Incremental Dilations Using CNN for Brain Tumor Classification

Sanjiban Sekhar Roy [1,2,3,†] [ID], **Nishant Rodrigues** [1,†] **and Y-h. Taguchi** [4,*] [ID]

1   School of Computer Science and Engineering, Vellore Institute of Technology, Vellore 632014, India; sanjibanroy09@gmail.com (S.S.R.); nishantrodrigues97@gmail.com (N.R.)
2   Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City 758307, Vietnam
3   Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City 758307, Vietnam
4   Department of Physics, Chuo University, 1-13-27 Kasuga, Bukkyo-ku, Tokyo 112-8551, Japan
*   Correspondence: tag@granular.com
†   These authors contributed equally to this work.

Abstract: Brain tumor classification is a challenging task in the field of medical image processing. Technology has now enabled medical doctors to have additional aid for diagnosis. We aim to classify brain tumors using MRI images, which were collected from anonymous patients and artificial brain simulators. In this article, we carry out a comparative study between Simple Artificial Neural Networks with dropout, Basic Convolutional Neural Networks (CNN), and Dilated Convolutional Neural Networks. The experimental results shed light on the high classification performance (accuracy 97%) of Dilated CNN. On the other hand, Dilated CNN suffers from the gridding phenomenon. An incremental, even number dilation rate takes advantage of the reduced computational overhead and also overcomes the adverse effects of gridding. Comparative analysis between different combinations of dilation rates for the different convolution layers, help validate the results. The computational overhead in terms of efficiency for training the model to reach an acceptable threshold accuracy of 90% is another parameter to compare the model performance.

Keywords: brain tumor; classification; dilation; convolutional; deep learning

1. Introduction

Tumors are a mass of abnormal tissue that ascends without palpable cause from body cells and have no crucial function. The uncontrollable growth of cells results in an increment of the size of the tumor. Brain tumor detection at the early stage and availing proper treatment can save the patient from any adverse damage to the brain [1]. Recently, computer-assisted techniques such as using deep learning for feature extraction, and classification techniques are being used intensively to diagnose the patients' brains to check if there are any tumors. The introduction of information technology and the e-healthcare system in the area of medical diagnosis has assisted clinical professionals in offering considerably better health care for patients. Different classification techniques, especially convolutional neural networks, have been proposed in recent years [1–6] however, these proposed techniques have failed to acquire high accuracy. Therefore, there is a need to develop new techniques for the detection of brain tumor. In this article, we have proposed the classic problem of detecting tumors from MRI images using a dilated deep convolutional neural network (CNN). Simultaneously we have bench-marked the performance of the proposed model with those of existing models such as Artificial Neural Network (ANN) and Convolutional Neural Network (CNN).

In convolutional neural networks (CNN) [2] the receptive field is too small to result in high accuracy. The fixed size of the sliding window in CNN fails to take advantage of the very architecture

of the CNN, such as convolution, pooling, and flattening techniques, hence considering a sizeable receptive field of the convolution kernel would help to increase the accuracy of the classification techniques. The parameters in the proposed model have the capability of learning features from the images.

A CNN with $C \times C$ layers of convolution with the exclusion of pooling can be represented as $l * (C-1) + C$ which is nothing but linear increments of the receptive field with number of layers l. This linear growth restricts CNN's performance on input images. Let G is a discrete function such that [7] $G : Z^2 \rightarrow R$; where Z is the set of all natural numbers and R is the set of all real numbers. Also by assuming a range $\phi_r = [-r, r]^2 \cap Z^2$ and the discrete filter k that is mapped to the real numbers defined as $k : \phi_r \rightarrow R$ which is a filter of size $(2r + 1)^2$. The operator $*$ is known as convolution operator. The convolution operator $*$ is defined in Equation (1), where 1-D dilated convolution with dilation rate $l = l$ convolves input image F with kernel or filter k. This 1-D convolution is called standard convolutional neural network. When the value of the dilation l increases then the network is referred as dilated convolutional neural network.

$$(F * k)(p) = \sum_{s+t=p} F(s) \times k(t) \tag{1}$$

Now we can introduce a dilation factor called l, and by generalizing this factor l that can be defined as,

$$(F *_l k)(p) = \sum_{s+lt=p} F(s) \times k(t) \tag{2}$$

Here l is referred to as the dilation rate of the convolution neural network. The value of $l = 1$ is a basic convolutional neural network.

Researchers have developed deep learning and other techniques to detect brain tumors. Developing a model with a high accuracy is a challenging task. Recent version of CNN models [8–10] have hardly focused on hyper parameters whereas we do so; the collection [2] of features that are locally available to the CNN are also a critical issue; moreover bluntly increasing the dilation rate may add to the failure of feature collections due to the sparseness of the kernel, affecting small object detection [11]. High dilation rates may affect small object detection. Therefore, in our proposed model, a gradual increase in the dilation rate (even-numbered arithmetic progression) has been carried out. This has helped to decrease sparsity on the dilated feature map which is able to extract more information from the area under analysis. Hence keeping all these in mind, the main contribution of this work include,

1. We propose a dilated convolution neural network with even-numbered increments of the dilation rate for the brain tumor classification along with data preparation (image pre-processing, data augmentation) and hyper-parameter tuning.
2. We critically discuss with the help of the proposed experimental results on why a small receptive field of CNN with dilation rate causes the poor accuracy in brain tumor classification.
3. We carry out an in-depth analysis of how the architecture of proposed dilated convolution with an enlarged receptive field of the kernel improves the efficiency of computation while also maintaining a high accuracy.
4. We analyze the relationship between the rate of dilation and image classification accuracy.
5. We also carry out a detailed comparative study between basic CNN and ANN; in both cases, the proposed dilated neural network has surpassed the other two.

Related Work

Medical image analysis is a vast area of research, and many researchers have added to the vast variety of subfields of it [12]. We have taken a look at past work on brain tumor classification. The majority of the work that has been carried out is based on the automatic segmentation of brain

tumors from MRI images [13,14]. Post segmentation of the tumor, it needs to go through different gradations of the classification; however, in the early research studies [15–17] the classification strategy is primarily for benign and malignant tumors. Kharrat et al. [15], have introduced a genetic algorithm and support vector machine (SVM), whereas Abdolmaleki et al. [16] have proposed a here-level backpropagation neural network for tumor classification. These two proposed methods have obtained a classification accuracy of 91% and 94% respectively for the classification of malignant and benign tumors from MRI images of 165 patients. Papageorgiou et al. [17] have implemented a fuzzy cognitive map (FCM) for hundred instances and obtained 90.26% accuracy for low-grade tumors of the brain. In addition to that, the multigrade classification of brain tumors was conducted by Zacharaki et al. [18]. A computer-assisted diagnosis (CAD) model was also proposed by Hsieh et al. [19] this CAD system has been applied to the malignancy of gliomas of 107 high and low-grade MRI images. It has obtained an accuracy of about 83%. Other works have also been carried out such as Sachdeva et al. [20], Cheng et al. [21] and Afshar et al. [4] which propose methods such as CAD model with GA (SVM+ANN), Bag-of-words (BoW) method, and capsule networks (CapsNets), respectively. These three methods have achieved an accuracy greater than 90%. Very recently, Özyurt et al. [22] has introduced a state of the art machine learning and deep learning application that consists of a Fuzzy C-Means and CNN merged with Extreme Learning Machine. Recent models include a symmetric neural network [23] by Chen et al. (2019), CNN combined with neutrosophic expert maximum fuzzy sure entropy by Özyurt et al. [24], and a big-data model brain tumor detection using deep CNN proposed by Amin et al. [25]. Zia et al. [26] proposed a generic classification model for the same with the wavelet transform as a feature extraction plus a PCA (principal component analysis) and SVM method to reduce the dimensionality and classification task. This kind of binary classification is not enough for the radiologist to make a solid decision on treatment for the patients. However, the essential work of brain tumor classifications remains focused on the binary classification of brain tumors. Also, the scarcity of data is a concern relating to these kinds of work. Very recently, deep learning-based techniques [10,27] have been adopted to address these issues. Techniques such as transfer learning [28] have also been implemented to improve model performance. We, therefore, proposed a model based on dilated deep convolutional networks that is better at detecting brain tumors.

2. Materials and Methods

2.1. Proposed Methodology

Our proposed model (Figure 1) makes use of Dilated CNNs hence adding another hyper-parameter to the mix, the dilation rate. Dilation is implemented by introducing zeros between filter elements. Dilation allows the network to cover more relevant information by increasing the receptive field of the filters. The CNN is designed to extract the most information out of the images per convolution layer. In our case, applying a 3×3 convolution layer allows the network to capture more detailed characteristics as 3×3 is the smallest filter to capture left/right, up/down, and center from the image. Instead of using large convolution filters (used in the underlying CNN architectures) such as 5×5 to detect coarse features such as the shape and contours, we make use of dilations in the convolution layers, and this allows the model to detect such coarse features without the additional computational overhead of using larger filters such as 5×5.

To analyze the performance in comparison to Basic CNN and Simple ANN, we make use of the same CNN architecture as depicted in Figure 1, but without applying any dilations (set *dilation_rate* = 1) to the convolution layers. The simple ANN is made up of fully connected (dense) layers of artificial neurons. Layer 1 consists of 1024 units and a dropout of 50%, Layer 2 contains 512 units and a dropout of 25%. The third layer contains 128 units and a dropout rate of 25% and the next layer consists of 32 units with a dropout of 15%. All the layers make use of the *Rectified Linear Unit (ReLU)* activation function. The final layer consists of a single unit and uses the *Sigmoid* activation function.
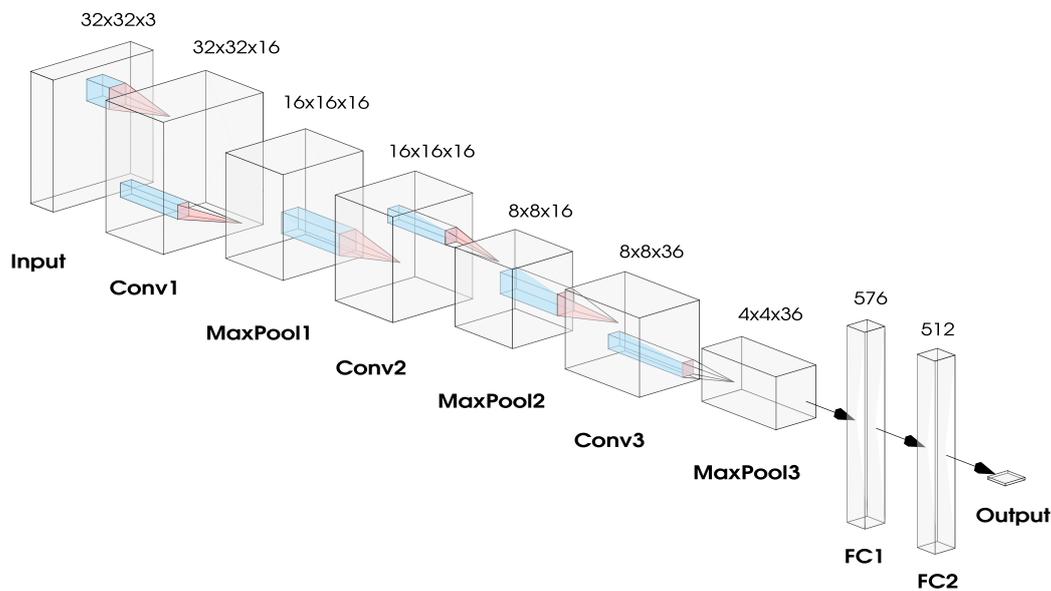
Figure 1. Proposed model architecture of Dilated CNN depicting the various Convolutional, Pooling, Flatten and Dense layers along with the shapes of their respective input and output tensors. The *Input* consists of a single image of shape 32×32 with 3 channels for colors RGB. Figure generated using NN-SVG [29].

2.1.1. Convolutional Layers

The proposed model architecture depicted in Figure 1 is trained on an input of RGB images. Each input tensor has a dimension of $(32, 32, 3)$. Three separate convolution layers use the same 3×3 filters, and the feature maps are generated wholly based on the dilation rate of each layer. The interior architecture is as simple as possible to test the effects of dilation rate on model performance and understand the gridding effect caused by the dilation technique. The layer *Conv1* generates 16 feature maps by applying a 3×3 filter and a dilation rate $d1$, layer *Conv2* generates 16 feature maps again by applying the same 3×3 filter and dilation rate $d2$ and the final convolution, and layer *Conv3* generates 36 feature maps using a 3×3 filter and dilation rate $d3$. The last convolution layer generates a more significant number of filters as the final layer must select more delicate features for higher-level reasoning for the upcoming layers. All three convolution layers make use of the *ReLU* activation function. The dilation rates $d1, d2$, and $d3$ are also used in the nomenclature for the Dilated CNNs such as $DilatedCNN(4, 2, 1)$ stands for a Dilated CNN model with dilation rates corresponding to $(d1 = 4, d2 = 2, d3 = 1)$.

2.1.2. Pooling Layers

The pooling layers are used to reduce the resolution of the generated feature maps. These layers are generally placed between convolution layers. To keep the model architecture simple and make the model more dependent on the dilation rate parameter, we have made use of simple *MaxPooling* layers with a pool size of 2×2. The three pooling layers namely *MaxPool1, MaxPool2* and *MaxPool3* depicted in Figure 1, all use the same pool size of 2×2.

2.1.3. Flattening and Dense Layers

Once the feature maps are generated, the model needs to be trained on high level reasoning. The feature maps are flattened into a 1 dimensional vector of size (576). A fully connected layer of shape (512) is added along with a dropout of 15% of the nodes. The dense layers can easily get biased

and dropout prevents the model from overfiitting on the dataset. To provide non linearity to the results, the *ReLU* activation function is used. Figure 1 shows the 2 fully connected layers *FC1* and *FC2* along with the output layer. For the final output, a dense layer of shape (1) is used with the *Sigmoid* activation function. Again to prevent overfitting, 15% of the nodes are dropped out from the previous layer. The model is trained to reduce the binary cross entropy loss with the help of the *Adam optimizer* [30].

2.1.4. Activation Functions

All the hidden layers use the *ReLU* activation function. The *ReLU* activation function provides more sensitivity to the activation sum and avoids easy saturation. It looks and acts like a linear function, but in fact is a nonlinear function allowing the network to learn complex non linear relationships. It is a piecewise linear function that is linear for half of the input domain and non linear for the other half of the domain. The final layer predicts the output of the binary classification. *Sigmoid* is chosen as it has several advantages over *Step* function and *Tanh* activation function. The *Sigmoid* function has a characteristic "S-shaped" curve. It bounds the values within the range $[0, 1]$ and allows for smoother training compared to using the *Step* function and helps to prevent bias in the gradients. *Tanh* is a rescaled logistic Sigmoid function such that its outputs range from $[-1, 1]$ and is centered around 0. For the final layer in a binary classification problem, *Sigmoid* activation function provides a softer gradient when compared to *Tanh*.

2.2. Dataset

No humans were directly involved in this study, the data is anonymous or generated synthetically from simulators. We have selected slices from various MRI scans and applied necessary preprocessing techniques to convert the images into a common JPEG format to maintain consistency throughout the dataset. The data is split into two categories namely "Normal" and "Tumour". The model is trained and tested on images curated from a number of publicly available sources. Kaggle provides an open dataset curated and maintained by Chakrabarty [31] which collects MRI images in 2 folders (tumor detected—"yes" and "no"), containing a total of 253 Brain MRI images. As MRI images contain personal information and require the assistance of specialized doctors for labeling, we have also made use of simulated brain images. The Brain wave [32–35] brain simulator provides a 3D simulation of the brain based on a range of user-defined parameters. The data is in 3D slices, and we can select a particular series of slices (top-down view) to add to the dataset. As these simulations are based on the anatomical model of a healthy brain, they serve as the ground truth for any analysis procedure. Another resource used is the Harvard Brain simulator [36], which provides many simulated brain MRI images that have been carefully selected and added to the dataset.

The next step consists of image pre-processing. We aim to remove additional data present around the main MRI brain scan, making sure that all the images are of the same type, and the focus is only on the central part of the brain. To carry out the mentioned preprocessing, we have used a relatively common method of using the extreme points of a contour. The simple step by step approach illustrated in Figure 2, combined with a few image processing methods such as converting the image to Grayscale, Thresholding, and Opening (Erosion followed by Dilation), as mentioned below in Algorithm 1, makes sure that the brain is in focus in each image. Finally, using the extreme points as a mask, it is a fairly simple task to crop out the parts of the image that do not add any value to the classifier model.

When preparing the data for training, we need to create a generalized dataset as deep learning algorithms are highly data-driven. This means that imbalanced sets and other skewed image properties in a particular class will create a bias in the model and results in improper classification. The Brain MRI dataset suffers from two main issues, firstly the size of the dataset and secondly, there is no single correct structural shape of a human brain. Each human brain is shaped uniquely and slightly different from the other. Using the Keras Image Data Generator [37], we augment the images over a set of parameters, and this adds a slight degree of randomness to the images. In terms of the entire training process, the model learns on a generalized dataset. The applied augmentation techniques include

Rescaling the image between a range of $[0, 1]$, *Rotation* randomly between $[-15°, +15°]$, *Shifts* in height and width to a maximum 10% of image dimensions, *Shear* range of 0.1 and a *Brightness* range between the bounds $[0.5, 1.5]$. As the brain MRI images are vertically aligned, it is not feasible to use a vertical flip. Preferably a *Horizontal flip* is used to create symmetrical data about the vertical axis. Figure 3 depicts the degree of randomness introduced by using the above mentioned image augmentation on a single sample MRI image.


Algorithm 1 Image Pre-processing.

Input   Raw Image

Output Cropped Image

1: **for** *iteration* $= 1, 2, \ldots, N_{images}$ **do**
2:     $Image_{Gray}$ $\leftarrow Grayscale(Image_{Input})$
3:     $Image_{Binary} \leftarrow Binary(Image_{Gray}, threshold[45–255])$
4:     ***Opening***
5:       $Image_{Binary} \leftarrow Erosion(Image_{Binary})$
6:       $Image_{Binary} \leftarrow Dilation(Image_{Binary})$
7:     $Image_{Contour} \leftarrow Contour(Image_{Binary})$
8:     $Image_{Mask}$ $\leftarrow Extremes(Image_{Contour})$
9:     $Image_{Output}$ $\leftarrow Crop(Image_{Binary}, Image_{Mask})$
10: **end for**
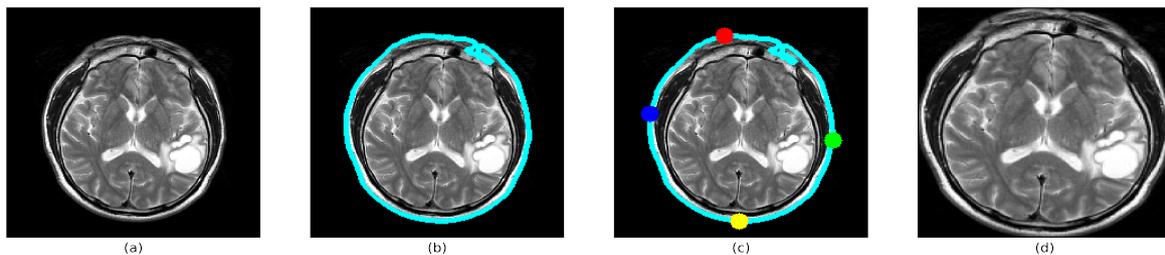
Figure 2. (**a**) Original Image: Section of brain from a MRI scan (**b**) Finding largest contour: Detecting the overall shape of the skull structure. (**c**) Calculating extreme points of contour: Selecting the best points to fit the entire brain into the frame with minimum loss of data (**d**) Cropping based on extreme points: The brain structure from the MRI image is the main focus area of the new image.
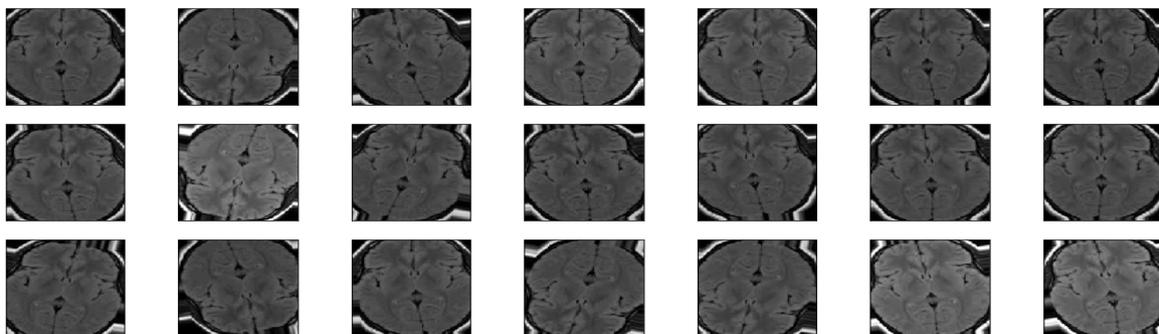
Figure 3. Augmented images: Each of the 21 displayed images are generated from the same input image by introducing slight variations such as rotation, sheer, zoom, brightness, shift and horizontal flip with the aim of building a generalized and robust dataset.

After preparing the data and applying the mentioned augmentation techniques, the data is split into training and validation sets. Each model is trained using backpropagation on the training set and after each *Epoch* (one iteration through the complete training set), the validation accuracy is calculated.

Using the validation accuracy, checkpoints are created. These checkpoints store the best weights that can later be used for model inferencing or for further training. The step by step flow, is mentioned in Algorithm 2.

Algorithm 2 Dilated CNN Classifier Algorithm.

    Input　Image Dataset

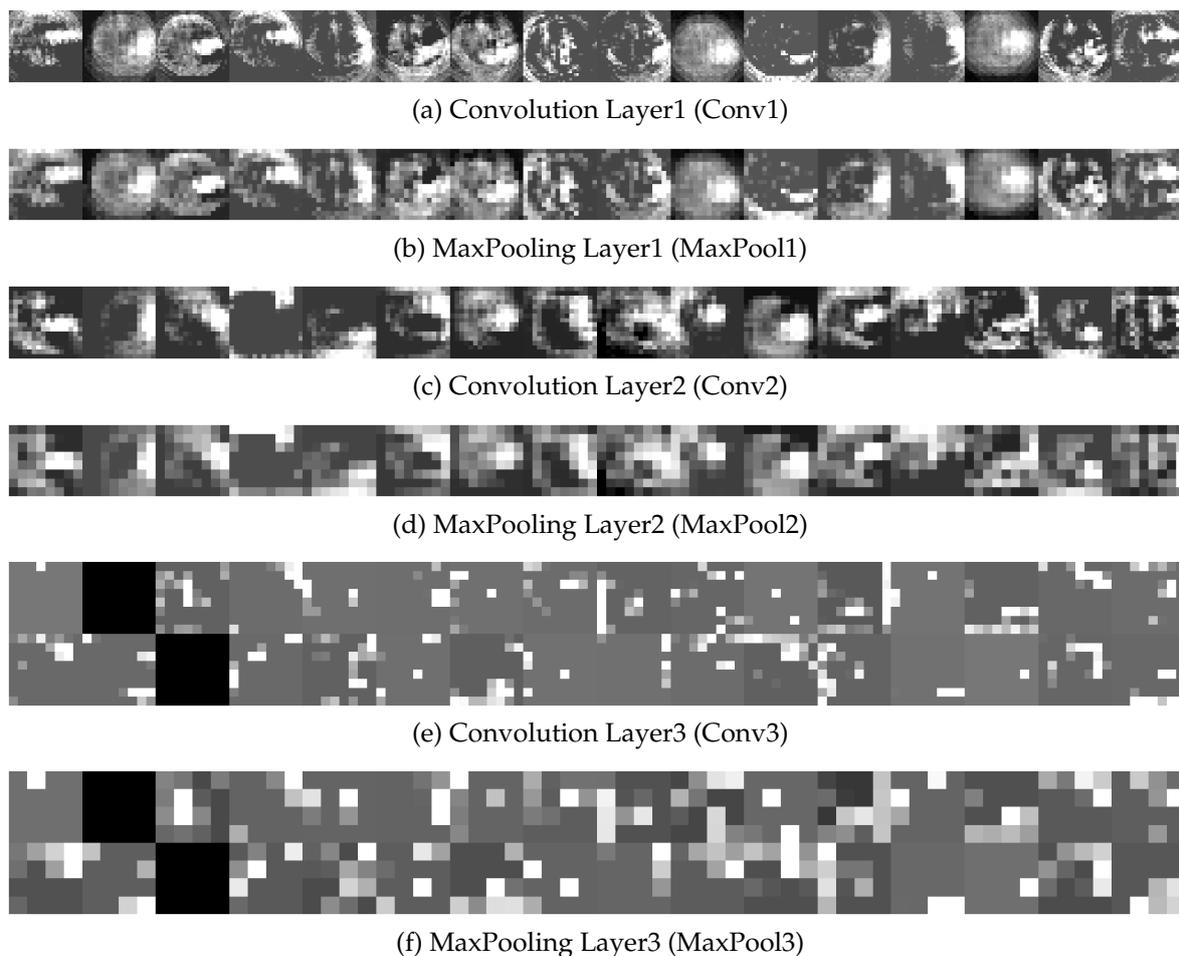   **Output** Trained Brain Tumor Classifier

1: MODEL TRAINING
2: Randomly initialize neural network with random weights
3: $Accuracy \leftarrow 0$
4: **for** $epoch = 1, 2, \ldots, N_{epochs}$ **do**
5:    **for** $image = 1, 2, \ldots, N_{batch\_size}$ **do**
6:       $Image_{data} \leftarrow Resize(Image)$
7:       $Image_{data} \leftarrow Augmentation(Image_{data})$
8:       Input layer $u(t)$ takes $Image_{data}$ and sends it to the hidden layers
9:       ***Hidden layers***
10:          Large *dilation_rate* Coarse features
11:          Small *dilation_rate* Finer features
12:       Output layer $w(t)$ returns diagnosed results
13:       Calculate error rate $e(t)$
14:       Update weights using *back_propagation*
15:       Using *Adam_Optimizer* to minimize $e(t)$
16:    **end for**
17:    MODEL CHECK-POINTING
18:    Using Validation set, Calculate *Validation_Accuracy*
19:    **if** *Validation_Accuracy* $\geq$ *Accuracy* **then**
20:       Checkpoint model, Save *Weights*
21:       *Best_Weights* $\leftarrow$ *Weights*
22:    **end if**
23: **end for**

---

3. Results

Once the model is trained, the best checkpoint is selected for model inferencing. The predicted classes are compared to the actual target classes to calculate the model accuracy, precision, recall and f-measure. Once the model has crossed the threshold accuracy of 90%, we try to understand the inner workings of the various layers of the model. Feature maps are generated for various input images. Feature maps help to determine the active areas of the image, i.e., The highlighted areas of the image that contribute to the classification decision. The activation maps for the various convolution and pooling layers are illustrated in Figure 4.

From the feature maps (Figure 4), we can deduce that the dilated CNN works on a well-described top-down approach. The outer layers (layer 1 and layer 2) focus on coarse features such as the shape of the brain or any problem areas (outliers–tumor locations). As we move through the layers, the granularity of the features decreases. The last layer generates features of fine granularity, and this means the focus is now on tiny sections of the image, that can determine small tumors. Finally, these generated feature maps are used to classify the MRI image using a Sigmoid activation function (binary classification).

(a) Convolution Layer1 (Conv1)

(b) MaxPooling Layer1 (MaxPool1)

(c) Convolution Layer2 (Conv2)

(d) MaxPooling Layer2 (MaxPool2)

(e) Convolution Layer3 (Conv3)


(f) MaxPooling Layer3 (MaxPool3)

Figure 4. Activation maps for various layers of the Dilated CNN. Each of the three convolutional layers depicts the granularity of the generated features with coarse features generated in Conv1 and fine features in Conv3. The MaxPooling layers reduce the resolution of the generated features using a 2×2 kernel.

Figure 5 shows the number of epochs versus classification accuracy graph for Dilated CNN, Basic CNN, and Simple ANN with Dropout. We can deduce from the graph that Dilated CNN has performed better than Basic CNN and Simple ANN with Dropout. Figure 5 clearly describes the power and efficiency of using Dilated CNN over Simple ANN and Basic CNN. The Dilated CNN model achieves the threshold accuracy of 90% within 10 epochs and an accuracy of 95% after 30 epochs. The model achieves a maximum accuracy of 97% after 50 epochs. The Basic CNN, on the other hand, is comparatively less effective in classifying brain tumors. The ANN model fails to achieve the set threshold of 90%. The model achieves an accuracy of 85%, after which it fails to improve further. Similarly, Figure 6 provides a comparative analysis of using different dilation rates for the various convolution layers. Using high dilation rates results in the gridding phenomenon that prevents the model from learning from finer features resulting in the low accuracy of the $(6, 6, 6)$ model. The best performing model is the $(4, 2, 1)$ model. Using these dilation rates allows the model to learn from the coarse features as well as the finer features while overcoming the additional computational overhead of using larger convolutional filters.
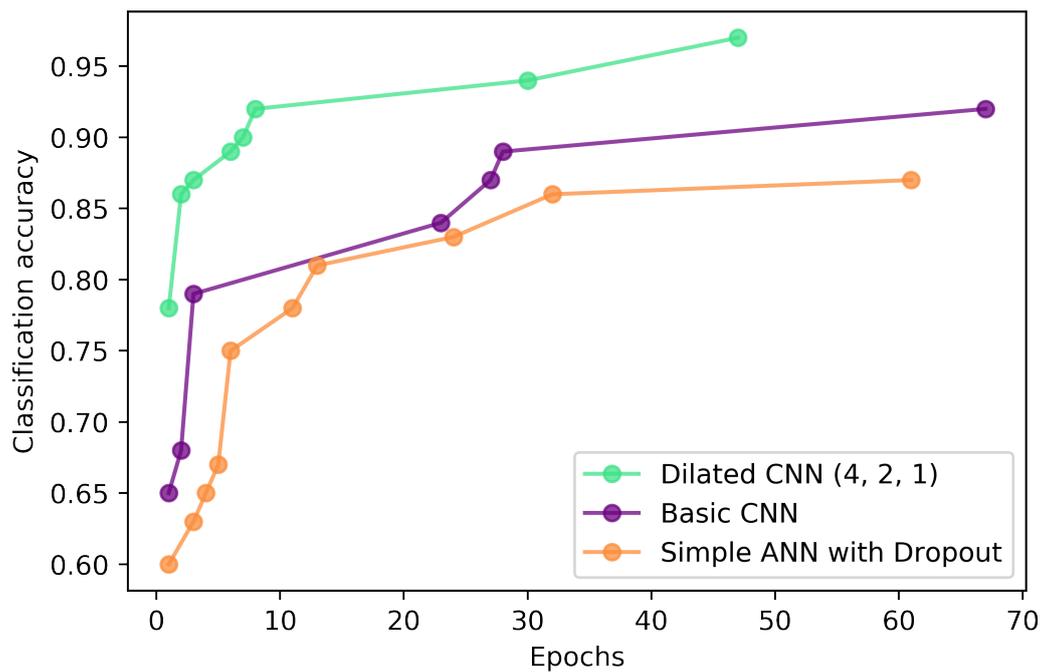
Figure 5. Performance comparison between the best Dilated CNN, Basic CNN and Simple ANN with Dropout architectures with respect to classification accuracy.
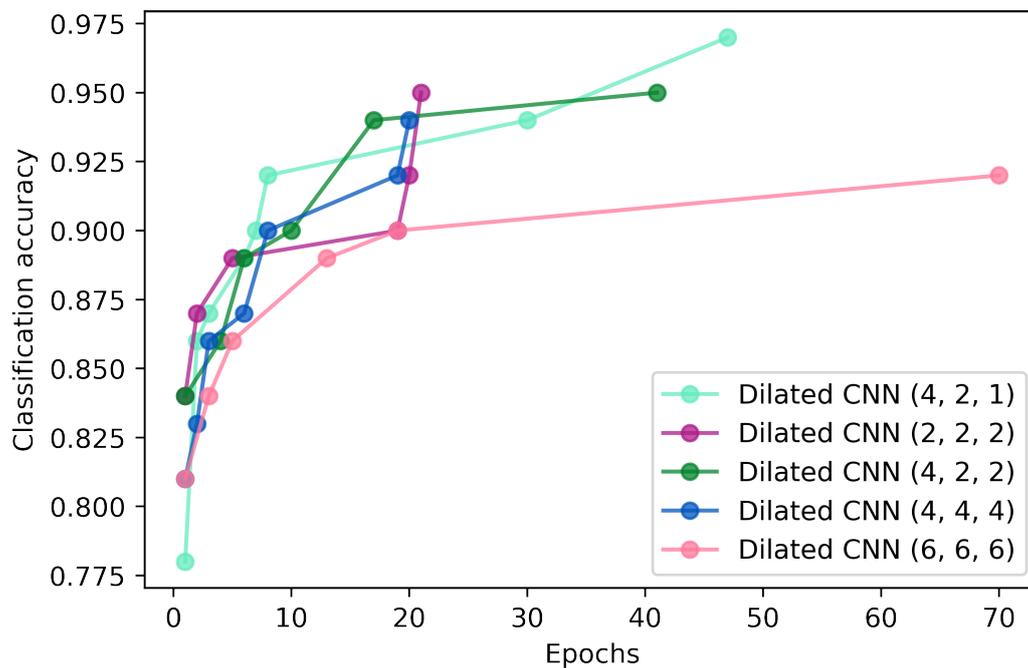
Figure 6. Comparative analysis of various dilation rate parameters based on the maximum validation (classification) accuracy achieved during training.

Comparative analysis between the various architecture is presented in Table 1. Analysis of the various classification metrics, namely the True Positive (TP) i.e., when the model correctly predicts the positive class, False Negative (FN) i.e., when the model wrongly predicts the negative class, False Positive (FP) i.e., when the model wrongly predicts the positive class, and True Negative (TN) i.e., when the model correctly predicts the negative class. These core metrics make up the confusion matrix and determine the core model performance. The CNN architectures outperform the Simple ANN

model. On further analysis, dilated CNN has a better False Positive (FP) rate compared to Basic CNN, and this is an essential factor when dealing with medical diagnosis. The Table 2 provides an in-depth analysis of the models' precision and recall metrics. Comparing the Basic CNN and the Dilated CNN, we can determine that the dilated CNN model has better precision compared to the Basic CNN model.

Table 1. True Positive, False Negative, False Positive and True Negative rates compared across Simple ANN, Basic CNN and Dilated CNN.

Model	TP	FN	FP	TN
ANN	0.34920635	0.04761905	0.07936508	0.52380952
Basic CNN	0.38095238	0.01587302	0.06349206	0.53968254
Dilated CNN (4, 2, 1)	0.38095238	0.01587302	0.01587302	0.58730159

Table 2. Accuracy, Precision, Recall and F-measure compared across Simple ANN, Basic CNN and Dilated CNN.

Model	Accuracy	Precision	Recall	F-Measure
ANN	0.87301587	0.81481481	0.88	0.84615384
Basic CNN	0.92063492	0.85714285	0.96	0.90566037
Dilated CNN (4, 2, 1)	0.96825397	0.96	0.96	0.96

The Table 3 displays a comparative side by side analysis of the various classification metrics, namely the True Positive, False Negative, False Positive, and True Negative rates. The above Table 4 provides an in-depth analysis of the gridding phenomenon and the effects of the various dilation rates. Using high dilation rates, the model cannot learn from the finer features, similarly using a low dilation rate, the model does not pick up on the coarse features. A well-balanced model should be able to learn both the coarse and fine features of the images. Using the $(4, 2, 1)$ model provides the best-case scenario.

Table 3. Comparing True Positive, False Negative, False Positive, True Negative rates across different configuration of dilation rates for the convolution layers.

Model	TP	FN	FP	TN
Dilated CNN (2, 2, 2)	0.38095238	0.01587302	0.03174603	0.57142857
Dilated CNN (4, 4, 4)	0.3968254	0.0	0.06349206	0.53968254
Dilated CNN (6, 6, 6)	0.34920635	0.04761905	0.03174603	0.57142857
Dilated CNN (4, 2, 1)	0.38095238	0.01587302	0.01587302	0.58730159
Dilated CNN (4, 2, 2)	0.36507937	0.03174603	0.01587302	0.58730159

Table 4. Comparing Accuracy, Precision, Racall and F-Measure across different configurations of dilation rates for the convolution layers.

Model	Accuracy	Precision	Recall	F-Measure
Dilated CNN (2, 2, 2)	0.95238095	0.92307692	0.96	0.94117647
Dilated CNN (4, 4, 4)	0.93650794	0.86206897	0.96	0.92592592
Dilated CNN (6, 6, 6)	0.92063492	0.91666667	0.88	0.89795918
Dilated CNN (4, 2, 1)	0.96825397	0.96	0.96	0.96
Dilated CNN (4, 2, 2)	0.95238095	0.95833334	0.92	0.93877551

The TPR (true positive rate) and FPR (false positive rate) are important AUC/ROC (Area Under The Curve/Receiver Operating Characteristics) [38] metrics that help to determine the amount of information learnt by the model and how well it is able to distinguish between the classes. In the ideal case, $TPR = 1$ and the $FPR = 0$. Table 5a compares the metrics across the various architectures and determines the Dilated CNN architecture with $TPR = 0.96$ and $FPR = 0.03$ is the best. Table 5b

compares the AUC/ROC metrics for the different dilation rates. Using an incremental dilation rate allows the model to learn the coarse as well as fine features, resulting in the maximum amount of information learnt by the model.

To compare the model performance in terms of computational resources required, we have designed the three models with a similar architecture in mind. This allows us to have a reasonably accurate analysis of the compute overhead (additional time required to setup the network architecture and data loaders) and efficiency of each network. The time taken for a single epoch (1 min 30 s approx.) is almost the same for each of the three networks. Using this as a benchmark, we can determine the computational effort required to achieve the threshold accuracy of 90%. Table 6a compares ANN, Basic CNN and Dilated CNN. As the dilated CNN requires the minimum time to achieve the threshold accuracy, it is set as the benchmark x and the performance of the other models is determined as a factor of x. As the ANN does not achieve the threshold accuracy, its performance cannot be determined. Table 6b shows the same analysis, for different combinations of dilation rates. The $(4, 2, 1)$ model (increasing dilation rates) performs the best and is selected as the benchmark. Other models with moderate dilation rates come in close behind. Using a small dilation rate $d = 2$ or large dilation rate $d = 6$ causes the gridding phenomenon, resulting in the models requiring additional computational effort to achieve the threshold.

Table 5. TPR (True Positive Rate) and FPR (False Positive Rate) measures compared across (**a**) Simple ANN, Basic CNN and Dilated CNN. (**b**) combinations of dilation rates for the different convolution layers.

Model	TPR	FPR
ANN	0.88	0.13157895
Basic CNN	0.96	0.1052632
Dilated CNN (4, 2, 1)	0.96	0.02631579
(**a**)		
Dilated CNN (2, 2, 2)	0.96	0.05263158
Dilated CNN (4, 4, 4)	0.96	0.10526316
Dilated CNN (6, 6, 6)	0.88	0.05263158
Dilated CNN (4, 2, 1)	0.96	0.02631579
Dilated CNN (4, 2, 2)	0.92	0.02631579
(**b**)		

Table 6. Epoch and comparative performance to determine the computational effort required for various the models to reach an acceptable accuracy of above 90%. x is the factor for the minimum time taken to achieve threshold accuracy (**a**) Simple ANN, Basic CNN and Dilated CNN. (**b**) combinations of dilation rates for the different convolution layers.

Model	Epochs	Performance
ANN	-	∞
Basic CNN	67	$9.57x$
Dilated CNN (4, 2, 1)	7	$1x$
(**a**)		
Dilated CNN (2, 2, 2)	19	$2.714x$
Dilated CNN (4, 4, 4)	8	$1.413x$
Dilated CNN (6, 6, 6)	19	$2.714x$
Dilated CNN (4, 2, 1)	7	$1x$
Dilated CNN (4, 2, 2)	10	$1.429x$
(**b**)		

4. Discussion

The primary purpose of this paper is to demonstrate the potential of Dilated CNN in comparison to other deep learning architectures, namely simple ANN with Dropout and Basic CNN in brain tumor detection. The paper has analyzed two aspects of the architecture of the models that is the classification accuracy and the computational resources required. For Dilated CNNs, we also have analyzed the effects of various dilation rates on the performance of the model. The classification accuracy is the highest (97%) for Dilated CNN (4, 2, 1), which has incremental, even number dilation rates, followed by Basic CNN. Simple ANN failed to break the threshold accuracy of 90%. In the case of the computing effort required by the models to attain a testing accuracy of more than 90%, the Dilated CNN outperformed the Basic CNN architecture by a considerable margin of 9.57 times where as ANN failed to achieve the threshold accuracy.

Finally, from the understanding of the gridding phenomenon and various values for the dilation rate parameter for each layer, the comparative study shows that an incremental dilation rate (4, 2, 1) provides the best results. Using dilation rates of (4, 2, 1) the model achieves an accuracy of 96.8% whereas the next best model (tied for dilation rates (2, 2, 2) and (4, 2, 2)) achieves an accuracy of 95.2%. This confirms the fact that the outer layers (higher dilation rates) focus on the coarse features, and the inner layers (lower dilation rates) learn from the finer features. This combination provides the best results. For future works, the experimental analysis can be carried out on other datasets to get a deeper understanding of the inner working of the network as well as the effectiveness of the dilation rate parameter.

References

1. Dai, Y.; Zhuang, P. Compressed sensing MRI via a multi-scale dilated residual convolution network. *Magn. Reson. Imaging* **2019**, *63*, 93–104. [CrossRef]
2. Lin, G.; Wu, Q.; Qiu, L.; Huang, X. Image super-resolution using a dilated convolutional neural network. *Neurocomputing* **2018**, *275*, 1219–1230. [CrossRef]
3. Mohsen, H.; El-Dahshan, E.S.A.; El-Horbaty, E.S.M.; Salem, A.B.M. Classification using deep learning neural networks for brain tumors. *Future Comput. Inform. J.* **2018**, *3*, 68–71. [CrossRef]
4. Afshar, P.; Mohammadi, A.; Plataniotis, K.N. Brain Tumor Type Classification via Capsule Networks. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 3129–3133.
5. Lakshmanaprabu, S.K.; Mohanty, S.N.; Shankar, K.; Arunkumar, N.; Ramirez, G. Optimal deep learning model for classification of lung cancer on CT images. *Future Gener. Comput. Syst.* **2019**, *92*, 374–382. [CrossRef]
6. Hemanth, D.J.; Anitha, J.; Naaji, A.; Geman, O.; Popescu, D.E.; Hoang Son, L. A Modified Deep Convolutional Neural Network for Abnormal Brain Image Classification. *IEEE Access* **2019**, *7*, 4275–4283. [CrossRef]
7. Yu, F.; Koltun, V. Multi-Scale Context Aggregation by Dilated Convolutions. *arXiv* **2016**, arXiv:1511.07122.
8. Rajinikanth, V.; Raj, A.N.J.; Thanaraj, K.P.; Naik, G.R. A Customized VGG19 Network with Concatenation of Deep and Handcrafted Features for Brain Tumor Detection. *Appl. Sci.* **2020**, *10*, 3429. [CrossRef]
9. Rehman, H.Z.U.; Hwang, H.; Lee, S. Conventional and Deep Learning Methods for Skull Stripping in Brain MRI. *Appl. Sci.* **2020**, *10*, 1773. [CrossRef]
10. Badža, M.M.; Barjaktarović, M.Č. Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 1999. [CrossRef]

11. Afshar, P.; Plataniotis, K.N.; Mohammadi, A. Capsule Networks for Brain Tumor Classification Based on MRI Images and Coarse Tumor Boundaries. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2018; pp. 1368–1372.

12. Wang, S.H.; Muhammad, K.; Phillips, P.; Dong, Z.; Zhang, Y.D. Ductal carcinoma in situ detection in breast thermography by extreme learning machine and combination of statistical measure and fractal dimension. *J. Ambient. Intell. Humaniz. Comput.* **2017**, 1–11. [CrossRef]

13. Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.M.; Larochelle, H. Brain tumor segmentation with Deep Neural Networks. *Med Image Anal.* **2017**, *35*, 18–31. [CrossRef] [PubMed]

14. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans. Med Imaging* **2015**, *34*, 1993–2024. [CrossRef] [PubMed]

15. Kharrat, A.; Garim, K.; Messaoud, M.B.; Benamrane, N.; Mohamed, A. A Hybrid Approach for Automatic Classification of Brain MRI Using Genetic Algorithm and Support Vector Machine. *Leonardo J. Sci.* **2010**, *17*, 71–82.

16. Abdolmaleki, P.; Mihara, F.; Masuda, K.; Buadu, L.D. Neural networks analysis of astrocytic gliomas from MRI appearances. *Cancer Lett.* **1997**, *118*, 69–78. [CrossRef]

17. Papageorgiou, E.; Spyridonos, P.; Glotsos, D.; Stylios, C.; Ravazoula, P.; Nikiforidis, G.; Groumpos, P. Brain tumor characterization using the soft computing technique of fuzzy cognitive maps. *Appl. Soft Comput.* **2008**, *8*, 820–828. [CrossRef]

18. Zacharaki, E.; Wang, S.; Chawla, S.; Yoo, D.; Wolf, R.; Melhem, E.; Davatzikos, C. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn. Reson. Med.* **2009**, *62*, 1609–1618. [CrossRef]

19. Hsieh, K.L.C.; Lo, C.M.; Hsiao, C.J. Computer-aided grading of gliomas based on local and global MRI features. *Comput. Methods Programs Biomed.* **2017**, *139*, 31–38. [CrossRef] [PubMed]

20. Sachdeva, J.; Kumar, V.; Gupta, I.; Khandelwal, N.; Ahuja, C.K. Segmentation, Feature Extraction, and Multiclass Brain Tumor Classification. *J. Digit. Imaging* **2013**, *26*, 1141–1150. [CrossRef]

21. Cheng, J.; Huang, W.; Cao, S.; Yang, R.; Yang, W.; Yun, Z.; Wang, Z.; Feng, Q. Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. *PLoS ONE* **2015**, *10*, e0140381. [CrossRef]

22. Özyurt, F.; Sert, E.; Avcı, D. An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine. *Med Hypotheses* **2020**, *134*, 109433. [CrossRef]

23. Chen, H.; Qin, Z.; Ding, Y.; Tian, L.; Qin, Z. Brain tumor segmentation with deep convolutional symmetric neural network. *Neurocomputing* **2019**. [CrossRef]

24. Özyurt, F.; Sert, E.; Avci, E.; Dogantekin, E. Brain Tumor Detection Based on Convolutional Neural Network with Neutrosophic Expert Maximum Fuzzy Sure Entropy. *Measurement* **2019**, *147*, 106830. [CrossRef]

25. Amin, J.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Big data analysis for brain tumor detection: Deep convolutional neural networks. *Future Gener. Comput. Syst.* **2018**, *87*, 290–297. [CrossRef]

26. Zia, R.; Akhtar, P.; Aziz, A. A new rectangular window based image cropping method for generalization of brain neoplasm classification systems. *Int. J. Imaging Syst. Technol.* **2018**, *28*, 153–162. [CrossRef]

27. Sultan, H.H.; Salem, N.M.; Al-Atabany, W. Multi-Classification of Brain Tumor Images Using Deep Neural Network. *IEEE Access* **2019**, *7*, 69215–69225. [CrossRef]

28. Deepak, S.; Ameer, P. Brain tumor classification using deep CNN features via transfer learning. *Comput. Biol. Med.* **2019**, *111*, 103345. [CrossRef]

29. Lenail, A. NN-SVG: Publication-Ready Neural Network Architecture Schematics. *J. Open Source Softw.* **2019**, *4*, 747. [CrossRef]

30. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.

31. Chakrabarty, N. Brain MRI Images for Brain Tumor Detection. *Kaggle* **2019**. Available online: http://xxx.lanl.gov/abs/https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection (accessed on 10 October 2019).

32. Cocosco, C.A.; Kollokian, V.; Kwan, R.K.S.; Pike, G.B.; Evans, A.C. BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. *NeuroImage* **1997**, *5*, 425.

33. Kwan, R.K.; Evans, A.C.; Pike, G.B. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans. Med Imaging* **1999**, *18*, 1085–1097. [CrossRef] [PubMed]

34. Collins, D.L.; Zijdenbos, A.P.; Kollokian, V.; Sled, J.G.; Kabani, N.J.; Holmes, C.J.; Evans, A.C. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med Imaging* **1998**, *17*, 463–468. [CrossRef] [PubMed]

35. Kwan, R.; Evans, A.; Pike, B. An Extensible MRI Simulator for Post-Processing Evaluation. In *International Conference on Visualization in Biomedical Computing*; Springer: Berlin/Heidelberg, Germany, 1996. [CrossRef]

36. Johnson, K.A.; Becker, A. The Whole Brain Atlas. 1995. Available online: http://xxx.lanl.gov/abs/www.med.harvard.edu/aanlib/home.html (accessed on 17 October 2019).

37. Chollet, F. Keras. 2015. Available online: https://github.com/fchollet/keras (accessed on 10 July 2020).

38. Mandrekar, J.N. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *J. Thorac. Oncol.* **2010**, *5*, 1315–1316. [CrossRef] [PubMed]

Sample Availability: The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.