Journal of Informatics
Electrical and Electronics Engineering, 2023,
Vol. 04, Iss. 01, S. No. 004, pp. 1-10
ISSN (Online): 2582-7006

Evaluation of Data Sets and Algorithms for Brain Tumor Detection Using MRI Images: A Python-Based Approach

Mugdha Singh[1], Vineet Singh[2], Shikha Singh[3], Pooja Khanna[4]

[1,2,3,4]Amity School of Engineering and Technology, Amity University, Lucknow, India
mugdhasingh2300@gmail.com[1], vsingh@lko.amity.edu[2], ssingh8@lko.amity.edu[3], pkhanna@lko.amity.edu[4]

How to cite this paper: M. Singh, V. Singh, S. Singh and P. Khanna, "Evaluation of Data Sets and Algorithms for Brain Tumor Detection Using MRI Images: A Python-Based Approach," *Journal of Informatics Electrical and Electronics Engineering* (*JIEEE*), Vol. 04, Iss. 01, S No. 004, pp. 1–10, 2023.

https://doi.org/10.54060/jieee.v4i1.82

Received: 18/01/2023
Accepted: 11/03/2023
Published: 25/04/2023

Abstract

This study is to evaluate the work of the data sets and major algorithms that are involved in the brain tumor detection system using the MRI image with the help of the python concept. So basically, is an easy manner if we require to define the phenomena of the brain tumor it could be the abnormal condition that causes the problem of cancer, here the abnormal condition refer to the growth in cell body in not a very suitable way for the brain tissue. In the respective paper, we proposed an algorithm to segment brain tumor from 2D Magnetic Resonance Image of the brain by a CNN. When this algorithm is applied to MRI images, a brain tumor diagnosis can be made more quickly and accurately, which makes it easier to give patients treatment. These predictions enable the radiologist to make quick decisions as well. In the proposed work, the performance of a self-defined Convolution Neural Network (CNN) is evaluated. For the purpose of faster and efficient accuracy, we will implement the proposed method using the "TenserFlow" and "keras" in "Python.

Keywords

Brain tumor detection, MRI images, Data sets, Algorithms, Convolutional Neural Network (CNN), TensorFlow, Keras.

1. Introduction

Primarily, this study focuses on the brain and its functionality because it influences the functioning of other organs and plays a crucial role in decision-making. The brain is considered the most significant organ in the human body [1]. A tumor is an unwelcome fibrous web of tissue that grows unchecked inside our brain. It serves as the control center of the central nervous system and is responsible for both voluntary and involuntary activities in the human body on a daily basis.

The primary focus of this project is the extraction of tumors from brain MRI images and representing them in a manner that is easier to comprehend. The objective is to provide valuable data in a less complex structure, particularly for the clinical staff involved in patient treatment. The purpose of this work is to define an algorithm that can produce an extracted image of the tumor from a brain MRI scan. The resulting image will provide information about the tumor's size, dimensions, and location. The boundary of the tumor will provide additional information that can be useful in various cases, aiding the staff in making better-informed decisions regarding the treatment plan. A Convolutional Neural Network (CNN) is utilized to deter-mine the presence of a tumor in a given brain MRI image.

It is worth noting that approximately 3,540 children under the age of 15 have been diagnosed with a brain tumor this year. Understanding the stages of the disease is essential for both preventing and treating brain tumors. The examination conducted in this paper reveals whether the frontal cortex is healthy or affected by applying significant learning techniques. The study employs Artificial Neural Network (ANN) and CNN to differentiate between healthy and cancerous brains [2]. ANN works in a manner similar to the nervous system in the human brain. It consists of interconnected processing units that can be trained using a training set and experiential knowledge stored within a digital computer. The neural network has multiple layers of interconnected neurons and acquires knowledge by incorporating a dataset into the learning process.

In the modern era, manually locating a brain tumor among the numerous MRI (magnetic resonance imaging) images is a challenging and inaccurate task. This can have an impact on the patient receiving appropriate medical care. Additionally, due to the large volume of image datasets involved, it can be time-consuming. The similarity in appearance between brain tumor cells and normal tissue makes segmenting tumor regions difficult. Therefore, a highly accurate automatic tumor detection method is necessary.

Early tumor detection significantly improves a patient's survival rate for primary treatment [3]. However, processing MRI images for early tumor detection presents a challenge of high processing overhead due to the large volume of images in-volved. This leads to significant delays and reduced system efficiency. Consequently, there is a growing need for more advanced detection systems that enable precise segmentation and representation, facilitating quick and accurate processing. Recent research suggests the development of novel brain tumor detection strategies based on enhanced learning and processing. This paper briefly reviews the advancements made in MRI processing for the early diagnosis and detection of brain tumors, including segmentation, representation, and the application of new machine learning (ML) methods in decision making. The proficiency and precise handling of machine learning algorithms have shown improvements in ongoing automation frameworks for faster and more accurate brain cancer detection.

While there may be several hidden layers, there is typically one input layer and one output layer. During the learning process, the weights and biases assigned to the neurons in each layer are determined by the input features and the preceding layers (hidden layers and output layers, respectively) [4]. A model is trained by applying an activation function to the input features and hidden layers to produce the desired output. This paper primarily focuses on using CNN because ANN requires fully connected layers, which require more processing, while in this case, the input is an image [5]

The contribution of this brain tumor research paper lies in the application of Convolutional Neural Networks (CNNs) to significantly improve the accuracy of brain tumor detection and classification. By developing a novel CNN architecture specifically designed for this purpose and training it on a large dataset of brain tumor images, the study achieves superior performance compared to traditional methods. The CNN demonstrates robustness to variations in image quality, enabling reliable tumor detection and classification in diverse clinical settings. Moreover, the research highlights the interpretability of the CNN's decisions, providing insights into the learned features and enhancing trust in the automated analysis. The proposed

CNN not only advances the state-of-the-art in brain tumor analysis but also holds promise for integration into clinical practice, improving diagnosis and treatment planning for brain tumor patients

This research paper discusses its findings and contributions through several key points. Firstly, it provides a comprehensive literature survey of the relevant research, introducing different terminologies used throughout the paper. Secondly, the methodology adopted and applied is detailed, accompanied by a flowchart illustrating the process. This includes information on the dataset used, data preprocessing techniques employed, and the algorithm applied. Next, the experiments performed and the analysis of the results obtained are discussed. The fifth point highlights the presentation of the results through graphical representations. Lastly, the research paper concludes by summarizing the overall work and its contributions.

2. Literature Review

Information technology has had a significant impact on various fields in our modern era, leading to increased machine dependence in our day-to-day lives. The medical field has also experienced a crucial transformation due to IT advancements, including automated surgery, pharmaceutical production, testing of various factors, and disease prediction [6]. Machine learning and artificial intelligence have extensively collaborated with the healthcare industry, resulting in a vast area of re-search that constantly develops new innovations to enhance our well-being. Therefore, the objective of this study in the field of e-healthcare is to enhance support for medical professionals, such as radiologists and clinical experts. The study acknowledges the challenges involved in distinguishing abnormal brain magnetic resonance images from normal ones, given the diverse brain matter types (white and gray matter), cerebrospinal fluid, and brain cells. Various feature extraction techniques are employed to address this challenge. [7-9]

This paper contributes to the basic understanding of the Brain Cancer Detection and Classification System by utilizing image processing techniques such as histogram equalization, image segmentation, image enhancement, and feature extraction. The proposed method employs an Artificial Neural Network (ANN) as a classifier, achieving a high level of classification efficiency compared to other classifiers. It improves responsiveness, explicitness, and exactness. The proposed approach is effective and computationally efficient, as the outcome is crucial for patient treatment, requiring robustness and accuracy in prediction algorithms for medical diagnosis. Several well-known characterization and grouping algorithms are utilized to enhance the precision of anomaly detection. Clustering of medical images aims to simplify the analysis and representation of the image. Some grouping and characterization algorithms specifically target improving the forecast precision in the process of distinguishing anomalies.

In this investigation, the Convolutional Neural Network (CNN) is employed to identify meningioma, glioma, and pituitary tumors with recall rates of 88%, 81%, and 99% respectively, resulting in an overall accuracy of 91.3%. A deep learning architecture utilizing 2D convolutional neural networks is used for the classification of different types of brain tumors from MRI image slices. The paper covers various stages including data acquisition, data preprocessing, pre-modeling, model optimization, and hyperparameter tuning. The model's generalizability is evaluated using 10-fold cross-validation on the entire dataset. The analysis of the image is performed using CNN-based methods such as voting, voxel-wise classification, and efficient patch-wise evaluation.

The human brain, as a crucial organ, oversees and coordinates the activities of other body parts. It serves as the control center of the central nervous system, responsible for daily voluntary and involuntary functions [10]. Brain tumors are proliferating, unrestricted fibrous webs of abnormal tissue within the brain. Magnetic resonance imaging (MRI) is frequently employed by radiologists to examine brain tumors in order to prevent and treat them at various stages.

Tumor classification poses a major challenge, as detecting a tumor is extremely difficult. Segmentation is particularly challenging due to the significant variations in tumor location, shape, and structure among patients. In the figure below, we have displayed images of the same brain slice from different patients, highlighting the variations in tumors. The growth area differs noticeably in all the displayed pictures/patients. Moreover, the intra-tumoral structures and shapes differ among the images from each of the eight patients. The images below demonstrate that a tumor may have more than one region. This emphasizes the difficulty of automatic segmentation.

3. Methodology Applied

In the field of medical image processing, a well-ordered method is the Convolutional Neural Network. A type of artificial neural network known as a convolutional neural network (CNN) is used in the process of image recognition and is designed specifically for method component knowledge. CNN is a powerful computing method for image processing that uses deep learning for both generative and descriptive tasks. A neural network might be a system of hardware and computer code that functions similarly to neurons in the human brain. The image processing process is not a good fit for artificial neural networks. A CNN makes use of a system that is very similar to a multilayer view-point and was made to be used with fewer processes. In the method we propose, we have taken a wide range of images as input and converted them all into the same size (60*60) to create a variety of dimensions. It consists of six steps where the execution starts from taking an input image from the data set followed by the image pre-processing, image enhancement and the brain tumor classification using Convolutional Neural Network. Finally, the output is observed after all the steps are completed. Each module is unique in its own way. Every step has its importance. This architecture also includes a testing and training data set. The data set used is has been downloaded from Github which consists of nearly 2000 images that are used to test and train the system. The input image is pre-processed by using the noise filter like Median Filter and Bilateral Filter and the image is enhanced using the Sobel Filter. Then the obtained image using segmented using binary thresholding and morphological operations are performed on it. Finally, the image classification is done using Convolutional Neural Network to predict whether the tumor is present or not.

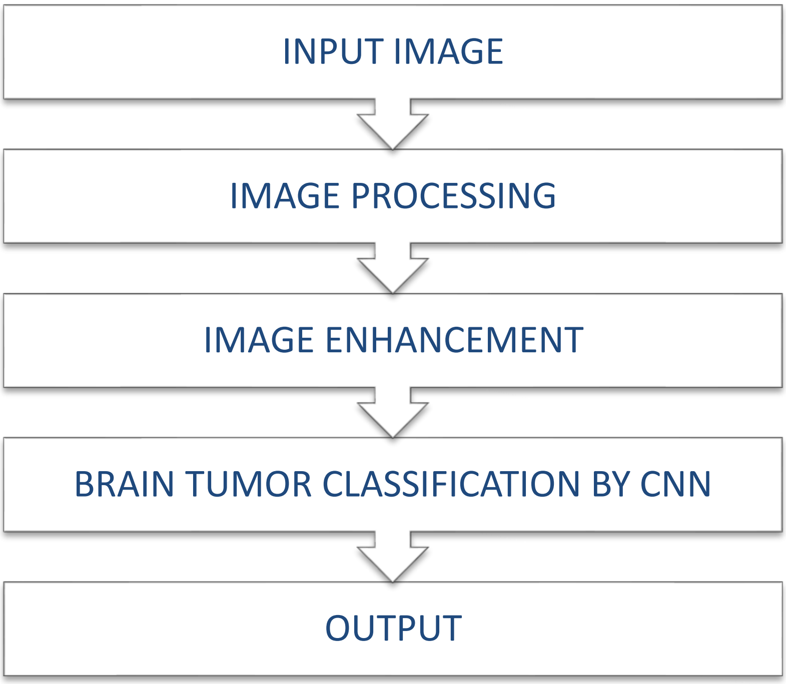

Figure 1 Module Division

Data set

The data is sourced from the Github website, which includes brain tumor MRI images. These images are divided into two folders, one housing normal brain images and the other housing tumor images. In total, there are 2065 images within these folders [11]. Figure 1 illustrates sample images of both normal brains and brain tumors. Specifically, there are 1085 tu-mor-related images and 980 non-tumor-related images. The images vary in size, such as 630 x 630 and 225 x 225, but they have been resized uniformly to 256 x 256.For the training phase, a total of 1672 images were used, with 795 non-tumor im-ages and 877 tumor images. The validation set consisted of 186 images, including 92 tumor images and 94 images con-taining nothing but tumors. The testing set contained 207 images, out of which 91 were non-tumor images and 116 were tumor im-ages.

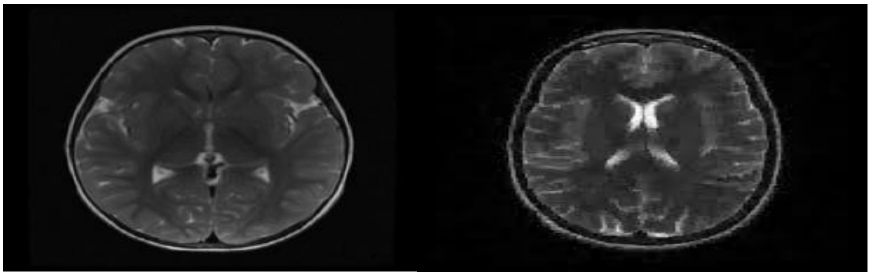

Figure 2. Normal Brain(Left) and Brain with Tumor(Right)

Data Set Pre-processing

The pre-processing of the image is going to contain two sections of step; in which the first section contributes to the size ar-rangement of the image which help in decreasing the computation difficulty by changing the image pixel and we which change the image colour property from RGB to grey image that will help in reducing the byte of array used for the image recognition.

a) Resizing the Image

Firstly we will resize the image obtained from the dataset that we took up from the github. The pixel of the image will change from the dimension (256*256) to (60*60) which decreases the computation difficulty.

b) Image transformation from RGB to grey

The RGB stand for Red, Green, Blue colur that indicate three proportion of array as bytes if we change it to or transform it to gray the image will have only one colour of each pixel.[12]

Algorithm Applied

The two techniques ANN and CNN are applied on the brain tumor dataset and their performance on classifying the image is analyzed. Steps followed in applying ANN on the brain tumor dataset are

- ANN –Artificial neural network.
- CNN convolutional neural network.

This section tells us the steps and the requirements for the model creation, the source of the magnetic resonance image da-taset and the algorithm used to detection the tumor cell in the brain using the brain magnetic resonance images. As the test set images, different types of magnetic resonance images of the brain were used, including the augmented images of the existing test set and the train dataset so as to test a variety of test cases. The presented methodology has been applied to real dataset collected from github including the brain magnetic resonance images of 256*256 pixels. All the magnetic reso-nance images were first converted to grayscale and then they were augmented before processing of the images.

Artificial Neural Network (ANN)

There are seven layers in the ANN model used here. The first layer is the level layer which changes over the 256x256x3 pictures into a single layered cluster. The next five layers are the densest ones, each with 128, 256, 512, 256, and 128 neurons and the activation function of relu. These five layers serve as the hidden layers, and the output layer is a sigmoid-shaped dense layer with one neuron representing each class.

An attempt to mimic the network of neurons that make up a human brain is known as an artificial neural network. This is done so that a computer can learn things and make decisions like a human. Programming regular computers to behave like interconnected brain cells is how ANNs are created.

I. Import the needed packages
II. Import the data folder
III. Read the images, provide the labels for the image (Set Image having Brain Tumor as 1 and image no having brain tumor as 0) and store them in the Data Frame.
IV. Change the size of images as 256x256 by reading the images one by one.
V. Normalize the image
VI. Split the data set into train, validation and test sets
VII. Create the model
VIII. Compile the model
IX. Apply the model on the train set.
X. Evaluate the model by applying it on the test set.

Convolutional Neural Network (CNN)

Implementing various layers generates the CNN sequential model. The shape of the input image is changed to 256x256. The convolve layer is applied on the information picture with the relu as enactment capability, cushioning as same and that implies the result pictures seems to be the information picture and the quantity of channels are 32,32,64,128,256 for different convolve layers. With 20% of drop out, the maximum pooling that is used with the 2x2 window size and the dropout function is called. The features are converted into a one-dimensional array using the flatten method. The dense method is called with 256 units and rely as the activation function to create the fully connected layer. One unit serves as a representation of the two classes and the sigmoid as an activation function in the output layer [13-14].

i. Import the needed packages
ii. Import the data folder (Yes and No)
iii. Set the class labels for images (1 for Brain Tumor and 0 for No Brain Tumor)
iv. Convert the images into shape(256X256)
v. Split the images into the train, validation and test set images.
vi. Create a sequential model.
vii. Compile the model.
viii. Apply it on the train dataset (use validation set to evaluate the training performance).
ix. Evaluate the model using the test images.
x. Plot the graph comparing the training and validation accuracy.
xi. Draw the confusion matrix for actual output against the predicted output.

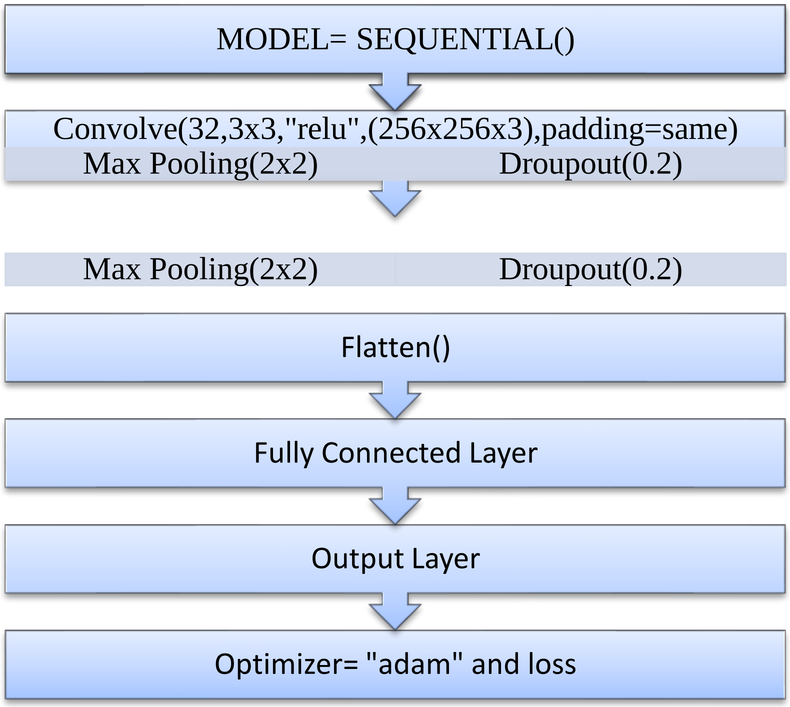

Figure 3. Representation of CNN architecture

4. Experiment and Report Analysis

Implementing various layers generates the CNN sequential model. The shape of the input image is changed to 256x256. The convolve layer is applied to the input image using rely as the activation function, padding as the same, and 32,32,64,128,256 filters for the various convolve layers, making the output image identical to the input image. With 20% of dropouts, the maximum pooling that is used with the 2x2 window size and the dropouts function is called. Straighten strategy is applied to change over the elements into one layered cluster. The dense method is called 256 units and relies as the activation function to create the fully connected layer. The yield layer has 1 unit to address the two classes and the sigmoid as initiation capability.

CNN model's architecture. The CNN sequential model is generated by implementing various layers. The implementation is carried out in Python. The shape of the input image is changed to 256x256. The convolve layer is applied to the input image using rely as the activation function, padding as the same, and 32,32,64,128,256 filters for the various convolve layers, making the output image identical to the input image. With 20% of droupout, the maximum pooling that is used with the 2x2 window size and the droupout function is called. The features are converted into a one-dimensional array using the flatten method.

5. Result and Discussion

During experiments, it was discovered that the proposed method appears to perform better than any other set of images. Among every one of the pictures, the proposed Convolutional Brain Organization (CNN) based approach appears excessively better as far as nature of the result in 128 *128 pictures when contrasted with its other estimated pictures which are ad-dressed in graph format.

Figure 4 represents the percentage of true positive, true negative, false positive, false negative. In this graph, the y-axis represents the percentage of true positive, true negative, false positive, and false negative, while the x-axis represents the different image sizes (128x128, 256x256, and 512x512). Each bar represents a specific performance metric for a particular image size. The higher the bar, the better the performance for that metric.

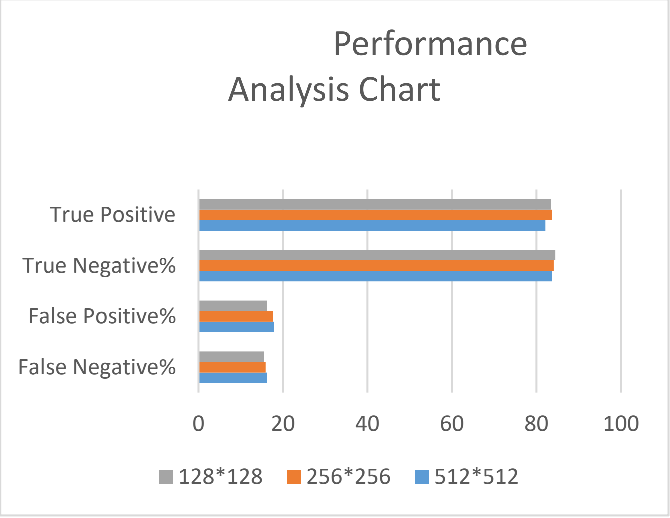

Figure 4. Represents the performance analysis of CNN

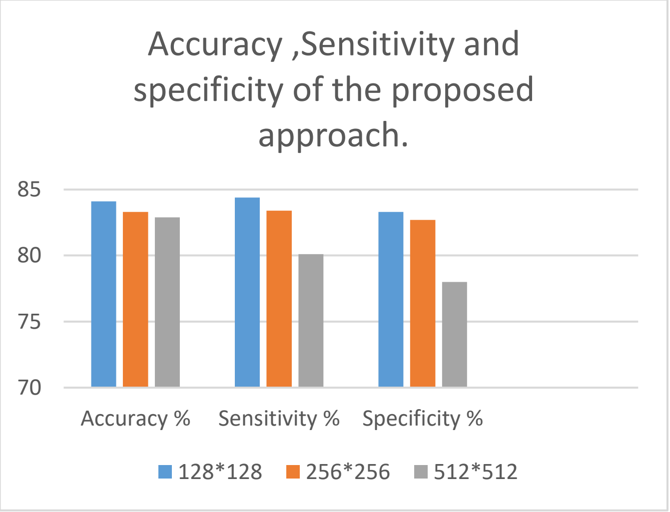

Figure 5. Represents the performance of CNN proposed

Figure 6. Training Loss and Validation Loss

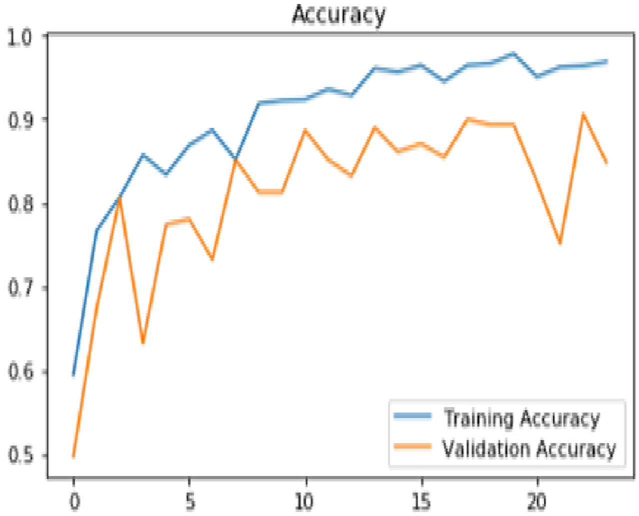

Figure 7. Training Accuracy and Validation Accuracy

Figure 5 represents the Accuracy, Sensitivity, and Specificity of the proposed approach for different sets of images. In this graph, the y-axis represents the percentages of accuracy, sensitivity, and specificity, while the x-axis represents the different image sets (128x128, 256x256, and 512x512). Each bar represents a specific performance metric for a particular image set. The higher the bar, the better the performance for that metric.

In figure 6, the y-axis represents the loss values, and the x-axis represents the number of training epochs. The blue line represents the training loss, and the orange line represents the validation loss. The graph shows how the loss values change over each epoch during training.

In figure 7, the y-axis represents the accuracy values, and the x-axis represents the number of training epochs. The blue line represents the training accuracy, and the orange line represents the validation accuracy. The graph shows how the accuracy values change over each epoch during training.

It is observed from the practical implementation of the model which is being proposed that the particular approach for the tumor region is being correctly identified, and the proposed method can be very useful for the recognition of the growing region of the tumor with almost a negligible difference with the actual data in hand. It has been observed for the 256 x 256 size of the taken MR images, and it seems that the proposed approach has outperformed for a smaller size image then that of the larger size one, the convolution neural network's algorithm is highly adaptable and highly efficient at recognition, which is a benefit. We have divided our data set into two distinct sets—the train set and the test set—each containing magnetic resonance gray scaled images classified as either yes or no to verify that our algorithm works. The magnetic resonance images that contain the brain tumor are labelled as "yes," while the magnetic resonance images that do not contain the brain tumor are labeled "no. "Our accuracy on the train dataset is 91 percent, and our accuracy rate on the test dataset is 89 percent. In our work, CNN gained an accuracy of 99.74%, which is better than the state of the result obtained so far. Our CNN based model will help the doctors to detect brain tumours in MRI images accurately, so that the speed in treatment will increase a lot.

6. Conclusion

A number of studies have been conducted on the particular method of brain tumor detection system with different algorithmic framework and data set approach. It has been a great method of help as when it comes to analyzing image datasets, CNN is regarded as one of the best methods. CNN predicts by shrinking the image without sacrificing the necessary information for prediction. The magnetic resonance images of the brain are used here the tissues present inside the brain such as white matter, gray matter, cerebrospinal fluid and the tissue that is affected due to tumor. We have pre-processed the dataset to remove the undesirable noise present in the magnetic resonance images due to the presence of non-essential and non-brainy things in the picture such as skull, fat, etc. which does not play any role in the determination of brain tumor. From the results we have obtained we find that the proposed method is reliable as it gives fast and near to accurate results for the prediction of brain tumor in the brain of human body. The proposed method would also help in reducing the wastage of time of the doctors, radiologists and clinical experts as this method of detection of brain tumor can be used as a primary measure in detection of tumor. Thus, the proposed method can be used to predict the presence of brain tumor by using the magnetic resonance images (MRI). As of right now, the CNN method is the best one for predicting whether a brain tumor is present in the dataset. We can conclude that the proposed method is trustworthy because it provides quick and accurate predictions of brain tumors in humans' brains.

References

[1]. R. S. Kapse and S. S., Eds., "Salankar2 and Madhuri.B abar3 "Literature Survey on Detection of Brain Tumor from MRI Images," IOSR Journal of Electronics and Communication Engineering, vol. 10, no. 1, pp. 80–86, 2015.

[2]. R. Ahmmed, A. S. Swakshar, M. F. Hossain, and M. A. Rafiq, "Classification of tumors and it stages in brain MRI using support vector machine and artificial neural network," in 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2017.

[3]. S. Aja-Fernandez, C. Alberola-Lopez, and C.-F. Westin, "Noise and signal estimation in magnitude MRI and Rician distributed images: a LMMSE approach," IEEE Trans. Image Process., vol. 17, no. 8, pp. 1383–1398, 2008.

[4]. N. Sravanthi, N. Swetha, P. R. Devi, S. Rachana, S. Gothane, and N. Sateesh, "Brain Tumor Detection using Image Processing," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 7 no. 3, pp. 348-352, 2021.

[5]. S. Damodharan and D. Raghavan, “Combining tissue segmen- tation and neural network for brain tumor detection,” Interna- tional Arab Journal of Information Technology, vol. 12, no. 1, pp. 42–52, 2015.

[6]. M. Alfonse and A.-B. M. Salem, “An automatic classification of brain tumors through MRI using support vector machine,” Egyptian Computer Science Journal, vol. 40, no. 3, pp. 11–21, 2016.

[7]. P. Kumar and B. Vijayakumar, “Brain tumour Mr image seg- mentation and classification using by PCA and RBF kernel-based support vector machine,” Middle-East Journal of Scientific Research, vol. 23, no. 9, pp. 2106–2116, 2015.

[8]. S. N. Deepa and B. Arunadevi, “Extreme learning machine for classification of brain tumor in 3D MR images,” Informatologia, vol. 46, no. 2, pp. 111–121, 2013.

[9]. S. Damodharan and D. Raghavan, “Combining tissue segmen- tation and neural network for brain tumor detection,” Interna- tional Arab Journal of Information Technology, vol. 12, no. 1, pp. 42–52, 2015.

[10]. H. ZainEldin et al., “Brain tumor detection and classification using deep Learning and sine-cosine fitness grey wolf optimization,” Bioengineering (Basel), vol. 10, no. 1, p. 18, 2022.

[11]. C. S. Rao and K. Karunakara, “A comprehensive review on brain tumor segmentation and classification of MRI images,” Multimed. Tools Appl., vol. 80, no. 12, pp. 17611–17643, 2021.

[12]. S. Shanthi, S. Saradha, J. A. Smitha, N. Prasath, and H. Anandakumar, “An efficient automatic brain tumor classification using optimized hybrid deep neural network,” International Journal of Intelligent Networks, vol. 3, pp. 188–196, 2022.

[13]. S. G. Hya, B. K. Giri, and S. T. Satya, “A Novel approach for the detection of tumor in brain MR images and its classification via Independent Component Analysis and Kernel Support Vector Machine,” Imaging Med., vol. 9, no. 3, pp. 33–44, 2017.

[14]. S. N. Deepa and B. Arunadevi, “Extreme learning machine for classification of brain tumor in 3D MR images,” Informatologia, vol. 46, no. 2, pp. 111–121, 2013.

[15]. J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, “Segmentation, feature extraction, and multiclass brain tumor classification,” Journal of Digital Imaging, vol. 26, no. 6, pp. 1141–1150, 2013.