How to Cite:
Mangla, R., & Chetna, C. (2022). Brain tumor detection and classification by MRI images using deep learning techniques. *International Journal of Health Sciences*, *6*(S3), 5783–5790. https://doi.org/10.53730/ijhs.v6nS3.7233

Brain tumor detection and classification by MRI images using deep learning techniques

Rashika Mangla
Department of Applied Mathematics, Delhi Technological University, New Delhi, India
*Corresponding author email: rashika123mangla@gmail.com

Chetna
Department of Applied Mathematics, Delhi Technological University, New Delhi, India
Email: chetnamor4family@gmail.com

***Abstract*---**MRI images play a significant influence in brain tumor classification and detection but instead of having detection and classification using the medical equipment which is a radiologists or clinical professionals do a time-consuming and laborious task where accuracy depends only on the experience only, it can be beneficial to detect and classify the brain tumor by deep learning techniques and algorithms. As a result, the employment of computer-assisted technology becomes increasingly important in order to overcome these constraints. In this paper, the early detected and diagnosed brain tumor images along with their csv data has been used to find out the accuracy of the CNN algorithm for tumor detection and SVM algorithm for tumor classification into benign and malignant. HOG has been used for the feature extraction. After performing the experiment, it was observed that CNN achieved the detection accuracy of 87.02% and further tumor classification by employing SVM, the highest accuracy achieved was 96.35%. The experiment proved a very good accuracy of detection and classification even after using three different methods in the whole procedure.

***Keywords*---**CNN (Convolutional neural network), HOG (Histogram of Oriented Gradients), SVM (Support Vector Machine), Magnetic Resonance Imaging (MRI).

Introduction

A brain tumor is an abnormal growth in the brain that can be harmful or even cancer-free. Uncontrolled cell proliferation and excessive brain damage cause

Manuscript submitted: 9 March 2022, Manuscript revised: 18 April 2022, Accepted for publication: 1 May 2022

tumors in the brain. Primary and secondary tumors are two types of brain tumors that result in malignant or benign brain tumors. Gliomas are the most widely used sort of tumor. In order to collect relevant clinical data, such as the presence of a tumor, location, and type, computer-assisted devices can be used to perform the procedure automatically. However, determining their composition, volume, parameters, plant detection, size and classification remains a difficult task. Because it provides high visual acuity to soft tissues and is non-invasive, magnetic resonance imaging (MRI) [8] is preferred over other therapies and diagnostic methods. Signs and symptoms of a tumor in the brain vary depending on the type and location of the tumor. Because different parts of the brain control different aspects of body processes, some plants do not show symptoms until they are extremely large and cause severe and rapid loss of health. Hearing problems, problems with balance, speech impairment, changes in vision, memory problems, mobility problems, personality changes, lack of concentration, and a weak point in one part of the body are some of the most common symptoms.

Gliomas are divided into three categories, including Oligodendroglioma, Astrocytoma, and Glioblastoma, regardless of the WHO classification or classification of tumor. Due to the remarkable diversity and structure of MRI data, Brain Tumor is a difficult task. Sound and image clarity are limits on MRI imaging. Unwanted information in pictures is called sound. Noise can sometimes affect the edges and details, reducing the brightness adjustment. Because it is difficult to find specific boundaries and to distinguish a tumor from noise, it is an emerging topic of image processing research. In this study, we compare and contrast an effective and efficient method of brain differentiation, detection, and differentiation. Many brain tumor classification techniques have recently been introduced, which may be categorized by machine learning and deep learning (DL) techniques such as CNN, GLCM, OTSU'S THRESHOLDING, SVM, and others, which we will use based on feature selection and learning methods.

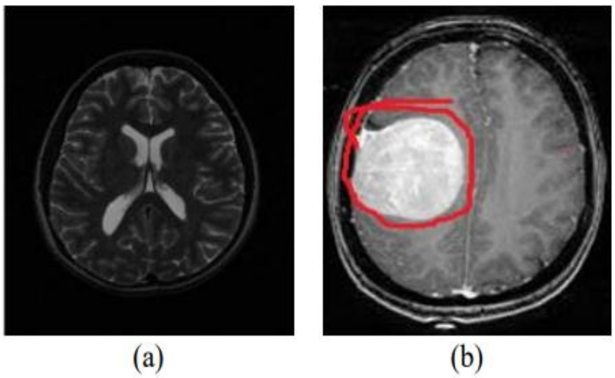

Fig 1. A healthy brain (a) and a tumor-affected brain (b) are depicted in these photos of a human brain.

Literature review

Machine learning techniques have been widely used in a variety of fields, including medical diagnosis and prevention. However, few studies have focused on the diagnosis of brain tissue using magnetic resonance imaging (MRI). Most DL methods use MRI data to train and test old DL algorithms. DL has recently been used to diagnose brain tumors in a number of techniques. G. Hemanth et al. [1] suggested Automated Brain Tumors Automatic Detection System. CNN is used for classifying and identifying tumors. Then a pixel-based detection method is used to obtain a picture of the brain and other affected areas. Then the separation using the extracted features is done using CNN. They found 91% accuracy in sections. Praveena Pedapati and Rama Vaishnavi Tanneedi [2] propose to use SVM and in-depth study to detect and differentiate brain cancer automatically. However, in this case, SVM failed to provide the highest level of accuracy. The accuracy of the features extracted from high glioma (Malignant) images without the use of histogram of the targeted gradients was 97 percent in high glioma images. Low glioma (Benign) images have an accuracy of 68.75 percent. In [3] the authors performed Brain tumor CT imaging classification into malignant and malignant images using SVM with kernel function. The best performance of the categories is obtained through WSVM, based on test data. Moreover, these findings suggest that the proposed method is effective and effective in predicting malignant and malignant tumors from CT scans of the brain, with 74% accuracy of SVM and 76% of WSVM.

The proposed method may be consistent with other types of imaging, such as MRI, in the future, and it may be used to classify and differentiate tumors from other parts of the body. The authors of another study [4] used the effectiveness of pressure strategies to improve accuracy as well as timing of CNN executions in differentiation by brain MRI. Before they were divided into sections, they presented the first step in processing. The method first uses the Probabilistic Neural Network (PNN) to differentiate the Genital Region (ROI) (especially the area of the brain tissue), and then uses Back Propagation Neural Networks to narrow the ROI (BPNN). Finally, compressed images are uploaded to CNN segregation. For comparison, results were obtained for three different grades of formats and the accuracy levels displayed were more than 90%. R-CNN's rapid method of successful detection and classification of brain MRI images was proposed by the authors of another paper [5]. Images are first embedded in a simple CNN, which produces a dynamic feature map, which is then changed to regional suggestions, which are, after that, resized into a feature vector by the integration stage of ROI.

Finally, by categories, this ROI vector feature is integrated into the fast R-CNN. SVM was also used to create the largest margin possible between classes, allowing the algorithm to classify images into categories with greater accuracy. Their algorithm had an accuracy rate of 95%. [6] proposes a flexible approach based on in-depth study of Generative Adversarial Networks (GANs). Their plan was to previously train CNN as a GAN racist using two databases. The productive part of GAN was aided by the development of data on the production of more realistic MRI scans of the brain. To separate, the final bed of CNN discrimination on GAN is changed by SoftMax. In the Figshare database, this precise CNN separator is fine-tuned over time. The outcomes reveal 88 percent precision. [7] Due to the complexity of imaging and there aren't any anatomical models that accurately represent the possible deformities in each and every component, the separation of medical imaging is a difficult task. When it comes to the original size of the collection, the proposed method is quite effective and collection centers. This function recommends a system that requires non-invasive human intervention to isolate tissue of the brain. The prime purpose of this suggested program is to assist human specialists or surgeons in locating patients in less time.

Methodology

The proposed work has been implemented in four phases. The experimental setup involved a system with 4GB Ram, Windows 10 and implemented via Python.

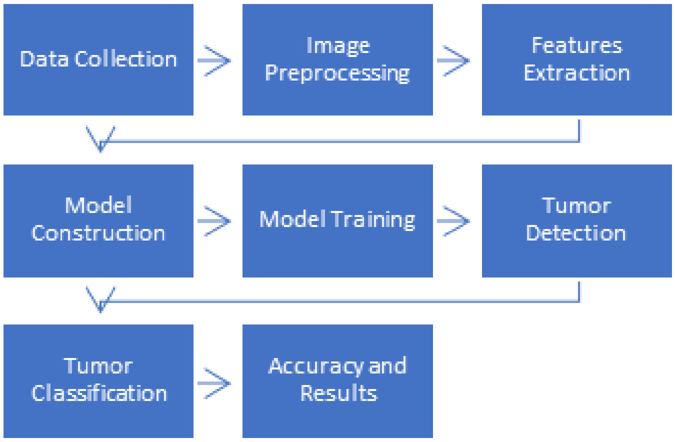

Fig 2. Steps followed for tumor detection and classification

Phase-I
Data Collection and Image Pre-processing

- Data collection - The standard available dataset [13] was collected for the implementation. The dataset contains 2212 brain MRI images in two folders: yes and no. Yes, there are 1112 tumorous brain MRI images in the folder, and no, there are 1100 non-tumorous brain MRI images in the folder.
- Data Pre-processing - The information we gathered came in a variety of shapes, sizes, and orientations. As a result, the photos had to be resized to a shape of (240, 240, 3) = (image width, image height, number of channels). As a result, all photographs should have the same shape to transmit as an input into the neural network. Normalization was the next step in the process. Normalization is a technique for altering the range of pixel intensity values in image processing. Normalization, also known as histogram stretching or contrast stretching, can be used on photos with poor contrast due to glare. The photos were scaled to the range 0-1 so that they could be viewed regularly. To do so, multiply all of the pixel values by the largest pixel. This equals 255.

Phase-II
Features Extraction using HOG

HOG [11] is a feature descriptor that is used in image processing to recognize objects. The feature descriptor's goal is to generalize an object in an image so that it offers the same feature descriptors in all images that contain that object, regardless of angle, illumination, distance, or other circumstances. In a detection window, or region of interest, the HOG descriptor approach counts occurrences of 22 gradient orientations. in a limited section of a picture (ROI). One of the most significant topics in computer vision is the gradient vector. A gradient vector is sometimes known as a picture gradient. Feature extraction and edge detection are two applications of gradient vectors. The difference of neighborhood values in both the horizontal and vertical axes is used to calculate the gradient vector for a certain pixel. In HOG, orientation is crucial. The term "orientation" refers to a shift in direction when the pixel intensity value changes. The direction of change can be along the X-axis or the Y-axis. Orientation = arctan (Y/X) is the formula for determining the orientation of a gradient vector. The foundation for histogram grouping and normalization is the grouping of cells into blocks. The normalized group of cells is called a block histogram. The feature descriptor is represented as a set of block histograms. A feature descriptor is a technical term that refers to a picture or a portion of an image that simplifies the original form of the image by removing the key data. In HOG, the distribution of gradient orientation is used as a feature. The picture, orientations, pixels per cell, cells per block, visualize, and multichannel parameters are used as input to the Hog() function.

Phase-III
Tumor Detection using CNN

A CNN is a Deep Learning system that can take an image as input and give significance to distinct aspects/objects in the image, as well as distinguish between them. In comparison to other partitioning algorithms, ConvNet just requires a little amount of computation. Despite the fact that the fundamental approaches need manual filter manipulation, ConvNets can learn these filters / symbols with enough training. ConvNet can successfully capture Local and Temporary dependencies on a picture by using the right filters. Architecture creates a high degree of image databases because of the reduced number of parameters involved and the reusability of weights. In other words, the network may be trained to recognize higher-quality image technologies. In order to detect brain cancers in multiple MRI scans of the brain we performed detection using various algorithms Like CNN, SVM, OTSU'S THRESHOLDING, KNN and Random Forest. But among all algorithms the best one found out with highest accuracy [12] was CNN. The formula used to determine the model's accuracy is –

$$\text{ACCURACY} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where, TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative. TP and TN mean the samples which have been detected accurately and FP and FN mean the samples which have not been detected accurately.

Phase-IV
Tumor Classification using SVM

A SVM algorithm's purpose is to find a hyperplane in an N-dimensional space that clearly separates data points. To separate the two sorts of data points, there are several hyperplanes to pick from. Our objective is to locate a plane with a large limit, or a large distance between data points from both groups. Increasing the gene range adds some support, making it simpler to succeed in distinguishing the following data points. SVM [10] has the following advantages: It works really well with a large separation margin. It works well in three-dimensional areas. If the maximum size exceeds the sample size, this method works best. Memory works well because it uses a set of training points (called supporting vectors) in the decision-making process. The HOG characteristics that we retrieved from the MRI images were saved in a.csv file. We employed a radial basis function to accomplish the processing, which involves mapping data into a higher-dimensional space. Kernelling is the term for this process, and the kernel function is the mathematical function that is utilized to alter the data. Then we have fitted our model and predicted the new values.

Results

For the CNN algorithm the data collected and the detection done by our algorithm is as follows –

Table 1
Positive and Negative brain tumor raw data

	No. of MRI Images	Positive MRI Images	Negative MRI Images
RAW DATA	2212	1112	1100

From all the images of dataset, the accuracy of detection of images is as follows –

Table 2
Positive and Negative brain tumor detected data

	TP	TN	FP	FN
No. of MRI Images detected	983	942	58	129

So, using the result the accuracy of our model CNN is 87.02%. For the SVM algorithm the data taken and the classification into benign and malignant done by our algorithm is as follows - In the dataset that we took for our CNN model - 983 were truly detected positive for brain tumor, this data was then split into training and testing data. 137 were kept into testing data and 846 were kept into the training dataset. Out of these 137-testing dataset, the data of 90 brain tumors were benign and 47 were malignant. But after evaluating the data in our SVM algorithm, out of these 137 tumors, only 85 were detected to be benign positively and 47 were detected to be malignant positively whereas 5 were detected to be false malignant and there were no false benign tumors.

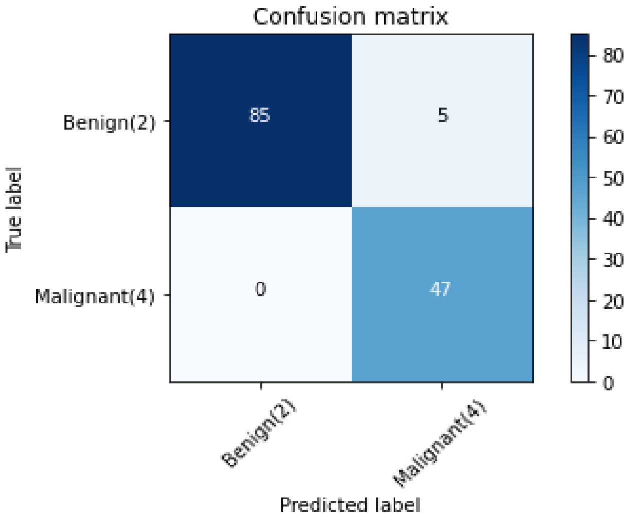

Fig 3. Confusion Matrix

From this matrix it is clear that the benign tumors truly detected are 85 and the malignant tumors truly detected are 47. While the false malignant tumors detected are 5 in number however there are no benign tumors detected falsely. So, the accuracy of our SVM model is 96.35%.

Conclusion

The most important aspect of our proposed work was how to use a CNN to detect the presence of a tumor and a HOG to extract features from brain tumor MRI images, followed by a classification method, SVM, to classify brain tumors into benign and malignant. An approach based on a combination of the feature extraction algorithm (HOG), CNN, and SVM for tumor identification and classification from brain pictures is provided. CNN has the ability to identify tumors. When it comes to picking an auto-feature in medical photographs, the CNN comes in handy. Clinicians classified the images obtained at the centers, and then tumor screenings were divided into two categories: normal and patient. The proportion of sick to healthy subjects was proportionate to the proportion of image categorization in two classes. After pre-processing, the images were fed into the CNN. CNN's accuracy is 87.02 percent for appropriately categorizing images into two normal and patient classes. The accuracy of the proposed algorithm increased to 96.35 percent on the test data while using the proposed approach of feature extraction and using SVM for categorization into benign and malignant tumors, which is an improvement over the CNN. Due to the importance of the physician's diagnosis, the model's accuracy can assist doctors in diagnosing the tumor and treating the patient, resulting in a high level of medical accuracy when utilizing the proposed technique. This model can be expanded in the future to further classify benign and malignant tumors into grades of different types (Grade 1, Grade 2, Grade 3, Grade 4). This could aid the medical system in providing more effective and timely treatments.

References

1. G. Hemanth, M. Janardhan and L. Sujihelen, "Design and Implementing Brain Tumor Detection Using Machine Learning Approach," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 1289-1294
2. Pedapati, Praveena, and Rama Vaishnavi Tannedi. "Brain Tumor Detection using HOG by SVM." (2018).
3. Rezaei, Kimia, Hamed Agahi, and David C. Wyld. "Segmentation and Classification of Brain Tumor CT Images Using SVM with Weighted Kernel Width." *Computer Science & Information Technology* 39 (2017).
4. El Boustani, Abdelhakim, and Essaid El Bachari. "MRI Brain Images Compression and Classification Using Different Classes of Neural Networks." In *International Conference on Model and Data Engineering*, pp. 122-134. Springer, Cham, 2019.
5. Poonguzhali, N., Kagne Raveena Rajendra, T. Mageswari, and T. Pavithra. "Heterogeneous Deep Neural Network for Healthcare Using Metric Learning." In *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1-4. IEEE, 2019.
6. Ghassemi, Navid, Afshin Shoeibi, and Modjtaba Rouhani. "Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images." *Biomedical Signal Processing and Control* 57 (2020): 101678.
7. Arif, Muhammad, F. Ajesh, Shermin Shamsudheen, Oana Geman, Diana Izdrui, and Dragos Vicoveanu. "Brain Tumor Detection and Classification by MRI Using Biologically Inspired Orthogonal Wavelet Transform and Deep Learning Techniques." *Journal of Healthcare Engineering* 2022 (2022).
8. https://www.nibib.nih.gov/science-education/science-topics/magnetic-resonance-imaging-mri
9. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
10. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47
11. https://towardsdatascience.com/hog-histogram-of-oriented-gradients-67ecd887675f
12. https://deepchecks.com/how-to-check-the-accuracy-of-your-machine-learning-model/
13. https://www.kaggle.com/jakeshbohaju/brain-tumor