RESEARCH ARTICLE OPEN ACCESS

Enhancing Brain Tumor Prediction with Deep Learning and Ensemble Machine Learning Approaches

Shambhavi Priya*, Dr. Himanshu Arora**
Department of CSE, Arya College of Engineering, Jaipur, Rajasthan, India

ABSTRACT
This paper introduces the new area of hybrid machine learning for brain tumor prediction. Hybrid approaches attempt to enhance forecast precision, robustness, and explainability by incorporating numerous methodologies. The survey assesses new research papers published, reviews methodology utilized in hybrid models, and compares their performance against conventional methods. The findings highlight the potential of hybrid machine learning for refining brain tumor prediction to improve patient outcomes. it can be seen that multilayer perceptions (MLP) and ensemble methods perform better than one classifier system in the case of brain tumor detection. Among all the models tested, Stacking Ensemble performed the best with an accuracy of 88.51, which reflects its ability to leverage the strength of the base learners and generate more predictions which are more accurate. The MLP also worked exceptionally well with a high accuracy rate of (86.37), precision rate of (86%), and balanced recall rate of (86%), indicating that it possesses good classification capabilities. Other algorithms (XGBoost, Light-GBM, Support Vector Classifier (SVC)) returned similar accuracies of between 87.9 however Light-GBM was slightly more accurate with a precision value of 84. The Voting Ensemble and Random Forest classifiers provided relatively poor but competitive results. In general, the obtained results prove that ensemble methods, where several models are used, can contribute to the improvement of the detection accuracy and the well-balanced performance in terms of precision, recall, and F1-score measurements, which are essential to the effective tumor classification in medical imaging.
Keywords — Diabetic Retinopathy, Deep Learning, CNN, Accuracy.

1. INTRODUCTION

Modern health practices involve Artificial Intelligence (AI), Information Technology (IT), and E-healthcare processes to build a smart system that helps doctors provide quality health services to patients [1], [2]. Brain tumors are a major malfunction in the human brain caused by an aberrant growth of cells. It can adversely affect cognitive function and can be a fatal condition [3], [4]. A brain tumor is a growth of abnormal cells within the brain or central canal of the spine. It often occurs within the nervous system and cerebellum [5] It can be either from the brain tissue itself, also known as primary tumors, or spread from the rest of the body, known as metastatic tumors. Brain tumors are a common form of cancer in human beings, and early diagnosis is critical to reduce mortality rates. Brain cancers have been diagnosed through medical imaging modalities like Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI) [6], [7]. Studies suggest that MRI is a common method used regularly for the diagnosis of brain tumors due to its capability to offer better contrast between MRI images of the brain and cancerous tissues [8].

In clinical practice, the characteristics and grade of a tumor significantly influence treatment decisions. The World Health Organization (WHO) states that malignant brain tumors categorized as Grade I and Grade II are referred to as low-grade gliomas (LGG), whereas those classified as Grade III and Grade IV are termed high-grade gliomas (HGG). The life expectancy of people with high-grade gliomas (HGGs) is approximately one to two years. The typical lifespan of an individual with LGGs ranges from 5 to 10 years. Figure 1 displays a sample of photos depicting various grades of brain tumors.

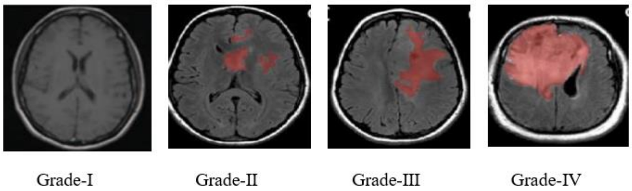

Figure 1: Brain Tumor Grades Images

Brain tumors are classified into two categories: malignant and benign. Benign tumors are non-neoplastic tissues that are harmless and lack the tendency to metastasize to other parts of the body. They can be removed surgically, but there is less chance of recurrence. Benign tumors also tend to have clearly defined borders and are not typically deeply embedded in brain tissue. As such, this feature makes their removal via surgery quite easy. Malignant tumors are cancers that arise from the brain. They grow faster than non-malignant tumors and are very harmful. Metastatic tumors spread to other parts of the body, also known as secondary tumors. This section reviews several methodologies and procedures used by various authors in the application of current brain tumor detection and classification techniques.

Seda Kazdal et al. (2015) [9] designed the detection of brain tumors in MR images. The paper suggested a computer-aided detection method based on reenactment morphological and statistical cancer identification using the morphological features of the region of interest (ROI). This research incorporated many phases that included preprocessing, segmentation, ROI specification, and tumor detection. All in all, 497 MR image slices from 10 patients were used in these processes, achieving a computer-aided detection system accuracy of 84.26%.

The MV-KBC deep approach, which is multi-view knowledge-based collaboration, was introduced in [10] to distinguish between potentially risky and potentially beneficial knobs with limited chest CT data. Their method investigates the features of 3D lung knobs by disassembling a 3D knob from nine different angles. A knowledge-based collaborative (KBC) subsystem is built for every scene based on three kinds of image segments designed to transform three pre-trained ResNet-50 models that capture the global features of the knobs, voxel variability, and shape variability, respectively. They collaborate to delineate lung knobs utilizing the nine KBC sub-models and an adaptable weighting scheme acquired by error backpropagation. This enables the MV-KBC modeling to be systematically created from inception to completion. Utilizing the punitive misfortune role, we attained a significant decrease in the fraudulent valuation with minimal impact on the broader application of the MV-KBC model method. This was executed to enhance the GA's optimal model.

R. Meena Prakash et al. (2019) [11] proposed a CNN-based automated approach for detecting cancer in cerebral images. The extensive image repository of ImageNet was employed for the pre-training of the CNN model. The brain scans were utilized for training this model. The retrieved high-level properties are regarded as identical to the input of the fully integrated layer following the softmax activation. The information obtained from Harvard Medical School contains MR brain pictures utilized for testing the approach. The VGG16, ResNet, and Inception were three pre-trained models utilized for the analysis. The accuracy achieved on the experimental database was assessed at 100%. The results indicated that categorization accuracy improved due to data augmentation.

T. A. Jemimma et al. (2018) [12] examined that one of the arduous difficulties was to identify brain cancer in the domain of clinical imaging. The segmentation of gray matter images, the reduction of cerebral attributes, and the classification of anomalies in the MRI brain image were included in the identification or detection of gray matter tumors. The advanced cancer localization techniques were endorsed by WDAPP-CNN. The malignant region was precisely segmented using watershed analysis. The completed features of the cerebrum were extracted by the dynamic point projection design, and the cancerous and non-cancerous regions of the MRI images of the brain were classified using CNN algorithms. The BRATS dataset was utilized effectively to assess and identify abnormalities in brain images.

2. METHODOLOGY

The methodology section explains the systematic approach, which will be used to design and implement the proposed brain tumor detector system based on deep learning and a graphical user interface (GUI). This methodology aims to combine an advanced deep learning technology in medical image classification with the interactive software application that would be able to support radiologists and healthcare specialists in making their decisions.

The workflow of this project is described in this section and includes the process of datasets acquisition, data preprocessing, deep learning model architecture design, training, evaluation, and the development of the GUI. Visual representations are represented in place by figures and tables, which can subsequently be provided to support the thesis documentation.

3.2 Research Workflow

The methodology is a systematic workflow broken into 6 larger steps:

1. **Dataset Collection and Understanding:** MRI images were obtained using benchmark brain tumor datasets which consist of tumor and non tumor images to create a rich basis to train and evaluate the classification model.
2. **Preprocessing and Data Augmentation:** All the images were normalized and scaled to a consistent size and were pre-processed with normalization and noise removal; data augmentation methods (such as rotation and flipping) were used to enhance data variety and make the model more general.
3. **Model Architecture Design (Deep Learning):** A Convolutional Neural Network (CNN) was created which helped to extract hierarchical characteristics of the MRI images, and layers like convolution, pooling, dropout, and fully connected layers were used in order to effectively classify the images.
4. **Training and Validation Strategy:** The dataset was divided into training, validation and test subsets and the model was trained under the supervised learning strategy with early-stopping and cross-validation strategies to avoid over-fitting and guarantee high performance.
5. **GUI Application Development:** A GUI was created in Python, based on the Tkinter library, to enable the user (e.g., a doctor) to upload MRI images, execute the trained model, and display the classification results in a form accessible to users.
6. **Evaluation and Performance Analysis:** To test the system, the usual measures of accuracy, precision, recall and F1-score, and the confusion matrix analysis were evaluated to thoroughly test the performance of the model and its practical use.

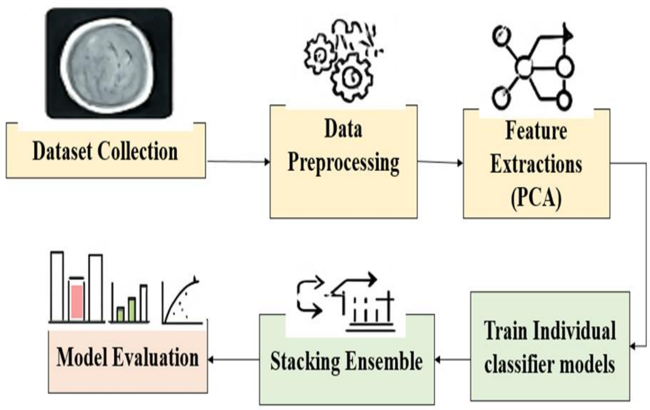

Figure 2: Overall Workflow of the Proposed Methodology

3. EVALUATION MATRICES

Evaluation Metrics

The standard classification measures were used to evaluate the model performance:

- **Accuracy** - Accuracy is the percentage of the correctly classified MRI images to all the samples considered. It measures the overall performance of the model in separating the tumor and non-tumor cases and the tumor subtypes. As much as precision is a simple metric, it might not be enough in unbalanced datasets where certain classes are overrepresented, so adding some metrics is essential. The measure of accuracy should be measured as in equation 1:

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total Number of Samples}} * 100 \quad (1)$$

- **Precision** – Precision is used to measure the accuracy of the positive predictions of the model and is calculated as the number of the true positive classifications against the total of those predicted to be positive. It demonstrates how minimally false positives the model has, which is especially valuable in medical diagnostics to prevent the needless alarm and invasive methods of follow-up. The accuracy is high so that as the model indicates the existence of a tumor, the prediction is accurate. Precision value was determined by equation 2.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}} \quad (2)$$

- **Recall (Sensitivity)** – In recall, which is also known as sensitivity measures the ability of the model to detect real tumor cases with the correct cases by computing the ratio of the true positives to all real positive cases. High recall is very important in clinical practice because it will mean that majority of the tumor patients will not be missed. It is a measure of how well the model eliminates false negatives, which increases diagnostic safety. Recall assessment performed based on equation 3:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}} \quad (3)$$

- **F1-Score** – The harmonic mean of precision and recall is the F1-score, which is a balanced measure considered to be a false positive and false negative. This composite metric as depicted in equation 4 is especially useful in imbalanced dataset where it helps to evaluate the trade-off between precision and recall. The fact that the F1-score is high means that the model is capable of reducing false tumor detection and false alarms on actual tumor cases.

$$\text{F1 Score} = \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} * 2 \quad (4)$$

- **Confusion Matrix** – The confusion matrix provides a class-wise table of all the predictions of the model and their true association with the labels, showing the true positives, true negatives, false positives and false negatives of each type of tumor. Such granular analysis allows the determination of certain misclassification patterns, including confusion of similar types of tumors, to further tune the model and interpret the clinical findings. It can be used as an overall diagnostic system to determine what is good and what is not so good in classification.

4. CONCLUSION

This paper comprehensively delineated the approach employed in this work to develop an effective system for detecting brain tumors using advanced deep learning algorithms and an intuitive graphical user interface (GUI). The entire dataset acquisition process and the unique features of the brain MRI images were clearly explained for a proper and relevant data base for model training and testing. The processes of preprocessing, such as resizing, normalization, noise removal, and data augmentation techniques, were conducted carefully to prepare the input data for normalization and to improve the model's capability to generalize new cases without overfitting. The evaluation techniques were comprehensive, assessing model accuracy, precision, recall, F1-score, and confusion matrix findings, so providing a thorough understanding of the classifier's strengths and flaws. The integration of a powerful CNN model with an intuitive GUI platform establishes a robust framework that not only advances research but also possesses significant potential for real-world application in the early and precise diagnosis of brain cancers. This paper focussed on the experimental data and analysis, illustrating the system's efficacy through quantitative metrics, visualization, and performance comparison with existing methodologies.

REFERENCES

[1] S. AlZu'bi, D. Aqel, and M. Lafi, "An intelligent system for blood donation process optimization-smart techniques for minimizing blood wastages," Clust. Comput. J. NETWORKS Softw. TOOLS Appl., vol. 25, no. 5, pp. 3617–3627, Oct. 2022

[2] Y. Pan, M. Fu, B. Cheng, X. Tao, and J. Guo, "Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform," IEEE ACCESS, vol. 8, pp. 189503–189512, 2020, doi: 10.1109/ACCESS.2020.3026214.

[3] J. Gowrishankar, I. Alam, V. S. Badiger, I. Soni, and M. S. Anwar, "A Hybrid CNN-RNN Model for Accurate Detection of Malignant Brain Tumors," pp. 1776–1781, 2025,

[4] D. Valluru et al., "An Enhanced Early Detection and Risk Prediction Of Brain Tumors Using MRI-CT Scans With Deep Learning Technique," J. Theor. Appl. Inf. Technol., vol. 102, no. 21, pp. 7780–7792, 2024.

[5] Jason, F. Venesius, Y. Sie, R. Fredyan, and H. Pranoto, "Deep Learning-Based Brain Tumor Prediction: An Analysis of Performance Evaluation of Convolutional Neural Network," in 2023 15th International Congress on Advanced Applied Informatics Winter (IIAI-AAI-Winter), 2023, pp. 205–208.

[6] E. Akpinar, B. Hangun, M. Oduncuoğlu, O. Altun, Ö. Eyecioğlu, and Z. Yalçın, "Quantum-Enhanced Classification of Brain Tumors Using DNA Microarray Gene Expression Profiles," 2025.

[7] J. Narasimharao, S. Mubeen, V. Santosh Kumar, M. Soujanya, M. Nagaraju Naik, and A. Veerababu, "A Unique Method Using Deep Learning to Detect Brain Tumors and Performance Enhancement," Lect. Notes Networks Syst., vol. 1229 LNNS, pp. 370–379, 2025.

[8] S. K. Mathivanan, S. Sonaimuthu, S. Murugesan, H. Rajadurai, B. D. Shivahare, and M. A. Shah, "Employing deep learning and transfer learning for accurate brain tumor detection," Sci. Rep., vol. 14, no. 1, p. 7232, 2024.

[9] S. Kazdal, B. Doğan, and A. Y. Çamurcu, "Computer-aided detection of brain tumors using image processing techniques," in 2015 23nd Signal Processing and Communications Applications Conference (SIU), 2015, pp. 863–866.

[10] K. Kamnitsas et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Med. Image Anal., vol. 36, pp. 61–78, 2017.

[11] R. M. Prakash and R. S. S. Kumari, "Classification of MR Brain Images for Detection of Tumor with Transfer Learning from Pre-trained CNN Models," in 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 2019, pp. 508–511..

[12] T. A. Jemimma and Y. J. Vetharaj, “Watershed Algorithm based DAPP features for Brain Tumor Segmentation and Classification,” in 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT), 2018, pp. 155–158.

[13] G. K. Soni, A. Rawat, S. Jain and S. K. Sharma, "A Pixel-Based Digital Medical Images Protection Using Genetic Algorithm with LSB Watermark Technique", Springer Smart Systems and IoT: Innovations in Computing. Smart Innovation, Systems and Technologies, Vol. 141, pp. 483-492, 2020.

[14] S. A. Saiyed, N. Sharma, H. Kaushik, P. Jain, G. K. Soni and R. Joshi, "Transforming portfolio management with AI and ML: shaping investor perceptions and the future of the Indian investment sector," Parul University International Conference on Engineering and Technology 2025 (PiCET 2025), pp. 1108-1114, 2025.

[15] A. Sharma and K. Gautam, "Flood prediction using machine learning technique," 2nd International Conference on Pervasive Computing Advances and Applications (PerCAA 2024), pp. 319-327, 2024.

[16] R. Ajmera et al., “Prediction analysis for diabetic patients using clustered based classification,” Journal of Emerging and Innovative Research, vol. 5, no. 7, pp. 770–775, Jul. 2018.

[17] H. Kaushik, I. Yadav, R. Yadav, N. Sharma, P. K. Sharma and A. Biswas, "Brain tumor detection and classification using deep learning techniques and MRI imaging," Parul University International Conference on Engineering and Technology 2025 (PiCET 2025), pp. 1453-1457, 2025.