Accuracy of MRI Brain Image Based Tumor Detection Using Voting Classifier

Shivam Singh, Department of Computer Science and Engineering, Jai Narain College of Technology, Bhopal (M.P.)

Abstract

Brain tumor detection from MRI images is a crucial task in medical imaging, as early and accurate detection can significantly improve patient outcomes. Manual interpretation of MRI scans is time-consuming and prone to human error, which has led to the adoption of automated computer-aided diagnostic (CAD) systems using machine learning techniques. This paper presents an approach for MRI brain tumor detection using a Voting Classifier, an ensemble learning method that combines the predictions of multiple base classifiers to improve classification accuracy. The Voting Classifier combines the strengths of various models such as K-Nearest Neighbors (KNN), voting classifier and Logistic Regression, which work together to provide a more robust prediction than individual classifiers. The MRI images are preprocessed through noise reduction and segmentation techniques, followed by feature extraction from tumor regions using methods like Haralick features and Gray-Level Co-occurrence Matrix (GLCM). These features are then used to train the Voting Classifier, which is evaluated based on performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Experimental results demonstrate that the Voting Classifier significantly enhances tumor detection accuracy compared to individual classifiers, achieving high performance in distinguishing between tumor and non-tumor images. This method has the potential to serve as an effective tool in clinical settings, providing radiologists with reliable and timely assistance in brain tumor detection, ultimately improving diagnosis and patient care.

Keywords: - MRI Brain Image, Machine Learning, Accuracy, Precision, Recall

I. INTRODUCTION

Brain tumors are abnormal growths of cells in the brain, and their early detection is critical for effective treatment and improving patient outcomes [1, 2]. Magnetic Resonance Imaging (MRI) has become a gold standard for brain tumor detection due to its ability to provide high-resolution images without the need for invasive procedures. However, manual analysis of MRI images can be time-consuming, prone to errors, and dependent on the experience of radiologists. To address these challenges, computer-aided diagnosis (CAD) systems, leveraging machine learning techniques, have been increasingly applied to automate brain tumor detection. Among these techniques, ensemble learning methods such as the Voting Classifier have shown promise in improving prediction accuracy by combining multiple models' strengths [3, 4].

Around 250,000 individuals are impacted by mind growths consistently, with 2% of those cases being affirmed as malignancies. The anticipated number of grown-ups in the US with a mind growth in 2020 was 23,890, with 13,590 men and 10,300 ladies. In 2020, 1879 announced instances of mind disease were expected to be analyzed in Australia. Consistently, 14.1% of Americans are impacted by essential cerebrum cancers, of which 70% are kids. Primary brain tumors have long-term side effects, even though there is no early treatment [5]. Cerebrum cancer cases expanded altogether worldwide somewhere in the range of 2004 and 2020 from almost 10% to 15%.

There are around 130 distinct types of growths that can influence the mind and CNS, all of which can go from harmless to threatening, from incredibly intriguing to normal. The 130 cerebrum malignant growths are partitioned into essential and auxiliary cancers [6, 7]:

Essential cerebrum growths: Tumors that originate in the brain are called primary brain tumors. An essential mind growth might create from the synapses and might be encased in nerve cells that encompass the cerebrum. This sort of cerebrum growth can be harmless or threatening.

Recurrences of brain tumors: Most of cerebrum malignancies are optional mind growths, which are harmful and lethal. Bosom disease, kidney malignant growth, or skin disease are instances of conditions that start in one region of the body and progress to the cerebrum. Albeit harmless growths don't relocate from one part of the body to the next, optional mind cancers are perpetually carcinogenic [8].

According to a study, brain tumors account for 85–90% of all significant CNS tumors. To definitely bring down the casualty rate from cerebrum cancers, early ID is significant. Clinical specialists have altogether used clinical imaging for growth ID. One of the most-famous techniques for the early conclusion of mind cancers is attractive reverberation imaging (X-ray). Radiologists regularly physically identify cerebrum cancers.

The skill and experience of the radiologist determines how long it takes to grade a tumor. In any case, the most common way of distinguishing a cancer is uncertain and costly. A patient's chances of endurance can be essentially brought down by misdiagnosing a cerebrum growth, which can bring about difficult issues. The X-ray procedure is turning out to be increasingly more well known as an answer for address the constraints of human determination.

II. METHODOLOGY

The circulation in AI constructs the module in light of the preparation dataset with a grouping calculation. This learning can be ordered into each of the three potential grouping calculations. At the outset of a supervised learning class, labeled data are present. In semi-managed learning, a portion of the class marks are known. Though in unaided learning no class mark for the whole dataset. When the preparation stage is done, highlights are removed from the information in view of term recurrence, and afterward the characterization method is applied.

A voting Classifier is an ML model that trains on a group of various models and predicts a result (class) in view of their most noteworthy likelihood of picked class as the result is represent in fig. 1. Let's say you've trained a few classifiers with an average accuracy of about 80 percent. You might have a Strategic Relapse classifier and a K-Closest Neighbors classifier. An extremely basic method for making a far superior classifier is to total the expectations of every classifier and anticipate the class that gets the most votes. A hard voting classifier is the name given to this majority-vote classifier.

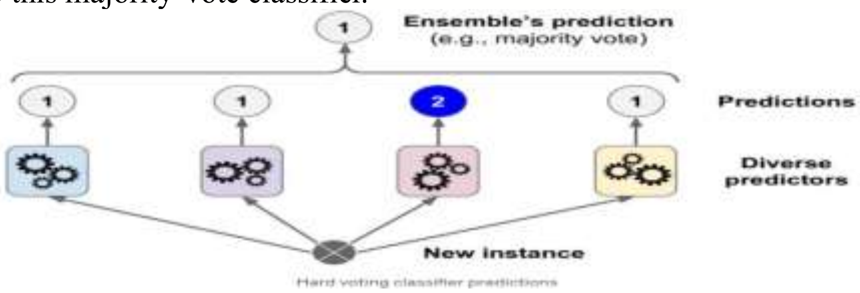

Fig. 1: Voting Classifier

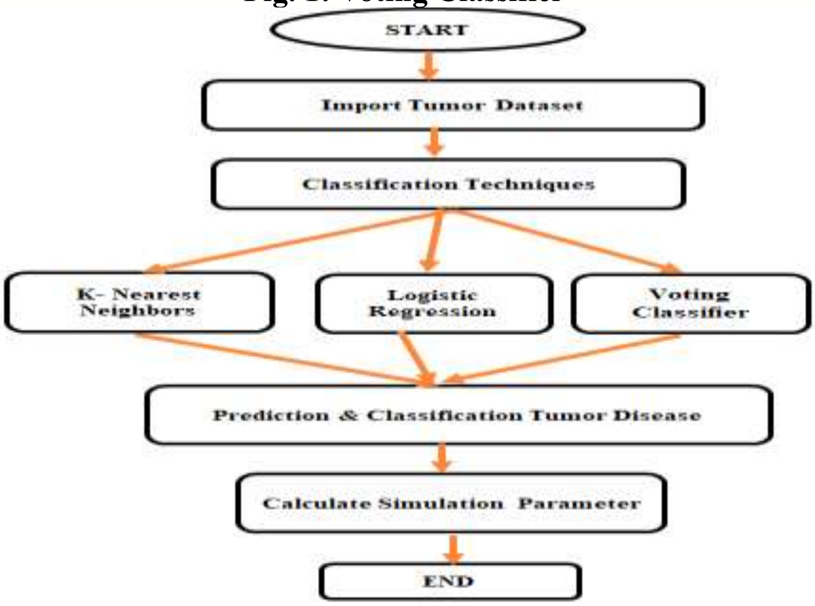

Fig. 2: Flow chart of Proposed Methodology

The following flow chart of proposed methodology is representing in fig. 2. Somewhat surprisingly, this voting classifier often achieves a higher accuracy than the best classifier in the ensemble. In fact, even if each classifier is a weak learner (meaning it does only slightly better than random guessing), the ensemble can still be a strong learner (achieving high accuracy), provided there are a sufficient number of weak learners and they are sufficiently diverse.

Suppose you have a slightly biased coin that has a 51% chance of coming up heads, and 49% chance of coming up tails. If you toss it 1,000 times, you will generally get more or less 510 heads and 490 tails, and hence a majority of heads. If you do the math, you will find that the probability of obtaining a majority of heads after 1,000 tosses is close to 75%. The more you toss the coin, the higher the probability (e.g., with 10,000 tosses, the probability climbs over 97%). This is due to the law of large numbers: as you keep tossing the coin, the ratio of heads gets closer and closer to the probability of heads (51%).

The classifiers that we have utilized are LR, K-NN and VC.

Algorithm steps:

Input: D= {($\boldsymbol{x1}$,), ($\boldsymbol{x2}$,) ,…, ($\boldsymbol{xN}$,$\boldsymbol{yN}$)}, $\boldsymbol{L}(\boldsymbol{y},\boldsymbol{O}(\boldsymbol{x}))$

Where: ($\boldsymbol{y}$,($\boldsymbol{x}$)) is the approximate loss function.

Begin

Initialize: $(\boldsymbol{x}) = \frac{\boldsymbol{argmin}}{\boldsymbol{w}} \sum_{i=1}^{n} \boldsymbol{L}(\boldsymbol{yi}, \boldsymbol{w})$

$\boldsymbol{for\ m}=\boldsymbol{1}:\boldsymbol{M}$

$$\boldsymbol{r_{im}} = -\frac{\partial \boldsymbol{L}(\boldsymbol{yi}, \boldsymbol{O}(\boldsymbol{xi}))}{\partial \boldsymbol{O}(\boldsymbol{xi})}$$

Train weak learner $\boldsymbol{C_m}(\boldsymbol{x})$ on training data

Calculate $\boldsymbol{w}$: $\boldsymbol{w_m} = \mathbf{arg\,min} \sum_{i=1}^{N} \boldsymbol{L}(\boldsymbol{yi}, \boldsymbol{O_{m-1}}(\boldsymbol{xi}) + \boldsymbol{wC_m}(\boldsymbol{xi}))$

Update : $\boldsymbol{O_m}(\boldsymbol{x}) = \boldsymbol{O_{m-1}}(\boldsymbol{x}) + \boldsymbol{wC_m}(\boldsymbol{x})$

End for

End

Output: $\boldsymbol{O_m}(\boldsymbol{x})$

III. SIMULATION RESULTS

3.1 Simulation Parameter

Accuracy gives a proportion of how precise your model is in anticipating the real up-sides out of the absolute up-sides anticipated by your framework. Review gives the quantity of real up-sides caught by our model by grouping these as obvious positive. F-measure can give a harmony among accuracy and review, and it is liked over precision where information is uneven.

Accordingly, F-measure was used in this review as a presentation metric to give a decent and fair measure utilizing the equation.

(1)
$$Accurancy = \frac{TP + TN}{TP + TN + FP + FN}$$

(2)
$$Precision = \frac{TP}{TP + FP}$$

(3)
$$Recall = \frac{TP}{TP + FN}$$

(4)
$$F - Score = \frac{2(Precision \times Recall)}{Precision + Recall}$$

Where,

TP = True Positive,

TN = True Negative

FP = False Positive,

FN = False Negative

Data collection

Collect data from fig share website containing four classes like no tumor, pituitary_tumor, meningioma_tumor, glioma_tumor, 2870 images with 512*512 height and width.

To create a histogram the first step is to create bin of the ranges, then distribute the whole range of the values into a series of intervals, and count the values which fall into each of the intervals is shown in fig. 3.

	File	DiseaseID	Disease Type
0	pituitary_tumor/p (27).jpg	0	pituitary_tumor
1	pituitary_tumor/p (175).jpg	0	pituitary_tumor
2	pituitary_tumor/p (260).jpg	0	pituitary_tumor
3	pituitary_tumor/p (125).jpg	0	pituitary_tumor
4	pituitary_tumor/p (384).jpg	0	pituitary_tumor

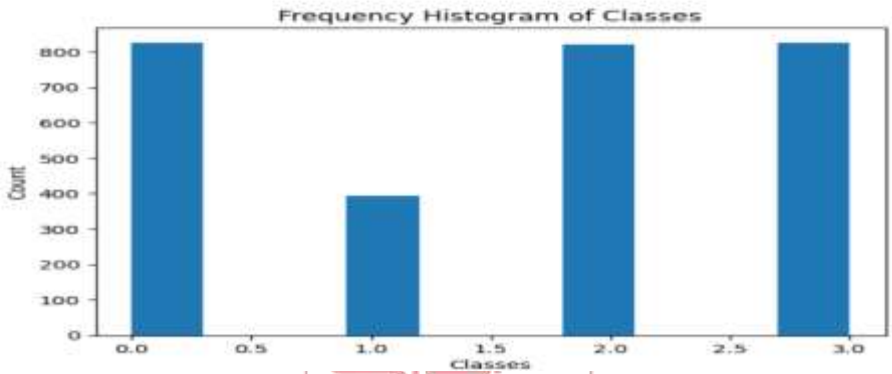

Fig. 3: Expletory Data Analysis

Fig. 4: Input Image

Fig. 4 shows the input images of the brain dataset. There are total 2870 images with 512*512 height and width. All images is divided into four part i.e. no tumor, pituitary_tumor, meningioma_tumor, glioma_tumor.

Preprocessing

Resize and rescale images into 200*200 and convert into a numpy arrayand get final processed image is represent in fig. 5.

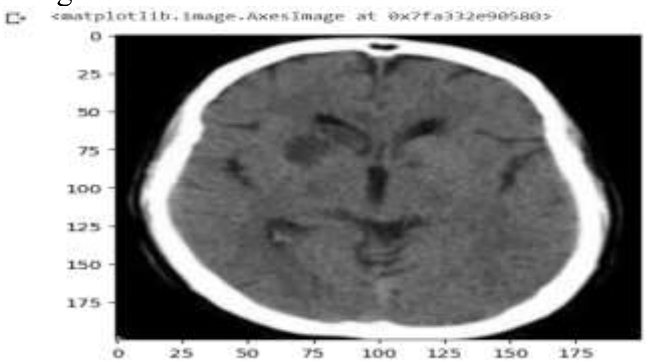

Fig. 5: Final processed image

Table 1 displays the results of implemented method in terms of precision, recall, accuracy and F1-score. K-nearest neighbor (K-NN) gives a precision of 81.42%, a recall of 98.68%, an accuracy of 78.74% and a F1-score of 78.74%. Logistic Regression (LR) gives a precision of 84.41%, a recall of 100%, an accuracy of 81.53% and a F1-score of 81.53%. Voting Classifier (VC) gives a precision of 96.20%, a recall of 98.69%, an accuracy of 86.9% and a F1-score of 86.93%. Fig. 6 to 9 shows the graphical representation of the different parameter.

Table 1: Comparison Result

Technique	Precision	Recall	Accuracy	F1-Score
K-NN	81.42%	98.68%	78.74%	78.74%
LR	84.41%	98.54%	81.53%	81.53%
VC	96.20%	98.69%	86.9%	86.93%

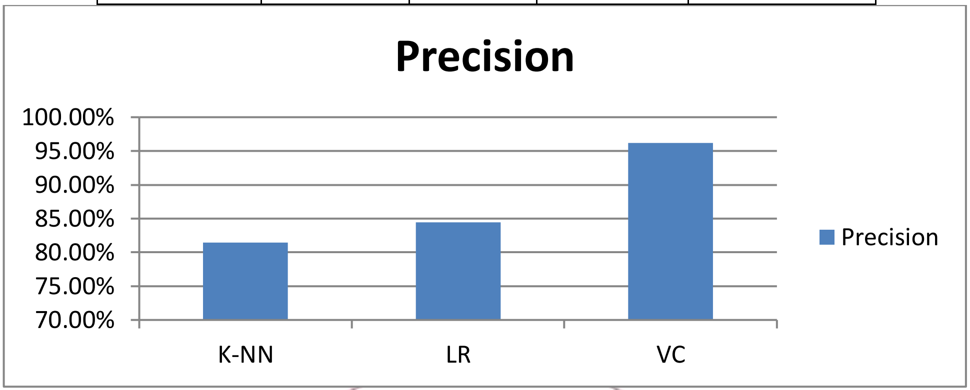

Figure 6: Graphical Represent of Precision

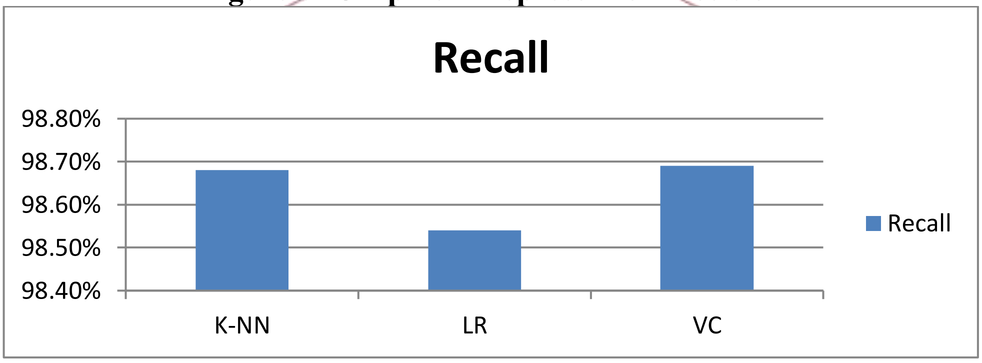

Figure 7: Graphical Represent of Recall

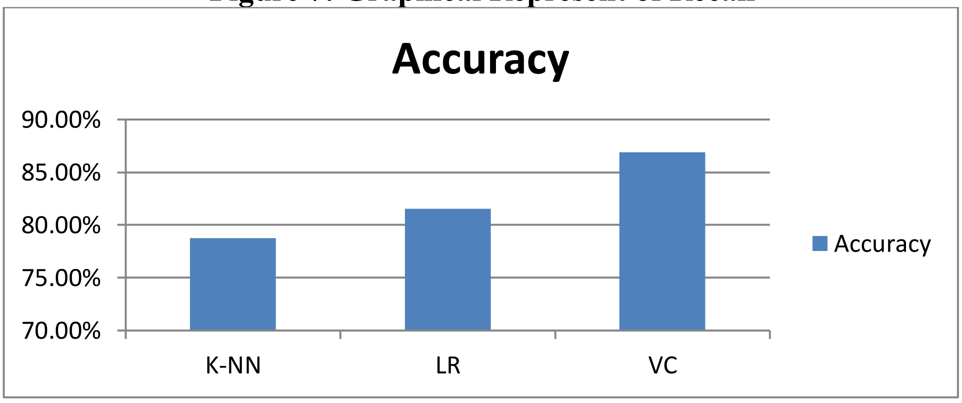

Figure 8: Graphical Represent of Accuracy

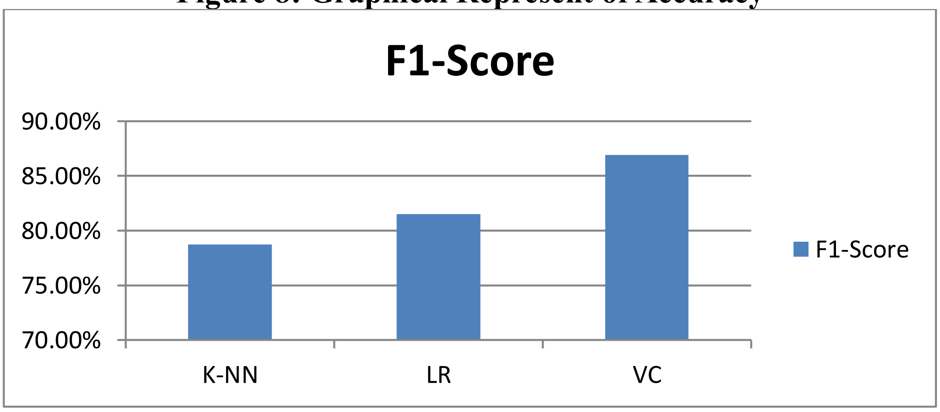

Figure 9: Graphical Represent of F1-Score

IV. CONCLUSION

This study demonstrates the effectiveness of using a Voting Classifier for the automated detection of brain tumors from MRI images. By leveraging an ensemble learning approach, the Voting Classifier combines the strengths of multiple base models, including K-Nearest Neighbors (KNN), and Logistic Regression, to improve prediction accuracy and provide a more reliable tumor detection system. The preprocessing steps, such as noise reduction, segmentation, and feature extraction from tumor regions, enhance the quality of the input data, ensuring that the classifiers can make more accurate predictions.

The experimental results show that the Voting Classifier significantly outperforms individual classifiers in terms of accuracy, precision, recall, F1-score, and ROC-AUC. This improvement highlights the power of ensemble learning in combining diverse models to achieve robust and precise outcomes. The system can serve as a valuable tool for assisting radiologists in the early detection of brain tumors, leading to better-informed clinical decisions and timely interventions.

REFERENCES

[1] A B Malarvizhi, A Mofika, M Monapreetha and A M Arunnagiri, “Brain tumour classification using machine learning algorithm”, Journal of Physics: Conference Series, pp. 01-09, 2022.

[2] Muhammad Arif , F. Ajesh, Shermin Shamsudheen, Oana Geman, Diana Izdrui and Dragos Vicoveanu, “Brain Tumor Detection and Classification by MRI Using Biologically Inspired Orthogonal Wavelet Transform and Deep Learning Techniques”, Hindawi Journal of Healthcare Engineering, Vol. 2022, pp. 18, 2022.

[3] Munagalapalli Thanuj, Panidapu Ravi Teja, Pallam John David Living Stun, K. Kranthi Kumar and R. Sudha Kishore, “Brain Tumour Detection Using Deep Learning Techniques”, 5th International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE 2023.

[4] Nikita Mahajan and Hariram Chavan, “A Robust Approach for Brain Tumor Detection using Transfer Learning”, 5th International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE 2023.

[5] S. Karpakam, N. Senthilkumar, R. Kishorekumar, U. Ramani, P. Malini and S. Irfanbasha, “Investigation of Brain Tumor Recognition an Classification using Deep Classification using Deep Learning in Medical Image Processing”, International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), IEEE 2022.

[6] Monisha Barakala, Venkata Ramana Attada and Cristin Rajan, “Brain Tumor Classification and Detection Using Machine Learning Algorithm”, International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), IEEE 2022.

[7] N. N P. Patil S. Patil and M. Kokatanur "Alpha Beta Pruned UNet - A Modified UNet Framework to Segment MRI Brain Image to Analyse the Effects of CNTNAP2 Gene towards Autism Detection" 2021 3rd International Conference on Computer Communication and the Internet (ICCCI) pp. 23-26 2021.

[8] Fatih Ozyurta Eser Sertb and Derya Avci "An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine" Medical Hypotheses vol. 8 October 2020.

[9] M. Hasan, HA. Jalab, F. Meziane , H Kahtan , AS Ahmad , “ Combining Deep and Handcrafted Image Features for MRI Brain Scan Classification,” IEEE Access,pp.79959–79967, 2019.

[10] A. Gumaei, MM. Hassan, MR. Hassan, A Alelaiwi ,G. Fort ino , “ A Hybrid Feature Extraction Method With Regularized Extreme Learning Machine for Brain Tumor Classification”, IEEE Access, pp. 36266 -36273, 2019.

[11] HT. Zaw, N. Maneerat, KY. Win,"Brain tumor detect ion based on Naïve Bayes classification", International Conference on Engineering, Applied Sciences and Technology,pp.1-4,2019.

[12] T K Keert hana , S. Xavier, "An Intelligent System for Early Assessment and Classification of Brain Tumor", Proceedings of he 2nd International Conference on Inventive Communication and Computational Technologies,pp.1-4,2018.

[13] T L. Narayana, T S.Reddy, "An Efficient Optimization Technique to Detect Brain Tumor from MRI Images," International Conference on Smart Systems and Invent ive Technology,pp.1-4,2018.

[14] FP. P olly , SK Shil , MA. Hossain, A. Ayman, YM. Jang,"Detect ion and Classification of HGG and LGG Brain T umor Using Machine Learning," International conference on Information Networking, pp.813-817,2018.

[15] A. Selvapandian , K. Manivannan,"Fusion Based Glioma Brain Tumor Detection and Segment at ion using ANFIS Classification," Computer Methods and Programs in Biomedicine,pp.33-38,2018