Identification and Classification of Brain Tumor

Deepthi Yadav G, Manisha L

Guide- Shwetha B N

Department of Computer Science and Engineering

Sapthagiri College of Engineering, #14/5 Chikkasandra,

Hesaraghatta main road, Bangalore-57 India

[1]deepthiyadav112@gmail.com

[2]manishalvaibhav76@gmail.com

[3]shwethabn@sapthagiri.edu.in

Abstract- **A brain tumor is a disease caused due to the growth of abnormal cells in the brain. There are two main categories of brain tumor, they are non-cancerous (benign) brain tumor and cancerous(malignant) brain tumor. Survival rate of a tumor prone patient is difficult to predict because brain tumor is uncommon and are different types. Treatment for brain tumor depends on various factors like: the type of tumor, how abnormal the cells are and where it is in the brain etc. With the growth of Artificial Intelligence, Deep learning models are used to diagnose the brain tumor by taking the images of magnetic resonance imaging (MRI). This project detects whether the brain tumor is present or not from MRI scans. It uses the CNN classification technique which is used to classify the type of brain tumor present. If the detected tumor is non-cancerous(benign), the further classification of the type of tumor is done. The implementation of the suggested model is applied in the Python and TensorFlow environment. Algorithms and methodologies used to solve specific research problems are included in the results and along with their strengths and limitations. This examines the quantitative characteristics of brain tumors, such as shape, texture, and signal intensity, to predict high accuracy with a low error rate.**

I. Introduction

The brain contains billions of cells and with lots of complex functions happening every second. When the brain is damaged, there will be many things affected. Unusual development of tissue in the human brain that frustrates the best possible working of the brain is known as the brain tumor. Because of the fact that it will boost the outcome of study in a space where ill-advised bad situations must be at a low point, computer assistance is strongly advanced in therapeutic organisations. The cost of two-dimensional MRI readings is exorbitant, but excellent programming to assist people in clinical settings is of great interest these days.

Human onlookers describe the tumor's highlights of traditional techniques. A mechanised symptomatic structure with certain anatomical highlights has been implemented to increase the accuracy of the present framework. To improve accuracy, a symptomatic system has been developed.[7] Manual segmentation takes sensible time and precision will be less and it prompts bury and intra rater blunders. Programmed segmentation is required because of this reason

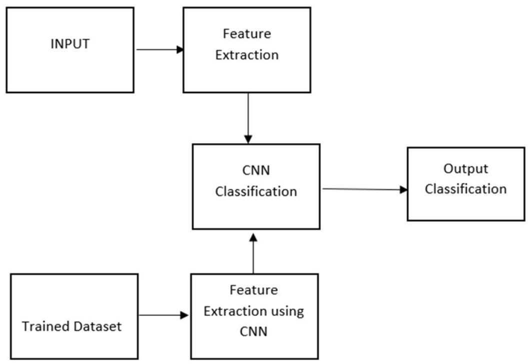
[5] Programmed segmentation gives the data about the encompassing tissues around the tumor. This is because of the intensity variation happening among the same groupings. Segmentation in MRI utilized in treatment checking and gaining up prominence with progress in picture guided surgical methodologies. Laying out of tumor contours is an essential advance. This technique depends on CNN and learn highlights that are explicit to gliomas detection and segmentation.

II. Methodology

In this project, we have augmented the dataset (MRI images of brain), performed certain data pre-processing steps to convert the raw data, further investigated two deep learning models namely CNN and VGG-16 and have presented the comparative analysis in the results section.

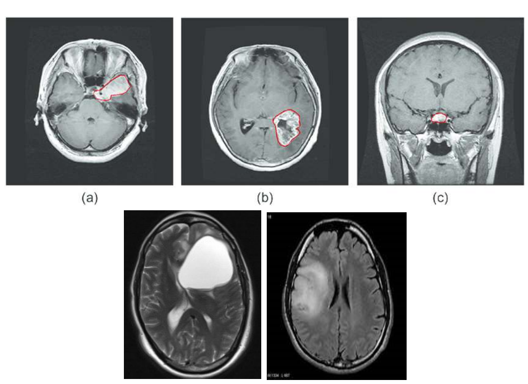
The main stages involved in this method are Data Collection, Pre-processing, Feature extraction, prediction model and evaluation.

A. Block Diagram



B. Gather the Dataset

The first component of building a deep learning network is to gather our initial dataset. The images themselves as well as the labels associated with each image are required. These labels should come from a finite set of categories, such as: glioma tumor, meningioma tumor, pituitary tumor and no tumor. Furthermore, the number of images for each category should be approximately uniform (i.e., the same number of examples per category) then the classifier will become naturally biased to overfitting into these heavily-represented categories. Class imbalance is a common problem in machine learning and there exist a number of ways to overcome it. As our system is mainly focusing on detection of brain tumor, we gathered our data as MRI images. The dataset used consists of several images in which tumor is present and in several images in which tumor is not present.



(d) (e)

Figure 1. (a) Glioma (b)Meningioma (c) Pituitary (d) Glioblastoma Multiforme (e) Oligodendroglioma Tumors

C. Image pre-processing

The primary target is to improve image highlights needed for additional processing.
The following techniques are used in the pre-processing step.

1. Greyscale Image Conversion

Grayscale brain tumor MRI image is cropped by applying vertical horizontal profile. Image cropping is performed first to improve image enhancement, otherwise background region affects intensity enhancement. In a grayscale image, each pixel is a scalar value between 0 and 255, where zero corresponds to “black” and 255 being “white”. Values between 0 and 255 are varying shades of gray, where values closer to 0 are darker and values closer to 255 are lighter.

2. Scaling and Aspect Ratios

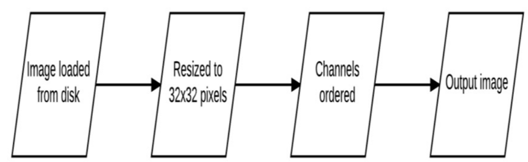


Figure 2. Image pre-processing pipeline that (1) loads an image from disk, (2) resizes it to 32_32 pixels, (3) orders the channel dimensions, and (4) outputs the image.

Scaling, or simply resizing, is the process of increasing or decreasing the size of an image in terms of width and height. When resizing an image, it’s important to keep in mind the aspect ratio

3. Histogram equalisation

histogram equalization normalizes the pixel intensities, thereby normalizing some of the illumination problems. The technique does a great job in normalizing the pixel intensities of the brain MRI image . The image belongs to the non-tumor class but the whiter patch on the top left part of the brain looks like a tumor although it is not. Histogram equalization reduced the visual difference between that area and the rest of the brain. This was an encouraging result and we can proceed to fit a Convolutional Neural Network (CNN) model on histogram equalized images.

4. Adaptive Thresholding

Thresholding is used to segment an image by setting all pixels whose intensity values are above a threshold to a foreground value and all the remaining pixels to a background value. Adaptive thresholding typically takes a grayscale or color image as input and, in the simplest implementation, outputs a binary image representing the segmentation. For each pixel in the image, a threshold has to be calculated. If the pixel value is below the threshold it is set to the background value, otherwise it assumes the foreground value.

5. Erosion and Dilation

Erosion is done using following steps:

A kernel (a matrix of odd size (3,5,7)) is convolved with the image. Then a pixel in the original image (either 1 or 0) will be considered 1 only if all the pixels under the kernel are 1, otherwise, it is eroded (made to zero). Thus all the pixels near the boundary will be discarded depending upon the size of the kernel.So the thickness or size of the foreground object decreases or simply the white region decreases in the image.

Dilation is done using following steps:

A kernel (a matrix of odd size (3,5,7) is convolved with the image. A pixel element in the original image is ‘1’ if at least one pixel under the kernel is ‘1’. It increases the white region in the image or the size of the foreground object increases.

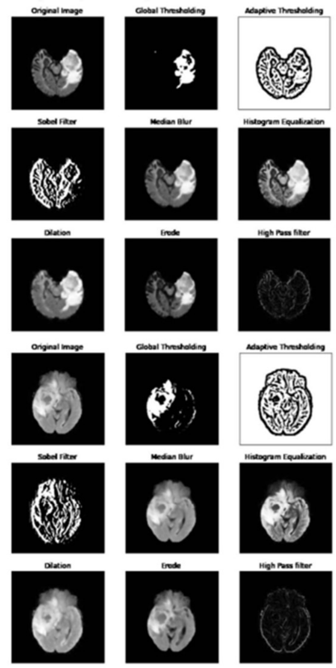


Figure 3. Image pre-processing employed MRI scans

D. Splitting the dataset

The dataset is split into training, testing and validation dataset**.** A training set is used by the classifier to learn what each category looks like by making predictions on the input data and then correct itself when predictions are wrong. After the classifier has been trained, we can evaluate the performing on a testing set.

It is important that the training set and testing set are independent of each other.

Neural networks have a number of knobs and levers (ex., learning rate, decay, regularization, etc.) that need to be tuned to obtain optimal performance. These are the called as the types of parameters hyperparameters.

Validation set is created because the test set is only used in evaluating the performance of your network. This set of the data (normally) comes from the training data and is used as “fake test data” so we can tune the hyperparameters. Only after determining the hyperparameter values using the validation set, we move on to collecting final accuracy results in the testing data.

E. Training

The training dataset is the one on which CNN is trained while after each epoch or iteration, the learned model till that iteration can be tested on a validation set. The validation set becomes a type of test data as the model was not trained on that; thereby becoming unseen data. CNN is utilized to get better result. The signal convolved with kernels to get include map. Past layers are interconnected with weights of the kernel. To upgrade the qualities of information image by back propagation calculation. Since feature maps of all units shared by the kernels. It will serve to reduces over fitting. Each data of neighbourhood is taken by utilizing kernels. Kernel is a major source of context information. Activation function is applied to the output of neural network

1. Convolutional Layers

Objective of the convolution layer is to take or extract the features from the input [image], just the part of picture is link to the following convolution layer.

2. Padding

Padding is incorporating a zero layer outside the input volume so the data on border won't be lost and we can get a similar dimension of output as input volume. Here we are using zero padding.

3. Activation Function

Non- linear activation function ReLU (Rectifier Activation function) is used to provide accurate results than classical sigmoid function.

4. Pooling Layer

It is used for combining spatially nearby features. Max-pooling is generally used to join features. It decreases the dimension of input image and controls over-fitting.

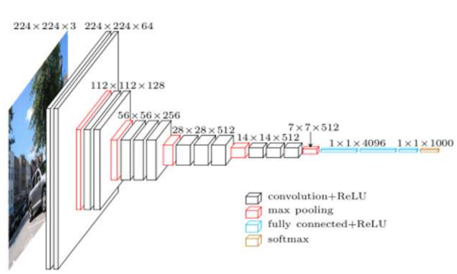


Figure 4. Visualization of VGG architecture

Each layer in a CNN applies a different set of filters, typically hundreds or thousands of them, and combines the results, feeding the output into the next layer in the network. During training, a

CNN automatically learns the values for these filters.

In the context of image classification, our CNN may learn to:

1. Detect edges from raw pixel data in the first layer.
2. Use these edges to detect shapes (i.e., “blobs”) in the second layer.
3. Use these shapes to detect higher-level features such as facial structures, parts of a car, etc.in the highest layers of the network.

F. Testing

Software testing is an investigation conducted to provide the information about the quality of the product or service under test. Software testing can also provide an objective, independent view of the software to allow the business to appreciate and understand the risks of software implementation. Test techniques include the process of executing a program or application with the intent of finding software bugs .

The testing steps are: Unit Testing ,Integration Testing, Validation Testing ,User Acceptance Testing ,Output Testing.

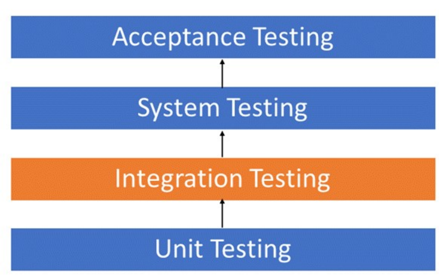


Figure 5. Steps of testing process

G. Conclusion

Brain tumor is comparatively infrequent i.e., 1.4% new cases per year, in developed countries. Fatalities due to brain tumors have increased over the past few decades. Hence, the scope of this project is widened. Brain tumors especially the malignant ones are considered almost incurable and fatal. The need for early detection arises from the fact that brain tumors can have symptoms that do not seem to be alarming at first. To identify the type of brain tumor, present a risky medical procedure Biopsy is done. This project proposes the method to detect and identify the type of brain tumor present based on the MRI (Magnetic Resonance Imaging) scans of the patient fed to the system. It also provides the precise identification of the brain tumor which is present based on CNN classification algorithm. Biopsy can be majorly avoided due to the accurate output of the system. This is considered as the low-cost method for solving the diagnosis problem of brain tumor. The previous systems have used various algorithm and methodologies to prove the presence of brain tumor by machine learning. Our system provides a clear picture of the precise tumor present in the patient by considering various factors such as location, size etc. We have referred more than 15 papers to know the need for this system. This method proves to be the most accurate and efficient method to diagnose the problem of brain tumor detection.

H. References

[1] Sobhaninia, Zahra, et al. "Brain tumor segmentation using deep learning by type specific sorting of images." arXiv preprint arXiv:1809.07786 (2018).

[2] Abhishek Anil, Aditya Raj, H Aravind Sarma, Naveen Chandran R, Deepa P L, "Brain Tumor detection from brain MRI using Deep Learning" International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE), Volume 3, Issue 2, DOI: 10.29027/IJIRASE.v3.i2.2019, 458-465, August 2019.

[3] Mrs. Shinde Apurva Swapnil, Ms. Vengurlekar Samidha Girish "Image Mining Methodology for Detection of Brain Tumor: A Review" Proceedings of the Fourth International Conference on Computing Methodologies and Communication (ICCMC 2020) IEEE Xplore Part Number:CFP20K25-ART; ISBN:978-1-7281-4889-2

[4] Mircea Gurbin * , Mihaela Lascu † , and Dan Lascu ‡ "Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines" 2019 42nd International Conference on Telecommunications and Signal Processing (TSP)

[5] Daisuke Hirahara "Preliminary assessment for the development of CADe system for brain tumor in MRI images utilizing transfer learning in Xception model" 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)

[6] Tamjid Imtiaz; Shahriar Rifat; Shaikh Anowarul Fattah "Automated Brain Tumor Segmentation from MRI Data Based on Local Region Analysis" 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON)

[7] Sobhaninia, Zahra, et al. "Brain tumor segmentation using deep learning by type specific sorting of images." arXiv preprint arXiv:1809.07786 (2018).

[8] Abhishek Anil, Aditya Raj, H Aravind Sarma, Naveen Chandran R, Deepa P L, "Brain Tumor detection from brain MRI using Deep Learning" International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE), Volume 3, Issue 2, DOI: 10.29027/IJIRASE.v3.i2.2019, 458-465, August 2019.

[9] Mrs. Shinde Apurva Swapnil, Ms. Vengurlekar Samidha Girish "Image Mining Methodology for Detection of Brain Tumor: A Review" Proceedings of the Fourth International Conference on Computing Methodologies and Communication (ICCMC 2020) IEEE Xplore Part Number:CFP20K25-ART; ISBN:978-1-7281-4889-2

[10] Mircea Gurbin * , Mihaela Lascu † , and Dan Lascu ‡ "Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines" 2019 42nd International Conference on Telecommunications and Signal Processing (TSP)

[11] Daisuke Hirahara "Preliminary assessment for the development of CADe system for brain tumor in MRI images utilizing transfer learning in Xception model" 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)

[12] Tamjid Imtiaz; Shahriar Rifat; Shaikh Anowarul Fattah "Automated Brain Tumor Segmentation from MRI Data Based on Local Region Analysis" 2019 IEEE International Conference on Biomedical Engineering.