



Automated Brain Tumour Classification using Hyperparameter Tuned Deep Learning Model

¹B. Devanathan, ²Dr. M. Kamarasan

¹Research Scholar, Department of Computer and Information Science, Annamalai University, Chidambaram, India.

²Professor, Department of Computer and Information Science, Annamalai University, Chidambaram, India.
devacisau@gmail.com, smkrasan@yahoo.com

Abstract

Radiology is a wide-ranging subject that requires further understanding and knowledge of medical science to accurately find tumors. The necessity for a tumor detection program, therefore, overcome the lack of trained radiologists. By utilizing magnetic resonance imaging, biomedical image processing make it easy to locate and detect brain cancers. This article develops a new differential evolution with deep learning based brain tumor classification (DEDL-BTC) model on MRIs. The presented DEDL-BTC model aims to identify the occurrence of BT using medical images. The DEDL-BTC model initially eradicates the node and U-Net model is employed to segment the MRI. Besides, depth wise separable convolution (DSC) network with long short term memory (LSTM) model can be employed as a feature extractor and classifier respectively. Eventually, the hyperparameter tuning of the DSC take place using the DE technique. The results indicated the DEDL-BTC method has shown improved performance than the other models interms of distinct measures.

Keywords: Brain tumor, Segmentation; Deep learning; Medical imaging; Computer aided diagnosis

1. Introduction

Brain tumor segmentation and detection copes with splitting the tumor portion from the normal tissues [1, 2]. It is performed automatically or manually via a software. But the human decision are time consuming and is not often accurate that results in computer analysis of tumor sample. Image processing plays a crucial role in the abovementioned technique. Object detection find the position of certain objects in the image and bound them through a bounding box [3, 4]. Where segmentation offer the exact outline of image by allocating class label for all the pixels of the object. Conventional approaches for image segmenting and detection [5]. But the current study shows that Machine Learning (ML) is a subdivision of artificial intelligence (AI) allows to effectively recognize region of interest (ROI) in information when compared to the abovementioned techniques [6]. The Deep learning (DL) which is a subclass of ML approaches has the capability to examine more than one levels of data demonstration for the detection and prediction tasks [7]. Deep neural network (DNN) includes Convolution neural network (CNN), efficiently works on image generating notable outcomes

that involves image classification, segmentation, and detection of an object. The study mainly focused on object detection and classification [8].

In [9], a novel automated DL technique was suggested for the categorization of multiclass brain tumors. Entropy–Kurtosis-based High Feature Values (EKbHFV) was primary method and another system was modified GA (MGA) related to meta-heuristics. The GAs selected features were again advanced by projected novel threshold operation. At last, MGA-related, as well as EKbHFV features, were merged with the help of non-redundant serial-related method and categorized with the help of multiclass SVM cubic classifier. Kang et al. [10] implement the TL concept and usages numerous pretrained deep CNN for extracting deep features from MR imageries. The mentioned features which were extracted were after assessed by numerous ML classifiers. The top 3 deep features that execute well on some ML classifications were chosen and concatenated as an ensembling deep feature that was after put into numerous ML classifiers for predicting the concluding output. Díaz-Pernas et al. [11] provide a fully automated brain tumor classifying and segmenting method with the help of Deep CNN which invokes a multiscale method. One major difference of this method in terms of preceding works was input images were processed in 3 spatial scales with distinct processing ways. This article develops a new differential evolution with deep learning based brain tumor classification (DEDL-BTC) model on MRIs. The presented DEDL-BTC model aims to identify the occurrence of BT using medical images. The DEDL-BTC model initially eradicates the node and U-Net model is employed to segment the MRI. Besides, depth wise separable convolution (DSC) network with long short term memory (LSTM) can be employed as a feature extractor and classifier respectively. Finally, the hyperparameter tuning of the DSC model take place using the DE technique. The results denoted the DEDL-BTC method has shown improved performance than the other models interms of distinct measures.

2. The Proposed Model

This article developed a new DEDL-BTC technique to classify BTs on MRIs. The presented DEDL-BTC model aims to identify the occurrence of BT using medical images. The DEDL-BTC model initially eradicates the node and U-Net model is employed to segment the MRI. Besides, the DSC network with LSTM can be employed as a feature extractor and classifier respectively. Finally, the hyperparameter tuning of the DSC model take place using the DE technique.

In this study, a new DEDL-BTC model recognizes the presence of BT using medical images. The DEDL-BTC model undergoes image filtering and U-Net image segmentation at the pre-processing stage. In addition, the DSC network with LSTM can be employed as a feature extractor and classifier. CNN is a widely applied DL algorithm that is used in classification tasks, for instance, image classification. Initially, CNN is characterized by an image, and computer vision is used as visual cortex.

An image tensor is convolved by the series of $d \times d$ kernels. This convolution (“Feature Map”) is stacked to characterize dissimilar features discovered by the filter in that layer and it can be explained in the following equation:

$$A_j = f \left(\sum_{i=1}^N I_i * K_{i,j} + B_j \right) \quad (1)$$

In every matrix, I_i is convolved with its $K_{i,j}$ corresponding kernel matrix, and bias of B_j [12]. To assure the model's accuracy and enhance its efficiency on mobile terminals, a sequence of lightweight convolution operators was proposed. Out of which, the most broadly utilized was a depthwise separable convolution. This convolution split up the convolution into 2 steps: firstly to convolve single channel parallelly with the kernel_size = 3 convolution kernel termed depthwise convolution and secondly for fusing the feature map by utilizing the 1×1 convolution kernel known as pointwise convolution.

At last, DE is used as hyperparameter optimizer for the DSC model. The DE approach is an evolutionary optimization algorithm [13]. DE using better global search capability employs mutation and crossover operations such as GA but with dissimilar mechanisms. In c dimension searching space, every individual in the DE population is generated as $X_j = (x_{j,1}, \dots, x_{j,c})$ from the best possible solution space. In every generation, mutation vector (X_{mut} , X_{mut1} , X_{mut2} and X_{mut3}) is produced for every individual based on the subsequent three approaches:

$$\begin{aligned} X_{mut} &= X_{p0} + F_1 \times (X_{p1} - X_{p2}) \\ X_{mut1} &= X_{p0} + F_1 \times (X_{p1} - X_{p2}) + F_2 \times (X_{best} - X_{worst}) \\ X_{mut2} &= F_3 \times (X_{best} - X_{worst}) + X_{p3} \\ X_{mut3} &= X_{p0} + F_1 \times (X_{p4} - X_{p3}) \end{aligned} \quad (2)$$

In Eq. (2), $p0$, $p1$, $p2$, $p3$, and $p4$ are the different random number ranges from $[1, N_{pop}]$ and N_{pop} was the size of population. To alter the differential vectors F_1 , F_2 and F_3 are designated as scaling factors arbitrarily chosen through the standard distribution.

Afterward employing mutation, trial vectors X_{new} , X_{new1} , X_{new2} and X_{new3} called crossover vector is generated by crossover function according to the X_j and X_{mut} , X_{mut1} , X_{mut2} , and X_{mut3} vectors as:

$$x_{new,m} = \begin{cases} x_{mut1,m} & \text{if rand} < CR \text{ or } m = z \\ x_j, m & \text{or else} \end{cases} \quad (3)$$

$$x_{new1,m} = \begin{cases} x_{mut3,m} & \text{if rand} < CR1 \text{ or } m = z \\ x_j, m & \text{or else} \end{cases} \quad (4)$$

$$x_{new2,m} = \begin{cases} x_{mut,m} & \text{if rand} < CR2 \text{ or } m = z \\ x_{mut2,m} & \text{otherwise} \end{cases} \quad (5)$$

$$x_{new3,m} = \begin{cases} x_{mut,m} & \text{if rand} < CR3 \text{ or } m = z \\ x_{mut1,m} & \text{otherwise} \end{cases} \quad (6)$$

$M=1 \dots c$. whereby x_j, m and x_{new}, m , x_{new1}, m , x_{new2}, m , x_{new3}, m indicates the m^{th} dimension component of the vector X_j and x_{new} , x_{new1} , x_{new2} , x_{new3} , correspondingly. CR , $CR1$, $CR2$, and $CR3$ indicate the

crossover constant that generally ranges from zero and one; z represents an integer that is chosen arbitrarily from $\{1, 2, c\}$ that ensures crossover vector gets minimum one variable from mutation vector. Fig. 1 shows the structure of LSTM model.

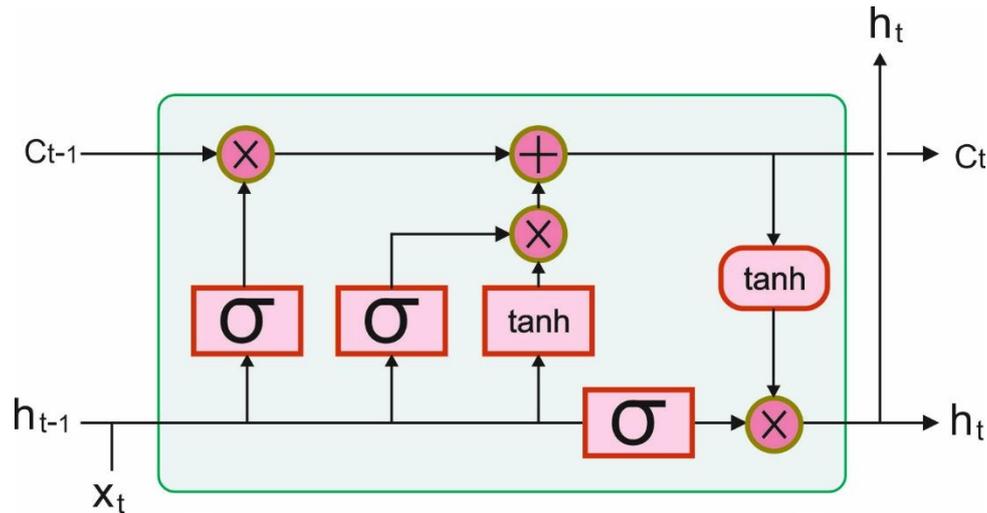


Fig. 1. Structure of LSTM model

The LSTM upgrades the cell state on the basis of gate activation function. The input gives an LSTM that is fed to distinct gates which controls the process are implemented on the cell memory: write (input gate), read (output gate) or reset (forget gate). The calculation of h_t hidden value of LSTM cell is upgraded at t time step as follows.

$$i_t = \sigma_i(W_{ai}a_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (7)$$

$$f_t = \sigma(W_{af}a_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (8)$$

$$c_t = f_t c_{t-1} + i_t \sigma_c(W_{ac}a_t + W_{hc}h_{t-1} + b_c) \quad (9)$$

$$o_t = \sigma_o(W_{ao}a_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (10)$$

$$h_t = o_t \sigma_h(c_t) \quad (11)$$

From the expression, c , i , f , & o , characterizes the cell activation vector, input, forget, and output gates, the equivalent size as h vector determines the hidden value. σ indicates nonlinear function. a_t characterize the input to the memory cell layer at t time. W_{ai} , W_{hi} , W_{ci} , W_{af} , W_{hf} , W_{cf} , W_{ac} , W_{hc} , W_{ao} , W_{ho} , & W_{co} shows weight matrixes, using subscript establishing from-to relationships. b_j , b_f , b_c , & b_o indicates bias vector.

3. Result Analysis

The presented model is tested using the MRI images from Kaggle repository [14]. The dataset encompasses of images under two classes namely normal and abnormal/tumor. Fig. 2 shows the sample images.

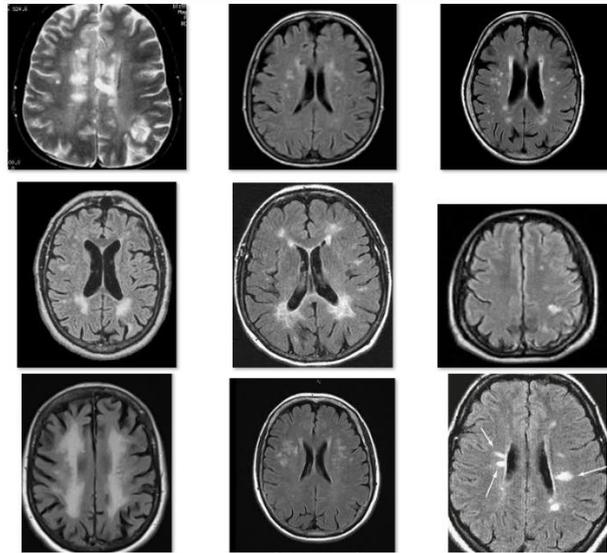


Fig. 2. Sample Images

Table 1 and Fig. 3 reports an overall BT classification results of the DEDL-BTC method. The results implied that the DEDL-BTC method has gained effectual output under all runs. On run-1, the DEDL-BTC method has offered $sens_y$, $spec_y$, and $accu_y$ of 97.87%, 98.13%, and 98.24% respectively. Also, on run-2, the DEDL-BTC method has offered $sens_y$, $spec_y$, and $accu_y$ of 98.19%, 98.24%, and 97.88% respectively. In addition, on run-3, the DEDL-BTC method has attained $sens_y$, $spec_y$, and $accu_y$ of 98.14%, 97.92%, and 97.73% respectively. Moreover, on run-4, the DEDL-BTC method has reached $sens_y$, $spec_y$, and $accu_y$ of 97.57%, 98.01%, and 98.20% respectively. Furthermore, on run-5, the DEDL-BTC method has offered $sens_y$, $spec_y$, and $accu_y$ of 98.08 %, 98.09%, and 97.84% respectively.

Table 1 BT classification outcomes of DEDL-BTC method

No. of Runs	Sensitivity	Specificity	Accuracy
Run-1	97.87	98.13	98.24
Run-2	98.19	98.24	97.88
Run-3	98.14	97.92	97.73
Run-4	97.57	98.01	98.20
Run-5	98.08	98.09	97.84
Average	97.97	98.08	97.98

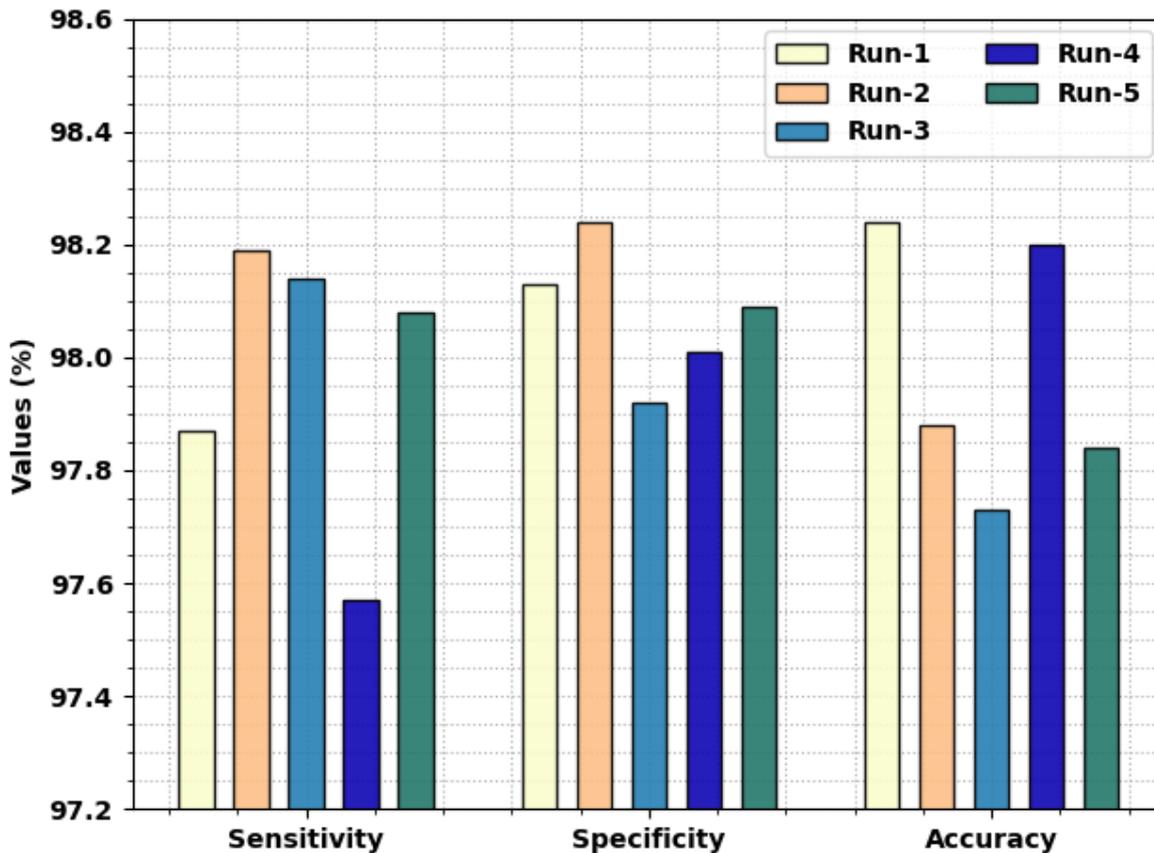


Fig. 3. Overall BT classification outcomes of DEDL-BTC method

To highlight the improved outcomes of the DEDL-BTC method, an average result examination under BT classification process is performed in Fig. 4. The figure showcased that the DEDL-BTC method has offered average $sens_y$, $spec_y$, and $accu_y$ of 97.97%, 98.08%, and 97.98% respectively.

The training accuracy (TA) and validation accuracy (VA) attained by the DEDL-BTC method on test dataset is demonstrated in Fig. 5. The experimental outcome implied that the DEDL-BTC method has gained maximal values of TA and VA. In specific, the VA seemed that superior to TA.

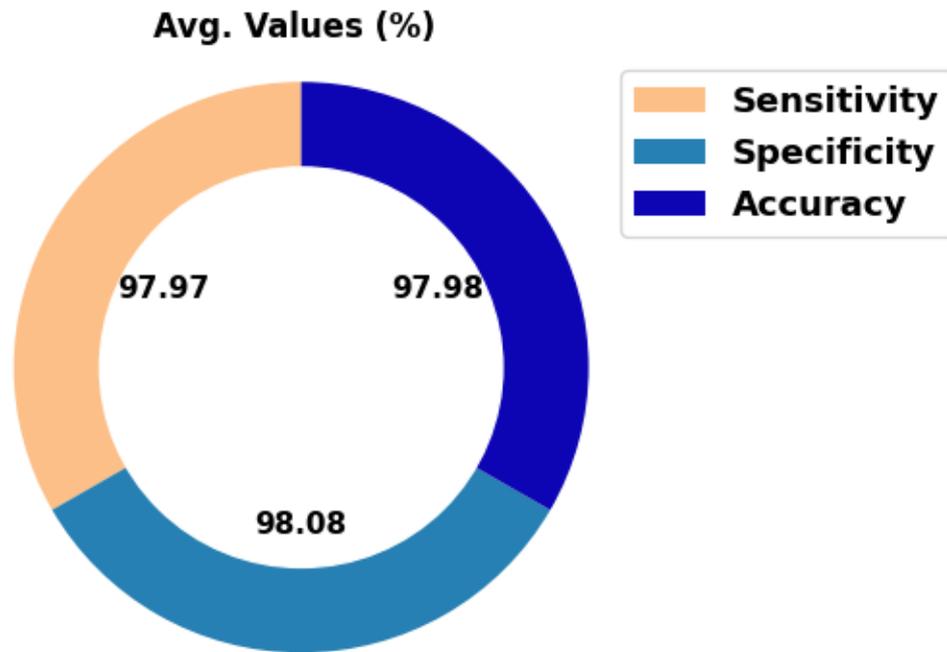


Fig. 4. Average classification outcomes of DEDL-BTC method

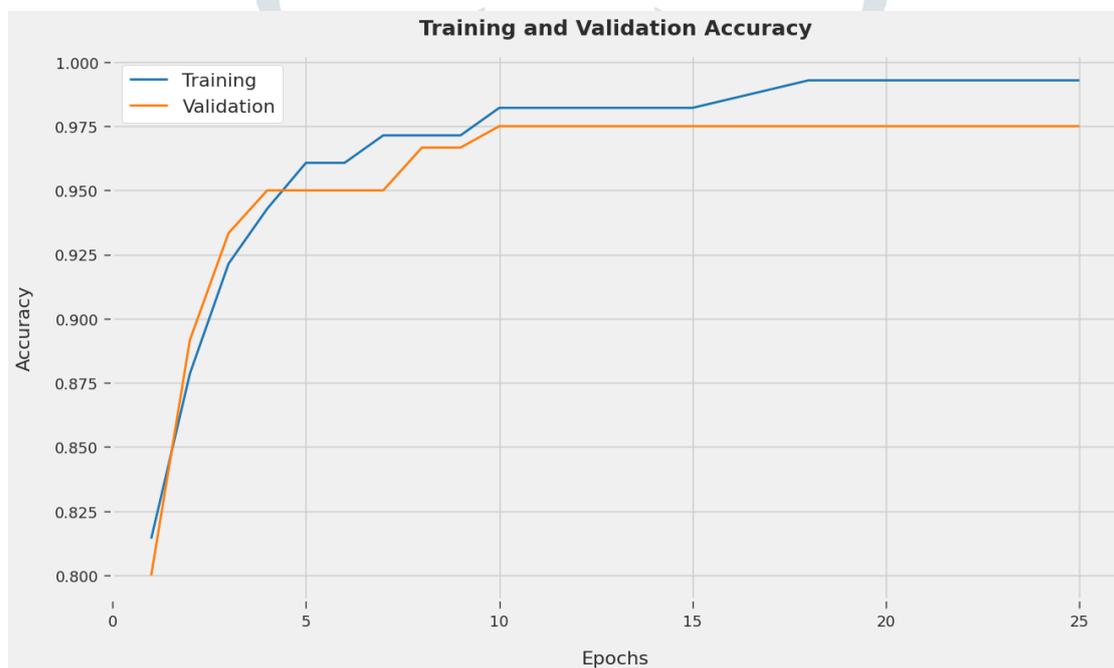


Fig. 5. TA and VA of DEDL-BTC model

The training loss (TL) and validation loss (VL) achieved by the DEDL-BTC method on test dataset are established in Fig. 6. The experimental outcome exposed that the DEDL-BTC method has been able to achieve the lowest values of TL and VL. In specific, the VL seemed lower than TL.



Fig. 6. TL and VL of DEDL-BTC model

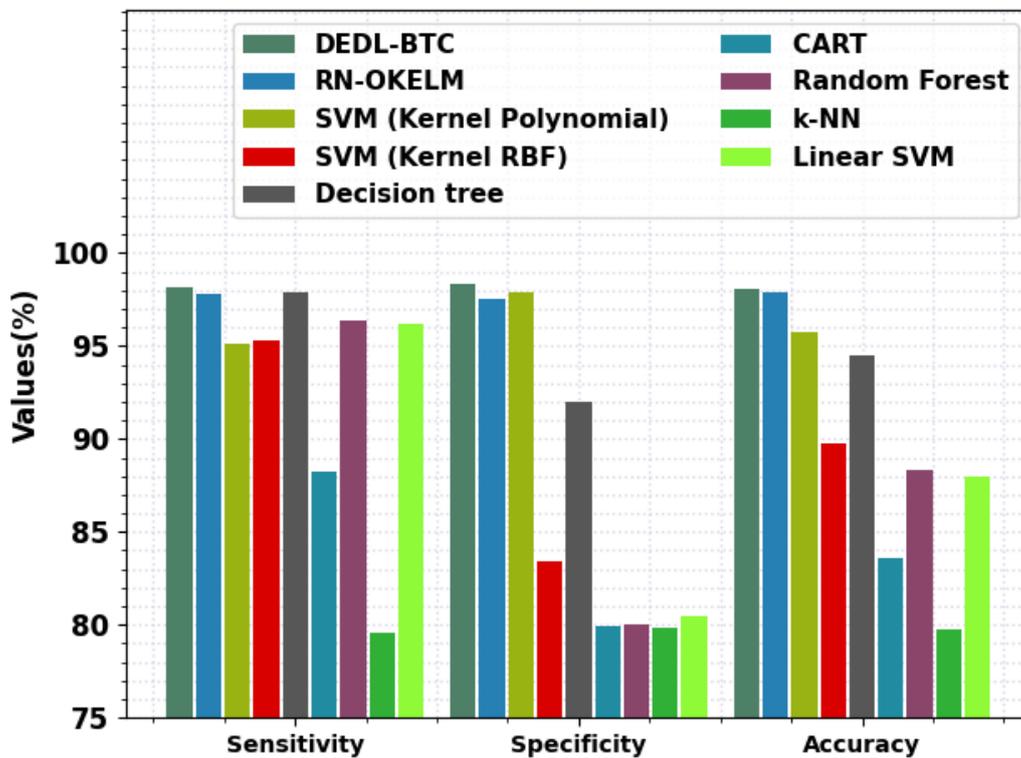


Fig. 8. Accuracy analysis of proposed with recent models

To assure the enhanced performance of the DEDL-BTC model, a comparison study is made with recent models [15, 16] in Fig. 7. The results implied that the CARD, RF, KNN-SVM, and linear SVM, and SVM (kernel RBF) models have obtained poor performance. But the DEDL-BTC model has offered higher performance with sensitivity of 97.97%, specificity of 98.08%, and accuracy of 97.98%.

4. Conclusion

In this study, a new DEDL-BTC method recognizes the presence of BT using medical images. The DEDL-BTC algorithm undergoes image filtering and U-Net image segmentation at the pre-processing stage. In addition, the DSC network with LSTM was utilized as a feature extractor and classifier. At last, DE is used as hyperparameter optimizer for the DSC model. The outcomes denoted the DEDL-BTC technique has shown improved performance than other models in terms of distinct measures. In future, hybrid DL classification models can be integrated to enhance the classification performance.

References

- [1] Abiwinanda N, Hanif M, Hesaputra ST, Handayani A, Mengko TR (2019) Brain tumor classification using convolutional neural network. *IFMBE Proc* 68(1):183–189.
- [2] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Vander Laak JAWM, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88.
- [3] Doğantekin A, Özyurt F, Avcı E, Koç M (2019) A novel approach for liver image classification PH-C-ELM. *Measurement* 137:332–338.
- [4] Banan R, Hartmann C (2017) The new WHO 2016 classification of brain tumors—what neurosurgeons need to know. *Acta Neurochir* 159(3):403–418.
- [5] Kabir Anaraki A, Ayati M, Kazemi F (2019) Magnetic Resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *Biocybern Biomed Eng* 39(1):63–74.
- [6] Çinar A, Yildirim M (2020) Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. *Med Hypotheses* 139:109684
- [7] Muhammad K, Khan S, Ser JD, Albuquerque VHC (2021) Deep learning for multigrade brain tumor classification in smart healthcare systems: a prospective survey. *IEEE Trans Neural Netw Learn Syst* 32(2):507–522.
- [8] Mehrotra R, Ansari MA, Agrawal R, Anand RS (2020) A Transfer learning approach for AI-based classification of brain tumors. *Mach Learn Appl* 2(9):1–12.
- [9] Sharif, M.I., Khan, M.A., Alhussein, M., Aurangzeb, K. and Raza, M., 2021. A decision support system for multimodal brain tumor classification using deep learning. *Complex & Intelligent Systems*, pp.1-14.
- [10] Kang, J., Ullah, Z. and Gwak, J., 2021. Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors*, 21(6), p.2222.
- [11] Díaz-Pernas, F.J., Martínez-Zarzuela, M., Antón-Rodríguez, M. and González-Ortega, D., 2021, February. A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. In *Healthcare* (Vol. 9, No. 2, p. 153). MDPI.
- [12] Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251-1258).

- [13] Price, K.V., 2013. Differential evolution. In *Handbook of optimization* (pp. 187-214). Springer, Berlin, Heidelberg.
- [14] <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>
- [15] Edalati-rad, A. and Mosleh, M., 2019. Improving Brain Tumor Diagnosis Using MRI Segmentation Based on Collaboration of Beta Mixture Model and Learning Automata. *Arabian Journal for Science and Engineering*, 44(4), pp.2945-2957.
- [16] Devanathan, B., and K. Venkatachalapathy. "An Optimal Multilevel Thresholding based Segmentation and Classification Model for Brain Tumor Diagnosis." In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1133-1138. IEEE, 2020.

