Turkish Journal of Computer and Mathematics Education Vol.12 No.13(2021), 2167-2183

Research Article

"Risk Stratification of Brain lesions in MRI Images using Cascaded Machine Learning Paradigm"

Ullas Agrawal[a], Pankaj Kumar Mishra[b]*, Bikesh Kumar Singh[c]

[a] PhD Scholar, Electronics & Communication, Rungta College of Engineering and Technology, Bhilai, INDIA
[b]Professor, Rungta College of Engineering & Technology, Bhilai,(C.G.) INDIA
[c]Assistant Professor, Department of Biomedical Engineering, NIT Raipur.
e-mail-[a]ullasagrawal361@gmail.com, [b]pmishra1974@yahoo.co.in, [c]bsingh.bme@nitrr.ac.in,

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 4 June 2021

Abstract: Brain lesion is the most severe among all type of lesion, which causes death of a huge number of patients every year. There are two types of brain lesion high-grade glioma andlow-grade glioma (LGG). Amongst both, LGG is fatal for normal life. Early detection of the LGG lesion can save huge amount of life. The article aims to build a cascaded computer aided diagnosis (CAD) system, which can detect the lesion and the severity of the lesion. In the study, a brain MRI dataset is achieved from The Cancer Genome Atlas (TCGA) dataset. Here 3D RGB brain MRI dataset is provided with brain lesion and normal images from 110 patients. Besides this dataset includes the death details of the patient, which is used for severity assessment. At first, the dataset is used to achieve feature for each classification model. Then, two groups of efficient features are selected by Mann-Whitney U test. Those efficient features are used to detect the brain lesion and the severity assessment by two cascaded classification models. Here k-fold, hold out andboth cross validation are applied. In the study, different neural network classifications are applied. Those are from support vector machine, k-nearest neighbour and ensemble classifiers. The results of the classifiers are used for performance evaluation. For the detection of lesion 94.4% accuracy and 95.3% sensitivity are achieved. For risk factor measurement, highest 100% accuracy and sensitivity are achieved..

Keywords: Brain lesion Classification, Risk factor classification, Brain MRI, SVM, KNN, Ensemble.

1. Introduction

Brain cancer is fatal among all types of cancers. In 2020, approximately 7,00,000 Americans have brain lesion, where 30% brain lesions are malignant. Out of those American patients, 55% were males and 45% were females. The survival rate is 36% of the patients with malignant lesions. The five-year survival rate of such a lesion is 6.8%. According to estimation, 18,020 people would die due to brain lesion in 2020. [1] According to American Cancer Society, 23,890 malignant lesions were detected including adults and children in USA. Out of these, 13,590 lesions and 10,300 lesions were detected from males and females, respectively. Approximately, 18,000 patients have severe lesion, which includes 10,190 males and 7830 females. The number would be higher with the benign brain lesions. [2] Brain lesion has two different classes. First is high-grade glioma (HGG) and low-grade glioma (LGG). HGG can be cured by treatment but LGG is fatal, which can cause death within very less time span. Early detection of cancer can save a life.

The treatment depends on the accurate detection and classification of the radiologist. The location, stage, severity werecrucial, while detecting the brain lesion. Brain MRI images is the best diagnostic images to detect brain lesion. There was inter-observer variability at the detection of the brain lesion. Wrong diagnosis leads to delay in treatment, which results in severity. Presently, computer aided diagnosis (CAD) systems were used for early detection of brain lesion. [3] There was also variability, which can be solved neural network based CAD system.

2.Related work

There were different applications of CAD system, where different features and various methods were used to detect brain lesions. In 2015, Cheng et al. [4] used a CAD model for ROI based brain lesion detection. In the study, meningioma, glioma, and pituitary lesion were classified in MRI images. Here, intensity histogram, grey-level co-occurrence matrix (GLCM), bag of words (BOW) was used for feature extraction. By using ROI based detection and region augmentation the accuracies were improved up to 87.54%, 89.72%, and 91.28%. Besides, Parveen et al. [5] applied a hybrid method of Support Vector Machine (SVM) and Fuzzy c-means (FCM) to detect brain lesion. Grey level run length matrix was used for feature extraction and performance evaluated with highest accuracy of 91.66%. In the same year, Kharrat et al.[6] applied 2D wavelet transform and spatial grey level dependence matrix to extract the feature from brain MRI images and simulated annealing to reduce feature size. Then generic algorithm and SVM were used to classify the lesion.

In 2016, Roy et al. [7] applied Adaptive Neuro Fuzzy Inference System, Back propagation neural network and K- Nearest Neighbour (KNN) model to classify the brain lesion from Brain MRI images.

In 2017, Kumar et al. [8] used brain MRI images from SICAS medical respiratory and applied discrete wavelet transform (DWT) to extract features. In the study generic algorithm (GA) was applied for featureselection and SVM to classify the benign and malignant lesion with the highest linear accuracy of 90.9%. In this year, Mohsen et al. [9] combined deep neural network, DWT and principle component analysis (PCA) to classify four classes of lesion, which were glioblastoma, sarcoma, metastatic bronchogenic carcinoma lesion and normal on 66 brain MRI images. Besides, Shree et al. [10] applied region growing for segmentation brain lesion. Then, morphological features including colour, shape, texture, contrast and DWT features were extracted. In the study, probabilistic neural network was used to classify the normal and abnormal tissues.

In 2018, Arun Kumar et al. [11] segmented 200 MRI cases images and based on the result texture and HOG feature were extracted. Then, the feature was used to classify normal and abnormal images of brain lesion. Besides, Bahadure et al. [12] applied watershed, FCM, Discrete cosine transformation, and Berkeley wavelet transformation to segment lesion. Then, GA was applied to classify the images and evaluated performance with 92.03% accuracy. In the same year, Ahmed et al. [13] applied pre trained convolutional neural network (CNN) to classify brain MRI images with the training images from Image Net ILSVRC and achieved the accuracy of 81.81%. In that year, Afshar et al. [14] applied capsule networks (Caps Net), CNN based various architecture to classify brain MRI images and evaluated the performance. In the study, Caps Net based architecture achieved the highest accuracy of 90.89%.

In 2019, Pugalenthi et al. [15] used BraTS2015 dataset to classify the Brain MRI images. In the study, Social Group Optimization algorithm based Fuzzy-TsallisThresholding was applied for pre-processing enhancement of the image and Level-Set Segmentation was used for post processing. Then, performance of achieved result was compared with Active-Contour and Chan-Vese techniques. After the part, GLCM was used to extract the feature and feature selection is achieved through statistical test. Then, SVM with Radial Basis Function is applied to classify the images and the result was compared with Random-Forest and KNN. The study achieved the highest accuracy of 94.33%. In the same year, Gumaei et al [16] proposed a hybrid feature extraction technique based on regularized extreme learning machine, PCA and normalized GIST feature. The feature were used to classify the brain MRI dataset and achieved the improved the performance from 91.51% to 94.233% by using holdout cross-validation. Besides, Anaraki et al. [17] proposed a hybrid approach based on CNN and GA to classify three types of lesion named Pituitary lesion, Meningioma and Glioma. The study achieved an accuracy of 94.2%.

In 2020, Ismael et al. [18] worked on 2D MRI dataset of 233 cancer patients of 3064 images of Pituitary lesion, Meningioma and Glioma. In the study, residual network was applied to classify the dataset and achieved the highest accuracy of 99% on that dataset. In the same year, Hussain et al. [19] applied a hybrid technique on BraTS 2013 dataset. In the study, curvelet transformation, ant colony optimization with Thresholding was used to segment the lesion. Then, PCA reduced skewness approach was used for feature selection and SVM were used to classify the dataset. Within selected features, top 50% features achieved the accuracy of 94.67%.

3.Contribution and outline of the study

The contribution of the study is as follows:-

- To find out the lesion existence in the brain MRI images they are classified in two groups of normal and abnormal. Then, three types of feature named COLOR, GIST and SIFT features are extracted and feature selection is used to get top features. The extracted features are used to classify normal and abnormal lesion by using various SVM, KNN and ensemble methods.
- Besides, severity assessment is also conducted. All the images are divided in two classes and all three types of features are extracted get feature vectors. Feature selection is applied on these feature vectors and top features are used to several SVM, KNN and ensemble method to classify the dataset.
- All the results are used to evaluate the seven-performance factor of the classification techniques of the both classification. Area under curve (AUC) of receiver operation characteristics (ROC) curve is also achieved for all the classification techniques.

The study contains four sections. Material and methods are described in the study. Results and discussions are explained in section three. Conclusion is drawn with the future scope in the section four.

4.Materials and Methods

In thissection, the details of the used dataset and various feature extraction and classification techniques are described.

4.1.Brain MRI image data

The brain MRI images are achieved from the Cancer Genome Atlas (TCGA) dataset. The images are achieved from 110 patients. The MRI dataset is in 3D RGB format with 256 levels. The images are 256×256 pixels. From this dataset, approximately 2500 images have been used for the study including normal and abnormal images. [22]In the dataset radiologist based classification is provided. In the study, 3D RGB images are used for feature extraction and classification.

Table 1

Properties of Brain MRI images.

Dataset characteristics	Category/value
Total number of patients	110
Number of Low-grade glioma (LGG) images	1250 (approximately.)
Number of normal images	2650 (approximately.)
Bit depth	8
Type	RGB
Format	Tag Image File Format (TIF)

4.2.Feature extraction

Feature extraction is the important crucial stage of classification. The dataset contains several brain MRI images, some of which contains brain lesion. Fig. 1 shows the types of images used in the study. In the brain MRI images, there are skull and different tissues. The lesion tissues are different from normal tissues. While using a large dataset then, the lesion tissue images from a brain may be same as the normal tissue of another patient in MRI images. Therefore, finding best feature is important to classify the lesion tissues for all MRI images. MRI images can contain lesion in different position. To find out complete data, complete unwrapped images are used for feature extraction in the study. Table 2 shows the details of the features extracted from a MRI image. To find out effective feature three types of features are extracted, those are followed.

a. In the RGB images, all tissues and skull have different pixel values in the RGB scale. The COLOR feature contains the tissue features related to colour pixel data, based on texture and contrast. Different tissues have different colour pixels in the RGB MRI images. [10]

b. The lesion has different surface data from normal tissues. GIST is a level set technique. GIST modernizes the level data based on partial differential equation and documents the image data by numerical analysis. The surface shape data of the 3D images are achieved from geometric surface deformation. It works by taking different levels of a 3D image. [20]

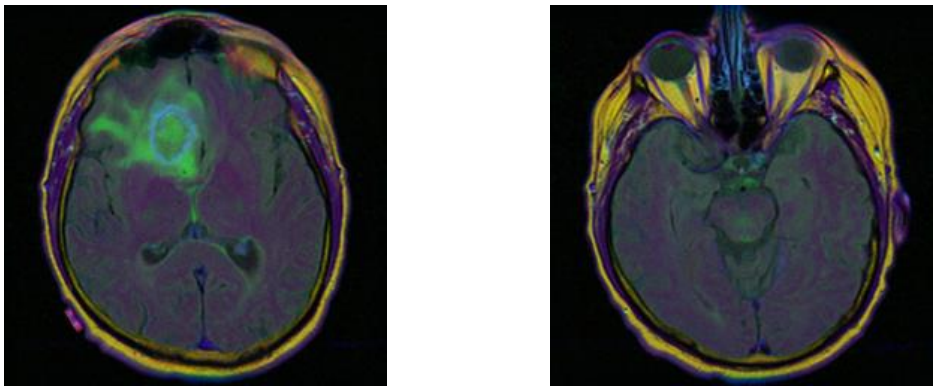
c. In RGB brain MRI images histogram plays an efficient part to classify the abnormal tissues. SIFT features are achieved from local extreme of scale-space Gaussian pyramid and descriptor of histogram achieved from gradient orientations. In the process, scale and rotation invariance's are achieved from descriptor and Gaussian pyramid respectively. Here the variation of gestational age is attenuated by scale invariance. [21]

From the literature it is found that, colour in the images related to texture, surface and histogram is very efficient for 3D RGB images. By using those three-feature techniques, 690 features are achieved from each image.

Table 2

Overview of the extracted feature from Brain MRI images.

Type of the features	Number of features	Properties of the feature	References
COLOR feature	50	Texture and contrast related data achieved by colour.	Shree et al. [10]
GIST feature	512	Level set surface data of geometric surface deformation.	Cates et al. [20]
SIFT feature	128	Scale invariance and rotational invariance.	Keraudren et al. [21]



(a) **(b)**

Fig.1. Brain MRI images of (a) LGG lesion and (b) normal.

4.3.Feature selection

There are some features in the images, which may not efficient for a particular classification due to same correlation. The number of huge feature also consumes more time and needs more memory due to complexity. The correlated features between the classes reduce the performance of the classification. The performance can be improved by using selected efficient features. In the study, statistical feature selection is used by Mann-Whitney U test with Monte Carlo Significance (2-tailed) at 95% confidence interval. From the large number of features (C1, C2…, C50; G1, G2…, G512; S1, S2…, S128) those features are selected, which have the significance value ≤ 0.05.Fig. 2 illustrates feature selection process. The feature selection techniques are followed,

a. Mann-Whitney U test: The test works on null hypothesis to classify two different classes. The statistical significance test evaluates the efficiency of a feature if the feature of a particular class is able to classify the sample with specified class and make a difference from another class. The significance value is 0.05. Any feature, which have larger significance value, that feature is not efficient to classify two different classes. [23]

b. Monte Carlo Significance: The test is used to control the error while calculating a statistical test i.e. positive false discovery rate. At the feature selection technique several iterations have been done. The test controls the error rate for each stage of calculation. If the statistics correlate, then no adjustment is made to generate the significant value. The tests uses joint distribution of the statistics for accurate calculation.[24]

Based on the feature selection techniques there are different selected features for two different classifications. Those are followed,

a. Feature selection for Abnormalitydetection: For the classification of brain lesion detection, two different classes are achieved i.e. normal and abnormal, which contains lesion. Out of 690 features, 512 features are selected for the classification. The features have the Mann-Whitney U test value ≤ 0.05 with Monte Carlo Significance (2-tailed) at 95% confidence interval.

b. Feature selection for Severity assessment: Toclassify the risk factor two different classes have been achieved based on the death i.e.if the patient is died or not. Total 547 features have been selected out of 690 features. The selected features have the significance value ≤ 0.05 for the Mann-Whitney U test with Monte Carlo Significance (2-tailed) at 95% confidence interval.

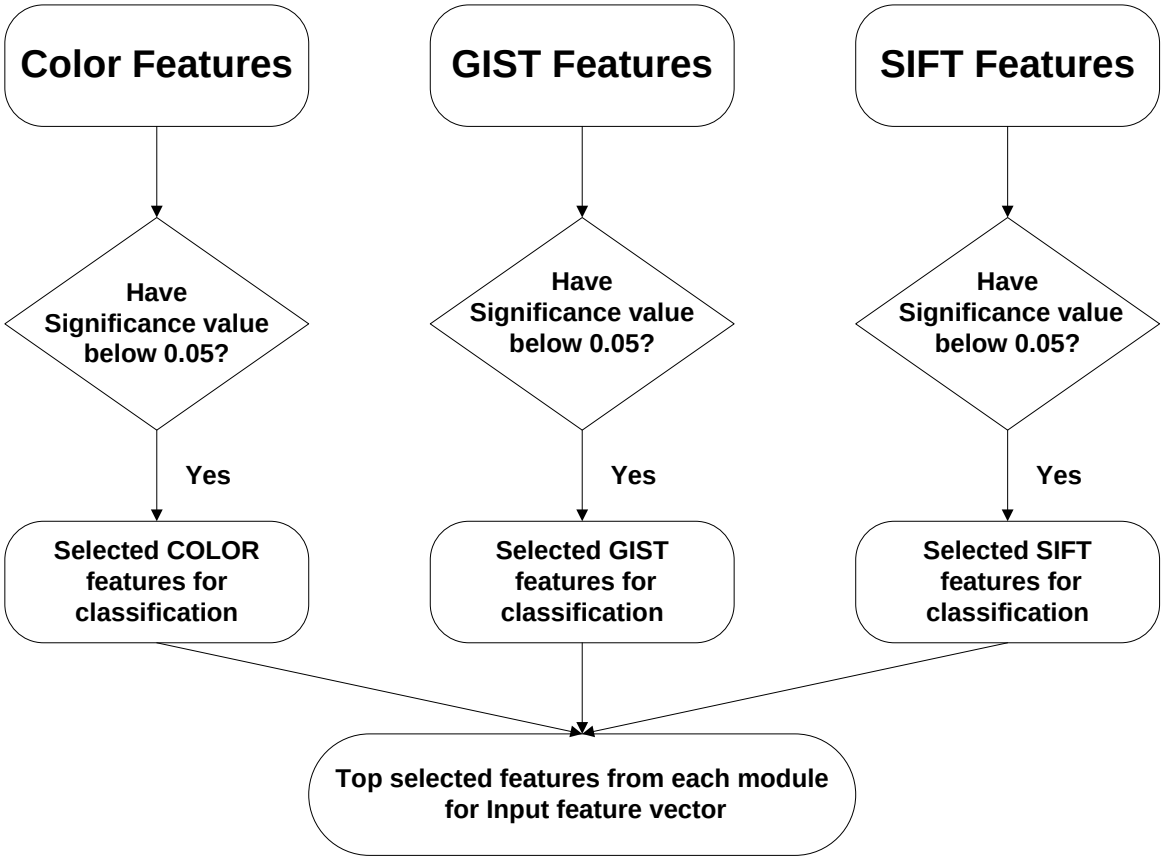


Fig.2. Feature selection procedure. Three feature selection modules are combined for input feature vector.

4.4.Classification

In the study, several types of classification techniques are applied to classify two different kinds of classification. For those classification 3D brain MRI images, features are used as feature vectors. Those are various kinds of support vector machine (SVM), k nearest neighbour (KNN) and ensemble methods.

4.4.1.SVM

SVM is a technique, which can minimize the classification error by maximizing the distance margin between the classes. SVM converts the data vector into a higher dimensional space. Those data of each class are separated by hyper planes as a margin. With increment of the distance between those two hyper plane, the error of the classifier is reduced. [25] For the data points $\{(U_1,V_1), (U_2,V_2), (U_3,V_3)\ldots, (U_N,V_N)\}$. Where, $V_N = 1/-1$ a constant determining the class where $U_N$belongs and N is number of sample. Here, U_N is a M dimensional input vector, where M is the number of patients in the classification model.

For a training data to separate hyper plane,

$$W * U + B = 0$$

Where, W is M dimensional vector and B is a scalar. B is offset parameter, which increases the distance between parallelhyper plane and vector W points perpendicular to the hyper plane. For $B = 0$ the hyperplane passes through the origin. The equation of parallel hyperplanes is,

$$W * U + B = 1$$

$$W * U + B = -1$$

The distance between the hyper plane is $2/|\ W\ |$. Minimum value $|\ W\ |$ is important for maximum value of distance between parallel hyperplane. Fig. 3 illustrates the SVM technique in a linear model.

$$W * U_i - B \geq 1\ or, W * U_i - B \leq -1$$

This can be expressed as,

$$V_i(W * U_i - B) \geq 1\ ;\ 1 \leq i \leq N$$

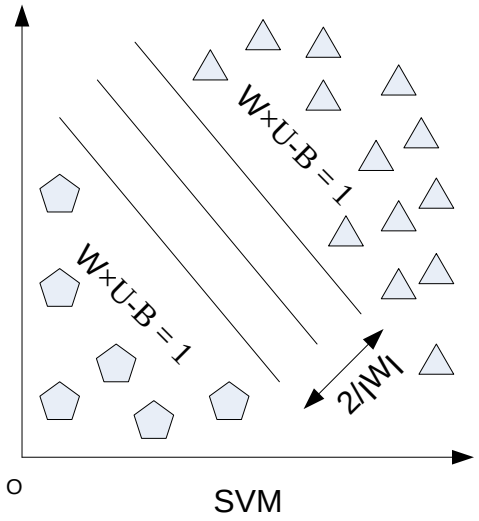


Fig. 3. SVM technique in graph.

There are different geometry of separating hyper plane based on the classification data. Based on different geometric figure there are various SVM techniques. Fig. 4 and Fig. 5demonstrate the various types of SVM techniques.

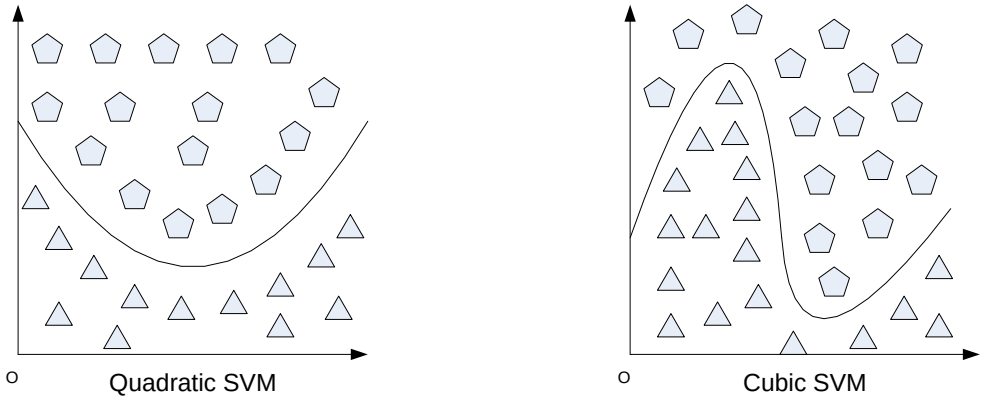


Fig. 4. Quadratic SVM and Cubic SVM in graph (From left).

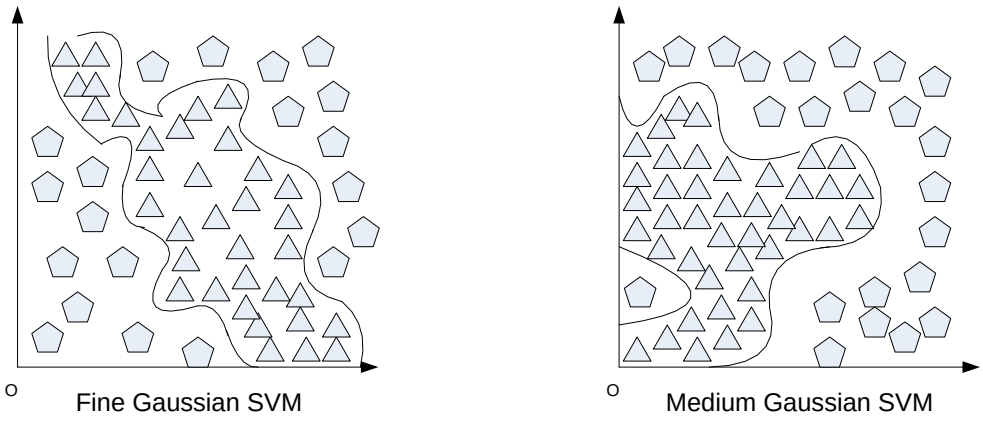


Fig. 5. Fine Gaussian SVM and Medium Gaussian SVM in graph (From left).

4.4.2.KNN

KNNclassifies the testing data based on Euclidian distance, which is achieved from the training data. For the sample U_i with Q features (U_{i1}, U_{i2}…, U_{iQ}), where total number of input samples is N. The Euclidian distance between two samples is,

$$d(U_i, U_l) = \sqrt{(U_{i1} - U_{l1})^2 + (U_{i2} - U_{l2})^2 + \cdots + \left(U_{iQ} - U_{lQ}\right)^2}$$

Where $i, l = 1, 2 \ldots, N$

If, U_i is a training sample, and U is a test sample, ω is the true class and ώ is the predicted class for $\omega, \acute{\omega} = 1, 2 \ldots, \Omega$. Here, $\Omega$is the total number of classes. During training ω is trained for each train sample and ώ is tested during testing. [26]

For one nearest neighbour U is same as the true class ω of its nearest neighbour and V_i is the nearest neighbour for the distance,

$$d(V_i, U) = min_j\{d\left(V_j, U\right)\}$$

For the K nearest neighbour, the ώ of the test sample U is one class between K training class samples. Fig. 7 illustrates the KNN techniques.

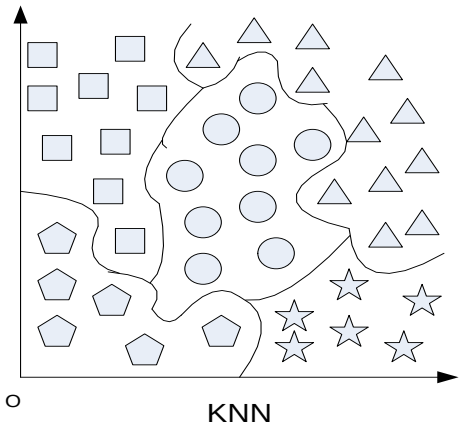


Fig. 6. KNN technique in graph.

4.4.3.Ensemble

Ensembles are one kind of generalized additive models. It predicts new data point by the achieved weighted summation of element models. It is a cascaded technique, which apply the base learning algorithm several times. Most of the ensembles are based on decision tree algorithm. Due to suffering from high variance, ensemble learning can be used instead of the decision tree. As base learning method suffers from various problem, i.e. statistical, computational and representational problems, where ensemble learning achieves the fundamental function more accurately. There are various ensemble-learning methods based on element models of an ensemble. [27]

4.4.4.Proposed CAD system

The proposed CAD system architecture for brain MRI lesion detection and severity assessment classification by cascaded machine learning technique isshownin Fig. 7. A dotted line divides two classification sections of the cascaded system. The TCGA dataset contains 3D RGB brain MRI images. Those images are used for feature extraction purpose to find input feature vector. Three different type of features are extracted and feature selection is achieved by filter method. For the lesion detection, part 512 features are selected out of 690 features. While, for the risk factor 547 image features are selected from 690 features. At first the selected features, training and testing samples are selected by cross validation. Then the training samples are used to train aabnormal detectectionof the cascaded model. The testing samples are used to test the models. For the second stage, feature vector samples for severity assessment are selected by matching the result of the abnormal samples from the result of the first stage. Then, cross validation is used to divide the samples into training and testing samples. After that, testing samples are used to achieve risk factor. At the end the result of both classification models are used for performance evaluation purpose

The images from the dataset are used for the following,

a. The dataset contains the ground truth of each patient sample. The levels are denoted based on the tissue texture, colour, contrast and other magnetic resonance data.

b. The dataset images are used for the cascaded CAD model for automatic detection. Feature extraction, feature selection and classification by various SVM, KNN and ensemble method to detect lesion and risk factor. The results of the classifiers are compared with the ground truth data. The architecture of SVM, KNN and ensemble methods are explained in section 2.4.1, 2.4.2, and 2.4.3 respectively.

Proper selection of classifiers for the first and the second stage is important for the cascaded cad system. All the classification techniques have the accuracy more than 84% and 96% for the first and second stage respectively. For the first stage classification KNN based ensemble achieved the highest result, while for second stage both SVM and KNN achieved the highest result.

4.4.5.Performance evaluation

In the study, hold out and k-fold both data division are applied for both lesion existence and severity factor classification. A dataset is divided for classification as training data and testing data. Hold out is such a method where the training and testing data can be divided. The process of the hold out cross validation in the study is followed,

a. The data samples are divided into training and testing sample with randomly 33% sample in testing data.
b. The training data with the selected features are trained to a classifier model.
c. The trained model is used to test the testing data.

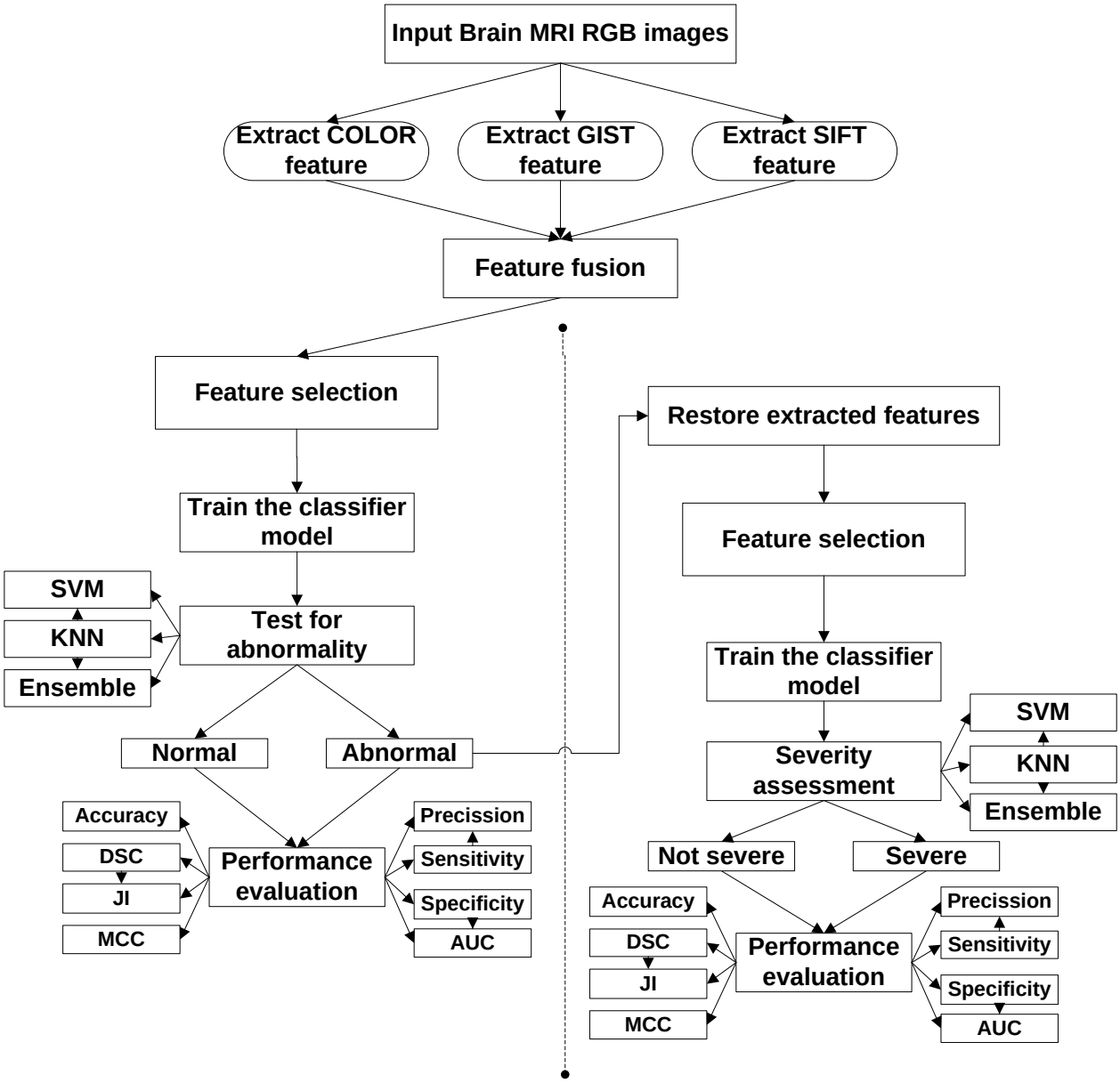


Fig.7. Proposed methodology.

However, for a particular training and testing data the classification is not accurate. The performance of the classification can be changed with the variation of the training and testing dataset. For k-fold cross validation 5 fold and 10 fold both cross validation are applied. For a particular value of k, the dataset is divided to k different part. Then, one of the all parts are tested while all of other dataset are used for training the classifier. The process is repeated k times and for each epoch, one different dataset is used for testing. In the study, 5 fold and 10 fold cross validation are applied. The process of the k-fold cross validation as follows

a. For k-fold, the data samples are divided into k different parts.
b. All dataset is used for training except leaving one dataset for testing.
c. The trained classifier models are applied to test the testing dataset.
d. Repeat the process k times by taking the different testing dataset for each epoch.

Table 3

Performance evaluation used in the study.

Performance evaluation	Explanation	Mathematical formula
Accuracy	Total percentage of correctly classified classes.	$(\tau_{tpos} + \tau_{tnef})/(\tau_{fpos} + \tau_{fnef} + \tau_{tpos} + \tau_{tnef})$
DSC	The probabilistic similarity of the classification model.	$(2 * \tau_{tpos})/(2 * \tau_{tpos} + \tau_{fpos} + \tau_{fnef})$
JI		$\mathrm{DSC}/(2 - \mathrm{DSC})$
MCC	Machine learning performance.	$\frac{(\tau_{tpos} * \tau_{tnef} - \tau_{fpos} * \tau_{fnef})}{\sqrt{((\tau_{tpos} + \tau_{fpos}) \times (\tau_{tpos} + \tau_{fnef}) \times (\tau_{tnef} + \tau_{fpos}) \times (\tau_{tnef} + \tau_{fnef}))}}$
Precision	The relevancy of positive prediction.	$\tau_{tpos}/(\tau_{tpos} + \tau_{fpos})$
Sensitivity	Total percentage of correctly classified positive class.	$\tau_{tpos}/(\tau_{tpos} + \tau_{fnef})$
Specificity	Total percentage of correctly classified negative class.	$\tau_{tnef}/(\tau_{tnef} + \tau_{fpos})$
AUC	Performance curve between sensitivity and specificity.	$\frac{1}{2}\left(\frac{\tau_{tpos}}{\tau_{tpos} + \mathrm{FN}} + \frac{\tau_{tnef}}{\tau_{tnef} + \tau_{fpos}}\right) \times 100\%$

DSC- Dice similarity coefficient; JI- Jaccard index; MCC- Mathew's correlation coefficient; AUC- Area under curve of receiver operating characteristics curve.

All the result data are achieved by the confusion matrix. Based on the confusion matrix the performance of a classifier model is evaluated for both classification technique. Table 3 illustrates the performance evaluation parameters. According to the confusion matrix, the result is divided into four different classes. Those are true positive (τ_{tpos}), true negative (τ_{tneg}), false positive (τ_{fpos}), false negative (τ_{fneg}).Those four classes are used to measure the performance parameter. The study is executed in MATLAB® R2018a.

5.Results and Discussions

Brain MRI images have noise, bias field, blur low contrast and partial volume effect (PVE). Due to the bias field, different tissues have low variation, which makes difficulty to find out the lesion accurately. The PVE multiple tissues are in a voxel. Due to that, effect accurate lesion detection from brain MRI is a challenging task. [28] To find out the solution of the problem, CAD system are proposed for automatic lesion detection. For the precise performance CAD system requires to proper operation of different parameters.

Classifier is the chief part of a CAD system. A CAD system have some difficulty due to number of images used to train the model, lack of efficient features, classifier complexity, time requirement, which reduces the performance of the CAD model. Some ensemble based and hybrid classifiers proposed to develop performance of the CAD system. [29, 30] The chief objective is to find out consistent CAD model. Different images from various patients with radiologist marking are used to make a robust CAD model. In the study, various classifiers of SVM, KNN and ensemble are used to find out accurate classifiers for each stage of the cascaded CAD system.

Table 4

Selected features for cancer existence classification.

Type of feature selection technique	Feature selection technique	Number of selected features	Selected features

Filter	Mann-Whitney U	20	C1, C4, C10-C13, C17, C20, C21, C24, C25, C30, C32, C33, C40, C45,C46,C47, C49, C50
Filter	Mann-Whitney U	398	G3, G5-G11, G13, G14, G15, G17, G21-G25, G27, G31, G33, G35, G37-G40, G42, G44, G45, G47, G50-G60, G63, G67, G69, G70, G71, G74, G75, G76, G79, G83-G92, G95, G99, G100, G102-G107, G111, G114-G123, G127, G130-G139, G142, G143, G146, G147, G149-G151, G153-G156, G158, G159, G162, G163, G165-G168, G170-G176, G178, G179-G188, G190, G191, G192, G194, G195, G197-G204, G206, G207, G210-G220, G222, G223, G227, G228, G230-G236, G238, G239, G242-G251, G254, G255, G258, G259, G261-267, G270, G271, G272, G274, G275, G277, G278, G279, G281-G284, G286, G287, G289, G290, G291, G293-G296, G298-G303, G306, G307, G309-G319, G322-G332, G334-G339, G341-G348, G350-G356, G358-G367, G370-G379, G382, G383, G384, G386-G396, G398-G412, G414-G431, G434-G448, G450-G460, G462-G476, G478-G496, G498-G508, G510, G511, G512
Filter	Mann-Whitney U	94	S1-S5, S8-S11, S13, S16, S17, S18, S20, S21, S22, S24, S25, S26, S28, S29, S30, S32-S38, S40-S46, S48, S49, S50, S53, S54, S56, S57, S60, S61, S62, S64-S70, S73-S78, S80, S81, S84, S85, S86, S88, S89, S90, S93, S94, S96-S102, S104, S105, S107-S110, S113, S114, S116-S119, S121, S122, S124, S125, S126, S128

Table 4, shows, feature selection techniques for the lesion detection classification of the cascaded CAD system. In the study, filter based feature selection is achieved. From COLOR, GIST and SIFT features 2, 398 and 94 features are selected. All the selected features are listed in the table, which are used for first stage of CAD system. Most of the GIST features are selected by the filter method. Table 4, shows GIST features as an important features than for the lesion detection classification, while less number of COLOR features are selected for stage of classification.

Table 5

Selected features for death risk classification.

Type of feature selection technique	Feature selection technique	Number of selected features	Selected features
Filter	Mann-Whitney U	27	C1, C2, C3, C9, C11, C13, C15, C17-C20, C22, C23, C25, C28, C29, C30, C35, C40, C41, C42, C44-C47, C49, C50
Filter	Mann-Whitney U	429	G1-G9, G11-G41, G43-G134, G136, G137, G140-G154, G156-G170, G172-G186, G188-G198, G200, G201, G202, G204-G214, G216, G217, G218, G220-230, G232-G246, G248-G262, G264, G265, G268, G269, G270, G273, G274, G275, G277, G278, G279, G281-G291, G293, G294, G295, G297, G298, G300-G313, G316-G326, G329, G332-342, G345, G346, G349-G357, G360-G363, G365-G370, G372, G373, G376-G379, G281, G382, G383, G385-G393, G396, G399-G403, G405, G406, G407, G409, G410, G414, G415, G417, G418, G420-G423, G425, G426, G432-G439, G441, G445-G454, G457, G458, G460-G470, G472, G473, G475, G477, G478, G79, G485, G487, G489-G492, G494, G495, G496, G498-G502, G505-G508, G511, G512
Filter	Mann-Whitney U	91	S1, S2, S3, S5, S7, S8, S9, S11, S13, S16, S17, S19, S21, S23-S27, S29, S31, S32, S33, S36, S37, S39, S41, S44, S45, S46, S49-S53, S55-S61, S63-S66, S69, S73, S74, S77, S80-S83, S85, S86, S87, S89-S103, S105-S110, S112-S121, S123, S125, S126, S127

Table 6

Performance evaluation for cancer existence classification by using 5-fold cross-validation.

Classification techniques	**Performance measures**							
	Accuracy	**DSC**	**JI**	**MCC**	**Precision**	**Sensitivity**	**Specificity**	**AUC**
Cubic SVM	0.923	0.923	0.857	0.846	0.921	0.925	0.921	0.923
Quadratic SVM	0.881	0.881	0.787	0.762	0.877	0.885	0.876	0.881
Fine Gaussian SVM	0.893	0.883	0.790	0.797	0.973	0.808	0.977	0.893
Medium Gaussian SVM	0.846	0.849	0.738	0.693	0.830	0.869	0.823	0.846
Fine KNN	0.943	0.943	0.891	0.885	0.943	0.943	0.943	0.943
Weighted KNN	0.856	0.852	0.743	0.713	0.873	0.833	0.879	0.856
Bagged tree ensemble	0.886	0.888	0.798	0.772	0.871	0.906	0.866	0.886
Subspace KNN ensemble	0.934	0.934	0.876	0.867	0.930	0.938	0.930	0.934

Table 7

Performance evaluation for cancer existence classification by using 10-fold cross-validation.

Classification techniques	**Performance measures**							
	Accuracy	**DSC**	**JI**	**MCC**	**Precision**	**Sensitivity**	**Specificity**	**AUC**
Cubic SVM	0.929	0.929	0.868	0.859	0.928	0.931	0.928	0.929
Quadratic SVM	0.893	0.891	0.804	0.785	0.899	0.884	0.901	0.893
Fine Gaussian SVM	0.894	0.884	0.791	0.799	0.974	0.808	0.979	0.894
Medium Gaussian SVM	0.862	0.864	0.760	0.724	0.852	0.87	0.848	0.862
Fine KNN	0.942	0.943	0.892	0.885	0.934	0.952	0.933	0.942
Weighted KNN	0.903	0.905	0.827	0.807	0.885	0.926	0.880	0.903
Bagged tree ensemble	0.880	0.881	0.788	0.761	0.873	0.890	0.871	0.880
Subspace KNN ensemble	0.944	0.944	0.894	0.887	0.935	0.953	0.934	0.944

Table 8

Performance evaluation for cancer existence classification by using 33% holdout cross-validation.

Classification techniques	**Performance measures**							
	Accuracy	**DSC**	**JI**	**MCC**	**Precision**	**Sensitivity**	**Specificity**	**AUC**
Cubic SVM	0.920	0.920	0.852	0.841	0.921	0.919	0.922	0.920
Quadratic SVM	0.881	0.881	0.787	0.762	0.883	0.879	0.884	0.881
Fine Gaussian SVM	0.889	0.877	0.782	0.794	0.982	0.793	0.986	0.889
Medium Gaussian SVM	0.859	0.863	0.759	0.719	0.837	0.890	0.827	0.859
Fine KNN	0.926	0.926	0.862	0.853	0.932	0.91	0.933	0.926
Weighted KNN	0.899	0.897	0.814	0.798	0.910	0.886	0.912	0.899
Bagged tree ensemble	0.886	0.890	0.802	0.774	0.858	0.924	0.848	0.886
Subspace KNN ensemble	0.931	0.931	0.870	0.862	0.935	0.926	0.936	0.931

Table 9

Performance evaluation for death risk classification by using 5-fold cross-validation.

Classification techniques	**Performance measures**							
	Accuracy	**DSC**	**JI**	**MCC**	**Precision**	**Sensitivity**	**Specificity**	**AUC**
Quadratic SVM	0.999	0.999	0.998	0.998	0.999	0.999	0.999	0.999
Fine Gaussian SVM	0.968	0.970	0.942	0.937	0.942	1.000	0.931	0.966
Medium Gaussian SVM	0.994	0.994	0.989	0.988	0.994	0.994	0.994	0.994
Cubic SVM	0.999	0.999	0.999	0.999	1.000	0.999	1.000	0.999
Fine KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Medium KNN	0.965	0.967	0.936	0.931	0.983	0.951	0.981	0.966
Cosine KNN	0.969	0.970	0.942	0.938	0.979	0.962	0.976	0.969
Weighted KNN	0.989	0.98	0.979	0.978	0.999	0.980	0.999	0.989
Boosted Trees ensemble	0.982	0.983	0.966	0.964	0.990	0.976	0.989	0.982

Bagged Trees ensemble	0.989	0.990	0.980	0.979	0.985	0.996	0.983	0.989
Subspace Discriminate Ensemble	0.973	0.974	0.949	0.946	0.992	0.957	0.991	0.974
Subspace KNN Ensemble	0.999	0.999	0.999	0.999	1.000	0.999	1.000	0.999

Table 5, shows the selected features for the severity assessment part of the cascaded CAD system. For the second stage of the CAD model, 27 CLOR features, 429 GIST features and 91 SIFT features are selected by the filter feature selection method. For this stage, most of GIST features are selected among all. But for the risk factor detection more number of COLOR are selected than, the previous stage while, less number of SIFT features are selected for that stage. For, severity assessment COLOR features are useful with the GIST features. Comparing to Table 4, more features are selected for this stage. The tables illustrates the importance of different features to classify different models. Besides, there are several features in both the classification models.

Table 6, illustrates the lesion existence classification performance. For the first stage, 5-fold cross validation is used to separate training and testing data. Here, the highest 93.4% accuracy and DSC are achieved by fine KNN and subspace KNN ensemble. For the selected cross validation, fine KNN achieved 89.1% JI, 94.3% precision, sensitivity, specificity and AUC while subspace KNN achieved 86.7% MCC as the highest performance.

Table 10

Performance evaluation for death risk classification by using 10-fold cross-validation.

Classification techniques	**Performance measures**							
	Accuracy	**DSC**	**JI**	**MCC**	**Precision**	**Sensitivity**	**Specificity**	**AUC**
Quadratic SVM	0.999	0.999	0.998	0.998	0.999	0.999	0.999	0.999
Fine Gaussian SVM	0.973	0.975	0.951	0.947	0.951	1.000	0.943	0.971
Medium Gaussian SVM	0.996	0.997	0.993	0.993	0.998	0.996	0.998	0.997
Cubic SVM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Fine KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Medium KNN	0.999	0.999	0.999	0.999	1.000	0.999	1.000	0.999
Cosine KNN	0.976	0.977	0.955	0.952	0.989	0.966	0.988	0.977
Weighted KNN	0.994	0.994	0.988	0.987	1.000	0.988	1.000	0.994
Boosted Trees ensemble	0.986	0.987	0.974	0.972	0.991	0.982	0.990	0.986
Bagged Trees ensemble	0.994	0.994	0.988	0.987	0.991	0.997	0.990	0.993
Subspace Discriminate Ensemble	0.976	0.977	0.955	0.952	0.992	0.962	0.991	0.977
Subspace KNN Ensemble	0.999	0.999	0.999	0.999	1.000	0.999	1.000	0.999

Table 7, shows the performance for lesion detection classification with 10-fold cross validation. From the table, subspace KNN ensemble method achieved the highest performance. The classifier achieves 94.4%

accuracy, DSC and AUC. Besides, 89.4% JI, 88.7% MCC, 93.5% precision, 95.3% sensitivity and 93.4% specificity is also drawn by the classifier. For the stage, all eight classifiers have attained the accuracy more than 85%.

Table 8, demonstrates the performance of first stage classification of the cascaded model with 33% hold-out cross validation. The module suffers due to limited number of samples. The highest 93.1% accuracy, DSC and AUC are achieved by subspace KNN. The classifier attains 87% JI, 86.2% MCC, 93.5% precision, 92.6% sensitivity and 93.6% specificity. For the cross validation eight classifiers reaches the accuracy more than 85%.

Table 11

Performance evaluation for death risk classification by using 33% holdout cross-validation.

Classification techniques	**Performance measures**							
	Accuracy	**DSC**	**JI**	**MCC**	**Precision**	**Sensitivity**	**Specificity**	**AUC**
Quadratic SVM	0.998	0.998	0.997	0.996	1.000	0.997	1.000	0.998
Fine Gaussian SVM	0.945	0.950	0.905	0.894	0.905	1.000	0.883	0.941
Medium Gaussian SVM	0.993	0.993	0.987	0.986	0.993	0.993	0.992	0.993
Cubic SVM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Fine KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Medium KNN	0.961	0.963	0.928	0.921	0.969	0.956	0.966	0.961
Cosine KNN	0.959	0.961	0.925	0.918	0.966	0.956	0.962	0.959
Weighted KNN	0.988	0.988	0.976	0.975	1.000	0.976	1.000	0.988
Boosted Trees ensemble	0.991	0.992	0.983	0.982	0.997	0.987	0.996	0.991
Bagged Trees ensemble	0.991	0.992	0.983	0.982	0.987	0.997	0.985	0.991
Subspace Discriminant Ensemble	0.977	0.978	0.957	0.954	0.986	0.970	0.985	0.977
Subspace KNN Ensemble	0.998	0.998	0.997	0.996	1.000	0.997	1.000	0.998

Table 9, clarifies the performance of severity assessment classification. For the stage of the cascaded CAD system, 5-fold cross validation is applied. For the stage, the performance is higher than the first stage classification. Here, the highest performance is accomplished by fine KNN, though all the classifiers have achieved the accuracy more than 96%. KNN reached 100% accuracy, DSC, JI, MCC, precision, sensitivity, specificity and AUC.

Table 10, illustrates the second stage classification performance with 10-fold cross validation. For the stage cubic SVM and fine KNN succeeds highest 100% accuracy, DSC, JI, MCC, precision, sensitivity, specificity, AUC. Here, quadratic SVM, medium KNN and subspace KNN ensemble achieves accuracy 99.9%. In the stage, all twelve classifiers have accomplished the accuracy more than 97% and other performance more than 95%.

Table 11, demonstrates the performance parameters for the risk factor classification. For the stage, cubic SVM and fine KNN achieves 100% performance for all performance parameters. Here, quadratic SVM, weighted KNN and subspace KNN ensemble achieves 99.8% accuracy while all the classifiers have reached the performance more than 90%.

The performance of the cascaded CAD system is varied with different classification models and different classifiers. For the drawback, different classifier can be used for the cascaded classification model. The performance is evaluated with the marking details provided with the dataset. There are different patients and different MRI images. To build a robust CAD model more patients can be included. Besides, the radiologist to involve different stages of classification can verify the cascaded model.

6.Conclusions and Future Scopes

The study detonates the prospect of multistage CAD model with the included radiologist marking. The proposed cascaded neural network approach is developed by combining different classifiers of SVM, KNN and ensemble based models for the lesion detection and risk factor classification. The performance of the cascaded CAD system is evaluated in each stage with various classifiers. Different features are selected for two different stage of classification. For the first and the second stage, 512 and 547 features are selected out of 690 features. The chief finding in the study are as follows (i) different feature selection for different classification; (ii) lesion detection classification by eight classification techniques; (iii) severity assessment classification by twelve classifiers; (iv) performance evaluation of each classifier for two stage of classification with the provided ground truth with the dataset.

The robustness and performance of the study can be improved in numerous methods: (i) the ground truth data can be used to implement classification for histopathology and biopsy; (ii) multi-model images can be used for robust CAD system. Brain CT scan, ultrasound images can be added with the brain MRI images;(iii) radiologist based CAD system can be implemented; (iv) live data can be used for testing purpose to implement integrated CAD system; (v) histopathological details also classified with the CAD system. We have not used live data due to unavailability of the public data with radiologist marking. Researchers on their own dataset implement most of the study. A public domain dataset is required with ground truth for implementation of all technique in a common platform.

References

1. Yearly rate of malignant brain lesion for American patients in 2020. Available online at https://braintumor.org/brain-tumor-information/brain-tumor-facts.[accessed30 December 2021]
2. American Cancer Society data for malignant brain lesion patients in USA. Available online at https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/about/key-statistics.html#:~:text=The%20American%20Cancer%20Society's%20estimates,in%20females)%20will%20be%20diagnosed. [accessed 30 December 2021]
3. Dandıl, E., Çakıroğlu, M., &Ekşi, Z. (2014, September). Computer-aided diagnosis of malign and benign brain tumors on MR images. In International Conference on ICT Innovations (pp. 157-166). Springer, Cham.
4. Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., ... & Feng, Q. (2015). Enhanced performance of brain tumor classification via tumor region augmentation and partition. PloS one, 10(10), e0140381.
5. Singh, A. (2015, February). Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM. In 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 98-102). IEEE.
6. Kharrat, A., Halima, M. B., &Ayed, M. B. (2015, December). MRI brain tumor classification using support vector machines and meta-heuristic method. In 2015 15th International Conference on Intelligent Systems Design and Applications (ISDA) (pp. 446-451). IEEE.
7. Roy, S., Sadhu, S., Bandyopadhyay, S. K., Bhattacharyya, D., & Kim, T. H. (2016). Brain tumor classification using adaptive neuro-fuzzy inference system from MRI. International Journal of Bio-Science and Bio-Technology, 8(3), 203-218.
8. Kumar, S., Dabas, C., &Godara, S. (2017). Classification of brain MRI tumor images: A hybrid approach. Procedia computer science, 122, 510-517.
9. 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.
10. Shree, N. V., & Kumar, T. N. R. (2018). Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. Brain informatics, 5(1), 23-30.
11. Arunkumar, N., Mohammed, M. A., Mostafa, S. A., Ibrahim, D. A., Rodrigues, J. J., & de Albuquerque, V. H. C. (2020). Fully automatic model-based segmentation and classification approach for MRI brain tumor using artificial neural networks. Concurrency and Computation: Practice and Experience, 32(1), e4962.

12. Bahadure, N. B., Ray, A. K., &Thethi, H. P. (2018). Comparative approach of MRI-based brain tumor segmentation and classification using genetic algorithm. Journal of digital imaging, 31(4), 477-489.
13. Ahmed, K. B., Hall, L. O., Goldgof, D. B., Liu, R., &Gatenby, R. A. (2017, March). Fine-tuning convolutional deep features for MRI based brain tumor classification. In Medical Imaging 2017: Computer-Aided Diagnosis (Vol. 10134, p. 101342E). International Society for Optics and Photonics.
14. Afshar, P., Plataniotis, K. N., &Mohammadi, A. (2019, May). Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1368-1372). IEEE.
15. Pugalenthi, R., Rajakumar, M. P., Ramya, J., & Rajinikanth, V. (2019). Evaluation and classification of the brain tumor MRI using machine learning technique. Journal of Control Engineering and Applied Informatics, 21(4), 12-21.
16. Gumaei, A., Hassan, M. M., Hassan, M. R., Alelaiwi, A., &Fortino, G. (2019). A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. IEEE Access, 7, 36266-36273.
17. Anaraki, A. K., Ayati, M., &Kazemi, F. (2019). Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybernetics and Biomedical Engineering, 39(1), 63-74.
18. Ismael, S. A. A., Mohammed, A., &Hefny, H. (2020). An enhanced deep learning approach for brain cancer MRI images classification using residual networks. Artificial Intelligence in Medicine, 102, 101779.
19. Hussain, U. N., Khan, M. A., Lali, I. U., Javed, K., Ashraf, I., Tariq, J., ... & Din, A. (2020). A Unified Design of ACO and Skewness based Brain Tumor Segmentation and Classification from MRI Scans. Journal of Control Engineering and Applied Informatics, 22(2), 43-55.
20. Cates, J. E., Lefohn, A. E., & Whitaker, R. T. (2004). GIST: an interactive, GPU-based level set segmentation tool for 3D medical images. Medical image analysis, 8(3), 217-231.
21. Keraudren, K., Kyriakopoulou, V., Rutherford, M., Hajnal, J. V., &Rueckert, D. (2013, September). Localisation of the brain in fetal MRI using bundled SIFT features. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 582-589). Springer, Berlin, Heidelberg.
22. Mazurowski, M. A., Clark, K., Czarnek, N. M., Shamsesfandabadi, P., Peters, K. B., &Saha, A. (2017). Radiogenomics of lower-grade glioma: algorithmically-assessedtumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with The Cancer Genome Atlas data. Journal of neuro-oncology, 133(1), 27-35.
23. McKnight, P. E., &Najab, J. (2010). Mann-Whitney U Test. The Corsiniencyclopedia of psychology, 1-1.
24. Lin, D. Y. (2005). An efficient Monte Carlo approach to assessing statistical significance in genomic studies. Bioinformatics, 21(6), 781-787.
25. Durgesh, K. S., &Lekha, B. (2010). Data classification using support vector machine. Journal of theoretical and applied information technology, 12(1), 1-7.
26. Peterson, L. E. (2009). K-nearest neighbor. Scholarpedia, 4(2), 1883.
27. Dietterich, T. G. (2002). Ensemble learning. The handbook of brain theory and neural networks, 2, 110-125.
28. Roy, S., Nag, S., Maitra, I. K., & Bandyopadhyay, S. K. (2013). A review on automated brain tumor detection and segmentation from MRI of brain. arXiv preprint arXiv:1312.6150.
29. Singh, B. K., Verma, K., &Thoke, A. S. (2016). Fuzzy cluster based neural network classifier for classifying breast tumors in ultrasound images. Expert Systems with Applications, 66, 114-123.
30. Singh, B. K., Verma, K., Thoke, A. S., & Suri, J. S. (2017). Risk stratification of 2D ultrasound-based breast lesions using hybrid feature selection in machine learning paradigm. Measurement, 105, 146-157.