International Journal of Engineering & Technology, 10 (2) (2021) 155-163

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Diagnosis of Medical Images Using Cloud-Deep Learning Model

Michael Jacobs, Ali Arfan, Alaa Sheta[1]

[1] *Computer Science Department,*
Southern Connecticut State University
501 Crescent Street, New Haven, CT 06516, USA
[*] *Corresponding author E-mail: shetaa1@southernct.edu*

Abstract

Diagnosis of brain tumors is one of the most severe medical problems that affect thousands of people each year in the United States. Manual classification of cancerous tumors through examination of MRI images is a difficult task even for trained professionals. It is an error-prone procedure that is dependent on the experience of the radiologist. Brain tumors, in particular, have a high level of complexity. Therefore, computer-aided diagnosis systems designed to assist with this task are of specific interest for physicians. Accurate detection and classification of brain tumors via magnetic resonance imaging (MRI) examination is a famous approach to analyze MRI images. This paper proposes a method to classify different brain tumors using a Convolutional Neural Network (CNN). We explore the performance of several CNN architectures and examine if decreasing the input image resolution affects the model's accuracy. The dataset used to train the model has initially been 3064 MRI scans. We augmented the data set to 8544 MRI scans to balance the available classes of images. The results show that the design of a suitable CNN architecture can significantly better diagnose medical images. The developed model classification performance was up to 97% accuracy.

Keywords: *Medical Imaging; Brain Tumor; Classification; Deep Learning; Convolutional Neural Network; Google Cloud; MRI*

1. Introduction

Healthcare is one of the critical sectors in the USA. The USA's federal government, by the year 2028, will spend $2.9 trillion on health care [1]. Eventually, these costs will continue to grow and consume a cumulative share of the USA federal resources. One of the essential processes in healthcare is diagnosing medical images, which offers a challenging mission for any physician. Nowadays, physicians address new health problems, unprecedented ailments, and many challenges that arise from many uncontrolled health problems. They capture benefits from up-to-date technologies to get over diseases and medical issues. The use of powerful image processing technologies, such as Magnetic Resonance Imaging (MRI) and X-ray and computed tomography, permits the physician to analyze medical images better. Magnetic resonance imaging (MRI) is a medical imaging technique that creates detailed images of the organs and tissues in a body. MRI machines create cross-sectional images that can be viewed from several angles. MRI is the most frequently used imaging method of the brain, and spinal cord [2]. Such technologies can provide an appropriate suggestion for the diagnosis. In the US between 2013 and 2017, there was an average of 16,249 deaths per year attributed to the primary malignant brain and other CNS tumors [3]. It is often necessary for surgeons to perform a biopsy for the identification of cancer. These biopsies involve invasive procedures and removal of tissue [4].

Brain tumors are one of the leading causes of death in the USA. It is estimated in 2021, that approximately 24,530 malignant tumors of the brain or spinal cord will be diagnosed [5]. It is also estimated that 18,020 adults (10,190 men and 7,830 women) will die from primary brain tumors this year. Currently, there are more than 120 types of brain and central nervous system (CNS) tumors. Most medical institutions use the World Health Organization (WHO) classification system to identify brain tumors which classifies brain tumors from the least aggressive (benign) to most aggressive (malignant) [6]. Survival rates for brain tumors vary widely depending on many factors, such as the type of brain or spinal cord and the person diagnosed. For

example, the survival rate for people younger than age 15 is around 74%, for ages 15 to 39 it is about 71%, and for people age 40 and over it is about 21% [7].

A brain tumor is a collection or mass of abnormal cells in one's brain. Since the skull that rims our brain is very rigid, any growth inside such an area can cause problems. Both malignant and benign tumors can grow to the extent that they can create enough pressure to cause life-threatening brain damage or cause some other physical inability [8]. Brain tumors can develop from brain cells, meninges, nerve cells, and glands. The various types of brain tumors affect people differently depending on their age group. For example, brain stem gliomas, medulloblastomas, and pineal tumors are more common in children than adults [9–11].

Many machine learning (ML) techniques have been used to handle tumor classification in the past few decades. For example, in [9], the authors presented a method that includes various phases such as pre-processing, feature extraction, association rule mining, and classification using a decision tree. Many brain tumor detection and classification techniques have been provided based around Support Vector Machine (SVM) classifiers [12]. SVM-based classification research was conducted by Javed et al. in [13]. Their proposed technique was divided into three steps to classify MRI Brain Imagery: feature extraction, weight assignment, and classification. The features of the proposed scheme were based on textural features and invariant moments. Textural features (coarseness, contrast, complexity, busyness, shape, directionality, and texture strength) were computed using Neighborhood Gray Tone Difference Matrix (NGTDM). A multi-class SVM was used to achieve superior classification performance compared to other algorithms. The proposed technique was verified using the Harvard medical brain database, which consists of MRI T2-weighted images (of 256×256 spatial resolution). The data set contains 48 normal and 25 brain images for each disease with seven moments and 20 orientations; their accuracy rate was about 96% accuracy. A method utilizing a logistic regression classifier as part of a machine learning algorithm was presented in [14]. Another type of ML classifier, naive Bayes, was used in combination with a multi-layer perceptron (MLP) network was presented in [15]. Fuzzy logic was used to assign weights to different feature values based on its discrimination capability to calculate the distance between the features of the test image and the mean feature vector of each class (normal, sarcoma, meningioma, and glioma).

The motivation behind this work is to develop a CNN model to diagnose MRI images via examining MRI scans. Accurate detection and classification of brain tumors is a critical area in medicine that affects thousands of people each year. Brain tumors have dramatic adverse effects on a person's health, including memory difficulties, cognitive impairments, and seizures, among other symptoms eventually resulting in death [16]. As a result, this motivates us to examine the performance of CNNs on the classification of brain tumor MRI images to improve early detection and identification. This work aims to offer a solution to brain tumor diagnosis using CNN when presented with MRI scans of different brain tumors. In this work, we explore several CNN architectures and collect the accuracy of each model based on each structure complexity. Also, we examine the effect of resolution variation of the input images have on the performance of different model complexities.

The paper is organized as follows. In Section 2, we provide a literature review of the brain tumor classification research. In Sections 8, we explore various aspects related to the CNN model design. We also explain the mission of multiple layers of CNN. The adopted data set and any pre-processing is described in Section 8.1. In Section 8.4, we explain the evaluation metrics that were used in this project. In Section 9, we illustrate the experimental results, and finally, in Section 10 we provide the conclusion and express future work.

2. Diagnosis Using CNN

Recently, CNN has become prominent for solving many image classification problems [17, 18]. A Deep Learning network to assist with this process would be especially useful as early detection of cancerous cells in the brain can help reduce the number of fatalities [8, 19]. Balasooriya *et. al* developed a deep learning algorithm using CNN's [10]. Their research was focused on identifying three tumor types: astrocytoma, glioblastoma, oligodendroglioma. They tested several models of varying depth to examine the performance of different-sized networks and fine-tune their network for MRI classification.

A transfer-based-learning CNNs was investigated by Rehman *et al.* in [20]. They investigated the performance of three pre-trained CNNs: AlexNet, GoogLeNet, and VGGNet. They used two different training strategies, which involved fine-tuning and layer freezing. They used MRI images and data augmentation techniques to generate more data. Khan *et al.* in [21] have used a small dataset of 253 brain MRI images to train a simple eight layers CNN model. They made a comparison between a proposed CNN model and a pre-trained VGG-16, ResNet-50, and Inception-v3 models using a transfer learning approach. Their adopted data augmentation to artificially increase the quantity and complexity of existing data. Their proposed model showed 96% accuracy on the training data and 89% accuracy on the validation dataset.

An overall comparison of the performance of several traditional classifiers, as well as CNN, was performed by Hossain *et al.* in [11]. They looked at the performance of SVM, K-Nearest Neighbor (KNN), MLP, logistic regression, naive Bayes, and random forest. Two models were used, one for the traditional classifiers and one for the CNN. In the first model, the tumor images were segmented using Fuzzy C-Means clustering algorithm and additional techniques. The features extracted from the segmented images include texture-based and statistical-based features. The second model used a five-layer CNN consisted of one convolutional layer, one max-pooling layer, a flattening layer, and two fully connected layers. They found the best performing traditional method was SVM with 92.42% accuracy and the CNN had 97.87% accuracy.

3. What is CNN Model?

A CNN is a type of artificial neural network (ANN) specialized for image processing. ANNs are computational systems inspired by the way biological nervous systems work, such as the human brain. They are comprised of a large number of interconnected computational nodes (neurons) which collectively learn from the input to optimize the output. ANNs have an input layer, any number of hidden layers, and an output layer [22].

CNN's are preferred over feed-forward ANNs for image processing problems because they retain context. They can capture

spatial and temporal dependencies in images that would be lost in a feed-forward ANN. Also, CNN's are better suited for handling the large amount of data presented by images. For example, if a 256×256 color image were the input for ANN, the first hidden layer would require 196,608 weights on a single neuron. The ANN would need to be extremely large to handle the input of this scale. The convolutional layers in a CNN are far better suited for handling input of this size.

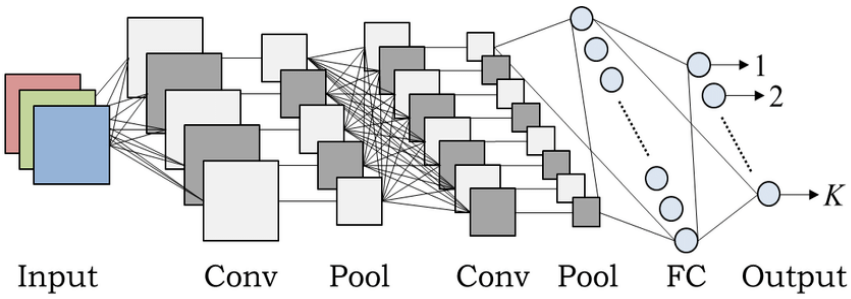


Figure 1: Overview of a CNN model.

The overall structure of a CNN is similar to an ANN. The primary difference being the inclusion of Convolutional Layers and Pooling Layers along with Fully-Connected Layers. These layers are stacked to form the CNN architecture (See Figure 1).

4. Convolutional Layer

Convolutional layers use learnable filters (also known as kernels) to create feature maps of an image [21,23]. These kernels are usually small in size, and the layer convolves each filter across the spatial dimensionality of the image to produce a 2D feature map. Each kernel in a convolutional layer has its feature map, and they are stacked to form the total output of the layer. Convolutional layers reduce the complexity of the model through optimization of its output through three hyperparameters: depth, stride, and zero-padding [11,24].

The depth is manually adjusted through the number of filters used in the layer. Stride is the number of pixels the filter shifts over the input image. Zero-padding is the process of padding the border of the image with zeros to control the dimensionality. The output dimensions can be calculated with the formula:

$$I_{Output} = \frac{W - K + 2P}{S} + 1 \tag{1}$$

W is the input dimension, K is the kernel dimension, P is the padding, and S is the stride.

5. Pooling Layer

The pooling layer is commonly applied after a convolution layer to reduce the spatial size of the input (see Figure 2. It is applied independently to each depth slice of the input volume. The volume depth is always preserved in pooling operation. The pooling layer is not trained during the backpropagation of gradients because the output volume of data depends on the input volume of data values. The advantages of the pooling layer are: 1) it reduces the number of training parameters and computation cost to control overfitting, 2) makes the model invariant to certain distortion. Types of Pooling layers include:

1. Max Pooling: the maximum value of each kernel in each depth slice is captured and passed on to the next layer.
2. Min Pooling: the minimum value of each kernel in each depth slice is captured and passed on to the next layer.
3. L2 Pooling: the L2 or the Frobenius norm is applied to each kernel.
4. Average Pooling: the average value of the kernel is calculated.

The pooling layer requires two hyperparameters, kernel/filter size K and stride S. For the pooling layer, it is not common to pad the input using zero-padding [25].

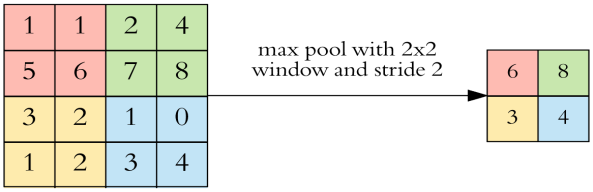


Figure 2: An example of a Max-Pooling operation.

6. Fully-connected Layer

Before being input into a Fully Connected (FC) Feedforward (FF) layer, the output of the convolutional and pooling layers needs to be flattened into a $m \times 1$ matrix, where m is the number of remaining pixels. FC layers contain neurons that are directly connected to the neurons of the two adjacent layers. CNN can include more than one fully connected layer, but the final layer will always be FC. The last fully-connected layer will have an output size equal to the number of possible classes.

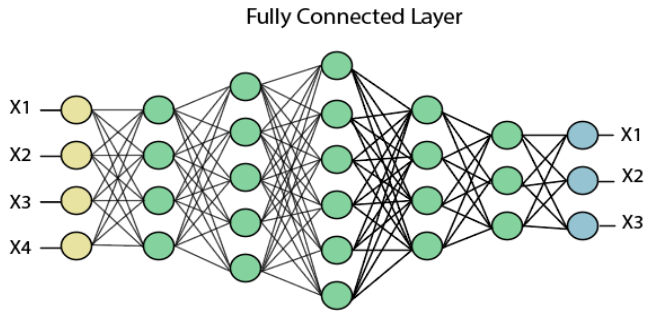


Figure 3: Fully connected network that accepts the feature maps vector $x_1, x_2, x_3, \ldots, x_n$ from previous layers [26].

7. Loss Function

The Loss Function is one of the essential components of any ANN. It is a prediction error of the ANN. The Loss is used to calculate the gradients. The gradients are used to update the weights of the ANN as a primary way to train the network. We will be using categorical cross-entropy. In this case, there must be the same number of output nodes as the classes. And the final layer output should be passed through a softmax activation so that each node output a probability value between (0–1) [27]. The Cross-Entropy Loss function is given as:

$$L_{CE} = -\sum_{i=1}^{n} t_i \log_2(p_i) \tag{2}$$

where t_i is the truth label and p_i is the Softmax probability for the i^{th} class.

8. Proposed Classification Process

The process we adopted in this research is depicted in Figure 4. We used this process to evaluate and improve the designs of our models. The proposed method is a CNN-based method. First, we collected our dataset. Next, we put the data through data augmentation to increase the size of the dataset. Then we split the dataset into 60% training, 20% validation, and 20% test sets. We selected a random collection of images for testing and validation, but each contains a balanced amount of images from each of the three classes. Next, the training set is used to train our models, and the validation set is used to analyze the training performance. The test set is kept hidden until we do a final performance evaluation.
Our method to improving our CNN models' performances involved adjusting the parameters of each layer, varying the number of convolutional layers in the model, and varying the input image resolution. We examine the performance of each architecture complexity with varying input image sizes. The three proposed CNN models are given in Table 1.

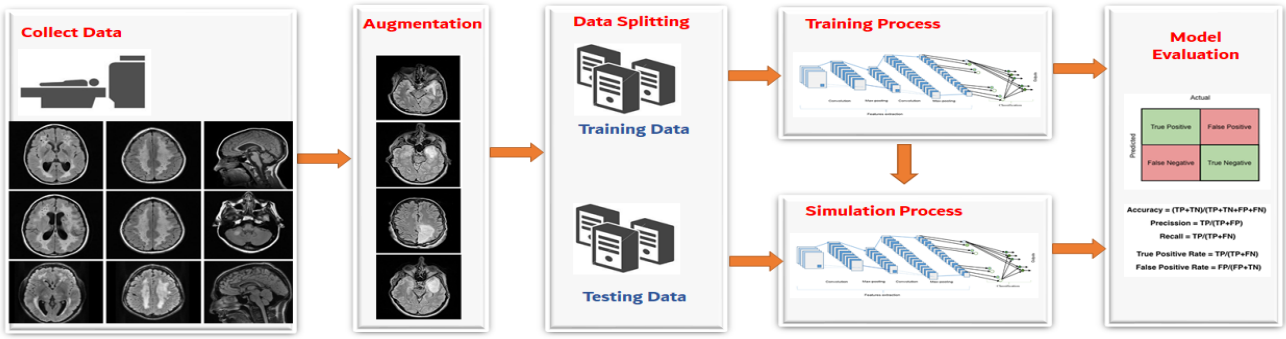


Figure 4: Proposed classification process using CNN

8.1. Dataset

The image database used in this research contains 3064 T1-weighted contrast-enhanced images collected from 233 patients with three kinds of brain tumors 1) Meningioma (708 slices), 2) Glioma (1426 slices), and 3) Pituitary tumor (930 slices).

Table 1: Architecture for the proposed CNN Models

Layer #	Model 1	Model 2	Model 3
1	Input Layer	Input Layer	Input Layer
2	Convolutional: (32 Kernels, size=(3x3), stride=(2 for 256 and 512)/(1 for 128))	Convolutional: (16 Kernels, size=(3x3), stride=(2 for 256 and 512)/(1 for 128))	Convolutional: (8 Kernels, size=(3x3), stride=(2 for 256 and 512)/(1 for 128))
3	ReLU	ReLU	ReLU
4	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)
5	Convolutional: (64 Kernels, size=(3x3), stride=(2 for 512)/(1 for 128 and 256))	Convolutional: (32 Kernels, size=(3x3), stride=(2 for 512)/(1 for 128 and 256))	Convolutional: (16 Kernels, size=(3x3), stride=(2 for 512)/(1 for 128 and 256))
6	ReLU	ReLU	ReLU
7	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)
8	Convolutional: (128 Kernels, size=(3x3), stride=1)	Convolutional: (64 Kernels, size=(3x3), stride=1)	Convolutional: (32 Kernels, size=(3x3), stride=1)
9	ReLU	ReLU	ReLU
10	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)
11	Fully-Connected: (in=25,088, out=6,272)	Convolutional: (128 Kernels, size=(3x3), stride=1)	Convolutional: (64 Kernels, size=(3x3), stride=1)
12	Dropout (p=0.5)	ReLU	ReLU
13	Fully-Connected: (in=6,272, out=3)	Max-Pooling(size=(2x2), stride=2)	Max-Pooling(size=(2x2), stride=2)
14	Classification Output	Fully-Connected: (in=4,608, out=1,152)	Convolutional: (128 Kernels, size=(3x3), stride=1)
15		Dropout: (p=0.5)	ReLU
16		Fully-Connected: (in=1,152, out=3)	Max-Pooling(size=(2x2), stride=2)
17		Classification Output	Fully-Connected: (in=512, out=256)
18			Dropout: (p=0.5)
19			Fully-Connected: (in=256, out=3)
20			Classification Output

The data contains images from three planes: sagittal (1025 images), axial (994 images), and coronal (1045 images) planes. An example of the brain tumor MRI images is shown in Figure 5. The dataset was published online in 2015, and the most recent version was released in 2017. The images were acquired from Nanfang Hospital, Guangzhou, China, and General Hospital, Tianjing Medical University, China, from 2005 to 2010 [28]. The images are provided initially with a resolution of 512×512. To test the effect of input image resolution on the performance of a model, we will test each of our three developed models with the original resolution of 512×512 and with images resized to 256×256, and 128×128.

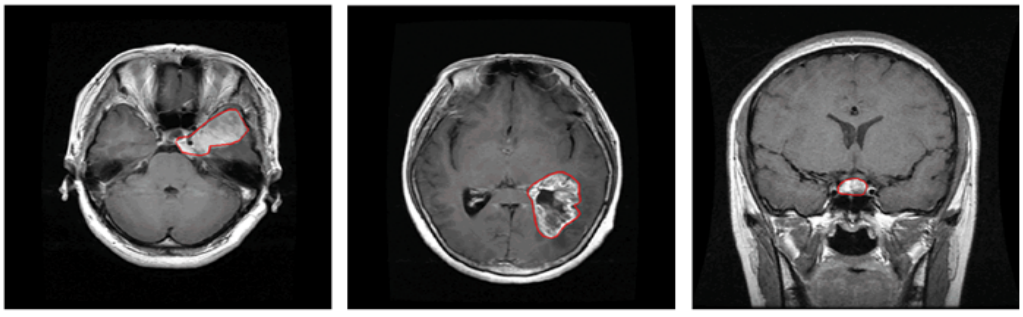


Figure 5: The adopted data set [29] has three types of tumors: (a) glioma; (b) meningioma; and (c) pituitary tumor.

To test the effect of input image resolution on the performance of a model, we will test each of our three developed models with the original resolution of 512×512 and with images resized to 256×256, and 128×128. The dataset was split into training, validation, and test sets with proportions of 60%, 20%, and 20%, respectively.

8.2. Data Augmentation

We increased the size of our dataset through two methods of data augmentation. The original dataset has significantly more glioma images than meningioma and pituitary images. To prevent our model from overfitting glioma samples, we increased meningioma and pituitary samples through data augmentation. We used the method of 90-degree rotation on meningioma and pituitary images to augment the data. Figure 6 shows an example of image rotation; this increased the total number of samples to 8,544. Next, because the dataset is still relatively small, we doubled the size by vertically flipping each of the images (shown in Figure 7); this way, we increased the size of the dataset to 8,544. The current size of the augmented dataset is:

- Meningioma (2832 slices),
- Glioma (2852 slices), and

- Pituitary tumor (2860 slices).

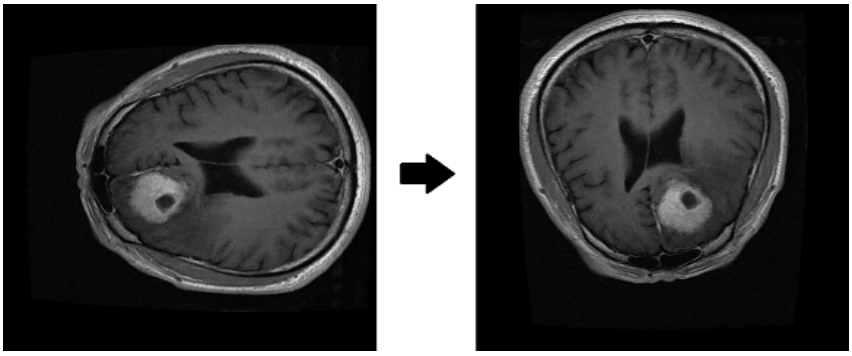


Figure 6: Augmentation via 90 degree rotation.

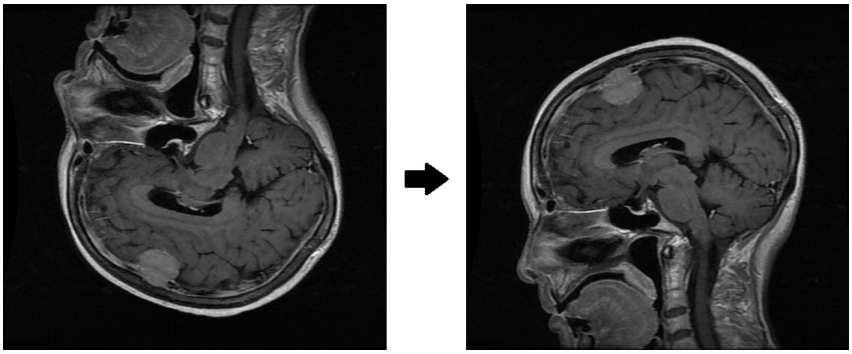


Figure 7: Augmentation via vertical flipping.

8.3. Google Colaboratory (Colab)

Tumor classification in our proposed method was done with a CNN model written in Python in Google Colaboratory [30]. Colab is a product from Google Research that allows anyone to write and execute Python code through the browser with zero configuration required, free access to GPUs, and easy sharing. Colab notebooks allow one to combine executable code and rich text in a single document, along with images, HTML, LaTeX, and more [31]. One can import the data into Colab notebooks from a Google Drive account, including spreadsheets, Github, and many other sources. We have imported the dataset from our Google Drive to the Colab. It is a Jupyter notebook environment that supports many popular machine learning environments [32]. Two deep learning libraries were utilized to construct the CNN, the PyTorch and the Fastai.

8.4. Model Evaluation

The most commonly used evaluation metrics are accuracy, precision, recall, and F1-score. We will use these metrics to evaluate the performance of our models. They are defined as the following:

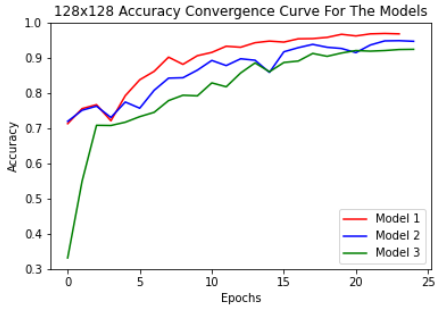
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

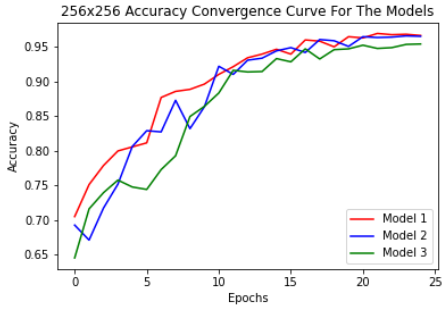
$$Recall = \frac{TP}{TP+FN} \tag{5}$$

$$F1 = \frac{2(Precision)(Recall)}{Precision+Recall} \tag{6}$$

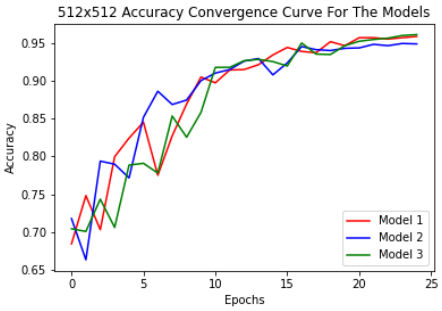
TP, FP, TN, and FN are the calculated values of true positive, false positive, and true negative and false negative.
Accuracy describes how closely a measured value corresponds to its true value. Precision is how close the measured values are to each other. The recall is the proportion of actual positives that were identified correctly. Recall indicates missed positive predictions. It gives a measure of how accurately the model can identify the relevant data. The F1 score is between 0 and 1 and is the harmonic mean of precision and recall. The F1 score combines the accuracy and recalls into a single metric by finding their harmonic mean. This metric is primarily used to compare the performances of two classifiers since it is difficult to compare two models with low precision and high recall or vice versa [33].



(a) Accuracy curve for CNN Model 1



(b) Accuracy curve for CNN Model 2



(c) Accuracy curve for CNN Model 3

Figure 8: Accuracy Curves for the three proposed CNN architectures

9. Experimental Results

In this work, we experimented to determine how the depth of a CNN affects the accuracy and performance. We developed three models with a different number of convolutional layers: Model 1 (3 Conv layers), Model 2 (4 Conv layers), and Model 3 (5 Conv layers). All three models had two fully connected layers before classification output. We tested each model with three different input image sizes. Table 3 reports the training and testing results for the three models. Figures 8(a), 8(b) 8(c) show the accuracy convergence curve of the models for each input size and Figures 9(a), 9(b) and 9(c) show the validation loss convergence curve of the models for each input size.
Our results show that the most acceptable CNN model for training and testing cases is CNN model 1. It appears that the depth of the CNN affects the accuracy depending on the input image resolution. Higher image resolution shall require deeper networks to achieve higher accuracy, while smaller input sizes may be hampered by increasing the depth of the CNN model. Higher image resolution also appears less severely affected by a change in depth than smaller input sizes. The confusion matrix for our best test is shown in Table 2.

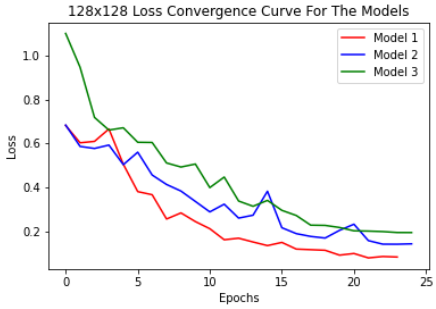
Table 2: Model 1 128×128 Test Confusion Matrix

		[Predicted]			
		glioma	*meningioma*	*pituitary*	
[Actual]	*glioma*	550	20	1	96.32%
	meningioma	16	545	6	96.12%
	pituitary	5	10	557	97.38%

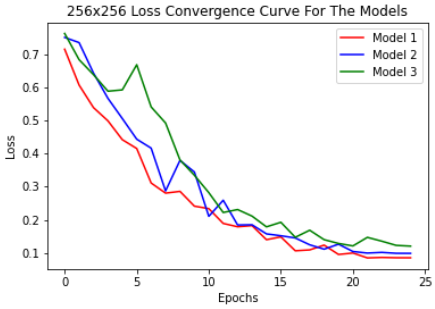
We also measured the time required to train each model for each input image size. Figure 10 shows the training time needed for each CNN model. Increasing the image resolution increases the time required for training with a possible. The model with an image resolution of 128×512 outperformed other models in the training case.

10. Conclusions

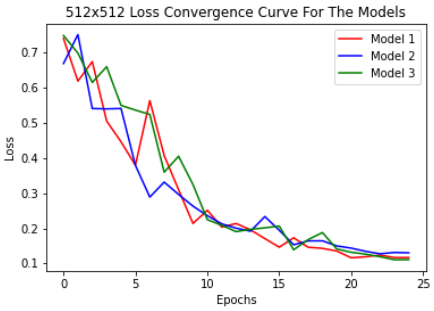
In this work, we proposed a CNN-based method to assist with brain tumor classification by examining MRI images. The main goal of the proposed method is to find the best-performing CNN architecture by varying the number of convolutional layers of the models. We also examined the effect of image resolution has on the performance of a model. We reviewed the performance of our models on a publicly available dataset consisting of 3,064 scans augmented to 8,544 scans. The results show that the depth of a CNN model affects the classification accuracy. Increasing the depth of a model may improve the accuracy of using larger input image sizes or decrease when using smaller input image sizes. Deeper models benefited larger resolutions but were a detriment for smaller resolutions. They were increasing the image size results in longer training times. The obtained results show that a CNN model with three convolutional layers and two fully connected layers using an input image size of 256×256 achieved the best accuracy.



(a) Accuracy curve for CNN Model 1



(b) Accuracy curve for CNN Model 2



(c) Accuracy curve for CNN Model 3

Figure 9: Loss Curves for the three proposed CNN architectures

Table 3: The three proposed CNN models results

		Training				Testing			
Model	Input Size	Accuracy	Precision	Recall	Time (mm:ss)	Accuracy	Precision	Recall	F1-Score
CNN Model 1	128×128	**0.9684**	**0.9685**	**0.9684**	33:02	**0.9632**	**0.9634**	**0.9632**	**0.9633**
	256×256	0.9666	0.9666	0.9666	41:29	0.9661	0.9662	0.9661	0.9661
	512×512	0.9590	0.9589	0.9589	80:08	0.9550	0.9551	0.9549	0.9550
CNN Model 2	128x128	0.9473	0.9475	0.9472	29:39	0.9456	0.9459	0.9456	0.9457
	256×256	0.9655	0.9657	0.9654	26:15	0.9579	0.9581	0.9579	0.9580
	512×512	0.9491	0.9490	0.9490	70:13	0.9567	0.9569	0.9567	0.9568
CNN Model 3	128×128	0.9245	0.9244	0.9244	25:43	0.9298	0.9296	0.9297	0.9296
	256×256	0.9543	0.9549	0.9543	29:17	0.9538	0.9541	0.9538	0.9539
	512x512	0.9614	0.9614	0.9613	63:38	0.9573	0.9573	0.9573	0.9573

References

[1] American Health Care, "American health care: Health spending and the federal budget," https://www.crfb.org/papers/american-health-care-health-spending-and-federal-budget, May 2018.

[2] Mayo Clinic Staff. Test procedures: Magnetic resonance imaging. [Online]. Available: https://www.mayoclinic.org/tests-procedures/mri/about/pac-20384768#:~:text=Magnetic%20resonance%20imaging%20(MRI)\%20is,large%2C%20tube%2Dshaped%20magnets.

[3] Central Brain Tumor Registry of the United States. CBTRUS fact sheet. [Online]. Available: https://cbtrus.org/cbtrus-fact-sheet-2020/#:~:text=This%20represents%20an%20average%20of,and%\2045%20deaths%20per%20day.&text=It%20was%20estimated%\20that%20there,and%207%2C830%20occurring%20in%20fema\les.

[4] American Cancer Society. (2020) Tests for brain and spinal cord tumors in adults. [Online]. Available: https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/detection-diagnosis-staging/how-diagnosed.html

[5] ——. (2021) Key statistics for brain and spinal cord tumors. [Online]. Available: https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/about/key-statistics.html

[6] National Brain Tumor Society. Tumor types: Understanding brain tumors. [Online]. Available: https://braintumor.org/brain-tumor-information/understanding-brain-tumors/tumor-types

[7] American Society of Clinical Oncology. (2020) Brain tumor statistics. [Online]. Available: https://www.cancer.net/cancer-types/brain-tumor/statistics

[8] E. Dandıl, M. Çakıroğlu, and Z. Ekşi, "Computer-aided diagnosis of malign and benign brain tumors on MR images," in *ICT Innovations 2014*, A. M. Bogdanova and D. Gjorgjevikj, Eds. Cham: Springer International Publishing, 2015, pp. 157–166.

[9] J. Naik and S. Patel, "Tumor detection and classification using decision tree in brain MRI," in *International Journal of Engineering Development and Research*, 2013, pp. 49–53.

[10] N. M. Balasooriya and R. D. Nawarathna, "A sophisticated convolutional neural network model for brain tumor classification," in *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*, 2017, pp. 1–5.

[11] T. Hossain, F. S. Shishir, M. Ashraf, M. A. Al Nasim, and F. Muhammad Shah, "Brain tumor detection using convolutional neural network," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp. 1–6.

[12] A. Kharrat, M. B. Halima, and M. B. Ayed, "MRI brain tumor classification using support vector machines and meta-heuristic method," *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, pp. 446–451, 2015.

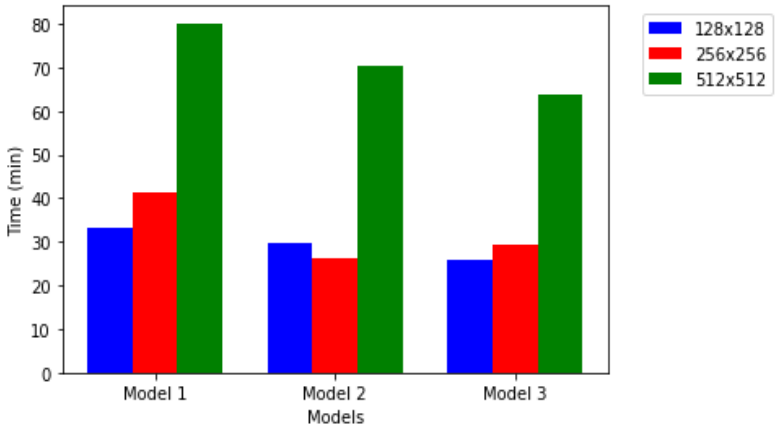


Figure 10: Computed run time for the three proposed CNN models

[13] U. Javed, M. Riaz, A. Ghafoor, and T. Cheema, "MRI brain classification using texture features, fuzzy weighting and support vector machine," *Progress In Electromagnetics Research*, vol. 53, pp. 1–6, July 2013.

[14] S. S. Saba, D. Sreelakshmi, P. S. Kumar, K. S. Kumar, and S. R. Saba, "Logistic regression machine learning algorithm on MRI brain image for fast and accurate diagnosis," *International Journal of Scientific Technology Research*, vol. 9, no. 3, pp. 7076–7081, Mar 2020.

[15] K. Sharma, A. Kaur, and S. Gujral, "Brain tumor detection based on machine learning algorithms," *International Journal of Computer Applications*, vol. 103, no. 1, pp. 7–11, Oct 2014.

[16] The Brain Tumor Charity. Side-effects of a brain tumour. [Online]. Available: https://www.thebraintumourcharity.org/living-with-a-brain-tumour/side-effects/

[17] A. Sheta, H. Turabieh, S. Aljahdali, and A. Alangari, "Pavement crack detection using convolutional neural network," in *Proceedings of 35th International Conference on Computers and Their Applications*, ser. EPiC Series in Computing, G. Lee and Y. Jin, Eds., vol. 69. EasyChair, 2020, pp. 214–223. [Online]. Available: https://easychair.org/publications/paper/t2BC

[18] E. Endri, A. Sheta, and H. Turabieh, "Road damage detection utilizing convolution neural network and principal component analysis," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 6, 2020. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2020.0110682

[19] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on MRI images and course tumor boundaries," *CoRR*, vol. abs/1811.00597, 2018.

[20] A. Rehman, S. Naz, M. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Systems, and Signal Processing*, vol. 39, pp. 757–775, Feb 2020.

[21] H. A. Khan, W. Jue, M. Mushtaq, and M. U. Mushtaq, "Brain tumor classification in MRI image using convolutional neural network," *Mathematical Bioscience and Engineering*, vol. 17, pp. 01–14, 09 2020.

[22] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *CoRR*, vol. abs/1511.08458, 2015.

[23] A. Bhandari, J. Koppen, and M. Agzarian, "Convolutional neural networks for brain tumour segmentation," *Insights into Imaging*, vol. 11, 12 2020.

[24] M. M. Badža and M. Barjaktarović, "Classification of brain tumors from MRI images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, 2020.

[25] A. K. Rana. (2020) Pooling layer — short and simples. [Online]. Available: https://ai.plainenglish.io/pooling-layer-beginner-to-intermediate-fa0dbdce80eb

[26] JavaTpoint. Convolutional neural network in pytorch. [Online]. Available: https://www.javatpoint.com/pytorch-convolutional-neural-network

[27] Shiva Verma. (2019, June) Understanding different loss functions for neural networks. [Online]. Available: https://towardsdatascience.com/understanding-different-loss-functions-for-neural-networks-dd1ed0274718

[28] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Correction: Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLOS ONE*, vol. 10, no. 12, pp. 1–1, Dec 2015.

[29] J. Cheng, "Brain Tumor Dataset," https://doi.org/10.6084/m9.figshare.1512427.v5, Apr 2017.

[30] T. Carneiro, R. V. Medeiros Da NóBrega, T. Nepomuceno, G.-B. Bian, V. H. C. De Albuquerque, and P. P. R. Filho, "Performance analysis of google colaboratory as a tool for accelerating deep learning applications," *IEEE Access*, vol. 6, pp. 61 677–61 685, 2018.

[31] M. J. Nelson and A. K. Hoover, "Notes on using google colaboratory in ai education," in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, ser. ITiCSE '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 533–534. [Online]. Available: https://doi.org/10.1145/3341525.3393997

[32] E. Bisong, *Google Colaboratory*. Berkeley, CA: Apress, 2019, pp. 59–64.

[33] M. Hossin and S. M.N, "A review on evaluation metrics for data classification evaluations," *International Journal of Data Mining Knowledge Management Process*, vol. 5, pp. 01–11, Mar 2015.