

MRI Classification and Segmentation for Brain Tumor Using Convolutional Neural Network and U-Net

Elsevier¹

Radarweg 29, Amsterdam

Elsevier Inc^{a,b}, Global Customer Service^{b,*}

^a1600 John F Kennedy Boulevard, Philadelphia

^b360 Park Avenue South, New York

Abstract

Classification and segmentation of brain tumors are critical for analyzing tumors and determining treatment options based on their classifications and appearance. Magnetic Resonance Image (MRI) is considered to be the mainly employed mechanism for viewing brain regions. The tissue contrast is normalized in MRI and it is regarded as the most compliant imaging technique for modeling the territory of attention in the brain, such as tumors. This paper describes a unique deep Convolutional Neural Network (CNN) strategy for the automatic classification of brain images into 4 classes (glioma, meningioma, pituitary, no tumor), as well as a U-net-based segmentation model. A total of six datasets (Dataset a, Dataset b, Dataset c, Dataset d, Merged dataset 1, Merged dataset 2) were used to test the classification model, and the segmentation was performed on a manually segmented dataset. Merged dataset 1 was manually segmented for classification purposes in order to compare classification performance by segmenting vs non-segmenting MRI images. The proposed classification model's performance is also compared to five transfer learning methods (VGG16, VGG19, ResNet152v2, EfficientNet B0, and EfficientNet B7). The proposed classification model outperforms other existing models in all six classification datasets due to its outstanding performance. The segmentation model is further evaluated using dice coefficient, dice loss, jaccard index, binary cross entropy dice loss, loss, recall, and accuracy metrics. For Merged dataset 1, the resulting masked images are also compared to the ground truths. Two approaches to classification were evaluated, one with segmented MRIs sent to the classification model and the other with the MRI pictures passed without segmentation, and their comparison was based on accuracy, recall, precision, and AUC. Merged dataset 1 has the greatest classification accuracy of 98.7%. The same classification model obtained an accuracy of 98.8% on the segmented Merged dataset 1 by delivering segmented MRIs to it. Dataset c achieves the greatest classification accuracy of 97.7% among the 4 individual datasets.

Keywords: Magnetic Resonance Imaging, Deep Convolutional Neural Network, Classification of brain MRI, Segmentation

1. Introduction

All of the functions in the human body are governed by the human brain, which is one of the most sophisticated processes in the body and this dominant organ consists of billions of neurons. The most frequent kind of brain disease is a tumor in the brain. In this kind of cancer, which is deadly, prompt and precise, diagnosis of brain tumors is critical [1]. The majority of primary Central Nervous System (CNS) cancers are tumors in the brain accounting for 85 percent to 90 percent of all cases. In 2020, about 308,102 persons worldwide are supposed to be detected with a spinal column or primary brain tumor [2].

It is caused due to abnormal and unregulated increase in brain cells' number [3]. There are more than 150 different forms of brain tumors thought to exist. They are categorized into two major groups: primary tumors and metastatic tumors.

Primary tumors are those that arise in the brain tissue or surrounding the brain. Primary tumors have been further divided into benign (noncancerous) and malignant (cancerous) categories. Metastatic tumors spread to the brain through the blood circulation after growing in other regions of the body (such the breasts or lungs). Meningiomas are a common kind of benign tumor (almost over 30%). Benign tumors in the brain can occasionally be life-threatening. Meningiomas, for example, can occasionally develop into malignant tumors. Meningiomas are slow-growing, benign (noncancerous) tumors that make up around 85% of all cases. Meningiomas are more frequently diagnosed in women than in men. They have a good probability of being surgically removed since they seldom disseminate to the adjacent brain tissue. Pituitary tumors originate in the pituitary glands and these glands govern hormones and physiological processes. Pituitary tumors are noncancerous tumors that do not spread to other organs. Although the majority of the benign type tumors are pituitary that seldom progress to cancer. Pituitary tumor problems can result in long-term hormone insufficiency as well as eyesight

*Corresponding author

Email address: support@elsevier.com (Global Customer Service)

URL: www.elsevier.com (Elsevier Inc)

¹Since 1880.

loss. Tumor cells that are malignant are abnormal cells that increase uncontrolled and irregularly. Normal tissues can be compressed, infiltrated, or destroyed by these tumors. The most prevalent variety of malignant brain tumors are noted to be gliomas. Gliomas make for about 33% of all brain tumors. They can rarely develop to the spine or other body organs, but they can develop quickly and can invade healthy tissues in the surrounding area [4].

The conventional practice for segmenting tumor areas is manual segmentation. Nevertheless, it is exorbitant, exhausting, and prone to inter-observer variability [5]. As a result, automated tumor segmentation [6] is favored, especially when dealing with massive amounts of data and the need for constant tumor observation and pliable treatment planning. Anyway, due to the wide range of tumor sites, forms, and architectures, effective automated tumor segmentation is typically difficult [7]. Deep learning models have been explored for the purpose of brain tumor detection over the recent years. MRI of brain tumors are used as the datasets for its high quality and non-ionizing radiation properties.

Due to the significant contributions that deep learning models may make to this topic, some of the pertinent papers have been extensively covered in this section.

Mohsen et al. accomplished a Deep Neural Network (DNN) classifier integrated with Principal Component Analysis (PCA) and Discrete Wavelet Transform (DWT) for categorizing MRIs in 4 classes [8]. Images were segmented using the Fuzzy C-Means (FCM) clustering. Though they performed classification with a very small dataset of 66 brain MRIs. Özyurt et al. implemented a method to categorize brain tumors as malignant or benign using a hybrid Neutrosophic Set – Expert Maximum Fuzzy-Sure Entropy Convolutional Neural Network (NS-EMFSE-CNN) classification technique. A comparison that was carried out for two classifiers gave an accuracy of 95.62% by Support Vector Machine (SVM) [9]. However, they only classified tumors as benign or malignant, no particular tumor type classification has been carried out.

Sajid et al. devised a method that implemented a hybrid CNN with 2 path CNN and 3 path CNN using a patch-based technique considering local and contextual information [10]. Glioma tumor regions are segmented using a patch-based approach. They segmented tumors into Higher Grade Glioma (HGG) and Lower Grade Glioma (LGG), but no classification is performed. Emrah Irmak put forward a model for classifying brain MRIs with Deep CNN and grid search optimizer-tuned hyper-parameters with the motivation to build three different models for the classification of tumors [11]. Though their second model which classifies tumors into five classes achieved less accuracy 92.66% compared to the other two models.

In a specific study, an approach offered using the CNN architecture for classifying brain MRIs into three distinct groups and distinguishing in different glioma grades [12]. Nevertheless, the dataset used for distinguishing glioma grades is smaller and must be evaluated on a bigger group of datasets. In the paper proposed by Havaei et al. [13], the CNN is a two-pathway architecture that is capable of extracting local features along with global features of the brain tumors concurrently. In the

two training phase, the CNN base output is fed again into the CNN subsequently. However the architecture is required to be observed on large datasets for efficiency of handling huge MRIs.

Badza et al. introduced a CNN approach to classify tumors in three classes: meningioma, glioma, and pituitary. Four evaluating approaches have been utilized by aggregating two databases and two 10 cross-validation approaches [14]. Both the 10 fold record and subject-wise cross-validation approach has been used on the original and augmented images for evaluation. In the paper by Amin et al. they divided the input image into N number of patches and using a pre-trained CNN model, the core pixel of each patch is calculated then combining all the predicted results. They classified the tumor into HGG and LGG with different tumor region. They claimed the average processing time is only 5.02 second [15].

An automated deep multi-scale 3D CNN is recommended by Mzoughi et al. [16]. Their proposed architecture may combine lower weight global and local contextual information. However they classified the tumor into two subclasses of glioma (HGG and LGG). Siar et al. proposed a strategy that used ImageNet feature extraction model. Their model showed good performance on a first order clustering algorithm and CNN softmax fully connected layer has been devised to distinguish fat tissue and tumor tissue which increased accuracy [17]. Though no specific tumor type categorization has been performed; only benign or malignant tumors have been classified. Mustafa R. Ismael combined the statistical characteristics derived from the 2D DWT and 2D Gabor extraction methods separately for feature extraction. In order to categorize three different forms of tumors, the features were fed to a traditional neural network trained by backpropagation.

A generic CNN was implemented by Irshaidat et al. MRI inputs were opened in grayscale mode and augmentation made the dataset fourteen times larger from the initial size. The model predicted the presence of a tumor or not with an accuracy of 96.7% and 88.25% in validation and test data respectively [18]. Though they simply did a binary classification and used a smaller dataset to test their model. Ayadi et al. came up with a CNN model that included multiple layers to classify brain tumor MRIs into meningioma, glioma, and pituitary [19] utilizing three publically available datasets: Figshare dataset, Radiopaedia dataset, and REMBRANDT dataset for comparison with the previous work.

Naser et al. incorporated a deep learning strategy in which CNN hinged on U-Net for segmenting tumor affected regions. A transfer learning VGG-16 model and classifiers has been developed to grade cancerous tumors. No independent dataset was available to compare with for testing and a moderately small LGG data has been used for validation [20]. Nevertheless, only glioma tumors have been categorized as LGG and HGG. Shahzadi et al. used a CNN based 3-D medical image classification cascading with Long Short Term Memory (LSTM) to identify LGG and HGG. [21]. They utilized VGG-16 for feature extraction and LSTM for the classification of glioma. Because they employed a small number of samples and didn't apply any data augmentation techniques to expand

the number of samples, their model has lesser accuracy than other research.

Noreen et al. implemented two separate multi-level architectures using Inception-v3 and DenseNet201 architectures with softmax classifier. In these pre-trained architectures, features are extricated from various modules and percolate through the softmax layer for classification after concatenation of the extracted features [22]. The Super Resolution- Fuzzy C Means (SR-FCM) technique has been proposed by Özyurt et al. in tumor detection [23]. Features were extracted from SqueezeNet architecture and then Extreme Machine Learning (ELM) classification technique was performed. Anyway tumors are mainly divided into benign and malignant categories in their work.

A method based on the Resnet-50 from CNN was proposed by Ahmet Çınar and Muhammed Yildirim [24]. The endmost five layers of the model are eliminated, and eight additional layers are inserted as a substitute. However, the type of tumor was not characterized, only binary classification has been implemented. A CNN of 22 layers is proposed for categorizing lung cancer into five categories [25]. Navid Ghassemi proposed a Generative Adversarial Network (GAN) based model where a deep CNN is used as the discriminator for detection of fake images that are generated by the generative model and a pre-trained CNN network is fine-tuned to perform as a classifier. The input size was limited to 64×64 on account of some GAN restrictions which retrains the application of a few well-performed architectures as the discriminator since the larger input size is a prerequisite [26].

A new method in deep learning features, as well as a fusion of hand-crafted for brain tumor detection, was established by Saba et al. [27] where input images went through segmentation employing the GrabCut algorithm. The incorporated VGG-19 with Histogram of Oriented Gradients (HOG) and Local Binary Pattern (LBP). In the end, multiple classifiers are used for the classification. Between glioma and healthy brain images, the classification is limited. Ahamed et al. developed a deep learning approach for detecting covid-19 cases using chest CT scans and X-ray pictures [28]. An end-to-end automatic incremental CNN was proposed by Naceur et al. for the segmentation of tumors in which they implemented deep learning models entitled 2Cnet, 3Cnet, and EnsembleNet using their proposed training strategy [29]. The EnsembleNet is an integrated model of the 2Cnet and 3Cnet taking into account ensemble learning. Though the model is solely suitable for the segmentation of MRIs. Aurna et al. also used the ensemble technique to classify brain MRIs [30].

A study by Chelghoum et al. involved exploiting nine pre-trained CNNs namely ResNet-50, AlexNet, VGG-16, ResNet-101, GoogleNet [31], VGG-19, ResNet-18, SENet and ResNet-Inception-v2 for comparative analysis [32]. However, all the models have been implemented over a single dataset. Nawab et al. made use of block wise VGG-16 network centered on transfer learning and fine-tuning [33] to classify MRIs as: glioma, meningioma, and pituitary. The model went through minimal preprocessing procedure and does not utilize any of the handcrafted features.

A segmentation approach employing a CNN was implemented followed by an extensive data augmentation strategy by Sajjad et al. to classify multi-graded brain tumors[34]. Due to the small-scale datasets, the segmented data has been augmented using eight different types of data augmentation strategies including various geometric transformations and noise invariances. Finally, data has been passed through the pre-trained VGG-19 model for classification. Elazab et al. devised a model named GP-GAN, which predicts the growth of glioma tumors at an early stage using stacked 3D GANs [35]. Using a 3D U-Net architecture, this model's generator was created. It is important to note that GP-GAN was used in this study which presupposes that tumors would always grow. In addition, the shrinking tumors as a result of treatment are not taken into account in the method.

Hamghalam et al. [36] used a multistage attention-GAN to increase the contrast of the tumor image. In their work, tumors are segmented as whole tumors, tumor core, and enhancing tumor portions. The High Tissue Contrast (HTC) synthetic image takes a significant amount of time to generate 27ms. Due to the multistage architecture of their models, the computational complexity and number of parameters rise as the number of Regions of Interest (ROI)s increases. Rezaei et al. gave out a conditional Generative Adversarial Network (cGAN) approach for segmentation into three distinctive sub-regions namely the whole tumor, the core tumor, and the enhancing tumor region with different labels that are used further to evaluate the survival days of patients after tumor diagnosis [37]. Their model learned a loss that makes it efficient to work on unseen data. Han et al. designed a model to detect brain metastases at desired position using Progressive Growing of GAN (PGGAN)s and a highly rough bounding box [38]. They divided the image into smaller regions and tried to predict the bounding box and achieved great sensitivity in tumor detection by working on random shape rather than appropriate segmentation, spontaneously at desirable positions and sizes. But they failed to achieve high sensitivity by adding more synthetic images.

Mohamed et al. employed Near Infrared Imaging (IR) mechanism for detecting brain tumors with a size of less than 3mm (random arbitrary value) that couldn't be detected using Computed Tomography (CT) or MRI scans. They sent the thermal information via WSN [39]. It is shown here that the feature vector, Gray Level Co-occurrence Matrix (GLCM), SVM, statistical features, and Back Propagation Neural Network (BPNN).

Table 1

Recent research using deep learning for the classification and segmentation of brain tumor MRIs.

Ref.	Contribution	Limitations
Mohsen et al. [8]	Accomplished a DNN classifier integrated with PCA and DWT for categorizing tumors into 4 classes. MRIs were segmented using the FCM clustering	Classified a very small dataset of 66 brain MRIs
Özyurt et al. [9]	Implemented a method to categorize brain tumors using a hybrid NS-EMFSE-CNN classification technique. They acquired an accuracy of 95.62% by SVM	No specific tumor type classification has been carried out
Sajid et al. [10]	Devised a method that implemented a hybrid CNN with 2 path CNN and 3 path CNN using a patch-based technique considering local and contextual information	Segmented MRIs of HGG and LGG and no classification was performed
Emrah Irmak [11]	Put forward a model for classifying brain MRI images with Deep CNN and grid search optimizer-tuned hyperparameters with the motivation to build three different models for the classification of tumors	The second model achieved less accuracy compared to the other two models
Sultan et al. [12]	Offered an approach of using the CNN architecture for classifying brain MRIs into three distinct groups and distinguishing in different glioma grades	The used dataset is smaller and must be evaluated on a bigger group of datasets
Havaei et al. [13]	Proposed a CNN which is a two-path architecture that is capable of extracting local features along with global features of the brain tumors concurrently	Required to be observed on large datasets for efficiency of handling huge MRIs
Irsheidat et al. [18]	A generic CNN was implemented which predicted the presence of a tumor or not with an accuracy of 96.7% and 88.25% in validation and test data respectively	A binary classification is performed and used a smaller dataset to test their model
Naser et al. [20]	Incorporated a deep learning strategy in which CNN hinged on U-Net for segmenting tumor affected regions. A transfer learning VGG-16 model and classifiers has been developed to grade cancerous tumors	No independent dataset was available to compare with for testing and a moderately small LGG data has been used for validation
Ahmet Çinar et al. [24]	A method based on the Resnet-50 from CNN was proposed. The endmost five layers of the model are eliminated, and eight additional layers are inserted as a substitute	Only binary classification has been implemented

Table 1 (continued)

Ref.	Contribution	Limitations
Han et al. [38]	Designed s to detect brain metastases at desired position using PGGANs and a highly rough bounding box	Failed to achieve high sensitivity by adding more synthetic images
Naceur et al. [29]	An end-to-end automatic incremental CNN was proposed for the segmentation of tumors in which they implemented deep learning models entitled 2Cnet, 3Cnet, and EnsembleNet using their proposed training strategy	The model is solely suitable for the segmentation of MRI
Elazab et al. [35]	Devised the , which predicts the growth of glioma tumors at an early stage using stacked 3D GANs	Presupposes that tumors would always grow and the shrinking tumors as a result of treatment are not taken into account

produce a better output than other methods. They did segmentation for a 2D image which would lose some information. Irfan et al. employed pre-trained Inception-v3 for extracting the features which were then incorporated with Dominate Rotate Local Binary Pattern (DRLBP) for superior texture analysis [40]. The feature vectors were optimized employing the Particular Swarm Optimization (PSO) algorithm and classified using softmax classifier. The following are the article's major contributions:

- A classification model for tumor MRIs multiclassification has been proposed and evaluated on four separate individual datasets and two merged datasets that are the composite of these individual datasets. The classification model was tested against 5 pre-trained models.
- A segmentation model that accepts input MRIs and creates masked pictures has been suggested. Although the model was trained using a manually constructed mask from Merged dataset 1, it can also produce masks for other datasets.
- In terms of performance metrics and time, a deep discussion has been held between the performance of the effect of segmentation before sending the MRIs to the classification model.
- Merged dataset 1 has been manually segmented with expert assistance and consists of a large number of images divided into four groups (glioma, meningioma, pituitary and no tumor)

The rest of the paper is structured as follows: Section 2 includes a comprehensive review of all datasets and techniques, including information on single, merged, augmented, and segmented datasets, data preprocessing, and classification and segmentation methods. Section 3 details the results as well as the additional training and validation methods. Finally, the conclusion is appended in Section 4.

2. Datasets and Methodology

The datasets have been utilized in accordance with the classification and segmentation methods. Data has undergone preprocessing steps. And also while doing the data augmentation, certain preprocessing techniques are used. Over the classification datasets, both the suggested model and pre-trained architectures have been employed. Two scenarios are taken into consideration, the first in which the classification is done without the segmentation being done first. And in the second method, segmentation is followed by sending the MRIs to the classification model. Sending the segmented MRI to the classification model minimizes computational complexity and processing time since segmentation only extracts the tumored portion of an MRI. Figure 1 illustrates the workflow of the proposed methodology. Comparing the classification performance of trained and suggested models using a variety of metrics, as well as the performance of segmentation model, are done. Finally, the two methods of classifying tumors while segmenting them and classifying them without segmentation are compared.

2.1. Classification datasets

A total of four individual datasets have been outlined, and two combinations of these datasets have been integrated to create two new merged datasets. As a result, our work has been applied to a total of six datasets (four individual and two merged datasets). The four datasets are referred to as Dataset a, Dataset b, Dataset c, and Dataset d throughout the paper. Merged dataset 1 and Merged dataset 2 will be used to refer to the two merged datasets respectively.

2.1.1. Dataset a

There are 3064 T1-weighted Contrast Enhanced (CE)-MRI images in this dataset. The images are taken from 233 patients having three types of tumors: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices). This dataset has been widely used in most research papers [12, 14, 19, 22, 32, 33]. The data was gathered from Nanfang Hospital in Guangzhou, China, and General Hospital, Tianjin Medical University in China between 2005 and 2010. The dataset is made publicly available in [41].

2.1.2. Dataset b

This dataset comprises 3264 T1, T2 and Fluid-Attenuated Inversion Recovery (FLAIR) MRI images. The images are split up into two directories: training and testing. Glioma, meningioma, pituitary, and normal brain MRI are the four subclasses in each directory. Glioma (100 slices), meningioma (115 slices), pituitary (74 slices), and no tumor (105 slices) are all included in the testing directory. Glioma (826 slices), meningioma (822 slices), pituitary (827 slices), and no tumor (395 slices) are all listed in the training directory. The dataset is publicly available in [42].

2.1.3. Dataset c

This dataset includes a total of 10000 images of 3 classes of tumor MRIs and MRIs with no tumor. There are glioma (2500 slices), meningioma (2500 slices), pituitary (2500 slices) and MRIs with no tumor (2500 slices). This dataset is also publicly available at [43].

2.1.4. Dataset d

There are a total of 4292 images in this dataset. The dataset is made publicly available at [44]. The dataset is divided into two directories: training and testing. Each of these two directories contain four subclasses of glioma, meningioma, pituitary and no tumor. There are glioma(1038 slices), meningioma (1318 slices), pituitary (1255 slices) and normal brain MRIs (681 slices).

2.1.5. Merged dataset 1

This dataset combines Dataset a and Dataset b. Glioma, meningioma, pituitary, and normal brain MRIs will all be included in the combined dataset. Because the number of normal brain MRIs is low in comparison to the other three classes, Normal brain MRIs have been taken from this dataset [45]. The dataset [45] includes MRIs of normal and tumor-affected brain tissue. As a result, the Merged dataset 1 contains a total of 7022 images. Glioma (1621 slices), meningioma (1645 slices), pituitary (1757 slices), and normal brain MRIs (2000 slices) are included in the Merged dataset 1. The dataset is available at []. The sample images from this dataset is shown in Figure 4a.

2.1.6. Merged dataset 2

The Merged dataset 2 has been created by combining the Merged dataset 1 with dataset c and dataset d. As a result, this Merged dataset 2 combines all four datasets used in our work (Dataset a, Dataset b, Dataset c, and Dataset d) along with the dataset [45] having normal tumor images. Glioma (5159 slices), meningioma (5465 slices), normal brain MRI (5181 slices), and pituitary MRI (5512 slices) are all included in this dataset.

2.2. Augmented Merged Dataset

MRI images for 4 classes, glioma, meningioma, no tumor, and pituitary, are contained in the Merged dataset 1, which is splitted into two folders to train and test. The amount of images in the testing and training folder for every class is not equal. In the event of training, each tumor class has 5000 images, but in the event of testing, each class contains 1000 images. So the Augmented Merged Dataset contains 24,000 images. Because the dataset comprises images of varied sizes, they have been shrunk in the augmentation stage to a constant size of 256 by 256. Rotation, zooming, height and width shifting, shearing, horizontal flipping, and mode filling are the preprocessing methods employed while augmenting the MRI images. To keep maximal features in the images, the settings for rotation, zooming, shifting, and shearing have been modified for different classes of tumor in the training and testing folders. It also includes augmenting the dataset such



Figure 1 The suggested methodology's workflow entails the selection of the dataset, MRI preprocessing, MRI segmentation, models shown for classification, MRI classification classes, and ultimately results analysis over the findings.

that it has 80% of images for training the model whereas 20% for testing and validation.

2.3. Segmentation dataset

Merged dataset 1 has been manually segmented because Merged dataset 1 contains MRIs from four different classes. As after the MRIs have been segmented, they will be forwarded to

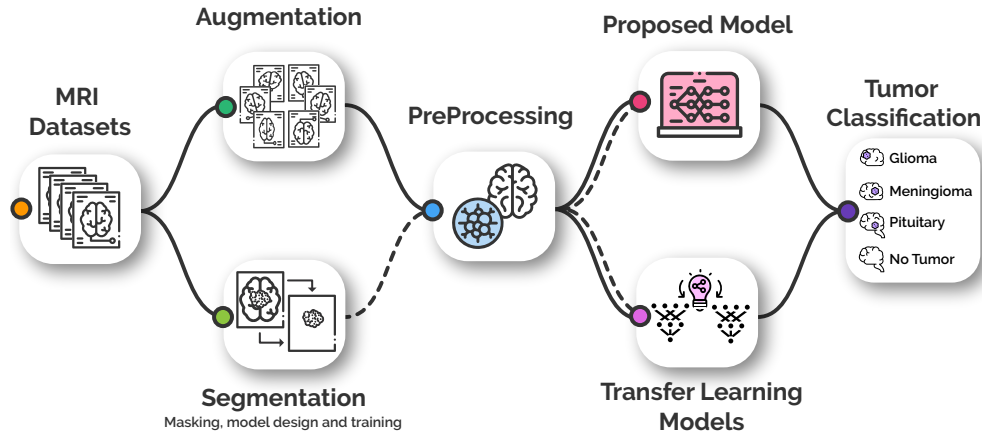


Figure 2 The classification technique depicted schematically, with MRIs segmented,preprocessed before passing them to proposed classification model and transfer learning models.

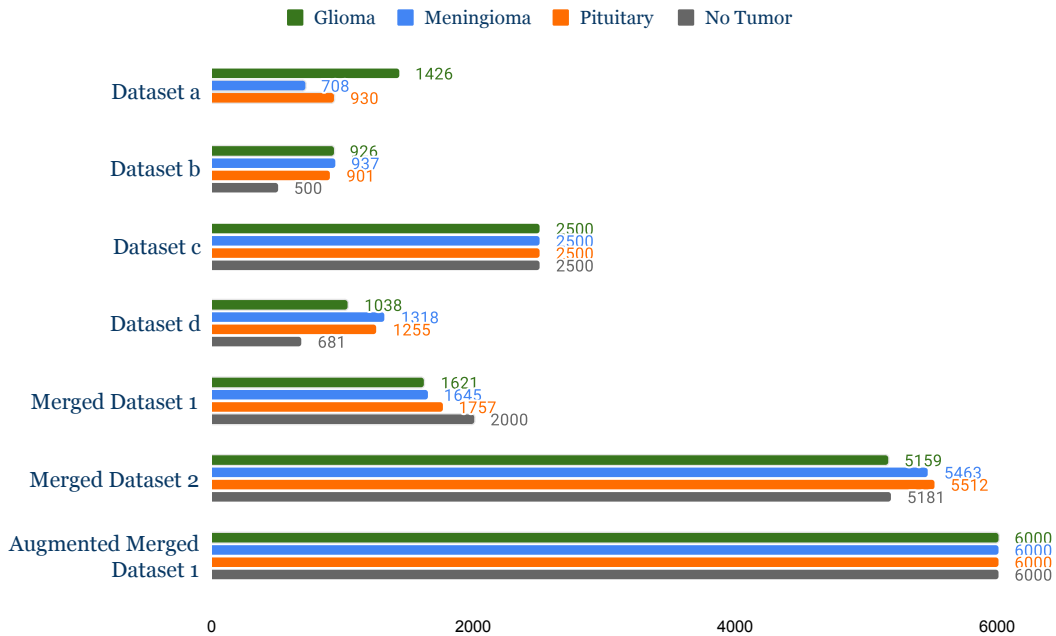


Figure 3 Dataset distribution for classification includes four individual datasets (Datasets a, b, c, and d), two merged datasets (Merged Dataset 1 combines three datasets, and Merged Dataset 2 combines Merged Dataset 1 and Datasets c and d).

a classification model where they will be divided into four categories: pituitary, glioma, and meningioma and no tumor. Four individuals manually segmented the MRIs of the four classes. The MRIs are inspected with an expert present before being segmented. For each MRIs, four individuals create four mask images. Then the masked MRIs have been provided back to the expert. Out of the four masks, the expert selects the one that is closest and then makes the required corrections. A corresponding generated mask exists for every MRI. The example images from this dataset are shown in Figure 4b.

2.4. Data preprocessing

Image resizing, RGB to grayscale conversion, unsharp masking, and sobel filtering [46] are some of the preprocessing techniques that have been used to the images. The MRIs are downsized to the size of $160 \times 160 \times 3$ in the preprocessing stage. They are transformed to grayscale using weighted average approach, which reduces its dimension to $160 \times 160 \times 1$. Unsharp masking with a radius of 100 pixels is applied to the grayscale MRI. Then the MRIs are subjected to the sobel filter.

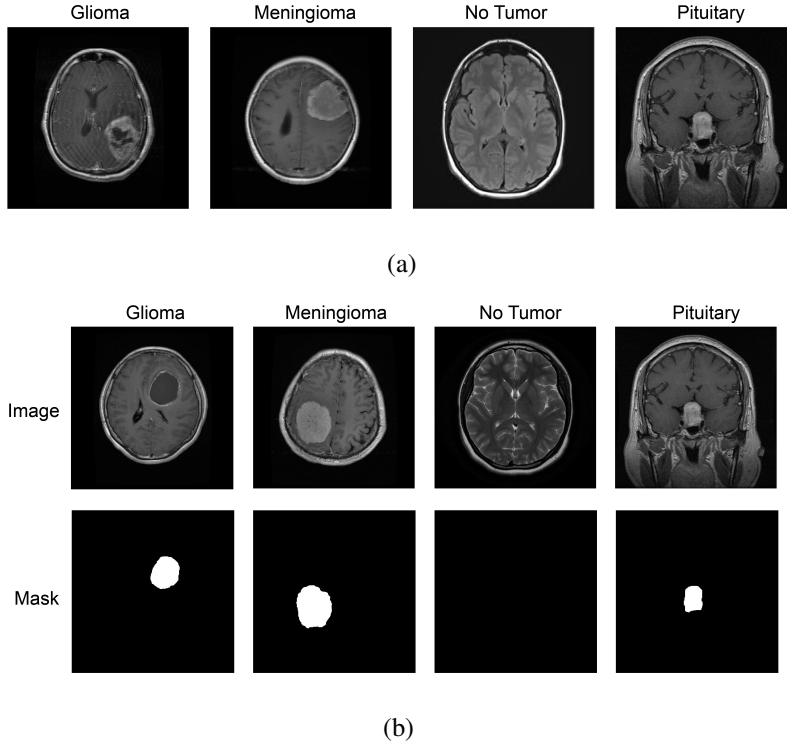


Figure 4 (a) Sample images from Merged Dataset 1 where MRIs are from four classes named: Glioma, Meningioma, No Tumor and Pituitary. (b) Segmentation dataset sample images where images with corresponding masks are generated manually.

2.4.1. Weighted average method

The equation representing the grayscale weighted average, Y

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

In equation (1), R , G and B are integers with values ranging from the 0 to 255 that indicate red (R), green (G) and blue (B).

2.4.2. Unsharp masking

Unsharp masking is a versatile and strong technique for improving image sharpness. A lower radius adds smaller-scale detail by affecting the size of the edges to be enhanced or the width of the edge rims. Threshold determines how much of a brightness shift will be sharpened. Blurred image is subtracted from original MRI, which is referred to as unsharp masking algorithm. An unsharp or blurred image is created by spatial filtering the original image using a Gaussian low-pass filter.

2.4.3. Sobel Filter

A basic 3×3 convolution is used in sobel filter. Separation of the sobel kernels is an additional optimization option. It is designed to operate on first order derivatives. It computes the MRI's first derivatives individually for the X and Y axes.

2.5. Classification models

A pre-trained model is a model that has already been trained on the ImageNet data set to perform a certain job. Since ImageNet includes 1000 classes, the pre-trained models have been

taught to handle a wide range of tasks. It takes less time and effort to create the architecture of a model that has already been trained. As the classification models, five pretrained architectures have been used. The pretrained architectures include VGG16 [47], VGG19 [47], EfficientNet B0 [48], and EfficientNet B7 [48], ResNet152v2 [49]. The pretrained models are shown in Figure 6. Along with the pretrained models, a proposed classification methodology has been also presented. Figure 7 shows the proposed model for classification.

2.5.1. VGG16

VGG16 is extensively applied convolutional neural network. This has achieved a lot of appeal in the research because of its straightforward methodology and the fact that pre-trained weights were publicly available online, allowing new tasks to be fine tuned in a simplest way possible. VGG16 has 16 convolutional layers and is well-liked for its relatively consistent architecture. 5 maxpooling layers, 2 fully connected layers, and 1 softmax layer make up the architecture in addition to the convolutional layers. For classification, a flatten layer, 2 batch normalization layers, 1 dropout with the dropout rate of 25%, and 2 dense layers are substituted for the fully connected layers and softmax layer in the modified VGG16 architecture shown in figure 6a, which accepts input of size 160×160 .

2.5.2. VGG19

One of the VGG [47] architectures, VGG19, has 16 convolutional layers, 5 maxpool layers, 3 fully connected layers, and

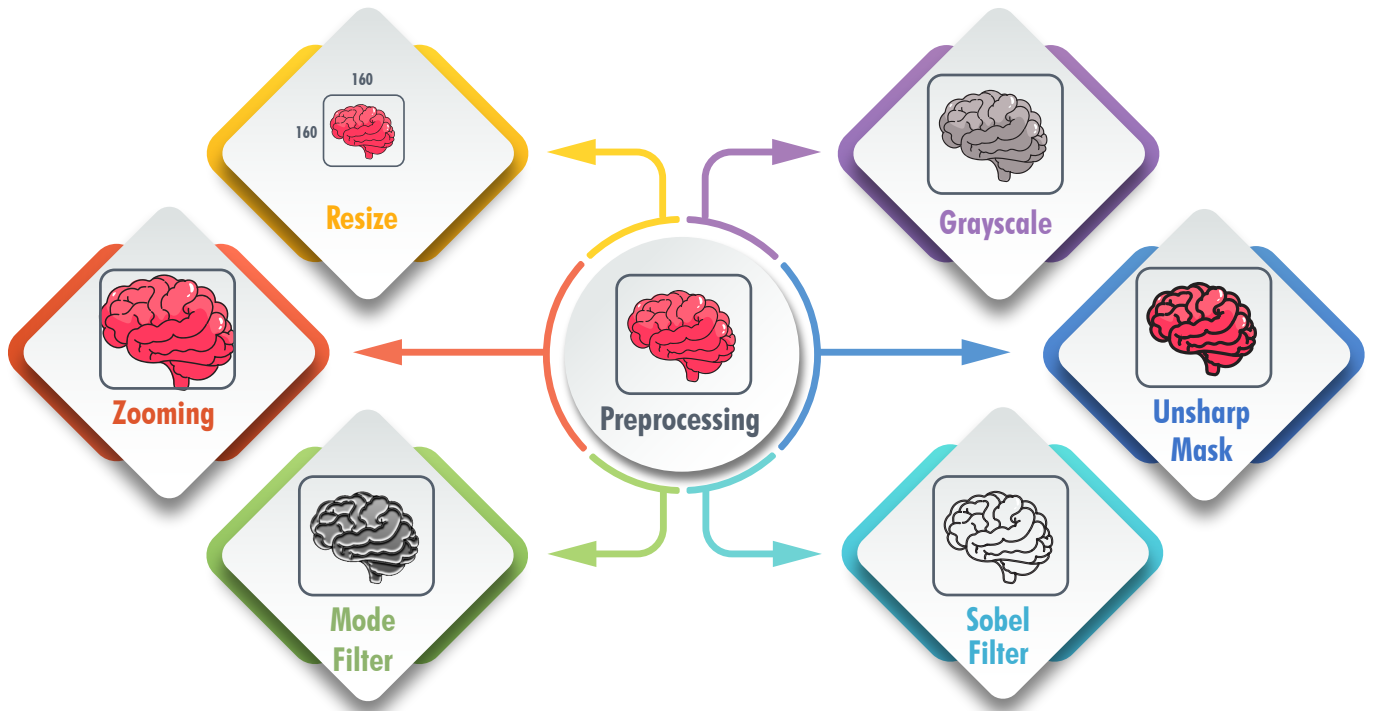


Figure 5 Input MRIs are preprocessed before being sent into the classification model and preprocessing methods are used during augmentation too.

1 softmax layer. The modified VGG19 architecture is made up of these convolutional and max pooling layers, and the fully connected layers and softmax layer are replaced with 6 layers that are arranged as flatten, batch normalization, dense, dropout layer with a dropout rate of 25%, and finally a dense layer to classify MRIs into four classes. Figure 6b depicts the modified architecture that has been tailored for MRI classification.

2.5.3. EfficientNet B0

In order to maximize accuracy and floating-point calculations, the researchers used a multi-objective neural network search to create EfficientNet B0 [48]. EfficientNet B0 to EfficientNet B7 have been developed using B0 as a baseline model, and they have attained the highest level of accuracy on ImageNet while being significantly more efficient than its rivals. There are 237 layers in EfficientNet B0. To classify MRIs using modified EfficientNet B0 architecture, the flatten, batch normalization, dense, batch normalization, dropout layer with a dropout rate of 25%, and lastly a dense layer have been added.

2.5.4. EfficientNet B7

EfficientNet B7 [48] is a convolutional neural network that makes use of compound coefficient for the uniform scaling of width, height and resolution. The scaling technique of EfficientNet B7 invariably and systematically enlarges the dimension of the network and resolution with a series of predefined scaling coefficients. Though in standard practice these factors are adjusted arbitrarily. There are 813 layers in EfficientNet B7. In the modified EfficientNet B7 architecture, the flatten, batch

normalization, dense, batch normalization, dropout layer, and finally a dense layer have been integrated to the base EfficientNet B7, and the input size is also taken into consideration to be 160×160 in this case.

2.5.5. Resnet152V2

ResNet [49] has become a game changer because it efficiently and considerably trains deep neural network. ResNet operate on the premise of building deeper networks than conventional simple networks while also determining the optimal amount of layers. These residual networks may get accuracy from far more depth and are simpler to tune. Before Resnet152V2, training operation for deep neural network was tough for vanishing gradient problem. The Resnet152V2 has 564 layers in all. The modified Resnet152V2 design adds 6 extra layers to the basic Resnet152V2 architecture shown in 6c.

2.5.6. Proposed classification model

Figure 7 shows the classification model that has been proposed. The model comprises 39 layers with two blocks between the input and output layer. The output layer predicts MRIs as glioma, meningioma, pituitary or no tumor. The input image is processed via convolution layers, activation functions followed by the convolutional layers which are used for selecting the features, maxpooling, batch normalization, and dropout layers. Dropout layers are employed to prevent overfitting. To forecast the output, the dense layers and softmax layer are utilized.

The first layer provides information about the input size. The input MRI size to the model is 160×160 . Following that is

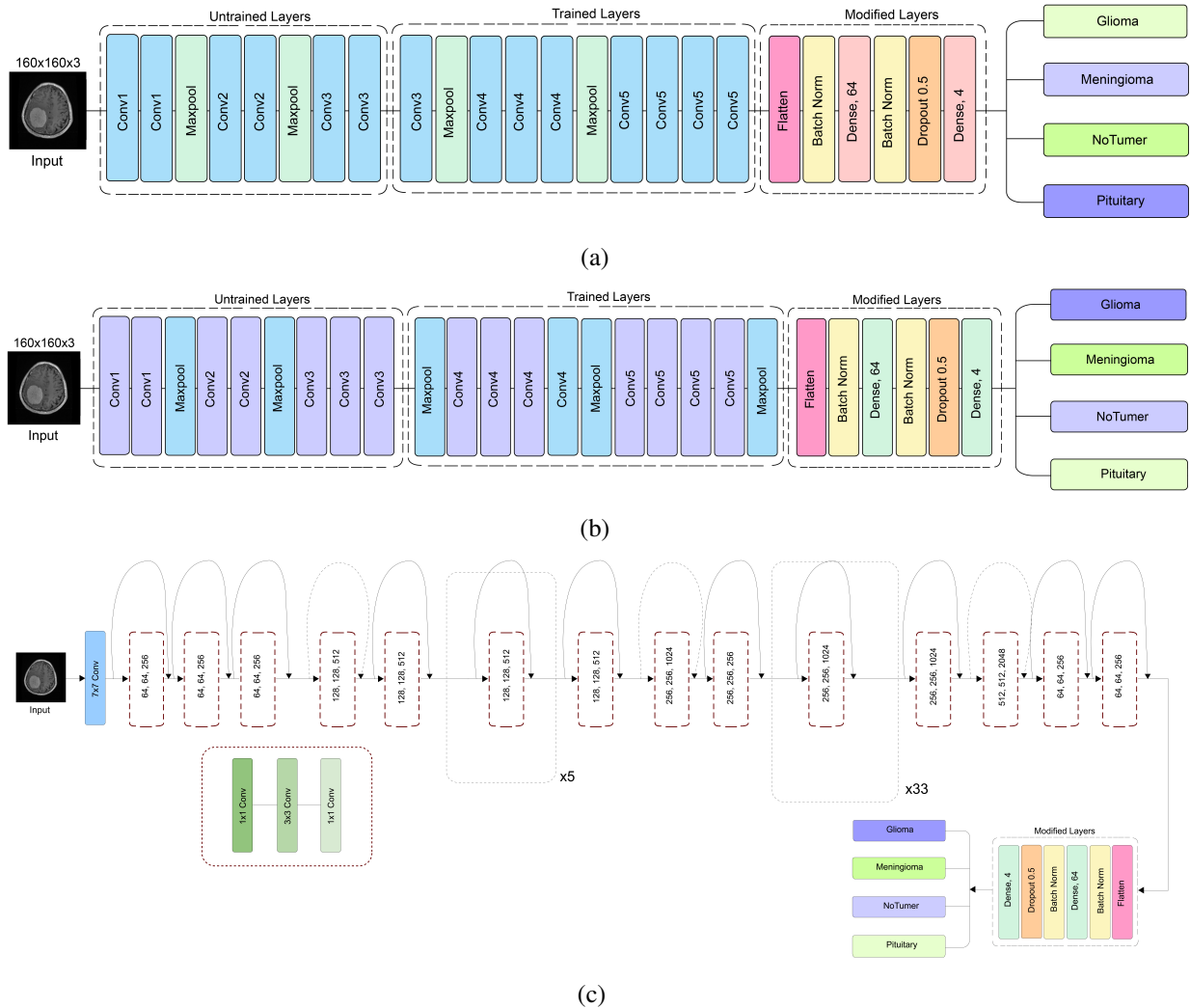


Figure 6 Pretrained architectures incorporated in this work, (a) Modified VGG16, (b) Modified VGG19, (c) Modified ResNet152V2

the convolution layer. A convolutional layer is operated by performing a convolution operation on the original image and the kernel. In the scenario of a 2D convolutional layer employs a number of filters, with the kernel moving horizontally or vertically over the image in a given number of steps known as strides. The filter size for the convolution layer in between input and the first block is 3×3 with 32 filters. After the convolution layer a maxpooling layer is added with a pool size of 2×2 . The maxpooling layer is used to downsample the image. Maxpooling is used to discard less significant data while simultaneously addressing the issue of overfitting. After the pooling layer, the batch normalization and dropout layer (dropout rate 25%) is added.

The architecture of the two blocks in the model is identical, with the exception that the first block uses 64 filters while the second block uses 128 kernels. Each of the blocks has 3 paths. Each path has a distinct kernel size for the convolution layers, which is 1×1 , 3×3 . For the max pooling operation, a 2×2 kernel is utilized, and one of the three pathways has a dropout

layer added with a 25% dropout rate. The paths guarantee that layers are applied at the same level. To make the model wider rather than deeper and to lessen the computing complexity of the model, layers with varying kernel sizes are used at the same level. The three pathways each extract a different set of features which are then combined at the end of each block.

Two convolution layers with 5×5 and 3×3 filter sizes are integrated between the second block and the output layer. An additional maxpool layer with a 2×2 filter size is added after the convolution layer. Then the dense layers, dropout, flatten, and batch normalization are included. The dropout rate for the latter two layers is 50 percent. Dropout layers are introduced to speed up the training process. The dense layer is followed by softmax. The dense layer's output is delivered to the softmax activation function. The softmax activation function is used to calculate the relative probabilities of having a certain type of tumor. The proposed model enables 2,051,872 worth of parameters to be trained.

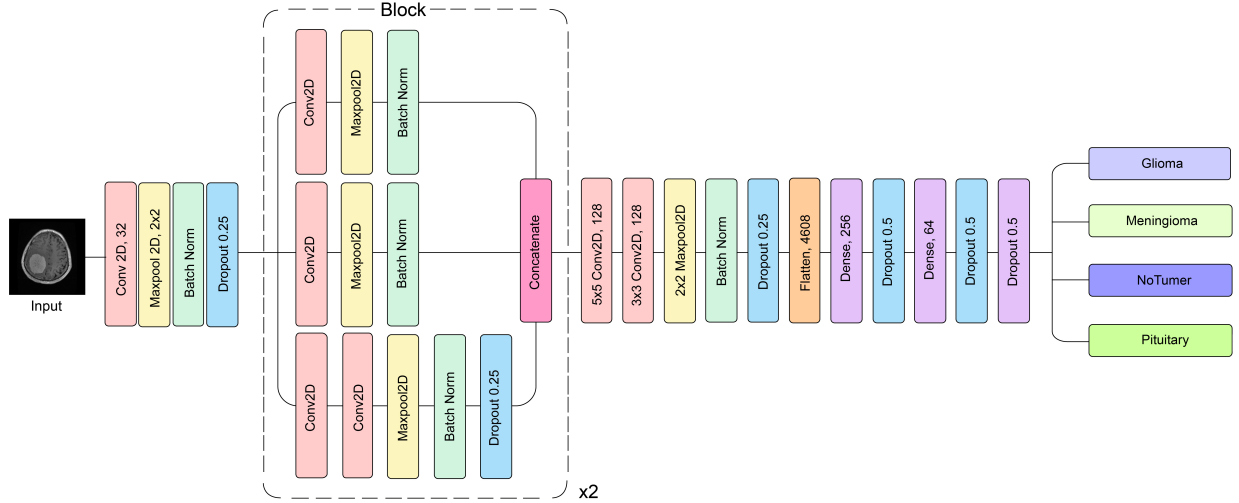


Figure 7 The proposed classification model which consists of 39 layers and includes two blocks, classifies MRIs into four classes: Glioma, Meningioma, No Tumor and Pituitary.

Table 2 The descriptions for each layer in Proposed classification model having two blocks and total of 39 layers.

Block No.	Layer name	Layer count	Filter count	Filter size	
Block 1	Convolution	1	32	3×3	
	Maxpooling	1		2×2	
	Batch Normalization	1			
	Dropout	1			
	Convolution	4	64	1×1, 3×3	
	Maxpooling	3		2×2	
	Batch Normalization	3			
	Dropout	1			
	Block 2	Convolution	4	128	1×1, 3×3
		Maxpooling	3		2×2
Batch Normalization		3			
Dropout		1			
Convolution		2	128	5×5, 3×3	
Maxpooling		1	128	2×2	
Batch Normalization		1			
Dropout		3			
Flatten		1			
Dense		3			

2.6. Hyper parameters for proposed classification model

A lot of trial and error is required when tuning the hyper-parameters. While the model is being trained, these hyper-parameters act as controls that may be changed. It is determined what these hyper-parameters should be set at in order to achieve the optimum results. The the number of layers, neurons, optimizer, input and output activation functions, batch size, the number of epochs and loss function have all been taken into consideration as model hyper parameters.

Adaptive Moment Estimation (Adam) [50] is chosen as the optimizer. When dealing with complex problems requiring a

lot of data or factors, this strategy is incredibly effective. And this strategy uses minimal memory also. In Adam, for a specific iteration (t) moving averages are dependent on the parameters: exponential decay rates for the first moment and second moment estimates which are denoted by β_1 and β_2 respectively, and gradient (gt). Equation (2) represents the bias correction formula for moving averages’.

$$\hat{p}_t = \frac{p_t}{1 - \beta_1^t}, \hat{q}_t = \frac{q_t}{1 - \beta_2^t} \quad (2)$$

In Equation (2), \hat{p}_t and \hat{q}_t represents first and second moment vector respectively. Then the weights and biases are updated accordingly. In order to find the best parameter configuration, the suggested model uses the Grid Search technique. All conceivable parameter combinations are tested using grid search.

The output layer uses softmax for multi classification while the hidden layers employ the Rectified Linear Unit (ReLU) [51] activation function. equation (3) and equation (4) represent ReLU and Softmax activation functions respectively.

$$f(x) = \max(0, x) \quad (3)$$

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (4)$$

In equation (3), any negative input causes the function to return 0, while any positive value x causes it to return that value. As a result, it generates an output whose range is 0 to infinite. In equation (4), the values from the output layer’s neurons are represented by the x. The non-linear function is represented by the exponential. After being normalized, these values are divided by the total of exponential values and then transformed into probabilities.

The batch size is fixed at 64 and the epoch number is 100. Categorical cross entropy have been used for multi classification. This loss is an excellent indicator of how easily two discrete probability distributions may be distinguished from one

another. Equation (5) denotes the cross entropy loss formula.

$$CELoss = - \sum_i^C gt_i * \log(s_i) \quad (5)$$

In equation (4), the score and ground truth are denoted by s_i and gt_i respectively for each class i in C .

All of the parameters are changed in an ad-hoc manner to figure out the best combination. Table 3 represents the hyperparameters utilized in the proposed model.

Table 3 Hyperparameters of the proposed classification model

Hyperparameters	Values
Number of trainable parameters	2,051,872
Optimizer	Adam
Searching strategy	Grid search
Output layer activation function	Softmax
Hidden layer activation function	ReLU
Batch size	64
Number of epochs	100
Loss function	Categorical cross entropy

2.7. Segmentation model

Ronneberger et al. established the U-Net structure [52]. The proposed segmentation model is based on U-Net. It has been commonly exploited for segmenting images in medical field and it has also showed competitive performance.

The U-Net architecture is divided by two paths: downsampling and upsampling. Figure 8 illustrates the architecture. The image’s context is captured via the downsampling route. The standard stack of convolutional and max pooling layers is all that is used for the downsampling procedure. The second method, also called as upsampling path, is symmetric expanding and allows for exact localisation using transposed convolutions.

The model’s first convolution layer accepts input of $160 \times 160 \times 3$. This convolution layer produces an output with a volume of $160 \times 160 \times 32$ by using 32 kernels of size 3×3 . Because padding is equal to 1 is maintained, the size of the output feature-maps matches that of the input feature-maps. Initializing the weights of the kernels uses a uniform distribution in the proposed segmentation. The area in the input volume that a certain feature extractor or kernel encompasses at a given time is known as the receptive field or context. The pooling process takes place between two layers in order to make the feature map smaller and transmit fewer parameters through the model. In the suggested model, the first max pooling layer uses a feature map with a size of $160 \times 160 \times 32$ and produces a feature map with a size of $80 \times 80 \times 32$. In this instance, the pool size is 2×2 . This procedure is carried out to keep the features that best capture the MRI’s context. In this manner, the pooling procedure in the downsampling route reduces the MRI size. The size of the MRI decreases as the network becomes deeper, while

the receptive field expands at the same time. The number of kernels is rising in each stage of the downsampling process to extract increasingly intricate features from the input MRI. The image is upsampled in the upsampling layer using the transposed convolution approach. The feature map is routed via a 2×2 deconvolutional layer with strides equal to 2. It is transmitted through two 3×3 convolutional layers, much like in the downsampling path, and is sequentially concatenated with the prior feature map. The aforementioned procedure is continued until an image of size $160 \times 160 \times 32$ is obtained, at which point a 1×1 convolution layer is applied to produce an output of size $160 \times 160 \times 3$. The detailed description for each layer of the proposed segmentation model is given in Table 4

The corresponding MRI and mask are used by this model from the manually masked segmentation dataset. For each of the four MRI classes that the dataset corresponds to, the U-Net model will be trained.

2.8. Hyper parameters for segmentation model

The optimizer, input and output activation functions, loss function, batch size, number of epochs, training-validation splitting, and early-stopping patience have all been taken into account as model hyperparameters for the proposed segmentation model.

The optimizer used is Adam [50]. ReLu [51] activation functions are used with convolution in the hidden layers of the model. The last convolution layer that generates the output makes use of the sigmoid activation function. The segmentation model’s last layer has a binary cross-entropy function and a sigmoid activation function. The batch size in the segmentation model is 32, and 200 epochs are taken into account. The proportion between training and validation is 80% and 20%. Early stopping patience is maintained at 100. The hyperparameter values are represented in Table 5.

2.9. Classification approaches

Two categorization methods were suggested by the methodology used in this research.

The first approach extracts the MRI from the dataset and sends it for preprocessing shown in Figure 5. The augmentation stage preprocessing also applies to the MRIs. The classification of MRIs into the four groups of glioma, meningioma, no tumor, and pituitary is then carried out by both the proposed model and the pretrained networks. In this method, segmentation is not done before the classification goal.

The conventional approach is to first segment the image before sending it for classification. As segmentation only pulls features from the MRI of the areas with the tumor. The time and computational cost are both decreased if only the tumored area of the MRI is supplied to the classification model as opposed to the entire image. Consequently, in the second method, the MRIs are delivered to the classification model after being segmented into the tumor sections.

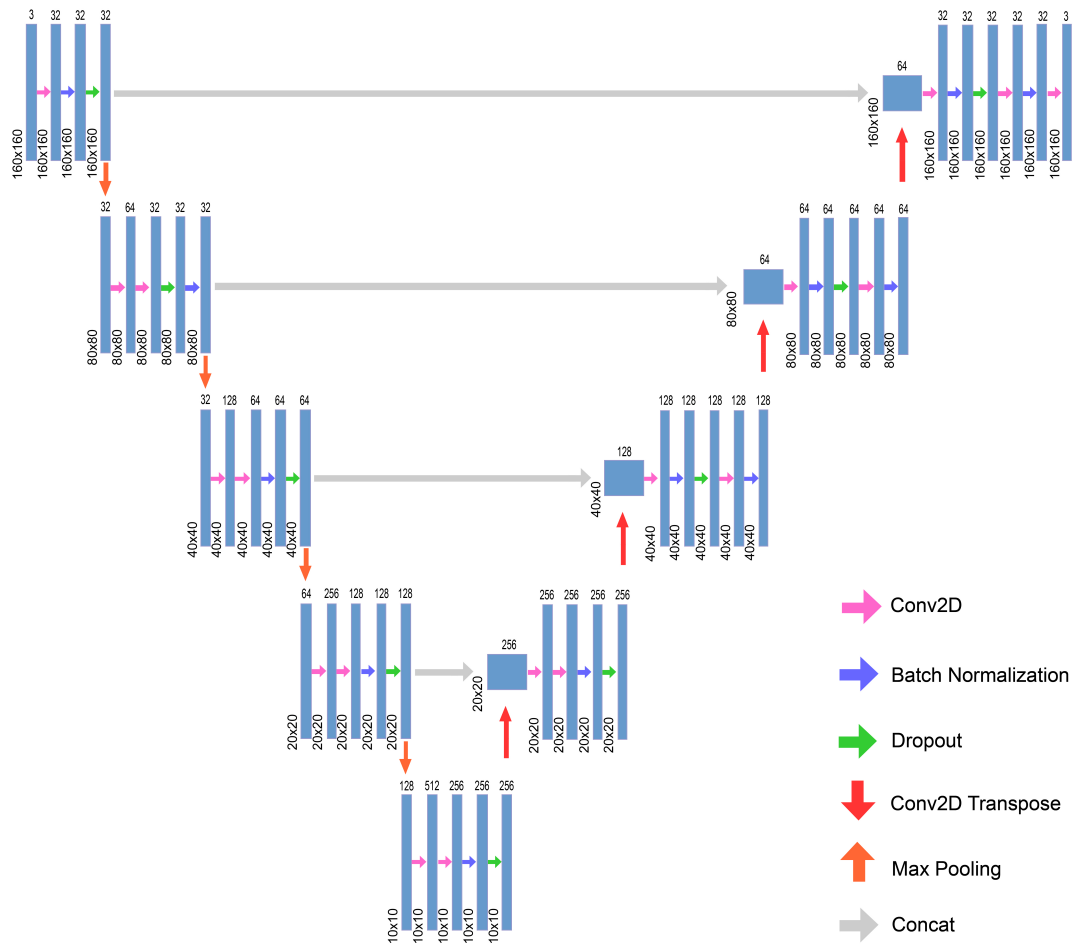


Figure 8 The suggested segmentation model is based on the U-Net architecture and contains a downsampling and an upsampling route that accepts input MRIs and outputs a segmented mask image.

3. Results

The experiments are divided into a number of sections, including comparisons and individual outcomes. Each DL model's and technique's effects have been examined independently, taking performance into account for each epoch. For training and testing, a total of 6 classification datasets were used. Among those, 4 are individual datasets (Dataset a [41], Dataset b [42], Dataset c [43], Dataset d [44]), and 2 are concatenated datasets (Merged dataset 1 and Merged dataset 2). However, one segmentation dataset has been employed during segmentation, which includes custom mask for the Merged dataset 1 mentioned before. Five transfer learning models, including the neural network model that has been proposed, are trained and tested on all of the classification datasets and a different segmentation model, has been used throughout the studies. Results of data augmentation and comparison between the results of segmentation in brain tumor classification also presented separately.

3.1. Experimental setup

The experiment has been run on both cloud services like google collab and using personal computing devices. Since

the scope of the experiment is quite huge, which includes executing multiple transfer learning and proposed neural networks on a variety of classification and segmentation dataset, simultaneously running them were crucial. The neural networks are built using python 3.9 with several packages such as Tensorflow, Sklearn, Numpy, and more.

3.2. Performance metrics

Accuracy is considered as a significant evaluation metric in the classification purpose. Precision indicates how much a classifier can be trusted when it indicates that an instance belongs to the positive class. A high precision rating indicates that there are very few false positives and the classifier is very strict in the criteria for classifying something as positive. The ratio of all instances properly classified in the positive class to the total number of real members of the positive class is defined as recall. In other words, it indicates how many of the total number of positive instances are classified properly. A metric for sensitivity and specificity is called ROC AUC. The ROC is a graph that depicts the relationship between True Positive Rate (TPR) and False Positive Rate (FPR). The resultant score, known as the

Table 4 The description for each layer of the proposed segmentation model based on the U-net architecture having downsampling and upsampling paths

Layer No.	Layer name	Filter count	Filter size	Dropout rate
Layer 1	Convolution+ReLU	32	3 × 3	10%
	Batch Normalization			
	Dropout			
	Maxpooling			
Layer 2	Convolution+ReLU	64	3 × 3	20%
	Convolution+ReLU			
	Dropout			
	Batch Normalization			
Layer 3	Convolution+ReLU	128	3 × 3	20%
	Convolution+ReLU			
	Batch Normalization			
	Dropout			
Layer 4	Convolution+ReLU	256	3 × 3	20%
	Convolution+ReLU			
	Batch Normalization			
	Dropout			
Layer 5	Convolution+ReLU	512	3 × 3	30%
	Convolution+ReLU			
	Batch Normalization			
	Dropout			
Layer 4	Conv2DTranspose	128	2 × 2	20%
	Convolution+ReLU			
	Convolution+ReLU			
	Batch Normalization			
Layer 3	Conv2DTranspose	64	2 × 2	20%
	Convolution+ReLU			
	Batch Normalization			
	Dropout			
Layer 2	Convolution+ReLU	64	3 × 3	20%
	Batch Normalization			
	Dropout			
	Convolution+ReLU			
Layer 1	Conv2DTranspose	32	2 × 2	10%
	Convolution+ReLU			
	Batch Normalization			
	Dropout			
Layer 1	Convolution+ReLU	3	1 × 1	
	Batch Normalization			

ROC AUC score, is the area under this ROC curve and broadly illustrates how well the model can predict classes.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (6)$$

Table 5 Hyperparameters of proposed segmentation model

Hyperparameters	Values
Number of trainable parameters	4,162,499
Batch size	32
Number of epochs	200
Optimizer	Adam
Hidden layer activation function	ReLU
Output layer activation function	Sigmoid
Early stopping patience	100
Training/validation split	80/20%
Loss function	Binary cross entropy

$$Precision = \frac{TP}{(TP + FP)} \quad (7)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (8)$$

The terms TP, TF, FP, and FN in equations refer to the number of true positive predictions, true negative predictions, false positive predictions, and false negative predictions, respectively.

The F1 score, also known as the F-measure, is specified as the harmonic mean of recall and precision and has been used as a classification evaluation metric. It is a statistical measure of a model's accuracy. It is mathematically represented as follows:

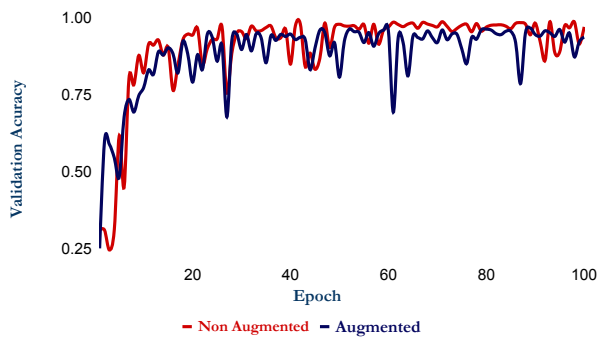
$$F1Score = \frac{2 * Precision * Recall}{(Precision + Recall)} \quad (9)$$

Dice coefficient, Dice loss, Binary cross entropy dice loss and Intersection Over Union (IOU), recall, precision has been considered as the evaluation metrics for the segmentation performance metrics. IOU also referred to as the Jaccard Index, is the area of overlap between the ground truth and the predicted segmentation and divided by the area of union between the ground truth and the predicted segmentation.

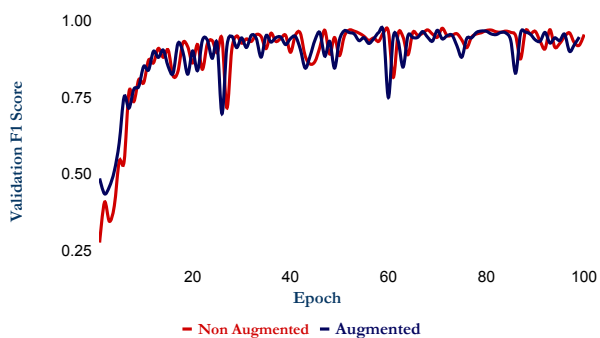
$$JaccardIndex = \frac{AreaOfOverlap}{AreaOfUnion} \quad (10)$$

3.3. Effect of Augmentation

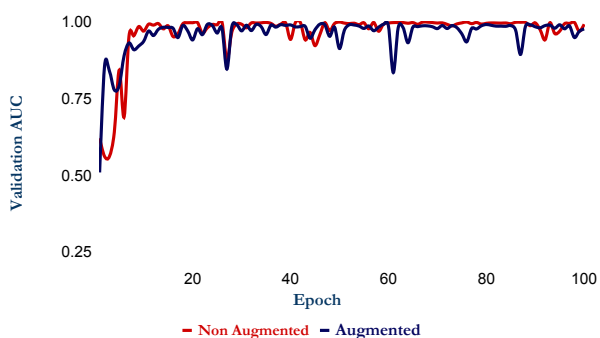
Data augmentation plays an important role in case of an imbalanced dataset or a lower quantity of data. Since brain tumor MRI data are hard to find and labeling these data need an expert's opinion, the amount of data along with the number of datasets are pretty rare. To observe the effect of augmentation, three merge dataset, which was created by concatenating three individual datasets and had imbalance ie: with different amount of class-wise MRI, was passed through several operations to create and augmented datasets. These operations include Rotation, zooming, height and width shifting, shearing, horizontal flipping, and mode filtering. To show the effect of augmentation, epoch-wise validation accuracy, F1 Score and AUC of the augmented and non-augmented dataset has been depicted in Figure 9a, Figure 9b and Figure 9c respectively. All of these



(a)



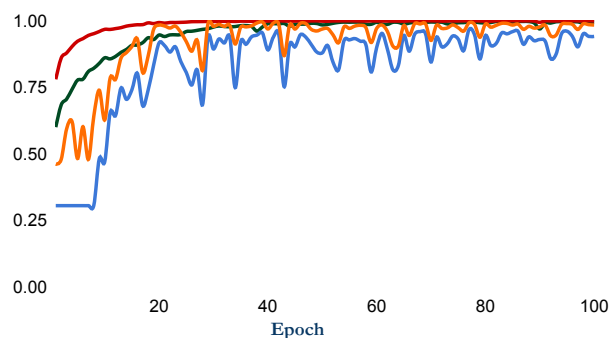
(b)



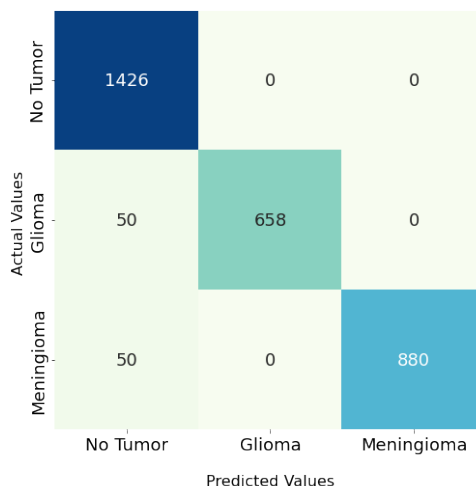
(c)

Figure 9 Line charts showing suggested classification model validation accuracy, F1 score, and AUC vs epochs using Merged dataset 1 and Augmented Merged dataset 1. (a) The chart showing the validation accuracy comparison for these two datasets (b) the validation F1 score comparison (c) the validation AUC comparison

figures represent similar trends and values while having more oscillations in the case of augmented data. Specially for validation accuracy Figure 9a augmented results are a little below the non-augmented ones, with additional fluctuations or drops, with 0% lower validation accuracy. The validation F1 score Figure 9b and AUC Figure 9c were almost identical epoch-wise with slight changes for AUC values in some epochs, and to our surprise, there is not much distinction between the augmented datasets and non-augmented ones, even in the last epochs the results of augmentations have deteriorated. Since the 5 merge



(a)



(b)

Figure 10 (a) Graph showing training and validation accuracy and AUC comparison for Dataset a of the proposed classification model (b) The confusion matrix for Dataset a showing the number of actual and predicted images for each class of the proposed classification model

augmentation has almost the same results that is not included in the section.

3.4. Impact of datasets in our proposed classification model

The proposed model has been tested and validated on all of the individuals and merged datasets. Throughout the study, test sets had been kept constant to ensure clear comparisons. This part of the experiments are concerned with the effect of different datasets in our proposed models. To represent that effect, handful of line chart and confusion matrix for each datasets validation part is given.

The proposed model has been trained and tested for 100 epochs. The line chart in Figure 10a illustrates the training and validation accuracy and AUC over the epochs for Dataset a. Since the proposed model is deployed for the classification task, AUC is given since it represents the separability of tumor classes. The highest validation accuracy is 96.73 percent with a training accuracy of 99.83 percent. On the other hand, both the

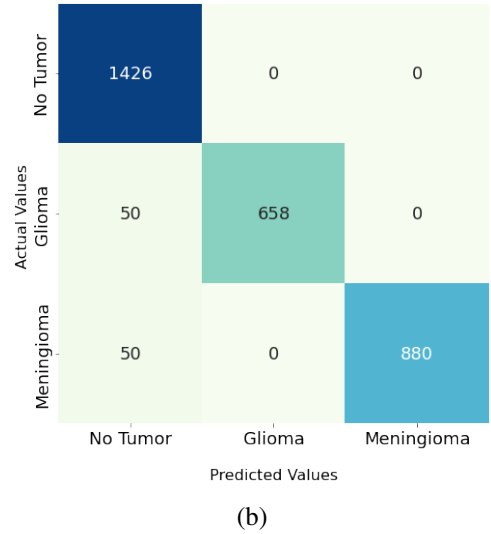
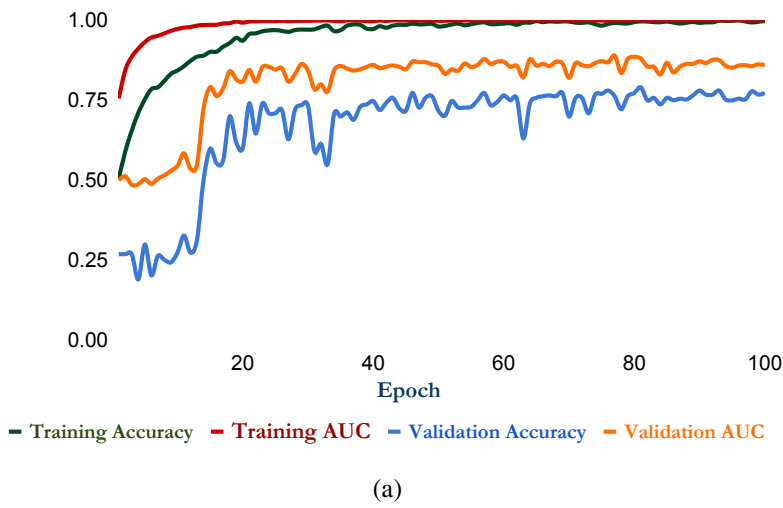


Figure 11 (a) Line chart comparing training and validation accuracy and AUC for the proposed classification model on Dataset b. (b) Dataset b’s confusion matrix, displaying the number of actual and predicted images for each class of the proposed model.

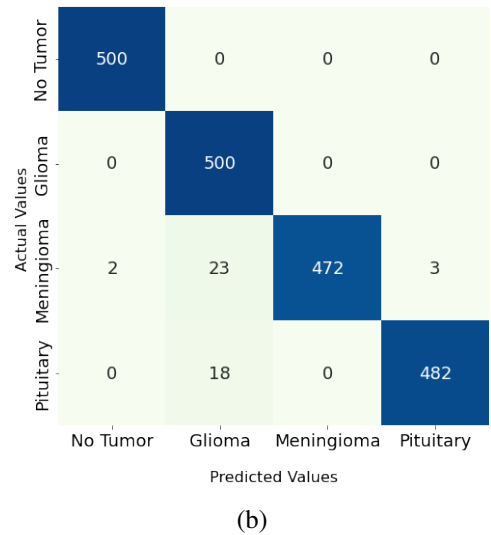
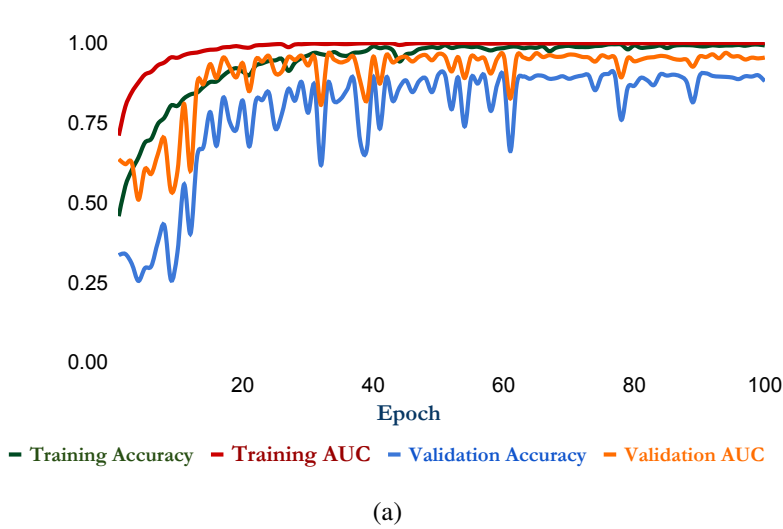


Figure 12 (a) Graph comparing training and validation accuracy and AUC for Dataset c of the proposed classification model. (b) The confusion matrix for Dataset c, which shows the number of actual and predicted images for each class.

training and validation AUC are quite high at 99.99 and 99.56 percent respectively. The training validation trends for both of these metrics are quite similar, with some minor oscillations. The next Figure 10b showcases the confusion matrix for the testing data where the actual vs the predicted value is given. Only a few of the glioma and meningioma were misclassified, although these numbers are very low compared to the correct classification.

Similarly, for Dataset b, epoch-wise accuracy and AUC curve have been illustrated in Figure 11a, which have a slight difference between the training and validation results. Since this dataset was imbalanced, without the augmentation process there was some deviation in each epoch, although the fluctuations were minimal. The maximum training accuracy is 99.93 percent and the AUC is 99.99 percent whereas the validation

AUC fell down to 88.85 percent. In Figure 11b the corresponding validation confusion matrix is given, where all of the classes except pituitary have some misclassifications.

Again in Figure 12a, same as previous, the accuracy and AUC curve is illustrated whereas Figure 12b delineated the confusion matrix for Dataset c. Similar to other single datasets there were some ups and downs in the validation curves as training progresses the maximum results were quite high, as the training and validation AUC are 99.9% and 96.88% respectively. Again in the confusion matrix, the correct classifications are way higher than the false ones. Dataset d also follows a similar pattern to Dataset c which is shown in Figure 13a and Figure 13b. Figure 13a shows similar oscillations, although that stabilize in the end, and the validation AUC has the value of 0.9688. On the other hand, Figure 13b depicts the confusing matrix for

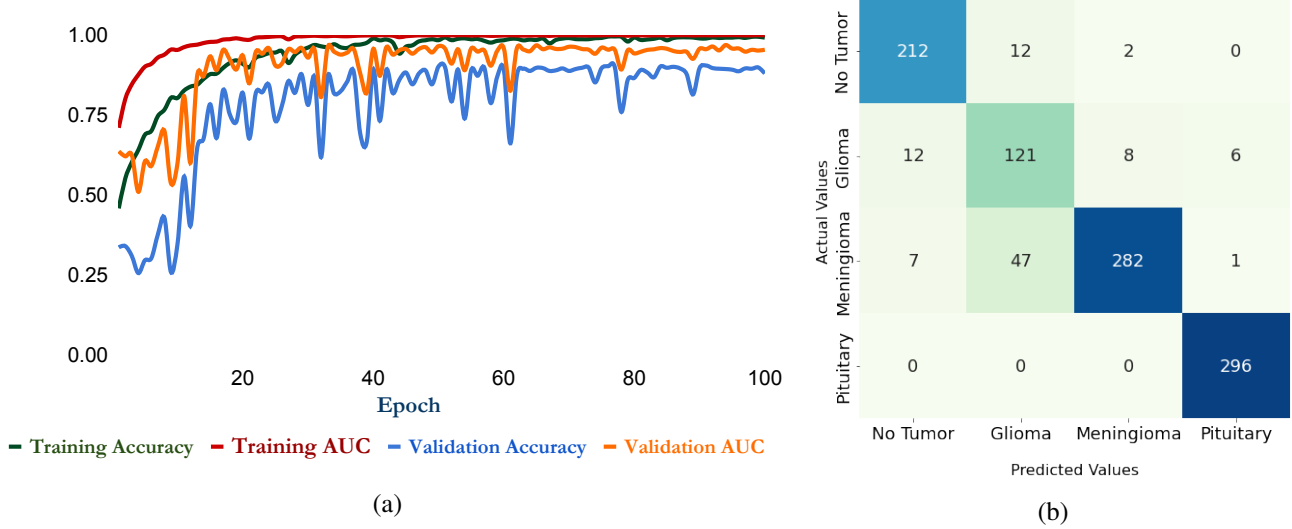


Figure 13 (a) Graph showing training and validation accuracy and AUC comparison for Dataset d of the proposed classification model (b) The confusion matrix for Dataset d showing the number of actual and predicted images for each class of the proposed classification model

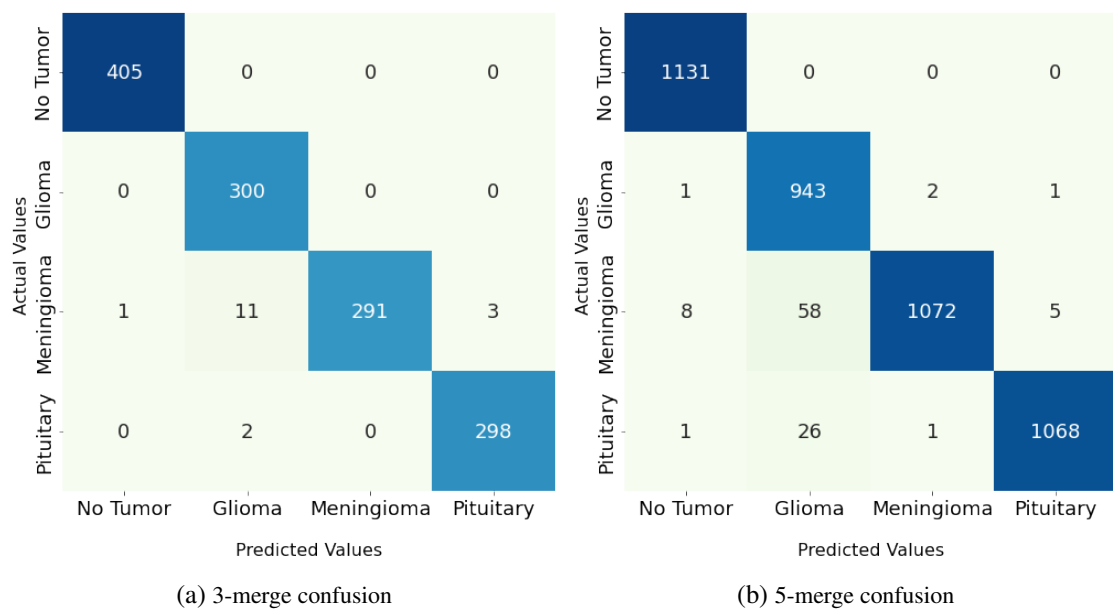


Figure 14 Confusion matrix illustrating the number of actual and predicted MRIs following application of the proposed classification model . (a) Merged dataset 1, (b) Merged dataset 2

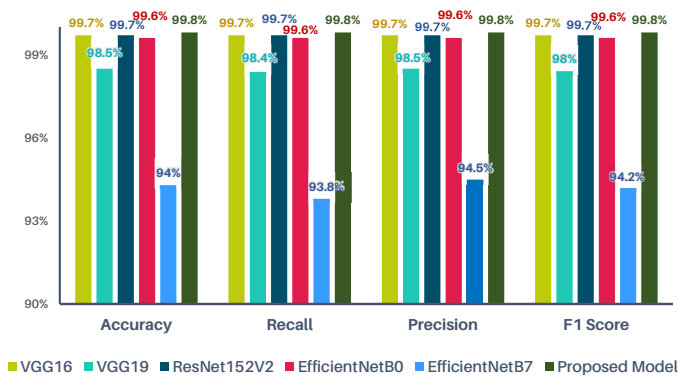
Dataset d, where glioma is the overall less correctly predicted class.

3.5. Classification comparison on two merged datasets

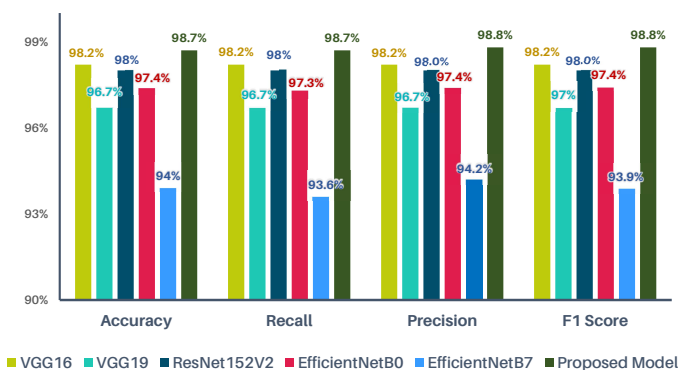
To better understand the impact on datasets for performance, two merged datasets were also validated on the proposed model. Figure 14a and Figure 14b demonstrate the validation confusion matrix for Merged dataset 1 and Merged dataset 2. Since the merge dataset is slightly imbalanced, there were differences in the count of separate classes, however most of the predicted classes are which can be seen from Figure 14a. Comparably, in

Figure 14b, there are only a few misclassifications, in total 93, compared to correct classification which is around 4200.

Comparative analysis of various evaluation metrics for different transfer learning models and the proposed model is illustrated in Figure 9 to Figure 19. The first comparison is done on the Merged dataset 1, where training and testing accuracy of proposed model along with VGG16, VGG19, Resnet152v2, EfficientNet B0, EfficientNet B7 is given. Although a lot more transfer learning models have been tested for this purpose, the above-mentioned models predicted surprisingly well, sometimes on par with the proposed model, compared to others. Fig-



(a)



(b)

Figure 15 Performance metrics for 5 transfer learning techniques vs the proposed approach (a) For the situation of training Merged dataset 1, where the metrics values are relatively near (b) For validation of Merged dataset 1, where the proposed model outperforms.

Figure 15a represents the comparisons of training metrics, where the proposed model achieved an astonishing 99.8 percent accuracy and F1 scores. The results of VGG16, ResNet152v2, EfficientNet B0 are also quite similar, touching the 99 percent mark. On the contrary, in case of testing accuracy in Figure 15b and F1 score, EfficientNet B0 comes down to 97 percent. Again, for testing metrics the proposed model achieved higher value than any other models, at a value of 98.7%, which is also really close to train accuracy, providing the fact that the model was well trained. The recall, precision and F1 score was also quite similar to accuracy, ranging from 98.7 to 98.8 percent.

The line chart in Figure 16 depicts the training and validation accuracy and AUC for proposed model and the best transfer learning model from previous comparison - VGG16, which indicates that the proposed model suffers from less oscillation than the VGG16 model. Also, the train and validation accuracy are almost close to each other, so there might not be any overfitting or underfitting issue.

Since we have also created a separate augmented dataset from the Merged dataset 1, another bar chart is given in Figure

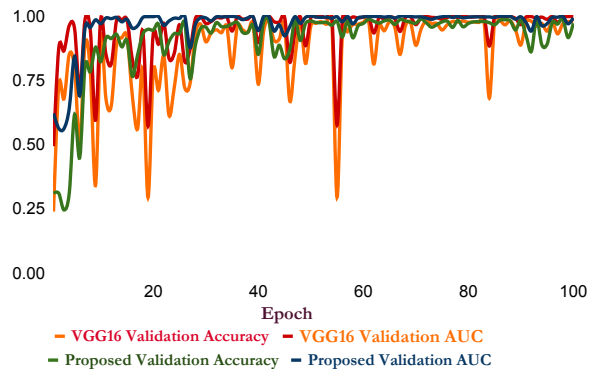


Figure 16 Graph comparing training and validation accuracy and AUC for the proposed classification model to VGG 16 for Merged dataset 1.

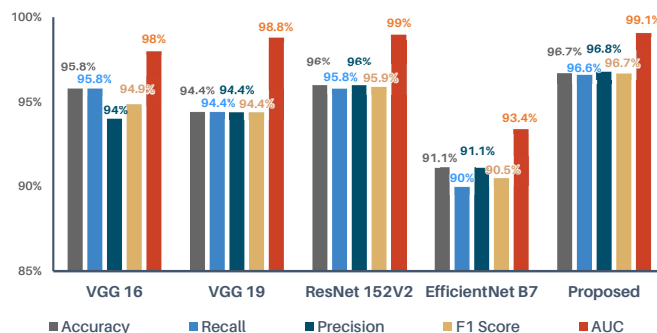
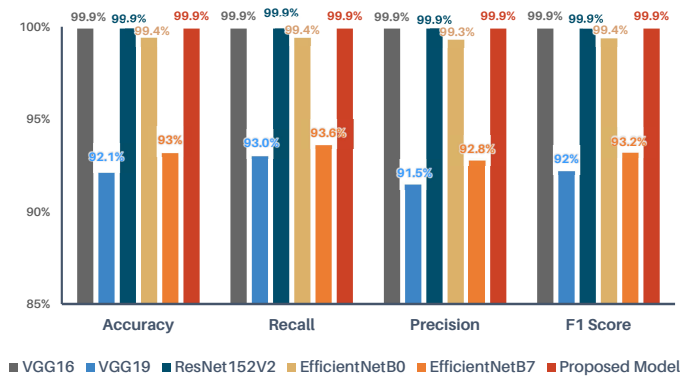


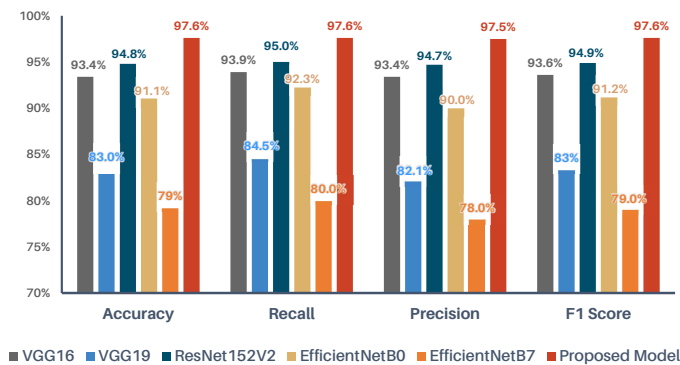
Figure 17 Comparison of proposed model with 5 transfer learning models on Augmented Merged dataset 1, where proposed models surpass

Figure 17 to showcase the validation performance comparisons, where our proposed model has higher performance than the other models in terms of results. The proposed model achieved validation accuracy and F1 score of 98.7 percent, and 99.1 percent AUC.

Again in Figure 18a and Figure 18b training and testing evaluation are given to create comparisons among models for Merged dataset 2. Likewise in the comparison of Merged dataset 1, the models achieve almost similar results in training, however VGG19 and EfficientNet B7 performs significantly worse than the proposed model which resulted in an accuracy of 99.9 percent. Although in validation performance, the proposed model outclasses the other models, with accuracy and F1 score of 97.6 percent, where all of the transfer learning models have these values less than 95 percent. Figure 19 represents another epoch-wise line chart comparing the VGG16 model performance with the proposed model. Similar to Figure 16, there were only minor fluctuations in the learning process of the proposed model. Although the proposed model has a lot less parameters than the transfer learning models, it performs better than in terms of both training and validation.



(a)



(b)

Figure 18 Metrics of performance for 5 transfer learning approaches against the suggested methodology (a) For the training Merged dataset 2 condition, where the metrics values are pretty close (b) For the validation Merged dataset 2 situation, where the proposed model outperforms.

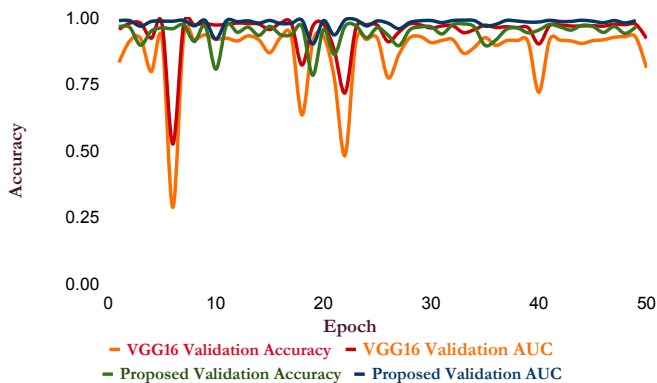


Figure 19 Graph comparing training and validation accuracy and AUC for the proposed classification model to VGG 16 for Merged dataset 2.

3.6. Segmentation result analysis

To separate the tumor portion from brain MRI images, a custom made U-Net was also trained on segmentation dataset. The Figure 20 illustrates proposed segmentation model output for different classes in the dataset with the ground truth image.

The input glsmri and predicted mask generated for Dataset d using the proposed segmentation model are shown in Figure 21. The graphic demonstrates how the model conveniently segmentates the tumor sections from the input MRI.

The segmentation model was also tested using a variety of statistical methods and image similarity measurements. Figure 22a and Figure 22b represents training and validation precision and cross entropy loss over 200 epochs for the segmentation model. Both loss values are quite low, near 0.1 and the precision are above 90 percent. There were almost no fluctuation in the training process.

Figure 23 represents the statistical and similarity measure for the validation data, the validation cross entropy loss is 6.6%, with 98.4 percent precision, and the recall is 71%.

From Table 6 the values of performance metrics are noted as the dice coefficient value is 0.89, the jaccard index is 0.81, dice loss and binary cross entropy dice losses are 0.10 and 0.79 respectively.

Table 6 Performance Metrics for the proposed segmentation model

Performance Metrics	Values
Dice Coefficient	0.89
Dice Loss	0.10
Binary cross entropy Dice loss	0.79
Intersection Over Union	0.81

3.7. Effect of segmentation

Another comparison is made between segmented and non segmented classification, to observe the effect. Since segmentation reduces the data size significantly, the over all time of classification models are reduced. Figure 24 illustrates the validation results of different classification models on segmented tumor image which is created from Merged dataset 1. The proposed model achieved 99 percent accuracy, 98.7 percent F1 score and 99.6 percent AUC for segmented tumor classification. Compared to proposed model, the best transfer learning model is Resnet152v2 with a validation accuracy of 96 percent, which is lower than the proposed model. In terms of the effect of segmentation, Figure 25 describes the difference between segmented and non-segmented tumor classification. The statistical metrics are quite similar, varies by only 0.1 percent utmost. The proposed models achieved 98.8 percent validation accuracy with the segmentation process whereas without cutting out the tumor portion from images, it achieved 98.7 percent validation accuracy. The time needed to train classification model with out segmentation was 21 min 66 seconds, compared to the with segmentation process, which need a total time of 58 minutes 33 seconds. Since in the second approach two model, one Unet for

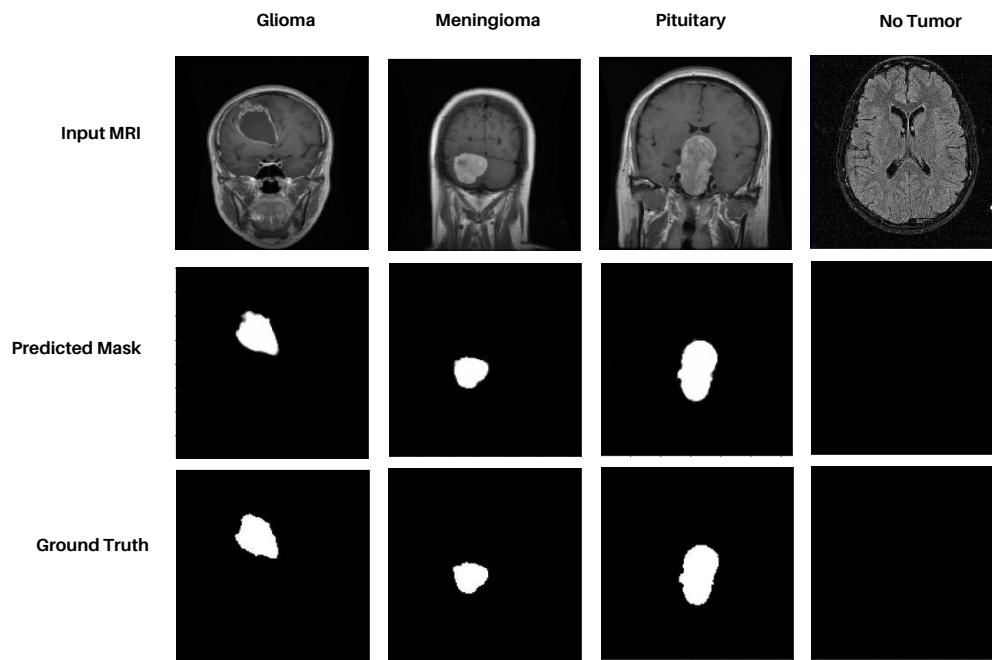


Figure 20 The proposed segmentation model’s projected mask and growth truth comparison, which is based on the U-net architecture and accepts input MRIs and outputs segmented mask image

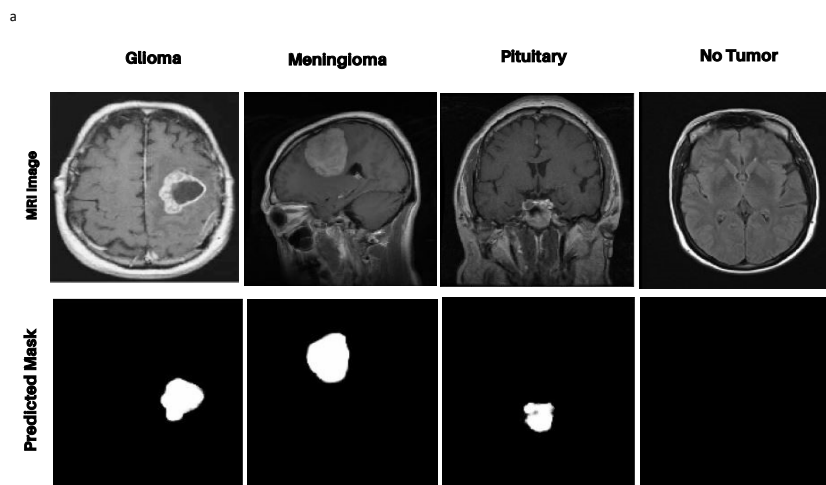


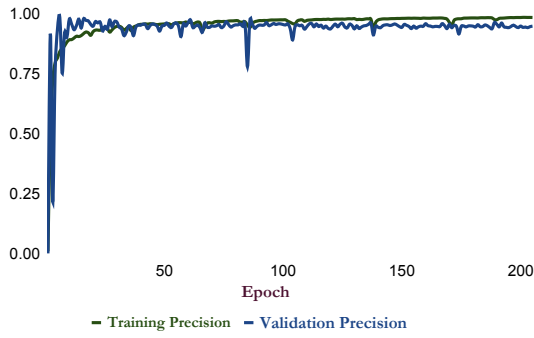
Figure 21 Predicted mask generated from input MRI using proposed segmentation model for classification Dataset d.

segmentation, and one classification model is needed to train, it had a higher training time. However in case of just classification after segmentation, the time reduces to only 13 minutes, which is 40 percent lower than the previous approach.

3.8. Comparison with State of the Art Papers

The proposed model is compared to several existing classification models in relation to dataset a. As a result, all of the

papers in Table 7 conducted MRI image multiclassification and classified the images as glioma, meningioma, pituitary, or no tumor. Ayadi et al. [19] also implemented a CNN approach and achieved an accuracy of 94.74%. Ghassemi et al. [26] proposed a GAN model where deep convolutional neural network is used as the discriminator for detection of fake images that are generated by the generative model and a pre-trained CNN network is fine-tuned to perform as classifier for brain tumor classifica-



(a)



(b)

Figure 22 The proposed segmentation model’s training and validation (a) precision vs. epoch (b) loss vs. epoch comparison, where the validation precision and loss both nearly match their respective training curves

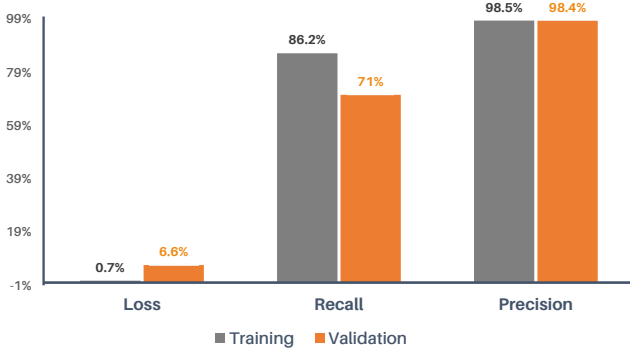


Figure 23 Comparison of training and validation performance measures (loss, recall, and accuracy) for the proposed segmentation model.

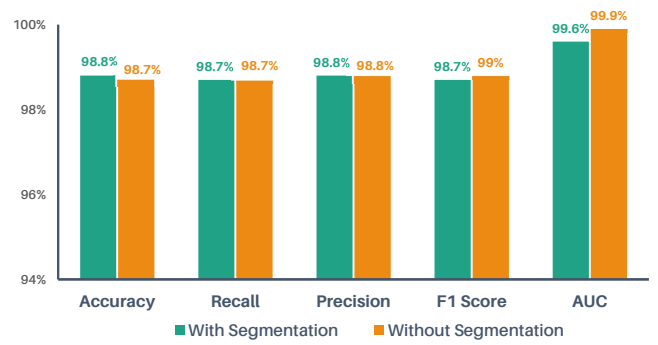


Figure 25 Proposed classification model validation performance metrics comparison for Merged dataset 1 when tumor MRIs are segmented using the proposed segmentation model and when they are not segmented.

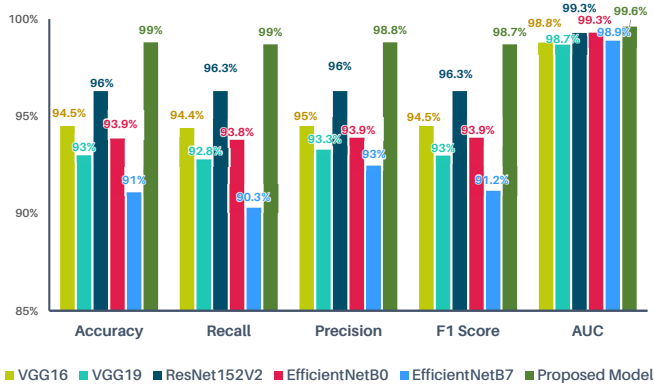


Figure 24 Validation performance comparison of proposed classification model with 5 transfer learning methods after segmenting tumor MRIs using the proposed segmentation model and transferring them to the classification model on Merged dataset 1.

and the model achieved an accuracy of 95.6%. Ismael et al. [53] combined the 2D DWT and 2D Gabor extraction meth-

ods for feature extraction, the features were fed to a traditional neural network and achieved an accuracy of 91.9%. Afshar et al. [54] employed Capsule Network on segmented tumor regions to address CNN’s shortcomings related to the loss of the active features at the specific location in the subsampling layers and poor training results regarding small datasets. Pashaei et al. [55] proposed a method that extracts brain tumor features using CNN and further classifies the obtained features into three classes of tumors such as meningioma, glioma and pituitary tumors using . With an accuracy of 96.7%, the proposed classification model beats the results of the previously described research using Dataset a.

4. Discussions

This research provides a comprehensive comparison of several tumor classification algorithms. First, the influence of data and datasets was evaluated across different accessible datasets

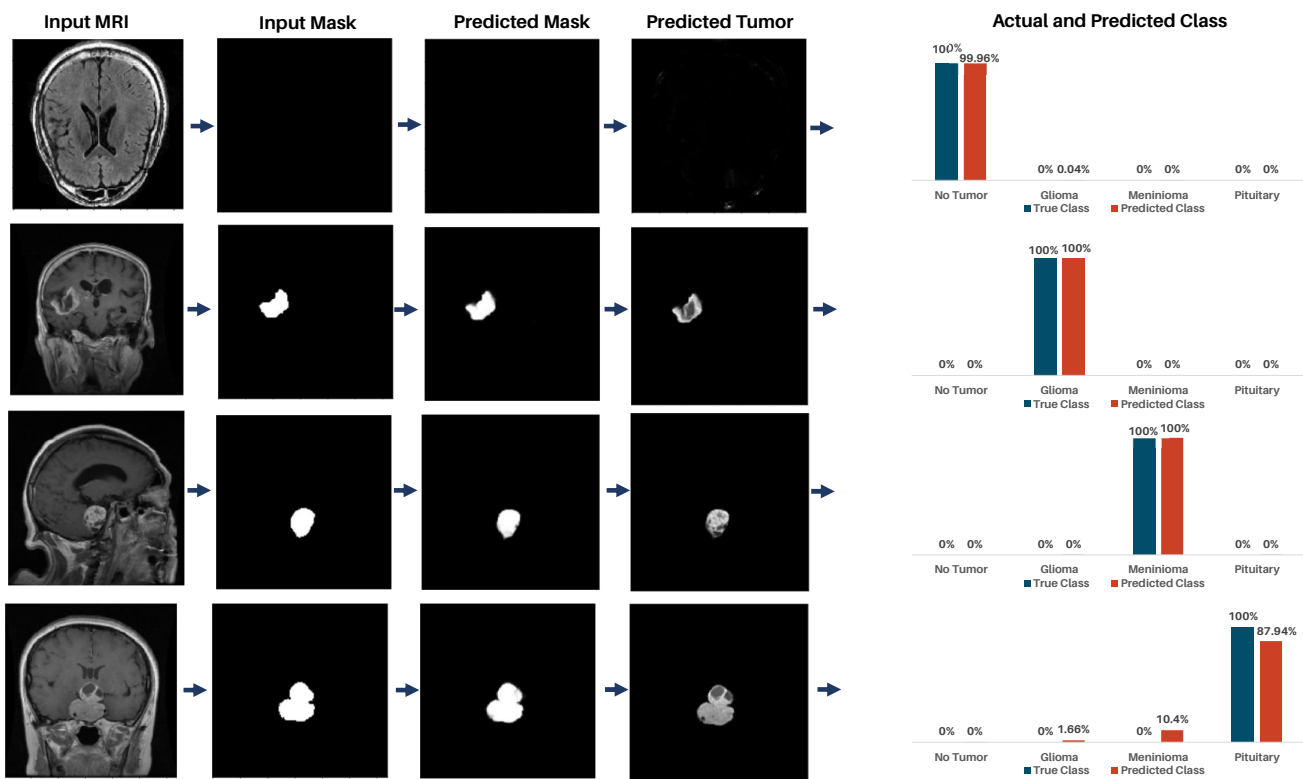


Figure 26 The whole classification process including the segmentation technique is shown, including the input MRI, input mask, generated mask, tumor area, actual class, and predicted class.

Table 7 Comparison among proposed model and State of The Art papers considering the Dataset a

Ref.	Year	Method	Accuracy (Percentage)
Ayadi et al. [19]	2021	CNN	94.74
Ghassemi et al. [26]	2020	GAN+CNN	95.6
Badža et al. [14]	2020	CNN	96.56
Nawab et al. [33]	2019	AlexNet, VGG16, VGG19	94.82
Sultan et al. [12]	2019	CNN	96.13
Pashaei et al. [55]	2018	CNN	93.68
Ismael et al. [53]	2018	Neural Network (NN)	91.9
Afshar et al. [54]	2018	Capsule Network (CapsNet)	86.56
Proposed Model		Deep CNN	96.7

to determine if it had any effect on classification. Because the size and output classes of this dataset vary, the findings

were fairly disparate. Nonetheless, the proposed strategy produced some promising outcomes. The suggested model was then trained and validated on two unique datasets developed by combining previously existing ones. There are three types of tumor MRI in this dataset, with normal MRIs. Attempts have been made to ascertain, if any, effect of augmentation in the process. Surprisingly, the effect of augmentation and even preprocessing was not discernible, which may be attributed to the proposed model's comprehensive feature extraction capacity. Then we tested our model to several transfer learning models to demonstrate that, even without a huge number of parameters, proposed model matches or outperforms the transfer learning models in most circumstances. The effect of segmentation was then demonstrated by training and validating both segmented and non-segmented MRI. The findings clearly illustrate that, whereas segmentation requires two steps of model training, classification obtained equivalent or even superior results in less time. So, if training time has no influence on the appli-

cation, segmentation can improve classification performance. However, if a minor variation in performance is not a problem, the suggested model may predict fairly well without any segmentation. Figure 26 depicts the segmentation process, as well as the actual classes, anticipated tumor classes, and actual MRI inputs. According to the Figure, the classification model with segmentation detects normal MRIs with 99.96% accuracy, glioma and meningioma classes with 100% accuracy, and pituitary tumors with 87.94% accuracy.

5. Conclusion

Because brain tumors may be exceedingly harmful, if not deadly, early identification can save people's lives. This study proposes an automated classification strategy for rapid, early, and accurate diagnosis in order to avert the disastrous consequences. A deep CNN model has been used to classify brain MRIs into four groups (glioma, meningioma, no tumor and pituitary). A segmentation model for automatic segmentation of brain MRIs from original input MRIs has been suggested. Effective automated tumor segmentation is often challenging due to the broad variety of tumor locales, shapes, and structures. For segmenting the tumor sections, we developed a U-Net architecture-based model. The classification techniques are not significantly different whether segmentation is used or not. However, classification without segmentation shortens the time required by the classification model. The impact of dataset augmentation has been studied. The classification model was evaluated across numerous datasets and with 5 pretrained models. Furthermore, the suggested model outperforms pretrained models and also when compared to state-of-the-art articles. The segmentation approach also generated a more accurate segmented mask, allowing every MRI picture from any dataset to be successfully segmented.

References

- [1] L. Arokia Jesu Prabhu, A. Jayachandran, Mixture Model Segmentation System for Parasagittal Meningioma brain Tumor Classification based on Hybrid Feature Vector, *Journal of Medical Systems* 42 (12) (2018) 251. doi:10.1007/s10916-018-1094-3. URL <https://doi.org/10.1007/s10916-018-1094-3>
- [2] Brain Tumor - Statistics, June 2012, <https://www.cancer.net/cancer-types/brain-tumor/statistics>, accessed: 27 February, 2022.
- [3] M. I. Razzak, M. Imran, G. Xu, Efficient brain tumor segmentation with multiscale two-pathway-group conventional neural networks, *IEEE journal of biomedical and health informatics* 23 (5) (2018) 1911–1919, publisher: IEEE.
- [4] E. Irmak, Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework, *Iranian Journal of Science and Technology, Transactions of Electrical Engineering* 45 (3) (2021) 1015–1036, publisher: Springer.
- [5] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, The multimodal brain tumor image segmentation benchmark (BRATS), *IEEE transactions on medical imaging* 34 (10) (2014) 1993–2024, publisher: IEEE.
- [6] J. Long, E. Shelhamer, T. Darrell, Fully Convolutional Networks for Semantic Segmentation 10.
- [7] M. A. Naser, M. J. Deen, Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images, *Computers in biology and medicine* 121 (2020) 103758, publisher: Elsevier.
- [8] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, A.-B. M. Salem, Classification using deep learning neural networks for brain tumors, *Future Computing and Informatics Journal* 3 (1) (2018) 68–71, publisher: Elsevier.
- [9] F. Özyurt, E. Sert, E. Avci, E. Dogantekin, Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy, *Measurement* 147 (2019) 106830, publisher: Elsevier.
- [10] S. Sajid, S. Hussain, A. Sarwar, Brain tumor detection and segmentation in MR images using deep learning, *Arabian Journal for Science and Engineering* 44 (11) (2019) 9249–9261, publisher: Springer.
- [11] E. Irmak, Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework, *Iranian Journal of Science and Technology, Transactions of Electrical Engineering* (2021) 1–22Publisher: Springer.
- [12] H. H. Sultan, N. M. Salem, W. Al-Atabany, Multi-classification of brain tumor images using deep neural network, *IEEE Access* 7 (2019) 69215–69225, publisher: IEEE.
- [13] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks, *Medical image analysis* 35 (2017) 18–31, publisher: Elsevier.
- [14] M. M. Badža, M. Barjaktarović, Classification of brain tumors from MRI images using a convolutional neural network, *Applied Sciences* 10 (6) (2020) 1999, publisher: Multidisciplinary Digital Publishing Institute.
- [15] J. Amin, M. Sharif, M. Yasmin, S. L. Fernandes, Big data analysis for brain tumor detection: Deep convolutional neural networks, *Future Generation Computer Systems* 87 (2018) 290–297, publisher: Elsevier.
- [16] H. Mzoughi, I. Njeh, A. Wali, M. B. Slima, A. BenHamida, C. Mhiri, K. B. Mahfoudhe, Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification, *Journal of Digital Imaging* 33 (2020) 903–915, publisher: Springer.
- [17] M. Siar, M. Teshnehlab, Brain tumor detection using deep neural network and machine learning algorithm, in: 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, 2019, pp. 363–368.
- [18] S. Irshaidat, R. Duwairi, Brain Tumor Detection Using Artificial Convolutional Neural Networks, in: 2020 11th International Conference on Information and Communication Systems (ICICS), 2020, pp. 197–203, ISSN: 2573-3346. doi:10.1109/ICICS49469.2020.239522.
- [19] W. Ayadi, W. Elhamzi, I. Charfi, M. Atri, Deep CNN for brain tumor classification, *Neural Processing Letters* 53 (1) (2021) 671–700, publisher: Springer.
- [20] M. A. Naser, M. J. Deen, Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images, *Computers in biology and medicine* 121 (2020) 103758, publisher: Elsevier.
- [21] I. Shahzadi, T. B. Tang, F. Meriadeau, A. Quyyum, CNN-LSTM: Cascaded framework for brain Tumour classification, in: 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), IEEE, 2018, pp. 633–637.
- [22] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, M. Shoaib, A deep learning model based on concatenation approach for the diagnosis of brain tumor, *IEEE Access* 8 (2020) 55135–55144, publisher: IEEE.
- [23] F. Özyurt, E. Sert, D. Avci, An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine, *Medical hypotheses* 134 (2020) 109433, publisher: Elsevier.
- [24] A. Çinar, M. Yildirim, Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture, *Medical hypotheses* 139 (2020) 109684, publisher: Elsevier.
- [25] N. Faruqui, M. Yousu, M. Whaiduzzaman, A. K. M. Azad, A. Barrosean, M. A. Moni, LungNet: A hybrid deep-CNN model for lung cancer diagnosis using CT and wearable sensor-based medical IoT data, *Computers in Biology and Medicine* 139 (2021) 104961. doi:10.1016/j.compbio.2021.104961.
- [26] N. Ghassemi, A. Shoeibi, M. Rouhani, Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images, *Biomedical Signal Processing and Control* 57 (2020) 101678, publisher: Elsevier.
- [27] T. Saba, A. S. Mohamed, M. El-Affendi, J. Amin, M. Sharif, Brain tumor detection using fusion of hand crafted and deep learning features, *Cognitive Systems Research* 59 (2020) 221–230, publisher: Elsevier.

- [28] K. Ahamed, M. Islam, M. A. Uddin, A. Akhter, B. K. Paul, M. Yousuf, S. Uddin, J. Quinn, M. A. Moni, A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images, *Computers in Biology and Medicine* 139 (2021) 105014. doi:10.1016/j.combiomed.2021.105014.
- [29] R. Saouli, M. Akil, R. Kachouri, Fully automatic brain tumor segmentation using end-to-end incremental deep neural networks in MRI images, *Computer methods and programs in biomedicine* 166 (2018) 39–49, publisher: Elsevier.
- [30] N. F. Aurna, M. A. Yousuf, K. A. Taher, A. K. M. Azad, M. A. Moni, A classification of MRI brain tumor based on two stage feature level ensemble of deep CNN models, *Computers in Biology and Medicine* 146 (2022) 105539. doi:10.1016/j.combiomed.2022.105539. URL <https://www.sciencedirect.com/science/article/pii/S0010482522003316>
- [31] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, USA, 2015, pp. 1–9. doi:10.1109/CVPR.2015.7298594. URL <http://ieeexplore.ieee.org/document/7298594/>
- [32] R. Chelghoum, A. Ikhlef, A. Hameurlaine, S. Jacquir, Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images, in: *IFIP International Conference on Artificial Intelligence Applications and Innovations*, Springer, 2020, pp. 189–200.
- [33] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J. Lu, Brain tumor classification for MR images using transfer learning and fine-tuning, *Computerized Medical Imaging and Graphics* 75 (2019) 34–46, publisher: Elsevier.
- [34] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, S. W. Baik, Multi-grade brain tumor classification using deep CNN with extensive data augmentation, *Journal of computational science* 30 (2019) 174–182, publisher: Elsevier.
- [35] A. Elazab, C. Wang, S. J. S. Gardezi, H. Bai, Q. Hu, T. Wang, C. Chang, B. Lei, GP-GAN: Brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR Images, *Neural Networks* 132 (2020) 321–332, publisher: Elsevier.
- [36] M. Hamghalam, T. Wang, B. Lei, High tissue contrast image synthesis via multistage attention-GAN: application to segmenting brain MR scans, *Neural Networks* 132 (2020) 43–52, publisher: Elsevier.
- [37] M. Rezaei, K. Harmuth, W. Gierke, T. Kellermeier, M. Fischer, H. Yang, C. Meinel, A conditional adversarial network for semantic segmentation of brain tumor, in: *International MICCAI Brainlesion Workshop*, Springer, 2017, pp. 241–252.
- [38] C. Han, K. Murao, T. Noguchi, Y. Kawata, F. Uchiyama, L. Rundo, H. Nakayama, S. Satoh, Learning more with less: Conditional PGGAN-based data augmentation for brain metastases detection using high-rough annotation on MR images, in: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019, pp. 119–127.
- [39] P. Mohamed Shakeel, T. E. E. Tobely, H. Al-Feel, G. Manogaran, S. Baskar, Neural Network Based Brain Tumor Detection Using Wireless Infrared Imaging Sensor, *IEEE Access* 7 (2019) 5577–5588, conference Name: IEEE Access. doi:10.1109/ACCESS.2018.2883957.
- [40] M. I. Sharif, J. P. Li, M. A. Khan, M. A. Saleem, Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images, *Pattern Recognition Letters* 129 (2020) 181–189, publisher: Elsevier.
- [41] Jun Cheng, Brain Tumor Dataset, https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/5, 2017, accessed: 9 April, 2021.
- [42] Sartaj Bhuvaji, Brain Tumor Classification (MRI), <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>, accessed: 12 April, 2021.
- [43] Pradeep, Brain MRI, <https://www.kaggle.com/datasets/pradeep2665/brain-mri>, accessed: 12 April, 2021.
- [44] MohamedMetwaySherif, Brain Tumor Dataset, <https://www.kaggle.com/datasets/mohamedmetwalysherif/braintumordataset>, 2020, accessed: 14 April, 2021.
- [45] Br35H :: Brain Tumor Detection, <https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection>, 2020, accessed: 14 April, 2021.
- [46] O. R. Vincent, O. Folorunso, A descriptive algorithm for sobel image edge detection, in: *Proceedings of informing science & IT education conference (InSITE)*, Vol. 40, 2009, pp. 97–107.
- [47] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv:1409.1556 [cs] (Apr. 2015). URL <http://arxiv.org/abs/1409.1556>
- [48] M. Tan, Q. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, in: *Proceedings of the 36th International Conference on Machine Learning*, PMLR, 2019, pp. 6105–6114, iSSN: 2640-3498. URL <https://proceedings.mlr.press/v97/tan19a.html>
- [49] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, USA, 2016, pp. 770–778. doi:10.1109/CVPR.2016.90. URL <http://ieeexplore.ieee.org/document/7780459/>
- [50] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, arXiv:1412.6980 [cs] (Jan. 2017). URL <http://arxiv.org/abs/1412.6980>
- [51] A. F. Agarap, Deep Learning using Rectified Linear Units (ReLU), arXiv:1803.08375 [cs, stat] (Feb. 2019). URL <http://arxiv.org/abs/1803.08375>
- [52] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.
- [53] M. R. Ismael, I. Abdel-Qader, Brain tumor classification via statistical features and back-propagation neural network, in: 2018 IEEE international conference on electro/information technology (EIT), IEEE, 2018, pp. 0252–0257.
- [54] P. Afshar, A. Mohammadi, K. N. Plataniotis, Brain tumor type classification via capsule networks, in: 2018 25th IEEE international conference on image processing (ICIP), IEEE, 2018, pp. 3129–3133.
- [55] A. Pashaei, H. Sajedi, N. Jazayeri, Brain tumor classification via convolutional neural network and extreme learning machines, in: 2018 8th International conference on computer and knowledge engineering (ICCKE), IEEE, 2018, pp. 314–319.