International Journal of Engineering & Technology, 7 (1) (2018) 140-146

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i1.8571
Research paper

Classification of brain tumor types using multiclass kernel-based Hellinger decision method for HD-Tree and HD-Forest

Nikita Singh [1] *, Naveen Choudhary [2]

[1] Department of Computer Science Engineering, College of Technology and Engineering
[2] Maharana Pratap University of Agriculture and Technology, Udaipur, 313001, India
**Corresponding author E-mail: nikisnikita@gmail.com*

Abstract

Currently, the radiologist needs to distinguish the medical imaging with their multiple classes. In this paper, we work on several steps: segmented ROI, feature extraction of ROI and classification. In this work, we proposed a multiclass kernel based Hellinger decision method HD-Tree and HD-Forest for the classification of brain tumor classes with respect to classification time and accuracy. The calculated features like patient symptoms, centroid, shape, etc. are used in the classification scheme. Total 97 MRI brain tumor images (Astrocytoma (22), Ganglioglioma (6), Glioblastoma (23), Epidermoid (3), Mixed Glioma (5) and Meningnet (41)) were used for the experiment. The Experimental result shows that kernel-based Hellinger HD-Tree was found to be 96.50 % of accuracy and HD-Forest was found to be 99.9%. In this paper, we compare our proposed method LA-SVM method. LA-SVM was found to be 96% of accuracy. We can see that HD-forest gives the best accuracy result.

Keywords: *MRI Brain Tumor; Feature Extraction; Classification; Kernel HD-Tree; Kernel HD-Forest; SVM.*

1. Introduction

Brain tumor on MRI is very important for all possible disease stages through clinical treatment. There are various types of medical imaging techniques are available for the identification of particular patient disease these are Magnetic resonance imaging (MRI), Computer Tomography (CT) and Ultrasound Sound (US) [10]. MRI imaging technique basically used to identify brain-related problems [9]. Because MRI can provide the information about human tissues, higher contrast resolution power, and multi-specialty characteristics [7], [8], [11], [18]. Tumor segmentation, Tumor type characterization, and classification are very difficult, because of the heterogeneous image profiles. The brain tumor includes primary and secondary images are difficult to segment or analysis or classified. To detect the abnormality in MR image the image processing based segmentation methods are used. The detection

Process is very challenging for automatic identification of abnormal image [15], [2], [3] complicated tissue or structures. There are two type brain tumor difficulties occurred, Primary and secondary tumors. This is Astrocytoma, Glioblastoma Multiforme, Ganglioglioma, Gliomas, Meningioma. Metastatic tumors. According to this difficulty, the objective of this work is to provide the computer-aided system tool for characterization and to analysis the multicasts classification for different types of Brain's tumor. These multi-classes include primary and secondary tumors into different classes. This objective is very important for the radiologist to make a decision on the patient-related problem and treatment [7], [9]. The accurate CAD segmentation and classification process provide the reliable differentiation when images are heterogonous and fast to analysis. This paper works on developing a CAD to analysis and radiologists in multiclass brain tumor detection and classification using MRI. Basically, CAD tool used for brain tumor classification with using the MRI image type data. These data types are T1, T2. Post contrasts T1. FLAIR provides the brain tumor information about image texture and intensity. Many studies [15], [4], [6], [19] have been used for analysis of classification of normal brain and abnormity. Very few state of the art studies worked on a multiclass classification [13-14], [5].
The aim of this paper is to work in the tumor region detection and classification. In this work, we present the segmented region. These regions are segmented using automatic localization, and level set based energy minimization technique [12]. In this paper, the author develops automatic segment the brain tumor MRI region [12]. After segmentation, extract the features of these regions then apply the classification module. We develop a method for brain tumor multiclass classification where the Hellinger distance decision -based method was developed for imbalance data set [1], [15], [17]. The method has never been used for multicasts medical imaging classification. In this paper, we proposed multiclass kernel based Hellinger Decision-Tree skewed splitting tree and skewed splitting forest Hellinger Decision - Forest for multiclass tumor classification for CAD system.

1.1. The framework of proposed CAD system for the characterization of the brain tumor

The proposed system developed for the usefulness of radiologists in classifying brain tumors in MR images is shown in Figure1. The system consists of three modules (i) After segmented tumor region using These regions are segmented using automatic localization, and level set based energy minimization technique [12] feature properties apply on this segmented tumor region feature (ii) secondly, to extract the feature properties of these tumors. Features such as centroid, shape, areas are calculated from the segmented tumor. And documented patient symptoms are taken from radiology department. The features are categorized in two modules such as patient tumor class properties and the patient symptoms also. The selected features are used as input attributes for the patient dataset (iii) classification module using proposed multiclass kernel-based HD-Tree and HD-Forest.

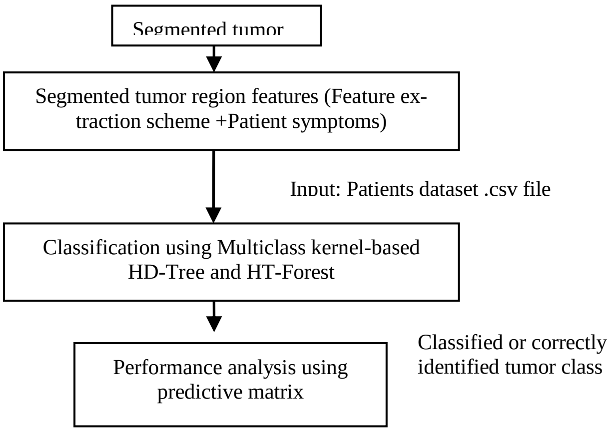

Fig. 1: Proposed Workflow for the Classification of Brain Tumor.

The paper is organized as follows: In Section II feature extraction, patient feature and ROI, features are used. For classification, we proposed Multiclass Kernel-based HD-Tree and the HD-Foreset classifier are used in Section III. Data set is illustrated in Section IV. In Section V experimental results are given. Discussions, including comparative analysis with the earlier method are given and Section VI respectively. The paper is concluded in Section VII.

2. Feature extraction

Proposed work extracts 2D features from the tumor to evaluate their effectiveness in differentiating between benign and malignant tumor and further types of tumor. 30 different features and symptoms are used in this work for tumor image analysis. We use total 97 ROI for feature characterization. The MRI parameter and region features extracted from the tumors shown in table 1 and table 2.

2.1. MRI parameters

Table 1: MRI Brain Tumor Type Image Features Category.

Brain tumor type Image features category	Description
Image type	primary T2-w(bright CSF), T1-w(dark CSF)
Photometric interpretation	Momochrome2 (Higher pixel values are displayed as white)
T1-w	Gray matter (dark gray) and white matter (lighter gray)
T2-w	Gray matter (lighter gray) and white matter (darker gray
Gray matter	810 (T1 ms) 100(T2 ms) multiple sclerosis
White matter	680 (T1 ms) 90 (T2 ms)
Internal echo	Homogeneity , heterogeneities

Table 2: MRI Feature Parameters Provided by MRI Division Lab (SMS Hospital)

Features category of various type of tumor	Description
Age	Min 1 and max 90
Gender	Male and female
Area	The number of pixels in the region- max 0.99
Pixel spacing	Between 0.1 to 1 mm in row n column (between the center of each pixel.).
Slice thickness	5.0, 6.0, 0.8mm
Flip angle	90 degree
Structure	Solid
shape	Irregular, circularity
Centroid	The center of mass of the tumor
Pixel bandwidth	100 to 500 max

2.2. Patient tumor feature

In this part, our proposed work working on two phases first is patient symptoms and second is automatic region properties.

2.2.1. Patient symptoms

After tumor detection, the tumor feature will find without the help of region properties. The patient symptoms are very important to characterize or to identify the features of the patient tumor class or what type of diseases can happen. Furthermore, the feature extraction scheme is an analysis the proper features of the tumor where a feature like a shape, area, location, etc. the fig shows the prediction rate 0 or 1 rang to analysis the symptoms of the detected tumors.

Symptoms are used to characterize the all patient disease-related symptoms like a headache (H), vomiting (V_N), seizures (S), Vision Problem (V_P), Memory Difficulties (M_D), Speech Difficulties (S_D), Personality Alterations (P_A), Hemiparesis (Hm), Drowsiness (D), Fatigue (F), Peripheral neuropathy (P_N), Hearing difficulties (H_D), Dizziness (Dz), Difficulty Swallowing (D_), S); Facial weakness (F_W), Trouble thinking (TT), Memory Loss (M_L), Speech problems (S_P) were used for tumor class identification in classification module.

2.2.2. Region properties

Regional properties are like patient tumor class properties are used to characterize the tumor features like gender, age shape, area, image intensity, centroid, solidity, entropy, orientation, perimeter, eccentricity, convex area, structure, internal echo, etc.

3. Classification

The multi-class classification problem is the main problem where data set consists of multiple classes. Only two-class classification is called absence or presents the possible results. According to this common condition, the multiple class classifications is very difficult to find out the multicasts classification of any medically related disease. Many researchers have developed two-class classification still multicasts problem is challenging for the classification of multiple classes medical imaging related problems.

3.1. Decision tree

Decision trees are one of the fundamental learning algorithms. The most popular of decision tree learning algorithm is Hellinger distance decision trees (HDDTs) [1] was developed for binary imbalance dataset. In this paper, we proposed multiclass kernel based HD-tree and HD-forest for the classification of MRI tumor class.

3.1.1. Proposed multiclass kernel based HD-Tree

The Hellinger distance splitting method is the best way to reduce the multiple class nodes into all binary class possibilities [17]. Hellinger distance can try to find out the best splitting classes between all possible conditions of positive and negative class [1, 14]] and thus split can play the main role of the finding the best multi-class problem. The proposed kernel based Hellinger HD-T algorithm [A]

- Let, training setT, feature$f_{r,s}$, tumor region properties t_r(n cell) and symptoms s_t(m cells) , set of classes t_c.
- For each pair of subsets of t_c (6 classes): $t_{c1} \subset t_c, t_{c2} = t_c \backslash t_{c1}, t_{c3} = t_{c1} \backslash t_{c2} \dots\dots.. t_{c6} = t_{c4} \backslash t_{c5}\ or$

classes in t_{c1} may be targeted disease class (positive) or classes in t_{c2} may be the different targeted class, when all classes t_{c2} as the targeted disease class and t_{c3} as the different targeted class (negative)class where t_{c1} as the negative and so on.

- For each value (rang) assume $\upsilon \in V_{fr,s}$ and Let $\varsigma = V_{fr,s}\ \upsilon$
- $\mathcal{M}_d$ $=\sum_{c=i} t_c\ (\sqrt{\frac{|\mathrm{T}_{fr,s,\upsilon,c+maximum\ value\ of\ personal\ feature\ (age)}|}{\mathrm{T}_+}} - \sqrt{\frac{|\mathrm{T}_{fr,s,\upsilon,c+minimum\ value\ of\ personal\ feature\ (age)}|}{\mathrm{T}_-}})^2 + (\sqrt{\frac{|\mathrm{T}_{fr,s,\upsilon,c+maximum\ value\ of\ personal\ feature\ (solidity)}|}{\mathrm{T}_+}} - \sqrt{\frac{|\mathrm{T}_{fr,s,\upsilon,c+minimum\ value\ of\ personal\ feature\ (solidity)}|}{\mathrm{T}_-}})^2$ (1)
- $|\mathrm{T}_{fr,s,\upsilon,c+}|$= the set of all positive attribute values that lie within frequency and personal features or say same class instance and $|\mathrm{T}_{fr,s,\upsilon,c-}|$= the set of all negative attribute values that do not lie within frequency and tumor class features or say same class instance.
- i=1..6
- If$\mathcal{M}_d$ > hellinger then Hellinger = $\mathcal{M}_d$
- End
- Return $\sqrt{\mathcal{M}_d(cur_\upsilon\_value t_c)}$

The following steps show the proposed MC-HD-T and HD-F process to distinguishing the various tumor class through the bin (A distance of partitioning the various class (max to min measurement of features) of tumor according to age, entropy etc.). This process works on the frequency and patient tumor class features. For the identification no. of the instance with tumor class presence and absence or vice a verse. The value or a range of tumor class will be defined as a value of frequency and symptoms with the condition of probability density function [B].

- Let bin B default size= n (bins passed $\mathcal{B}$)
- Put $\mathrm{T}_{fr,s,\upsilon,c+,patient\ personal\ feature}$ and $\mathrm{T}_{fr,s,\upsilon,c,calculated\ tumor\ tc\ features}$ feature values
- Estimated bin $\mathcal{B}$ size = ($|\mathrm{T}_{fr,s,\upsilon,cpatient\ minpersonal\ or\ calcylated\ feature}|$ + $|\mathrm{T}_{fr,s,\upsilon,calculated\ maxpersonal\ or\ calculated\ features}|$)/ $\mathcal{B}$ B=$\sqrt{no. of\ all\ features\ or\ values\ or\ no\ of\ cells}$
- Let T_+=number of instance with class label if = 1 (no. of instance with tumor class t_c are classified) Or Let T_-=number of instance with class label if = 0 (no. of instance with tumor class t_c are not classified)
- Create $\mathcal{B}_n$ = [$\mathcal{B}\{-1,1\}$],
- For each attribute $t_{class,frequancy,value}$ ($|\mathrm{T}_{region,symptoms,}value,class-|$+ $\mathcal{B}$)
- $t_{c,f,\upsilon}$ ($|\mathrm{T}_{fr,s,\upsilon,c-}|$+ $\mathcal{B}$,{($|\mathrm{T}_{fr,s,\upsilon,c-}|$+ $\mathcal{B}$)+B},………
- bins points expressed as a P($f_{,r,c,r}$, $f_{,r,c,c}$)
- ($f_{,r,c,r}$, $f_{,r,c,c}$)= $t_{c,f,\upsilon}$=[r 1 c]
- Xbins= bins ranges lie on a row and column
- *Lets* value (rang) be like $V_{fr,s} = f_{,r,c,0}$, $f_{,r,c,1}$, $f_{,s,c,2}$.
- ($f_{,r,c,0}$, $f_{,r,c,1}$)=$f_{,r,c,1}$, ($f_{,r,c,1}$, $f_{,s,c,2}$) = $f_{,s,c,2}$
- If pdf of t_cx= $f_{,r,c,0} < f_{,r,c,1}$not more than normal distribution range
- t_cy> $f_{,r,c,1}$ & less < $f_{,s,c,2}$

The condition shows the patient class with frequency and feature selection. For example, age 1 between 90 years (rang 1) that will be using for separation of between class with entropy (rang 2) and special case of memory loss (rang 3)

3.1.2. Workflow of proposed multiclass HDT-F

- Select 97 tumor regions
- Training class (set of 6 classes) –measure the separability of tumor classes in classification
- Tc (tumor class) positive subset of all and another negative class- identify objects of a specific class amongst all tumor [A]
- Apply Bhattacharya coefficient for multiple classes see in (1)
- Attribute selected from dataset 30*cell feature and symptoms ex-age cell(85 max-1min=84)
- Find feature size n*m cells
- Calculate the per-feature selected cell column values with respect to tumor classes [B] $\sqrt{bins}$=$\sqrt{84} = 9.1$
- The Distance of both partitions 1+9.1= 10.1, 10.1+9.1=19.2.. … 84.3
- Using [B] apply decision discretization splitting tree method: First to make a histogram of the data sample values or width samples then bins can work on above points may be expressed as a P(X, Y) lets X-axis (1,2) (2,3) (3,4) (4,5)… then Y axis counts the (1, 2) value and so on.
- The maximum no. of observation for learning 97*30 cell
- Final tumor class identification of multiple Tc*cell
- Performance evolution matrix using no. of observation matrix

The maximum no of the instance that's called learner observation symptom attribute calculated through algorithm [B]. After this, discrete the node branch through splitting criteria. If the symptoms not found in the disease class then further apply splitting criteria. If class found the disease then calculates the prediction of no of disease class.

3.1.3. Multiclass kernel based Splitting HD-tree and HD-forest classification parameters and attributes selection

F=Predicted classes = predict tree (model, attributes)
Description: Predict labels using trained Distance Decision Tree
Parameters: model_: a trained Distance Decision Tree model_
Attributes: I X F numeric matrix where I is the number of instances and F is the number of attributes. Each row represents one tr instance and each column represents the value of one of its corresponding attributes
Output: Predicted classes: I X 1 matrix where each row represents a predicted label of the corresponding attribute set. To classify the brain tumor class using HD-tree form describes the tumor class characterization in tree form where function F shows the all predicted classes with model and attributes (97*30 cell features and symptoms). The leaf node of the tree is based on (IX F) numeric matrix where I is the number of instances (tumor classes) and F is the number of attributes (30 *cell features and symptoms). Each row represents one tr instance and each column represents the value of one of its corresponding attributes (30*cell).the number of instances (tumor classes) and a number of features are based on the total size of the attributes (symptoms and features) if

no of the instance is 0 the result shows none. HD-Forest algorithm reduces the (if and else) conditions. The prediction of tumor class follows the no of trees where no of the tree is 1:1 then prediction of attribute follows the tree nodes The forest has tree node and child nodes with less no of (if and else) condition.

We proposed multiclass kernel based HD-T and HD-Forest. In HD splitting forest is basically used to reduce the (if, else) conditions in the decision tree because of the top to down connected components are calculated in this classification that's why the algorithm gives the best accuracy and reduces the classification time shown in splitting HD-Tree and HD-Forest classification scheme. In addition, less percent of pruning is occurring during classification. There is no over fitting occur. It is work like a number of decision trees or no. of node's conditions in order to improve the classification rate in splitting forest form. Decision splitting tree process works as a shape of the connected leaf node to multiple nodes. The method can deal with the multicasts data set for classification of brain tumors type.

4. Dataset

In this paper, we work on 97 patient MRI brain tumors at MRI division lab, Sawai Mansingh SMS hospital Jaipur, India and Kamaljeet MRI division lab, Punjab, India, total November 2015 to December 2016 time taken to collect these patient data. These images include- (Astrocytoma, Ganglioglioma, Glioblastoma, Epidermoid, Mixed Glioma and Meningnet) (5type of tumors) were used for the experiment.

We can take an example of meaning tumor 35-year-old men patient (kesher) region segmented using automated localization method [12] shown in Fig 2 and region properties shown in Table 3

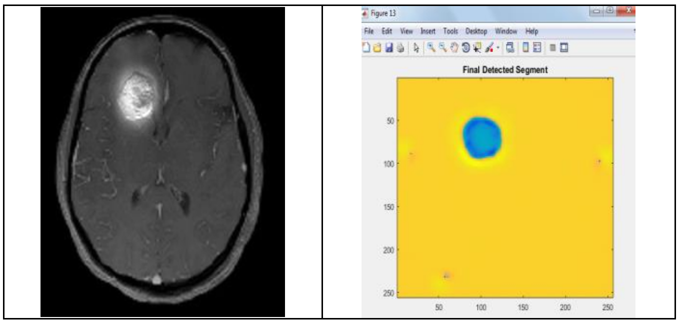

Fig. 2: Segmented Tumor Region.

Table 3: Region Properties of 35-Year-Old Patient

Region properties	Pixel range
Centroid	[17.00, 89.00]
Area	2098
Convex area	2368
Orientation	76.5
Perimeter	190.4
Solidity	0.885
Entropy	0.204
Eccentricity	0.385

Table 4: Trained Classifier with Tumor Parameters through multiclass Kernel Based HD-T and HD-F (Patient Dataset)

Parameters	Trained attributes and class
Y (Response)	97*1 cell
X(observation for learner)	97*30 cell
W (no. of observation)	97*1
Predictor name	1*30 cell (attributes)
Class	6*1 cell
Prior	[0.2165,0.0309,0.0619,0.2268,0.4124]
Trained weight	30*1

5. Experimental results

The experimental result shows the best classification results with the help of our proposed multiclass HD-Tree and Forest classification method. In this paper, we have to compare our proposed work with currently available classification method that is called Support Vector Machine.

Table 5: Total 97 MRI Tumor Images are used as Classifiers.

Total number of images	97
Astrocytoma (A)	21
Ganglioglioma (GA)	6
Glioblastoma (GL)	22
Epidermoid (EP)	3
Mixed glioma (MG)	5
Meningnet (M)	40

Kernel HD-T and F Predicted tree and forest expanded predictor = 1*30 cell attribute, model parameter 1* ensemble parameter gender, age and symptoms etc. (patient database) shown in Table 4.

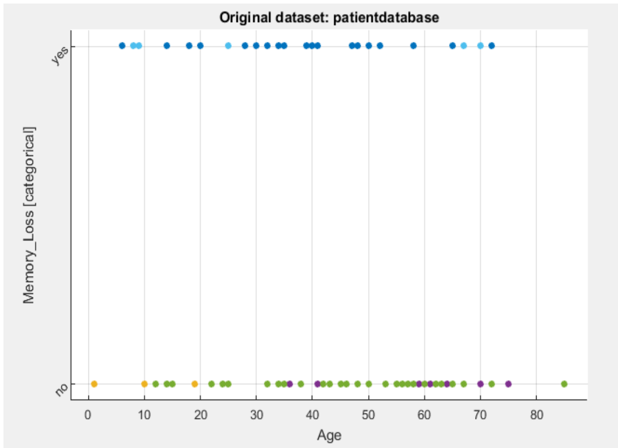

Fig. 3: Analysis of Predictive Values for A Particular Class Using A Decision Tree. where Astrocytoma (Blue Dots) and Migexd Glioma (Light Blue Dots).

Predictors X= Age, Y= Memory loss, Prediction speed: ~940 obs/sec, Training Time: 2.1602 sec shown in fig 3. In this graph, the predictive features of tumor class are shown.

5.1. Comparative analysis

Total training and testing time per tumor classes during classification of MRI tumors through proposed Multiclass kernel-based HD-Tree and HD-Forest algorithm and LA-SVM [6] shown in fig 4.

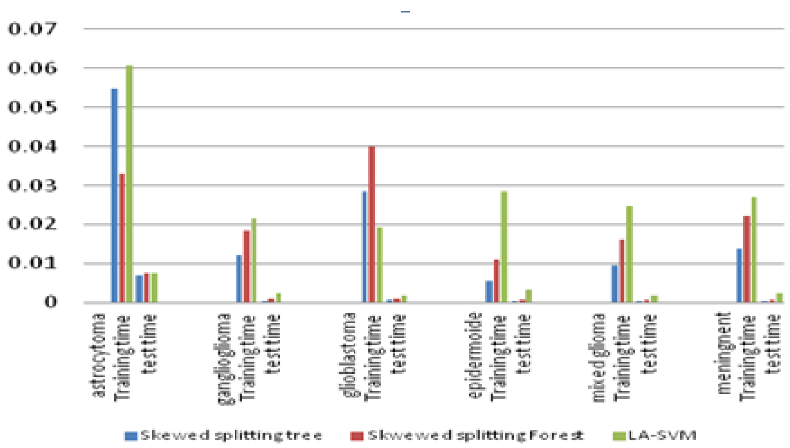

Fig. 4: Training and Test Time per Class Classification Result.

Total running time per tumor classes are taken by proposed multiclass kernel-based HD-Tree, HD-Forest and LA-SVM [6] shown in fig 5.

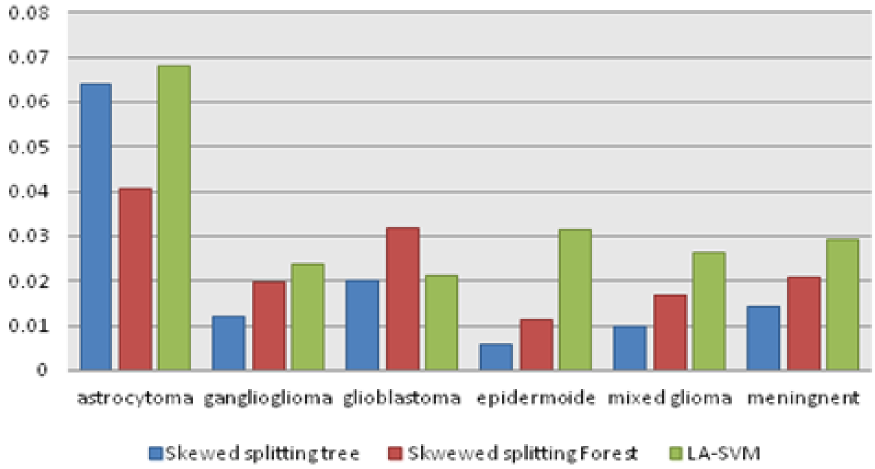

Fig. 5: Total Time Is Taken by Per Class Classification Result.

Table 6: Shows the Per-Class Classification Accuracy Result

Tumor class	Kernel-based HD accuracy %	LA-SVM [6] Accuracy %
Astrocytoma	69%	65%
Ganglioglioma	99.7%	98.7%
Glioblastoma	60%	56%
Epidermoid	95%	90%
Mixed glioma	100%	84%
Meningnet	50%	50%

In this work, we have to work on total 97 images for the classification shown in Table 5.The separate class accuracies for multiclass kernel-based Hellinger method are Astrocytoma (A) 69%, Ganglioglioma (GA) 99.7%, glioblastoma (GL) 60%, epidermoid (EP) 95%, mixed glioma (MG) 100% and meningnet (M) 50% shown in Table 6 also, total time taken by per class classifiers shown in fig 5. Accuracy for HD-T (96.5%) and HD-F (99.9%) shown in fig 6. and Table 6 was found. The total time taken by HD-T (0.004294) and HD-F (6.69) was found. The HD-T take less time just because of more (if, else) are used. This may be creating less purring. HD-F based Skewed splitting forest is basically used to reduce the (if, else) conditions in decision tree because of the top to down connected components are calculated in this classification that's why the 99.9% of accuracy results was found for the classification of brain tumor MRI image and reduce the classification time. Decision splitting tree process works as a shape of the connected leaf node to multiple nodes. The splitting tree results showed less accuracy as compare to splitting forest because of the multiple if and else conditions take more running time for brain tumor class classification. When we compare to LA-SVM based method the separate class accuracies for LA-SVM are Astrocytoma (A) 60%, Ganglioglioma (GA) 98.7%, glioblastoma (GL) 56%, epidermoid (EP) 90%, mixed glioma (MG) 84% and meningnet (M) 50% found shown in Table 6. This method gives less accuracy for per-class classification rate as compared to our proposed Hellinger's method. LA-SVM methodology [6] has an overfitted and pruning problem which is not useful. Classification accuracy of LS-SVM was found 96%shown in fig 6.

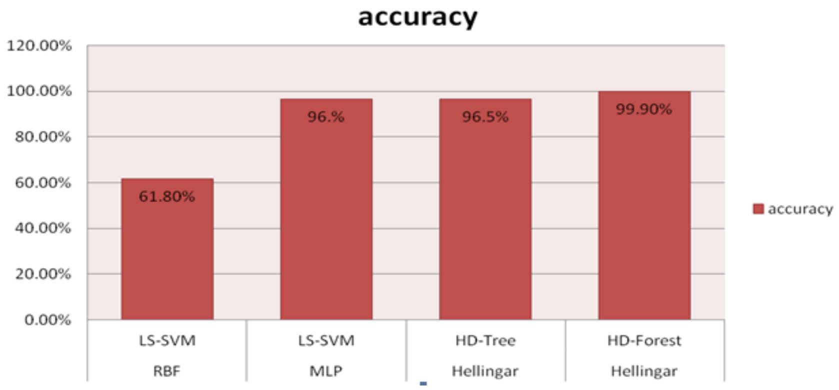

Fig. 6: Accuracy of Our Proposed Hellinger Based HD-T, HD-F, and LA-SVM Based RBF, MLP [6].

Table 7: Kernel Method Based Proposed Classifiers and LA-SVM [6] Comparison with Respect to PPR and FNR Value Percentage

Kernel methods	PPR values	FNR values
HD-Tree	93%	4%
HD-Forest	97%	0%
SVM-RBF	61%	36%
SVM-MLP	93%	4%

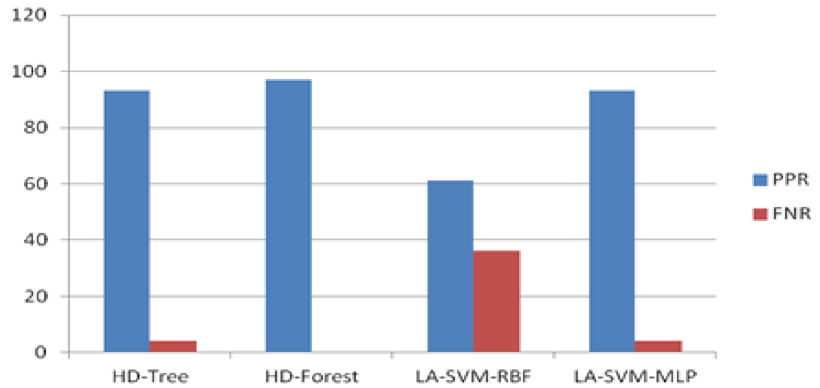

Fig. 7: Shows the Positive Predicted Rate (PPR) and False Negative Rate (FNR) for Our Proposed Kernel Based HD-Tree and HD-Forest Based Classifiers and LA-SVM Based Classifiers [6].

Table 8: Comparative Analysis of proposed multiclass HD-Tree and HD-Forest and LA-SVM [6] True Positive Value for All Classes

True positive value	HD-Tree	HD-Forest	LA-SVM-RBF	LA-SVM-MLP
Astrocytoma	21	21	17	21
Epidermoid	0	3	0	0
Ganglioglioma	5	6	4	5
Glioblastoma	22	22	12	22
Meningnet	40	40	25	40
Mixed glioma	5	5	3	5

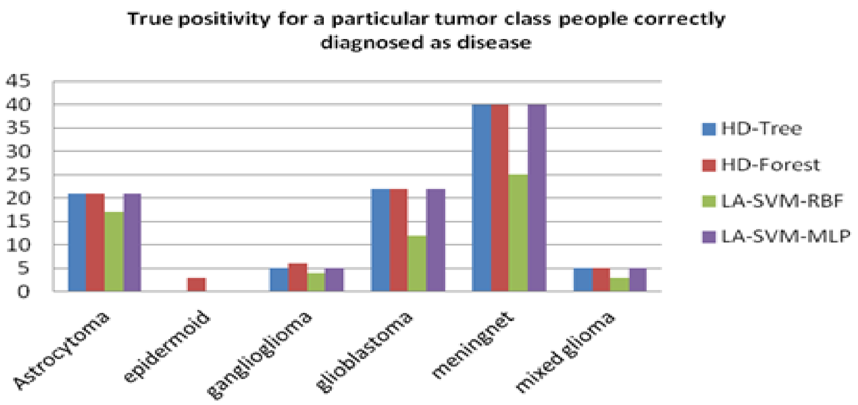

Fig. 8: Comparative Analysis of True Positive Value for All Tumor Classes.

The positive rate of proposed multiclass kernel-based HD-Tree and 93% to be found and negative rate 4% to be found also, 97% to be found in HD-Forest method. SVM 90% to be found the positive rate and 7% to be found negative rate we can see that the proposed multiclass kernel-based HD method shown in table 7 can positively found classes with higher percentage as compared to the LA-SVM [6] shown in fig 7. Also, the comparison analysis of true positive values for all tumor classes are shown in table 8 and fig 8 The ROC curve of various classifiers with respect to True positive rate and false positive rate shown in fig 9.

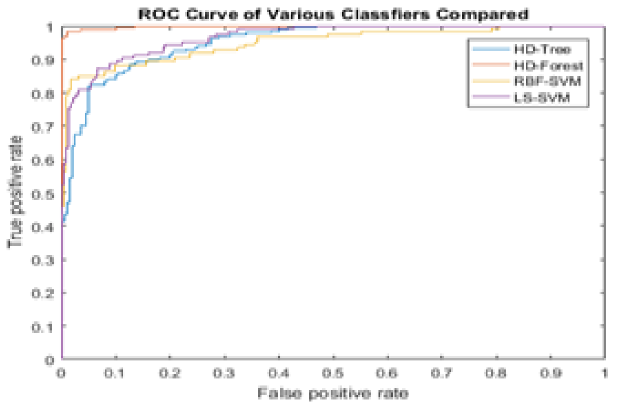

Fig. 9: ROC Curve for Our Proposed multiclass HD-Tree, HD-Forest and LA-SVM [6] Classifiers.

Table 8: Comparison Analysis of Proposed multiclass Hellinger Based HD-T, HD-F, and LA-SVM Based [6] Classifier

Kernel Hellinger	Specificity	Accuracy	AUC	SE
HD-T	99%	96.50%	0.9618	0.0153
HD-F	100%	99.9%	0.9985	0.0087
SVM-RBF	40%	61.8%	0.936	0.0233
SVM-MLP	81.33%	96%	0.9467	0.0210

Table 9: Comparison between Our Proposed Method and LA-SVM [6] Classifier

Classification result points	Proposed classifier	Previous LA-SVM classifier
Pruning effect	less	Pruning problem
Overfitting effect	No	Overfitting problem
Database support	Easily Classify Large dataset	Results affected during classification
Class classification	Support multiclass classification	Support bi-class classification
Classification time	Take less time	Take more time
Type of dataset support	Multiclass imbalance	Binary class

6. Conclusion

In this paper, we proposed a multiclass kernel based Hellinger decision tree and forest were used for multicasts classification of brain tumor classes. In this paper, we work on both conversancy time and accuracy for all tumor classes. There are total 97 tumor regions were segmented. The experimental result shows the individual class accuracy. The individual class accuracies are Astrocytoma (69%), ganglioglioma (99.7%), glioblastoma (60%), epidermoid (95%), mixed glioma (100%) and meningnet (50%) shown in Table 6. In this study, we proposed a multiclass kernel based Hellinger HD-T and HD-F for multicasts classification. Hellinger based method can deal with the over fitting and pruning problems. HD-F can reduce the (if, else) conditions in the decision tree because of the top to down connected components are calculated in this classification that's why the 99.9% of accuracy results was found under the classification of brain tumor MRI image and reduce the classification time. Nevertheless, the LA-SVM has a problem of over fitting and pruning that is the main reason for less accuracy and more conversancy time. The comparative analysis of both methods shown in Table 9 that our proposed Hellinger based method gives more accurate results and gives less error for a particular class as compared to the previous LA-SVM method [6] shown in Table 8.

References

[1] David A, Cieslak, Ryan Hoens T, Nitesh V, Chawla W (2012), Philip Kegelmeyer. Hellinger distance decision trees are robust and skew-insensitive. *Springer Knowl Disc*, 24:136–158 https://doi.org/10.1007/s10618-011-0222-1.

[2] Liew AWC and Yan H (2006), Current methods in the automatic tissue segmentation of 3D magnetic resonance brain images. *Current Med Imaging Rev*, 91–103, pp.2 (1).

[3] Pham DL, Xu C, and Prince JL (2002), Current methods in medical image segmentation. *Annu Rev Biomed Eng* 2:315–337 https://doi.org/10.1146/annurev.bioeng.2.1.315.

[4] Dahshan El. A. El, Hosny T and Salem A M (2010), Hybrid Intelligent Techniques for MRI Brain Image Classification. *Elsevier Digital Signal Processing*, vol.20, pp. 433–441. https://doi.org/10.1016/j.dsp.2009.07.002.

[5] Zacharaki El, Wang S, Chawla S, Yoo D S, Wolf R, Melhem E R, and Davatzikos C (2009), Classification of Brain Tumor Type and Grade using MRI Texture in a Machine Learning Technique.' *Magn. Reson. Med*, vol.62, pp.1609-1618. https://doi.org/10.1002/mrm.22147.

[6] Praveen G B and Agarwal A (2015), Hybrid Approach for Brain Tumor Detection and Classification in Magnetic Resonance Images. *International Conference on Communication, Control and Intelligent Systems* (CCIS), 2015, 978-1-4673-7541-2. https://doi.org/10.1109/CCIntelS.2015.7437900.

[7] Kong J, Wang J, Yinghu Li Y, Zhang Lu Z and Zhang B (2006), A novel approach for segmentation of MRI brain image. *IEEE Melecon*, Benalmádena (Málaga), Spain.

[8] Marroquin J L, Vemuri B C, Botello S, Calderon F, and Fernandez-Bouzas A (2002), An accurate and efficient bayesian method for automatic segmentation of Brain MRI. *IEEE Transactions on medical imaging*, pp. 21(8).

[9] Clarke L P, Velthuizen R P, Camacho M A, Heine J J, Vaidyanathan M L, Hall O, Thatcher R W and Silbiger M L (1995), MRI segmentation: methods and applications. *Magnetic Resonance Imaging*, 343-368, pp.13 (3).

[10] Singh N and Jindal A (2012), A Survey of different types of Characterization Technique in Ultra- sonograms of the Thyroid Nodules. *International journal of computer science and informatics*, 2231-5292, pp. 1(4).

[11] Singh N and Choudhary N (2014), A Review of Brain Tumor Segmentation and Detection Techniques through MRI. *International Journal of Computer Applications*, 0975 – 8887, pp. 103(7).

[12] Singh N and Choudhary N (2017), Automatic localization and level set based energy minimization for MRI brain tumor. *IEEE International Conference on Computer, Communications and Electronics (Comptelix)*, 978-1-5090-4708-6. https://doi.org/10.1109/COMPTELIX.2017.8003951.

[13] Georgiardis P, Cavouras D, Kalatzis I, Daskalakis A, Kagadi D C, Malamas M, Nikifordis G, Solomou E (2007), Non-linear least square feature transformations for improving the performance of probabilistic neural networks in classifying human brain tumors on MRI. *Lecture Notes on Computer Science*, pp. 239-47. https://doi.org/10.1007/978-3-540-74484-9_21.

[14] Georgiardis P, Cavouras D, Kalatzis I, Daskalakis A, Kagadi D C, Malamas M, Nikifordis G, Solomou E (2008), Improving brain tumor characterization on MRI by probabilistic neural networks on non-linear transformation of textural features. *Comput Meth Prog Bio*., vol. 89, pp.24-32. https://doi.org/10.1016/j.cmpb.2007.10.007.

[15] Lyon R J, Brooke J M, Knowles J D and Stoppers B W (2014), Hellinger Distance Trees for Imbalanced Streams. *22nd International Conference on Pattern Recognition. https://doi.org/10.1109/ICPR.2014.344*.

[16] Chaplot S, Patnaik L M and Jagannathan N R (2006), Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomedical Signal Processing and Control*, vol. 1, pp. 86–92. https://doi.org/10.1016/j.bspc.2006.05.002.

[17] Hoens T R, Qian Qi, Nitesh Chawla V, and Zhi-Hua Zhou (2012), Building Decision Trees for the Multi-class Imbalance Problem. *Springer-Verlag Berlin Heidelberg*, pp.122-134 https://doi.org/10.1007/978-3-642-30217-6_11.

[18] Song Y, Zhang C, Lee J, Wang F, Xiang Z, and Zhang D (2009), Semi-supervised discriminative classification with application to tumorous tissues segmentation of MR brain images. *Pattern Anal Application*, 99–115, pp. 12. https://doi.org/10.1007/s10044-008-0104-3.

[19] Zhang Y, Dong Z, Wu L, Wang S (2011), A Hybrid Method for MRI Brain Image Classification. *Expert system with applications*, vol.38, pp.10049-10053. https://doi.org/10.1016/j.eswa.2011.02.012.