Disease Detection (Brain Tumor)

R Suganya
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
suganya.ramamoorthy@vit.ac.in
ORCID: /0000-0003-1874-6479

Anuska Ray
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
anuska.ray2020@vitstudent.ac.in

T Subbulakshmi School of Computer Science and Engineering
Vellore Institute of Technology
Cheannia, India
subbulakshmi.t@vit.ac.in

Abstract—In the domain of medical diagnosis, tumor detection is of paramount importance, and machine learning methods have emerged as valuable tools for enhancing accuracy and efficiency. This abstract reviews and compares the performance of four popular machine learning algorithms, namely logistic regression, support vector machines (SVM), decision trees, and random forest classifiers, in the context of tumor detection.
Logistic regression, a widely used linear classifier, estimates the probability of an instance belonging to a specific class. In contrast, SVM, a non-linear classifier, determines the optimal hyperplane for segregating data points of distinct classes.
Recent studies have been scrutinized, focusing on logistic regression, decision trees, and random forest classifiers for tumor detection. Both decision trees and random forest classifiers have exhibited commendable performance in this context, with the latter generally surpassing the former. Additionally, logistic regression and SVM have been utilized and compared, with SVM demonstrating superior accuracy, particularly in dealing with complex datasets. Nevertheless, logistic regression stands out due to its computational efficiency and interpretability.
Moreover, this abstract highlights advances in leveraging multiple machine learning algorithms, such as logistic decision trees and random forest classifiers, to further enhance tumor detection accuracy. These developments hold significant potential for transforming medical image analysis and ultimately improving patient outcomes.

Keywords—Brain Tumor, Image Processing, Malignant, Benign, Magnetic Resonance Imaging(MRI), Pree-Processing ,Region of Interest, SVM, Decision Tree, Machine Learning

I. Introduction

Tumor detection in medical diagnosis is increasingly reliant on machine learning, aiming to enhance early detection and patient outcomes through the analysis of medical images from modalities like MRI, CT, and ultrasound. While deep learning methods, particularly convolutional neural networks (CNNs), have excelled in tumor detection, traditional machine learning algorithms such as logistic regression, decision trees, support vector machines (SVMs), and random forests remain effective.
The primary objective is to classify medical images into normal or abnormal categories, with abnormal images indicating the presence of tumors. Machine learning models are trained on labeled datasets, learning patterns that distinguish normal from abnormal images.
The application of machine learning in tumor detection holds the potential to revolutionize medical diagnosis, improving accuracy and efficiency. However, challenges persist, including data imbalance, tumor shape and size variability, and the need for extensive annotated datasets. This paper reviews the current state of machine learning-based tumor detection, addresses challenges, explores future directions, and highlights recent advancements in this vital field.

II. Literature Survey

This literature survey encompasses five research papers on medical image analysis and disease detection using various machine learning techniques. The papers address the detection of lung cancer in CT images, diagnosis of diabetic retinopathy, detection of brain tumors, and advancements in cancer detection, primarily focusing on lung and breast cancer.
Key findings from these papers include:

- A lung cancer detection system achieved an accuracy of 98.50% using image processing techniques on CT images.
- The diagnosis of diabetic retinopathy achieved an accuracy of 93.33% through the application of convolutional neural networks (CNNs) on the Kaggle Diabetic Retinopathy Detection database.
- Brain tumor detection was improved using a hybrid approach, combining K-Means clustering and Support Vector Machine (SVM) with an accuracy of 95.4%.
- The breast cancer detection model achieved 95.4% accuracy when compared to state-of-the-art models and employed DenseCNN.
- Recent advancements in cancer detection, including lung cancer, showcased high accuracy, recall, and precision, especially through the use of machine learning and image analysis, with accuracy figures reaching up to 98.50%.

Overall, these studies highlight the effectiveness of machine learning and image processing techniques in disease detection, offering promising results for early diagnosis and improving patient outcomes.

979-8-3503-2820-2/$31.00 ©2024 IEEE

III. METHODOLOGY AND IMPLEMENTATION

A. *Image and Dataset Description*

This paper mainly focus on image segmentation of a MRi Scan of brain and tumor dataset from Github.

The original MRI Scan of the brain "Fig 3.1". The tumor dataset which has classified 2 stages of tumor M - malignant and B - Begin.There is total 357 begin tumor stage and 212 malignant tumour stage and there is total 539 rows and 31 columns. The dataset also contains the following information-

Ten real-valued features are computed for each cell nucleus:

- radius (mean of distances from center to points on the perimeter)
- texture (standard deviation of gray-scale values)
- perimeter
- area
- smoothness (local variation in radius lengths)
- compactness (perimeter2 / area — 1.0)
- concavity (severity of concave portions of the contour)
- concave points (number of concave portions of the contour)
- symmetry
- fractal dimension ("coastline approximation" — 1)

B. *Algorithm and Technologies*

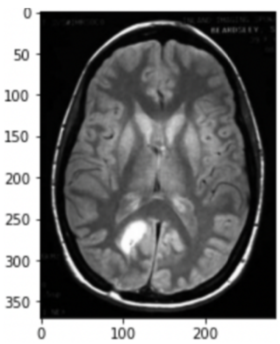


Fig – 3.1

In this paper, four machine learning algorithms are employed: Support Vector Machine (SVM), Decision Tree, Logistic Regression, and Random Forest Classifier along with image segmentation using MRI brain images. These algorithms serve classification and regression tasks. Python, a high-level, interpreted language, is used for programming, and Jupyter Notebook is adopted as an open-source web-based application for interactive code execution and data exploration, allowing for efficient experimentation and collaboration.

C. *Implementation (MRI SCAN)*

We are starting with importing necessary libraries like skimage, skimage. color, matplotlib, pyplot,cv2 and few basic packages. Before applying filters, the image is pre-processed. This may include resizing, converting to grayscale, or denoising. Now, different filters is applied for image segmentation. The image processing is continued with more segmentation techniques like Threshold Segmentation and Gaussian Thresholding. Further the value of threshold is increased to get better features of the tumor so that the stage of the tumor is detected. Therefore with the help of different filtering methods the brain tumor is detected. The output images after filtering can be viewed in "Table 3.2" and the classified tumor is detected.

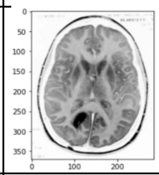
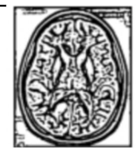
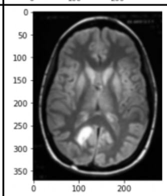
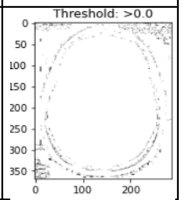
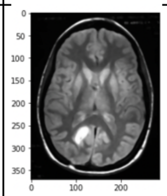
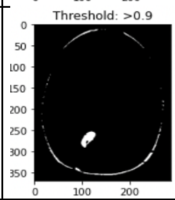
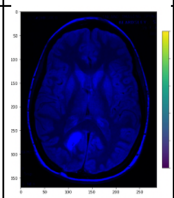
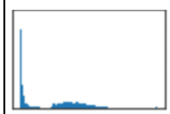
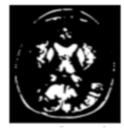
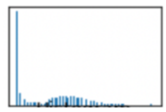
FILTER	OUTPUT IMAGE	FILTER	OUTPUT IMAGE
Negative Filter		Adaptive Mean Thresholding	
Average Filter		increasing the threshold value to get better features of the tumor so that we can understand what stage of the tumor it is in. Threshold >0.0	
Median Filter		Threshold >0.9	
Threshold Segmentation (HSV Format)		Histogram for Qstu's Thresholding	
Global Thresholding(v=127)		Histogram for Thresholding(v=127)	

Table 3.2

D. *Implementation(TUMOR DATASET)*

The tumor dataset is imported and preprocessed "Fig 3.3" and made suitable for applying the Ml algorithms. From the dataset count plot is generated which gives the diagnosis of the count of Malignant = 212 and Begin = 357 and a Correlation Heatmap is generated which can be used to to find potential relationships between variables and to understand the strength of these relationships "Table 3.4".

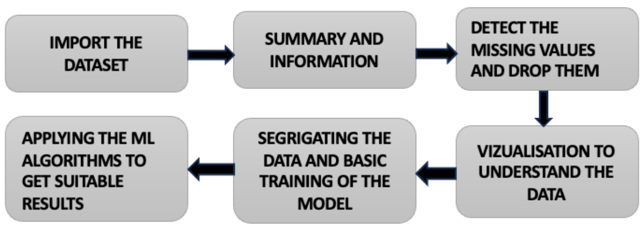


Fig 3.3

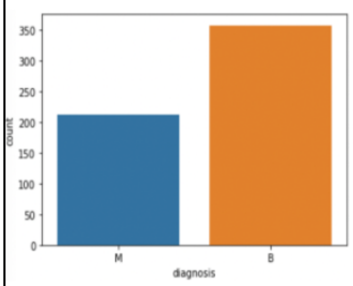
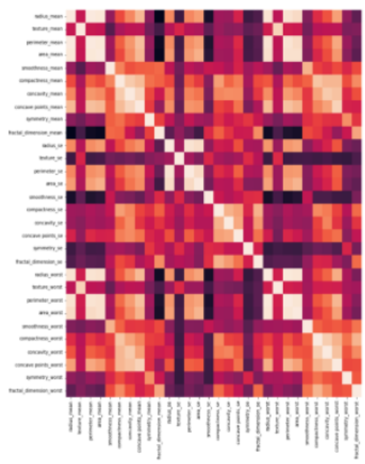
<table>
<tr><th>VISUALISATIONS</th><th>OUTPUT IMAGE</th></tr>
<tr><td>Counter Plot</td><td>
</td></tr>
<tr><td>Correlation Heatmaps</td><td>
</td></tr>
</table>

Table 3.4

To understand more about the dataset and to understand what tupe and features like size ,diameter, area etc. of brain tumor the patient is having four different ML model is incorporated. The accuracy of the models “Table 3.5” are calculated to find the best fit.

S.NO	ALGORITHM	ACCURACY
1.	Logistic Regression Method	0.976608
2.	Decision Tree Classifier Method	0.923977
3.	Random Forest Classifier Method	0.953216
4.	Support Vector Classifier Method	0.976608

Table 3.5

After going through the accuracy of the above-used machine learning algorithms, we can conclude that these algorithms will give the same output every time if the same data set is fed. we can also say that these algorithms majorly provide the same output of prediction accuracy even if the data set is changed.
From the above table, we can conclude that through SVM Model and Logistic Regression Model were the best-suited models for tumor detection.

IV. Conclusion

Brain tumor detection is a crucial aspect of medical diagnosis and treatment, where machine learning algorithms such as logistic regression, SVM, random forest classifiers, decision trees and image segmentation have shown promise when applied to categorical datasets as well as images like MRI scans. The effectiveness of these algorithms relies on factors like dataset quality, size, feature selection, and parameter optimisation. The choice of algorithm should align with specific needs, including accuracy, interpretability, and computational efficiency. However, it's essential to underscore that these algorithms should complement, not replace, expert medical judgment. Trained medical professionals should always interpret results, considering clinical factors beyond what the algorithm captures. In conclusion, while machine learning can significantly assist in brain tumor detection, its application must be responsible and accompanied by clinical expertise. Ongoing research and development hold the potential to enhance algorithm accuracy, ultimately benefiting patient outcomes.

V. Acknowledgment

The author expresses gratitude to Dr. R. Suganya, the course faculty, for guidance and patience, and acknowledges the university's leadership, including Chancellor Dr. G. Viswanathan and Vice-Chancellors. Special mention is made of Dean, Dr. Ganesan R, and HOD, Dr. Siva Balakrishnan M, for support. Thanks extend to faculty, staff, parents, and friends for their encouragement and assistance.

VI. References

[1] Abdur Rafiq”Classification of Tumors and It Stages in Brain MRI Using Support Vector Machine and Artificial Neural Network by Rasel Ahmmed Anirban Sen, Md Swkshar

[2] Prabira Kumar Sethy, Baby Barnali Pattnaik, Srikanta Dash, "Brain Tumour Localization using Image Processing Techniques", 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), pp.753-756, 2022.

[3] Mandala Jagadeesh, P. Chitra, K. Srilatha, M. Sumathi, I. Rexiline Sheeba, "Brain Tumour Classification using CNN Algorithm", 2022 International Conference on Electronics and Renewable Systems (ICEARS), pp.1813-1816, 2022.

[4] Po-Hsiang Tsui, Yin-Yin Liao, Chien-Cheng Chang, Wen-Hung Kuo, King- Jen Chang and Chih-Kuang Yeh, "Classification of Benign and Malignant Breast Tumors by 2-D Analysis Based on Contour Description and Scatterer Characterization", IEEE Transactions on Medical Imaging, vol. 29, no. 2, pp. 513-522, February 2010.

[5] Kimia Rezaei, Hamed Agahi et al., "Segmentation and Classification of Brain Tumor CT Images Using SVM with Weighted Kernel Width", ITCS SIP CST ARIA NLP – 2017, pp. 39-50, 2017.

[6] Nowak RD (1999) Wavelet-based Rician noise removal for magnetic resonance imaging. IEEE Trans Image Process 8(10):1408–1419

[7] Behin A., Hoang-Xuan K., Carpentier A.F., Delattre J.-Y. Primary brain tumours in adults. Lancet. 2003;361(9354)

[8] Hollon T.C., Pandian B., Adapa A.R., Urias E., Save A.V., Khalsa S.S.S., Eichberg D.G., D’Amico R.S., Farooq Z.U., Lewis S., et al. Near real-time intraoperative brain tumor diagnosis using stimulated raman histology and deep neural networks. Nat Med. 2020

[9] Kavitha AR, Chitra L, kanaga R. Brain tumor segmentation using genetic algorithm with SVM classifier. Int J Adv Res Electr Electron Instrum Eng 2016

[10] Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 2009

[11] Ahmad M, Hassan M, Shafi I, Osman A. Classification of tumors in human brain MRI using wavelet and support vector machine. IOSR J Comput Eng 201

[12] Sobhangi Sarkar Avinash Kumar Sabyasachi Chakraborty Satyabrata Aich Jong-Seong Sim Hee-Cheol Kim 2020 A CNN based Approach for the Detection of Brain Tumor Using MRI Scans Test Engineering and Management

[13] Abdusalomov AB, Mukhiddinov M, Whangbo TK. Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. Cancers (Basel). 2023 Aug 18

[14] Ramtekkar, P.K., Pandey, A. & Pawar, M.K. Accurate detection of brain tumor using optimized feature selection based on deep learning techniques. *Multimed Tools Appl* (2023).

[15] A. B. Mitta, A. H. Hegde, A. R. K. P. and G. S., "Brain Tumor Detection: An Application based on Transfer Learning," *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 2023, pp. 1424-1430, doi: 10.1109/ICOEI56765.2023

[16] U. Raghavendra, Anjan Gudigar, Aritra Paul, T.S. Goutham, Mahesh Anil Inamdar, Ajay Hegde, Aruna Devi, Chui Ping Ooi, Ravinesh C. Deo, Prabal Datta Barua, Filippo Molinari, Edward J. Ciaccio, U. Rajendra Acharya,Brain tumor detection and screening using artificial intelligence techniques: Current trends and future perspectives,2023.

[17] Saeedi, S., Rezayi, S., Keshavarz, H. *et al.* MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Med Inform Decis Mak* 23, 16

[18] Takowa Rahman, Md Saiful Islam,MRI brain tumor detection and classification using parallel deep convolutional neural networks,2023

[19] Chetan Swarup, Kamred Udham Singh, Ankit Kumar, Saroj Kumar Pandey, Neeraj varshney, Teekam Singh. Brain tumor detection using CNN, AlexNet & GoogLeNet ensembling learning approaches[J]. Electronic Research Archive, 2023

[20] B. Shetty, R. Fernandes, A. P. Rodrigues and P. Vijaya, "Brain Tumor Detection using Machine Learning and Convolutional Neural Network," *2022 International Conference on Artificial Intelligence and Data Engineering (AIDE)*, Karkala, India, 2022, pp. 86-91, doi: 10.1109/AIDE57180.2022