International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# Bayesian Weighted Convolutional Neural Network for Brain Tumor Classification

**Gayathri Govindappa Nalina[1*], Channakrishna Raju[1]**

**Submitted**: 07/05/2023 **Revised**: 13/07/2023 **Accepted**: 07/08/2023

**Abstract:** Brain Tumor Classification (BTC) in medical image processing is crucial for physicians to make precise diagnoses and treatment decisions. Brain tumours are classified as normal or malignant features by employing Magnetic Resonance Imaging (MRI) images. The extraction of features from the MRI images is crucial, because it identifies the object based on its color, name, shape, and other characteristics. Deep learning methods are currently used in BTC which facing overfitting and vanishing gradient issues. Here, this research is evaluated by the Contrast Enhancement – MRI (CE-MRI) dataset to test the efficiency of the model. After collecting the dataset, pre-processing is done by using min-max normalization which is applied to enhance the quality of the image and the difference between pixel values in the model. After pre-processing, Long Short-Term Memory (LSTM) model is used to extract the features. Finally, the classification stage is done by using Convolutional Neural Network (CNN) to classify the accurate features from the extracted images. For both the extraction and classification stages, optimal weights are updated using Bayesian Weighted (BW) which overcomes the overfitting and vanishing gradient problem in the network. The result analysis, clearly shows that the proposed BW-CNN model has 98.83 % accuracy which is better than the existing Xception model accuracy which has 98.04 %.

***Keywords:*** *Bayesian Weighted, Brain Tumor Classification, Convolutional Neural Network (Cnn), Feature Extraction, And Long Short-Term Memory*

## 1. Introduction

The abnormal collective behavior of brain cells is one of the most frequent causes of brain tumors today, which are among the most dreadful disorders. There are two types of tumors, one is primary and another one is secondary. Primary tumours are considered for tiny portions, while secondary tumors are considered for larger portions of the brain [1]. Computer-aided diagnosis helps for early-stage detection of brain tumors and reduces diagnosis time. The early-stage diagnosis of a tumor plays a significant role in treatment because it increases the rate of patient survival [2]. Image classification and segmentation techniques achieve huge importance in various applications such as interpreting, extracting features, understanding, and analyzing medical images [3]. Various efforts are made to develop a robust solution and accurate classification of brain tumors. This is a challenging problem due to contrast variations, texture, and high inter and intra shape [4]. Machine learning techniques with handcrafted features retains robust solution for this model. To provide noticeably greater performance, deep learning approaches gather significant features [5]. But deep learning models took a lot of time to classify and locate brain tumours in large datasets [6].

Recently, deep learning techniques provides excellent performance in the field of medical analysis and image processing. Some of the difficulties in classifying brain tumors in MRI are complex brain structure, tissues intertwining, and classifying brain tumors due to the brain's high density nature [7]. The early detection, classification, and analysis of brain tumors are critical for effective treatment. Several CAD systems have recently been launched in the territory of medical imaging to assist doctors and radiologists in diagnosing various kinds of illnesses and health conditions. As a result, ML approaches were considered to be the basis for taking over categorization and mining tasks [8]. Recently, medical data analysis was struggling to develop novel approaches for brain tumor detection to enhance classification accuracy. As a result, DL has gained popularity to create accurate prediction models employing large amounts of data such as photos and text [9]. Various deep learning techniques such as CNN, and LSTM were exploited for MRI classification and have restrictions of overfitting issues [10]. The objectives and contributions of this research are discussed below:

1. Initially, the dataset is collected from the CE-MRI dataset. Then the pre-processing is done using Min-Max Normalization to eliminate the noises/errors.
2. The LSTM is exploited for the feature extraction and the CNN model is used for the classification. The Bayesian optimization is applied for LSTM and CNN models to update the weight value.
3. The Bayesian Optimization technique is applied to find the optimal weight value for the neurons in the CNN and LSTM network. The Bayesian weight value is

[1] *Department of Computer Science and Engineering, Sri Siddhartha Academy of Higher Education, Tumkur, India*
** Corresponding Author Email:gaytri.infosea@gmail.com*

updated in the network to avoid overfitting and vanishing gradient problems.

4. The BW-CNN model has greater efficiency in the BTC than conventional techniques due to the efficient weight update process.

The paper is organized as follows: Section 2 presents the literature survey, and Section 3 provides an explanation of the BW-CNN model. Section 4 explains the finding, and Section 5 provides the conclusion of this research work.

## 2. Literature Review

Effective and Early diagnosis of the BTC supports efficient to manage the patients and reduces the mortality rate. Recent, deep learning-based techniques applied for the BTC were studied in this part.

Kumar et al. [11] developed the ResNet-50 deep network model and global average pooling to decrease overfitting and vanishing gradient issues. The ResNet-50 model is evaluated on the CE-MRI dataset involving 3064 images. The ResNet-50 model shows considerable improvement in solving the vanishing gradient problem. The ResNet-50 model has higher accuracy than existing BTC techniques in CE-MRI dataset. A Mean Accuracy of 97.08 % is achieved for augmented data and 97.48 % is achieved without data augmentation. As this model shows augmented results create the overfitting problem in the classification.

Shaik and Cherukuri [12] proposed a multi-level attention technique for BTC tasks based on the Xception model. The cross-channel attention and spatial information are applied in Multi-level Attention Network (MANet) to focus on prioritizing tumor regions and maintain temporal dependencies of cross-channel in semantic feature sequence from the backbone of Xception. The CE-MRI and BraTS benchmark datasets were exploited to assess the presence of the developed system. The spatial attention and cross-channel attention block were used to generalize better efficiency with less constraints.

Bodapati et al. [13] established Xception and InceptionResNetV2 networks of convolution blocks for feature extraction, and a pooling-based technique is proposed for vectorizing the features. An attention technique was applied to focus on tumor regions and less focus on non-tumor regions which helps to differentiate tumor types. The two-channel mode allows joining the two sets of training and end-to-end manner of tumor image representations. The Xception and InceptionResNetV2 were evaluated on two datasets such as CE-MRI and BraTS 2018 datasets in terms of some layers and generalization to fine-tune deep CNN models. The Xception and InceptionResNetV2 model helps to avoid pre-processing and data augmentation process. The Xception and InceptionResNetV2 model has an overfitting problem and degrades the efficiency of the model.

Alhassan and Zainon [14] applied histogram descriptors to describe contour and edge features of the images related to other feature descriptors. The normalization technique and oriented gradient histogram was applied to increase the visual level of MRI images and extract feature vectors from normalized images. CNN model was applied with extracted features to classify into pituitary, gliomas, and meningiomas tumors. The RELU activation function based on the hard swish was utilized in CNN to effectively improve classification performance and learning speed. The Swish-RELU model has restrictions of overfitting and vanishing gradient issues.

Dixit and Nanda [15] applied an improved Whale Optimization Algorithm (IWOA) in Radial Basis Neural Network (RBNN) to increase the accuracy and convergence speed. The IWOA-RBNN model was applied for pre-processing in input MRI images. Fuzzy C Means (FCM) clustering has been exploited to find tumor regions in image segmentation. The wavelet transforms, entropy, mean, and Principal Component Analysis (PCA) was used for feature extraction for classification. The IWOA technique has the restriction of local trap and RBNN contains an overfitting problem.

## 3. Proposed Method

The BW-CNN model is evaluated in CE-MRI dataset to test the efficiency in BTC. The normalization process is exploited to reduce the difference between the pixel values and improve the image quality. The Bayesian optimization is applied to update the weight values in the CNN and LSTM networks. The LSTM is exploited for feature extraction and the CNN model is exploited for classification. The block diagram of the proposed model is revealed in Fig. 1.

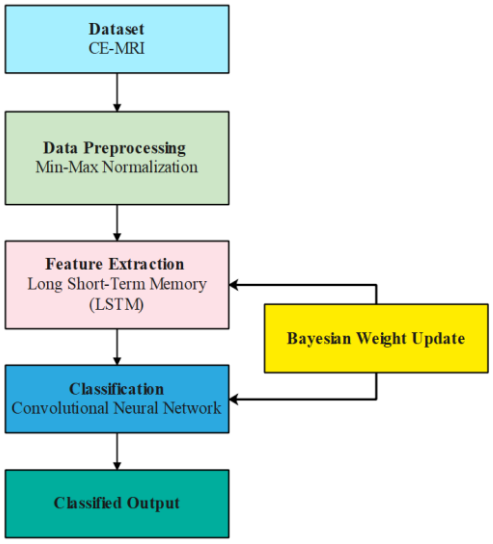


**Fig. 1.** Block diagram for the proposed method

### 3.1. Dataset Description

CE-MRI contains a small number of brain images with a sizable slice gap which are normally available in clinical settings [16]. But there is difficulty while creating a 3D model with limited information. In the CE-MRI dataset, 233 patients from Nanfang Hospital and Tianjing Medical University, China donated 2D slices in the year between 2005 to 2010. Meningiomas (708 slices), gliomas (1426 slices), and pituitary tumors (930 slices) are among the 3064 segments in this dataset's common views (sagittal, coronal, and axial). Moreover, this dataset offers 5-fold cross-validation indices. Using this data, 20% (612 images) of the images are used to evaluate performance, and 80% (2452) of the images are used for training. The pixels in the images are $0.49 \times 0.49\ mm^2$ in area, with an in-plane resolution of $512 \times 512$. The slice gap is $1\ mm$, and the slice width is $6\ mm$. Three qualified radiologists painstakingly marked the tumour border.

### 3.2. Pre-processing

#### 3.2.1.Min-Max Normalization

Once the data collection is done, the collected images are processed to the pre-processing stage for eliminating the noises and unnecessary variance from the data. Additional techniques, like Min-Max Normalization, are needed to get around the limitations of the network paradigm [17]. In the pre-processed stage, the image contains a significant value in the range of [0, 255]. As a result, the Min-Max normalization rule [18] is used to convert the brain images into intensity brain images in the [0, 1] range. Additionally, the Min Max normalization technique measures the standard value in the provided features, which are depicted in (1) and (2) using the minimum and maximum values in the given features.

$$x_{std} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

$$x_{scaled} = x_{std} \times (max - min) + min \quad (2)$$

Here, the feature range is signified as $min$, and $max$. The $x_{scaled}$ is employed as $X$ in the network. Once the pre-processing is done, those data are given as input to the feature extraction stage, those details are explained in the following sub-section.

### 3.3. Feature Extraction

#### 3.3.1.Long Short-Term Memory

The LSTM consists of $x_t$ input data at time $t$, $h_t$ cell output, $c_t$ cell value memory, and $h_{t-1}$ is the previous step cell output [19, 20]. The LSTM unit procedure is specified in (3).

Equation (3) measures $h_t$ LSTM unit output

$$h_t = o_t * tanh\ (c_t) \quad (3)$$

Equation (4) measure $o_t$ output gate using the output gate of $o_t$ state value, $b_0$ bias, and $W_0$ weight matrix.

$$o_t = \sigma(W_0.[h_{t-1}, x_t] + b_0) \quad (4)$$

Equation (5) measures the LSTM state value using $c_{t-1}$ last state value and $c_t$ current moment.

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c_t} \quad (5)$$

Where " × " is the dot product, the input and output gate cell state value is controlled by the input and output gate.

Equation (6) provides forget gate process that controls historical data using $W_f$ weight matrix, $b_f$ bias, $f_t$ forget gate.

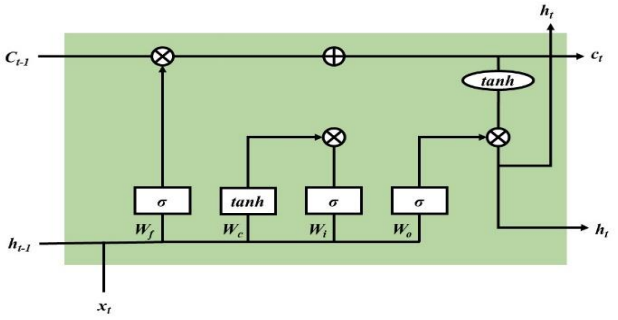


**Fig. 2.** LSTM unit cell for feature extraction

$$f_t = \sigma(W_i.[h_{t-1}, x_t] + b_f) \quad (6)$$

Equation (7) provides the input gate process, input gate updates the current input state value of the memory cell using $i_t$ input gate, $W_i$ weight matrix, $\sigma$ sigmoid function, and $b_i$ bias.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (7)$$

Equation (8) provides a candidate memory cell $\tilde{c_t}$ using $b_c$ bias, and $W_c$ weight matrix.

$$\tilde{c_t} = tanh\ (W_c.[h_{t-1}, x_t] + b_c) \quad (8)$$

CNN model has higher popularity in different domains due to learning in image representation. Kernel convolution is applied to input data in convolution layers of the CNN model to learn from input data. The filter is performed using a non-linear activation function, as in (9).

$$a_{i,j} = f\left(\sum_{m=1}^{M}\sum_{n=1}^{N} w_{m,n}\,.\,x_{i+m,j+n} + b\right) \quad (9)$$

Where $x_{i+m,j+n}$ are upper neurons in the neuron $(i, j)$, $a_{i,j}$ is corresponding activation, $w_{m,n}$ is convolution weight matrix, $f$ is a non-linear function, and $b$ is a bias value.

Rectified Linear Unit (ReLU) is applied in (10) of convolution layers that are applied to measure non-linear functions and feature maps.

$$\sigma(x) = max\ (0, x) \quad (10)$$

The LSTM unit cell for feature extraction is shown in Fig. 2.

### 3.4. Convolutional Neural Network based Classification

Deep neural networks are also referred to as Convolutional Neural Networks (CNNs) which easily categorize and segment the images. The machine that interprets images has many hidden intricacies. CNN always provides better performance in the area of image processing, detection, localization, segmentation, and classification. The model is widely used primarily due to the great CNN efficiency in image categorization. The CNN model uses a convolutional layer with trainable weights and biases that are modeled after animal neurons. More convolution kernels are provided in the model and more hidden features are applied in input samples [21-23]. Two convolution layers are applied in the CNN model, 64 convolutional kernels are used in feature extraction and 1 × 5 convolution kernel is applied in the first convolution layer. The 128 convolution kernels are used for deeper feature extraction and 2 sliding steps are used for the convolution window. Each convolution kernel of 1 × 3 size is applied and 1 step size of convolution window is applied in the layer. Max-pooling layer is used for two convolutional layers for down-sampling. Two processes are used to filter noise interference and the dominant feature is used to reduce features.

#### 3.4.1.Convolutional Layer

The visual cortex's neuronal cells are involved in extracting characteristics from pictures in animal brains. Each neural cell extracts different characteristics that aid in the comprehension of a picture. Convolutional layers are used to simulate neuronal cells, and this allows for the extraction of properties including gradient direction, texture, colours, and edges. In convolutional layers and size, convolutional filters or kernels are learning filters is $n \times m \times d$, where image depth is $d$. During the forward pass, the Kernels are twisted over the input volume's height and width, and Input and filter entries are generated using the dot product. CNN learns which filters to use for texture, colour, edge, etc. An activation function layer is applied using the convolution layer's output.

#### 3.4.2.Activation Function

Since most of the real-world data is non-linear, Utilizing non-linear data transformation's activation functions. This guarantees that input space representation is translated to various output spaces by requirements.

This requires real-value number $x$ and converting it into a range of 0 and 1. Large positive and negative inputs in particular are placed near 0 and unity, respectively. This is expressed in (11).

$$f(x) = \frac{1}{1+e^{-x}} \tag{11}$$

A real value $x$ is taken into account on non-linear functions. A common non-linear function that takes less time to compute and performs better than the ReLU activation function. While $\tan h$ functions are shown in (12).

$$f(x) = max\,(0, x) \tag{12}$$

#### 3.4.3.Pooling Layer

Many features in the system of activation map provide overfitting problems and computational burden. A pooling layer of sub-sampling non-linear is performed to reduce features in the system. Pooling is carried out translation invariance and two pooling methods of average pooling and max-pooling are applied. Average pooling uses average value and max pooling uses the max value in each pooling region.

#### 3.4.4.Output Layer

Fully connected layer and softmax classifier in the output layer presented in the system. The final layer is utilized to extract features from the network. Each node of the fully connected layer with upper layer nodes is merged to extract features. The CNN framework is revealed in Fig. 3.

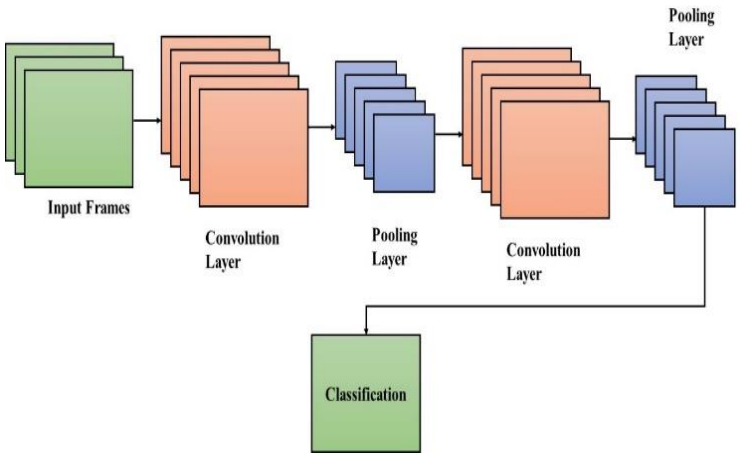


**Fig. 3.** CNN model for the classification

#### 3.4.5.Steps for CNN based Classification

1. In 1st layer, use a convolution filter
2. Smoothing the convolution filter (also known as subsampling) lowers the sharpness of the filter

3. The activation layer regulates the transmission of signals from one layer to another.
4. Shorten the duration of the training time by using rectified linear units
5. All the neurons are linked to one another in the subsequent layer.
6. To deliver feedback during training, a loss layer is included in the end.

### 3.5. Bayesian Weighted Process

Bayesian optimization is used in solving functions and effective technique for computationally expensive identifying extrema [24, 25]. This is used to solve closed-form expressions, and non-convex functions, evaluate hard derivatives, and calculate expensively. The optimization aim is to discover maximum values at the sampling point of unknown function $f$, as shown in (13).

$$X^{+} = \underset{x \in A}{argmax} f(x) \tag{13}$$

Where $x$ search space is denoted as $A$. Bayes' theorem is used to derive Bayesian optimization i.e., the prior probability of $P(M)$ of multiplied $M$ of overserving $E$ of likelihood $P(E|M)$ and given evidence data $E$, as in (14):

$$P(M|E) \propto P(E|M)P(M) \tag{14}$$

Bayesian optimization core idea is given in (14). The Bayesian optimization principle is to combine function $f(x)$ prior distribution with information sample to obtain the posterior of the function and according to criterion, posterior information is used to identify function $f(x)$.

## 4. Results and Discussion

The Bayesian weight is added to the LSTM and CNN model to update the optimal weight in the network. The BW-CNN is evaluated with other classifiers in the CE-MRI dataset. For processing this analysis, the system required Intel i7 processor, 64 GB RAM, 22 GB graphics, and Windows 11 operating system to execute the proposed model. The following are the standard classifiers used for the comparison, they are SVM, RF, KNN, LSTM and CNN. Those classifiers are compared with the proposed BW-CNN which is mentioned in Table 1.

**Table 1.** Comparison with standard classifiers

| Methods | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|
| SVM | 90.04 | 88.9 | 88.88 |
| RF | 90.59 | 92.52 | 89.15 |
| KNN | 91.16 | 93.11 | 92.31 |
| LSTM | 92.01 | 93.42 | 92.73 |
| CNN | 94.35 | 94.3 | 92.91 |
| BW-CNN | 98.83 | 98.72 | 98.81 |

**Table 2.** Deep learning model comparison

| Methods | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|
| AlexNet | 95.6 | 94.7 | 93.2 |
| ResNet | 95.8 | 95.3 | 94.6 |
| VGG-16 | 96.2 | 95.6 | 97.3 |
| BW-CNN | 98.83 | 98.72 | 98.81 |

The BW-CNN model is evaluated in the CE-MRI dataset and related to standard classifiers, as revealed in Fig. 4. The BW-CNN has the advantage of updating the weight in the network using Bayesian optimization which helps to avoid overfitting problems. The optimal weight updates help to increase the sensitivity of the classification. The KNN is sensitive to outliers, RF has overfitting and SVM contains an imbalance data. The BW-CNN model has 98.83 % accuracy, and 98.72 % sensitivity, and the CNN model has 94.35 % accuracy and 94.3 % sensitivity in classification.

The deep learning models such as VGG-16, ResNet, and AlexNet are compared with the BW-CNN model, as shown in Table 2 and Fig. 5. The existing deep learning models have an overfitting issue which decreases the efficiency. The BW-CNN updates the weight values based on Bayesian optimization which supports to rise in the performance of the system. The BW-CNN has 98.83 % accuracy, and 98.72 % sensitivity, and the VGG-16 model has 96.2 % accuracy and 95.6 % sensitivity.

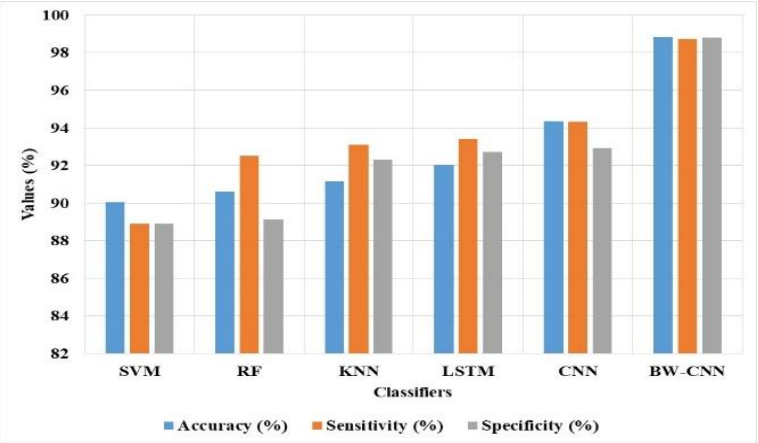


**Fig. 4.** Classifier comparison in CE-MRI dataset

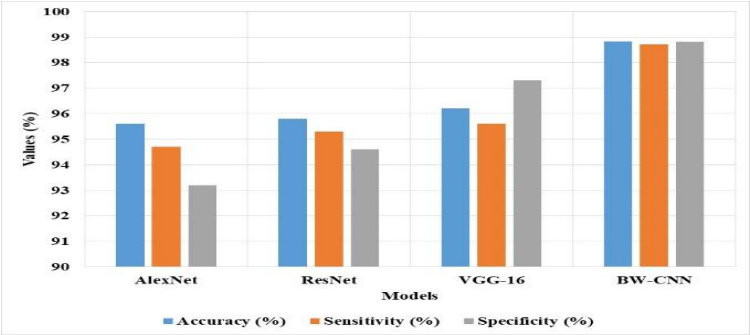


**Fig. 5.** Deep learning model comparison in CE-MRI dataset

### 4.1. Comparative Analysis

The BW-CNN is evaluated with existing models in Brain tumour classification in CE-MRI.

As seen in Table 3 and Fig. 6, the CE-MRI dataset compares the BW-CNN to prior research. When it comes to classification, the BW-CNN is more accurate than current methods. The previous models suffer from overfitting issues that reduce their effectiveness. The BW-CNN comprises the advantage of updating the weights of the network which helps to avoid overfitting problems. The BW-CNN model has 98.83 % accuracy, and the Histogram descriptor [14] has 98.6 % accuracy in the classification.

**Table 3.** BW-CNN comparison with existing technique

| Methods | Accuracy (%) |
|---|---|
| ResNet-50 [11] | 97.48 |
| MANet [12] | 96.51 |
| Xception [13] | 98.04 |
| Histogram descriptor [14] | 98.6 |
| BW-CNN | 98.83 |

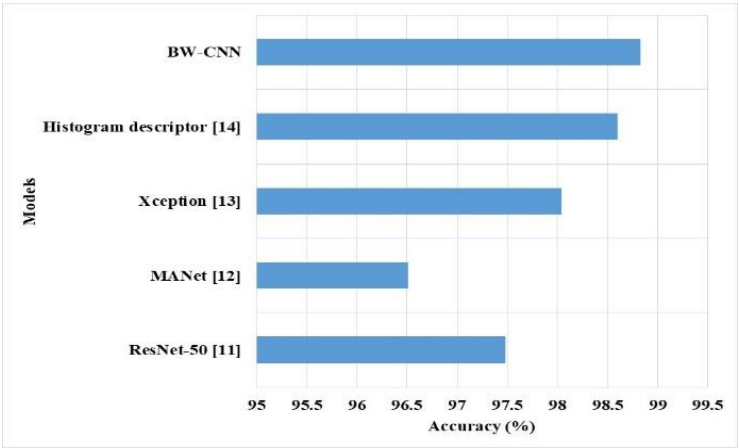


**Fig. 6.** The BW-CNN comparison with the existing technique

## 5. Conclusion

The automatic and efficient BTC technique helps in the diagnosis and early treatment of patients. To prevent overfitting and vanishing gradient issues, this study suggests using the BW-CNN model to modify the weights in the CNN framework. The optimal weights are updated using the Bayesian optimization method. For feature extraction and classification, LSTM and CNN models are employed to improve the accuracy. In the CE-MRI dataset, the BW-CNN model is assessed and contrasted with other classification methods. In terms of classification accuracy and sensitivity, the CNN model has achieved 94.35% accuracy and 94.3% sensitivity, while the proposed BW-CNN model has accomplished 98.83% accuracy and 98.72% sensitivity. The comparative analysis of the results indicates that the proposed BW-CNN model has higher accuracy (98.83%) than the conventional Xception model which has 98.04%. In the future, this research will be further extended by applying the attention layer to select the relevant features.

### Conflicts of Interest

The authors declare no conflict of interest.

### Author Contributions

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization have been done by first author. The supervision, review of work and project

## References


[1] M. I. Sharif, M. A. Khan, M. Alhussein, K. Aurangzeb, and M. Raza, "A decision support system for multimodal brain tumor classification using deep learning," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3007–3020, Aug. 2022. https://doi.org/10.1007/s40747-021-00321-0

[2] S. Kokkalla, J. Kakarla, I. B. Venkateswarlu, and M. Singh, "Three-class brain tumor classification using deep dense inception residual network," *Soft Comput.*, vol. 25, no. 13, pp. 8721–8729, Jul. 2021. https://doi.org/10.1007/s00500-021-05748-8

[3] V. V. S. Sasank, and S. Venkateswarlu, "Brain tumor classification using modified kernel based softplus extreme learning machine," *Multimedia Tools Appl.*, vol. 80, no. 9, pp. 13513–13534, Apr. 2021. https://doi.org/10.1007/s11042-020-10423-5

[4] J. Kang, Z. Ullah, and J. Gwak, "MRI-Based Brain Tumor Classification Using Ensemble of Deep Features and Machine Learning Classifiers," *Sensors*, vol. 21, no. 6, p. 2222, Mar. 2021. https://doi.org/10.3390/s21062222

[5] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, M. A. Pramanik, S. Aryal, M. A. Almoyad, K. F. Hasan, and M. A. Moni, "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning," *Expert Syst. Appl.*, vol. 26, p. 120534, Nov. 2023. https://doi.org/10.1016/j.eswa.2023.120534

[6] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and classification using machine learning: a comprehensive survey," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3161–3183, Aug. 2022. https://doi.org/10.1007/s40747-021-00563-y

[7] I. A. E. Kader, G. Xu, Z. Shuai, S. Saminu, I. Javaid, and I. S. Ahmad, "Differential Deep Convolutional Neural Network Model for Brain Tumor Classification," *Brain Sci.*, vol. 11, no. 3, p. 352, Mar. 2021. https://doi.org/10.3390/brainsci11030352

[8] S. Rinesh, K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Y. A. Waji, "Investigations on brain tumor classification using hybrid machine learning algorithms," *J. Healthcare Eng.*, vol. 2022, p. 2761847, Feb. 2022. https://doi.org/10.1155/2022/2761847

[9] R. Singh, A. Goel, and D. K. Raghuvanshi, "Computer-aided diagnostic network for brain tumor classification employing modulated Gabor filter banks," *Visual Comput.*, vol. 37, no. 8, pp. 2157–2171, Aug. 2021. https://doi.org/10.1007/s00371-020-01977-4

[10] T. Tazin, S. Sarker, P. Gupta, F. I. Ayaz, S. Islam, M. M. Khan, S. Bourouis, S. A. Idris, and H. Alshazly, "A robust and novel approach for brain tumor classification using convolutional neural network," *Comput. Intell. Neurosci.*, vol. 2021, p. 2392395, Dec. 2021. https://doi.org/10.1155/2021/2392395

[11] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-class brain tumor classification using residual network and global average pooling," *Multimedia Tools Appl.*, vol. 80, no. 9, pp. 13429–13438, Apr. 2021. https://doi.org/10.1007/s11042-020-10335-4

[12] N. S. Shaik, and T. K. Cherukuri, "Multi-level attention network: application to brain tumor classification," *Signal, Image Video Process.*, vol. 16, no. 3, pp. 817–824, Apr. 2022. https://doi.org/10.1007/s11760-021-02022-0

[13] J. D. Bodapati, N. S. Shaik, V. Naralasetti, and N. B. Mundukur, "Joint training of two-channel deep neural network for brain tumor classification," *Signal, Image Video Process.*, vol. 15, no. 4, pp. 753–760, Jun. 2021. https://doi.org/10.1007/s11760-020-01793-2

[14] A. M. Alhassan, and W. M. N. W. Zainon, "Brain

tumor classification in magnetic resonance image using hard swish-based RELU activation function-convolutional neural network," *Neural Comput. Appl.*, vol. 33, no. 15, pp. 9075–9087, Aug. 2021. https://doi.org/10.1007/s00521-020-05671-3

[15] A. Dixit, and A. Nanda, "An improved whale optimization algorithm-based radial neural network for multi-grade brain tumor classification," *The Visual Computer: International Journal of Computer Graphics*, vol. 38, no. 11, pp. 3525–3540, Nov. 2022. https://doi.org/10.1007/s00371-021-02176-5

[16] F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A Deep Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network," *Healthcare*, vol. 9, no. 2, p. 153, Feb. 2021. https://doi.org/10.3390/healthcare9020153

[17] H. Mzoughi, I. Njeh, A. Wali, M. B. Slima, A. BenHamida, C. Mhiri, and K. B. Mahfoudhe, "Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification," *J. Digital* Imaging, vol. 33, no. 4, pp. 903–915, Aug. 2020. https://doi.org/10.1007/s10278-020-00347-9

[18] T. Sadad, A. Rehman, A. Munir, T. Saba, U. Tariq, N. Ayesha, and R. Abbasi, "Brain tumor detection and multi-classification using advanced deep learning techniques," Microsc. Res. Tech., vol. 84, no. 6, pp. 1296–1308, Jun. 2021. https://doi.org/10.1002/jemt.23688

[19] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," Inf. Process. Manage., vol. 58, no. 1, p. 102435, Jan. 2021. https://doi.org/10.1016/j.ipm.2020.102435

[20] I. E. Livieris, N. Kiriakidou, S. Stavroyiannis, and P. Pintelas, "An Advanced CNN-LSTM Model for Cryptocurrency Forecasting," Electronics, vol. 10, no. 3, p. 287, Jan. 2021. https://doi.org/10.3390/electronics10030287

[21] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing," ISPRS J. Photogramm. Remote Sens., vol. 173, pp. 24-49, Mar. 2021. https://doi.org/10.1016/j.isprsjprs.2020.12.010.

[22] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, "Classification of Remote Sensing Images Using EfficientNet-B3 CNN Model With Attention," IEEE Access, vol. 9, pp. 14078–14094, Jan. 2021, doi: 10.1109/ACCESS.2021.3051085.

[23] K. He, L. Ji, C. W. D. Wu, and K. F. G. Tso, "Using SARIMA–CNN–LSTM approach to forecast daily tourism demand," Journal of Hospitality and Tourism Management, vol. 49, pp. 25–33, Dec. 2021. https://doi.org/10.1016/j.jhtm.2021.08.022

[24] M. A. Amou, K. Xia, S. Kamhi, and M. Mouhafid, "A Novel MRI Diagnosis Method for Brain Tumor Classification Based on CNN and Bayesian Optimization," Healthcare, vol. 10, no. 3, p. 494, Mar. 2022. https://doi.org/10.3390/healthcare10030494

[25] L. Du, R. Gao, P. N. Suganthan, and D. Z. W. Wang, "Bayesian optimization based dynamic ensemble for time series forecasting," Inf. Sci., vol. 591, pp. 155–175, Apr. 2022. https://doi.org/10.1016/j.ins.2022.01.010

[26] Shanthi, D. N. ., & J, S. . (2021). Machine Learning Architecture in Soft Sensor for Manufacturing Control and Monitoring System Based on Data Classification. Research Journal of Computer Systems and Engineering, 2(2), 01:05. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/24

[27] Ms. Mayuri Ingole. (2015). Modified Low Power Binary to Excess Code Converter. International Journal of New Practices in Management and Engineering, 4(03), 06 - 10. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/38

[28] Timande, S., Dhabliya, D. Designing multi-cloud server for scalable and secure sharing over web (2019) International Journal of Psychosocial Rehabilitation, 23 (5), pp. 835-841.