Received: March 14, 2021. Revised: May 21, 2021.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Brain Tumor Classification in MRI Images Using En-CNN

Hapsari Peni Agustin Tjahyaningtijas[1, 2] **Dewinda Julianensi Rumala**[1] **Cucun Very Angkoso**[3]
Nurul Zainal Fanani[4] **Joan Santoso**[5] **Anggraini Dwi Sensusiati**[6] **Peter M.A van Ooijen**[7]
I Ketut Eddy Purnama[1,8] **Mauridhi Hery Purnomo**[1,8*]

[1]*Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
[2]*Department of Electrical Engineering Universitas Negeri Surabaya, Surabaya, Indonesia*
[3]*Department of Informatic Engineering, Universitas Trunojoyo, Madura, Indonesia*
[4]*Department of Engineering. Politeknik Negeri Jember. Indonesia*
[5]*Department of Information Technology, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, Indonesia*
[6]*Faculty of Medicine, Airlangga University, Surabaya, Indonesia*
[7]*Department of Radiation Oncology, University Medical Center Groningen, Groningen, Netherlands*
[8]*Departement of Computer Engineering, UCE AIHeS, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
* Corresponding author's Email: hery@ee.its.ac.id

Abstract: Brain tumors are among the most common diseases of the central nervous system and are harmful. Early diagnosis is essential for patient proper treatment. Radiologists need an automated system to identify brain tumor images successfully. The identification process is often a tedious and error-prone task. Furthermore, brain tumor binary classification is often characterized by malignant and benign because it involves multi-sequence MRI (T1, T2, T1CE, and FLAIR), making radiologist's work quite challenging. Recently, several classification methods based on deep learning are being used to classify brain tumors. Each model's performance is highly dependent on the CNN architecture used. Due to the complexity of the existing CNN architecture, hyperparameter tuning becomes a problem in its application. We propose a CNN method called en-CNN to overcome this problem. This method is based on VGG-16 that consists of seven convolutional networks, four ReLU, and four max-pooling. The proposed method is used to facilitate the hyperparameter tuning. We also proposed a new approach in which the classification of brain tumors is done directly without priorly doing the segmentation process. The new approach consists of the following stages: preprocessing, image augmentation, and applying the en-CNN method. Our new approach is also doing the classification using four MRI sequences of T1, T1CE, T2, and FLAIR. The proposed method delivers accuracy on the MRI multi-sequence BraTS 2018 dataset with an accuracy of 95.5% for T1, 95.5% for T1CE, 94% for T2, and 97% for FLAIR with mini-batch size 128 and epoch 200 using ADAM optimizer. The accuracy was 4% higher than previous research in the same dataset.

Keywords: A brain tumor, Deep learning, Data augmentation, Hyperparameter.

1. Introduction

Brain tumor refers to the uncontrolled growth of clusters of brain cells in a particular area. Between 2012 and 2016, the Central Brain Tumor Registry of the United States (CBTRUS) recorded an annual occurrence of benign and malignant brain tumors and central nervous system (CNS) tumors of 23.41 percent or 100,000 [1]. In all brain and other CNS tumors, the most common form is malignant, called glioblastoma, which accounts for approximately 14.6 percent of all tumors and often affects males. By contrast, the most prevalent benign tumor is meningioma, which affects 37.4 percent of women [1].

Brain tumors are categorized as benign or malignant based on the pattern of cell development. Benign tumors are noncancerous and grow slowly, and therefore do not spread to other areas of the brain. Malignant, on the other hand, is cancerous, exhibits rapid cell proliferation, and invades other brain organs. Glioma is a tumor that arises from glial brain cells, accounting for 20% of brain and central nervous system tumors and 80% of malignant tumors. According to the World Health Organization, gliomas are categorized into four classes, from type I to IV [2]. Grade I tumors typically occur in children and are benign with a similar glial cell texture. Grade II tumors with prolonged growth differ slightly in texture. Grade III oligoastrocytoma is a malignant oligodendroglioma that combines oligoastrocytoma and anaplastic oligodendroglioma. Glioblastoma is a grade IV tumor that is the most serious type of gliomas and tissue defects visible to the naked eye [3]. Tumors in stages III and IV are malignant, while those in stages I and II are benign (low-grade glioma).

Imaging computer technology methods, specifically Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) provide a detailed brain tumor image [4]. MRI is one of the most frequently used noninvasive approaches for diagnosing and treating brain disorders because it determines the size and location of contaminated tissue [5]. MRI greatly increases the amount of knowledge available to determine normal anatomy. It is more effective than all other imaging techniques at detecting brain tumor tissues than any other imaging technique due to the high contrast nature of soft tissue, the high spatial resolution, and the lack of harmful radiation [6]. The diagnosis is made in this study using a variety of MRI sequences, including T1 MRI, T1-CE MRI, T2 MRI, and fluid-attenuated inversion recovery (FLAIR) MRI. Brain tumor examination via imaging modalities has gained prominence in the radiology department.

Appropriate brain tumor classification is important for determining a patient's prognosis and treatment. A screening technique involving a radiologist visualizing the brain's MRI is important. Due to the time and subjective nature of the visual diagnosis, an accurate computer-aided diagnostic (CAD) method has been developed [5]. Machine learning and deep learning have emerged as new tools for medical image processing.

Medical image processing has advanced significantly as a result of the rapid advancement of computers and deep learning. Machine learning (ML) is an application of artificial intelligence (AI) that uses statistical techniques to automatically generate a model from a collection of data, thus enabling computers to "read." It enables computers to learn from a variety of data and generate a model for performing the input-output process without the need for explicit software code. The ML algorithm has been widely used as part of AI in medical imaging. Machine learning is classified into two types: supervised and unsupervised. The supervised algorithm is used to test the input variable mapping function to predict novel subject labels using output labels. The algorithm's primary objective is to discover inherent patterns in training data. Several algorithms can be used to accomplish this, including K-Nearest Neighbors (KNN) [7], Support Vector Machine (SVM) [8, 9], and Artificial Neural Network (ANN) [10]. Unsupervised learning is the process of analyzing previously undetected trends in datasets without the use of labels or human control, for example, using fuzzy c-means [9] and Self-Organization Map (SOM) [11].

Extracting features from a training image is critical for the learning process. After segmentation, a grayscale, texture, and statistical feature extraction is performed. This is referred to as handcrafted feature extraction, and it needs an expert with strong expertise while understanding the features is critical. Additionally, the use of large-scale data increases time consumption and error rates.

Deep learning (DL) is a sub-part of machine learning concerned with multiple representation layers and their associated hierarchical features. The DL algorithm uses several layers of nonlinear identities for extracting features. Each sequence layer's output corresponds to the preceding layer's input [12]. Convolutional neural networks (CNN) are used in deep learning to process visual images in optimal 2D and 3D formats. It is motivated by the way the brain works and how it deals with large amounts of data. Le Cun launched CNN in 1998 as a network called "Lenet." However, it became prevalent when Krizhevsky won the CNN Imagenet Image Classification Competition with Alexnet in 2017.

Machine learning and deep learning techniques are often used on massive MRI brain tumor datasets that provide ground reality for a variety of practice cases. Classification problems are solved using artificial intelligence and historical data. While designing the optimal classifier, accuracy, performance, algorithm complexity, and computational resources must all be considered.

Prior to the classification point, researchers used the segmentation method in their previous approaches. Segmentation is a repetitive, time-consuming, and subjective procedure that is dependent on the medical staff's expertise. Bahadure *et al.* [13] proposed using Berkeley wavelet transformation (BWT) and SVM to segment and identify MRI brain tumors. Additionally, they extracted the function using the statistics feature formula.

Alfonse and Salem recommend performing a fast Fourier transformation to extract features for automatic classification using the SVM in enhancing the classifier's performance [9]. The reduction of features employs approaches based on minimum redundancy and maximum relevance. Additionally, Kumar and Vijay Kumar [14] developed a theory factor analysis method for feature reduction and the radial base function of the SVM kernel base for classification. Ramdlon *et al.* [15] categorized brain tumors into three types: Astrocytoma, Glioblastoma, and Oligodendroglioma. They performed segmentation using the watersheds approach and feature extraction using shape-based features using a KNN classifier. Meanwhile, the stage used a filled area and a centroid to extract features.

In another publication, Amin *et al.* [16] segmented brain tumors before classifying them using the Potential Field (PF) clustering process. To boost classification accuracy, the Local Binary Pattern (LBP) and wavelet transform features are used. Besides the disadvantage of performing segmentation before classification, the current method has a disadvantage in that the efficiency of the classifier machine is highly dependent on the combination and selection of features. This complicates the optimization of the classifier's efficiency, especially during the preprocessing stage. If it is assumed that selecting useful features is no longer necessary, then self-learning features are needed during the machine learning preprocessing stage. As a result, the best approach for the function to self-learn is deep learning via the CNN method.

Compared to traditional approaches, which use feature extraction as an input for classification, CNN analyzes complex features that directly reflect the data in real-time. CNN takes the image's collected patches as input and extracts increasingly complex feature hierarchies using trainable constitutional filters and local subsampling. Amin *et al.* [17] implemented the CNN tumor classification system and preprocessed the data using the discrete wavelet transform (DWT) fusion process method. This technique enhances the accuracy of MRI images by suppressing noise. Additionally, global thresholding was used to isolate tumors for CNN classification.

Sajjad *et al.* [18] used multimodal CNN-based tumor classification, segmenting the brain tumor using CNN. The image augmentation process was used to complement the pre-trained process data used as input to the VGG-19 model. The results obtained from the selected data help to boost the overall performance of the system. Vijh *et al.* [19] determined the optimal segmentation threshold using adaptive particle swarm optimization (PSO) and the Otsu process. Additionally, anisotropic diffusion (AD) filtering was used to eliminate noise and increase image quality, and the resulting features were used as CNN inputs.

Sharif *et al.* [20] suggested a saliency-based tumor segmentation method and an inception V3 pre-trained CNN for MRI classification of brain tumors. The inception V3 network is composed of 95 layers of batch normalization, 94 layers of convolutional layers, and layers of rectified linear unit (ReLU). Prior to classification, these studies conducted segmentation and feature extraction, which raises the probability of additional classification levels. Several additional methods for classifying brain tumors have been created. Ghasemi *et al.* [21] used six layers of CNN to complement and extract features using a generative adversarial network (GAN) method.

Khan *et al*. [22] suggested a five-step research process for binary tumor classification of malignant and benign brain tumors using the VGG network and MRI images of multi-sequence brain tumors (T1, T2, T1CE) and FLAIR. Preprocessing edge-based histogram equalization and discrete cosine transformation (DCT) have been used for linear contrast stretching. VGG-16 and VGG-19 networks were used to extract features at the second stage. An extreme learning machine (ELM) was used to choose the best features at the third stage. A robust partial least squares (PLS) covariant was used in one matrix in the fourth step and then fed to ELM for final classification. Gashemi [21] and Khan [22] both used feature extraction and feature selection in their study. Additionally, they employ a sophisticated CNN architecture and conduct direct classification on a variety of MRI sequences, including T1, T1CE, T2, and FLAIR, lowering classification accuracy.

However, the current approach used a CNN network, which is very complex, and also required segmentation and feature extraction prior to classification. Given that tuning hyperparameters such as the number of convolutional layers, optimizers, large quantities of training data, and dropouts all affect deep learning efficiency, CNN has struggled with hyperparameter tuning.

Thus, in this paper, we propose a new method called en-CNN to classify brain tumors into benign (Low-Grade Glioma or LGG) and malignant (Glioblastoma or GBM). The en-CNN stands for “enhanced CNN.” This approach is based on VGG-16 architecture but simpler architecture, with fewer layers and parameters. Seven convolutional networks, four ReLU networks, and four max-pooling networks are used to allow hyperparameter tuning. Additionally, we suggest a method in which brain tumor classification is performed directly without first performing segmentation. This approach consists of the following stages: preprocessing, image augmentation, and applying the en-CNN method. Our new approach is also doing the classification using four MRI sequences of T1, T2, T1CE, and FLAIR.

The contributions of this paper are as follow:

1. Propose a new method called en-CNN that is based on VGG-16 but has simpler architecture by omitting several convolutional layers, which then simplifies the process of hyperparameter tuning
2. Applying classification method directly without doing segmentation process before the classification
3. Applying the proposed architecture using four sequences of MRI images: T1, T1CE, T2, and FLAIR

The remainder of this paper is organized as follows. Section 2 of this paper is material and methods, covering the research description and methods used, starting from the dataset to the network. Section 3 is Theory and Calculation, which describes CNN and how it works. Sections 4, 5, and 6 are results, discussion, and conclusions, respectively.

2. Data set

As shown in Fig. 1, this research used the Brain Tumor Segmentation (BraTS) dataset, which consists of 3T multi-sequence MRI GBM and LGG scans. It is updated regularly by trained radiologists in collaboration with regular clinicians and is confirmed by pathological diagnosis.

Preprocessing of the BraTS dataset included recording directly with the same anatomical template, interpolation to the exact resolution (1 mm3), and skull stripping [23]. The BraTS 2018 dataset included multi-sequence MRI images of 101 GBM and 66 LGG patients from various perspectives, including axial, coronal, and sagittal, as illustrated in Fig. 2.

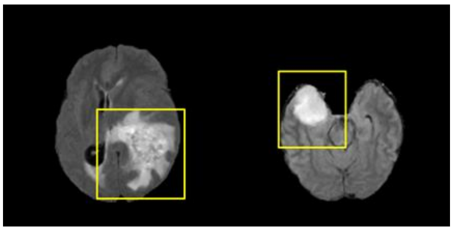


Figure. 1 From left to right, different grades of gliomas BraTS dataset GBM and LGG. The tumor has a yellow square

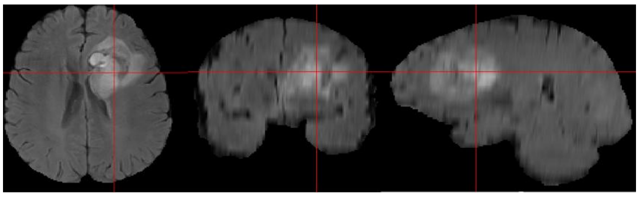


Figure. 2 From left to right, different acquisition of Glioblastoma (Axial, Coronal, and Sagittal). Tumors are localized with a red cross line

3. Convolutional neural network (CNN)

3.1 CNN architecture

CNN architecture is constructed from three layers, consisting of Convolutional, Sub-Sampling layer, and fully connected (FC) layers, and each layer can be modified onto several layers. The first layer of CNN, the convolutional layer, is liable to detect local features extracted from input image locations. Subsequently, those layers provide a connecting layer that converts inputted images through feature maps wrapped with filters. The convolution representation is depicted in Eq. (1) [24].

$$z^l = h^{l-1} * W^l \tag{1}$$

where z^l is pre-activation layer l, h^l is an activation of layer l function, and W^l learnable parameter. The second layer reduces the features map dimension by choosing pixel values as output based on specific regulations. Max pooling has been used as an algorithm on subsampling layers as in Eq. (2) below:

$$h^l_{xy} = max_{i=0,\dots,a,j=0\dots,s} h^{l-1}_{(x+i)(y+j)} \tag{2}$$

FC layers are capable of classifying data and performing nonlinear transformations to obtain the desired output value. The first and second layers correspond to the learning section, while the third layers (FC layers) correspond to the classification section.

$$z^l = W^l h^{l-1} \tag{3}$$

Apart from the filter and kernel sizes, the convolutional layer has parameters for stride (S) and zero-padding (ZP). Suppose the stride value is 1, the convolution filter shifts by 1 pixel horizontally, followed by a 1-pixel vertically. The smaller the step, the more detailed information input can be obtained, even when it affects the computing process. Additionally, the output of the convolutional layer becomes the input of the subsequent convolutional layer. The output is always smaller than the original

Received: March 14, 2021. Revised: May 21, 2021.

condition's input dimension, wasting much information (except using a stride one 1x1 filter). However, ZP, a parameter that specifies the number of pixels on either side of the convolutional layer's input (which contains the value 0), is used to circumvent this limitation. Additionally, it acts as a manipulator, ensuring that the output of the convolutional layer matches the dimensions of the input. By utilizing a deeper convolutional layer to extract additional features, the ZP parameter improves CNN performance and makes CNN increasingly useful for extracting and identifying multi-dimensional data.

Fig. 3 shows a 3x3 kernel implementation on a 3x3-size image with zero paddings and producing a 3x3 convolved image with S = 1. The activation function is before the pooling layer and after the convolution process.

An activation function is the resulting convolution value at this point. The activation function is ReLU, where the output value of the neuron can be represented as 0 when the input is negative. If the activation function input is positive, the neuron output equals the value of the activation function input. The activation mechanism of ReLU is depicted in Fig. 4.

$$f(z) = \max(0, z) \tag{4}$$

The pooling layer decreases the size of the image dimension, a process known as downsampling. The downsampling method divides the image into small rectangles, such as 2x2, traverses the image in steps of 2x2, and determines the maximum value of the four elements. Fig. 5 illustrates a max-pooling layer in action.

Another often used parameter in CNN is the dropout function. Dropout is a regularization technique for neural networks in which multiple neurons are randomly chosen and ignored during training, though they can be spelled out randomly. The network temporarily disables the input of discarded neurons, and the new weight cannot be used during backpropagation.

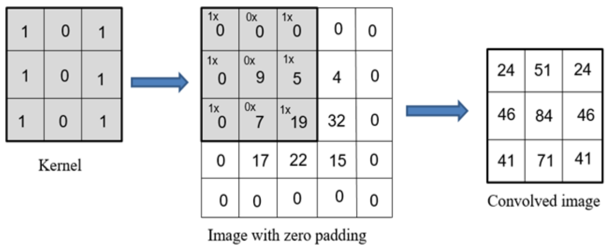


Figure. 3 Convolutional layer example with 3×3 input image, 1 zero paddings, 3×3 kernel size, stride:1, and output: (3×3)

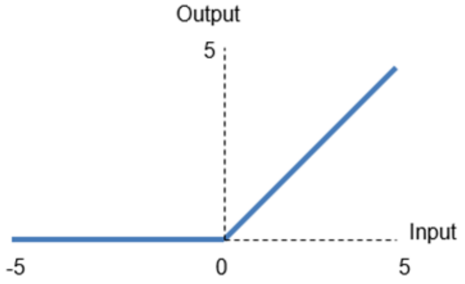


Figure. 4 ReLU activation function

Dropout refers to the process of eliminating secret or observable neurons from a network in order to avoid overfitting and accelerate learning. Removing a neuron from the network means momentarily removing it from the network. A neuron is randomly selected and given a probability value between 0 and 1. In the proposed architecture, 10% dropout probabilities were found to be the most appropriate for dropout layers 1 and 2.

Fig. 6 (a) illustrates a simple neural network with two hidden layers. In part (b), the neural network used the dropout regularization strategy, which eliminated the need for multiple activation neurons. This technique was applied to CNN models and had an impact on their training efficiency in addition to eliminating overfitting.

Additionally, three advanced layers were used: FC, sigmoid, and classification layer. The FC is used to bind each neuron in a layer to those in the preceding and following layers. The sigmoid layer is used to suppress all predicted classes between 0 and 1. This layer's performance can be determined as follows:

$$y(z_k)j = \frac{1}{1+e^{-z_k}} \tag{5}$$

The probability of any class (j) can be calculated over (k) different classes as a function y (z). Finally, the classification layer was evaluated using cross-entropy loss. Additionally, these classifications assign a final binary prediction mark to each image that was used. Eq. (6) is used to represent loss, where p(z) denotes the target label vector, q (z) denotes the contribution from the sigmoid layer, and C denotes the number of classes.

$$H(p, q) = -\sum_{i=1}^{C}(p_i(z)\log(q_i(z))) \tag{6}$$

3.2 Regularization and optimization techniques

Regularization is used to eliminate generalization errors while minimizing training losses. The generalization error is the difference between the sampling error obtained when testing the model using previously observed data and the error obtained when testing the model using previously unknown data. With a narrow generalization error, the model created by the learning process should be well generalized. When the generalization error is high, the model tends to overfit because its capacity is too high. The model is restricted to the training data given and cannot be generalized to previously unseen data. The inverse of overfitting is underfitting, which occurs when the resulting model cannot extract the appropriate function from the data.

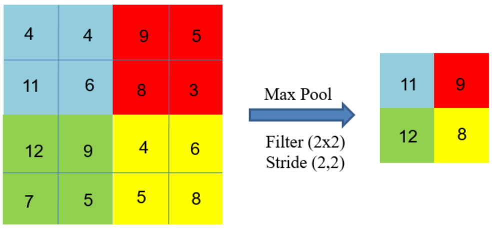
Numerous techniques are used to mitigate over- or under-fitting, including data augmentation by geometric and original image color adjustments. Additionally, dropout layers are used to stochastically erase hidden unit weights. Regularization with L2 applies a penalty to the cost function and decays the weights, as shown in the following equation [25].

$$Cost\ function' = Loss + \lambda \sum_{i=1}^{k} w_i^2 \qquad (7)$$

Where λ is regularization parameter and w is the corresponding weight(s) for i = 1, . . . , k. In an idealistic situation, it is critical to achieving the global minimum by optimizing parameters and minimizing the loss feature. Convergence reaches when small steps towards a negative gradient are taken. This study shows that the "adaptive moment estimation" is the optimal optimizer for the proposed structure.

4. The proposed method

Fig. 7 illustrates the proposed method's block diagram. The system begins by extracting the raw brain tumor data image set and then labeling the image. The extracted picture and its mark are fed into the system. Considering that, preprocessing and augmentation are performed, proceeded by the dataset being divided into preparation, validation, and test datasets.



4	4	9	5
11	6	8	3
12	9	4	6
7	5	5	8

11	9
12	8

Figure. 5 Max-pooling layer 2x2 example

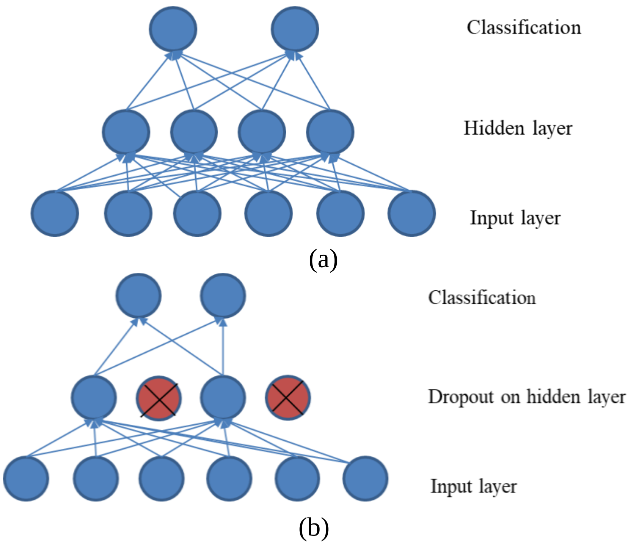


Figure. 6 Dropout layer example with 50% probability

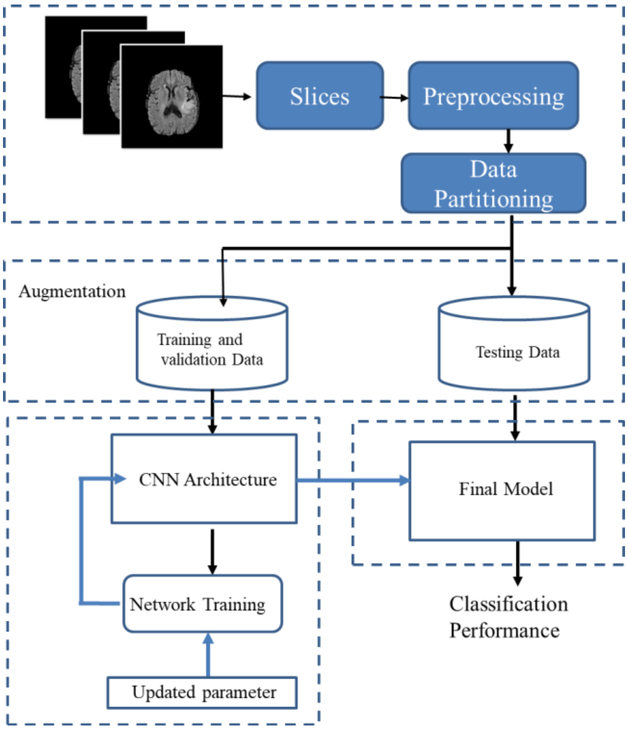


Figure. 7 The CNN architecture block diagram

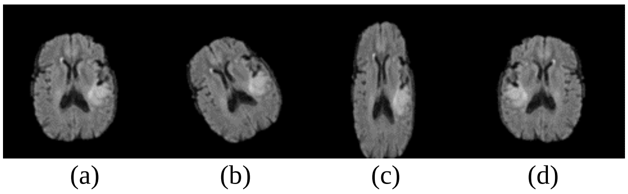


Figure. 8: (a) The original image, (b) rotation, (c) zoom, and (d) horizontal flip (From the left to right)

Regarding that, a deep learning model was built using the proposed network, followed by hyperparameter tuning, regularization techniques,

Received: March 14, 2021. Revised: May 21, 2021.

and algorithm optimization. Even after that, network training and computing efficiency were discussed.

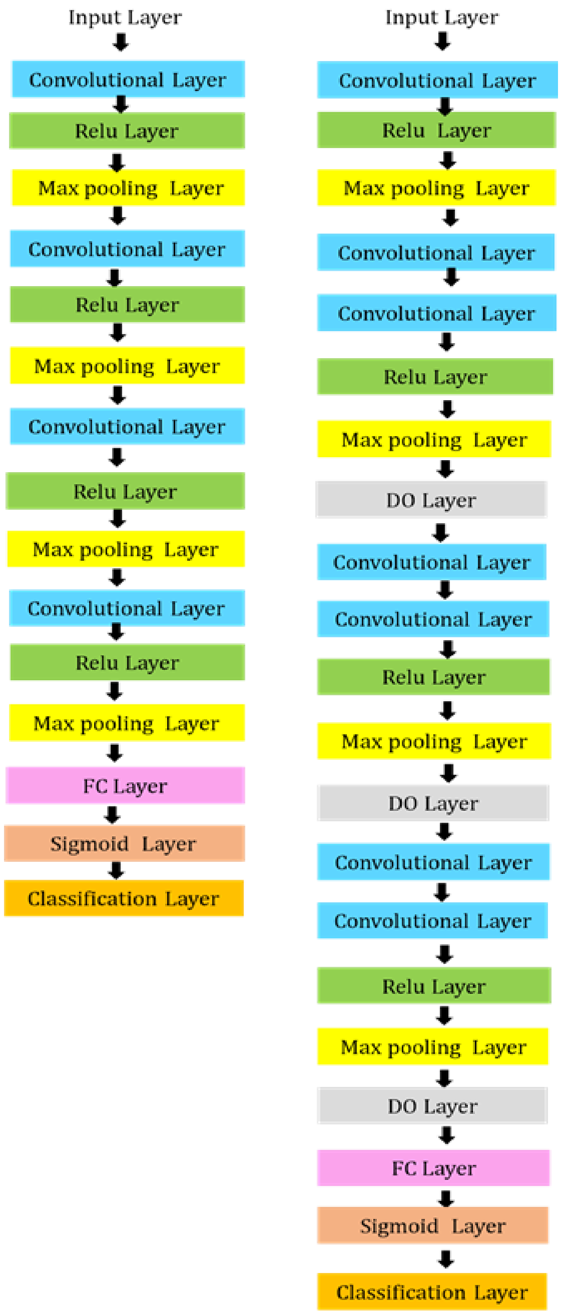
4.1 Preprocessing stage

The classification stage used only MRI slices with tumors for pre-processing. Slices are chosen in accordance with the BraTS ground-truth image. Apart from shrinking the original image from 255x255x1 to 224x224x1, the dataset is resized to reduce processing time. Prior to splitting, data is shuffled to preserve the process, train with the original data, and focus on the dataset's narrow band. The data will be segmented into three groups: training, validation, and testing (64 % for training and 36 % for testing and validation). The next step is to augment image data in order to reduce overfitting and improve model accuracy. As shown in Figure 8, augmentation involves multiplying the dataset for the machine to accept the image as fresh.

Algorithm 1. Pseudocode for preparing the data and augmentation

```
Input  : Image X, Labels Y
Output: Image X_generated and Labels
          Y_generated

Get X as array of images
Get Y as array of each image label
Set num_operations as number of
augmentation operations
Initialize BATCH_SIZE
For each i in num_operations
        Initialize augmentation operation
        {shear, zoom, horizontal flip}
For each slice in BATCH_SIZE
        Generate data based on
        augmentation operations:
        X_aug[i][slice] as result of
        generated image data
        Y_aug[i][slice] as the image label
        of generated image
Initialize generated_data array to save all
        generated images and labels
For each i in num_operations
        For each slice in BATCH_SIZE
            Combine X_aug[i][slice] and
            Y_aug[i][slice] to the
            generated _data
Shuffle the data inside the generated_data
Create X_generated array
Create Y_generated array
For each images and labels in generated_data
        X_generated = images
        Y_generated = labels
```



(a) (b)

Figure. 9 (a) CNN Network, (b) en-CNN (Proposed Network)

This study analyzed 1000 slices of GBM from 101 patients and 1000 slices of LGG from 66 patients. Each of the shear [26], zoom [27], and horizontal flip processes [28] adds a multiplier factor of two to the results. This means that the datasets for the GBM and LGG classes were each 6000.

4.2 Proposed CNN architecture

Fig. 9 (a) illustrates a 16-layer CNN network, beginning with the input, which is the image added in the previous preprocessing step. The activation function is used to select features, and downsampling is used to reduce the size of the CNN layers' arrays through sequential convolution, ReLU, and pooling. The dropout layer, followed by a completely linked and sigmoid layer, eliminates overfitting. Classification is used to forecast future grades. The 2D layer performs convolution on an image by applying a convolution kernel (filter) of a defined size (MxN).

Additionally, the machine obtains new representative data by multiplying portions of the image with the filter and calculating the dot product of the weights (kernel weights) and the input. The filter alters the picture vertically and horizontally, which is referred to as stride (S). The padding or zero padding parameter determines how many pixels are inserted to each side of the input and is used to control the convolution layer (feature map). The output measurements can be calculated to ensure that they remain consistent with the input or that they are not significantly reduced using padding. As a result, a deeper convolutional layer is used to remove additional features effectively, as shown in Fig. 9. (a).

VGG Net is a simpler CNN architecture meant to stack convolutional layers by increasing filter size. Suppose the convolution layer 1 has 13 filters. The next one should have the same filter size, 13 or more filters. Furthermore, all the filter sizes used are the exact dimensions of 3 x 3. The size of these two 3x3 filters is cheaper than one 5x5 filter.

Since VGG-16 has thirteen convolutions, five max-pooling, three fully connected (FC) layers, and one Softmax as a classification layer, the process of hyperparameter tuning are not accessible. An en-CNN (Fig. 9(b)) was developed based on VGG-16 with seven convolution, four max-pooling layers, and one fully connected layer. A sigmoid layer was used after a fully connected layer instead of softmax for binary classification. After each layer, a dropout layer was added to reduce overfitting. By simplifying the VGG network (en-CNN), the hyperparameter tuning process becomes more manageable.

The en-CNN model is implemented using a higher level Python library, Keras, which runs over an open-source deep learning framework Tensor Flow as a backend to classify the brain tumor MRI images.

5. Experiment result

The accuracy and loss of multimodal MRI brain tumors show in Fig. 10. The training accuracy is nearly 100% after the 200 epochs, while the resulting loss is close to zero for all MRI sequences. MRI sequences T1 and T2 have more fluctuations in the loss. Fig. 11 shows validation accuracy and losses for multimodal brain tumors. Accuracy shows a plateau of nearly 100 percent, with loss decreasing to zero. According to Fig. 11, the accuracy and loss of validation on T1 and T2 sequences are more unstable than T1CE and FLAIR. In T1CE and FLAIR sequences, the curve tends to be smoother after epoch 150.

The models performance can be measured by finding its accuracy and recall. Some terms

.Algorithm 2. Building the en-CNN model

```
Training Algorithm
Input   : Image , label
Output: Trained Model

1. Define the epoch number
2. Define the input shape of the network
3. Initialize number of blocks = 4
4. for each number of block
        If number == 0
            Get feature maps by
                convolutional process
                using Eq. (1)
        Else
            Get feature maps by doubling
            the convolutional process
        Do nonlinearity operation with Relu
        layer using Eq. (4)
        Maxpool the feature maps using Eq.
        (2)
        Dropout some number of layer
        outputs
5. Flatten the feature maps
6. Feed flattened feature maps through fully
   connected layer using Eq. (3)
7. Get output class through sigmoid layer using
   Eq. (5)
8. For each epoch
        Update the weights and loss using
        Eq. (6) and Eq. (7) respectively
```

Algorithm 3. Classification algorithm for en-CNN

```
Testing Algorithm
Input   : TestImage, label
Output: Prediction Result
1. Feed the test set with n number of data to the
   trained network
2. Get prediction score
3. For each test data
        If prediction score > 0.5
            Prediction result is "HGG"
        Else
            Prediction result is "LGG"
```

commonly used to measure classification model performance include positive and negative predicted cases, the basis for the search for true positive (Tp), true-negative (Tn), false-positive (Fp), and false-negative values (Fn). These terms are usually summarized as a confusion matrix, as shown in Fig. 12. Precision, Sensitivity, Specificity, and Accuracy are calculated, as shown in Eq. (8).

$$Sensitifity = \frac{Tp}{Tp+Fn}.100\%$$
$$Specifity = \frac{Tp}{Tn+Fn}.100\%$$
$$Accuracy = \frac{Tp+Tn}{Tp+Fn+Tn+Fp}.100\%$$
$$Precission = \frac{Tp}{Tp+Fp}$$
$$F1\ score = \frac{2(Precision)(Recall)}{Precision+Recall} \qquad (8)$$

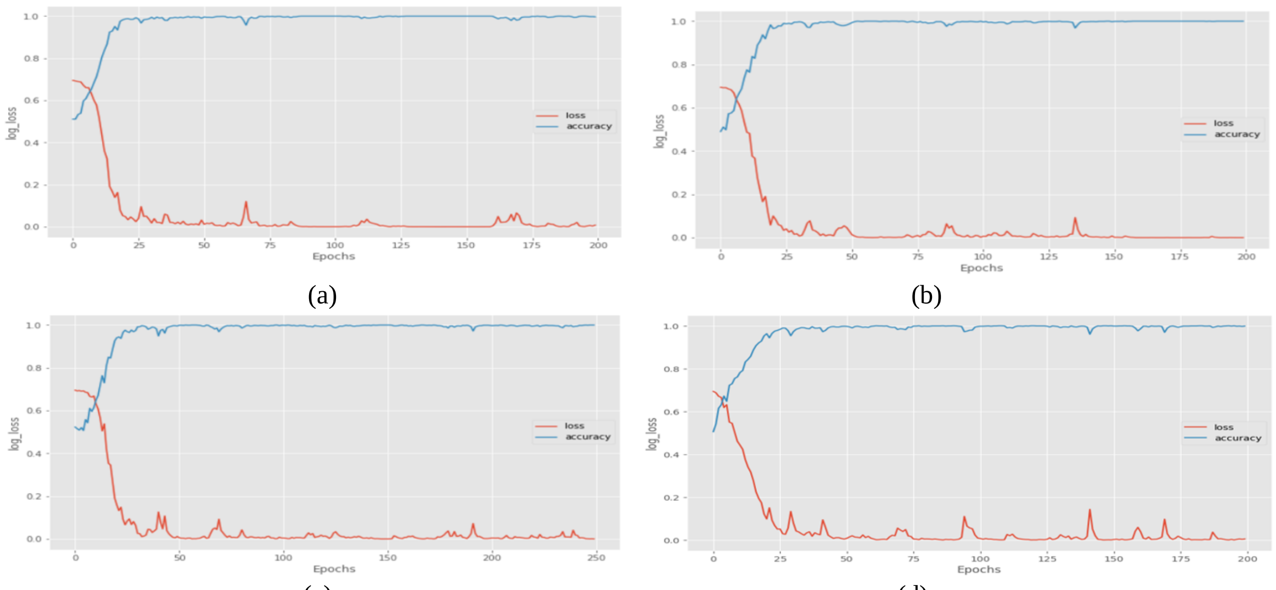
True positive and negative values refer to positive and negative predicted cases that are classified correctly by the model. In comparison, false positives and negatives are negative and positive predicted cases classified as positive and negative classes by the model. The confusion matrix on the image shows that the classification models performance can be calculated, as shown in Table 1.

In this study, various hyperparameter tunings were assessed to obtain the appropriate tuning and high accuracy results, as shown in Table 2. The highest classifier performance was achieved at epoch 200, as shown in Table 3 and Fig.12. The batch value was constant at 128 with a dropout layer, variable number of epochs, and an optimizer function in the design trials. Table 3 shows the test results, where the more extensive the epoch, the higher the accuracy.

The en-CNN architecture provides accuracy values that increase from epoch 30 to 300. The best performance is obtained at epoch 200 with 97 % accuracy, 97.03 % sensitivity, 96.97 % specificity, 0.97 precision, and 0.97 F1 scores. FLAIR MRI brain tumor image dataset obtained the highest performance value. In Fig. 12, the proposed network was tested with variations in optimizer function and epoch value from 30 to 300 epoch and batch size 128.

The optimizer value that best matches the FLAIR is The adaptive moment estimation, or "ADAM," it was performed on an MRI brain tumor dataset. According to Figure 13, the ADAM optimizer function achieves a higher accuracy of 97 percent as compared to Root Mean Square Propagation (RMSProp) at 92.6 percent and Stochastic Gradient Descent (SGD) at 87.5 percent. According to Kingma [29], the ADAM optimizer excels at dealing with massive datasets and space parameters with a high dimension. Additionally, the Adam optimizer has been used to accelerate network convergence [30]. This is consistent with the brain tumor dataset we use, which is a large collection of high-dimensional parameters designed to aid in the computation process.

The following table compares pre-existing layers to en-CNN. According to Table 4, the en-CNN architecture has fewer convolutional layers, a smaller filter size, a ReLU layer, a max-pooling layer, and an FC layer than the VGG-16 and VGG-19 architectures.



(a)

(b)

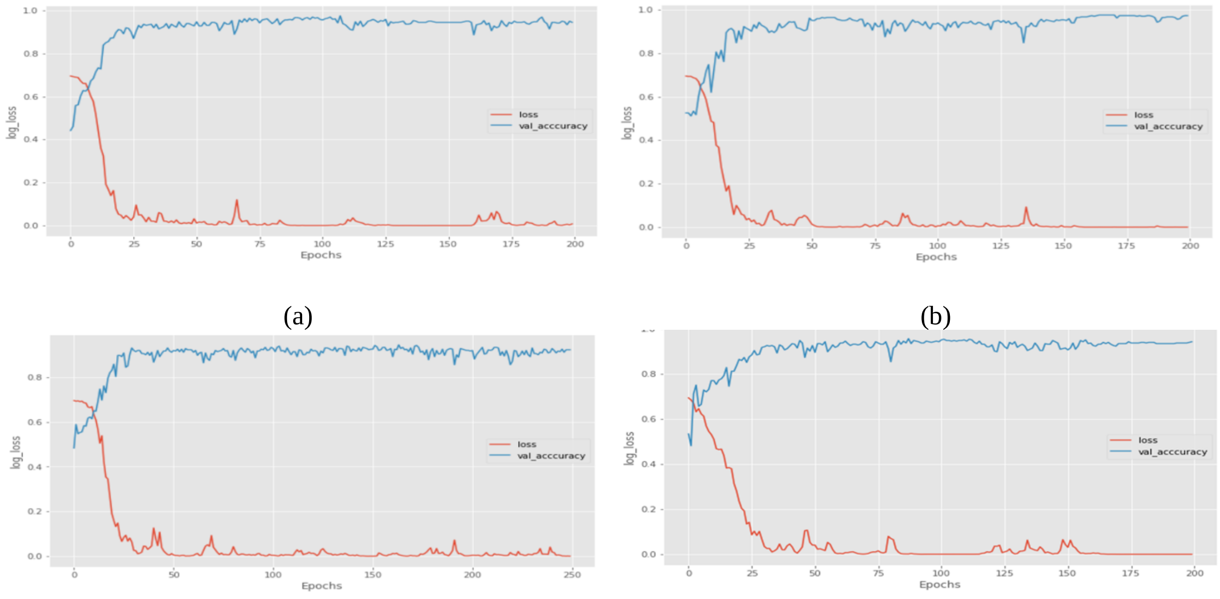
(c)

(d)

Figure. 10 en-CNN Network Training Accuracy and Loss for: (a) T1, (b) T1CE, (c) T2, and (d) FLAIR brain tumor MRI image classification

Table 1. Proposed classifier performance of brain tumor multimodal MRI sequences

Multimodal	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision	F1 Score
T1	95.5	96.08	94.9	0.96	0.95
T1CE	95.5	95.7	95.33	0.95	0.95
T2	94	97	91	0.94	0.94
FLAIR	**97**	**97.09**	**96.91**	**0.97**	**0.97**



(a) (b)

(c) (d)

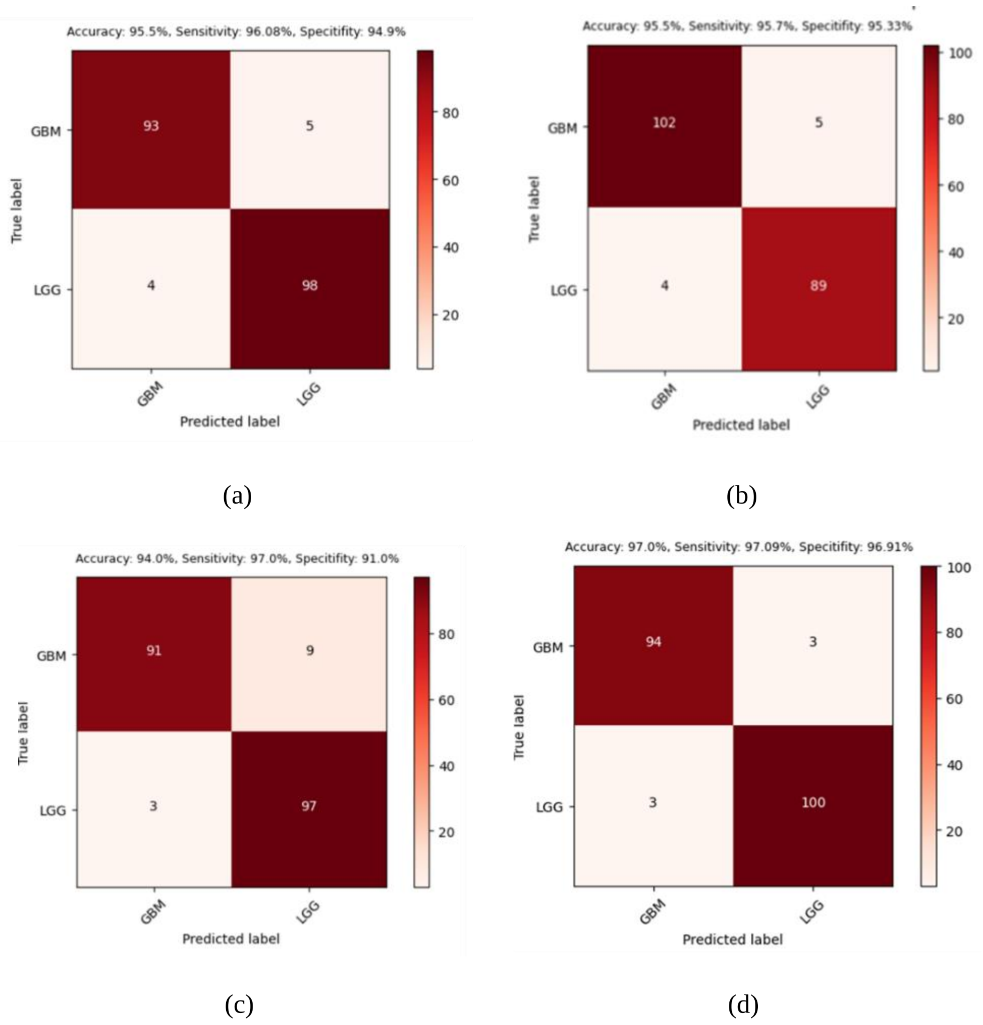
Figure. 11 en-CNN Network Validation Accuracy and Loss for: (a) T1, (b) T1CE, (c) T2, and (d) FLAIR brain tumor MRI image classification.

Table 2. Hyperparameter and architecture before the final model

Factor(s)	Values
Number of convolutional + ReLU layers	1,2,3,4,5,6
Number of drop out layers	1,2,3
Maximum epochs	30,50,100,150,200,250,300
Number of fully connected layers	1,2
Optimizer	ADAM, RMSprop, SGD.
Mini batch size	32, 64, 128
Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8

Table 3. Classifier performance for brain tumor FLAIR MRI according to various epoch

Epoch	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision	F1 Score
30	89	84	94	0.89	0.89
50	89	92.86	86.21	0.89	0.89
100	92.5	92.05	92.86	0.92	0.92
150	96	94.39	97.85	0.96	0.96
200	**97**	**97.03**	**96.97**	**0.97**	**0.97**
250	96	98.96	93.27	0.96	0.96
300	97	97.09	96.91	0.97	0.97



(a)

(b)

(c)

(d)

Figure. 12 en-CNN prediction Confusion Matrix for: (a) T1, (b) T1CE, (c) T2, (d) and FLAIR brain tumor MRI image classification.

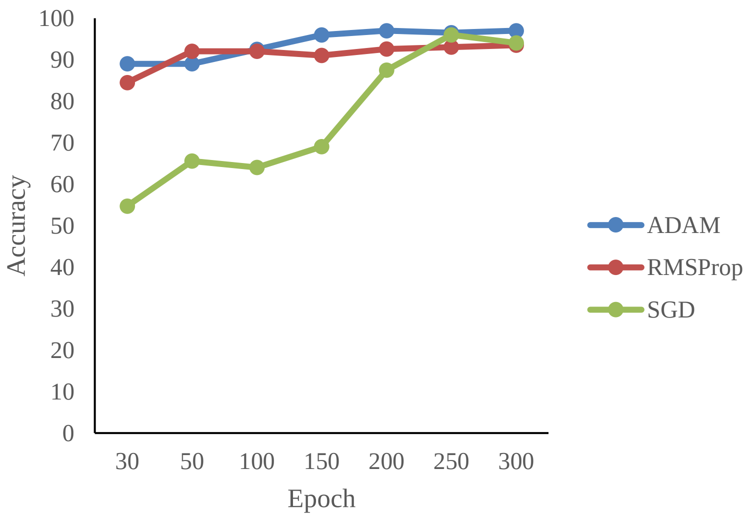


Figure. 13 Proposed network accuracy based on optimizer function for FLAIR brain tumor image classification

Table 4. Layer comparison of VGG-19, VGG-16, and en-CNN

Layers	VGG-19	VGG-16	en-CNN (Proposed)
Image Input Size	224x224 pixel	224x224 pixel	224x224 pixel
Convolutional Layer	16	13	7
Filter Size	64,128, 256,512	64,128	32,64,128
ReLU	18	5	4
Max Pooling	5	5	4
FC Layer	3	3	1

Table 5. The proposed model and previous related comparison

Model	Dataset	MRI. Modal	Classification Type	Classification Method	Accuracy (%)
Bahadure [13]	Brain Web	T1	Multi	SVM	96.51
Ramdlon [15]	The Cancer Imaging Archives (TCIA)	T1	Multi	KNN	89.5
Sajjad [18]	Radiopedia and Nanfang Hospital and General Hospital, Tianjin Medical University.	T1CE	Multi	VGG-19	95.5
J. Amin [16]	**BraTS 2015**	**T2, FLAIR**	**Binary**	**SVM**	**94**
Sharif [20]	**BraTS 2018**	**T1, T1CE, T2, FLAIR**	**Binary**	**Inception V3**	**92.5**
Khan [22]	**BraTS 2018**	**FLAIR**	**Binary**	**VGG-16**	**92.79**
Proposed	**BraTS 2018**	**T1, T1CE, T2, FLAIR**	**Binary**	**en-CNN**	**97**

Meanwhile, Table 5 shows a comparison of the proposed methods against existing research. Table 5 shows that the proposed method has higher accuracy in the classification of brain tumors.

6. Discussion

This study explained how to classify brain tumors using a deep learning technique called en-CNN. The Proposed Approach is more reliable than the SVM methods proposed by Bahadure [13] and Amin [16]. This is because, in addition to several multi-modalities, deep learning classifications were used, while the previous two studies focused exclusively on a single sequence modality and machine learning methods.

Additionally, the proposed method outperforms Ramdlon [15], as we used a deep learning-based classification method without segmentation or feature extraction stages in this analysis.

Moreover, the proposed approach outperforms Sajjad [18], Sharif [20], and Khan [22] as compared to other approaches in the deep learning community. This is because Sajjad [18] uses fewer sequences, specifically two, whereas other researchers' proposed method is more effective while using the same number of sequences. This is because they conducted their research using a complex network, namely VGG-16 and Inception V3. Meanwhile, the approach we suggest is more reliable due to the simpler design, which makes determining the architecture's tuning hyperparameter easier.

As input, the network receives an MRI brain tumor image. In contrast to machine learning, the proposed classification approach does not require manual feature extraction; rather, a filter on a convolutional network performs feature extraction automatically. Similar to some current approaches, the proposed classification method is non-segmented. This research experimented with different CNN model parameters before settling on the optimal value in the previous architecture. The most difficult part of the process is the CNN preparation, which takes weeks to complete without overfitting or underfitting brain tumor datasets. As Triwiyatno [31] states, the efficiency of deep learning is highly dependent on hyperparameters. The more complex the network, the more challenging the tuning of the hyperparameters becomes.

We addressed this issue in this work by simplifying the CNN network called en-CNN. En-CNN was able to distinguish two distinct types of brain tumors correctly. A confusion matrix with a precision of greater than 95% indicates the accuracy achieved. As illustrated in Table 4, en-CNN appeared to have a simpler architecture than VGG-16 and

VGG-19 but did not degrade classification results. We used VGG-16 as a comparison; while VGG-16 used thirteen convolutional layers, en-CNN used seven. If VGG-16 used five layers of ReLU and max-pooling, en-CNN used four layers of ReLU and max-pooling. Three FC layers were used in VGG16, while en-CNN used a single FC layer.

Additionally, we compared our results to those of previous research, as shown in Table 5. We compare many reported findings to proven methods. Amin [16] analyzed the BraTS 2015 dataset using machine learning techniques. The authors used a handcrafted feature in their analysis, achieving a classification accuracy of 94% using SVM for the LBP and wavelet transform fusion features. Khan [22] classified brain tumors using the BraTS dataset using a combination of pre-trained VGG networks. This approach achieved a classification accuracy of 92.5 %. Sharif [20] classified brain tumors using the BraTS 2018 dataset and a 92.5 % accurate Inception V3 network. In comparison to previous research using the same dataset, our findings indicated a 4% improvement in classification accuracy.

The BraTS data set was used in this analysis. While we are still in the segmentation stage of the BraTS datasets usage problem, we have not ruled out the possibility of using datasets for brain tumor classification. Table 5 contrasts brain tumor datasets from a variety of sources in light of these similarities. Apart from comparing our findings to the same data set, we compared the results of this analysis to a variety of other datasets. Despite the use of diverse datasets, it can be concluded that deep learning is generally superior to machine learning for classification.

7. Conclusions

This article introduced a new approach called en-CNN that is based on VGG-16 but has a more straightforward architecture for classifying brain tumors visible in MRI images. Seven convolutional layers, four ReLU layers, and four Maxpooling layers comprise the en-CNN. Additionally, we used an optimizer to tune the hyperparameters, which included a dropout layer, followed by completely connected and sigmoid layers to predict the class performance.

Although the dataset was insufficiently large, the new approach achieved a high level of accuracy. Additionally, hyperparameter tuning was used to achieve a 97 % accuracy value with the ADAM optimizer for the FLAIR series, which is greater than the T1, T1CE, and T2 sequences in the same dataset, using the values of 200 epoch, 128 mini-batch size, and 0.1 dropouts.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization, Hapsari Peni; methodology, Hapsari Peni; software, Dewinda Julianensi Rumala and Nurul Zainal Fanani; validation, Anggraini Dwi Sensusiati; Formal analysis, Joan Santoso; writing—original draft preparation, Hapsari Peni; writing—review and editing, I Ketut Eddy Purnama and Peter M.A van Ooijen; visualization, Cucun Very Angkoso; supervision, Mauridhi Hery Purnomo; Funding acquisition, I Ketut Eddy Purnama.

Acknowledgments

This research has been sponsored by the Indonesian Education Scholarship (BPPDN) and partially funded by Lembaga Pengelola Dana Pendidikan under the scheme of Riset Inovatif Produktif (RISPRO) - Invitasi 2019 Grant, contract number: PRJ-41/LPDP/2019.

References

[1] Q. T. Ostrom, G. Cioffi, H. Gittleman, N. Patil, K. Waite, C. Kruchko, and J. S. Barnholtz-Sloan, "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016", Neuro-Oncology, vol. 21, no. S5. Oxford University Press, pp. V1–V100, 01-Oct-2019.

[2] D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, P. C. Burger, A. Jouvet, B. W. Scheithauer, and P. Kleihues, "The 2007 WHO classification of tumours of the central nervous system", Acta Neuropathologica, Vol. 114, No. 2. Springer, pp. 97–109, Aug-2007.

[3] L. M. DeAngelis, "Medical progress: Brain tumors", N. Engl. J. Med., Vol. 344, No. 2, pp. 114–123, 2001.

[4] G. Mohan and M. M. Subashini, "MRI based medical image analysis: Survey on brain tumor grade classification", Biomed. Signal Process. Control, Vol. 39, pp. 139–161, 2018.

[5] J. E. Villanueva-Meyer, M. C. Mabray, and S. Cha, "Current clinical brain tumor imaging", Clin. Neurosurg., Vol. 81, No. 3, pp. 397–415, Sep. 2017.

[6] S. C. Thust, M. J. van den Bent, and M. Smits, "Pseudoprogression of brain tumors", J. Magn.

Reson. Imaging, Vol. 48, No. 3, pp. 571–589, Sep. 2018.

[7] P. Thanh Noi and M. Kappas, “Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery”, *Sensors (Basel).*, Vol. 18, No. 1, Dec. 2017.

[8] W. L. Al-Yaseen, A. Jehad, Q. A. Abed, and A. K. Idrees, “The Use of Modified K-Means Algorithm to Enhance the Performance of Support Vector Machine in Classifying Breast Cancer”, *International Journal of Intelligent Engineering and Systems*, Vol. 14, No. 2, pp. 190-200, 2021.

[9] M. Alfonse and A.-B. M. Salem, “An Automatic Classification of Brain Tumors through MRI Using Support Vector Machine”, 2016.

[10] L. Zhang, L. Li, M. Tang, Y. Huan, X. Zhang, and X. Zhe, “A new approach to diagnosing prostate cancer through magnetic resonance imaging,” *Alexandria Eng. J.*, Oct. 2020.

[11] H. P. Bahareh Shahangian, “Automatic brain hemorrhage segmentation and classification algorithm based on weighted grayscale histogram feature in a hierarchical classification structure”, *Biocybern. Biomed. Eng.*, Vol. 36, No. 1, pp. 217–232, Jan. 2016.

[12] L. Deng and D. Yu, “Deep learning: Methods and applications”, *Found. Trends Signal Process.*, Vol. 7, No. 3–4, pp. 197–387, 2013.

[13] N. B. Bahadure, A. K. Ray, and H. P. Thethi, “Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM”, *Int. J. Biomed. Imaging*, Vol. 2017, pp. 1–12, Mar. 2017.

[14] P. Kumar and B. Vijay Kumar, “Brain tumor MRI segmentation and classification using ensemble classifier”, *Int. J. Recent Technol. Eng.*, Vol. 8, No. 1 Special Issue4, pp. 244–252, 2019.

[15] R. H. Ramdlon, E. Martiana Kusumaningtyas, and T. Karlita, “Brain Tumor Classification Using MRI Images with K-Nearest Neighbor Method”, in *IES 2019 - International Electronics Symposium: The Role of Techno-Intelligence in Creating an Open Energy System Towards Energy Democracy, Proceedings*, 2019, pp. 660–667.

[16] J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, “Brain tumor detection using statistical and machine learning method”, *Comput. Methods Programs Biomed.*, Vol. 177, pp. 69–79, Aug. 2019.

[17] J. Amin, M. Sharif, N. Gul, M. Yasmin, and S. A. Shad, “Brain tumor classification based on DWT fusion of MRI sequences using convolutional neural network”, *Pattern Recognit. Lett.*, Vol. 129, pp. 115–122, Jan. 2020.

[18] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, “Multi-grade brain tumor classification using deep CNN with extensive data augmentation”, *J. Comput. Sci.*, Vol. 30, pp. 174–182, Jan. 2019.

[19] S. Vijh, S. Sharma, and P. Gaurav, “Brain tumor segmentation using OTSU embedded adaptive particle swarm optimization method and convolutional neural network”, in *Lecture Notes on Data Engineering and Communications Technologies*, Vol. 32, Springer, 2020, pp. 171–194.

[20] M. I. Sharif, J. P. Li, M. A. Khan, and M. A. Saleem, “Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images”, *Pattern Recognit. Lett.*, Vol. 129, pp. 181–189, Jan. 2020.

[21] N. Ghassemi, A. Shoeibi, and M. Rouhani, “Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images”, *Biomed. Signal Process. Control*, Vol. 57, p. 101678, Mar. 2020.

[22] M. A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S. A. C. Bukhari., “Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists”, *Diagnostics*, Vol. 10, No. 8, Aug. 2020.

[23] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M. A. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, Ç. Demiralp, C. R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharuddin, R. Jena, N. M. John, E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H. C. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, and K. Van Leemput, “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”, *IEEE Trans. Med. Imaging*, Vol. 34, No. 10, pp. 1993–

2024, Oct. 2015.

[24] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Learning features for offline handwritten signature verification using deep convolutional neural networks", *Pattern Recognit.*, Vol. 70, pp. 163–176, Oct. 2017.

[25] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-Classification of Brain Tumor Images Using Deep Neural Network", *IEEE Access*, Vol. 7, pp. 69215–69225, 2019.

[26] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification", *Neurocomputing*, Vol. 321, pp. 321–331, Dec. 2018.

[27] J. Nalepa, M. Marcinkiewicz, and M. Kawulok, "Data Augmentation for Brain-Tumor Segmentation: A Review", *Front. Comput. Neurosci.*, Vol. 13, p. 83, Dec. 2019.

[28] M. M. Badža and M. C. Barjaktarović, "Classification of brain tumors from mri images using a convolutional neural network", *Appl. Sci.*, Vol. 10, No. 6, 2020.

[29] D. P. Kingma and J. Lei Ba, "Adam: A method for stochastic optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp 1-15, 2015.

[30] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. S. Kruthiventi, and R. V. Babu, "An Introduction to Deep Convolutional Neural Nets for Computer Vision," in *Deep Learning for Medical Image Analysis*, Elsevier Inc., 2017, pp. 25–52.

[31] T. Triwiyanto, I. P. A. Pawana, and M. H. Purnomo, "An Improved Performance of Deep Learning Based on Convolution Neural Network to Classify the Hand Motion by Evaluating Hyper Parameter", *IEEE Trans. Neural Syst. Rehabil. Eng.*, Vol. 28, No. 7, pp. 1678–1688, Jul. 2020.