribution in activating downstream neurons is shut down on the forward pass and during the backward pass, any weight updates are not applied on the neurons.

TABLE I. TRAINABLE PARAMETERS FOR CNN MODEL

Layers	Parameters
Conv2D (C1)	320
Conv2D (C2)	18,496
Conv2D (C2)	73,856
Dense (1)	9,437,312
Dense (2)	387
Total	**9,530,371**

In second convolution layer (C2), 64 convolution filters are used on the previous layer's features map that result in 64 convolved features map each with a size of 53×53. Hence, total parameters for this operation are (32×3×3+1) ×64 = 18,496. These features are fed to second subsampling layer (S2) that also uses max pooling with window size of 2×2. Therefore, 64 pooled features map each with 26×26 are generated. Then, a dropout layer with 25% dropout rate is used.

In the third and final convolution layer (C3) 128 filters are used which results in 128 features map each with size 24×24. Here, total parameters are (64×3×3+1) ×128 = 73,856. Then we used another dropout layer with a dropout rate of 40% and flatten the output. Then the generated features are fed to a dense layer with 128 nodes. Here, total parameters are (24×24×128+1) ×128 = 9437312. Then, again a dropout layer with 30% dropout rate is used. Finally, we used another dense layer with softmax activation and 3 nodes for the classification. Hence, total parameters will be (128+1) ×3 = 387. The total trainable parameters for the model are shown in Table I.

D. Performance Evaluation

The performance of the system is analyzed using confusion matrix, precision, recall, f1-score, support and ROC (Receiver Operating Characteristic) curve. For calculating precision, recall, f1-score the following formulas are used.

$$Precision = \frac{True\ Positive}{True\ positive + False\ Positive} \quad (2)$$

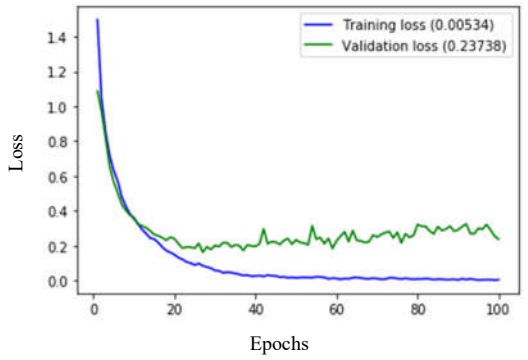
$$Recall = \frac{True\ Positive}{True\ positive + False\ Negative} \quad (3)$$

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

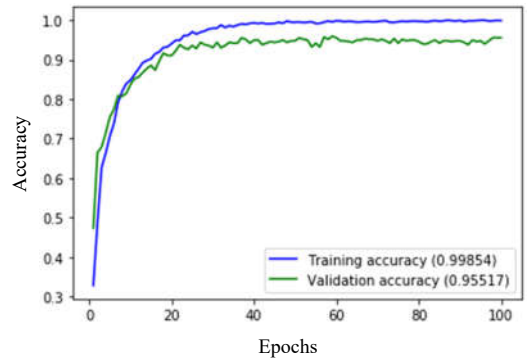
ROC curve represents a plot of the true positive rate (Sensitivity) versus the false positive rate (100-Specificity) for different cut-off points. In micro-average ROC scheme, the necessary individual value for different classes is summed up to compute the average while a macro-average ROC curve will calculate the required values for each class independently and then take the average. The area under the ROC curve (AUC) is a measure of how well the model is distinguishing between different classes. An area of 1 is considered to be the best in evaluating test cases.

IV. Result Analysis

After developing the CNN model, the model is compiled with Adam optimizer. It is a very powerful optimization algorithm used for training deep neural networks. It combines the advantages of two optimization methods i.e. AdaGrad (Adaptive Gradient Algorithm) and RMSProp (Root Mean Square Propagation) [10]. For fitting the training data with the model, a batch size of 256 and 100 epochs are used. Fig. 4 illustrates the loss and accuracy of the training set and validation set for 100 epochs. After evaluating the model on test data, we achieved a loss of 28.16% and an accuracy of 94.39%. Fig. 5 depicts the ROC curve. The closer the ROC curve is to the upper left corner, the higher the overall accuracy [11]. It also includes micro-average and macro-average ROC curve. Furthermore, it shows that the area under the curve ranges from 0.99 to 1 for the different classes which are desirable for the test cases.Table II represents the confusion matrix for the CNN model. For testing the model, 500



(a). Loss vs Epochs



(b). Accuracy vs Epochs

Fig. 4. Loss and Accuracy curve for 100 epochs

TABLE II. CONFUSION MATRIX FOR THE MODEL

	Class 1 (Glioma)	Class 2 (Meningioma)	Class 3 (Pituitary)
Class 1 (Glioma)	88	12	3
Class 2 (Meningioma)	13	213	0
Class 3 (Pituitary)	0	1	170

images are used. The model identifies class 3 (pituitary tumor) more accurately than other classes. From 171 pituitary tumor images, it can successfully recognize 170 as pituitary while locating 1 as meningioma and 0 as glioma. The ability of the model to predict the other tumor class is also satisfactory.

Table III illustrates some performance measure indices for the model. It shows that precision, recall, f1-score and support are used to evaluate the model. The model achieved the precision of 88%, 94%, 98% for Class 1(Glioma), Class 2 (Meningioma) and Class 3 (Pituitary) respectively which gives the average precision of 93.33%.

A comparison is exhibited in Table IV between several methods from different papers which used the same dataset. It shows that our model gives the highest average precision of 93.33% among the different methods. The CBIR method with BoVW and REML [6] also performs well with the average precision of 93.1%.

TABLE III. PERFORMANCE MEASURE INDICES

Class	Precision	Recall	F1-score	Support
Class 1 (Glioma)	0.88	0.85	0.87	103
Class 2 (Meningioma)	0.94	0.95	0.94	226
Class 3 (Pituitary)	0.98	0.99	0.99	171

TABLE IV. COMPARISON WITH EXISTING METHODS IN TESTING PHASE

Methods	**Average Precision (%)**
CBIR [4]	89.3
CBIR with BoVW [5]	91
CBIR with BoVW and REML [6]	93.1
Proposed model using CNN	**93.33**

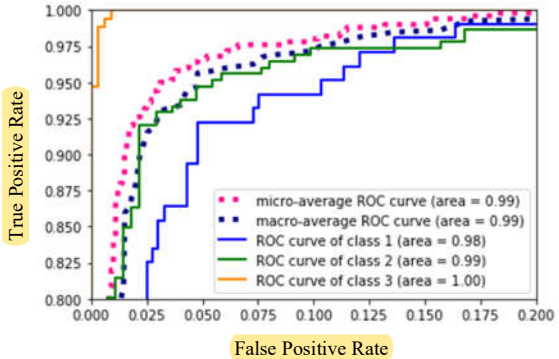


Fig. 5. ROC (Receiver Operating Characteristic) curve of the system

V. Conclusions

Brain tumor classification is very crucial in the domain of medical science. In this paper, we concentrated on developing a CNN classifier which classifies among three important tumor classes (glioma, meningioma, pituitary). Initially, the proposed system preprocesses the image data. The preprocessing includes filtering images using Gaussian filter and applying histogram equalization technique on the filtered images. Then the system classifies the images using the CNN model. The number of parameters of the model is too high, and the model is trained on a significantly small amount of data. Hence, there is a possibility of overfitting. To prevent the overfitting, a regularization technique, i.e. dropout regularization is used on the model. It helps the model to concentrate on the most prominent patterns during the training phase. Therefore, there is a better chance of generalization which keeps the model stable. The model ends up with the accuracy of 94.39% and an average precision of 93.33%. Hence, the CNN classifier will be very significant in the medical field and in saving precious lives.

References

[1] cancer.org, 'Key Statistics for Brain and Spinal Cord Tumors', January, 2019. [Online]. Available: https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/about/key-statistics.html, [Accessed: Jan. 9, 2019].

[2] Q. T. Ostrom, H. Gittleman, J. Xu, C. Kromer, Y. Wolinsky, C. Kruchko, and J. S. Barnholtz-Sloan, "Cbtrus statistical report: primary brain and other central nervous system tumors diagnosed in the united states in 2009–2013," *Neuro-oncology*, vol. 18, no. suppl5, pp. v1–v75, 2016.

[3] mayoclinic.org, 'Brain tumor'. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084, [Accessed: Dec. 15, 2018].

[4] W. Yang, Q. Feng, M. Yu, Z. Lu, Y. Gao, Y. Xu, and W. Chen, "Content-based retrieval of brain tumor in contrast-enhanced mri images using tumor margin information and learned distance metric," *Medical physics*, vol. 39,no. 11, pp. 6929–6942, 2012.

[5] M. Huang, W. Yang, M. Yu, Z. Lu, Q. Feng, and W. Chen, "Retrieval of brain tumors with region-specific bag-of-visual-words representations in contrast-enhanced mri images," *Computational and mathematical methods in medicine*, vol. 2012, 2012.

[6] M. Huang, W. Yang, Y. Wu, J. Jiang, Y. Gao, Y. Chen, Q. Feng, W. Chen, and Z. Lu, "Content-based image retrieval using spatial layout information in brain tumor t1-weighted contrast-enhanced mr images," *PloS one*, vol. 9, no. 7, p. e102754, 2014.

[7] J. Cheng, 'brain tumor dataset', 2017. [Online]. Available: https://figshare.com/articles/brain_tumor_dataset/1512427, [Accessed: Oct. 5, 2018].

[8] T. Liu, S. Fang, Y. Zhao, P. Wang, and J. Zhang, "Implementation of training convolutional neural networks," *arXiv preprint arXiv:1506.01195*, 2015.

[9] 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.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[11] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine." *Clinical chemistry*, vol. 39, no. 4, pp. 561–577, 1993.