AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
E-ISSN: 3069-0110
Vol.4, No. 3(2025)
www.ajmimc.com

"BRAIN TUMOR CLASSIFICATION FROM MRI IMAGES: A HYBRID APPROACH WITH PRE-PROCESSING AND FEATURE EXTRACTION"

Satyam Singh
Student of Computer Science & Engineering (CSE), Millennium Institute of Technology Bhopal, M.P, India.
satyameyt@gmail.com

Parvin Mohane
Professor of Computer Science & Engineering (CSE), Millennium Institute of Technology Bhopal, M.P, India.
pkmohane9910@gmail.com

ABSTRACT

A brain tumor is an abnormal growth of cells within the brain that can disrupt its normal functions. Early detection is vital, as tumors can become cancerous. MRI and CT scans are commonly used for diagnosis, allowing timely treatment. This research aims to identify and locate brain tumors using MRI images through a three-step process: pre-processing, segmentation, and classification. Pre-processing involves converting images to grayscale and removing noise. Segmentation then isolates the tumor region, and feature extraction highlights key attributes for accurate classification.

Keywords— CNN (Convolutional neural network), RFC (Random forest classifier), hybrid model CNN-RFC, MRI (Magnetic-Resonance-Imaging), brain tumor classification

1.INTRODUCTION

The brain, as the control center of the central nervous system, manages numerous functions through complex neural networks. Brain tumors, resulting from abnormal cell growth, can be life-threatening without early detection. Primary tumors originate in the brain, while secondary ones spread from other organs. Diagnosing these tumors using medical imaging is labor-intensive, requiring skilled radiologists to analyze CT, MRI, or PET scans—often taking hours.

To address this challenge, computer-aided diagnostic (CAD) tools have emerged, improving early tumor detection. This study focuses on classifying three brain tumor types: Glioma, Meningioma, and Pituitary Tumor. Gliomas arise from glial cells, Meningiomas from the meninges, and Pituitary tumors affect the pituitary gland, typically in a benign form.

A CNN-based model is proposed to automate feature extraction from MRI images. After this, classification algorithms like SVM, KNN, RFC, DT, and NB are applied. Hybrid models—such as CNN-RFC—combine deep learning with traditional classifiers, enhancing accuracy and efficiency. Among them, the CNN-RFC model shows superior performance by selecting the most relevant features automatically.

2.RELATED WORK

Ghazanfar Latif et al. (2017) proposed a method to classify and segment low- and high-grade Gliomas using multimodal MRI images. They used 3D discrete wavelet transform (DWT) for feature extraction and applied a Random Forest classifier. Segmentation was achieved by reconstructing images from classified blocks.

Xiuming Li et al. (2015) introduced a geodesic level set model combining global and regional intensity through a signed pressure force (SPF) function, improving segmentation by reducing intensity variations and allowing flexible contour initialization.

Aminee Ben Slama et al. (2017) developed a fully automatic vestibular nystagmus (VN) diagnostic method using pupil detection and a geodesic active contour model for accurate pupil segmentation.

A. Shenbagarajan et al. (2016) proposed an MRI brain image classification method involving pre-processing, segmentation via region-based active contours, shape feature extraction, and classification using an ANN with the Levenberg-Marquardt algorithm.

Theodoros N. Sergentanis et al. (2017) conducted a meta-analysis exploring the association between obesity and brain tumors (meningiomas and gliomas), using random-effects models and meta-regression on data from PubMed.

RESEARCH OBJECTIVES

This study aims to improve brain tumor detection by using image segmentation with machine learning for clustering and classification, enhancing accuracy while reducing diagnosis time. The following is the major goal of the research project:

- To segment the brain MRI images with higher accuracy and less error rate
- To classify MR images intonormal, glioma or meningioma with minimal error rate, Random Forest Classifier is used.

4.PROPOSED METHOD

We implement this proposed work in Windows 10 with Intel Core i5 7th Gen and 8 GB of RAM configuration. We used the Google Colab server for better computational power. In the Google Colab server, we have used the GPU as the runtime, 12 GB of RAM, and 100 GB of Disc space for running the code efficiently. We have used some libraries such as TensorFlow version 2.x, Pandas, NumPy, Seaborn, Matplot library, Sklearn library, and Keras.

4.1 Dataset Description

In the fields of deep learning and machine learning, datasets play a critical role in prediction. We will use two Brain Tumor MRI Images Dataset datasets from the kaggle platform. MRI scans of brain tumours are included in this dataset. On these datasets, we used Machine Learning and Deep Learning algorithms, calculating the accuracy of each approach and comparing the results.

4.2 Algorithms used

4.2.1 Logistic Regression

To calculate the results, we employed logistic regression the result of a categorization rule. As a result, the output will have to be discrete or classification. It can be Yes or No, 0 or 1, right or wrong, and so on, but rather than specific numbers like 0 and 1, it offers probability.

Because it can provide estimates and categories additional knowledge using both continuous and discrete datasets, logistic regression is an essential machine learning approach. Logistic regression can immediately identify the most useful classification criteria and can classify results based on a range of data sources.

Using the train test split method of the Scikit learn module, we split each divide the data into two sections: training data and testing data. The Logistic Regression Module is then imported from the Scikit-Learn package. The parameters were then estimated using the fit method of Logistic Regression. It accepts two arguments: training input (xtrain) and training output (ytrain) (ytrain).After that we simply calculated the accuracy of Logistic Regression by using score method and the accuracy of Logistic Regression comes out to be 97.1428% for first dataset and 82.3529% for second dataset.

4.2.2 Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a supervised machine that stands for Support Vector Machine learning approach for classification and regression problems. Its prediction is regarded as one of the most reliable. The notion of Hyperplane is employed in SVM. In n-dimensional space, A hyperplane is a straight line. SVM attempts to find a hyperplane that separates the dataset's classes. The target point is then classified according to its position on the positive or negative side of the Hyperplane.

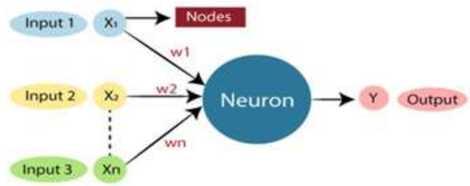
4.2.3 Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) is a type of deep learning algorithm and system that is similar to the human brain. ANN has neurons that are interconnected to each other in multiple layers of networks, similar to how the human brain has neurons coupled to each other. Nodes are the neurons in Artificial Neural Networks. The

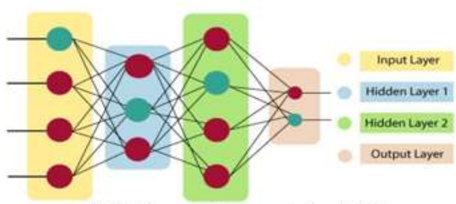
AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
E-ISSN: 3069-0110
Vol.4, No. 3(2025)
www.ajmimc.com

artificial neural network (ANN) enables machines to understand and make decisions in the same way that people do.

The diagrammatic representation of ANN is given in fig(1):



Fig(1): Diagrammatic representation of ANN



Fig(2): Diagrammatic representation of ANN Layers

There are 3 layers in the architecture of an Artificial neural network:

(a) Input Layer: This layer accept the inputs in various formats given by the user.

(b) Hidden Layer: This layer is present between Input layer and Output layer. It is responsible for modifying the input data to find hidden patterns and features.

Output Layer: This layer is responsible for collecting the calculated results from hidden layer and produce the final results which is called as model's prediction. The diagrammatic representation of ANN layers is given in fig(2):

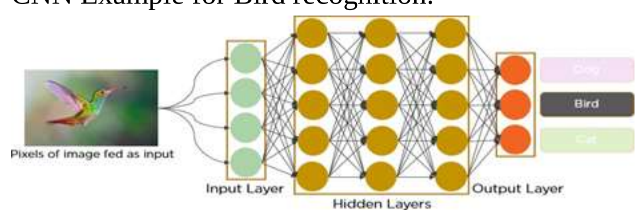
We implemented an ANN using a Sequential model with hyperparameterized input shape and ReLU activation. Trained over 100 epochs, the model achieved 97.96% accuracy on the first dataset and 80.39% on the second.

Convolutional Neural Network(CNN) +Random Forest Classifier(RFC)

For hybrid models such as CNN+RFC, we take our model first layer and last output layer with name dense_1. Our model then predicted 80% of the training dataset and complete testing dataset. After that, we implement each classification model separately such as SVM (Support-Vector-Machine-Classifier), KNN (K-Nearest-Neighbour-Classifier), RFC (Random-Forest-Classifier), DT (Decision-Tree-Classifier), and NB (Naïve-Bayes-Classifier)

Convolutional neural networks are made up of many levels of artificial neurons.

CNN Example for Bird recognition:



Fig(3): Example of CNN Representation

5. RESULT AND DISCUSSION

Our project's purpose is to create a model that improves the accuracy of existing models. The following is a table of the accuracy of all the algorithms used:

For Machine Learning:

Algorithms	**Accuracy on 1st Dataset**	**Accuracy on 2nd Dataset**
Logistic Regression	97.1428%	82.3529%
Support Vector Machine	96.3265%	72.5490%

Table 5.1: Machine Learning Algorithms Accuracy

Deep Learning:

Algorithms	**Accuracy on 1st Dataset**	**Accuracy on 2nd Dataset**
Artificial Neural Network	97.9591%	80.3921%
Convolutional Neural Network	87.096% (Combined)	

Table 5.2: Deep Learning Algorithms Accuracy

The Loss and Accuracy Graphs in Convolutional Neural Network:

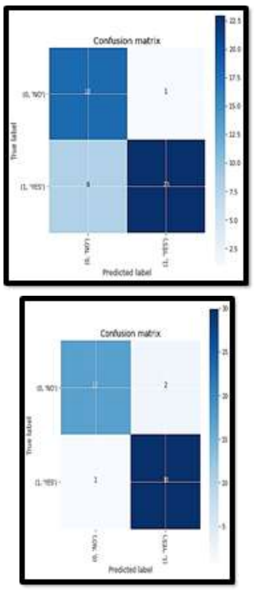


Fig 4 : Loss and Accuracy Graphs in Convolutional Neural Network

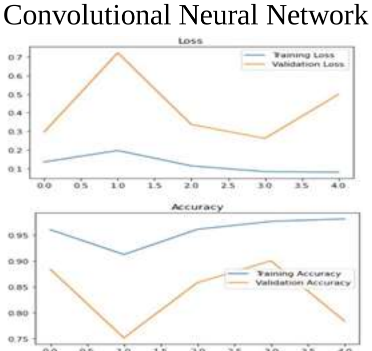


Fig 5: Confusion matrix plots (fromtopleft)–VGG16, ResNet50 and Inceptionv3

6. CONCLUSION AND FUTURE WORK

In this model we used several methodology of Deep Larning and Machine Learning and almost all techniques gives good accuracy in which Artificial Neural Network have highest accuracy of 97.9591%. So this model is very helpful for Doctors in determining accurately that whether the patient has malignat tumor or benign tumor.As our model gives less accuracy for Convolutional Neural Network, so we can work on that part to increase the accuracy of Convolutional Neural Network by increasing the no of epochs, hypertuning factors and several other criterion also. Also we will think to apply tactics of Long Short Term Memory in order to get higher accuracy with the help of Deep Learning concept.

REFERENCES

1. A. Haseena Thasneem, R. Mehaboobathunnisa, M. Mohammed Sathik and S. Arumugam, "Comparison of Different Segmentation Algorithms for Dermoscopic Images", ICTACT Journal on Image and Video Processing, Volume 05,
2. A. Shenbagarajan, V. Ramalingam, C. Balasubramanian and S. Palanivel, "Tumor Diagnosis in MRI Brain Image using ACM Segmentation and ANN-LM Classification Techniques", Indian Journal of Science and Technology, Volume 9, Issue 1, January 2016.
3. Annegreet Van Opbroek, Hakim C. Achterberg, Meike W. Vernooij and Marleen De Bruijne, "Transfer Learning for Image Segmentation by Combining Image Weighting and Kernel Learning", IEEE Transactions on Medical Imaging, Volume 38, Issue 1, January 2019.
4. Annegreet van Opbroek, M. Arfan Ikram, Meike W. Vernooij and Marleen de Bruijne, "Transfer Learning Improves Supervised Image Segmentation across Imaging Protocols", IEEE Transactions on Medical Imaging, Volume 34, Issue 5, May 2015.

5. Annemie Ribbens, Jeroen Hermans, Frederik Maes, Dirk Vandermeulen and Paul Suetens, "Unsupervised Segmentation, Clustering, and Groupwise Registration of Heterogeneous Populations of Brain MR Images", IEEE Transactions on Medical Imaging, Volume 33, Issue 2, February 2014.
6. Ayşe Demirhan and İnan Güler, "Combining stationary wavelet transform and self-organizing maps for brain MR image segmentation", Engineering Applications of Artificial Intelligence, Elsevier, Volume 24, Issue 2, March 2011.
7. Ayşe Demirhan and Inan Guler, "Image Segmentation using Self-Organizing Maps and gray level co-occurrence matrices," Journal of the Faculty of Engineering and Architecture of Gazi University, Volume 25, Issue 2, June 2010.
8. Ayşe Demirhan, Mustafa Toru and Inan Guler, "Segmentation of tumor and edema along with healthy tissues of brain using wavelets and neural networks", IEEE J Biomed Health Informatics, Volume 19, Issue 4, July 2015.
9. Bhumika N. Parmar and Tejas Bhatt, "Volume Visualization using Marching Cubes Algorithms", International Journal of Innovative Research in Technology (IJIRT), Volume 2, Issue 11, April 2016.
10. Boying Wu and Yunyun Yang, "Local- and Global-Statistics-Based Active Contour Model for Image Segmentation", Mathematical Problems in Engineering, Hindawi Publishing Corporation, Volume 2012, 2012.