Brain Tumour Detection Using Machine Learning

Conference Paper *in* Journal of Electronics and Informatics · December 2021

DOI: 10.36548/jei.2021.4.005

Manav Sharma
Sharda University

Journal of Electronics and Informatics (ISSN: 2582-3825)
www.irojournals.com/iroei/

Brain Tumour Detection Using Machine Learning

Manav Sharma[1], Pramanshu Sharma[2], Ritik Mittal[3], Kamakshi Gupta[4]

[1,2,3]Department of Computer Science & Engineering, SET, Sharda University, India.
[4]Assistant professor, Department of Computer Science & Engineering, SET, Sharda University, India.

E-mail: [1]2018007623.manav@ug.sharda.ac.in, [2]2018002014.pramanshu@ug.sharda.ac.in,
[3]2018006102.ritik@ug.sharda.ac.in, [4]Kamakshi.gupta@sharda.ac.in

Abstract

This paper presents a model which is based on machine learning algorithms to detect brain tumours from magnetic resonance images with high accuracy. A Convolutional Neural Network (CNN) has been used as the algorithm for feature extraction, and segmentation. The dataset used has been acquired from an internet website. The results show that this technique is promising and the accuracy of 97.79% has been achieved.

Keywords: Image segmentation, CNN, Augmentation, Image classification, MRI

1. Introduction

This research paper has discussed the different stages of brain tumours. Cerebrum cancer division is a significant assignment in clinical picture handling. Early determination of mind cancers assumes a significant part in further developing therapy prospects and expands the endurance pace of the patients. Manual division of the mind growths for disease finding, from the enormous measure of MRI pictures produced in clinical daily practice, is a troublesome and tedious errand. Additionally, mind growth analysis requires an acute level of precision, where a minor mistake in decision making may result in a calamity. Consequently, cerebrum cancer division is difficult for clinical purposes.

Among the right now proposed mind division strategies, cerebrum cancer division techniques dependent on conventional picture handling isn't sufficiently ideal. In customary strategy, an MRI is produced by utilizing attractive field radiation through which a two-dimensional picture (predominantly dependent on a particular dark scale) is created and afterwards that picture is handled and inspected by a clinical expert. This makes a chance of

DOI: https://doi.org/10.36548/jei.2021.4.005
Received: 15.03.2022, received in revised form: 01.04.2022, accepted: 17.04.2022, published: 27.04.2022

human mistake and increments the general danger element of a clinical case which can at times prompt lamentable conceivable outcomes. That is the reason there is a requirement for programmed mind cancer picture division.

Current models that are based on deep learning algorithms are facing a big issue and that is their accuracy. And accuracy plays a crucial role in health care intelligent systems, hence for solving this issue, this model which is highly accurate has been developed.

1.1 Convolutional Neural Network

It is a deep learning algorithm that is used for image processing. This algorithm uses an image as an input and differentiates it on different bases or features.

1.2 Advantages

- Brain tumours are detected from MRI images.
- No human intervention and hence human errors are removed.
- Human life can be saved from earlier detection of the tumour.
- Artificial intelligent systems are more reliable.

1.3 Disadvantages

- System requirements for the proper functioning of the model are high.
- Time taken to train the dataset is high.
- Highly accurate but not completely accurate.

2. Related Work

In [1], the Fuzzy segmentation method (FCM) was applied to separate tumour and non-tumour regions of the brain. Wavelet features were also extracted using a multilevel discrete wavelet transform (DWT). Finally, deep neural networks (DNNs) were incorporated to classify brain tumours with high accuracy. This technique was compared with the methods of KNN classifier, Linear Discriminant Analysis (LDA) and Sequential Minimum Optimization (SMO). The accuracy rate was 96.97 in the DNN-based brain tumour classification analysis. But the complexity was very high and the performance was very poor. In [2], a new biomechanical model of tumour growth was presented for step-by-step analysis of patient tumour growth. It will be applied to gliomas and individual fringed solid tumours to capture a significant tumour effect. Discrete and continuous methods were combined to model tumour growth. The proposed scheme provides the possibility of implicit segmentation of atlas-based registry-based tumour-bearing brain images. This technique was mainly used to segment brain tissue. But the computation time was high.

Paper [3], exploited the new multi-feature feature (Multi FD) and improved the AdaBoost classification scheme used for brain tumour detection and segmentation. Structures of brain tumour tissue were extracted using the Multi FD feature extraction scheme. Advanced AdaBoost classification was used to determine whether the donated brain tissue is tumour tissue or non-tumour. Paper [4], explained a highly complex work that the Local Independent Projection (LIPC) based classifier was used to classify brain voxels. Also, the path function was extracted in this method.

In [5], a new method of segmenting granular tumours using the Cellular Automata (CA) technique was presented, which is compared with the histogram-based segmentation method. Seed selection and volume of interest (VOI) were calculated for efficient segmentation of brain tumours. Segmentation of tumour sections was also incorporated into this work. Thus, the complexity was less but the accuracy was also less. In [6], a brain tumour segmentation method, also known as multimodal brain tumour segmentation diagram was introduced. Also, it combined different segmentation algorithms to achieve high performance compared to the existing method. But the complexity was high.

In [7], studies on brain tumour segmentation were presented. It discussed different segmentation methods like Area Based Segmentation, Threshold Based Segmentation, C means Fuzzy Segmentation, Map-Based Segmentation, Margo Random Field Segmentation (MRF), Modelling formattable, geometry deformable model, Accuracy, robustness, validity, analysed for different types of models. In [8], Hybrid feature selection by ensemble classification was applied to the diagnostic process of brain tumours. It used the GANNIGMAC, decision tree, and bagging C-based wrapper approaches to get the decision rules. It also simplified decision rules with hybrid feature selection that includes a combination of (GANNIGMAC + MRMR C + Bagging C + Decision Tree).

Paper [9], used a Convolutional neural network for their algorithm and the data set used was BRATS 2015. The limitation was that the computation time was high. In [10], the algorithm used was KNN (k- nearest neighbour) but the accuracy achieved while using the proposed method was 62.07%. [11] used a Convolutional neural network for their algorithm and the data set used was BRATS 2015. The limitation was that the computation time was high. In [12], the data set was acquired from the internet website GitHub. 2 different algorithms ANN (Artificial neural network) and CNN (Convolutional neural network) were used. The final result achieved was CNN is more accurate than ANN.

In [13], the dataset was acquired from sources University of Pavia, Kennedy space centre, Indian Pives. The algorithm used was CNN and the accuracy achieved was 88.75%. In [14], the dataset was acquired from The Cancer Imaging Archive" (TCIA). The algorithm used was KNN and classifiers are SVM, RF, LOG, MLP and PCA. The accuracy achieved by the proposed method was 83%. In [15], the dataset was acquired from Fig share Cheng. The algorithm used was CNN (Convolutional neural network) and the accuracy achieved was 84.19%.

3. Data Set

The dataset selected for our model is acquired from the internet. It contains 253 MRI images. In this dataset, there were different folders with a different set of images one with a healthy brain image and the other with a brain tumour image. Then the entire dataset is trained. Took 50 images for validation and 25 for testing. The MRI images present in the dataset are of different dimensions. This dataset is selected because acquiring data sets from hospitals is not a very simple task. The data set used is available on the internet website Kaggle and the link to reach there is provided below.

https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection.

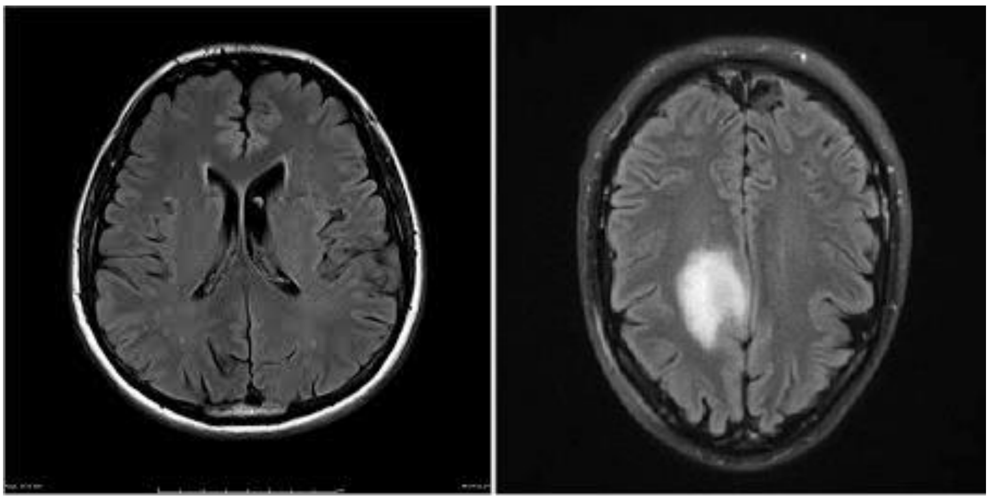


Figure 1. MRI images of the brain without tumour and with tumour

Table 1. Set of images

Number of images	Folder directory
253	Training
25	Testing
50	Validating

- Validation set – It is the set of images that will be used during training for adjusting the parameters.
- Testing set – It is the set of images that will not be involved until checking the final performance of the model.

4. Methodologies

In this section, the basic methodologies used to develop the model are discussed.

4.1 Data acquisition

The data collected had been separated into two categories as healthy and non-healthy ones. Further, the images are of different dimensions so they are converted into the same dimensions of 224*224.

4.2 Pre-processing

In this stage noise removal will be done from the MRI images to increase the accuracy of the model. MRI images often consist of noise which will increase the redundancy and hence decrease the accuracy of the model. There is a high chance of a tumour not getting detected because of the noise present on the borders of an MRI. Hence affects the accuracy of the model. Pre-processing was done by scaling, reducing and converting them into grayscale. Image Pre-processing is done to enhance the quality, look and characteristics of the image.

4.3 Image smoothing

This is an act of simplifying images while preserving important information. The aim is to reduce unnecessary noise or detail without creating too much distortion to simplify subsequent analyses.

4.4 Feature extraction

Feature extraction is the process to extract useful information out of the image. Pixel-based feature extraction is used to extract the information and get it classified into tumour or non-tumour.

4.5 Classification

Classification of brain MRI images from tumour to non-tumour is done using the Convolutional neural network. The classifier used for classification is done by CNN itself. It is highly accurate while dealing with image related datasets. It is used for classifying tumour or non-tumour MRIs.

5. Proposed Work

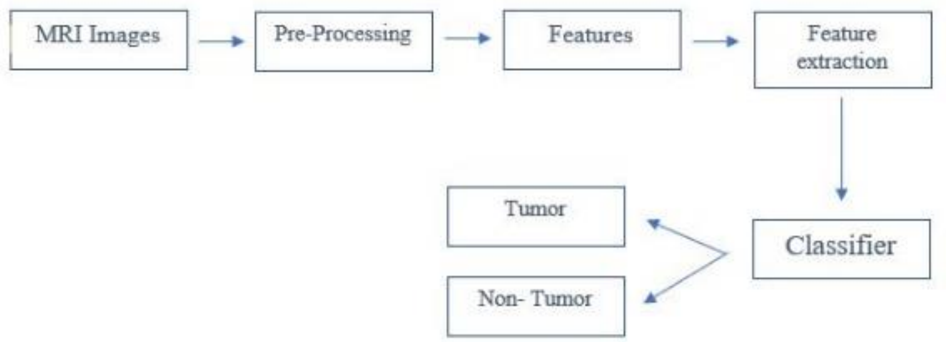


Figure 2. Flow chart of the proposed work

The technique used for making this model work is CNN (Convolutional Neural Network). The different stages by which CNN is applied on the dataset are

1. At first, the required packages are imported.
2. Then the folder where the dataset is stored is imported.
3. Image reading is done after that it is labelled (such as 0 non-tumour and 1 for tumour image) and then the images are stored in the data frame.
4. Then the size of the image is changed to 224*224 and the shape of the image.
5. Image normalization is completed.

6. The data was split into 3 parts –
   - Training
   - Validation
   - Test
7. Sequential model creation is completed.
8. The model compilation is completed.
9. Then the model is applied to the training dataset and a validation dataset was used for evaluating the model.
10. Then the accuracy of the model is evaluated using the test images.
11. The loss graph and the accuracy graph are plotted.

The entire implementation is done using python 3.8 and are executed on Google Colaboratory and the model is stored in Google Drive.

6. Results

When the model is applied to the testing data set for 10 epochs, a validation accuracy of 82.86% is obtained and the validation loss is also less.

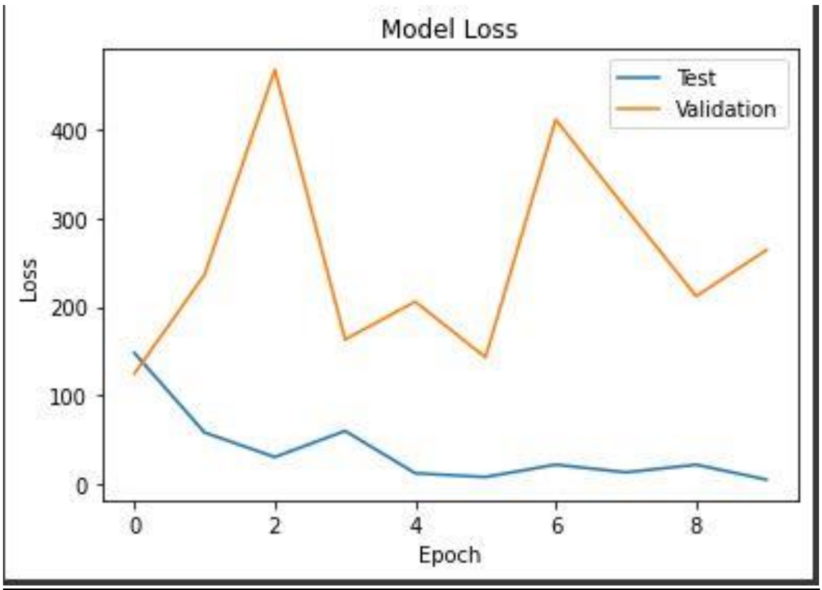


Figure 3. Model loss

As seen in figure 3, when the model is applied to the validation, then a high loss is obtained but once applied to the testing set, the loss gradually decreases with the increasing number of epochs.

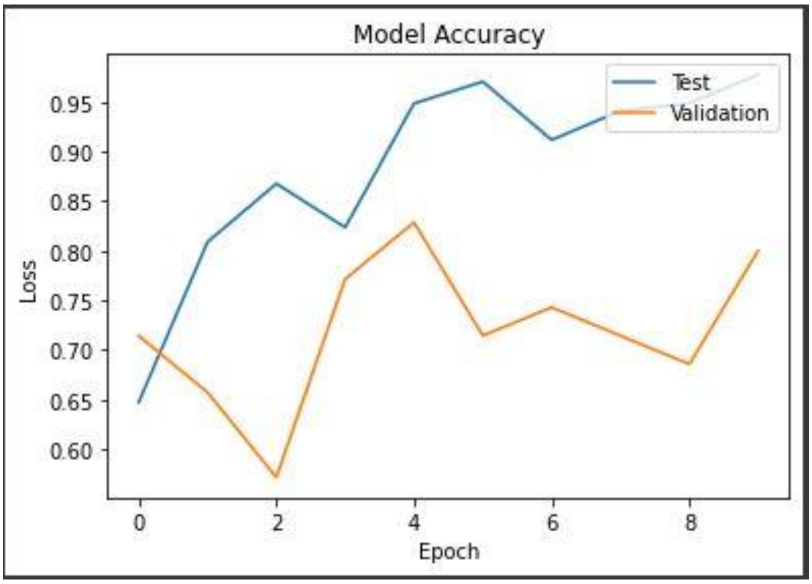
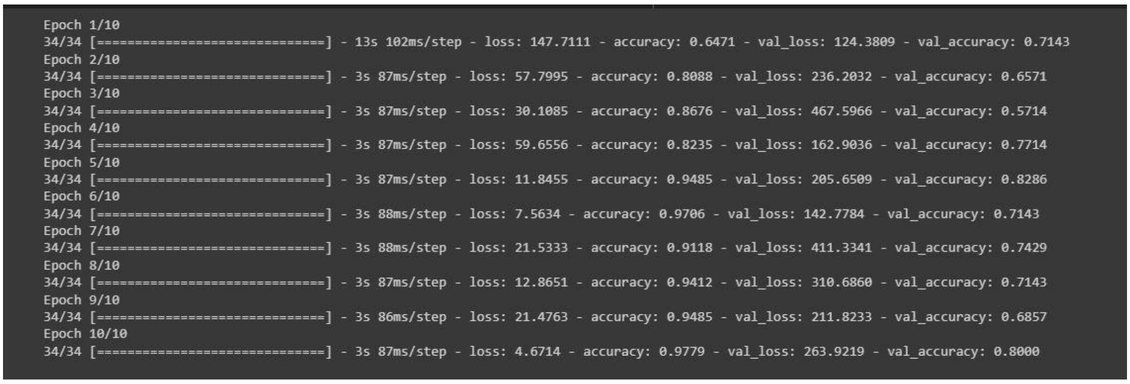


Figure 4. Model accuracy

The accuracy of the convolutional neural network model achieved after applying it to the testing set was 97.79%. with a very minimal loss with increasing epochs. The difference in model accuracy can be seen between the validation dataset and the training dataset in Figure 4.



```
Epoch 1/10
34/34 [==============================] - 13s 102ms/step - loss: 147.7111 - accuracy: 0.6471 - val_loss: 124.3809 - val_accuracy: 0.7143
Epoch 2/10
34/34 [==============================] - 3s 87ms/step - loss: 57.7995 - accuracy: 0.8088 - val_loss: 236.2032 - val_accuracy: 0.6571
Epoch 3/10
34/34 [==============================] - 3s 87ms/step - loss: 30.1085 - accuracy: 0.8676 - val_loss: 467.5966 - val_accuracy: 0.5714
Epoch 4/10
34/34 [==============================] - 3s 87ms/step - loss: 59.6556 - accuracy: 0.8235 - val_loss: 162.9036 - val_accuracy: 0.7714
Epoch 5/10
34/34 [==============================] - 3s 87ms/step - loss: 11.8455 - accuracy: 0.9485 - val_loss: 205.6509 - val_accuracy: 0.8286
Epoch 6/10
34/34 [==============================] - 3s 88ms/step - loss: 7.5634 - accuracy: 0.9706 - val_loss: 142.7784 - val_accuracy: 0.7143
Epoch 7/10
34/34 [==============================] - 3s 88ms/step - loss: 21.5333 - accuracy: 0.9118 - val_loss: 411.3341 - val_accuracy: 0.7429
Epoch 8/10
34/34 [==============================] - 3s 87ms/step - loss: 12.8651 - accuracy: 0.9412 - val_loss: 310.6860 - val_accuracy: 0.7143
Epoch 9/10
34/34 [==============================] - 3s 86ms/step - loss: 21.4763 - accuracy: 0.9485 - val_loss: 211.8233 - val_accuracy: 0.6857
Epoch 10/10
34/34 [==============================] - 3s 87ms/step - loss: 4.6714 - accuracy: 0.9779 - val_loss: 263.9219 - val_accuracy: 0.8000
```

Figure 5. Experimental results

By the use of figure 5, it can be confirmed that the accuracy increases with the increase in the number of epochs and there is a decrease in loss of the testing set.

7. Conclusion

The aim of this paper is to create a model with high accuracy to determine brain tumours from the MRI images. The dataset used consists of 253 brain MRI images and was sufficient to check the performance of the model. The model is based on the machine learning algorithm CNN (Convolutional Neural Network). It helps to predict just by reducing and resizing the image without losing any important information that will be used for predicting. The created model achieves an accuracy of 97.79% when applied to the training set and an accuracy of 82.86% when applied to the validation set. The loss gradually starts decreasing with the increase in the number of epochs. The model loss is very less when applied to the training set whereas it is high when applied to the validation set. In future, different datasets would be applied to this model, to further increase the overall accuracy.

References

[1] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty and A.-B. M. Salem, “Classification using deep learning neural networks for brain tumors,” Future Computing and Informatics Journal, pp. 68-71, 2018.

[2] S. Bauer, C. May, D. Dionysiou, G. Stamatakos, P. Buchler and M. Reyes, “Multiscale modeling for Image Analysis of Brain Tumor Studies,” IEEE Transactions on Biomedical Engineering, vol. 59, no. 1, pp. 25-29, 2012.

[3] Islam, S. M. Reza and K. M. Iftekharuddin, “Multifractal texture estimation for detection and segmentation of brain tumors,” IEEE Transactions on Biomedical Engineering, pp. 3204-3215, 2013.

[4] M. Huang, W. Yang, Y. Wu, J. Jiang, W. Chen and Q. Feng, “Brain tumor segmentation based on local independent projection-based classification,” IEEE Transactions on Biomedical Engineering, pp. 2633-2645, 2014.

[5] Hamamci, N. Kucuk, K. Karaman, K. Engin and G. Unal, “Tumor-cut: Segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications,” IEEE Transactions on Medical Imaging, pp. 790-804, 2012.

[6] H. B. Menze, A. Jakab, S. Bauer, K. Farahani and K. Justin, “The Multimodal Brain Tumor Image Segmentation Benchmark,” IEEE Transactions on Medical Imaging, pp. 1993-2024, 2014.

[7] Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu and Yi Pan, "A survey of MRI-based brain tumor segmentation methods," Tsinghua Science and Technology, pp. 578-595, 2014.

[8] S. Huda, J. Yearwood, H. F. Jelinek, M. M. Hassan, G. Fortino and M. Buckland, "A hybrid feature selection with ensemble classification for Imbalanced Healthcare Data: A case study for brain tumor diagnosis," IEEE Access, pp. 9145-9154, 2016.

[9] J. Seetha and S. S. Raja, "Brain tumor classification using Convolutional Neural Networks," Biomedical and Pharmacology Journal, pp. 1457-1461, 2018.

[10] P. Meena and V. Janani, "IMAGE SEGMENTATION FOR TUMOR DETECTION," A Monthly Journal of Computer Science and Information Technology, pp. 244-248, 2013.

[11] M. Aarthilakshmi, S. Meenakshi, A. P. Pushkala, N. B. Prakash and V. R. Ramalakshmi, "Brain Tumor Detection Using Machine," INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, pp. 1976-1979, 2020.

[12] P. Gokila Brindha, M. Kavinraj, P. Manivasakam and P. Prasanth, "Brain tumor detection from MRI images using Deep Learning Techniques," IOP Conference Series: Materials Science and Engineering, p. 012115, 2021.

[13] Y. Chen, H. Jiang, C. Li, X. Jia and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on Convolutional Neural Networks," IEEE Transactions on Geoscience and Remote Sensing, pp. 6232-6251, 2016.

[14] E. Alberts, G. Tetteh, S. Trebeschi, M. Bieth, A. Valentinitsch, B. Wiestler, C. Zimmer and B. H. Menze, "Multi-modal image classification using low-dimensional texture features for genomic brain tumour recognition," Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics, pp. 201-209, 2017.

[15] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani and T. R. Mengko, "Brain tumour classification using Convolutional Neural Network," IFMBE Proceedings, pp. 183-189, 2018.

Author's biography

Manav Sharma is going to receive a Bachelor's degree in Computer engineering with a specialization in Data Science in the year 2022 from Sharda University. His research area includes machine learning, deep learning and data analytics.

Pramanshu Sharma is going to receive a Bachelor's degree in Computer engineering with a specialization in Data Science in the year 2022 from Sharda University. His research area includes machine learning, deep learning and data analytics.

Ritik Mittal is going to receive a Bachelor's degree in Computer engineering with a specialization in Data Science in the year 2022 from Sharda University. His research area includes machine learning, deep learning and data analytics.

Kamakshi Gupta received a bachelor's degree in Computer engineering from AKTU in 2012 and master's degree from Guru Gobind Singh Indraprastha University in the year 2019. She is currently working as an assistant professor at the department of computer science at Sharda University. Her area of research includes Cryptography and Networking.