

Brain Tumor Segmentation and Classification using CNN model

Srujan K S¹, Shakthy Shivakumar², Kalifkhan Sitnur³, Omkar Garde⁴, Poornima PK⁵

^{1,2,3,4}Student, Dept. of Information Science, R.V. College of Engineering, Bangalore, Karnataka, India

⁵Assistant Professor, Dept. of Information Science, R.V. College of Engineering, Bangalore, Karnataka, India

Abstract - Brain tumor segmentation and classification is a critical task for assessing tumors and making a choice of treatment by grade. Many machine learning algorithms and segmentation techniques were used and tested for the tumor classification purpose. Satisfiable accuracy was not obtained by the above mentioned techniques. With the help of Deep Learning, many classification and segmentation problems which could not achieve good accuracy were solved. In this paper, using two publicly available datasets, a Deep Learning system based on a convolutional neural network is implemented to classify different types of brain tumors. Tumors; meningioma, glioma and pituitary tumor are included in the former classification. The other differentiates between the three grades of glioma; Grade II, Grade III, and Grade IV. The proposed convolutional neural network consists of 16 layers. Activation functions such as ReLU and optimizers like Adam optimizer are used in the neural network. The findings suggest the model's potential for multi-classification purposes of brain tumor. An accuracy of 95.36% is obtained for the proposed model. The system will also determine an estimate of the patient's survival time in months, based on the results of the brain tumor grade and level.

Key Words: Brain tumor, Segmentation, Convolutional neural network, Magnetic resonance imaging, Deep Learning

1. INTRODUCTION

Unusual development of tissue in the human brain that frustrates the best possible working of the brain is known as the brain tumor. Meningiomas and gliomas are the common type of brain tumors that occur in adults. 81% of the malignant brain tumor cells in grown-ups are gliomas. Detection and analysis of brain tumor is viewed as one of the troublesome assignments in the clinical field. Clinical imaging systems are for the most part utilized in recognizing even the smallest variation from the norm in the human tissues. Computed Tomography scans (CT scans), X-rays and Magnetic Resonance Imaging (MRI) are some the best known medical imaging techniques. Such techniques allow us to recognize even the slightest defects in the human body. The main purpose of the analysis of medical images is to distinguish correct and relevant details using pictures with the least possible error. MRI is primarily used to acquire images of the human body and cancerous tissues as opposed to other imaging techniques, owing to its high resolution and clearer images. Identification of brain tumors by MRI images is a difficult

task considering the anatomy of the brain. Tumor cells are the group of cells in the central nervous system that arises after the brain cells have been differentiated. The cells inside an adult's central nervous system will trans differentiate to other cell types even under normal circumstances. The differentiated cells may be either stable or a tumor cell. It is discovered from experiments that undifferentiated cells are more susceptible after differentiation to become tumor cells. These results in small, differentiating molecules becoming a fair component of tumor therapy. MRI brain image cancer classification was first investigated on the basis of Rough Set Theory and Feed Forward Neural Network. The feature extraction is done using Rough Set theory from the sample MRI images and are then given as a sample for the Feed Forward Neural Network Classifier. Eventually, the Feed Forward Neural Network classifier is used for the execution of two tasks. The first is that they differentiate between normal and abnormal. The second task is to identify the abnormality as either benign or malignant tumor. Brain Tumor Image Segmentation Benchmark, an experiment done to identify the suitability of algorithm for tumor classification remains as a continuous benchmarking tool till date. 20 brain tumor segmentation algorithms are shortlisted. 65 high grade and low grade MR scan images are obtained. The selected algorithms are made to run on the data obtained and results are analyzed. Specific algorithms have been shown to function well on different areas of the brain. There was no standard model for all regions of the brain. The analysis concluded that multiple good algorithms are mixed to achieve a positive outcome instead of a single algorithm. Later in the studies, some machine learning methods have been used for the classification of patients affected by the tumor. In this analysis the dataset used is the same as in BraTS 2017. Specific types of characteristics have been identified and trained using various ML methods. The characteristics included volumetric, statistical and strength texture, histograms and depth characteristics; ML techniques included support vector machines (SVM), k nearest neighboring (KNNs), linear discriminant, tree, ensemble techniques and logical regression. Accuracy and precise results were not obtained using above ML methods. To tackle the above short-comings we have developed a system based on deep learning methods to classify brain tumor MR images. The rest of the paper includes methodology used to develop the model, results obtained in the experiments and conclusion.

2. METHODOLOGY

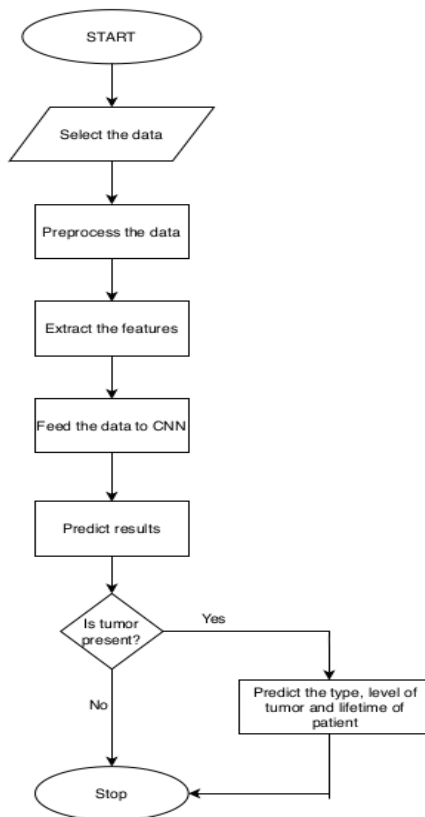


Figure 1: FLOWCHART

Flowchart is the pictorial representation which shows the sequential workflow of the model being developed.

The workflow first starts with the collection of the data. Data may contain the inaccurate and inconsistent attributes in it which may result in the wrong predictions. So preprocessing of the data remains an utmost important task. Next step is to preprocess the data taken as input. The input data which was preprocessed is fed to the convolutional neural network explained above. The outcome of this step is the predicted results. In the obtained result, an assertion is made to check whether the tumor is present or not. It is represented by the diamond symbol in Figure 1. If it is found that a tumor is present, further the workflow continues. In the next step, prediction is made to guess the level and type of the tumor present in the image. Lifetime of the patient is also found out in this step and the system stops. If the tumor is not present, the system stops with a message saying the image is normal.

2.1 Collecting data

The dataset gathered incorporates MRI pictures of patients gathered far and wide. Around 3000 pictures from 230 patients are gathered from The Cancer Imaging Archive, which was distributed and overhauled in 2017.

The dataset has T1-weighted complexity pictures of cerebrum tumor patients named with Three classes of tumor; meningioma, glioma, and pituitary tumor. Three unique perspectives in particular hub, coronal, and sagittal perspectives on the pictures are available.

2.2 Preprocessing data

Data processing is a crucial step until the deep learning model feeds the data. Preprocessing removes inconsistencies and inaccuracies in the data collected so that the neural network is adequately trained. The data that is obtained is in the format of .mat. The images collected are translated to the .jpg format using python libraries to enable further study and processing. The data is resized to smaller resolutions to reduce the processing time. Mixing up of the data is done so that the model does not get trained to a narrow range of the data. Better results are obtained when we train the network on unsorted data. Data augmentation is performed to increase the count of images. Noise is added to the original image. Upon augmentation and salting the number of photographs is increased by a factor of 5. Thus, a sum of 15,000 images are available for neural network training and testing.

2.3 Feeding the data to CNN model

The images preprocessed move through a 16-layer CNN model. Images of 128x128 pixels in resolution are fed to the input layer. This then goes through 3 layers of convolution, a fully connected layer, a layer of dropout, a layer of softmax and a layer of classification. Multiple features are identified that are useful in classifying MRI images into unique tumor types. The convolution layer consists of a function for activation, a function for normalization and a max pooling layer. Within the convolution layer different kernel functions and strides are used which help extract different features. Reduction of the dimensionality is achieved in the layers of max pooling and activation. The classification layer uses categorical cross entropy loss function to measure error. Layer wise explanation of the CNN model is given below:

Layer 1: Input layer which takes preprocessed images as input

Layer 2: Convolution layer with 64 filters, size 3x3, stride 1

Layer 3: Rectified Linear Unit(ReLU) layer which removes negative activations

Layer 4: Normalization layer which scales the activations

Layer 5: Max Pooling layer which does downsampling and reduces the computation in the network

Layer 6: Convolution layer with 64 filter, size 3x3, stride 1

Layer 7: Rectified Linear Unit(ReLU) layer

Layer 8: Max pooling layer

Layer 9: Dropout layer randomly removes some activations to avoid overfitting

Layer 10: Convolution layer with 64 filter, size 2x2, stride 1

Layer 11: Rectified Linear Unit(ReLU) layer

Layer 12: Max pooling layer

Layer 13: Dropout layer

Layer 14: Fully connected layer which collects the results from all the neurons

Layer 15: Softmax layer converts the obtained results into probabilities by fitting the values between 0 and 1

Layer 16: Classification layer returns the class of the tumor to which the image belongs

3. RESULTS

Preprocessing module:

The preprocessing module preprocesses the MRI scan images which are of .mat format and converting them into .jpg format that can be used by the CNN model that we have built.

CNN model building and training module:

The output of the CNN model building and training module is to produce a model that will be able to predict and identify the class; Meningioma, glioma and pituitary and the grade of the tumor.

Prediction module:

The output of the prediction module is to predict and identify the class; Meningioma, glioma and pituitary and the grade of the tumor.

Survival time module:

The output of the lifetime module is to predict the time the patient will survive based on the identified class and the grade of the tumor.

The experiments are carried out in a machine with Intel core-i7 processor. CPU used is of 2.71 GHz frequency. The 64 bit-Windows operating system is used for the experiment.. The RAM capacity of the system is 16 GB DDR4. Anaconda virtual environment is used as the operating environment for the experiment. Python version 3.7 is used for the development purpose.

Different combinations of the hyper parameters are used to check where the maximum accuracy occurs. Maximum accuracy is achieved by the model when 74% of the input data is for training the model and 26% of the remaining data is used for testing and validation purposes. Out of 3064 images obtained, 2267 images are used for training the model and 797 images are used for testing images. Optimizer used for training is Adam optimizer. Categorical cross entropy is the loss function used when the model achieves maximum accuracy. A total of 35 epochs are used for training the model. After the 1st epoch, an accuracy of 56.67% is obtained. After the 10th epoch, 82.12%

accuracy is obtained. Finally after the 35th epoch, the highest accuracy of 95.36% is obtained. Table 2 shows the accuracy obtained at different epochs of the training.

EPOCH	ACCURACY(%)
1	56.67
5	77.10
10	82.12
20	89.98
30	93.98
35	95.36

TABLE 1: EPOCH VS ACCURACY TABLE

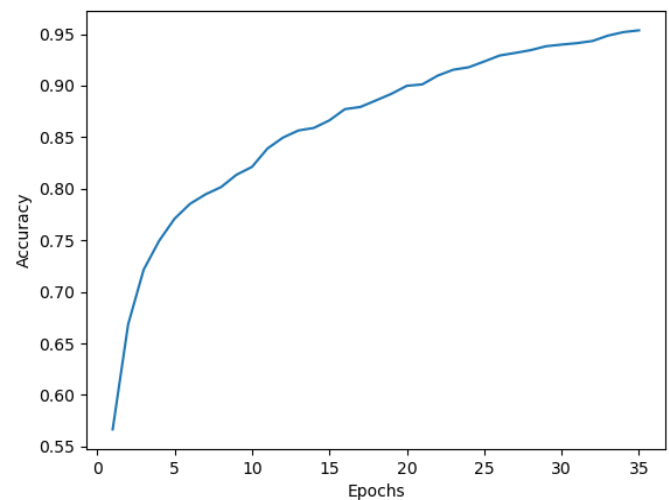


FIGURE 2: GRAPH OF EPOCH VS ACCURACY

3. CONCLUSIONS

This project was initiated because of the shortcomings of the current programme. Such limitations have been converted into new criteria, and to address existing device limitations a new project "Brain tumor segmentation and classification using CNN model" has been developed and implemented. Segmentation of the brain tumor and classification resolution will be an updated version, using the CNN model. It would also help the health care and medical research centre also save time and money by following the plan. This project also intends to calculate the surviving time of the individual based on the class and the grade of the tumor.

LIMITATIONS

Although it will give better performance than the existing system using the CNN model approach. No CNN model can give an accuracy of 100 per cent. The actual model's

accuracy is limited by how clear the MRI data is one of the limitations is that it takes more time. The size of the dataset is very less. The unavailability of the images in larger numbers is one of the limitations of testing the model.

REFERENCES

- [1] Lapointe S, Perry A, Butowski NA. "Primary brain tumors in adults. *Lancet*". 2018 Aug 4; doi: 10.1016/S0140-6736(18)30990-5
- [2] T Rajesh, R. Suja Mani Malar, "Rough set theory and feed forward neural network based brain tumor detection in magnetic resonance images", *IEEE Xplore*, 26 July 2013, 10.1109/ICANMEET.2013.6609287
- [3] Ketan Machhale, Hari Babu Nandpuru, Vivek Kapur, Laxmi Kosta, "MRI brain cancer classification using hybrid classifier (SVM-KNN)", *IEEE*, 09 July 2015, 10.1109/IIC.2015.7150592
- [4] M. Shasidhar V Sudheer Raja, B. Vijay Kumar, "MRI brain image segmentation using modified fuzzy C-means clustering algorithm", *IEEE*, 29 July 2016, DOI: 10.1109/CSNT.2016.102
- [5] Suchita Goswami, Lalit Kumar P. "Brain tumor detection using unsupervised learning based neural network", *IEEE*, 8 April 2016, DOI: 10.1109/CSNT.2016.123
- [6] Samina Khalid, Tehmina Khalil, Shamila Nasreen "A survey of feature selection and feature extraction techniques in machine learning", 09 October 2014 , DOI: 10.1109/SAI.2014.6918213
- [7] Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme", December 2017.
- [8] El-Sayed Ahmed, El-Dahshan Tamer, Hosny, Abdel-Badeeh M.Salem, "Hybrid intelligent techniques for MRI brain images classification", 9 July 2015
- [9] Cheng J, Huang W, Cao S, Yang R, Yang W, et al. (2015) "Enhanced performance of brain tumor classification via tumor region augmentation and partition", 9 Dec 2016
- [10] Ertosun MG, Rubin DL, "Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks", % November 2017
- [11] L. M. De Angelis, "Brain tumors," *New England J. Med.*, vol. 344, no. 2, pp. 114–123, Jan. 2001.
- [12] B. W. Stewart and C. P. Wild, *World Cancer Report 2014*. Lyon, France: IARC, 2014.
- [13] Brain, Other CNS and Intracranial tumors Statistics. Accessed: May 2019. [Online]. Available: <https://www.cancerresearchuk.org/>
- [14] A. DrevelegG. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
- [15] M.L.Goodenberger and R.B.Jenkins, "Genetics of adult glioma," *Cancer Genet.*, vol. 205, no. 12, pp. 613–621, Dec. 2012.
- [16] C. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer-Verlag, 2006.
D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 World Health Organization classification of tumors of the central nervous system: A summary," *Acta Neuropathol.*, vol. 131, no. 6, pp. 803–820, Jun. 2016.