

A Comparative Study of Enhanced Machine Learning Algorithms for Brain Tumor Detection and Classification

This paper was downloaded from TechRxiv (https://www.techrxiv.org).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

23-10-2021 / 26-10-2021

CITATION

GHOSH, ANKIT; KOLE, ALOK (2021): A Comparative Study of Enhanced Machine Learning Algorithms for Brain Tumor Detection and Classification. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.16863136.v1

DOI

10.36227/techrxiv.16863136.v1

A Comparative Study of Enhanced Machine Learning Algorithms for Brain Tumor Detection and Classification

Ankit Ghosh, M.E. Student, School of Nuclear Studies and Application, Jadavpur University, Kolkata, India, Alok Kole, Professor, Department of Electrical Engineering, RCC Institute of Information Technology, Kolkata, India

Abstract—The improvement of Artificial Intelligence (AI) and Machine Learning (ML) can help radiologists in tumor diagnostics without invasive measures. Magnetic resonance imaging (MRI) is a very useful method for diagnosis of tumors in human brain. In this paper, brain MRI images have been analyzed to detect the regions containing tumors and classify these regions into three different tumor categories: meningioma, glioma, and pituitary. This paper presents the implementation and comparison of various enhanced ML algorithms for the detection and classification of brain tumors. A brain tumor is the growth of abnormal cells in the human brain. Brain tumors can be cancerous or non-cancerous. Cancerous or malignant brain tumors can be life threatening. Hence, detection and classification of brain tumors at an early stage is extremely important. In this paper, enhanced ML algorithms have been implemented to predict the presence or the absence of brain tumors using binary classification and to predict whether a patient has brain tumor or not and if he does, detect the type of brain tumor using multiclass classification. The dataset that has been used to perform the binary classification task comprises of two types of brain MRI images with tumor and without tumor. Here nine ML algorithms namely, Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbor (KNN), Naïve Bayes (NB), Decision Tree (DT) classifier, Random Forest classifier, XGBoost classifier, Stochastic Gradient Descent (SGD) classifier and Gradient Boosting classifier have been used to classify the MRI images. A comparative analysis of the ML algorithms has been performed based on a few performance metrics such as accuracy, recall, and precision, F1-score, AUC-ROC curve and AUC-PR curve. Gradient Boosting classifier has outperformed all the other algorithms with an accuracy of 92.4%, recall of 94.4%, precision of 85%, F1-score of 89.5%, AUC-ROC of 97.2% and an AUC-PR of 91.4%. To address the multi-class classification problem, four ML algorithms namely, SVM, KNN, Random Forest classifier and XGBoost classifier have been employed. In this case, the dataset that has been used consists of four types of brain MRI images with glioma tumor, meningioma tumor, and pituitary tumor and with no tumor. The performances of the ML algorithms have been compared based on accuracy, recall, precision and the F1-score. XGBoost classifier has surpassed all the other algorithms in terms of accuracy, precision, recall and F1-score. XGBoost has produced an accuracy of 90%, precision of 90%, and recall of 90% and F1-score of 90%.

Index Terms—artificial intelligence, binary classification, brain tumor classification, comparative analysis, machine learning, MRI image, multi-class classification

I. INTRODUCTION

THE applications of Artificial Intelligence (AI) and Machine Learning (ML) have been expanding at an exponential rate [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. There has been a substantial growth in the applications of AI and ML techniques in various radiological imaging tasks [11], [12]. AI and ML have been playing a major role in uncovering hidden insights into clinical decision making [13], [14]. ML is used in clinical practice including digital pathology and chest, brain, cardiovascular and abdominal imaging. Merging AI and ML with the competency of medical practitioners, it is feasible to enhance the productivity and potentially improve the accuracy [15], [16], [17].

It is essential to promote early diagnosis of brain tumors because they are the most common cause of cancer-related deaths in children and people up to 40 years of age. Therefore, it is necessary to devise strategies to accelerate early diagnosis of brain tumors. An early diagnosis of brain tumor implies faster response in treatment, thereby increasing the surviving rates of patients. A process designed to automatically detect, locate and classify brain tumors is desirable. AI and ML have gained prominence in almost every field of decision-making and can be successfully implemented for the detection and classification of brain tumors.

The objective of this paper is to investigate the use of ML classification algorithms to detect the presence of brain tumors and also distinguish between different types of brain tumors such as glioma, meningioma and pituitary tumors from brain MRI images. A computer-aided classification method is more reliable for brain tumor diagnostics. The proposed scheme involves a few steps including data collection, data preprocessing (data labeling and image pre-processing), classification based on enhanced ML techniques and finally a comparative analysis of the implemented models.

II. RELATED WORK

A lot of research work has been done in the field of Artificial Intelligence (AI) and Machine Learning (ML) application in the field of medical imaging. Noreen et al. [18] have proposed the use of two pre-trained deep learning models i.e. Inception-v3 and DensNet201 for developing a multi-level feature extraction and concatenation method for the early detection of brain tumors and their classification. At first, they have extracted the features from different Inception modules from the pre-trained Inception-v3 model. Then they have passed those features to the softmax classifier to perform the classification of the brain tumors. Secondly, they have used a pre-trained DensNet201 to extract features from various DensNet blocks. Then they have concatenated those features and passed them to the softmax classifier to classify the brain tumors. The dataset that they have used comprised of three classes of brain tumors and it is available publicly. Their proposed methodology has produced exceptional results and has outperformed all the existing state-of-the-art ML and Deep Learning (DL) models for brain tumor detection and classification.

In [19] Naik and Patel have used the decision tree classification algorithm for the detection and classification of brain tumor from MRI images. In the pre-processing step they have used the median filtering process and texture feature extraction technique has been used to extract the features. Their proposed model has exhibited improved efficiency in comparison to the traditional image mining methods. The results that they have obtained have been compared with the Naïve Bayesian classification algorithm. The decision tree classification algorithm has achieved a precision of 100%, Sensitivity of 93%, Specificity of 100% and Accuracy of 96%.

Tandel *et al.* [20] have proposed a transfer-learning-based AI paradigm using a Convolutional Neural Network (CNN) for brain tumor classification using MRI data. The transfer-learning-based CNN model has been benchmarked against six different ML classification algorithms, namely Decision Tree, Linear Discrimination, Naive Bayes, Support Vector Machine, K-nearest neighbour and Ensemble. Their proposed model has proven to be very useful in multiclass brain tumour grading and has yielded better results in comparison to the other ML models.

Sarhan [21] has presented a computer-aided detection (CAD) technique for the classification of brain tumors in MRI images. The features from the brain MRI images have been extracted by utilizing the Discrete Wavelet Transform (DWT). The extracted features have then been applied to a CNN to classify the input MRI image. His proposed approach has produced an overall accuracy of 98.5%.

Mohsen *et al.* [22] in their research work have proposed the development of a Deep Neural Network (DNN) classifier for the classification of brain tumors on a dataset comprising of 66 brain MRI images of 4 types of brain tumors, namely, normal, glioblastoma, sarcoma and metastatic bronchogenic carcinoma tumors. They have combined the classifier with DWT for feature extraction and principal components analysis (PCA). The DNN classifier yielded extremely good results with an average classification rate of 96.97%, average recall of 0.97, average precision of 0.97, average F-Measure of 0.97 and average area under the ROC curve (AUC) of 0.984 of all

the four classes (normal, glioblastoma, sarcoma and metastatic bronchogenic carcinoma tumors).

In [23] Rehman *et al.* have conducted three studies using three architectures of convolutional neural networks (AlexNet, GoogLeNet, and VGGNet) to perform the classification of brain tumors such as meningioma, glioma, and pituitary. Then they have explored the transfer learning techniques, i.e., finetune and freeze using MRI slices of brain tumor dataset—Figshare. They have applied data augmentation techniques to the MRI images to generalize the results, increase the dataset samples and reduce the chance of over-fitting. The proposed fine-tune VGG16 architecture has attained the highest accuracy up to 98.69% in terms of classification and detection.

III. DATASETS

Two different datasets have been used:

- 1. Dataset-A (binary classification)
- 2. Dataset-B (multi-class classification)

A. Dataset-A (binary classification)

Dataset-A has been used for binary classification. It comprises of 982 brain MRI images of patients with tumor and 493 images with no tumor. Thus, a total of 1475 images are present. A collection of 18 brain tumor images from Dataset-A are shown in Fig.1.

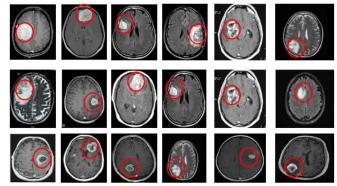


Fig. 1. Collection of 18 brain tumor MRI images from Dataset-A

15 images with no brain tumor from Dataset-A are shown in Fig.2.

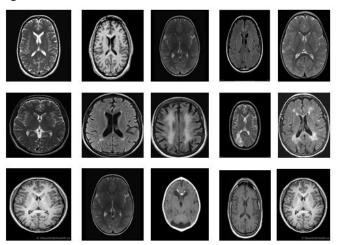


Fig. 2. Collection of 15 brain MRI images with no tumor from Dataset-A

B. Dataset-B (multi-class classification)

Dataset-B has been used for multi-class classification. It consists of 826 brain MRI images of glioma tumor, 822 images of meningioma tumor, 827 images of pituitary tumor and 395 images with no tumor. Thus, a total of 2870 images are present. A collection of 8 glioma tumor images from Dataset-B are shown in Fig.3.

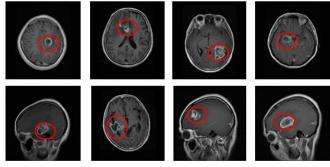


Fig. 3. Collection of 8 glioma brain tumor MRI images from Dataset-B

Fig.4 shows a collection of eight meningioma brain tumor images from Dataset-B.

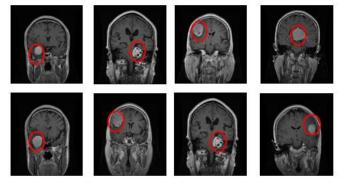


Fig. 4. Collection of 8 meningioma brain tumor MRI images from Dataset-B

A collection of eight pituitary brain tumor images from Dataset-B are shown in Fig.5.

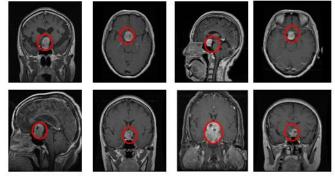


Fig. 5. Collection of 8 pituitary brain tumor MRI images from Dataset-B

Fig.6 shows a collection of eight images with no tumor.

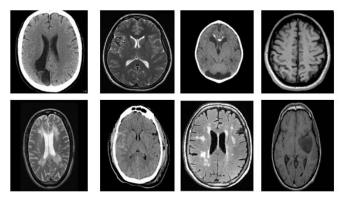


Fig. 6. Collection of 8 brain MRI images with no tumor from Dataset-B

The presence of a white spot as marked by the red circles in Fig. 1, 3, 4, and 5 is an indication of the abnormal growth of tissues in the human brain. There is an aggregation of abnormal cells in some tissues of the brain in the above mentioned section. The tumors in these images are quite critical. Manual classification of these images is rather difficult. So, ML algorithms have been employed for the efficient detection and classification of brain tumors from these MRI images.

IV. STATE-OF-THE-ART MACHINE LEARNING ALGORITHMS

In this section some of the state-of-the-art ML algorithms that can be used for the detection and classification of brain tumors have been discussed.

Support Vector Machine (SVM) is a supervised ML classification algorithm designed by a separative hyper-plane. The main purpose of the SVM is to segregate the data in the best possible way [24]. Therefore, SVM is a frontier which best segregates the two classes. Logistic Regression is a supervised ML algorithm which is used to predict a binary outcome based on a set of independent variables. The main aim of logistic regression is to find the best fitting model to describe the relationship between the outcome and a set of predictor variables [25]. K-Nearest Neighbor (KNN) is a supervised ML algorithm. It is used to solve binary classification problems. KNN predicts whether a given data point belongs to a particular class or the other by calculating the distance between the given data point and the other points [26]. The given data point belongs to that class whose data points are nearest to it. K in KNN refers to the number of points to be selected in the vicinity of the given data point. Naïve Bayes (NB) is a supervised ML algorithm used mostly for binary classification. It is based on Bayes' theorem with an assumption of independence among the predictors [27]. NB classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Decision Tree (DT) is a supervised learning model which is used to solve binary classification problems. DTs learn from simple decision rules inferred from the data features and predict the value of a target variable [28]. Random Forest is an ensemble ML algorithm which builds multiple DTs and then merges those together [29]. Hence, it produces results which are more accurate. In case of DTs, if the dataset is too large

there are chances of over-fitting. So, Random Forest is used to avoid over-fitting of data. Random Forest can be used to solve both classification and regression problems. Stochastic Gradient Descent (SGD) is an efficient optimization algorithm which minimizes the cost function by altering the values of the parameters or coefficients of a function [30]. SGD Classifier implements a SGD learning routine to support various loss functions to perform classification tasks. Extreme Gradient Boosting, XGBoost is a member of the family of boosting algorithms. It is an efficient implementation of the Gradient Boosted Trees algorithm which is a supervised learning method. It is an ensemble ML technique and uses Gradient Boosting framework for prediction [31]. Boosting is an ensemble learning technique. It combines predictors with low accuracy and converts them into a model with an improved accuracy [32]. In gradient boosting the errors made by the predecessors is corrected by the predictor itself resulting in a strong model with high accuracy.

V. METHODOLOGY

The proposed methodology has been illustrated in Fig.7.

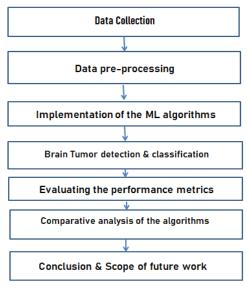


Fig. 7. Flowchart of the proposed workflow

A. Methodology (binary classification)

In this section, the methodology that has been used to predict whether a patient has brain tumor or not from the brain MRI images in Dataset-A has been described.

1) Data pre-processing

Data labeling

The images of brain tumor have been labeled as '1' and the images with no brain tumor as '0'.

Image pre-processing

The images have been read in the gray scale (2D). To build a classifier using ML algorithms all the images have been converted into the same dimension. So, each image has been resized into 200*200 pixels.

For instance, the original image as shown in Fig. 8 has a dimension of 630*630 pixels. Its dimension has been transformed into 200*200 pixels. Similarly, each and every

image in the dataset has been converted into a dimension of 200*200 pixels.

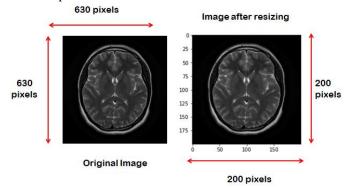


Fig. 8. Image pre-processing (Dataset-A)

2) Splitting the dataset

The entire dataset has been split into training and testing data with a test size of 25%.

3) ML models used

The following ML algorithms have been implemented to perform the binary classification task:

- SVM
- Logistic Regression
- KNN
- NB
- DT
- Random Forest
- SGD classifier
- XGBoost
- Gradient Boosting classifier

B. Methodology (multi-class classification)

The methodology for multi-class classification of the brain MRI images in Dataset-B has been described in this section.

1) Data pre-processing

Data labeling

The images with no tumor have been labeled as '0', images of glioma brain tumor as '1', images of meningioma as '2' and pituitary tumor as '3'.

Image pre-processing

Each and every image in Dataset-B has been resized into a dimension of 200*200 pixels. For instance, the original image shown in Fig.9 has a dimension of 350*350 pixels. It has been converted into a dimension of 200*200 pixels.

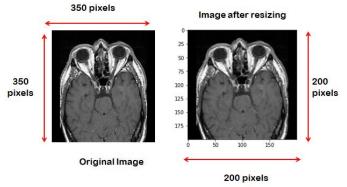


Fig. 9. Image pre-processing (Dataset-B)

2) Splitting the dataset

75% of the dataset has been used for training and the remaining 25% has been used for testing purposes.

3) ML models used

The following ML algorithms have been implemented to perform the multi-class classification task:

- SVM
- KNN
- Random Forest
- XGBoost

VI. EXPERIMENTAL RESULTS AND DISCUSSION

A. Binary classification results

A comparative analysis of the nine algorithms has been done based on the following performance metrics:

- Accuracy
- Recall
- Precision
- F1- score
- Area Under Curve —Receiver Operating Characteristics (AUC-ROC)
- Area Under Curve Precision Recall (AUC-PR)

For the evaluation of the accuracy, recall, precision and F1 score, the following 4 attributes have been used in the measurement:

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

The attributes for determining the performance metrics for each of the 9 ML algorithms has been demonstrated in Table I.

Table I
Attributes for the calculation of performance metrics of the proposed ML algorithms

Algorithm	TP	TN	FP	FN
SVM	95	219	24	31
Logistic	92	220	23	34
Regression				
KNN	102	210	33	24
NB	79	205	38	47
DT	104	215	28	22
Random	120	215	25	6
Forest				
SGD	88	222	21	38
XGBoost	112	218	25	14
Gradient	119	222	21	7
Boosting				

From the values of the attributed described in Table 1, it can be observed that Gradient Boosting classifier has detected 119 data as true positives and 222 true negatives accurately. It has misclassified only (21+7=) 28 data.

Accuracy= (119+222)/(119+222+21+7) = 0.924

Misclassification= (21+7)/(119+222+21+7) = 0.076

Therefore, it can be concluded that Gradient Boosting has exhibited an accuracy of 92.4% and misclassification of 7.6%.

To get a clear picture, the evaluation metrics of all the ML algorithms have been computed and tabulated. The performance comparative analysis of the different ML algorithms based on accuracy, precision, recall and F1-score have been depicted in Table II.

Accuracy= (TP+TN)/ (TP+FP+FN+TN)

Recall= TP/(TP+FN)

Precision= TP/ (TP+FP)

F1-score= (2*Precision*Recall)/ (Precision + Recall)
Table II

Comparison of the different ML algorithms based on accuracy, recall, precision and F1-score

model	accuracy	recall	precision	F1-score
SVM	0.851	0.754	0.798	0.775
Logistic	0.846	0.730	0.800	0.763
Regression				
KNN	0.845	0.804	0.756	0.779
NB	0.769	0.627	0.675	0.650
DT	0.865	0.825	0.788	0.806
Random	0.908	0.952	0.811	0.876
Forest				
SGD	0.840	0.709	0.807	0.755
XGBoost	0.894	0.888	0.818	0.852
Gradient	0.924	0.944	0.850	0.895
Boosting				

From the test results shown in Table 2, it can be inferred that Gradient Boosting classifier has exhibited the highest accuracy among all the other ML models that have been implemented. Gradient Boosting has produced an accuracy of 0.924, recall of 0.944, precision of 0.850 and F1-score of 0.895. Random Forest follows Gradient Boosting with an accuracy of 0.908, recall of 0.952, precision and F1-score of 0.811 and 0.876 respectively. After Gradient Boosting and Random Forest, XGBoost classifier has produced the third highest accuracy of 0.894. It has a recall, precision and F1-score of 0.888, 0.818 and 0.852 respectively. DT classifier has an accuracy, recall, precision and F1-score of 0.865, 0.825, 0.788 and 0.806 respectively. SVM has exhibited an accuracy of 0.851, recall of 0.754, precision of 0.798 and F1-score of 0.775. Logistic Regression has generated an accuracy of 0.846 and recall, precision and F1-score of 0.730, 0.800 and 0.763 respectively. KNN has an accuracy of 0.845 which is less than that of Logistic Regression by only 0.001. It has produced recall, precision and F1-score of 0.804, 0.765 and 0.779 respectively. SGD classifier follows KNN with an accuracy of 0.840, recall of 0.709, precision of 0.807 and F1-score of 0.755. NB classifier has been the least accurate with an accuracy of 0.769. It has produced a recall, precision and F1-score of 0.627, 0.675 and 0.650 respectively. Thus, it can be said that the performances of all the classification algorithms have been satisfactory with Gradient Boosting outperforming all the other classifiers.

To enable a clear interpretation of the binary classification models, ROC curves and PR curves have also been plotted. This has been described in the following sections.

ROC curves of the different algorithms for performance comparison

The ROC curves corresponding to SVM, Logistic Regression, KNN, NB, DT classifier, Random Forest classifier, SGD classifier, XGBoost classifier and Gradient Boosting classifier have been shown in Fig. 10, 11, 12, 13, 14, 15, 16, 17 and 18 respectively.

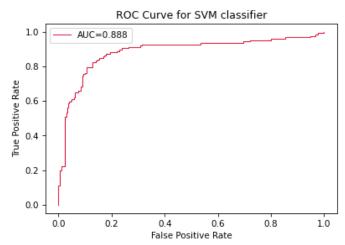


Fig. 10. ROC curve of SVM classifier

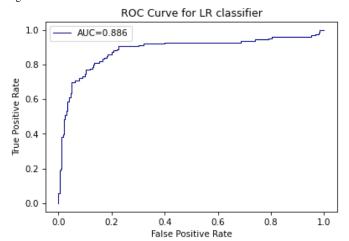


Fig. 11. ROC curve of Logistic Regression classifier

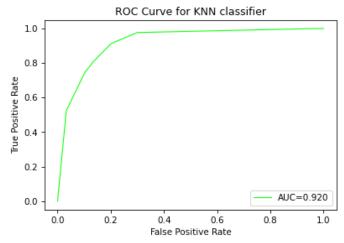


Fig. 12. ROC curve of KNN classifier

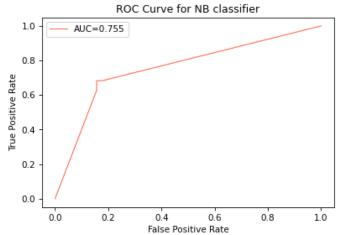


Fig. 13. ROC curve of NB classifier

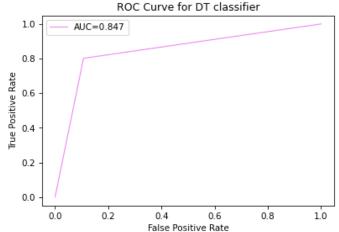


Fig. 14. ROC curve of DT classifier

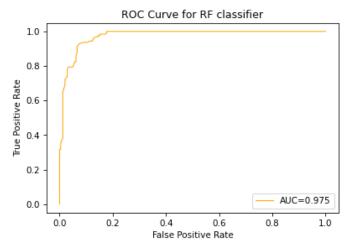


Fig. 15. ROC curve of Random Forest classifier

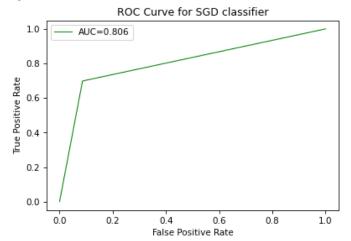


Fig. 16. ROC curve of SGD classifier

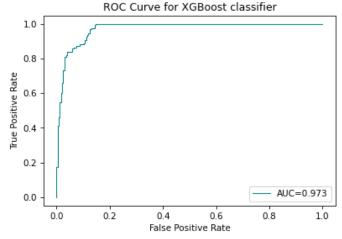


Fig. 17. ROC curve of XGBoost classifier

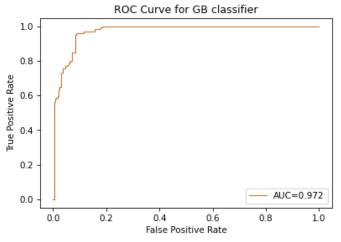


Fig. 18. ROC curve of Gradient Boosting classifier

The AUC-ROC scores of all the ML classification algorithms have been depicted in Table III to enable a performance comparative analysis. The AUC-ROC curve is an important evaluation metric for binary classification problems. The area under the curve is a measure of the ability of a classifier to distinguish between the two classes. An excellent model should have an AUC near to 1 which is an indication that the model has a good measure of separability.

Table III
Performance comparison of the models based on AUC-ROC

model	AUC-ROC
SVM	0.888
Logistic Regression	0.886
KNN	0.920
NB	0.755
DT	0.847
Random Forest	0.975
SGD	0.806
XGBoost	0.973
Gradient Boosting	0.972

From the test results in Table III, it can be observed that the AUC-ROC scores of the Random Forest, XGBoost and Gradient Boosting classifiers are extremely close. While Random Forest has an AUC-ROC score of 0.975, XGBoost has an AUC-ROC of 0.973 and Gradient Boosting has an AUC-ROC of 0.972. KNN follows Random Forest, XGBoost and Gradient Boosting classifiers with an AUC-ROC score of 0.920. The AUC-ROC scores of SVM and Logistic Regression have also been quite promising. SVM and Logistic Regression have AUC-ROC scores of 0.888 and 0.886 respectively. DT has an AUC-ROC score of 0.847 and AUC-ROC score for SGD classifier is 0.806. NB has the least AUC-ROC score of 0.755. Thus, it can be observed that the performances of Random Forest, XGBoost and Gradient Boosting classifiers stand out, each having AUC-ROC scores very near to 1. While Gradient Boosting classifier has the highest accuracy among the classifiers followed by Random Forest and XGBoost, Random Forest and XGBoost both have slightly higher AUC-ROC scores than Gradient Boosting.

AUC-ROC score of Random Forest exceeds that of Gradient Boosting by 0.003. The AUC-ROC score of Gradient Boosting is less than that of XGBoost by 0.001. Therefore, in order to conclude which classifier is the most accurate in detecting the presence or absence of brain tumor, the AUC-PR scores of all the algorithms have been computed. This has been described in the following section.

PR curves for performance comparison among the ML algorithms

The PR curves corresponding to SVM, Logistic Regression, KNN, NB, DT classifier, Random Forest classifier, SGD classifier, XGBoost classifier and Gradient Boosting classifier have been shown in Fig. 19, 20, 21, 22, 23, 24, 25, 26 and 27 respectively.

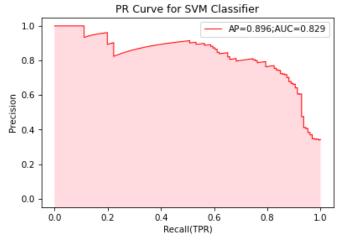


Fig. 19. PR curve of SVM classifier

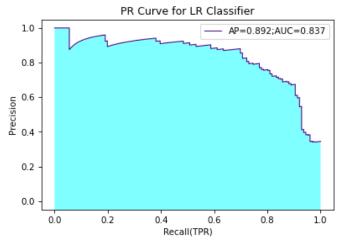


Fig. 20. PR curve of Logistic Regression classifier

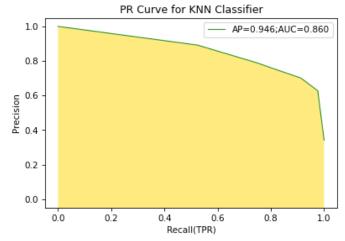


Fig. 21. PR curve of KNN classifier

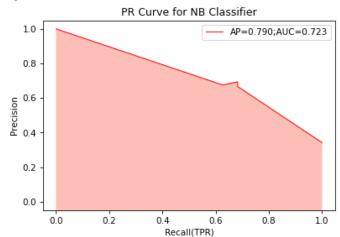


Fig. 22. PR curve of NB classifier

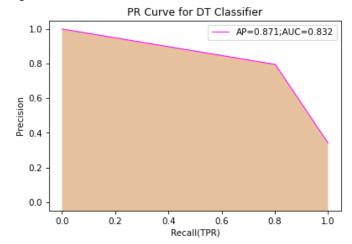


Fig. 23. PR curve of DT classifier

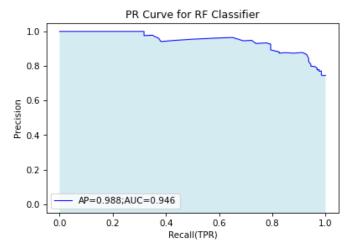


Fig. 24. PR curve of Random Forest classifier

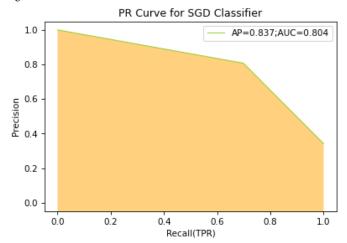


Fig. 25. PR curve of SGD classifier

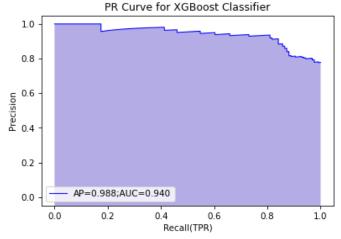


Fig. 26. PR curve of XGBoost classifier

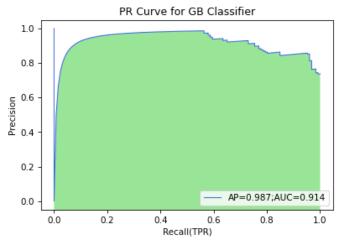


Fig. 27. PR curve of Gradient Boosting classifier

Similar to the ROC curve, the PR curve is used for evaluating the performance of the binary classification algorithms. The AUC-PR is constructed by plotting the precision against the recall for a single classifier at various threshold values. The nearer the AUC-PR score is to 1, the better the performance of the classifier.

The AUC-PR scores and the weighted average precision (AP) across all the thresholds of the proposed ML algorithms have been enlisted in Table IV.

Table IV
Performance comparison of the models based on AUC-PR and AP

model	AUC-PR	AP
SVM	0.829	0.896
Logistic	0.837	0.892
Regression		
KNN	0.860	0.946
NB	0.723	0.790
DT	0.832	0.871
Random Forest	0.946	0.988
SGD	0.804	0.837
XGBoost	0.940	0.988
Gradient Boosting	0.914	0.987

From the test results demonstrated in Table IV, it can be observed that Random Forest has the highest AUC-PR score of 0.946 followed by XGBoost with an AUC-PR score of 0.940 and Gradient Boosting with an AUC-PR score of 0.914. A good classifier should maintain both high precision and high recall across the graph. Thus, it can be inferred performances of Random Forest, XGBoost and Gradient Boosting classifiers have been noteworthy. SVM, Logistic Regression, KNN, NB, DT and SGD have AUC-PR scores of 0.829, 0.837, 0.860, 0.723, 0.832 and 0.804 respectively. While Random Forest has the highest AUC-PR score NB classifier has the lowest.

Therefore, it is evident that the performances of Random Forest, XGBoost and Gradient Boosting classification algorithms have been quite encouraging. The evaluation metric values of these three classifiers have been tabulated in

Table V in order to conclude which of these classifiers has exhibited the best overall performance.

Table V Evaluation metrics of Random Forest, XGBoost and Gradient Boosting classifiers

metric	Random	XGBoost	Gradient
	Forest		Boosting
Accuracy	0.908	0.894	0.924
Recall	0.952	0.888	0.944
Precision	0.811	0.818	0.850
F1-score	0.876	0.852	0.895
AUC-ROC	0.975	0.973	0.972
AUC-PR	0.946	0.940	0.914
AP	0.988	0.988	0.987

As depicted in Table V, Gradient Boosting classifier has an accuracy of 0.924 which is higher than that of Random Forest classifier by (0.924-0.908=) 0.016 and exceeds the accuracy of XGBoost classifier by (0.924-0.894=) 0.030. The precision of Gradient Boosting is 0.850. Precision of XGBoost classifier is 0.818 which is less than that of Gradient Boosting by (0.850-0.818=) 0.032. The precision of Random Forest is 0.811 which falls short of (0.850-0.811=) 0.039 from the precision of Gradient Boosting classifier. The F1-score of Gradient Boosting classifier is 0.895 which is higher than that of XGBoost and Random Forest. XGBoost and Random Forest have F1-scores of 0.852 and 0.876 respectively. However, the AUC-ROC scores and AUC-PR scores of both Random Forest and XGBoost are higher than that of Gradient Boosting. Random Forest has an AUC-ROC score of 0.975 which is higher than that of Gradient Boosting by (0.975-0.972=) 0.003. The AUC-ROC score of XGBoost is higher than that of Gradient Boosting by (0.973-0.972=) 0.001 only. Gradient Boosting has an AUC-PR score of 0.914 which is less than that of Random Forest by (0.946-0.914=) 0.032 and that of XGBoost by (0.940-0.914=) 0.026. Therefore, after comparing all the evaluation metrics it can be concluded that Gradient Boosting classifier has exhibited the best performance altogether with an accuracy, recall, precision, F1-score, AUC-ROC and AUC-PR scores of 0.924, 0.944, 0.850, 0.895, 0.972 and 0.914 respectively.

Also, the training time and prediction time of all the ML algorithms have been evaluated. This has been described in Table VI.

Table VI Comparison of training time and prediction time among the ML algorithms

model	training time (s)	prediction time (s)
SVM	39.325	12.24
Logistic	12.523	0.05
Regression		
KNN	13.079	47.655
NB	2.362	1.393
DT	24.41	0.101
Random Forest	13.268	0.172

SGD	2.629	0.061
XGBoost	145.068	0.65
Gradient Boosting	1180.088	0.172

As demonstrated in Table VI, NB classifier takes the least training time among all the other algorithms. It takes a training time of 2.362s. Gradient Boosting takes the maximum time for training. The training time for Gradient Boosting classifier is 1180.088s. Its prediction time is however 0.172s. The prediction time for KNN is the highest. It takes 47.655s for predicting the outcome. Logistic Regression with a prediction time of 0.05s is the fastest among the ML algorithms in predicting the outcome.

The test results of the multi-class classification problem have been described in the following sections.

B. Multi-class classification results

A comparative analysis of the 4 ML algorithms has been done based on the following evaluation metrics:

- Accuracy
- Recall (weighted average)
- Precision (weighted average)
- F1-score (weighted average)
- AUC-ROC

The performance comparison of the proposed algorithms on the basis of accuracy, recall, precision and F1-score has been demonstrated in Table VII.

Table VII Performance comparison of the proposed algorithms

model	SVM	KNN	Random	XGBoost
			Forest	
Accuracy	0.85	0.77	0.89	0.90
Recall(Weighted	0.85	0.77	0.89	0.90
Average)				
Precision(Weighted	0.87	0.82	0.89	0.90
Average)				
F1-score(Weighted	0.85	0.78	0.89	0.90
Average)				

As depicted in Table VII, XGBoost has outperformed the other models in terms of accuracy, recall, precision and F1-score. XGBoost has produced an accuracy, recall, precision and F1-score of 0.90 respectively.

In order to visualize the performance of the multi-class classifiers, the AUC-ROC curves of the four ML algorithms have also been plotted.

ROC curves of the multi-class classifiers for performance comparison

The ROC curves corresponding to SVM, KNN, Random Forest and XGBoost classifiers have been shown in Fig. 28, 29, 30 and 31 respectively.

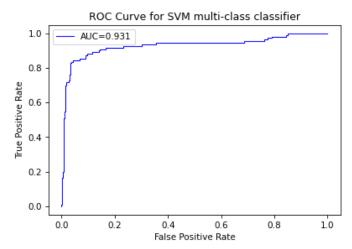


Fig. 28. ROC curve of SVM multi-class classifier

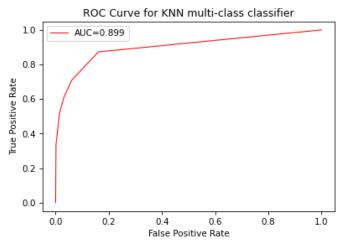


Fig. 29. ROC curve of KNN multi-class classifier

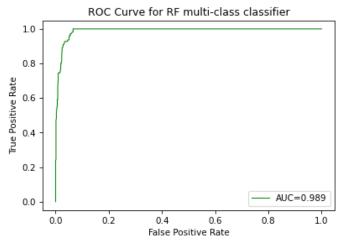


Fig. 30. ROC curve of Random Forest multi-class classifier

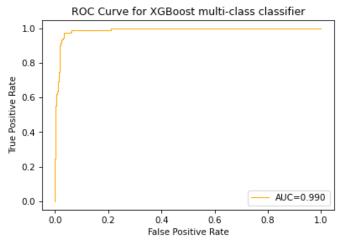


Fig. 31. ROC curve of XGBoost classifier

The AUC-ROC scores of the multi-class classification algorithms have been tabulated in Table VIII.

Table VIII AUC-ROC scores of the proposed models

model	AUC-ROC
SVM	0.931
KNN	0.899
Random Forest	0.989
XGBoost	0.990

From the test results shown in Table VIII, it can be observed that the AUC-ROC score of XGBoost is the highest. XGBoost has exhibited an AUC-ROC score of 0.990. Random Forest has produced an AUC-ROC score of 0.989, followed by SVM and KNN with AUC-ROC scores of 0.931 and 0.899 respectively.

Hence, after comparing the evaluation metrics of all the four ML algorithms it can be concluded that XGBoost classifier has outperformed the other models with accuracy, precision, recall, F1-score and AUC-ROC of 0.90, 0.90, 0.90, 0.90 and 0.99 respectively. Therefore, XGBoost is the best model to accomplish the task of multi-class classification of the images from Dataset-B.

VII. CONCLUSION AND SCOPE OF FUTURE WORK

This paper presents a detailed overview of how ML algorithms can be used for medical image processing. ML has improved and paved the way for efficient diagnosis, recognition and prediction in numerous domains of healthcare, brain tumor detection and classification being one of them. Nine ML algorithms have been used to predict whether a patient has a brain tumor or not based on a dataset comprising of brain MRI images. The ML algorithms that have been used are Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbor (KNN), Naïve Bayes (NB), Decision Tree (DT) classifier, Random Forest classifier, XGBoost classifier, Stochastic Gradient Descent (SGD) classifier and Gradient Boosting classifier. A performance comparison of the different ML algorithms has been conducted based on a few

performance metrics such as accuracy, recall, precision, F1-Score, AUC-ROC curves and AUC-PR curves. After the evaluation of the test scores, it has been concluded that Gradient Boosting is the best classifier among all the other ML classifiers that have been used. Also, multi-class classification has been performed on a different dataset comprising of brain MRI images of glioma, meningioma, pituitary and no tumor using SVM, KNN, Random Forest and XGBoost classifier. The ML algorithms have been compared based on accuracy, recall, precision, F1-score, AUC-ROC score and it has been observed that XGBoost classifier has exhibited the best results. In future, one of the most important improvements that can be made is adjusting the architecture so that it can be used during brain surgery, for classifying and accurately locating the tumor. Detecting the tumors in the operating theatre can be performed in real-time conditions; thus, in that case, the improvement would also involve adapting the network architecture to a 3D system. By keeping the network architecture simple, detection in real time can be made possible.

REFERENCES

- [1] Lars Kunze, Nick Hawes, Tom Duckett, Marc Hanheide and Tomáš Krajník, "Artificial Intelligence for Long-Term Robot Autonomy: A Survey", *IEEE Robotics and Automation Letters*, vol.3, issue.4, pp.4023-4030, 2018.
- [2] Li Deng, "Artificial Intelligence in the Rising Wave of Deep Learning: The Historical Path and Future Outlook [Perspectives]", IEEE Signal Processing Magazine, vol.35, issue.1, pp.180-187, 2018.
- [3] Alok Kole, "Design and Stability Analysis of Adaptive Fuzzy Feedback Controller for Nonlinear Systems by Takagi-Sugeno Model based Adaptation Scheme", *International Journal of Soft Computing*, vol.19, issue.6, pp.1747-1763, 2015.
- [4] Ruimin Ke, Yifan Zhuang, Ziyuan Pu and Yinhai Wang, "A Smart, Efficient, and Reliable Parking Surveillance System With Edge Artificial Intelligence on IoT Devices", *IEEE Transactions on Intelligent Transportation Systems*, vol.22, issue.8, pp. 4962 – 4974, 2021.
- [5] P. P. Bhattacharya, Alok Kole, Tanmay Maity and Ananya Sarkar, 'Neural Network Based Energy Efficiency Enhancement in Wireless Sensor Networks', *International Journal of Applied Engineering Research*, vol. 9, no.22, pp. 11807-11818, 2014.
- [6] Huimin Lu, Yujie Li, Min Chen, Hyoungseop Kim and Seiichi Serikawa, "Brain Intelligence: Go beyond Artificial Intelligence", Mobile Networks and Applications, vol.23, pp.368-375, 2018.
- [7] Sanjeevani Bhardwaj and Alok Kole, 'Review and Study of Internet of Things: It's the Future', in Proc. IEEE International Conference on Intelligent Control, Power and Instrumentation (ICICPI-2016), Kolkata, India, 2016, pp.47-50.
- [8] Daniel E. O'Leary, "Artificial Intelligence and Big Data", IEEE Intelligent Systems, vol.28, issue.2, pp.96-99, 2013.
- [9] Jiaying Liu, Xiangjie Kong, Feng Xia, Xiaomei Bai, Lei Wang, Qing Qing and Ivan Lee, "Artificial Intelligence in the 21st Century", *IEEE Access*, vol.6, pp. 34403 – 34421, 2018.
- [10] Chinmaya Kumar Pradhan, Shariar Rahaman, Md. Abdul Alim Sheikh, Alok Kole and Tanmoy Maity, 'EEG Signal Analysis Using Different Clustering Techniques', in *Proc. International Conference on Emerging Technologies in Data Mining and Information Security*, Kolkata, West Bengal, 2018, pp.99-105.
- [11] Alejandro F. Frangi, Sotirios A. Tsaftaris and Jerry L. Prince, "Simulation and Synthesis in Medical Imaging", *IEEE Transactions on Medical Imaging*, vol.37, issue.3, pp. 673 – 679, 2018.
- [12] Norio Nakata, "Recent technical development of artificial intelligence for diagnostic medical imaging", *Japanese Journal of Radiology*, vol.37, pp.103-108, 2019.
- [13] Subhamoy Mandal, Aaron B. Greenblatt and Jingzhi An, "Imaging Intelligence: AI Is Transforming Medical Imaging Across the Imaging Spectrum", *IEEE Pulse*, vol.9, Issue.5, pp. 16 – 24, 2018.

- [14] Ankit Ghosh, Purbita Kole and Alok Kole, "Automatic Identification of Covid-19 from Chest X-ray Images using Enhanced Machine Learning Techniques", *International Research Journal of Engineering and Technology (IRJET)*,vol.8, issue.9, no.115, pp.765-772, 2021.
- [15] Miles N. Wernick, Yongyi Yang, Jovan G. Brankov, Grigori Yourganov and Stephen C. Strother, "Machine Learning in Medical Imaging", *IEEE Signal Processing Magazine*, vol.27, issue.4, pp.25-38, 2010.
- [16] Sanjay Saxena, Neeraj Sharma and Shiru Sharma, "Image Processing Tasks using Parallel Computing in Multi core Architecture and its Applications in Medical Imaging", *International Journal of Advanced Research in Computer and Communication Engineering*, vol.2, issue.4, pp.1896-1900, 2013.
- [17] Chetanpal Singh, "Medical Imaging using Deep Learning Models", European Journal of Engineering and Technology Research, vol.6, issue.5, pp.156-167, 2021.
- [18] Neelum Noreen, Sellappan Palaniappan, Abdul Qayyum, Iftikhar Ahmad, Muhammad Imran and Muhammad Shoaib, "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor", IEEE Access, vol.8, pp. 55135 – 55144, 2020.
- [19] Janki Naik and Sagar Patel, "Tumor Detection and Classification using Decision Tree in Brain MRI", IJCSNS International Journal of Computer Science and Network Security, vol.14, no.6, pp.87-91, 2014.
- [20] Gopal S. Tandel, Antonella Balestrieri, Tanay Jujaray, Narender N. Khanna, Luca Saba and Jasjit S. Suri, "Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm", Computers in Biology and Medicine, vol.122, pp.103804-103860, 2020.
- [21] Ahmad M. Sarhan, "Detection and Classification of Brain Tumor in MRI Images Using Wavelet Transform and Convolutional Neural Network", *Journal of Advances in Medicine and Medical Research*,vol.32, issue.12, pp.15-16, 2020.
- [22] Heba Mohsen, El-Sayed A. El-Dahshan, El-Sayed M. El-Horbaty and Abdel-Badeeh M. Salem, "Classification using deep learning neural networks for brain tumors", *Future Computing and Informatics Journal*, vol.3, issue.1, pp.68-71, 2018.
- [23] Arshia Rehman, Saeeda Naz, Muhammad Imran Razzak, Faiza Akram and Muhammad Imran, "A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning", Circuits, Systems, and Signal Processing, Springer, vol.39, pp.757–775, 2020
- [24] Hyun-Chul Kim, Shaoning Pang, Hong-Mo Je, Daijin Kim and Sung Yang Bang, "Constructing support vector machine ensemble", Pattern Recognition, vol.36, issue.12, pp. 2757-2767, 2003.
- [25] Sandro Sperandei, "Understanding logistic regression analysis", Biochemia Medica, vol. 24, no. 1, pp.12-18, 2014.
- [26] Shichao Zhang, Xuelong Li, Ming Zong, Xiaofeng Zhu and Ruili Wang, "Efficient kNN Classification With Different Numbers of Nearest Neighbors", *IEEE Transactions on Neural Networks and Learning Systems*, vol.29, issue.5, pp. 1774 – 1785, 2018.
- [27] Saurabh Mukherjee Dr., Neelam Sharma, "Intrusion Detection using Naive Bayes Classifier with Feature Reduction", *Procedia Technology*, vol.4, pp.119-128, 2012.
- [28] S.R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology", *IEEE Transactions on Systems, Man, and Cybernetics*, vol.21, issue.3, pp. 660 – 674, 1991.
- [29] Angshuman Paul, Dipti Prasad Mukherjee, Prasun Das, Abhinandan Gangopadhyay, Appa Rao Chintha and Saurabh Kundu, "Improved Random Forest for Classification", *IEEE Transactions on Image Processing*, vol.27, issue.8, pp.4012 – 4024, 2018.
- [30] N. Deepa, B. Prabadevi, Praveen Kumar Maddikunta, Thippa Reddy Gadekallu, Thar Baker, M. Ajmal Khan and Usman Tariq, "An AI-based intelligent system for healthcare analysis using Ridge-Adaline Stochastic Gradient Descent Classifier", The Journal of Supercomputing, vol.77, pp.1998–2017, 2021.
- [31] Shenglong Li and Xiaojing Zhang, "Research on orthopedic auxiliary classification and prediction model based on XGBoost algorithm", Neural Computing and Applications, vol.32, pp.1971–1979, 2020.

[32] Djavan De Clercq, ZongguoWen and FanFei, "Determinants of efficiency in anaerobic bio-waste co-digestion facilities: A data envelopment analysis and gradient boosting approach", Applied Energy, vol.253, pp. 113570- 113572, 2019.