Dogo Rangsang Research Journal
ISSN : 2347-7180

UGC Care Group I Journal
Vol-12 Issue-09 No. 03 September 2022

Brain Tumor Classification from MRI Images Using Machine Learning Techniques

Maneesha Pandula[1], P.Rajesh Kumar[2], T.Roopa Rechal[3]
Affiliation 1:P.G. Student, Department of Electronics and Communication Engineering, Andhra University College of Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India, pmaneesha1999@gmail.com
Affiliation 2:Professor& HOD, Department of Electronics and Communication Engineering, Andhra University College of Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India, rajeshauce@gmail.com
Affiliation 3: Department of Electronics and Communication Engineering, Andhra University College of Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India, trrachel.rs@andhrauniversity.edu.in

ABSTRACT

Brain tumor is one of the major issues right now, it means growing different cells in brain. Brain tumor detection is the major task that can be performed by using magnetic resonance image (MRI) scans. The development of tumors in the human brain is a prominent cause of widespread loss of life. Tumor is the hostile assortment of compartments that increase in the humanoid body. The cancers develop in the brain are called as brain cancers. Each day large number of human lives are invisible due to this disease. The idea behind this paper is to detect the tumor of brain by finding the affected part of the brain MRI scans with the use of machine learning methods. In this method, main tumor image features are gathered by transferring them through a Convolutional Neural Network model, which is VGG16 model. By using different classification models like K- Nearest Neighbour, Logistic regression classifier, Decision Tree classifier, Gaussian Naïve Bayes classifier, and SVM classifiers brain tumor detection is done. The results are used for the study of the tumor image classification. Observe that K-Nearest Neighbor gives better accuracy than Logistic Regression, Support Vector Machine, Decision Tree, Gaussian Naïve Bayes. K Nearest Neighbour classifier gives better accuracy of 91%, precision 91%, recall 95% and f1 score 93% respectively.

Keywords:

Brain MRI images, Machine Learning, Image acquisition, Extraction, Image Processing, KNN Classifier, SVM Classifier, Decision Tree, Logistic Regression, Naïve Bayes, Classification.

1.INTRODUCTION

Now a day's medical image Processing is trending in the competitive field and is rapidly growing. Some methodologies are used for tumor detection and giving treatment. Brain tumor is one of the major issues at present and it means growing of the abnormal tissues in the brain. Brain plays major role in our body. MRI can provide plentiful of information about human soft tissues anatomy as well as helps diagnosis of brain tumor. MR images are used to analyze and study behavior of the brain.

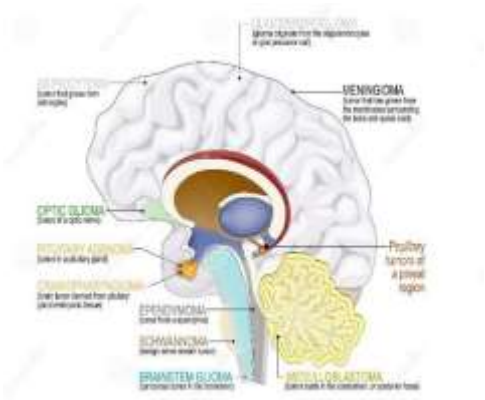

Fig 1: Primary brain tumors

Brain tumor detection is the major task that can be performed by using magnetic resonance image scans.

To detect the tumor, preform preprocessing, segmentation, feature extraction and then classification. To identify the tumor stage, extract features from the image. Before that must apply preprocessing by using different approaches to remove the unwanted data as well as noise and then segmentation next to extract the segmented part from the image. Next to that perform feature extraction based on that feature classification is going to be performed. In this era, apply the Naïve

Bayes algorithm, Decision Tree, support vector machine, Logistic Regression, K-Nearest Neighbour algorithms for the persistence of classification. By using different methods tumor identification and detection is done.

Identification of Brain tumor includes bodily inspection of CT scan, X-Ray, or MRI scans. Analysis through MRI images is the greatest trustable and decent approach as it is not involving in the revealing of the body cells to any radioactivity. If the cancer is not established, then discovery the size and location of tumor is needed.

Manually this process is not effective worthy, and computerization would be too cooperative to doctors recognize the tumors in the brain. In this model, a simple instrument founded on machine learning technique to recognize the presence of tumor orabsence of the tumor in the MRI image of the human brain. Following are main modules of the proposed model:

- Design an optimal and simple classification technique for identifying the tumor in MRI images.
- Extracting the features in the MRI images by using feature extraction techniques.
- Usage of various machine learning based classifiers to get accuracy, precision, recall, f1-score, and support.

2. RELATED WORK

In this section, conservative models are available in this works for brain cancer classification

Computer assisted system for Brian Tumor Detection and Segmentation, 2011.Segmentation using global verge value is a modest and time efficient technique. Drawback: The threshold value is chosen physically and less accurate.

Irregular tissue extraction in MRI Brain medical images, 2011. Expectation expansion algorithm is an iterative procedure to find maximum likelihood estimates of strictures in statistical model. Drawbacks: The algorithm is very composite. Similar effectiveness is realized with less complexity.

Efficient breakdown methods for tumor detection in MRI images, 2014. Capable study among three brain tumor detection methods.

This part of this segment gives classification of expectable models available in the works for the job of tumor identification [2-5, 7, 12, 13]. Many methods on detection and taxonomy of the type of tumors [2-5, 13].

Gabor filters are used to excerpt texture-based features from the images [2]. To evade over-suitable, undesirable feature elimination is done by feature collection representations like rank- based models and normal feature removal methods.

The problem of tumor classification is preserved as a voxel classification job as the voxel class is dependent on their neighboring voxels [3]. Conditional random field (CRF) methods are used to signify spatial relationships among thevoxels. Various filters have been applied to remove sound from the images [4], then wavelet features are removed and finally the images are classified using support vector-based classificationmodels.

Feature selection methods are used to excerpt prominent features from the features extracted from areas of interest [5]. KNN classifier is used to favor low grade neoplasm to high grade neoplasms.

Arrangement-based structures are utilized to piece the tumor exaggerated region from the given brain cancer images [7]. Four different types of arrangement-based features, intensity and spatial regularized images, even muscle spatial priors, expectable intensity spatial maps and flattened spatial mind mask and left to right uniformity have measured in this effort. The excellence of segmentation is measured by Jaccard similarity amount.

The early examination provides a positive result (means tumor exists) extract the tumor part; simple k-means algorithm is applied. An SVM based classification is feature that uses LBP features extracted from the MRI imageries [13].

In the next part of this unit, it present different deep learning models [9, 11, 14, 15] that are available in the literature. Few works [9, 15] focus on either discovery of classification or classification of the kind of tumors. The region of image is segmented before MRI scan is being classified[8, 11, 14].

A fully convolutional neural network (FCNN) withContinuous random filed (CRF) is proposed by Zhao et al. [9]to segment the affected region from the given image. DWT based features are

extracted, and deep learning models are used as classification approaches to categorize the kind of tumor [15].

A deep learning-based method has been proposed [14]. C-means clustering, a fuzzy technique is used to part the tumorpart. Initially, tumor part is segmented, and features are extracted using DWT and then for size reduction (Principal component analysis) PCA is applied over the features to evade over-fitting. Lastly, a deep structure is trained to categorize the data.

To segment the cancer part from the assumed MRI images they planned a deep learning-based technique [8]. This deep learning system can extract together local as well as global contextual features simultaneously. Different deep learning- based procedures and copies were studied concentrating on both breakdown and classification [11]. Most of the current methods utilize conventional models for classification and hand-made features for feature extraction.Some approaches use deep models that are educated as part of the deep model that is used for classification. In case of earlier models, the model is not vigorous and in the final case, the representations are vigorous, but it wants huge datasets.

3. PROPOSED METHODOLOGY

In this proposed methodology, Machine Learning techniques like SVM, Decision tree, Logistic Regression, K-Nearest Neighbour, and Naïve Bayes classifiers are used to detect the harshness of tumor in the brain basedon the givenMRI images.

It doesn't need more pre-processing of the data. It decreases the time throughout testing. On behalf of relating difficult preprocessing stages, topographies like energy, mean, variance and some of features are extracted from the dataset. And after that classification is done by usingdifferent machine learning techniques.

Detailed information about the dataset:

To show the effectiveness and capacity of the projected method and the new studies have used the dataset presented at Kaggle. And dataset link https://www.kaggle.com/navoneel/brain-mri- images-for-brain-tumor/cancer-detection. The name of the dataset is "Brain MRI images Dataset".

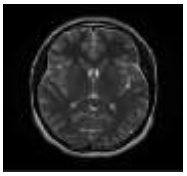

Fig 2: Brain MRI having a tumor(Left)**Fig 3:** Brain MRI not having a tumor(Right)

This dataset consists of number of MRI scans occupied under different types of imagery positions. The dataset contains a total number of 253 images. It has two classes. One is belonging to the YES class and the second is belonging to the NO class, both are tumorous and non-tumorous images. From the images, 98 images are non- tumor images, and155 images are having tumor.

DataVisualization:

In this stage, analyzing the data of the MRI which is taken from the dataset takes place. In this situation, out of 253 MRI images, 155 images are named as "yes", defines that there is a tumor present in the images and the remaining 98 images are named as "no", which indicate that no tumor presence in the images.
Training a neural network, VGG16 model to extract the features in the dataset.

Step 2: Data Augmentation

In this Data Augmentation step, by taking one MRI image and performing different image enhancement techniques as crop, effect, rotate, mirror and flip to get many images. Applying further augmentation to the classes to get almost equivalent amount of images to both classes.

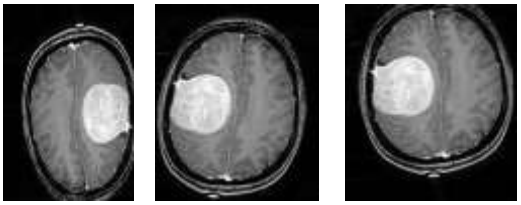

Fig 4: Data Augmentation of MRI scans

Step 3: Splitting the data

In this step, splitting the augmented data to training set and test set. Among 2064 images, 80 percent (1651) images are sent to training set and used by the VGG16 model to get trained. The 20 percent (413) images will be sending to test set, to understand the classification and they are classified to check the accuracy of the Neural

Network, VGG16 model. Deep models like VGG16, ResNet152, VGG19 model, Inception V3 model which are used in feature extraction. The architecture of the proposed model is divided into 5 different steps namely, picture acquisition, preprocessing stage, Segmentation process, Extraction of features, training, and testing of model.

Image Acquisition: Firstly, the brain MRI images are analyzed, acquired and then those images are set as input to the pre-processing stage.

Preprocessing: M R I pictures in the dataset are of various dimensions. This preprocessing step resizing MRI images of brain in the dataset, then every part of the images of brain in the whole dataset will power be of same dimension.

Segmentation Step: Segmentation of the image is the process of splitting the image into multiple segments. Medical Image Segmentation is the process of identifying of boundaries within a two Dimensional or 3Dimensional image.

Feature extraction: During this step, pre-trained model of neural network VGG16 is used to get the deep features. Feature extraction is a process by which certain features within an image are identified and representative for the next process, identification.

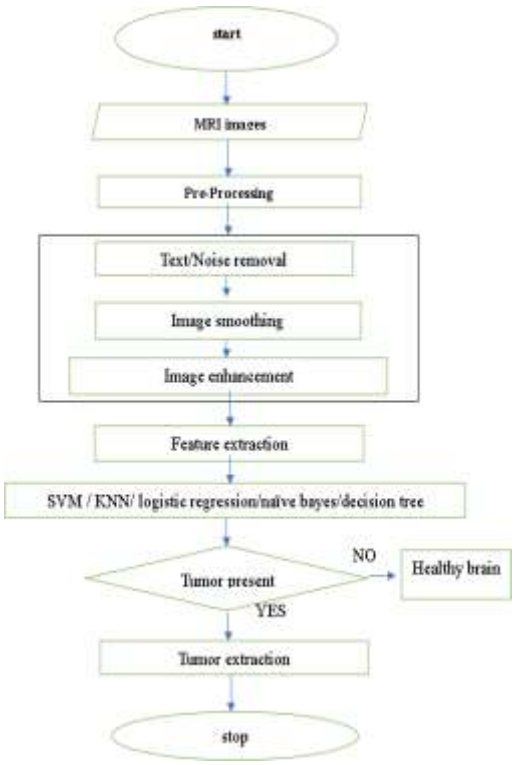

Fig 5: Flow Chart for Proposed system

Training and evaluation of model:

Through this module,train the method on the qualified data, which is undergoes through training phase, which is obtained utilizing extracted topographies in the altered space, and relate the method on the train data, calculate the accuracy of the model on the tested data. Figure 4 shows all the particulars of this process. This approach is accurate and modest as the method can be any effective model and this is powerful as the t u m o r features a r e extracted from the neural network models.

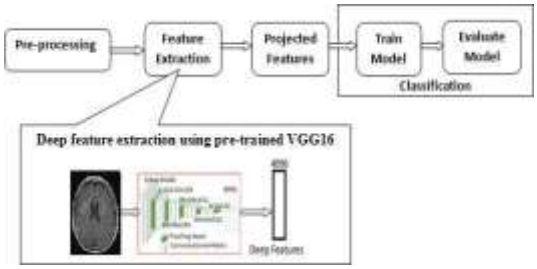

Fig 8: Architecture of the usage of VGG16 model

Classification Step:

The classification was done by using different classification algorithms those are naïve bayes classifier, decision tree, k-nearest neighbour classifier, support vector machine, neural networks. KNN is the best that it's efficient and accurate, it gives better performance when comparison with other algorithms. KNN is simple and accurate for implementing and easy to understand. It could learn all the boundaries when used for both the classification and regression.

No time during training for classification and regression. The KNN algorithm has no direct and clear training phase, and total work happens at prediction. During the time of classification, training and testing the data, 70 percent of it is used for training and 30 percent of it is used for testing.

4. COMPARISION AND RESULTS

The results of different classification models are carried out, to understand the ability and capability of this proposed model. Firstly, there is a need to present the dataset which is used in our revisions after the complete representation of the results of the proposed or advanced method for the involuntary finding of brain tumor by means of MRI scans of the brain

Table 1. Summary of the Brain MRI imagesdataset

Total images in the Dataset	253
Tumor images (YES)	155
Tumor images(NO)	98
Total pixels in the image	2622
Total train images	203
Total train images	50
After Augmentation Train Images	1651
After Augmentation Test Images	413

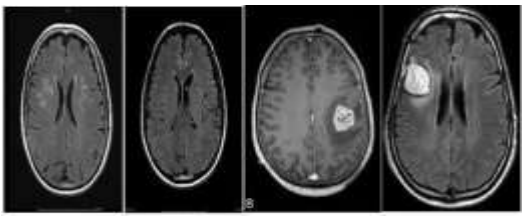

Fig 6: MRI without tumor

Fig 7: MRI with tumor

After the feature extraction step, classification is done on the dataset by using the machine learning techniques and then results are obtained.

4.1 Logistic Regression:

Logistic Regression is the most useful technique in classification problems.

Results for Logistic Regression:

```
model = LogisticRegression()
model.fit(X_train,y_train)
y_pred = model.predict(X_test)

print("Accuracy: ",accuracy_score(y_pred,y_test))
print("Confusion Matrix:\n", confusion_matrix(y_pred,y_test))
print("Classification Report:\n",classification_report(y_pred,y_test))
```

```
Accuracy:  0.9086460032626428
Confusion Matrix:
 [[269   2  22]
 [  0 170   6]
 [ 24   2 110]]
Classification Report:
              precision    recall  f1-score   support

           0       0.92      0.92      0.92       293
           1       0.98      0.97      0.97       184
           2       0.80      0.81      0.80       136

    accuracy                           0.91       613
   macro avg       0.90      0.90      0.90       613
weighted avg       0.91      0.91      0.91       613
```

Accuracy: 0.9086460032626428

4.2 SVM Classifier:

It is the utmost efficient and approved algorithm used for classification in machine learning. SVM is used for face observation, intrusion observation. It can hold both classification and regression process.

Results for SVM (linear SVC) classification:

```
model = LinearSVC()
model.fit(X_train,y_train)
y_pred = model.predict(X_test)

print("Accuracy: ",accuracy_score(y_pred,y_test))
print("Confusion Matrix:\n", confusion_matrix(y_pred,y_test))
print("Classification Report:\n",classification_report(y_pred,y_test))
```

```
Accuracy:  0.890701468189233
Confusion Matrix:
 [[265   2  22]
 [  3 177  12]
 [ 25   3 104]]
Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.92      0.91       289
           1       0.97      0.92      0.95       192
           2       0.75      0.79      0.77       132

    accuracy                           0.89       613
   macro avg       0.88      0.88      0.88       613
weighted avg       0.89      0.89      0.89       613
```

Accuracy: 0.890701468189233

4.3 Gaussian Naive Bayes:

Naive Bayes classifiers are the collection of algorithms which are used in classification based on Bayes Theorem.

Results for naïve bayes classifier:

```
model = GaussianNB()
model.fit(X_train,y_train)
y_pred = model.predict(X_test)

print("Accuracy: ",accuracy_score(y_pred,y_test))
print("Confusion Matrix:\n", confusion_matrix(y_pred,y_test))
print("Classification Report:\n",classification_report(y_pred,y_test))
```

```
Accuracy:  0.8221859706362153
Confusion Matrix:
 [[226   3   8]
 [  8 165  17]
 [ 59  14 113]]
Classification Report:
              precision    recall  f1-score   support

           0       0.77      0.95      0.85       237
           1       0.91      0.87      0.89       190
           2       0.82      0.61      0.70       186

    accuracy                           0.82       613
   macro avg       0.83      0.81      0.81       613
weighted avg       0.83      0.82      0.82       613
```

Accuracy: 0.8221859706362153

Results for decision tree classification:

```
model = DecisionTreeClassifier()
model.fit(X_train,y_train)
y_pred = model.predict(X_test)

print("Accuracy: ",accuracy_score(y_pred,y_test))
print("Confusion Matrix:\n", confusion_matrix(y_pred,y_test))
print("Classification Report:\n",classification_report(y_pred,y_test))
```

```
Accuracy:  0.7846655791190864
Confusion Matrix:
 [[247  12  36]
 [  8 148  16]
 [ 38  22  86]]
Classification Report:
              precision    recall  f1-score   support

           0       0.84      0.84      0.84       295
           1       0.81      0.86      0.84       172
           2       0.62      0.59      0.61       146

    accuracy                           0.78       613
   macro avg       0.76      0.76      0.76       613
weighted avg       0.78      0.78      0.78       613
```

Accuracy: 0.784665579119086

4.4 K-Nearest neighbor:

The k-nearest neighbor algorithm, also called as KNN, and it is a distribution free statistic, machine learning classifier. It uses closeness to make predictions and classifications regarding the collection of every data point.

Results for KNN classification:

```
model = KNeighborsClassifier(n_neighbors=7)
model.fit(X_train,y_train)
y_pred = model.predict(X_test)

print("Accuracy: ",accuracy_score(y_pred,y_test))
print("Confusion Matrix:\n", confusion_matrix(y_pred,y_test))
print("Classification Report:\n",classification_report(y_pred,y_test))
```

```
Accuracy:  0.9151712887438825
Confusion Matrix:
 [[268   1  14]
 [  0 181  12]
 [ 25   0 112]]
Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.95      0.93       283
           1       0.99      0.94      0.97       193
           2       0.81      0.82      0.81       137

    accuracy                           0.92       613
   macro avg       0.91      0.90      0.90       613
weighted avg       0.92      0.92      0.92       613
```

Accuracy: 0.9151712887438825

4.5 Decision tree:

Decision Tree is one of the classifiers in Machine Learning, in which the data is continuously splitting based on a certain parameter.

Table 2: Accuracy of different classification methodsin machine learning in the detection of brain tumor

Classification Model	Accuracy
Gaussian Naïve Bayes	82%
Logistic Regression	90%
K-Nearest neighbour	91.5%
decision tree	77%
Support Vector Machine	89%

Table 3: Precision of different classification methodsin machine learning in the detection of brain tumor

Classification Model	Precision
Logistic Regression	92%
Naïve Bayes'	77%
K-Nearest neighbor(K-NN)	91%
Decision tree	84%
SVM	90%

Table 4: Recall of different classification methods in machine learning in the detection of brain tumor

Classification Model	Recall
Logistic Regression	92%
Naïve Bayes'	87%
K-Nearest neighbor(K-NN)	95%
Decision tree	84%
SVM	92%

Table 5: f1-Score of different classification methodsin machine learning in the detection of brain tumor

Classification Model	f1- Score
Logistic Regression	92%
Naïve Bayes'	85%
K-Nearest neighbor(K-NN)	93%
Decision tree	84%
SVM	91%

Table 6: Support of different classification methods inmachine learning in the detection of brain tumor

Classification Model	Support
Logistic Regression	293
Naïve Bayes'	237
K-Nearest neighbor(K-NN)	283
Decision tree	295
SVM	289

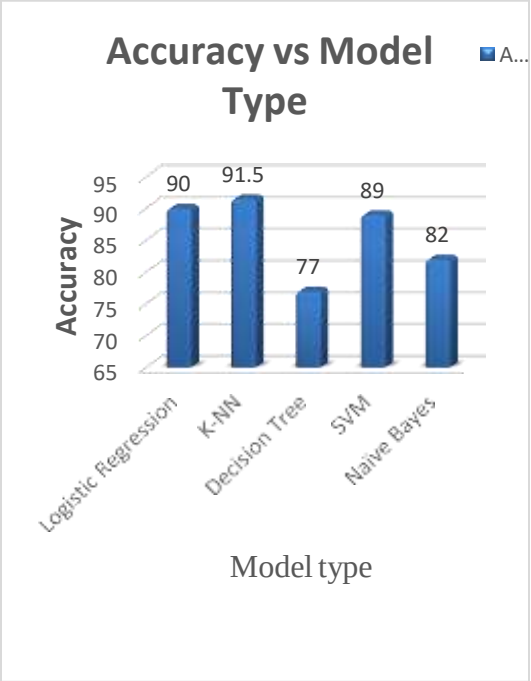

Fig 9: accuracy of different classification modelsto detect brain tumor

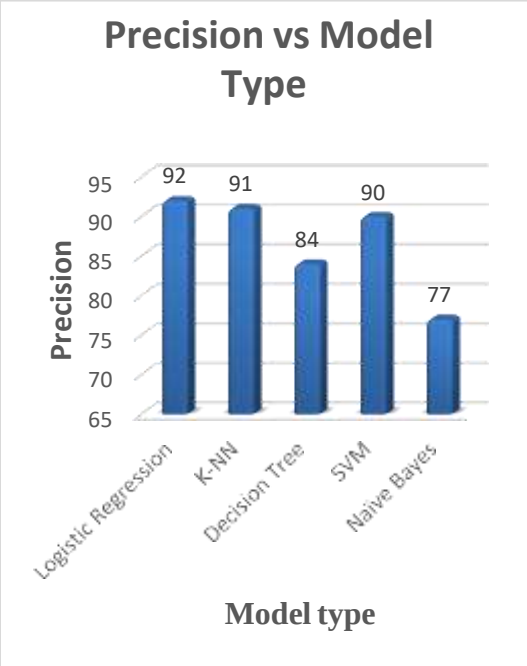

Fig 10: Precision of different classificationmodels to detect brain tumor

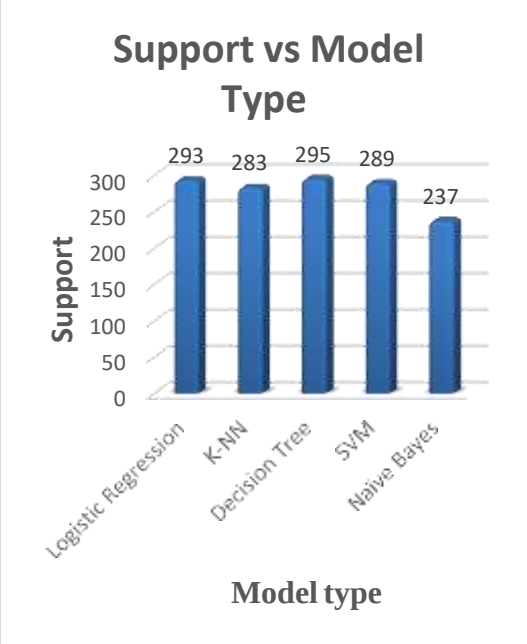

Fig 11: Support of different classification modelsto detect brain tumor

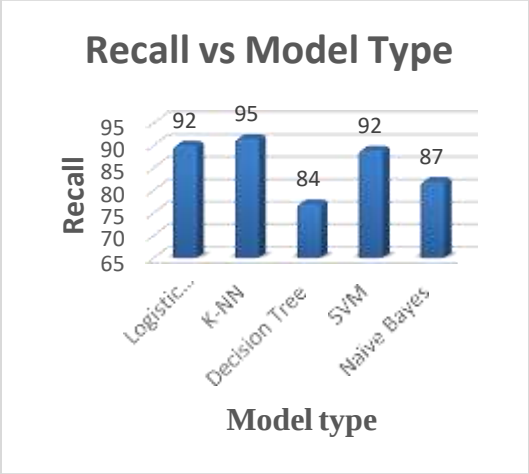

Fig 12: Performance of different classificationmodels to detect brain tumor

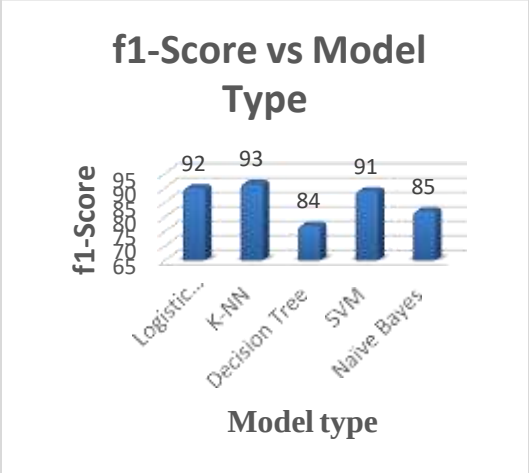

Fig 13: Performance of different classificationmodels to detect brain tumor

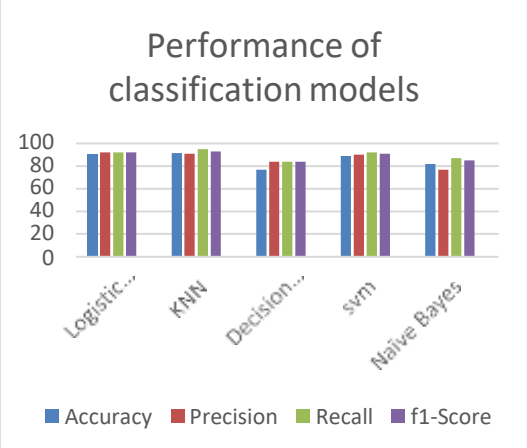

Fig 14: Performance of different classificationmodels to detect brain tumor

Compared to all classifiers KNN gives more accuracy, recall, f1-Score and precision to find the tumor in the brain.

CONCLUSION

The main aim of this approach is improving a method that can robotically detect presence of cancer in the brain MRI scans. This method is humble at efficient. Firstly, features are extracted from pretrained neural network model, VGG16 model to describe the additional important structures from the brain MRI images. On applying logistic regression, naïve bayes, decision tree, SVM, and K-nearest neighbor algorithms and based on the outcomes obtained, it can conclude that k-nearest neighbor classifier gives more accuracy to find the tumor presence. Here the proposed method classifies and detects the tumor in the brain MRI images without having misrepresentations.

REFERENCES

[1] Stephanie, S., Brix, T., Varghese, J., Warneke, N., Schwake, M., Brokinkel, B., Ewelt, C., Dugas, M., Stummer, W. (2019). Adverse events in brain tumor surgery: Incidence, type, and impact on current quality metrics. Acta Neurochirurgica, 161: 287-306. https://doi.org/10.1007/s00701- 018-03790-4

[2] Zacharaki, E.I., Wang, S.M., Chawla, S., Yoo, D.S., Wolf, R., Melhem, E.R., Davatzikos, C. (2009). Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 62(6):1609-1618. https://doi.org/10.1002/mrm.22147

[3] Bauer, S., Nolte, L.P., Reyes, M. (2011). Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Berlin, Heidelberg, pp. 354-361.

[4] Gurusamy, R., Subramaniam, V. (2017). A machine learning approach for MRI brain tumor classification. Computers, Materials & Continua, 53(2): 91-108. https://doi.org/10.3970/cmc.2017.053.091

[5] Zacharaki, E.I., Kanas, V.G., Davatzikos, C.

(2011). Investigating machine learning techniques for MRIbased classification of brain neoplasms. International Journal of Computer Assisted Radiology and Surgery, 6: 821-828. https://doi.org/10.1007/s11548-011-0559-3

[6] Zhou, J., Chan, K.L., Chong, V.F.H.,Krishnan, S.M. (2006). Extraction of brain tumfrom MR images using one-class support vectormachine. 2005 IEEE Engineering in Medicineand Biology 27th Annual Conference, Shanghai,China. https://doi.org/10.1109/IEMBS.2005.1615965

[7] Schmidt, M., Levner, I., Greiner, R., Murtha, A., Bistritz, A. (2005). Segmenting brain tumors using alignment-based features. Fourth International Conference on Machine Learning and Applications (ICMLA'05), Los Angeles, CA, USA. https://doi.org/10.1109/ICMLA.2005.56

[8] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. Medical Image Analysis, 35: 18-31. https://doi.org/10.1016/j.media.2016.05.004

[9] Zhao, X.M., Wu, Y.H., Song, G.D., Li, Z.Y., Zhang, Y.Z., Fan, Y. (2018). A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Medical Image Analysis, 43:98-111.

https://doi.org/10.1016/j.media.2017.10.002

[10] Zhou, M., Scott, J., Chaudhury, B., Hall, L., Goldgof, D., Yeom, K.W., Iv, M., Ou, Y.,Kalpathy-Cramer, J., Napel, S., Gillies, R.,Gevaert, O., Gatenby, R. (2018). Radiomics inbrain tumor: image assessment, quantitativefeature descriptors, and machine-learningapproaches. American Journal of Neuroradiology,39(2):208-216. https://doi.org/10.3174/ajnr.A5391

[11] Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D.L., Erickson, B.J. (2017). Deep learning for brain MRI segmentation: State of the art and future directions. Journal of Digital Imaging, 30(4): 449-459. https://doi.org/10.1007/s10278-017-9983-4

[12] Ain, Q.U., Mehmood, I., Naqi, S.M., Arfan Jaffar, M. (2010). Bayesian classification using DCT features for brain tumor detection. International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15387-7_38

[13] Deepika, K., Jyostna Devi, B., Ravu, K.S. (2019). An efficient automatic brain tumor classification using lbp features and SVM-based classifier. Proceedings of International Conference on Computational Intelligence and Data Engineering, Springer, Singapore. https://doi.org/10.1007/978-981-13-6459-4_17

[14] Mohsen, H., El-Dahshan, E.S.A., El-Horbaty, E.S.M., Salem, A.B.M. (2018). Classification using deep learning neural networks for brain tumors. Future Computing and Informatics Journal, 3(1): 68-71. Httweps://doi.org/10.1016/j.fcij.2017.12.001

[15] Bodapati, J.D., Kishore, K.V.K., Veeranjaneyulu, N. (2010). An intelligent authentication system using wavelet fusion of K-PCA, R-LDA. International Conference on Communication Control and Computing Technologies, Ramanathapuram, India. https://doi.org/10.1109/ICCCCT.2010.5670591