AN EFFECTIVE IDENTIFICATION AND ANALYSIS FOR
BRAIN TUMOR DIAGNOSIS USING AN EFFICIENT
MACHINE LEARNING TECHNIQUE

REDROUTHU VAMSI KRISHNA¹, K V V SATYANARAYANA²

¹Research Scholar, Department of computer science and Engineering,
Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.
²Professor, Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

Email: 2101050004@kluniversity.in, kopparti@kluniversity.in

ABSTRACT

The surgeon would be assisted in clinical diagnosis by machine learning diagnostic image detection, but the effectiveness of the detection sample depends on a huge amount of labelled data. The most effective and accurate strategy to prevent the epidemic of brain tumor disease was to diagnose it, given the disease's rapid rise. In earlier studies, reducing the input data's dimension resulted in a reduction in data storage space. Therefore, a novel Tiger-based Support Vector Machine (TbSVM) technique was proposed for diagnosing brain tumor disease. Initially, this system performed preprocessing, feature extraction, prediction, and segmentation for diagnosing brain tumors. Here, utilizes the wiener filter to eliminate the noise during the preprocessing stage, and the GLCM feature to extract the features. Finally, the prediction and segmentation process were preceded with the help of a designed model to provide an accurate result for diagnosing the brain tumor. The suggested model was compared against other methods in the results section using several metrics.

Keywords: Tiger Algorithm, Optimization, Feature Extraction, Wiener Filter, Machine Learning.

1. INTRODUCTION

Brain tumors consist of two types: benign and malignant. The patient's condition may deteriorate due to the malignant tumors fast spread to other brain areas. Most damaged cells are eliminated and replaced with new cells when this occurs [1, 2]. It can cause several issues when the old cells are not removed to produce new cells. Brain tumor detection is extremely difficult and intricate because of the tumor's size, location, form, and type. Since the tumor's size and resolution cannot be exactly evaluated in the early stages of the tumors formation, diagnosing a brain tumor can be difficult [3, 4]. Obtaining crucial clinical information on the size, location, and kind of the tumor is the main objective of computerized brain tumor diagnostics. Any impending procedures can be directed and regulated by clinical imaging data to enable precise tumor diagnosis and therapy [5, 6]. Early tumor detection is essential for better treatment options. A number of brain tumor techniques, such as magnetic resonance imaging (MRI), single-photon emission computed tomography (SPECT), positron emission tomography (PET), computed tomography (CT), and magnetic resonance spectroscopy (MRS), can be used to assess the size, shape, and type of a brain tumor [7]. Using radio frequency signals and a powerful magnetic field, MRI produces images of human tissues that reveal details about the anatomy of the tissue [8].

Segmentation provides the foundation for classifying brain tumors, however feature extraction and classification issues are less significant [9, 10]. As a result, the researchers use machine learning (ML) approaches to concentrate on the classification process. Machine learning techniques have recently enhanced the effectiveness of computer-aided medical diagnostics for brain tumor diagnosis [11, 12]. ML methods are essential in the treatment of acute illnesses including kidney failure and the identification of lung cancer. In the detection and
management of gliomas, image segmentation is essential. By classifying each pixel or voxel of the input image into the pre-set tumor region, the segmentation technique uses the image to separate the tumor area from the surrounding normal tissues [13, 14]. The segmentation result for the brain tumor diagnosis is then established. An appropriate segmentation approach provides quantitative and qualitative data on the brain tumor, which is used to develop effective treatments for the patients [15, 16]. Multiple algorithms and strategies are needed for both manual and completely automated tumor segmentation in MRI images [17].

According to the pixels with similar characteristics, segmentation divides the image into pieces. However, manually separating the abnormal from the normal brain tissues takes a lot of time and doesn't yield reliable findings [18, 19]. Create a computer-aided method to automate the diagnosis procedure with the aid of ML techniques to assist the doctor in segmenting the brain tumor location from MRI data [20]. In this study, a novel Tiger-based Support Vector Machine (TbSVM) method is proposed to address the challenge of medical image processing for the accurate segmentation of brain tumor diagnosis.

The following are some of the work's key ideas:

- First, train the system with the brain MRI dataset, which includes both normal and tumor affected images.
- Meanwhile, a novel TbSVM was developed with the required feature analysis and segmenting the brain tumor process.
- Execute the preprocessing step; the noise from the input image has been removed using a wiener filter.
- Additionally, extracting features for the aim of detecting the brain tumor utilising the suggested model's tiger fitness features.
- Moreover, apply suggested TbSVM model to the prediction and segmentation processes.
- Lastly, validate the performance’s accuracy, recall, precision, and F-score.

Section 2 of current study expands on recent related brain tumor segmentation work. Also, section 3 detailed the system model and problem statement. Moreover, elaborate on the proposed solution to the discussed problem in Section 4. The gained outcome of the developed framework is detailed in Section 5, and the research arguments ended in Section 6 conclusion.

2. RELATED WORK

Some of the recent literature related to brain tumor diagnosis has been described as follows:

Amin et al. [21] proposed the Fused Feature Vector Method to diagnose a brain tumor. In this method, the input data was enhanced by the Weiner filter’s removal of noise using several wavelet bands. Then, using Potential field (PF) clustering, isolate subgroups of tumor pixels. Then, in the T2 MRI and Fluid Attenuated Inversion Recovery, isolate the tumor location using a range of global thresholds and mathematical techniques. Finally, the Gabor Wavelet Transform (GWT) and Local Binary Pattern (LBP) efficiently discriminated various forms of brain cancers. However, segmenting the brain tumor requires additional time.

Manogaran et al. [22] presented an improved orthogonal gamma distribution-based machine learning system for analysing the over and under-segmented brain tumor regions to detect the faulty function in the Region of Interest (ROI). This technique was used to calculate the tissues edge enhancement and organise the matching done by the orthogonal gamma distribution. Then, estimate the variance for the edge detection from the enhanced image. Due to high noise and cluster sensitivity issues, it was difficult to segment the brain tumor.

Brunese et al. [23] discussed the Ensemble Learning technique was utilised to locate brain tumors having radiomic features. The ensemble learner is dedicated to insight into the brain tumor grades by noninvasive radiomic features. This considered feature belongs to five groups. Finally, once the effectiveness of the features had been evaluated by hypothesis testing and decision boundaries, the best classifier for the ensemble learner to segment the brain tumor was selected. This method requires increasing the diagnosis accuracy of brain tumors.

Rehman et al. [24] suggested Brain tumor Localization technique was used to recognize brain tumors from MR images. Initially, apply the bilateral filtering to remove the noise and create the text on the map using the Gabor filter. The low-level features were then recovered from the images after they had been split from the map's
text. The text's histogram level was then determined for each super pixel level on the map. Finally, three classes—background, tumor, and non-tumor region were predicted by the classifier. The image quality in image processing, however, was inadequate for identifying a brain tumor.

Kang et al. [25] proposed a group of machine learning classifiers and deep features for the classification of brain tumors. The original method consists of three steps: Pre-trained CNN models were used to extract the deep features in order to retrieve meaningful information. Then, the features were selected by fined-tuned ML models and combined with the ensemble model for obtaining the brain tumor classification from brain MRI. However, it occurred due to the time complexity involved in segmenting the brain tumor result.

3. SYSTEM MODEL AND PROBLEM STATEMENT

Consequently, initialize the data in the training phase. Preprocessing, feature extraction, segmentation, and classification were the four steps in the segmentation process for brain tumors. In pre-processing stage, there would be noise in the data then removed from the data by using the filter. The feature was then extracted using the data that had been subjected to noise removal as the input. Then, segmentation and classification were determined using the optimized process. Brain tumor segmentation was needed to predict and segment the brain tumor accurately.

The analysis of brain tumors is impacted by artefacts including intensity inhomogeneity and partial volume effect in MR images. Then, define the system model in the segmentation of brain tumor application in fig.1.

4. PROPOSED TbSVM FOR BRAIN TUMOR DIAGNOSIS

This study performed a procedure of diagnosing brain tumor condition. The development of a novel Tiger-based Support Vector Machine (TbSVM) method to distinguish the brain tumor's location from the MR images was a key objective of this diagnosis approach. This process involves four stages: preprocessing, feature extraction, prediction, and segmentation. Initially, the unnecessary noise was removed from the input image by the wiener filter. Next, the feature used in the extraction process was GLCM features, with the help of the tiger fitness feature; extract the features for prediction and segmentation. The tumor and normal region could be segmented through the proposed TbSVM model.

Fig.2 The basic diagram of Proposed TbSVM

The dataset used in the suggested strategy for the diagnosis of the brain tumor disease was the Brain Tumor MRI Dataset. This dataset includes the whole segmentation of brain tumor disease using MRI data. Fig. 2 defines the novel framework that is being presented.

4.1 Preprocessing

Brain MRI data was determined as $S$ the dataset training process was described $F(S)$ and $\{S_1, S_2, S_3, \ldots, S_n\}$ denoted as the trained $n$
amount of data. The data initialization process is exposed in Eqn. (1).

\[
F(S) = \{S_1, S_2, S_3, \ldots, S_n\}
\]  
(1)

Preprocessing was the first stage of the segmentation procedure for brain tumors. In this case, the wiener filter would inverse the blurring while also removing the noise. Use this filter as part of the preprocessing to eliminate unwanted noise features. The preprocessing function is indicated in Eqn. (2)

\[
P = \mu + \frac{\sigma^2 - \sigma^2}{\sigma^2}[g(s,t) - \mu]
\]  
(2)

Where, \( P \) represented as the preprocessing variable, \( \sigma \) and \( \mu \) denoted as the mean and variance for the input images, \( g(s,t) \) represented as the pixel value of the image. Then, use the wiener filter to get higher efficiency for removing the impulse noise in the preprocessing stage.

4.2 Feature Extraction

The necessary step in brain tumor segmentation was feature extraction. In this research, the input for this stage was the noise removal data that aimed to extract the feature. Here, the GLCM feature was extracted for the feature extraction process using the fitness features of the Tiger algorithm [29]. In tiger algorithm, the Tiger (input images) sees its food source (GLCM feature), and it gets its food source with the help of its claw (fitness function). The input image and food source comparison are based on estimating the fitness value in Eqn. (3).

\[
D = D_0 e^{-\eta x}
\]  
(4)

Where \( D_0 \) denoted as the coefficient of the direct distance between a food source and the input image. To avoid zero division, use a formula with Gaussian approximation. Then, calculate the fitness function between the tigers \( i \) and \( j \) \( X_i, X_j \) and find the best solution in Eqn. (5).

\[
FF = |X_i - X_j| = \sqrt{\sum_{k=1}^{d_n}(X_{i,k} - X_{j,k})^2}
\]  
(5)

The feature extraction process was earned by the proposed model, which was detailed in Eqn. (6).

\[
E = P + \phi_0 \text{Feature}(S_n)
\]  
(6)

Where \( P \) denoted as the Preprocessing variable \( \text{feature}(S_n) \) represented the number of features in the dataset, determine the monitoring process variable was \( \phi_0 \), and the feature extraction parameter was defined \( E \).

4.3 Prediction and Segmentation

After completing the feature extraction process, the prediction and segmentation process was performed. The Support Vector Machine (SVM) classifier should then be provided with the retrieved characteristics to determine whether the MR images are normal or tumor. If brain tumor was present, segment the tumor region using the TbSVM model. Hence, the disease-affected region had been determined by Eqn. (7).

\[
Pr = E + \sum_{i=1}^{n}(A_i, Nb_i)
\]  
(7)
Here, the extracted feature was defined as $E$, $N_b$, the normal MRI brain images, $A_i$ denoted as the abnormal brain MRI image, and $Pr$ was the prediction process. Here, execute the prediction process in the image region to recognize whether it had a brain tumor in Eqn. (8).

$$\text{Segmentation} = \begin{cases} \Pr \leq 0 & \text{Tumor region} \\ \Pr \geq 1 & \text{normal region} \end{cases}$$

(8)

Hence, if the predicted region was less than equal to 0, then segment the tumor region. If the prediction region was greater than equal to 1, then the region was not affected by the tumor, which was normal.

**Algorithm 1. TbSVM**

**Input:** MR Brain images

**Output:** Brain Tumor or Normal

**Start**

1. $F(S) = \{S_1, S_2, S_3, ..., S_n\}$
   // Initialization of dataset

2. Preprocessing() 
   
   - $\mu, \sigma, g, s, t$
     // Initializing Preprocessing variable
   - $P \rightarrow \mu + \sigma^2 (g - \mu)$
     // Noise features had been filtered

3. Feature Extraction() 
   
   - $E, P, \phi_0, S_n, D(x), D_i$
     // Feature analysis variables were initialized
   - $E = \text{extract}(\phi_0(S_n))$
     // Features were extracted

4. Prediction and Segmentation() 
   
   - $Pr, E, A_i, N_b$
     // Prediction parameters were initialized
     predict $\rightarrow$ disease features were predicted using eqn. (7)

   ```
   Segmentation()
   {
   if ($Pr \leq 0$)
   {
   Segment Tumor Region
   // if the condition was satisfied, then perform the segmentation process
   }else (Normal region)
   }
   }
   
   Stop
   ```

**Start**

1. Dataset Training TbSVM

2. // Designing proposed framework

3. ThSVM

4. Preprocessing

5. Feature Extraction using GLCM features

6. Prediction phase on TbSVM

7. Tracing affected region

8. // The brain tumor was predicted

9. Segmentation

10. // Exact brain tumor disease affected region segmentation

11. // Evaluating the segmentation condition

12. if ($Pr \leq 0$)

13. Segment Tumor region

14. No

15. Normal

16. Yes

**Fig. 3 Flow of TbSVM**
The detailed steps and processes are presented in the designed model in algorithm one and fig. 3. Based on these step processes, execute the Python code, and verify the results. The algorithm incorporated all mathematical function parameters in the pseudocode format.

5. RESULTS AND DISCUSSION

The Python tool tested the proposed TbSVM technique on the Windows 10 platform. The dataset considered for this research was Brain Tumor MRI Dataset, which includes normal and affected region images. Also, consider that this present study's training and the testing percentage is 80% of training and 20% of testing.

Table 1. Parameter Specification

<table>
<thead>
<tr>
<th>Operating system</th>
<th>Windows 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming platform</td>
<td>Python</td>
</tr>
<tr>
<td>Version</td>
<td>3.7.6</td>
</tr>
<tr>
<td>Database</td>
<td>Brain Tumor MRI Dataset</td>
</tr>
</tbody>
</table>

5.1 Comparative Analysis

The proposed TbSVM model considered various performance metrics with accuracy, precision, recall, and F-score, providing accurate results for brain tumor diagnosis. Then, examine the suggested model with several models already being used such as Hybrid Fuzzy K Mean-Self Organization Mapping (FKM-SOM) [26], Convolutional Neural Network- K-Nearest Neighbor (CNN-KNN) [27], Contrast limited
fuzzy adaptive histogram equalization (CLFAHE) technique [28].

Table 2. Testing Process Outcome

<table>
<thead>
<tr>
<th>Input images</th>
<th>Pre-processed images</th>
<th>Segmented images</th>
</tr>
</thead>
</table>

5.1.1 Accuracy

In the proposed model, accuracy (A) was determined as the rate of correctly classified cases, i.e., predicting the presence of brain tumor correctly, as determined in Eqn. (9). Thus, the accuracy would be high when providing the correct detection.

\[ A = \frac{TN + TP}{TP + FP + FN} \]  

\[
(9)
\]

Fig.4 Accuracy Of Various Segmentation Methods For Brain Tumor

The accuracy of the framework FKM-SOM was scored at 90%, record the CNN-KNN model at 96.25%, and CLFAHE attained 92.15%. The presented scheme gained the finest outcome, 99.074%. Here, the new TbSVM strategy has achieved a greater accuracy rate than other approaches shown in fig. 4, demonstrating its efficiency.

5.1.2 Precision

In addition, the precision (PREC) was accurately determined to measure the value of positive scores in prediction or segmentation outcomes. Also, this performance metric focused on the positive class rather than the negative class. Hence, represent the estimation of the precision metrics in Eqn. (10).

\[ PREC = \frac{TP}{TP + FP} \]  

\[
(10)
\]

Fig.5 Precision Of Various Segmentation Methods For Brain Tumor

The precision of different existing models, such as FKM-SOM, had gained 97.17%, CNN-KNN had attained 96.67%, and CLFAHE scored 98.14%. In this case, the suggested TbSVM model gained 99.076%, providing a better precision value than the other available mechanisms and depicting the precision comparison in fig. 5.

5.1.3 Recall

The recall (R) metrics were calculated based on the measure of stability range in the false segmentation case. Moreover, the formulation of recall was illustrated by Eqn. (11).
The recall of the model FKM-SOM scored 76%, the CNN-KNN mechanism recorded 95.83%, and the CLFAHE model reported 77.94%. Hence, compared to the previous studies, the developed model scored the highest recall value at 99.074%. Hence, describe the statistics of recall in fig.6

5.1.4 F-score

The harmonic average between precision and recall was used to define the F-score. It could be calculated from the values from the dataset collection. The computation of the F-score was determined in Eqn. (12).

\[
F - score = 2 \frac{PREC \times R}{R + PREC}
\]

(12)

Figure 7 depicts a comparison of the F-score. The F-score obtained with the current methods, such as FKM-SOM, CNN-KNN, and CLFAHE, is 86%, 96.25%, and 86.88%. The F-score of the proposed model gained 99.074%, respectively.

Table 3 Performance Analysis Of Different Techniques Under Several Metrics

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FKM-SOM[26]</td>
<td>90</td>
<td>97.17</td>
<td>76</td>
<td>86</td>
</tr>
<tr>
<td>CNN-KNN[27]</td>
<td>96.25</td>
<td>96.67</td>
<td>95.83</td>
<td>96.25</td>
</tr>
<tr>
<td>CLFAHE[28]</td>
<td>92.15</td>
<td>98.14</td>
<td>77.94</td>
<td>86.88</td>
</tr>
<tr>
<td>Proposed TbSV M</td>
<td>99.074</td>
<td>99.076</td>
<td>99.074</td>
<td>99.074</td>
</tr>
</tbody>
</table>

5.2 Discussion

Researchers looked at the potential of the ML approach to help doctors diagnose brain tumors, which requires MR scans. The presented model recorded the finest segmentation outcome from the performance assessment, which was relatively better than the previous studies. Hence, this performance improvement was verified by performing the comparative studies in the last section.
In this study, feature analysis and noise removal in the preprocessing step were used to predict and segment the brain tumor disease. Finally, the extraction of features that distinguish classes and reduce the diagnostic efficiency of brain tumors was performed from the strict perspective of the ML system. Table 4 displays the overall performance of the suggested model.

6. CONCLUSION

This research aims to evaluate a novel Tiger-based Support Vector Machine (TbSVM) for brain tumor diagnosis using trained data from the Brain Tumor MRI Dataset. Initially, the noisy data had to be removed by the wiener filter and perform the feature extraction process, and the prediction process was done to segment the brain tumor region. The efficacy and practicality of the recommended TbSVM technique were assessed using performance metrics like accuracy, F-score, recall, and precision. In that, the proposed model had recorded the finest accuracy of 99.074%, compared to other models; it had enhanced the segmentation accuracy by 3%. Furthermore, the proposed model's precision, recall, and F-score were 99.076%, 99.074%, and 99.074%. Using MR images, this proposed model precisely segmented the brain tumor illness. In future, creating a hybrid machine learning model will allow for a flexible result.

REFERENCE


