

## ARTICLE

## BRAIN TUMOR DETECTION AND CLASSIFICATION MODEL USING OPTIMAL KAPUR'S THRESHOLDING BASED SEGMENTATION WITH DEEP NEURAL NETWORKS

Bakthavachalam Devanathan<sup>\*</sup>, Kaliyappan Venkatachalapathy

Department of Computer and Information Science, Annamalai University, Chidambaram, TN, INDIA



## ABSTRACT

Brain Tumor (BT) detection and classification is a crucial process for radiologists and the detection of BT at the earlier stage can lead to high survival rate. Therefore, an automated BT detection and classification model is needed to detect the tumor regions at a faster rate. This paper presents a novel brain tumor detection and classification model using an Optimal Kapur's Thresholding based Segmentation with Deep Neural Network, abbreviated as OS-DNN model. The proposed OS-DNN model operates on five major stages, namely preprocessing, segmentation, feature extraction, feature reduction, and classification. Initially, preprocessing of MRI brain images takes place in three ways namely bilateral filtering (BF) based noise removal, contrast limited adaptive histogram equalization (CLAHE) based contrast enhancement and skull stripping. Then, optimal kapur's thresholding based segmentation process takes place and grey wolf optimization (GWO) algorithm is used to optimize the threshold value. Next to that, Discrete Wavelet Transforms (DWT) is applied as a feature extraction technique and kernel principle component analysis (KPCA) is employed as a feature reduction. Finally, DNN model is applied to classify the MRI brain images into normal and abnormal MRI images. The experimentation of the OS-DNN model is carried out using a benchmark Kaggle dataset and the results are examined under different aspects. The simulation outcome ensured that the OS-DNN model is superior to other models with the maximum sensitivity of 97.94%, specificity of 98.08% and accuracy of 98.02%.

## INTRODUCTION

## KEY WORDS

Brain Tumor,  
Classification, Deep  
learning, Feature  
Extraction,  
Segmentation

Unlike other tumors, detection of brain tumor using invasive method is very difficult. Therefore, image based diagnosis and classification using Magnetic resonance imaging (MRI) scans is generally practiced. However, computer-aided methods for improved diagnosis and classification of MRI scans are of demand that include better accuracy in tumor detection, segmentation, and classification [1]. Image processing technologies are frequently applied for the purpose of tumor prediction. The key objective of image segmentation is to divide a picture into homogeneous blocks and identifies the structure of every region. Magnetic resonance imaging (MRI) and Computed tomography (CT) scans were employed for examining the presence of BT. MRI scan is highly productive when compared with CT scan and it is not constrained with radiation power [2]. Tumor might be developed with various tissues, single MRI is inadequate to provide to complete information of abnormal tissues. By integrating unique complementary data could extend tumor region which are segmented. A feature of MRI has been applied for segmentation with weighted 3 pictures for all slice's axial. The segmentation principles are highly beneficial mostly at the time of measuring affected tissues. Segmentation is measured as vital procedure while computing MRI which enhances the model's performance. Segmentation contains image partitions that have similar parameters like color, brightness, texture and intensity [3]. Human segmentation of images is assumed to be tedious and time-utilizing operation. Hence, automated segmentation models are needed to develop for effective diagnosis of brain tumor.

The combination of Artificial Bee Colony (ABC) and Genetic Algorithm (GA) methodologies finds useful to predict the threshold values. Additionally, statistical features are utilized with support vector machine (SVM) approach as a classifier [4]. El-Dahshan and colleagues have developed feedback pulse coupling network (FPCNN) for image segmentation of BT diagnosis. The DWT, PCA and artificial neural network (ANN) were employed for feature extraction, FS and classification, correspondingly [5]. Similarly, a novel kernel clustering method for MR images segmentation was also developed [6]. These methods have replaced the previous isotropic Gaussian kernel using anisotropic kernel equated by Mahalanobis distance. As fuzzy C-means (FCM) segmentation is applied for detecting affected area in brain MRI. Statistical features as well as SVM are mainly used to process the feature extraction and classification.

Ali et al (2014) defined a new MR image segmentation that depends upon the morphological pyramid with FCM clustering. Initially, wavelet multi-resolution is applied for retaining the spatial context among pixels. Morphological pyramid is utilized to combine final multi-resolution images with actual image which tends to improve the sharpness and reduce the noise in computed image. Consequently, FCM is used to divide the processed images [7]. Ahmadvand and Kabiri implied a novel segmentation approach according to the Markov random field as well as feature vector for the combination of spatial and spectral data for MRI image segmentation [8]. In 2017, Cabria and Gondra provided a segmentation technique named as Potential Field Segmentation (PFS) [9]. It is evident that, the developed method is capable of classifying MR images effectively. Also, the MRI segmentation techniques are viewed as successful image segmentation schemes which are applied in this literature.

\*Corresponding Author

Email:  
devacisau@gmail.com

Received: 2 June 2020  
Accepted: 10 July 2020  
Published: 14 July 2020

Cigaroudy and Aghazadeh applied a new multiphase segmentation technique which depends on the combination of intensity function, eigenvector of Hessian matrix and Curvelet [10]. Similarly, Akbarizadeh developed a new model named as Synthetic Aperture Radar (SAR) for image segmentation which depends upon the simple segmentation processes. The presented method compounds the intensity data and input image texture with the help of Cellular Learning Automata (CLA) for segmenting SAR images [11]. Rahmani and Akbarizadeh developed an unsupervised feature learning approach according to the unification of spectral clustering and sparse coding SAR image segmentation [12]. Spectral clustering is defined as an image segmentation model which makes feasible combination of features and cues. Sparse coding is referred as an unsupervised learning that identifies the patterns of high-level data semantics. Furthermore, it provides alternate collaboration of texture and color features for the segmentation of polarimetric SAR (PoSAR) image. Though various segmentation techniques are available, still there is a need to develop automated segmentation and classification models for BT diagnosis for achieving maximum classification performance.

This paper introduces novel BT detection and classification model using an Optimal Kapur's Thresholding based Segmentation and DNN model, abbreviated as OS-DNN model. The proposed OS-DNN model initially undergoes preprocessing to improve the classification performance by noise removal, contrast enhancement, and skull stripping. Then, optimal Kapur's thresholding based segmentation process takes place and grey wolf optimization (GWO) algorithm is used to optimize the threshold value. Then, discrete wavelet transform (DWT) based feature extraction and kernel principal component analysis (KPCA) based feature reduction process takes place. At last, deep neural network (DNN) model is applied to classify the MRI brain images into normal and abnormal MRI images. The experimentation of the OS-DNN model is carried out using a benchmark Kaggle dataset and the results are examined under different aspects.

### METHODS

The overall working procedure involved in OS-DNN method is illustrated in Fig. 1. As depicted, the projected OS-DNN model undergoes pre-processing, GWO-KT based segmentation, DWT relied feature extraction, KPCA based feature reduction, and classification using DNN model. These processes are explained in the following sections.

**Preprocessing:** Initially, the input image undergoes preprocessing in three stages namely bilateral filtering (BF) based noise removal, contrast limited adaptive histogram equalization (CLAHE) based contrast enhancement, and skull stripping.

**GWO-KT based Segmentation:** In early decades, Kapur's image segmentation was mainly applied for segmenting gray scale images with the help of histogram's entropy. It identifies an optimal  $T$  value that enhances the entire entropy [13]. Suppose  $T = [t_1, t_2, \dots, t_{k-1}]$  be a vector of image threshold. Next, Kapur's entropy would be represented as

$$J_{\max} = f_{\text{kapur}}(T) = \sum_{i=1}^k H_j^c \text{ for } C\{1, 2, 3\} \quad (1)$$

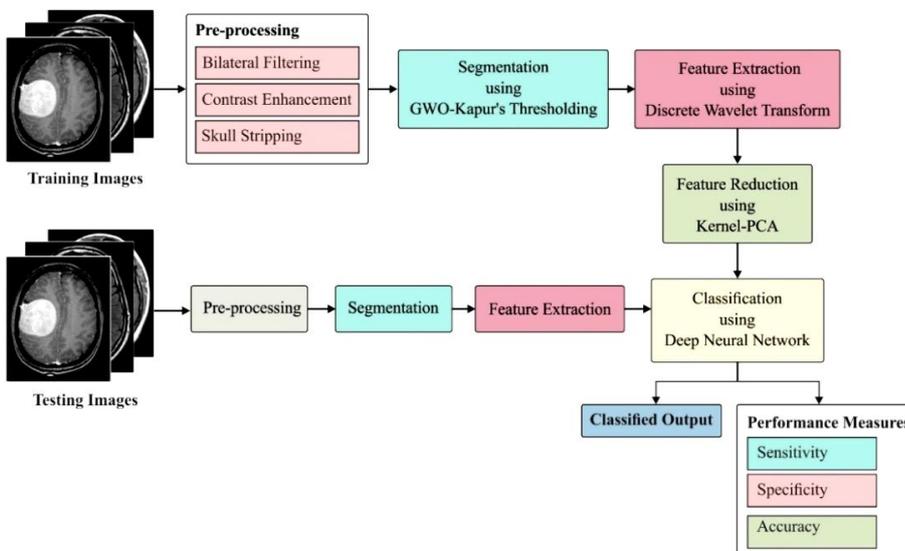


Fig. 1: Working Principle of OS-DNN Model

Basically, entropy is determined in an independent manner according to the  $t$  value. The multi-level thresholding issue can be demonstrated as,

$$\begin{aligned}
 H_1^C &= \sum_{j=1}^{t_1} \frac{Ph_j^C}{\omega_0^C} \ln \left( \frac{Ph_j^C}{\omega_0^C} \right) \\
 H_2^C &= \sum_{j=t_{1+1}}^{t_2} \frac{Ph_j^C}{\omega_1^C} \ln \left( \frac{Ph_j^C}{\omega_1^C} \right) \\
 &\vdots \\
 H_k^C &= \sum_{j=t_{k+1}}^L \frac{Ph_j^C}{\omega_{k-1}^C} \ln \left( \frac{Ph_j^C}{\omega_{k-1}^C} \right)
 \end{aligned} \tag{2}$$

where  $Ph_j^C$  denotes the probability distribution and  $\omega_0^C, \omega_1^C, \omega_{k-1}^C$  denotes the probability action for  $K$  levels. At the time of multi-level thresholding process, it is needed to decide the optimum threshold value  $T$  that maximize the objective function  $f(T)$ . Here, the maximization of  $(f(T))$  is done by the use of GWO algorithm.

GWO method is deployed by [14], which has been evolved from the hunting nature of grey wolves. These wolves are assumed to be top-level predators and reside in a group of 5-12 wolves. Based on the hunting nature, the wolves are categorized into alpha ( $\alpha$ ), beta ( $\beta$ ), delta ( $\delta$ ), and omega. Initially, dominant wolves are named to be original dominators for the team, as it is capable of making decisions according to the residing location, hunting, and so on. It is termed as dominant wolves. Additional wolves follow the rules of alpha wolves. The next level wolves are beta that makes decisions in absence of alpha wolves. It is given with some rights of finding solutions for alpha wolves and offers a response accordingly. Thirdly, the delta wolves are the consecutive level wolves, called as subordinate wolves. It comes under the type of elders, sentinels, hunters, scouts, and caretakers. The delta wolves monitor the task of alpha and beta wolves and manage the subsequent levels known as omega wolves. At last, omega is final ranking wolves and treated as scapegoat. It applies the process recommended by alpha wolves. The omegas are ineffective wolves, and it help is resolving the internal problems.

Hunting process of wolves are categorized into tracking and chasing, pursuing and surrounding until it terminates moving, and attacking the prey. GWO is operated using exploration and exploitation phase. The exploitation phase performs the task of exploring best solutions in local search space. The grey wolves surround and attack the prey at the time of searching better solutions from local search space. When the prey is surrounded, the wolves identify prey's location and attack them. In this approach, the position vectors of a prey have been implied and find the agents that change the location according to accomplished optimal solution which is illustrated as:

$$\vec{D} = |\vec{C} \cdot \vec{Q}_p(n) - \vec{Q}(n)| \tag{3}$$

$$\vec{Q}(n+1) = \vec{Q}_p(n) - \vec{A} \cdot \vec{D} \tag{4}$$

Where  $n$  is the recent iteration,  $\vec{A}$  and  $\vec{C}$  are coefficient vectors, position vector of a prey as  $\vec{Q}_p(n), \vec{Q}$  denotes the position vector,  $||$  implies absolute value and  $\cdot$  shows an element by element improvement. The vectors  $\vec{A}$  and  $\vec{C}$  are expressed as:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{5}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{6}$$

In hunting phase, grey wolves are triggered by  $\alpha$  and only the contributions are provided to  $\beta$  and  $\delta$ . Initially, the optimum solution cannot be found as it consumes maximum amount of energy in hunting phase,  $\alpha$  is the primary better candidate solution, beta defines upcoming optimal candidate solution and lastly delta represents consequent better candidate solution. The 3 solutions are reached and extended to change the position of lower ranking solution which demonstrates the hunting phase:

$$\begin{aligned}
 \vec{D}_\alpha &= |\vec{C}_1 * \vec{Q}_\alpha - \vec{Q}| \\
 \vec{D}_\beta &= |\vec{C}_2 * \vec{Q}_\beta - \vec{Q}| \\
 \vec{D}_\delta &= |\vec{C}_3 * \vec{Q}_\delta - \vec{Q}|
 \end{aligned} \tag{7}$$

Where  $\vec{D}_\alpha, \vec{D}_\beta$ , and  $\vec{D}_\delta$  are changed distance vector among the  $\alpha, \beta$  and  $\gamma$  positions to alternate wolves and  $\vec{C}_1, \vec{C}_2$ , and  $\vec{C}_3$  are 3 coefficient vector applied in varying distance vector.  $\vec{Q}$  refers the position of vector of grey wolf (omega).

$$\begin{aligned}
 \vec{Q}_1 &= \vec{Q}_\alpha - \vec{A}_1 * \vec{D}_\alpha \\
 \vec{Q}_2 &= \vec{Q}_\beta - \vec{A}_2 * \vec{D}_\beta \\
 \vec{Q}_3 &= \vec{Q}_\delta - \vec{A}_3 * \vec{D}_\delta
 \end{aligned} \tag{8}$$

where  $\vec{Q}_1$  is novel position vector by exploiting the  $\alpha$  position  $\vec{Q}_\alpha$  and distance vector  $\vec{D}_\alpha$ ,  $\vec{Q}_2$  means a novel position vector is accomplished under the application of  $\beta$  position  $\vec{Q}_\beta$  and distance vector  $\vec{D}_\beta$ ,  $\vec{Q}_3$  shows the novel position vector determined by applying delta position  $\vec{Q}_\delta$  as well as distance vector  $\vec{D}_\delta$ , and  $\vec{A}_1, \vec{A}_2,$  and  $\vec{A}_3$  are 3 coefficient vectors are evaluated.

$$\vec{Q}(n+1) = \frac{\sum_{i=1}^k \vec{Q}_i}{n} \quad (9)$$

where  $\vec{Q}(n+1)$  signifies new position vector is estimated by using  $\alpha, \beta$  and  $\delta$  ( $k=3$ ).

**DWT based Feature extraction:** In general, DWT is an efficient deployment of WT with the help of dynamic scales as well as positions. A basic DWT is established in [15]. Assume that  $u(t)$  refers square-integrable function, and continuous WT of  $u(t)$  related to provided wavelet  $\psi(t)$  is expressed as

$$W_\psi(a, b) = \int_{-\infty}^{\infty} u(t) \psi_{a,b}(t) dt \quad (10)$$

Where

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-a}{b}\right) \quad (11)$$

In this approach, the wavelet  $\psi_{a,b}(t)$  is acquired from native wavelet  $\psi(t)$  using translation and dilation, where  $a$  implies a dilation factor and  $b$  denotes translation parameter. The Harr wavelet is a significant wavelet, which is simple and mostly employed in many domains.

Eq. (11) is discretized by retaining  $a$  and  $b$  for discrete lattice ( $a = 2^j$  &  $a > 0$ ) to provide DWT as described below:

$$\begin{aligned} ca_{j,k}(n) &= DS \left[ \sum_n u(n) g_j^*(n - 2^j k) \right] \\ cd_{j,k}(n) &= DS \left[ \sum_n u(n) h_j^*(n - 2^j k) \right] \end{aligned} \quad (12)$$

Here  $ca_{j,k}$  and  $cd_{j,k}$  means the coefficients of approximation components and detail components, correspondingly;  $g(n)$  and  $h(n)$  are low-pass filter and high-pass filter;  $j$  and  $k$  shows the wavelet scale as well as translation factors, and DS is the down sampling.

In case of 2D images, the DWT is used for all dimensions. It is composed with 4 sub-bands such as LL, LH, HH, and HL images at every scale. The sub-band LL is applied for consecutive 2D DWT. The LL sub-band is named as approximation unit of an image, while LH, HL, and HH sub-bands are known to be expanded image components. The decomposition level has been enhanced, and then coarser approximation component can be reached. Therefore, wavelets provide easy hierarchical approach to interpret the image details.

**Feature Reduction using KPCA:** The existence of massive features leads to maximum processing time and requires more storage. Besides, it makes the classification process difficult, known as curse of dimensionality and is needed to minimize the feature count. KPCA is an extended version of PCA, which has been evolved by kernel models. Under the application of a kernel, the actual linear task of PCA was processed in regenerating kernel Hilbert space. The component of PCA can be established for clustering, and it is monitored that, if  $N$  points could not be a general term, then it is linearly classified in  $d < N$  dimensions, it is prominently separated in a linear manner in  $d \geq N$  dimensions. The given  $N$  points  $x_i$ , when it is mapped into  $N$ -dimensional space with

$$\Phi(u_i) \text{ where } \Phi: \mathbb{R}^d \rightarrow \mathbb{R}^N, \quad (13)$$

It is simply manufactured as a hyper-plane which classifies the points into random clusters. Obviously,  $\Phi$  develops the autonomous vectors randomly, thus no covariance in eigen decomposition could be explicitly processed in linear PCA.

**DNN based Classification:** Once the feature reduction process is done, classification process takes place on the feature values or vectors. Normally, classification is represented as a boundary among the classes for labeling the classes depending upon the determined features. Here, DNN is applied as a classification model. The DNN generally operated on FFNN and is defined as an unsupervised pre-training model which uses greedy layer wise training. In this approach, the data flow from input layer to output layer with no looping function. The benefit in DNN classification is the nature of classifying the feasibilities of missing values is minimum. It shows a single layer in unsupervised pre-training level [16]. The DNN

assigns a classification score  $t(x)$  over a prediction time. Each input data sample  $x = [x_1, \dots, x_N]$  is a forward pass. Then,  $f$  signifies the function where a series of layers for processing, that is implied in Eq. (14)

$$Z_{ij} = X_i W_{ij}; Z_j = \sum_i Z_{ij} + b_j; X_j = g(Z_j), \tag{14}$$

where input layer is depicted as  $x_i$ , output layer is  $x_j$ , and  $w_{ij}$  are modelling parameters and  $g(Z_j)$  analyze the pooling function. Layer-wise relevance propagation degrades the classification result  $f(x)$  with respect to relevance's  $r_i$  with input component  $x_i$ , and applied on classification process as defined in Eq. (15):

$$f(x) = \sum_i r_i \tag{15}$$

where  $r_j > 0$  is the positive evidence that supports the classification decision and  $r_j < 0$  shows the negative evidence of classification; else, it named as neutral evidence, even though related attribute  $r_i$  is evaluated using Eq. (16).

$$r_i = \sum_j \frac{z_{ij}}{\sum_i z_{ij}} \tag{16}$$

The typical structure of DNN is shown in Fig. 2. The DNN is capable to analyze the unwanted feature correlation of input. It offers a hierarchical feature learning technology. Hence, higher-level features can be accomplished from minimum level features using greedy layer wise unsupervised pre-training data. Therefore, the main aim of DNN is to resolve the tedious functions and results in high-level abstraction.

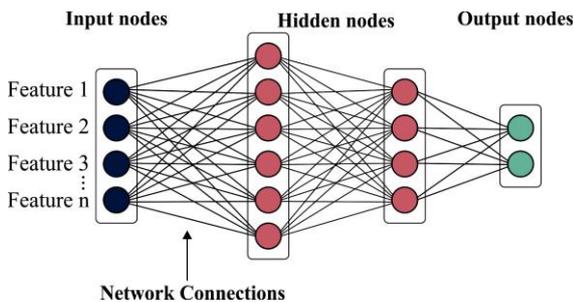


Fig. 2: Architecture of Deep Neural Networks (DNN)

## RESULTS AND DISCUSSION

**Dataset used:** The working function of projected approach is verified with the application of standard dataset from Kaggle [17]. The dataset comprises a total of 98 normal MRI brain images and 155 abnormal MRI brain images. The size of the images varies between 192\*192 and 630\*630.

**Results Analysis:** A sample visualization of the results offered by the presented OS-DNN model on the applied images is illustrated in Fig. 3. The original test images are shown in Fig. 3a, contrast enhanced image is depicted in Fig. 3b, skull stripped image is provided in Fig. 3c, segmented image in Fig. 3d, and finally, classified image in Fig. 3e. The figures clearly ensured that the OS-DNN model has effectively classified the images into its appropriate class labels.

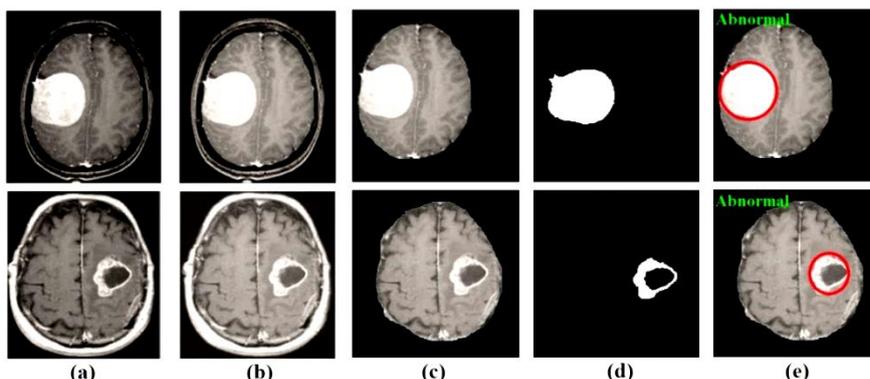
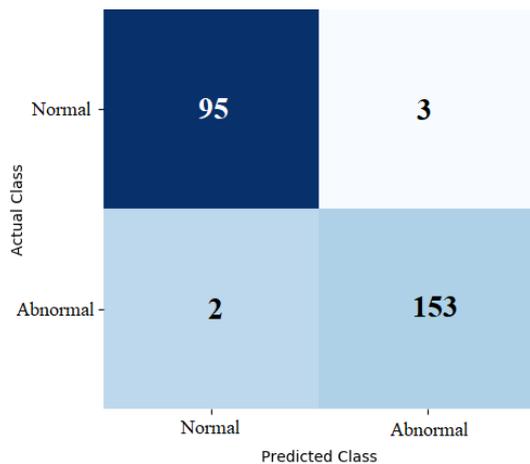


Fig. 3: a) Original b) Contrast Enhanced c) Skull stripped d) Segmented e) Classified



**Fig. 4:** Confusion Matrix of OS-DNN

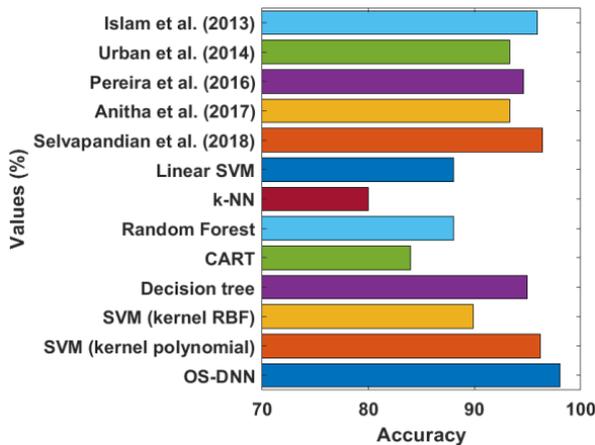
[Fig. 4] displays the confusion matrix of the OS-DNN model generated at the time of execution time. The figure clearly stated that the proposed OS-DNN model has properly classified a set of 95 images as normal out of 98 and 153 images as abnormal out of 155. These values provided that the OS-DNN model has effectively classified the MRI brain images.

A detailed comparative study of the results offered by the OS-DNN and existing models is shown in Table 1.

**Table 1:** Result Analysis of Existing with Proposed Method in term of Sensitivity, Specificity, Accuracy

Methods	Sensitivity	Specificity	Accuracy
Proposed OS-DNN	97.94	98.08	98.02
SVM (kernel polynomial)	94.73	97.59	96.18
SVM (kernel RBF)	95.62	83.71	89.88
Decision tree	97.88	91.71	94.95
CART	88.00	80.00	84.00
Random Forest	96.00	80.00	88.00
k-NN	80.00	80.00	80.00
Linear SVM	96.00	80.00	88.00
Anitha et al. [18]	91.20	93.40	93.30
Urban et al. [19]	92.60	93.00	93.30
Pereira et al. [20]	94.20	94.40	94.60
Islam et al. [21]	94.30	95.10	95.90
Selvapandian et al. [22]	96.20	95.10	96.40

Fig. 5 depicts the classification result analysis of OS-DNN with related models with respect to accuracy. The figure refers that the k-NN model has demonstrated poor classifier results with the accuracy of 80%. Besides, the CART approach has showcased considerable accuracy of 84%. On the other hand, the RF as well as Linear SVM models have shown better and identical accuracy values of 88%. Additionally, the SVM has displayed considerable accuracy of 89.88%. Simultaneously, the approaches used by Anitha et al. (2017) [18] and Urban et al. (2014) [19] have exhibited even more and identical accuracy values of 93.30%. Additionally, the model deployed by Pereira et al. (2016) [20] and DT models attempted to indicate applicable results with the accuracy of 94.60% and 94.95% correspondingly. In addition, the method deployed by Islam et al. (2013) [21] has accomplished an accuracy of 95.90%. Moreover, the SVM (kernel polynomial) technique has showcased better accuracy of 96.18%. In line with this, the Selvapandian et al. (2018) [22] framework has exhibited near optimal results with the accuracy of 96.40%. At last, the OS-DNN model has showcased optimal outcome than related models with the higher accuracy value of 98.02%.



**Fig. 5:** Accuracy analysis of OS-DNN with other models

From the Table-1 and Fig-5, it is obvious that the OS-DNN model can be applied as an effective diagnosis and classification model for BT from MRI brain images.

## CONCLUSION

Effective detection and classification of BT has received maximum attention among research community and medical institutions. This paper has presented novel BT detection and classification model using OS-DNN model. The proposed OS-DNN model initially undergoes preprocessing to improve the classification performance by noise removal, contrast enhancement, and skull stripping. Followed by, GWO-KT based segmentation, DWT based feature extraction, KPCA based feature reduction, and DNN based classification processes are carried out. The proposed OS-DNN model effectively detects and classifies the affected regions in the MRI image. The experimentation of the OS-DNN model is carried out using a benchmark Kaggle dataset and simulation outcome ensured that the OS-DNN model is superior to other models with the maximum sensitivity of 97.94%, specificity of 98.08% and accuracy of 98.02%. In future, the performance of the OS-DNN model can be improved by the use of hyper parameter tuning models to effectively choose the parameters involved in DNN.

### CONFLICT OF INTEREST

There is no conflict of interest.

### ACKNOWLEDGEMENTS

None

### FINANCIAL DISCLOSURE

None

## REFERENCES

- [1] Abd-Allah MK, Awad AI, Khalaf AAM, Hamed HFA. [2019] A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. *Magn Reson Imaging*, 61:300-318.
- [2] Patel J, Doshi K. [2014] A study of segmentation methods for detection of tumor in brain MRI. *Advance in Electronic and Electric Engineering*, 4(3):279-84.
- [3] Logeswari T, Karnan M. [2010] An improved implementation of brain tumor detection using segmentation based on hierarchical self-organizing map. *International Journal of Computer Theory and Engineering*, 2(4):591.
- [4] Beno MM, Rajakumar BR. [2014] Threshold prediction for segmenting tumour from brain MRI scans. *International journal of imaging systems and technology*, 24(2):129-37.
- [5] El-Dahshan ES, Mohsen HM, Revett K, Salem AB. [2014] Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. *Expert systems with Applications*, 41(11):5526-45.
- [6] Liao L, Zhang Y. [2011] MRI image segmentation based on fast kernel clustering analysis. *Frontiers of Electrical and Electronic Engineering in China*, 6(2):363-73.
- [7] Ali H, Elmogy M, El-Daydamony E, Atwan A. [2015] Multi-resolution mri brain image segmentation based on morphological pyramid and fuzzy c-mean clustering. *Arabian Journal for Science and Engineering*, 40(11):3173-85.
- [8] Ahmadvand A, Kabiri P. [2016] Multispectral MRI image segmentation using Markov random field model. *Signal, Image and Video Processing*, 10(2):251-258.
- [9] Cabria I, Gondra I. [2017] MRI segmentation fusion for brain tumor detection. *Information Fusion*, 36:1-9.
- [10] Cigaroudy LS, Aghazadeh N. [2017] A multiphase segmentation method based on binary segmentation method for Gaussian noisy image. *Signal, Image and Video Processing*, 11(5):825-8231.
- [11] Akbarizadeh G. [2013] Segmentation of SAR satellite images using cellular learning automata and adaptive chains. *J. Remote Sens. Technol*, 1(2):44-51.
- [12] Rahmani M, Akbarizadeh G. [2015] Unsupervised feature learning based on sparse coding and spectral clustering for segmentation of synthetic aperture radar images. *IET Computer Vision*, 9(5):629-638.
- [13] Rajinikanth V, Satapathy SC, Fernandes SL, Nachiappan S. [2017] Entropy based segmentation of tumor from

- brain MR images—a study with teaching learning based optimization. *Pattern Recognition Letters*, 94:87-95.
- [14] Mirjalili S, Mirjalili SM, Lewis A. [2014] Grey wolf optimizer. *Advances in engineering software*, 69:46-61.
- [15] Zhang Y, Wang S, Ji G, Dong Z. [2013] An MR brain images classifier system via particle swarm optimization and kernel support vector machine. *The Scientific World Journal*, DOI: 10.1155/2013/130134.
- [16] Suresh R, Rao AN, Reddy BE. [2019] Detection and classification of normal and abnormal patterns in mammograms using deep neural network. *Concurrency and Computation: Practice and Experience*, 31(14):5293.
- [17] <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>
- [18] Anitha R, Raja DSS. [2017] Segmentation of glioma tumors using convolutional neural networks. *International Journal of Imaging Systems and Technology*, 27(4):354-360.
- [19] Urban G, Bendszus M, Hamprecht FA, Kleesiek J. [2014] Multimodal brain tumor segmentation using deep convolutional neural networks. *MICCAI Brain Tumor Segmentation (BraTS) Challenge. Proceedings*, 31– 35, Boston, MA, USA.
- [20] Pereira S, Pinto A, Alves V, Silva CA. [2016] Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Transactions on Medical Imaging*, 35: 1240- 1251.
- [21] Islam A, Reza S, Iftakharuddin KM. [2013] Multifractal texture estimation for detection and segmentation of brain tumors. *IEEE Trans. Biomed. Eng.*, 60:3204–3215.
- [22] Selvapandian A, Manivannan K. [2018] Fusion based glioma brain tumor detection and segmentation using ANFIS classification. *Computer methods and programs in biomedicine*, 166: 33-38.