# Role of Data Sets in Analyzing the Accuracy of Deep Learning Methods Applications

**Purushottam Das,**
Asst. Professor, Department of Computer Science, Graphic Era Hill University,
Dehradun, Uttarakhand India 248002,

**Abstract**
After the embedding the deep learning approaches with the medical sciences techniques, it becomes easy to detect and diagnosis the disease then earlier. Most of the comfort happens by the invention of the imaging technologies with deep learning. In this article we discuss the deep learning approaches or applications in detecting different type of brain tumors at early stages and the advantages, which helps in saving the life of patient. Through deep learning we can cover multiple areas of medical sciences viz. image detection, classification, segmentation, localization registration etc.
**Keywords-** image detection, classification, segmentation, deep learning, brain tumor.
**DOI Number: 10.48047/NQ.2022.20.1.NQ22355** **NeuroQuantology2022;20(1): 752-758**



## I Introduction

The potential of Deep Learning techniques have been deeply invested in several fields of medicine form drug discovery to clinical decision making by significantly altering the way of practicing medicine. As most of the medical records have been digitalized now-a-days the use of these deep learning techniques and algorithms has paved a best path in medical image analysis. By readily increasing the advancements in medical sciences, the usage of electronic health records (EHR) quadrupled from 11.8% to 39.6% in the US from 2007 to 2012 [1] among office-based physicians. Patient's electronic health records have many medical images ass mandatory which are being analysed for fatigue and experience by human radiologists. It takes a lot of effort and years of experience to train the radiologists in detecting the disease which also involve financial cost. It is observed that some of the health care systems outsource radiology which report to lower-cost and developing countries like India via tele-radiology. Any delay in the detection of disease or erroneous diagnosis can cause harm to the patient. Therefore, it is really worthy for automated, accurate and efficient deep learning techniques to be used in medical image. Through our model we are going to predict the type of brain tumor based on the back-propagation methodology in Convolutional Neural Networks. There is rapid development at image classification techniques and image segmentation which increase the frequency of using the automated systems in medical image diagnosis. Smith-Bindman *et al.* [2] has observed the medical records used from 1996 to 2010 across six large integrated healthcare systems in the United States, which involve 30.9 million imaging examinations. After long tedious efforts and many years of study the researchers have found that the usage of CT scans, MRI images and PET have been increased by 7.8%, 10% and 57% respectively.

The development of rule-based, expert systems in 1970s was one of the greatest achievements in disease diagnosis. The MYCIN system by Shortleaf [3], was the early discovery in developing automated health care systems which provide expert-level solutions to several complex problems by suggesting various regimes of antibiotic therapies for patients. Parallel to those developments, AI algorithms gradually evolved from heuristics- based techniques to manual, handcrafted feature extraction technique and then to supervised learning techniques followed unsupervised machine learning methods. Even though there are many techniques which are being discovered, the majority of algorithms from 2015- 2017 have reported the usage of supervised learning algorithms.

Currently, Convolutional Neural Networks (CNNs) are mostly used in medical image analysis apart from machine learning algorithms [4]. CNNs will carefully preserve the spatial relationships when filtering the medical images which are passed as input. Spatial relationships have utmost importance in radiology as they find out minute changes in body like how the edge of a bone joins with muscle, or where normal lung tissue interfaces with cancerous tissue or detect any growth of tumors on various parts of body. As shown in the Figure-1., a CNN takes an image with raw pixels and then transforms it through several Convolutional layers, Rectified Linear.



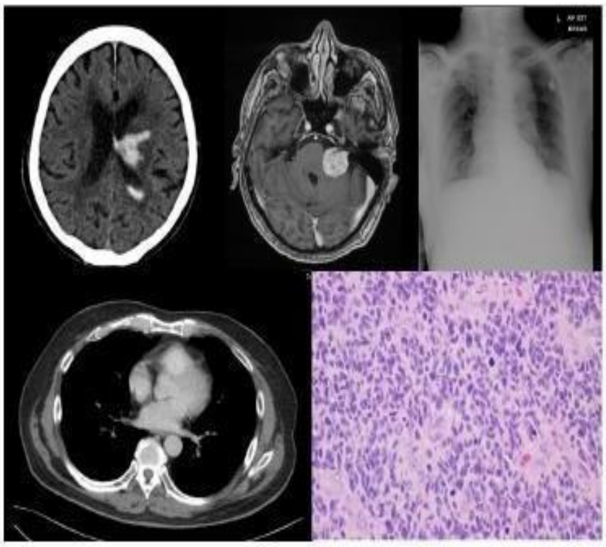


Figure-1 CNN raw pixels and transformation

**Methodology**

Numerous works in the diagnosis of brain tumor using deep learning techniques have motivated this work. A Brain Tumor Prediction System built with the assistance of deep learning techniques, Convolutional Neural Networks, Image classification and segmentation is discussed here. There may be a variation in the results obtained by applying various techniques in determining the accuracy. Brain tumor prediction system can answer complex "what if" queries which conventional decision support system scan not be proposed [2]. This system is capable of responding queries that the traditional decision support systems were not able to. It involves classification by considering crucial knowledge such as patterns, relationships amid medical factors connected with brain tumor diagnosis. This brain tumor prediction system remains well-being web- based, user-friendly, reliable, scalable and expandable. Multiple experiments were carried out on a sampled data set of patient's records by training the system with many sample MRI images. The Neural Network is trained and tested with many attributes like weighted average of RGB colors, size of the raw pixel, position of the tumor, orientation of tumor, ambient light conditions, age of patient and many others. The supervised network has been advised for diagnosis of brain tumor. Training was carried out with the help of back propagation algorithm which is typical concept of Convolutional Neural Networks.

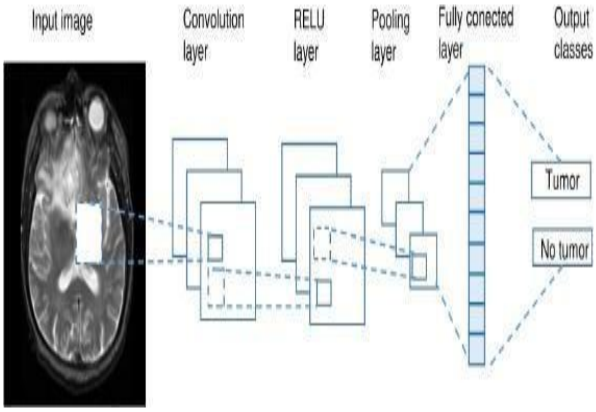
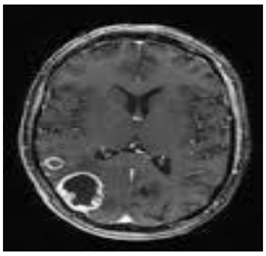


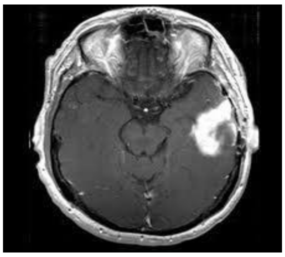

Figure 2-Brain Tumor detection method

After the system is trained by set of sample MRI images even though if an unfamiliar data was inserted by the doctor, the system predicts the type of tumor based on previous experience and determines the type of tumor whether it is cancerous or non-cancerous. In the article "Disease Forecasting System Using Data Mining Methods", the preprocessed data is clustered using clustering algorithms as K-means to gather relevant data in a database i.e. to train the system. Maximal Frequent Item set Algorithm (MAFIA) is applied for determining maximal frequent model in brain tumor database. Using the C4.5 algorithm as training algorithm using the concept of information entropy the regular patterns can be classified into different classes. This model shows that the designed prediction system is capable of predicting whether the type of brain tumor is Benign or Malignant. In this research paper, the usage of image segmentation, feature extraction and classification techniques are used for prediction of type of brain tumor whether it is cancerous or non-cancerous. In traditional medical scoring systems some of the limitations are that there is a presence of intrinsic linear combinations of attributes in the input set, and hence they are not skilled at modeling nonlinear complex interactions in medical domains. This limitation is handled in this research by use of classification techniques which can detect complex nonlinear relationships between independent and dependent variables. Also the model has the ability to identify all possible interactions between predictor variables. In this research paper Predict the Diagnosis of Brain Tumor Patients Using Classification Mining Techniques", the data mining is mainly utilized for the prediction of brain tumor. The main objective of our work is to predict the type of brain tumor by performing several segmentation and classification techniques involved in Convolutional Neural Networks.

In this brain tumor detection we discuss several image processing techniques, feature extraction techniques and classification techniques in deep learning. We will train the system by providing several MRI images and predict the type of tumor based on different features. We perform image segmentation and thus analyze whether the tumor is cancerous or non -cancerous. By using Convolutional Neural Networks in deep learning we can predict it accurately. There are many types of tumors like Medulloblastoma, Meningioma, and Acoustic Neuroma but in our model we focus mainly on cancerous and non-cancerous tumors as shown in figure 3.

Types of tumors-





a) Benign Tumor b) Malignant Tumor

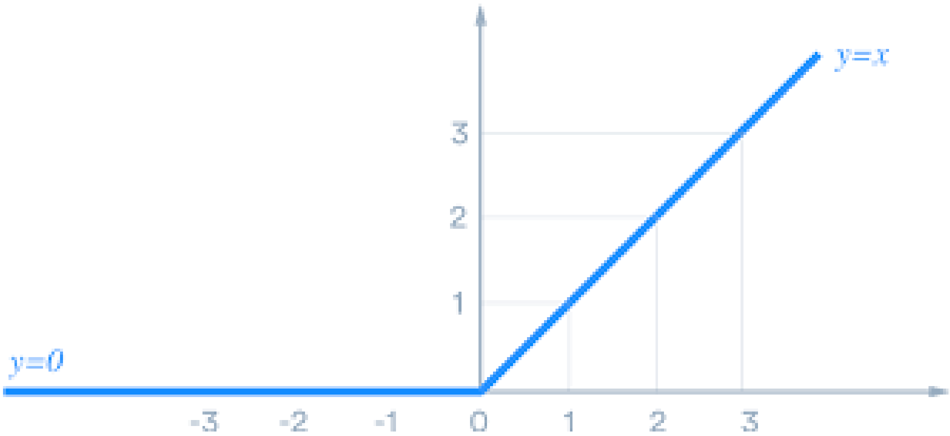
Figure 2. cancerous and non-cancerous tumors

Past image processing techniques involved human interaction by observing medical images, CT scan images and MRI images. These systems are digitally backward which involves human observance in detecting flaws, anomalies and tumors in parts of the body. To detect brain tumor radiologists used to observe the MRI images and predict the tumor. It would take days to check the tumor test it and then predict the type of tumor whether it is Benign or Malignant.

In the candidate model we build a neural network by using CNN [11] and train the system by providing MRI images of brain which are divided into two types Benign and Malignant. Benign tumors are usually non-cancerous tumors and they do not spread across other body parts. Malignant tumors are cancerous tumors which are spread rapidly across other body parts and may be dreadful. CNN also deals with back-propagation which is vital when working with images and feature extraction techniques. We initially import deep learning libraries like Keras and Tensor flow using Anaconda Navigator. These are the two high level libraries dealing with complex numeric calculations and image classification functions. The activation functions used in this model are ReLU and sigmoid functions. ReLU stands for rectifier linear unit is a kind of activation function especially used in CNN. ReLU is especially used to convolute the image into matrix and then sets all negative values in the matrix to zero and all other values are kept constant.



Figure 3- Graph for ReLU activation function



The model is initially pre-processed by importing all the CNN libraries like keras and tensor flow. The MRI image of brain tumor is passed as an input to the neural network and then it is passed through many transformation layers. Firstly the image is processed through the Convolutional layer where image is pre-processed and then it is forwarded to ReLU layer where the feature extraction is done. ReLU activation function takes the features of the image dynamically unlike Machine Learning algorithms which have the features constant. Here some features like weighted average of RGB pixel value, age of person, position of tumor and light incidence are considered.

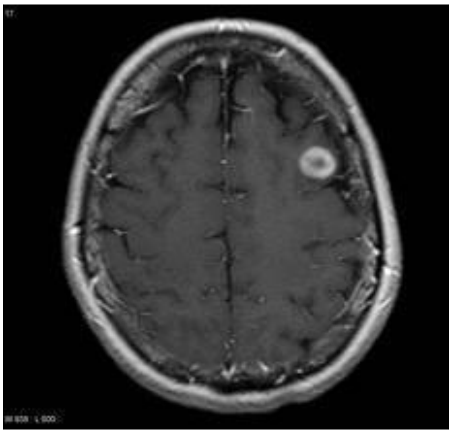
The image is set to the Pooling layer which progressively reduces the spatial size of the representation to reduce the amount of parameters and computation in the network. Here we used an approach called max pooling. The obtained result is then forwarded to Fully Connected Layer. In this layer each neuron receives input from every element of the previous layer. The input area of a neuron is called its receptive field. As we cannot predict the life of a person by taking only one decision we need to cross check it multiple times with different combinations by the use of Fully Connected Layer. The obtained output is classified to determine the type of tumor whether it is Benign or Malignant.

We have used CNN methodology here because it allows back-propagation which takes numerous computations in determining the solution for complex problems with high accuracy.

**III Results**

The results we have obtained when provided MRI images of brain tumor as input to our system have yielded the following results:



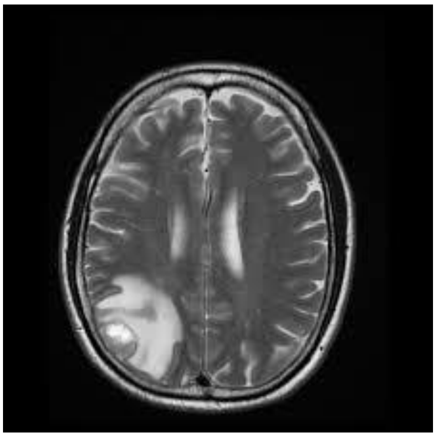


a) Benign tumor image

```
if result[0][0] == 0:
    prediction = 'Benign'
else:
    prediction = 'Malignent'
print("Detected tumor type is %s"%prediction)

Detected tumor type is Benign
```

b) Benign tumor result



c) Malignant tumor image

```
if result[0][0] == 0:
    prediction = 'Benign'
else:
    prediction = 'Malignent'
print("Detected tumor type is %s"%prediction)

Detected tumor type is Malignent
```

d) Malignant tumor result

Figure 4- results

**IV Conclusion**

A recurring theme in deep learning is that the limit imposed by the shortage of labeled datasets, which hampers training and task performance. Conversely, it's acknowledged that more data improves performance, as Sun et al. [85] shows using an enclosed Google dataset of 300 million images. Normally computer vision tasks, attempts are made to avoid limited data by using smaller filters on deeper layers with novel CNN architecture combinations or hyper parameter optimization. In medical image analysis, the shortage of data is two-fold and more acute: there's general lack of publicly available data, and top of the range labeled data is even scarcer. Most of the datasets presented during this review involve fewer than 100 patients. Yet true may not be as dire because it seems, as despite the little training datasets, the papers during this review report relatively satisfactory performance within the assorted tasks. The question of what percentage images are necessary for training in medical image analysis was partially answered by Cho et al. He ascertained the accuracy of a CNN with Google Net architecture in classifying individual axial CT images into one in every of 6 body regions: brain neck, shoulder, chest, abdomen, and pelvis.

**References**

[1] C.-J. Hsiao, E. Hing, and J. Ashman, „„Trends in electronic health record system use among office-based physicians: United states, 2007–2012,"" Nat. Health Stat. Rep., vol. 75, pp. 1–18, May 2014.

[2] R. Smith-Bindman et al., „„Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996–2010,"" JAMA, vol. 307, no. 22, pp. 2400–2409, 2012.

[3] E. H. Short life, Computer-Based Medical Consultations: MYCIN, vol. 2.New York, NY, USA: Elsevier, 1976.

[4] G. Litjens et al. (Jun. 2017). „„A survey on deep learning in medical image analysis.""
Available:
https://arxiv.org/abs/1702.05747

[5] Bikku, Thulasi, and Radhika Paturi. "A novel somatic cancer gene based biomedical document feature ranking and clustering model." Informatics in Medicine Unlocked (2019): 100188.

[6] F. Rosenblatt, „„the perceptron: A probabilistic model for information storage and organization in the brain,"" Psychol. Rev., vol. 65, no. 6, pp. 365–386, 1958.

[7] D. H. Hubel and T. N. Wiesel, „„Receptive fields, binocular interaction and functional architecture in the cat‟s visual cortex,"" J. Physiol.,vol. 160, no. 1, pp. 106–154, 1962.

[8] K. Fukushima and S. Miyake, „„Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition,"" in Competition and Cooperation in Neural Nets.Berlin, Germany: Springer, 1982, pp. 267–285.

[9] Y. LeCun et al., „„Backpropagation applied to handwritten zip code recognition,"" NeuralComput., vol. 1, no. 4, pp. 541–551, 1989.

[10] D. E. Rumelhart, G. E. Hinton, and R. J.Williams, „„Learning representations by back- propagating errors,"" Nature, vol. 323, pp. 533– 536, Oct. 1986.

[11] https://towardsdatascience.com/acomprehensi

[12] Bikku, Thulasi. "A new weighted based frequent and infrequent pattern mining method on realtime E-commerce." Ingenierie des Systemesd'Information 23.5 (2018): 121.