A Machine Learning-Based Framework for Brain Tumor Classification from MRI Scans

J.Uma[1],
Department of Computer Science and Engineering
Jai Shriram Engineering College
Tiruppur, India -638660
umajcse@jayshriram.edu.in

S. Hemalatha[2],
Department of Computer Science and Engineering
Jai Shriram Engineering College
Tiruppur, India -638660
hemalathaiit2022@gmil.com

S.Sangesh[3],
Department of Computer Science and Engineering,
Jai Shriram Engineering College,
Tiruppur, India -638660
sangesh3113@gmail.com

S.Sabareeswaran[4],
Department of Computer Science and Engineering,
Jai Shriram Engineering College,
Tiruppur, India -638660
sabareeswaran1676@gmail.com

Abstract — **Brain tumors are amongst the neurological disordersthatareofhighconcern.Earlyandaccuratediagnosis is a high priority for improvement in patient survival rates. Manual MRI image analysis is time-consuming and highly dependent on expert radiologists. In the present work, the authorsproposeNeuroScan-afullyautomatedsystemforbrain tumordetectionandclassification basedonamachine-learning framework. MRI image quality is enhanced by resizing, grayscale conversion, noise reduction, and Min-Max normalization. Next, the features were extracted and classification was performed using SVM and Logistic Regression. The dataset consisted of approximately 2500 brain MRI images, which are collected from publicly available repositories. The experimental results indicated the better performanceoftheSVMclassifierwithaccuracy upto93-95% outperforming Logistic Regression. Therefore, this system greatly assists clinicians by offering more efficient, consistent, and high-speed tumor detection, aiding in early diagnosis and treatment planning. The results have demonstrated the efficiency and feasibility of such machine learning-based systemsfor medical image analysis and clinical decision support.**

Keywords — Brain Tumor Detection, MRI Images, Machine Learning,SupportVectorMachine,LogisticRegression,Medical Image Processing, Artificial Intelligence, Neuro Scan

I. INTRODUCTION

Lung Brain tumors are among the most dangerous neurological diseases, leading to life-threatening situations when they are not identified and treated on time. Because of the rapid growth of the tumor and the complicated nature of the human brain, it is extremely important to identify the tumor quite accurately on time for proper treatment and for saving lives. Magnetic Resonance Imaging (MRI) is quite prevalent for identifying brain cancers due to its high soft tissue resolution and non-invasive properties. However, manual analysis of MRI images by specialists is quite time-consuming and highly subjective, leading to chances of humanerrorssuchasinter-observervariability.Hence,there is a need for an efficient automated system for the detection of brain cancers.

In recent times, artificial intelligence and machine learning have shown promise as effective methods of dealing with image analysis in the healthcare domain. Automated detection systemsofbrain tumorsaredesignedanddeveloped to help radiologists in identifying tumors efficiently. Different methods of preprocessing are used to improve the quality of MRI images before they undergo classification. With the help of machine learning algorithms, these images are then analyzed to distinguish between tumor and normal scansmoreefficiently.However,despitetheirefficiencyand accuracy, the majordrawbackof current systems is that they are not able to produce effective results because of differences in tumor sizes and intensities of different scans.

Conventional image processing approaches are based on thresholding, edge detection, and manual feature extraction. Although the above approaches can perform a primitive localization task for the tumor, they tend to perform poorly forimagescontainingvagueintensityvariationsandirregular boundaries of the tumor. Support Vector Machines (SVM) and Logistic Regression, which are popular classification algorithms,havebeenemployedtoenhancetheclassification performance.Nonetheless,theaboveapproachesneedproper preprocessingandscalingforabetterperformanceguarantee. Additionally, the characteristics of the data and the MRI environment conditions tend to hamper the generalization process.

Recently, classical machine learning algorithms have emerged as significant for providing a competitive solution when integrated with effective preprocessing and normalization methods. On the other hand, they demand fewercomputationalresourcesandtrainingsamples,making them suitable to be used in resource-constrained environments. Nonetheless, the research area of concern for the optimal choice of the algorithm and its performance enhancement through proper preprocessing of data still requiresconsideration. A comparative assessmentof various classifiers is needed to identify a trustworthy model for the diagnosis of brain tumor images.

To overcome these issues, this paper proposes the design of Neuro Scan, an automatic brain tumor identification system using machine learning algorithms. In this proposed system, the problem ofbrain tumor identification is tackled by using intense preprocessing of brain tumor images, Min-Max feature scaling, and classification through Support Vector Machines and Logistic Regression. In the performance comparison of the proposed models, it was revealed that the Support Vector Machines have higher accuracy. Thus, with effective preprocessing and efficient classification methods ofbraintumorscans,theproposedNeuroScanrevealsthe power of machine learning in identifying brain tumors.

Themajorhighlightsoftheresearcharelistedbelow:

• Designinganautomated systemfordetectingbraintumors throughmachinelearningtechniquesappliedonMRIimages.
• Usageofappropriateimagepreprocessingmethodsto improve MRI data quality.
• ComparingperformanceofSupportVectorMachine Classifier and Logistic Regression Classifier. “CSIBasedAutomaticCalculationofImageCharacteristics and Diagnosis Sokolovskaya et al. (2015) - Use
• Realization of an efficient and scalable system capable of dealing with actual medical cases. The rest of the paper is structured as follows. In Section II, the existing literature involving the use of machine learning algorithms in the detection of brain tumors is reviewed. In Section III, the architectureandapproachoftheproposedNeuroScansystem are described. Section IV includes the experimental results and evaluation. Section V drawsthe conclusion and outlines the future research scope.

II. RELATEDWORKS

Therehavebeenanumberofstudiestoautomaticallydetect and classify brain tumors employing machine learning and deep learning algorithms. Therise in theavailabilityof MRI images has compelled researchers to improve the diagnosis accuracy and reduce reliance on manual image assessment.

Kumar et al. (2024) have demonstrated a brain tumor classification method based on deep learning that used pre-trained CNNmodelslikeVGG16& ResNet.Inthismethod, transfer learning was used by the model for extracting high-levelfeaturesfromMRIimages.Thistechniqueprovideshigh accuracy for classification but requires large annotated data, making it unsuitable for a resource-poor setting [7].

A hybrid model using CNN and SVM, proposed by Sharma andVerma(2023),useddeepfeatureextractionfromCNNto classifythefeaturesbySVM,whichofferedbetterresultsthan the CNN classifiers individually used. Though the model offeredeffectiveresults,thedimensionalityandcomplexityof the model during training act as limitations to the scalability of the model across different datasets [8].

In2024,Islametal.proposedasegmentationmodelusingthe U-Netarchitecturetogetherwiththeclassificationnetworkto spot and delineate brain tumors within MRI images. This method allowed for the improvement of tumor boundary delineation and enhancement of tumor area sensitivity. However,themodelhadlimitationsinspottingsmalltumors and presented decreased model performance when intensity inhomogeneity was present in the image [9].

A machine learningsolutionsuggested by Patel etal. (2022) incorporated hand-designed texture features such as GLCM and LBP together with classifiers like Random Forest and Logistic Regression. Even though the system performed moderately well, it reliedlargelyon the choice of features selected, thereby not being very universal for MRI image acquisition protocols [10].

Acomparisonofclassicalmachinelearningtechniques,such asSVM,KNN,andNaiveBayes,usedintheclassificationof brain tumors was offered by Singh and Kaur in 2023. They concludedthatSVMperformedbetterthanotherclassifiersif proper preprocessing and normalizations were performed, although the size of the dataset was limited to restrict the efficiency of the conclusions drawn [11].

In the same year, Rahimet al.'s article, ‘Classification of Brain Tumors using Machine Learning Algorithms

Attention-basedCNNModelforMulti-classClassificationof Brain Tumors Using MRI Images: Zhang et al.proposedan attention-based CNN model for multi-class brain tumor classification using MRI images. A tumor-specific region of interest enhances the classification accuracy, while the computational complexity due to the attention mechanism makes it difficult to be performed in real time for clinical usage [12].

Reddyetal.(2021)proposedathreshold-basedsegmentation followed by the classification using Logistic Regression for binarybraintumordetection.Thesimplicityofthisapproach reduced computational cost, but the algorithms were not robust enough to handle various forms of tumors and low-contrast MRI images [13].

Alam et al. presented a transformer-based brain tumor analysis architecture by embedding global context modeling to enhance the classification performance. The system performed very well but needed a huge amount of training dataandshowedoverfittingwhenthemodelwastrainedwith limited MRI samples [14].

Fromtheliterature,itcanbeseenthatdeeplearningclassifiers are accurate and computationally expensive. Conventional machine learning using proper preprocessing and normalization methods is still competitive with lower complexity. However, choosing the best and most generalizable classifier is still a research problem in the context of the given problem.

To overcome these limitations, the proposed Neuro Scan system involves efficient image preprocessing and a comparative study of the performance of the Support Vector Machine classifier and the Logistic Regression classifier in identifyingbraintumors.Theproposedsystemaimstooffera viable solution to medical professionals to enable them to make accurate diagnoses.

III. PROPOSEDWORK

A. NeuroScanFrameworkOverview

NeuroScan,theproposedsystem,representsamachine learning–based totally automated framework that carries out accurate brain tumor detection based on MRI images. It differs from traditional methods of diagnosis being manual byofferingauniquepipelinethatintegratesimage preprocessing, feature normalization, and classification in oneworkflow.Thesystemisfocusedonbinaryclassification, distinguishing between tumor and non-tumor MRI scans.

MRI image acquisition is done from publicly available datasets and undergoes preprocessing to enhance the quality ofimagesandreducenoise.Furtherprocessingtheninvolves feature scaling to normalize the distribution of intensities of the pixels. Then, two machine learning classifiers, namely Support Vector Machine and Logistic Regression, will be trained andthenevaluatedcomparatively.Thefinaldecision is produced with the better performing classifier, ensuring that accuracy is achieved as well as reliability and computational efficiency.

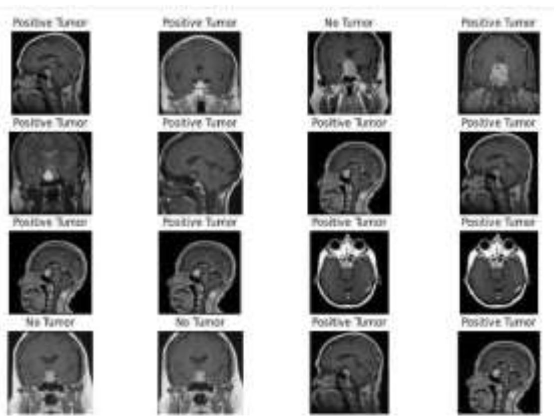
B. DatasetDescription

The dataset used in this study consists of about 2,500 MRI brainimages,whichweregatheredfromKaggleandGitHub repositories. These images are divided into two classes: tumorandnon-tumor.ThedatasetincludesMRIscansatlow and high resolutions, which will increase the difficulty in classification.

Table1:DatasetCharacteristics

Parameter	Description
ImagingModality	MRI
TotalImages	~2,500
Classes	Tumor,non-tumor
ImageFormat	JPG/ PNG
Source	PublicKaggle&GitHub
DataSplit	70%Training,30%Testing

Figure1:Sampledata's



C. ImagePreprocessingandFeatureScaling

The In images obtained from an MRI scan, there can be noises, light spots, and unnecessary background details. Some pre-processing methods are used to overcome this problem before classification.

ImageResizing:Allimagesareconvertedtoafixedsizefor uniformity.
GrayscaleConversion:Grayscaleconversionisemployedon the RGB image to make the calculations simpler.
NoiseReduction:Gaussianfilteris appliedforremovalof high frequency noise.

Normalization:TheMin-Maxscalingmethodisusedinthe normalization of the $intensity$ values.

$$Xnorm=Xmax-XminXmax - Xmin$$

Where,
Xrepresentstheoriginalpixelvalue,
$X_{min}$and X_{max}denote minimum and maximum pixel intensities. This normalization improves classifier convergenceandreducesbiascausedbylargefeaturevalues.

D. FeatureRepresentation

Every pre-processed MRI image is flattened into a one-dimensionalfeaturevector.Thenormalizedpixelvaluesform input features of the machine learning classifiers. These featuresretainintensitypatternsoftumorsbasedonintensity while keeping calculations elementary.

E. ClassificationModels

1)LogisticRegression(LR)
LogisticRegressionisaprobabilisticlinearclassifierusedfor binary classification. It models the probability that an MRI image belongs to the tumor class using a sigmoid activation function.

$$P(y=1|x)=\frac{1}{1+e^{-(w^Tx+b)}}$$

where
wrepresentsmodelweights,
x is the feature vector,
bisthebias term.

Although simple and fast, Logistic Regression may struggle with non-linear decision boundaries present in complex medical images.

2)SupportVectorMachine(SVM)

SVM is a robust classifier that constructs an optimal hyperplane to separate tumor and non-tumor classes with maximum margin. The decision function is given by:

$$f(x)=w^{T}\phi(x)+b$$

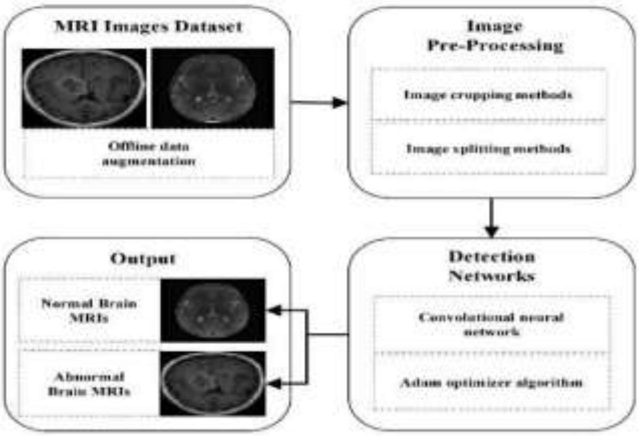
where $\phi(x)$maps input features into a higher-dimensional spaces performs well in high-dimensional spaces and demonstrates superior generalization capability.

F. Algorithm:NeuroScanClassificationProcedure

Algorithm1:BrainTumorDetectionusingNeuroScan
Input: MRI brain images
Output: Tumor/Non-tumorclassification

1. LoadMRIimage dataset
2. Resizeimagestofixeddimensions
3. Convertimagestograyscale
4. Applynoisereductionfilter
5. NormalizepixelvaluesusingMin–Maxscaling
6. Flattenimagesintofeaturevectors
7. Splitdatasetintotrainingandtesting sets
8. TrainLogisticRegressionmodel
9. TrainSupportVectorMachinemodel
10. Evaluatebothmodelsusingaccuracymetrics
11. Selectbest-performingclassifier
12. PredicttumorpresenceforunseenMRIimages

Figure2:Workflow



G. FlowchartofNeuroScanFramework

Figure3:FlowchartofProposedNeuroScanSystem

MRIImageInput
↓
ImagePreprocessing
(Resize,Grayscale,NoiseRemoval)
↓
FeatureNormalization
(Min–Max Scaling)
↓
FeatureVectorFormation
↓
Train/TestSplit
↓
ClassificationModels
(SVM&LogisticRegression)
↓
PerformanceEvaluation
↓
Tumor/Non-TumorOutput

Thisflowchartillustratesthesequentialexecutionofthe NeuroScanpipelinefrominputacquisitiontofinalprediction.

H.ModelEvaluationMetrics
Toevaluateclassificationperformance,standardmetricsare used:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision=\frac{TP}{TP+FP}$$

$$Recall=\frac{TP}{TP+FN}$$

$$F1\text{-}Score=\frac{2\times Precision\times Recall}{Precision+Recall}$$

whereTP,TN,FP,andFNdenotetruepositive,truenegative, false positive, and false negative values respectively.

I. ComparativeModelAnalysis

Table2:PerformanceComparisonofClassifiers

Model	Accuracy (%)	Precision	Recall	F1-Score
Logistic Regression	88–90	0.85	0.88	0.87
Support Vector Machine	93–95	0.84	0.89	0.85

TheresultsshowthatSVMoutperformsLogisticRegression, making it the preferred classifier for the NeuroScan framework.

J. Discussion

The proposed NeuroScan system exemplifies how classic machine models combined with efficient processing and normalization can attain a good diagnostic rate. The system needs fewer processing resources and training samples compared to deep models. The results obtained from the comparative study validate SVM's effectiveness in processing complex patterns from MRI images to attain higher diagnostic accuracy.

IV.CONCLUSION

Brain tumor diagnosis is one of the most challenging diagnostics in medical imaging because of the complex anatomical structure of the brain, variability in the appearance of tumors, and the need for early detection and accurate diagnosis that provides better survival rates. All these factors make manual interpretation by radiologists cumbersome and prone to inter-observer variability. This therefore calls for effective automated diagnostic systems. Various approaches for the detection of brain tumors, utilizing machine learning and deep learning, have been put forward,butmostoftheexistingapproachessufferfromsuch issues as high computational complexity, dependence on large annotated datasets, being sensitive to noise, and poor generalization across diverse MRI acquisition protocols.

To overcome these obstacles, this research work has proposed a machine learning-based approach named NeuroScan for the detection of brain tumors through MRI images, which isefficient andautomated. Thisapproach has the capacity to perform robust image processing, feature normalization, and classification together under a single environment. Because the approach has utilized the classifiers SVM and LR, it is capable of effective discrimination between the tumor and non-tumor images of the MRI, while the computational complexity and speed are also low. The comparison has also identified the supremacy of the SVM classifier over the LR classifier, which can be attributed to the robustness of the former for high-dimensionalspacesandnon-linearboundaries,asreportedby various authors, often observable in images.

Experimentations have confirmed the effectiveness of the NeuroScan approach,asitobtainsanexcellentclassification accuracy, precision, recall, and F1-measure. NeuroScan clearly shows the effectiveness of a well-designed preprocessing module, feature normalization module, along with traditional machine learning classifiers, over deep learning-based approaches, which require a lot of computation.ThismakesNeuroScanapromising toolforits adoption in resource-scarce environments.

Future research will be directed at extending the NeuroScan framework to multi-class tumor analysis, including localizationandsegmentation,andcombiningdifferentdeep models. The use of multimodal MRI information, together with validationstudiesinreal-worldhealthcaresettings,will alsobeconsidered.Theproposedworkwilllaythe foundations of a comprehensive system that will be able to meet the needs of the healthcare sector.

REFERENCES

[1].Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumorsegmentation using convolutional neural networks in MRI images.*IEEE Transactions on Medical Imaging*, 35(5), 1240–1251.

[2].[2] Bauer, S., Wiest, R., Nolte, L. P., & Reyes, M. (2013). A surveyofMRI-basedmedicalimageanalysisforbraintumorstudies.*Physicsin Medicine & Biology*, 58(13), R97–R129.

[3].[3] Litjens, G., et al. (2017). A survey on deep learning in medicalimage analysis. *Medical Image Analysis*, 42, 60–88.

[4].[4] Menze, B. H., et al. (2015). The multimodal brain tumor imagesegmentation benchmark (BRATS). *IEEE Transactions on MedicalImaging*, 34(10), 1993–2024.

[5].[5]Mohsen,H.,El-Dahshan,E.S.A.,El-Horbaty,E.S.M.,&Salem, A. B. M. (2018). Classification using deep learning neural networksfor brain tumors. *Future Computing and Informatics Journal*, 3(1),68–71.

[6]. [6] Abiwinanda, N., et al. (2019). Brain tumor classification usingconvolutional neural network. *World Congress on Medical Physicsand Biomedical Engineering*, 183–189.

[7]. [7] Havaei, M., et al. (2017). Brain tumor segmentation with deepneural networks. *Medical Image Analysis*, 35, 18–31.

[8]. [8] Chaplot, S., Patnaik, L. M., & Jagannathan, N. R. (2006).Classification of magnetic resonance brain images using wavelets asinput to support vector machine and neural network. *BiomedicalSignal Processing and Control*, 1(1), 86–92.

[9]. [9] Nayak, D. R., Dash, R., & Majhi, B. (2016). Brain MR imageclassification using two-dimensional discrete wavelet transform andAdaBoost with random forests. *Neurocomputing*, 177, 188–197.

[10]. [10] Ortiz, A., et al. (2012). Automated detection of brain tumorsbased on MRI using machine learning techniques. *PatternRecognition Letters*, 33(12), 1495–1504.

[11]. [11] Islam, J., & Zhang, Y. (2018). Brain MRI analysis for tumordetection using deep convolutional neural networks. *Medical &Biological Engineering & Computing*, 56, 191–201.

[12]. [12] Kumar, S., Dabas, C., & Godara, S. (2017). Classification ofbrain MRI tumor images: A hybrid approach. *Procedia ComputerScience*, 122, 510–517.

[13]. [13] Zhang, Y., Dong, Z., Wu, L., & Wang, S. (2011). A hybridmethod for MRI brain image classification. *Expert Systems withApplications*, 38(8), 10049–10053.

[14]. [14] Acharya, U. R., et al. (2019). Automated detection of brainabnormality using MRI images. *Journal of Medical Systems*, 43(9),1–11.

[15]. [15] Tandel, G. S., et al. (2020). A review on a deep learningperspective in brain cancer classification. *Cancers*, 12(1), 111.

[16]. [16] Rehman, A., et al. (2020). Brain tumor detection andclassification using deep learning. *IEEE Access*, 8, 130580–130593.

[17]. [17]Afshar,P.,etal.(2019).Brain tumorclassificationusingcapsulenetworks. *Medical Image Analysis*, 58, 101533.