Research Article

Detecting Brain Tumour in Early Stage Using Deep Learning

Probuddha Konwar*, Julius Bhadra, Manash Jyoti Dutta, Jintu Dowari

Department of Information Technology, School of Computing Science, Kaziranga University, Jorhat, Assam, India

ABSTRACT

In brain tumors treatment planning and quantitative evaluation, determining the tumors extent is a major challenge. Noninvasive Magnetic Resonance Imaging (MRI) has developed as a front-line diagnostic technique for brain malignancies without the use of ionizing radiation. Segmenting the extent of a brain tumor manually from 3D MRI volumes is a time-consuming process that relies greatly on operator competence. For correct tumors extent evaluation, a reliable fully automated brain tumors segmentation approach is required in this scenario. We present a completely automated brain tumors segmentation method based on U-Net deep convolutional networks in this paper. The Multimodal Brain Tumor Image Segmentation (BRATS 2015) datasets were utilized to test our approach, which included 220 high-grade brain tumors.

Keywords: Magnetic resonance imaging; Deep learning; U-Net deep; 3D MRI

INTRODUCTION

The brain, which contains billions of cells, is one of the most complex organs in the human body. When cells divide uncontrollably, they form an abnormal group of cells around or inside the brain, which is known as a brain tumor. Benign tumors are non-cancerous and therefore considered less aggressive. They start in the brain and grow slowly, and they cannot spread to other parts of the body [1]. Malignant tumors, on the other hand, are cancerous and have no defined boundaries. In both the detection and treatment phases, researchers have relied on brain Magnetic Resonance Imaging (MRI) as one of the best imaging techniques for detecting brain tumors and modeling tumors progression. Because of the high resolution of MRI images, they have a significant impact on automatic medical image analysis. Using brain MRI images as data sets and the CNN algorithm to extract and test data, we will apply the deep learning concept to perform automated brain tumors detection [2,3].

The segmented brain tumors extent can eliminate confounding structures from other brain tissues and thus provide a more accurate classification for the sub-types of brain tumors and inform the subsequent diagnosis; the accurate delineation is crucial in radiotherapy or surgical planning, from which not only the brain tumor extend but also surrounding healthy tissue has been outlined. Segmentation is still done manually by human operators in current clinical practice [4]. Manual segmentation is a time-consuming task that typically involves slice-by-slice procedures, and the results are highly reliant on the operators' knowledge and subjective decision-making. Furthermore, even when using the same operator, consistent results are difficult to achieve. A fully automatic, objective, and repeatable segmentation method is highly desired for a multimodal, multi-institutional, and longitudinal clinical trial [5].

This project's main goal is to create a cutting-edge Convolutional Neural Network (CNN) model for classifying skin lesion images into cancer types. On the dataset made available by the International Skin Imaging Collaboration, the model is trained and tested (ISIC). The model can be used to examine a lesion image and determine whether or not it is dangerous at an early stage.

For image and audio applications, such as [6-8], etc., unsupervised feature learning methods and deep learning have been widely used. These techniques have shown a lot of promise in these domains for automatically representing the feature space using unlabeled data to improve the accuracy of subsequent classification tasks. These capabilities have been extended to allow learning in very high dimensional feature spaces using additional data properties. A method to scale unsupervised feature learning and deep learning methods to high dimensional and full sized images by using image characteristics such as locality and stationary of images. Similarly, an unsupervised feature learning method (specifically Reconstruction Independent Subspace Analysis) to classify histological image signatures and tumor architecture in the context of cancer detection

[9]. Unsupervised feature learning methods have not been applied to gene expression analysis to our knowledge (it should be noted that Le's method is still used for images, not gene expression). The extremely high dimensionality of gene expression data, the lack of sufficient data samples, and the lack of global known characteristics such as locality in gene expression data are some of the reasons for this, which limit the applicability of techniques such as convolution or pooling, which have been highly successful in the above-mentioned image data applications [10,11].

We try to solve the dimensionality problem in gene expression data with the method proposed here. In our method, we first use PCA to reduce the dimensionality of the feature space, and then apply the result of PCA as a compressed feature representation that still encodes the data available in the sample set, as well as some randomly selected original gene expressions (i.e. original raw features) as a more compact feature space to either a one or a multi-layered sparse auto-encoder to find a sparse representation for data that will be used for classification (Figure 1) depicts the overall process of developing and training a system to detect and classify cancer based on gene expression data [12-15]. The approach proposed here, as shown in the figure, is divided into two parts: feature learning and classifier learning.

A file containing one or more records is referred to as a data set. A programed operating on z/OS uses records as the basic unit of information [16].

A data set is a named collection of records. Data sets can contain information such as medical records or insurance records for use by a system application. Data sets are also used to store information that is required by applications or the operating system, such as source programmers, macro libraries, and system variables or parameters [17,18]. You can print or show data sets that contain readable text on a console if they contain readable text (many data sets contain load modules or other binary data that is not really printable). Cataloging data sets is possible.

METHODOLOGY

In the last few decades, medical image analysis and classification has become a burgeoning field, with health-related applications that could aid medical experts and radiologist in diagnosing chronic diseases like chest cancer, brain tumors, diabetes prediction, and cardiac murmur evaluation [19]. A hybrid approach based on k-means clustering and a fine-tuned CNN model with synthetic data augmentation is presented in this paper. The proposed technique consists of three main stages: (a) normalizing and unbiased the intensities of MR modalities; (b) segmenting tumor area using k-means clustering; and (c) synthetic data augmentation followed by deep features extraction from ROI and classification of brain tumor into relevant benign or malignant categories using a fine-tuned CNN model. Convolution, ReLU Layer, pooling, flattening, and full connection were all used in the CNN model's training [20-23].

Dataset Brain MRI images of 153 patients, including normal and brain tumor patients, were used in this study. The images collected included brain images of 80 healthy patients after the doctor's examination and diagnosis. There are 1321 images total, with 56 for testing and 515 for train data. 73 patient tumors include 571 images, with 170 test images and 1151 train images. Patients with brain tumors account for a small percentage of the total number of patients.

86 women and 68 men, ranging in age from 8 to 66 years old, were diagnosed with the disease. 1892 images were collected from a total of 153 patients, 1666 for train data and 226 for test data. The original size of the collected images was 512512 pixels.

Some areas of fat in photographs are mistakenly identified as tumors, or tumors may not be visible to the physician; the most precise diagnosis is entirely dependent on the physician's skill. The CNN was used to detect tumors using brain images in this paper. The images gathered from imaging centers had additional
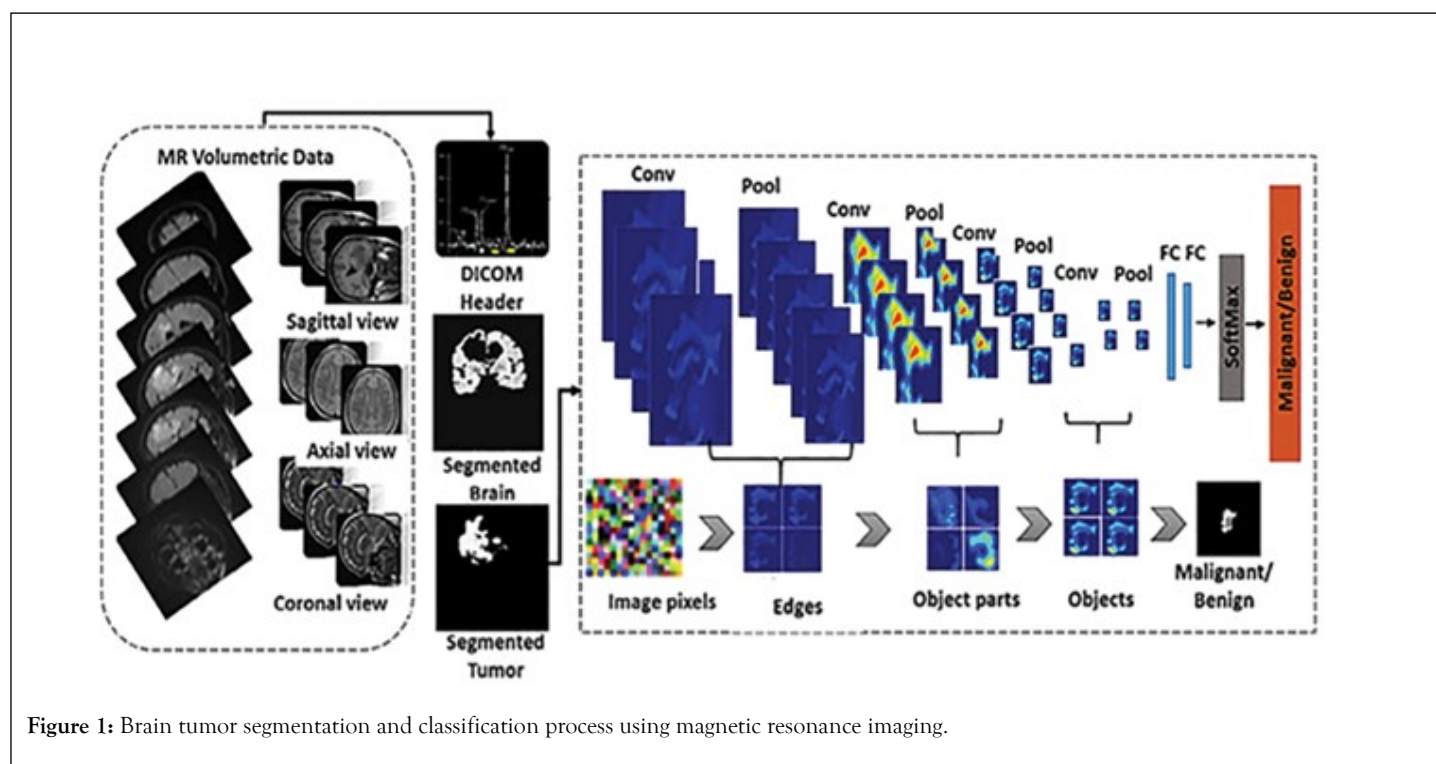

Figure 1: Brain tumor segmentation and classification process using magnetic resonance imaging.

margins. To keep the images from becoming noisy, these margins were cropped. One of the primary motivations for combining the feature extraction technique with the CNN is to retrieve image feature extraction in order to improve the network's accuracy [24]. According to the results of the CNN on the initial images, a new method combining the Clustering algorithm for feature extraction and CNN is proposed in this study to improve the network accuracy.

Method of obtaining features (C) A clustering method called central clustering. This algorithm has a duplicate procedure that iteratively tries to get points as cluster centers, which are actually the same mean points belonging to each cluster, for a fixed number of clusters. Assign each sample data set to a cluster with the shortest distance between the data and the cluster's center. The cluster centers are first chosen at random in the basic version of this method [25,26]. According to the degree of similarity, the points are assigned to the cluster centers, resulting in new clusters. The first-order clustering algorithm is used to extract features from the data in this paper. The image produced by applying the clustering algorithm to the image is shown in Figures 1 and 2.

D. Neural network convolutional initially, no feature extraction methods were used when the images were fed into the CNN. The input images are 227 227 pixels wide at first. To identify and classify the images, the Alex net architect was used, which consisted of 5 Convolutional layers and 3 layers of Sub-sampling layers, Normalization layers, Normalization layers, Fully Connected layers, and the classification layer [21]. 4096 neurons are present in the fully connected layers. This layer has two classes: brain tumor patients and healthy people.

RESULTS AND DISCUSSION

Deep learning methods were used instead of traditional machine learning in a large portion of the recently published brain tumour classification literature. For the classification of brain tumour tasks, many researchers used pre-trained CNN architectures that were fine-tuned. Abrasion-based architectures have been proposed/ tested and trained by some data scientists. The majority of the CNN architectures were 2D, but some were 3D. The main issue in the age of deep learning is the lack of large enough labeled MRI data for CNN training. Synthetic data augmentation was used to solve this problem. For brain tumour classification, proposed three convolutional auto encoders (for three MR sequences), fusion, and finally a fully linked layer. After extensive data augmentation, Sajjad, et al. fed the tumor area to a pre-trained VGG-19 CNN, which was fine-tuned for grade classification [6]. Rehman, et al. [4] proposed combining pre-trained CNN models (VGG16 VGG19) to classify multimodal automated brain tumors using linear contrast stretching, a transition of learning-based extraction functions, and correntropy-based collection [4]. Finally, to detect a brain tumor, the fused matrix was fed to an extreme learning machine (ELM). For BraTs 2015, BraTs 2017, and BraTs 2018, the proposed methodology was validated on the BraTS data sets, with precision of 97.8, 96.9, and 92.5%, respectively (Figures 3 and 4). Figure 3 depicts the model effect after the end of a successful run, whereas Figure 4 depicts the model effect throughout.

Various types of features were extracted and trained using various machine learning techniques to improve the prediction accuracy of overall survival for glioma patients in MRI. Deep feature extraction using pre-trained Alex Net and Linear Discriminant yielded the highest classification accuracy. Despite the fact that texture features are widely used in tumor classification, their accuracy is only 46%. For two classification classes, histogram features had a precision of 68.5 percent; this low precision could be due to noise. The number of coils in the imaging system determined the noise model for MRI images. In a single-coil system, the noise is represented by a Rician distribution, while in a multi-coil (parallel) imaging system, it is represented by a non-central chi distribution. The data provided was noisy, as shown in Figures 5 and 6. The single-coil system assumes that an MRI image's real and imaginary parts are uncorrelated Gaussian distributions with zero mean and equal variance. The Rician distribution is, in fact, a subset of the non-central chi distribution [20-22].
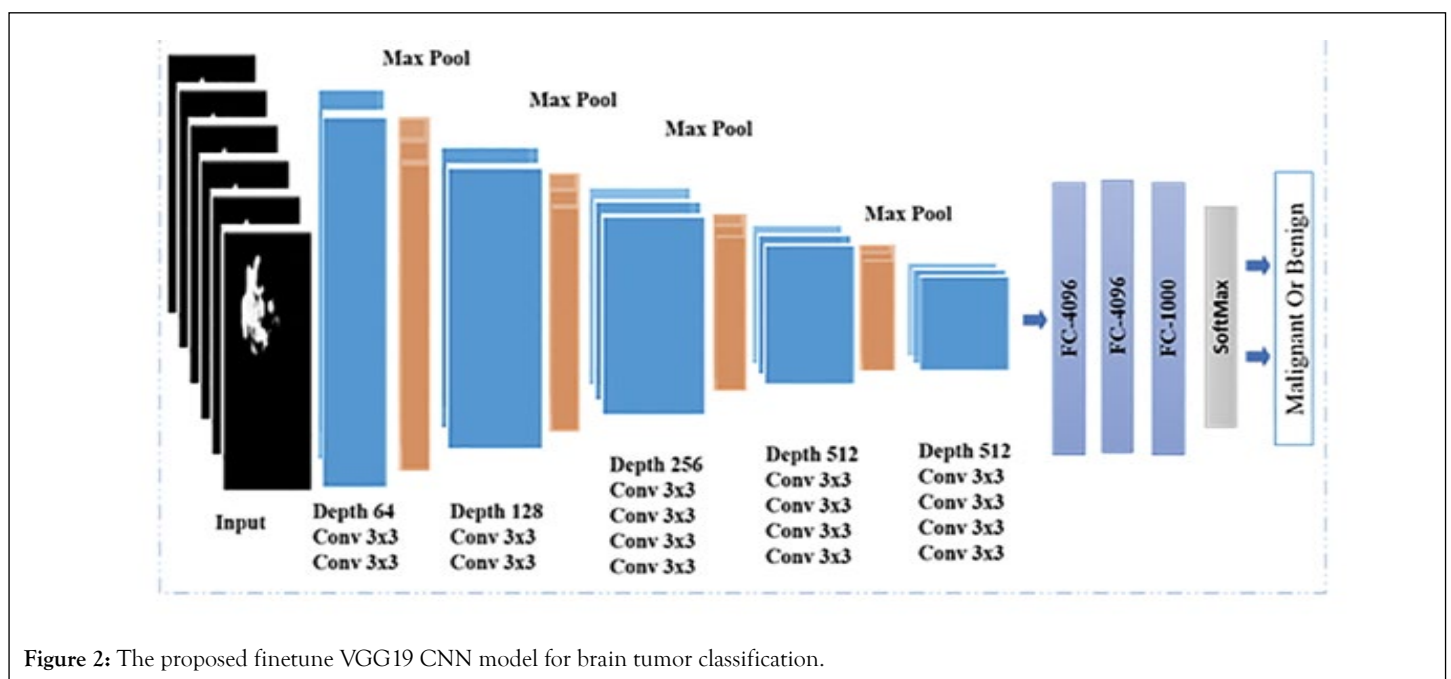

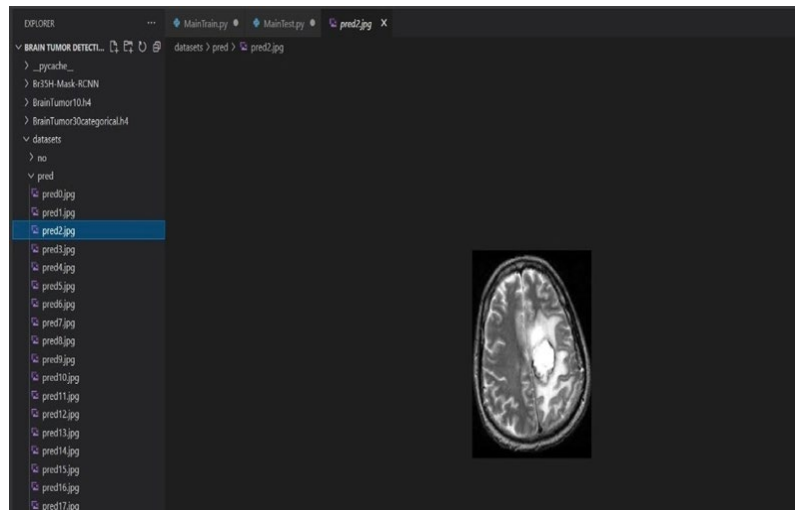**Figure 2:** The proposed finetune VGG19 CNN model for brain tumor classification.

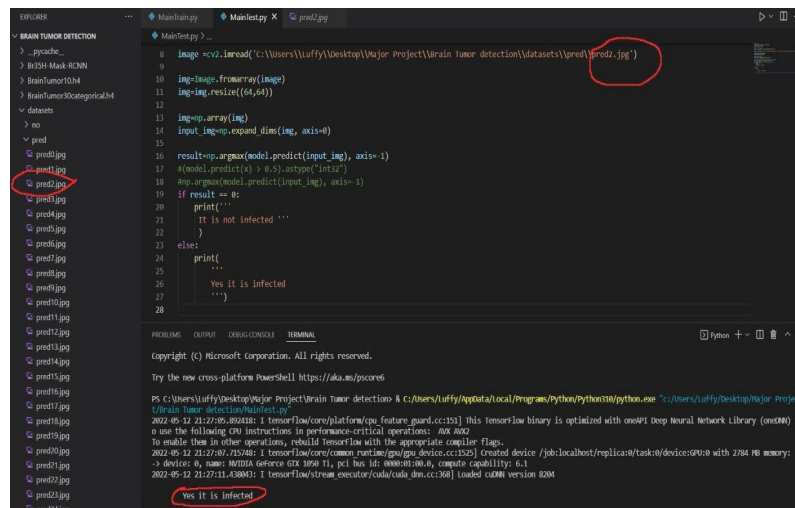**Figure 3:** Data set of infected brain.
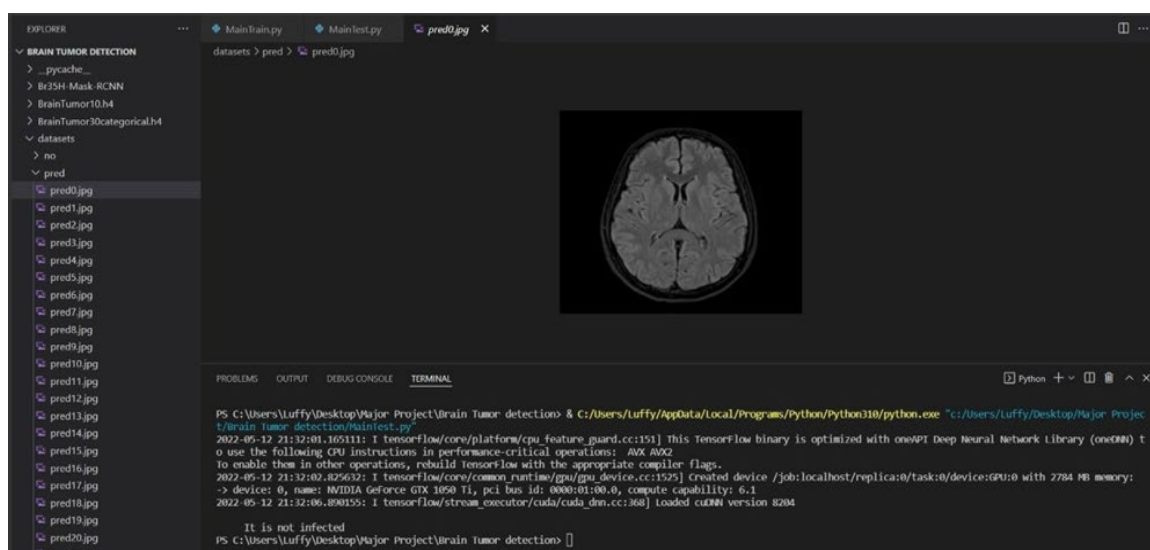

Figure 4: Model showing the result of infected.

Figure 5: Data set of infected brain and depicts the model effect after the end of a successful run.

Figure 6: Model showing the result of not infected.

CONCLUSION

DE noising the data and testing the histogram features, as well as other types of proposed features, is one recommendation for future research. In addition, texture features for each sample were extracted using specific slices that contained the three tumor regions. To improve deep feature classification accuracy, the specific patch size that contains the tumor could be used only to extract the 2D and 3D deep features, based on classification for three classes; then classification for two classes could be tested as well.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017;67(1):7-30.

2. American Brain Tumor Association statistics on brain tumors. 2019.

3. Razzak MI, Imran M, Xu G. Efficient brain tumor segmentation with multiscale two-pathway-group conventional neural networks. IEEE J. Biomed. Health Inform. 2018;23(5):1911-9.

4. Rehman A, Naz S, Razzak MI, Akram F, Imran M. A deep learning-based framework for automatic brain tumors classification using transfer learning. Circuits Syst Signal Process. 2020;39(2):757-75.

5. Deepak S, Ameer PM. Brain tumor classification using deep CNN features *via* transfer learning. Comput Biol Med. 2019;111:103345.

6. Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, Baik SW. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. J Compu Sci. 2019;30:174-82.

7. Brain Tumor Dataset, Figshare. 2019.

8. Gu Y, Lu X, Yang L, Zhang B, Yu D, Zhao Y, et al. Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Comput Biol Med. 2018;103:220-31.

9. Yousefi M, Krzyżak A, Suen CY. Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning. Comput Biol Med. 2018;96:283-93.

10. Naseer A, Rani M, Naz S, Razzak MI, Imran M, Xu G. Refining Parkinson's neurological disorder identification through deep transfer learning. Neural Comput & Applic. 2020;32(3):839-54.

11. Zuo H, Fan H, Blasch E, Ling H. Combining convolutional and recurrent neural networks for human skin detection. IEEE Signal Processing Letters. 2017;24(3):289-93.

12. Charron O, Lallement A, Jarnet D, Noblet V, Clavier JB, Meyer P. Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network. Comput Biol Med. 2018;95:43-54.

13. Shao L, Zhu F, Li X. Transfer learning for visual categorization: A survey. IEEE Trans Neural Netw Learn Syst. 2014;26(5):1019-34.

14. Zhou L, Zhang Z, Chen YC, Zhao ZY, Yin XD, Jiang HB. A deep learning-based radiomics model for differentiating benign and malignant renal tumors. Translational Oncology. 2019;12(2):292-300.

15. Deniz E, Şengür A, Kadiroğlu Z, Guo Y, Bajaj V, Budak Ü. Transfer learning based histopathologic image classification for breast cancer detection. Health Inf Sci Syst. 2018;6(1):1-7.

16. Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. In International Conference on Artificial Neural Networks. 2018:270-279.

17. Talo M, Baloglu UB, Yıldırım Ö, Acharya UR. Application of deep transfer learning for automated brain abnormality classification using MR images. Cogn Syst Res. 2019;54:176-88.

18. Afshar P, Mohammadi A, Plataniotis KN. Brain tumor type classification *via* capsule networks. IEEE International Conference on Image Processing (ICIP). 2018:3129-3133.

19. Abiwinanda N, Hanif M, Hesaputra ST, Handayani A, Mengko TR. Brain tumor classification using convolutional neural network. In world congress on medical physics and biomedical engineering 2019: 183-189.

20. Ismael MR, Abdel-Qader I. Brain tumor classification *via*

statistical features and back-propagation neural network. In2018 IEEE International Conference on Electro/Information Technology (EIT) 2018:0252-0257.

21. Pashaei A, Sajedi H, Jazayeri N. Brain tumor classification *via* convolutional neural network and extreme learning machines. International Conference on Computer and Knowledge Engineering (ICCKE) 2018: 314-319.

22. Pereira S, Meier R, Alves V, Reyes M, Silva CA. Automatic brain tumor grading from MRI data using convolutional neural networks and quality assessment. In understanding and interpreting machine learning in medical image computing applications 2018:106-114.

23. Anaraki AK, Ayati M, Kazemi F. Magnetic resonance imaging-based brain tumor grades classification and grading *via* convolutional neural networks and genetic algorithms. Biocybernetics and Biomedical Engineering. 2019;39(1):63-74.

24. Zhou Y, Li Z, Zhu H, Chen C, Gao M, Xu K, et al. Holistic brain tumor screening and classification based on densenet and recurrent neural network. In International MICCAI Brainlesion Workshop. 2018:208-217.

25. Ahammed Muneer KV, Rajendran VR, Joseph PK. Glioma tumor grade identification using artificial intelligent techniques. J Med Syst. 2019;43(5):1-2

26. Banerjee S, Mitra S, Masulli F, Rovetta S. Deep radiomics for brain tumor detection and classification from multi-sequence MRI. arXiv preprint arXiv:1903.09240. 2019.