Classification and Detection of Brain Tumors by Aquila Optimizer Hybrid Deep Learning Based Latent Features with Extreme Learner

Amishi Agrawal*

1Greenwood High International School, Bangalore, 560087, India

Abstract. Brain cancer is a potentially fatal illness that affects the brain. To preserve lives, early tumour detection is now crucial. Imaging in medicine is one method for diagnosing brain tumours. To help find brain tumours, researchers are turning to deep learning. Error in individual early diagnosis of the condition has been demonstrated to be reduced using deep learning algorithms. In the case of brain tumours, even a slight misdiagnosis might have serious consequences. When it comes to processing medical images, spotting brain tumours is still a difficult task. It's difficult to spot the tumour on a brain scan. The precision of the image is impacted by many disturbances and a delay. We used deep learning methods to describe brain disorders in our paper. Brain disease detection utilizing deep learning methods is related to the study of new information. Proposed TL-based DenseNet121 model achieved accuracy, sensitivity, specificity, F1Score, precision, and IoU of 98.38, 97.33, 99.1, 98.23, 98.62, and 96.62 respectively. The results obtained on the brain tumor data set demonstrate that the proposed method outperforms others in terms of F1-score, Precision, Sensitivity, Accuracy, Specificity, and IoU.

1 Introduction

The advancements in both computer vision and machine learning have paved the road for new types of research and the creation of revolutionary algorithms. Several fields, including autonomous vehicles, healthcare, education, and the Internet of Things, have benefited greatly from it (Internet of Things). Anomaly detection, a biomedical aspect of artificial intelligence and machine learning is particularly captivating researchers at the moment. The increasing prevalence and mortality rate associated with brain tumours [1–5] make it one of the world's worst diseases. According to [31], this is the second largest cause of cancer in the country of India. Around 24,000 individuals have been diagnosed with a brain tumour in the year 2020, and an expected 19,000 have lost their lives to the disease, based on the American Cancer Society’s recently released report, "Cancer Statistics 2020" [32]. The rising use of electronic devices like smartphones and tablets has also made this a problem among young people.

* Corresponding author: amishiagrawal2710@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).
Because of the brain's complexity, detecting any of the roughly 120 different forms of tumours that have been identified so far can be a challenge. Positron emission tomography (PET), Computed tomography (CT), magnetic resonance imaging (MRI), and magnetoencephalography (MEG) among others, are only few of the medical imaging modalities used to diagnose brain irregularities. Due to its ability to distinguish between architecture and tissue mostly on grounds of contrast levels, MRI multimodality imaging is the most common and effective approach routinely utilised in the diagnosis of brain tumour [27]. Clinicians presently spend considerable time manually detecting anomalies and segmenting tumours for therapy and surgical purposes using magnetic resonance imaging (MRI). This method is also dangerous because it is done manually and mistakes are possible. Because of these obstacles, researchers have turned their attention to machine learning (ML) and deep learning (DL) strategies for cancer diagnosis in computers. To create a fully automatic, semiautomatic, or hybrid model which can identify and segment the tumour with little time and highest accuracy, deep learning, a branch of ML, has been increasingly applied in recent years.

1.1 The requirement for Detection

In recent years, automatic defect detection in medical imaging has developed as a potentially useful field for a range of diagnostic applications. Tumor detection and monitoring using MRI is critical because it provides information about aberrant tissues that can be used to inform therapy decisions. Detecting brain cancers via MRI is difficult because there are so many different kinds of tumours and ways, they can manifest themselves. In today's digital world, we can now collect, organize, and analyse medical images digitally [23]. Modern medical imaging technology does not solve the time and accuracy issues that arise from accurate interpretation of medical pictures. The challenge is greatest in regions with anomalous colour and shape, which radiologists must identify for scientific study. Image analysis techniques necessary for detecting and diagnosing brain tumours can benefit from the application of Deep learning technologies [7]. In view of the present knowledge state and the high need for Deep Learning [6-10], this study proposes data training to deal with the complexity involved in recognizing brain tumours while proposing solutions.

1.1.1 Brain Tumor

As previously discussed, [11,12,18,24], a brain tumour is a clumpy growth of living and dead brain cells. The prognosis for people with brain tumours varies according on the type and stage of the disease. According to their point of origin, brain tumours can be classified as either primary or secondary. Tumors that grow elsewhere within the body and subsequently spread to the brain are called secondary tumours [19,20]. Primary brain tumours are those that start in the brain itself.
Brain tumours can be either benign or malignant, depending on the degree of malignancy [22]. In contrast to malignant tumours, which can be fatal due to their rapid development and atypical shape, benign tumours grow slowly [25], look normal, and have clear borders. Malignant tumours are classified into four stages by the World Health Organization (WHO) based on their chemical and physical characteristics [17].

1.2 The Deep Learning Approach

The field of artificial intelligence and machine learning known as deep learning algorithms (DLAs) is built on the premise that they can learn just like people do by modelling human learning processes [28]. Deep Learning and Convolutional Neural Networks are given prime consideration while studying the evolution of brain MRI and related computer interventions. Traditional neural networks have matured into Deep Neural Network (DNN) methods. This outstanding performance and pinpoint accuracy can be attributed to the data-driven interconnections in such networks and the autonomously developed approach. Actually, this is a deep learning algorithm made up of different neural network–based algorithms that automatically recognise aspects and qualities in incoming data, then use that information to create interventions [14,21,22,28].

Convolutional Neural Network: A convolutional neural network (CNN) is a DL method that needs and prioritises several components in an image to enable differentiation between them. Pre-processing time for a ConvNet is much lower compared to that of alternative classification methods. While basic filtering techniques do the job, a well-trained ConvNet can pick up on useful additions to the original algorithm. Each feature in a convolution layer is represented by a neuron, and each neuron's output is determined by the surrounding pixels. The perceptron [[[26]]] is an area of dense pixels that can affect a neuron's output. The more convolution operations incorporated into a neuron's design, the more efficient it will be as a computer. The CNN is built with a hierarchical structure of layers. The components of a CNN include the input, the hidden units, the convolutional features, and so on. It also has convolution layers and batch normalising built in. Each CNN has its own unique architecture depending on the number of layers it has, its size, and the activation techniques it employs. The parameters of CNNs are derived from and backed by empirical evidence. The Receptive Field is the region of visual space where neurons respond to stimuli.

VGG 16: For a 16-layer CNN model, you can use VGG 16. It is still thought of as one of the most successful models used today. In place of a large number of parameters, the VGG 16 model design prioritises ConvNet layers with a 3 x 3 kernel size. The model's values can be freely downloaded and used in one's own applications and systems which is a major reason for its popularity and practicality. Its ease of use is praised in contrast to those of other advanced comprehensives. The smallest 3x3-channel input image size this model can handle is 224x224. Optimization algorithms are utilized in neural networks to determine whether or not a neuron needs to be activated based on the total amount of input. When the output neuron
is made non-linear, the necessity for a kernel function becomes apparent. The neurons in a neural network are only as effective as the weight, bias, and training technique that surrounds them. The output error is used to fine-tune the connection weights of the neurons. Artificial neural networks are able to learn and perform complex tasks thanks to the non-linearity introduced by the input layer and the activation function.

Ensemble Model: Ensemble modelling is a cutting-edge answer to many machine-learning problems since it combines the predictive abilities of several different models into a single, more accurate one. Ensemble modelling combines the use of multiple models that make use of various training data sets or modelling approaches to produce a single predicted performance for the unlabeled data. As long as the base models are differentiated and autonomous, the prediction error lowers when using several models. To do this, Deep Learning Synthesis takes the best parts of various models and combines them into one overall projected feature.

2 Related Work

The most cutting-edge techniques for detecting brain tumours are discussed.

2.1 Machine Learning Techniques

Cognitive health evaluation, cancer identification (breast, cervical, and others), tumour detection, and other medical diagnoses all make use of machine learning in their respective approaches (Mohiyuddin et al., 2022; Javed et al., 2021). Rehman et al. (2020) employ the machine learning methods SVM, RF, Andrus Boost, and AdaBoost1 to pinpoint the location of the tumour on FLAIR scans of the brain. The BraTs 2012 dataset is used to test these methods on both syntactic and natural image categories; the top results show an accuracy of 0.99%, sensitivity of 0.92, specificity of 0.96, precision of 0.88, and a dice score of 0.88. Kumar et al. (2021) offers an automatic brain tumour classification system that uses the K-nearest neighbour technique to categorise MRI scans as normal or abnormal. Tumor areas are separated using fuzz C-means clustering. Experiments are conducted using BRATS and MICCAI datasets, with 96.5% accuracy, 100% sensitivity, and 93% specificity being achieved. Zhou et al. (2020) describe the manual model optimization by a ML expert and compare it to the automatic Tree-Based Pipeline optimizing tool as a means of assessing the performance of the model. Using MRI scans from 288 patients, the suggested model achieved an AUC of 0.94% and an accuracy of 0.88%.

2.1.1 Deep Learning Techniques

Despite being used in a variety of industries, deep learning models have yet to be modified before being applied to sensitive industries like diagnostic imaging. The novel model is suggested by authors in Amin et al. (2022) and uses quantum variational classifiers and ensemble transfer learning to detect brain cancer (QVR). The inceptionv3 model extracts the detailed features, and the softmax score vector is utilised with the QVR to distinguish between pituitary tumours, no tumours, meningiomas, and gliomas. The suggested model successfully detected over 90% of the samples in three separate datasets, including 2020-BRATS, local pictures, and Kaggle. Qureshi et al (2022) proposed an automated Ultra-Light Brain Tumor Detection (ULBTD) system, textural characteristics that were taken from the Gray Level Co-occurrence Matrix were combined with the new Ultra-Light Learning Architecture for in-depth features. To use a support vector machine, it developed the Hybrid
Feature Space to find the brain tumour.

The suggested method has been put into use on a T1-weighted MRI data and has a 99.23% average detection rate, with an F1 measure of 0.99%. Alsaif et al. (2022) include a thorough analysis of CNN architectures, including the features of different modeling techniques including ResNet, AlexNet, and VGG. The VGG framework produced a high value with a 0.93% F1-score, 0.93% accuracy, 0.93% recall and 0.94% precision, when implemented to the MRI dataset to detect the brain tumour. Zailan et al. (2022) identifies the most suitable DL algorithm for small datasets. The outcomes of the performance evaluation are based on factors such as the accuracy, confusion matrix, recall, and precision, among others. Because its recall value is 86.00%, the classification performance of the MobileNet-V2 is usually higher than those of those certain models. Inception-V3 had the second largest accuracy, 84.00%, and VGG-16 had the lowest accuracy, 79.00%. Using Convolutional Neural Networks and Machine Learning classification, Rathod & Khan (2021) suggest automatic brain tumour diagnosis. Small kernels handle the in-depth architecture design. The neuron is reported to have a negligible mass. It has been shown that CNN has the most accuracy with the least amount of complexity.

In order to better diagnose and categorise brain tumours through MRI scans, Nazir et al. (2020) give a comprehensive study of the existing research and findings. Specialists in deep learning who want to apply their skills to the classification and detection of brain tumours will find this analysis especially useful. Satpute et al. (2020) suggests a CNN-based strategy to segment MRI images of brain malignancies. In order to diagnose a brain tumour, a number of classification and prediction methods already exist. Presents a comprehensive analysis of the strengths and weaknesses of current methods for diagnosing brain tumours. A convolutional neural network (CNN) based classifier is proposed as a solution to these restrictions. The best result is obtained by comparing the training and testing data using a CNN-based classifier. Kabir et al. (2020) introduced a method for brain tumour identification and classification, which can accurately determine whether an MRI picture has a tumour or not (for 217 BRATS images), and whether the tumour is benign or malignant for a dataset constructed from 3D BRATS images. Image pre-processing, image segmentation, image enhancement with multi-valued threshold and the feature extraction, feature selection, Chan-Vese algorithm, and classification are all parts of the suggested technique. Both the original BRATS database and the enhanced version of the same dataset are used to evaluate the suggested algorithm. The experimental analysis achieves an accuracy of over 98%, which surpasses that of a large number of currently available methods. To this end, Swapnil & Girish (2020) give a comprehensive evaluation of the Deep Learning-based research reported between 2015 and 2020 that focuses on the classification and detection of brain tumour MRI images into tumour and non-tumor classes. While many useful and effective algorithms have been devised to far, they all fall short in some way because there is a lack of industry-wide standardization.

3 The Proposed Method

3.1 Hybrid Net-Based Classification

This section describes the concept of ResNet and Inception modules used in this work.

3.1.1 Aquila Optimizer (AO) model

In recent times, all applications are switching to an optimization method due to its problem-solving behaviour. The optimization method is used to choose the best solution among entire
solutions. A meta-heuristic model is a tool to provide decision-making when an accurate solution cannot be processed. In the meta-heuristic model, there are many real-time issues that are solved without any complexity. Therefore, an Aquila Optimizer-based meta-heuristic model is the recent method that can be able to solve an issue effectively. The AO model is based on the hunting behaviour of an Aquila that is given in the following section.

3.1.2. ResNet Model

The ResNet-18 is the mostly used DNN model that belongs to the attention mechanism and residual structure. The ResNet-18 structure is presented in the Fig. 2.

The ResNet model is based on DL that consists of a squeeze-and-excitation Basic Block (SEBB) module. It consists of 22 layers such as CV1, SEBB, CV2_x, CV3_x, CV4_x, CV5_x, a global average Pooling layer, and a Fully Connected layer. The CV1 comprises a Convolution layer, a ReLU activation function, Batch Normalization layer, and a maximum pooling layer. The Convolution layer has a configuration of 7x7 kernel size, Padding-3, and stride-2. Then the maximum pooling layer has its configuration of 3x3 kernel size, padding-1 and stride-2. On involving the maximum pooling layer, the parameters and dimensions are minimized. This method has to increase the receptive fields and reserve significant feature information. The SEBB module is attached next to the ResNet-18 which has a residual basic block and a Squeeze-and-excitation (SE) module. The next part of ResNet-18 has two Convolution layers. The SE module is combined into a residual block to constitute a SEBB module. The SEBB module consists of two Convolution layers with the stride-1 and 3x3 kernel.

Fig. 2. ResNet18 Model

In this model, the CV-1 layer is added next to the ReLU and Batch Normalization layer and then the CV-2 layer is just followed by a Batch Normalization layer. The SE module contains two-steps such as squeeze and excitation. The squeeze step consists of a global average pooling layer which transfers a feature map into a vector. Excitation process comprises two Fully Connected layers, ReLU activation, and Sigmoid activation function. The Fully Connected layer has a 1x1xC input and also a 1x1xCx1/r output. In this condition, r represents a scaling parameter that is used to minimize the number of channels with calculation of reduction. The next Fully Connected layer has an input of 1x1xCx1/r, and also an output of 1x1xC. In this method, after getting the 1x1xC vector, the feature map is initialized and the 1x1xC vector is scaled. The real feature map size is equated as WxHxC.
where each channel output weight of the SE module is multiplied by a 2D matrix. This method is used to perform the corresponding feature map of channel to attain a final solution. The CV-3 to CV-6 is connected with a global average pooling layer. This layer is also known as the Adaptive AvgPool function that fitted an output into 1x1 kernel size. At last, the fully connected layer provides a ResNet classification output which is to fit as 7. Also, the corresponding data can be learned and classified with various dataset types.

### 3.1.3. Inception Model

Inception-v3 is a 48-layer deep pre-trained convolutional neural network model. It is a version of the network that has already been trained on over a million images from the ImageNet database. Inception-v3 is made up of two parts: feature extraction with a convolutional neural network and classification with fully-connected and softmax layers. It also includes RMSProp Optimizer, Factorized 7x7 convolutions, Batch Normalizations and Label Smoothing. Figure 4 shows the architecture of Inception-v3.

![Fig. 3. Inception Module](image)

![Fig. 4. Inception-v3 Model](image)

### 3.1.4. Hybrid Classification
The classification is done by combining a DL-based ResNet model and Inception Net model for higher prediction accuracy. The ResNet model is unique and is based on a deep CNN architecture with 152 layers. The DL algorithms have a major demerit of gradient vanishing which is completely eliminated by the skip connections concept. Also, the Inception net model is an advance of GoogleNet DL model. The Inception net model can be used to overcome all the issues of GoogleNet model. In the Inception net model architecture, it has a network layer, pooling layer, and convolution layer. Every layer is performed sequentially in a classical approach. In the Inception Net method, every layer is processed in parallel to minimize the number of parameters as well as functions. Because of the parallel function, the processing cost and memory are decreased.

Moreover, in comparison with the other methods, the InceptionNet and ResNet models combination is matched well and also provides major benefits. This hybrid model has assured the capability to scale up to thousand layers. The residual networks can provide several numbers of residual blocks to determine the mapping. There are numerous convolutional networks provided in the Inception network model. The convolution function and conversion of input data using a Rectified Linear Unit (ReLU) operation are processed in convolutional layers. The hybrid model architecture comprises three convolutional layers as given in Figure 5. Every layer has nine filters to extract the nine feature maps. Thus, the function of every filter is derived as,

\[ D_i = \varphi(F_i \ast D_{i-1} + \beta_i) \quad (1) \]

Where,

* Indicates the convolution function
\( F_i \) indicates the filter
\( D_{i-1} \) indicates a data of input \( \beta \) represents the bias
\( D_i \) represents an output feature map

Thus, the function of upper residual block is calculated as:

\[ D_1 = \varphi(F_1 \ast D_0 + \beta_1) \quad (2) \]
\[ D_2 = \varphi(F_2 \ast (D_0+D_1) + \beta_2) \quad (3) \]
\[ D_3 = \varphi(F_3 \ast (D_0+D_1+ D_2) + \beta_3) \quad (4) \]
\[ D_4 = \text{AvgP}(D_3) \quad (5) \]

Thus, the function of lower residual block is calculated as:

\[ D''_1 = \varphi(F_1 \ast D''_0 + \beta'_1) \quad (6) \]
\[ D''_2 = \varphi(F_2 \ast (D''_0+D''_1) + \beta'_2) \quad (7) \]
\[ D''_3 = \varphi(F_3 \ast (D''_0+D''_1+ D''_2) + \beta'_3) \quad (8) \]
\[ D''_4 = \text{AvgP}(D''_3) \quad (9) \]

According to the process of an inception net model, both the lower and upper blocks are parallely evaluated. The pooling and drop-out functions are also evaluated and executed. With a filter size of 2, the pooling layer calculates the average pooling operation.
Fig. 5. Hybrid ResNet and Inception Net Classification Model

### 4 Algorithm: Pseudocode of AO based TL model

**Input:** Size of Population, Number of Iterations, Pre-trained Model, Data sets and Image labels  
**Output:** best solution  
**Start Initialize** Optimizers  
  - Drop out ratio Batch size  
  - Performance metrics Fitness values $\leftarrow \left[ \right]$  
  - data $\leftarrow$ datasplit (train, validation)  
**initialize** AO populations  
solution $\leftarrow$ Fit Solution (population, np)  
TL $\leftarrow$ Train CNN (Model, solution, Validation, test, performance metrics) metrics $\leftarrow$ Test CNN (Model, Metrics, test, train)  
population **increment end**  
sorting the solutions  
population $\leftarrow$ update AO (sorted solutions, tuned parameters) //utilize an AO four steps to update best solutions)  
**Increment end**
5 Result Analysis

5.1 Result of AO Optimized TL Based Deep Convolutional Neural Network Model

As discussed in Section 4.4, the proposed TL model employs four DL methods namely DenseNet121, MobileNetV2, ResNet18V2, and AlexNet. Aquila Optimization which is based on hunting behaviour of Aquila is used for hyperparameter tuning. The AO model is used for a finetuning of hyperparameters of four CNN based TL architectures. The brain tumor data sets are collected with a few abnormal tumor types namely Metastases, Meningioma, Glioma, and Astrocytoma. In this section, the AO-based TL classification model is evaluated and compared with the previous methods. For every architecture, the results are analyzed for an average of sensitivity, specificity, accuracy, F1Score, precision, and IoU coefficient after applying 15 iterations. All of these DL models' performance was evaluated using F1Score, sensitivity, specificity, accuracy, precision, and IoU coefficient. Thus, the performance values for these metrics are tabulated in Table 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>F1Score</th>
<th>Precision</th>
<th>IoU</th>
</tr>
</thead>
<tbody>
<tr>
<td>DenseNet121</td>
<td>98.38</td>
<td>97.33</td>
<td>99.10</td>
<td>98.23</td>
<td>98.62</td>
<td>96.62</td>
</tr>
<tr>
<td>MobileNetV2</td>
<td>97.78</td>
<td>96.18</td>
<td>98.15</td>
<td>98.91</td>
<td>97.73</td>
<td>95.59</td>
</tr>
<tr>
<td>ResNet18V2</td>
<td>97.46</td>
<td>97.37</td>
<td>97.60</td>
<td>97.14</td>
<td>97.60</td>
<td>95.71</td>
</tr>
<tr>
<td>AlexNet</td>
<td>95.24</td>
<td>95.83</td>
<td>96.84</td>
<td>95.81</td>
<td>94.82</td>
<td>94.96</td>
</tr>
</tbody>
</table>
Based on Table 1 and Figure 7, the performance results of four DL models showed that the DenseNet121 based TL model is more efficient than other methods. The result of DenseNet121 metrics based on Accuracy, Sensitivity, Specificity, F1Score, Precision, and IoU are 98.38, 97.33, 99.1, 98.23, 98.62, and 96.62 respectively. The results of MobileNetV2 model in all iterations for the Accuracy, Sensitivity, Specificity, F1Score, Precision, and IoU are 97.78, 96.18, 98.15, 98.91, 97.73, and 95.59 respectively. The results of the ResNet18V2 model in all iterations for Accuracy, Sensitivity, Specificity, F1Score, Precision, and IoU are 97.46, 97.37, 97.60, 97.14, 97.60, and 95.71 respectively. The results of AlexNet in all iterations for the Accuracy, Sensitivity, Specificity, F1Score, Precision, and IoU are 95.24, 95.83, 96.9, 95.81, 94.82, and 94.96. It is clear that the DenseNet121 model achieved higher performance than all other models.
From Table 2 and Figure 8, the performance analysis and performance chart are shown. The performance comparison between the proposed TL model classifier and the prior TL models is evaluated. Finally, the AO-based TL model obtained a higher performance in all the metrics such as Accuracy, Sensitivity, Specificity, F1 Score, Precision, Recall, and IoU. From the result, it is concluded that the brain tumor prediction is performed effectively using the proposed AO-based TL model.

6. Conclusion

Proposed work for brain tumor detection is presented. In this work, the AO based TL model is proposed for an efficient brain tumor prediction. The main motive of this model is to predict the tumor effectively and treat a patient very earlier as soon as possible. This work is the hybrid of AO and TL models of AlexNet, ResNet-18, DenseNet and MobileNet method. The AO model is used for a finetuning of hyperparameters of four CNN based TL architectures. The optimized model is very supportive for the experts to handle an automation process. The method is evaluated and tested in brain tumor MRI data set image. The results showed that the AO based TL model obtained higher performance than other popular existing models.

References


