Research Article

A Systematic Approach for MRI Brain Tumor Localization and Segmentation Using Deep Learning and Active Contouring

Shanaka Ramesh Gunasekara ⬤, H. N. T. K. Kaldera ⬤, and Maheshi B. Dissanayake ⬤

Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Kandy 20400, Sri Lanka

Correspondence should be addressed to Shanaka Ramesh Gunasekara; shanakag@eng.pdn.ac.lk
and Maheshi B. Dissanayake; maheshid@eng.pdn.ac.lk

One of the main requirements of tumor extraction is the annotation and segmentation of tumor boundaries correctly. For this purpose, we present a threefold deep learning architecture. First, classifiers are implemented with a deep convolutional neural network (CNN) and second a region-based convolutional neural network (R-CNN) is performed on the classified images to localize the tumor regions of interest. As the third and final stage, the concentrated tumor boundary is contoured for the segmentation process by using the Chan–Vese segmentation algorithm. As the typical edge detection algorithms based on gradients of pixel intensity tend to fail in the medical image segmentation process, an active contour algorithm defined with the level set function is proposed. Specifically, the Chan–Vese algorithm was applied to detect the tumor boundaries for the segmentation process. To evaluate the performance of the overall system, Dice Score, Rand Index (RI), Variation of Information (VOI), Global Consistency Error (GCE), Boundary Displacement Error (BDE), Mean Absolute Error (MAE), and Peak Signal to Noise Ratio (PSNR) were calculated by comparing the segmented boundary area which is the final output of the proposed, against the demarcations of the subject specialists which is the gold standard. Overall performance of the proposed architecture for both glioma and meningioma segmentation is with an average Dice Score of 0.92 (also, with RI of 0.9936, VOI of 0.0301, GCE of 0.004, BDE of 2.099, PSNR of 77.076, and MAE of 52.946), pointing to the high reliability of the proposed architecture.

1. Introduction

Medical image classification and segmentation is a field, where deep learning can make a huge impact with promising results. It facilitates the automation of noninvasive imaging-based diagnosis. Interestingly, computer-aided brain tumor diagnosis has effectively utilized the advances in medical image processing in the past and has opened up many promising research activities in the domain of deep learning, with the expectation of developing entirely computerized automatic accurate diagnostic systems for physicians.

A brain tumor is a mass or growth of abnormal cells in the brain which might be cancerous (malignant) or noncancerous (benign). The early, comprehensive diagnosis and proper treatments are essential for a patient's survival in brain tumor management. During the past decades, more than 120 types of brain tumors were discovered by medical scientists. These brain tumors can be broadly categorized into two main groups, namely, primary brain tumors, which originate in the brain itself and secondary deposits in the brain, where the primary tumor is elsewhere in the body [1]. Typically, noninvasive medical imaging techniques such as Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) are favoured as brain tumor identification tools at the initial stages, over incursion invasive procedures like tissue biopsies [2, 3]. The authors in [4] found that CT, MRI, and Positron Emission Tomography (PET) usage has increased by 7.8%, 10%, and 57%, respectively, during the period of 1996–2010. Furthermore, as of healthcare resource statistics of the EU for 2020 [5], the EU Member States have shown a widespread increase in the availability of medical imaging technology and equipment for diagnosis in the

recent decades. Moreover, according to [6], the overall employment of radiologic and MRI technologists grows faster than the average for all occupations in the USA. All these findings confirm that medical image-based diagnosis is favoured in the modern healthcare system.

Image classification and segmentation have shown rapid growth during the past two decades with the introduction of machine learning and computer vision techniques. Deep learning has found applications in medical imaging as in identifying local anatomical characters, detecting organs and body parts, and identifying cells of different shapes and sizes [2]. A review of the related works [4] shows that a considerable portion of the latest research on image analysis uses a deep convolutional neural network (DCNN) for both image classification and segmentation [7]. Multimodal Brain Tumor Segmentation (BRATS) challenge is the main competition on brain tumor classification which is organized by the Perelman School of Medicine at the University of Pennsylvania, Centre for Biomedical Image Computing and Analytics (CBICA) from 2012 onwards. The BRATS challenge focuses on automating brain tumour detection and the survival rate estimation techniques and algorithms. Each year, the dataset is updated and the overall performance of the proposed algorithms has shown a tremendous improvement over time. On the whole, the accuracy of the algorithms proposed using the BRATS dataset falls around 90% [8–10]. Some of these algorithms were developed using classical CNN architecture whereas some are developed using improved CNN algorithms like U-net [11] and superpixel-based extremely randomized trees [12].

Another popular and publicly available brain tumour dataset is the Figshare MRI dataset [13, 14] which is the dataset employed in this paper. Due to the easy accessibility and the ready availability, the Figshare MRI brain tumour dataset also has been used in many brain tumor classification and segmentation related research [15–18]. The dataset, which was initiated in 2015 and last updated in 2017 [13, 16], carries an average classification accuracy in the range of 90–95% [14, 16, 19, 20]. The authors in [16] achieved a classification average of 95% accuracy by using a modified CNN architecture while the authors in [15] achieved around 96% accuracy with an automatic content-based image retrieval (CBIR) system. A deep network was enhanced by employing cross channel normalization (CCN) and parametric rectified linear unit (PRELU) in [18] for brain tumor segmentation.

Although the convolutional neural network (CNN) has achieved a considerable performance gain in the medical image classification [14] and segmentation tasks, it is associated with a significant increase in the computational cost, especially when high-resolution images are analysed. In general, an object detection algorithm draws a bounding box around the object of interest in the image. In a classical automated system, this can be achieved by using a typical convolutional network, followed by a fully connected layer. However, when the number of objects required to be detected is not a fixed number, it is difficult to proceed with the above approach as it requires defining the length of the fully connected layer at the initial design stage.

In the previous work by the authors, a region proposal algorithm is proposed to address the problem of selecting a random number of objects in a single region [21, 22]. In the proposed method, instead of searching the entire image for the number of objects, the algorithm searches for objects in several selective areas of the image, while treating each subregion as an independent subimage. In [21], a fully autonomous learning algorithm was constructed using Region-based Faster Convolutional Neural Network (Faster R-CNN) to localize the meningioma tumor regions in MRI. Once the tumor is localized, Prewitt and Sobel edge detection algorithms are applied to the localization output, with the expectation of detecting the exact tumor boundary. Both of these techniques compute an approximate tumor boundary using the gradient intensity function of the image [23]. As MRIs on the whole consist of Rician noise and edges are not defined only by gradient, these algorithms underperform in this task. In practice, the effectiveness of the developed deep learning models to make informed decisions are evaluated through accuracy and system loss. Going beyond simple accuracy, standard mathematical objective parameters such as precision, recall, and Dice Score are utilized to choose the best model for the given problem. Furthermore, graphical representations such as confusion matrix and receiver operation characteristic (ROC) too are utilized to evaluate the performance of the deep learning model. A confusion matrix is a two-dimensional matrix which summarises the performance of the classification algorithm. One dimension of the matrix represents the true classes of an object while the other represents the class that the classifier predicts [24, 25]. The ROC curve is also a two-dimensional plot which illustrates how well a classifier system works as a discrimination cut-off value is changed over the range of predictor variable [25].

In the research presented, we outline an automated systematic approach to classify, segment, and extract the exact tumour boundary from MRI images. The key contribution of this research can be summarised as follows:

(1) We present a simplified CNN architecture based on a small number of layers and faster R-CNN, for the classification of axial MRI into glioma and meningioma brain tumors and produce a bounding box of the tumor with a 94% of accuracy confidence level [21, 22]. One of the key challenges in medical image analysis is the scarcity of the labelled data. Hence, in this research, we specifically focus on applying R-CNN based tumor localization for a scenario with a lower amount of annotated data. Specifically, we have used a dataset with less than 500 axial MRI images for our research. Furthermore, we kept our network simple, to reduce the total number of trainable parameters.

(2) The exaction of the exact tumor boundary was performed by means of an unsupervised active contour detection method, developed by means of Chan and Vese [26] algorithm. One of the inherent drawbacks of using active contour algorithms for segmentation is the requirement of the initial search

area. Without a correct initial search area, the algorithm is at the risk of segmenting unwanted regions in the brain MRI, such as the orbital cavity and lateral ventricles. In this research, we used tumor bounding box coordinates obtained at faster R-CNN based localization step as the initial search area for the Chan–Vese algorithm to obtain a refined segmentation of the tumor. This in turn reduces the search area of the Chan–Vese algorithm and helps the algorithm to converge to an accurate boundary in a shorter time frame. The resultant segmented boundary carries a high accuracy according to objective quality evaluation measures, compared to the state of the art.

In this article, we were able to present an end to end complete systematic approach for both meningioma and glioma tumor detection and segmentation using MRI. The complete system comprises 3 main subsystems: brain tumor classification using a simple CNN algorithm, Faster R-CNN based network for tumor localization, and finally the Chan–Vese algorithm for exact tumor segmentation. All three algorithms were connected in a cascade manner, with the final deliverable as the exact tumor boundary for segmentation purpose for any given axial brain MRI. The rest of the paper is organized as follows: In Section 2, we present a summarised overview of the proposed framework. Section 3 briefly outlines the theoretical background of the research. The methodology adopted for the dataset preparation, classification, segmentation, and contouring is presented in Section 4. Section 5 presents the performance analysis, using objective matrices, whereas Section 6 discusses and compares our proposed architecture against the existing works in the literature. Finally, Section 7 concludes the paper.

2. Architecture of the Proposed Algorithm

In this paper, we propose a threefold complete architecture to classify and segment brain tumors using a T1 weighted MRI sequence. The proposed system architecture consists of three cascaded algorithms, namely, in the order of application, convolutional neural network for classification, Faster R-CNN for tumor localization, and Chan–Vese algorithm for precise tumor segmentation. The flow diagram of the complete architecture is illustrated in Figure 1. More details of the proposed system can be found in Appendix.

Initially, a downsampled MR image is fed into a typical CNN to be classified as Meningioma or Glioma. Then, at the second stage, the classified image goes through a faster RCNN for tumor localization. The faster RCNN model adopted uses the pretrained weights generated using COCO net data for its feature extracting CNN, which is followed by a Region Proposal Network (RPN) and a classifier. The output of this second step is a boundary box around the tumor in $128*128$ downsampled image. As the third step, these boundary box coordinates are mapped to $512*512$ image with the original resolution, and the Chan–Vese algorithm is applied only for the boundary box area. This approach assures that the Chan–Vese algorithm converges
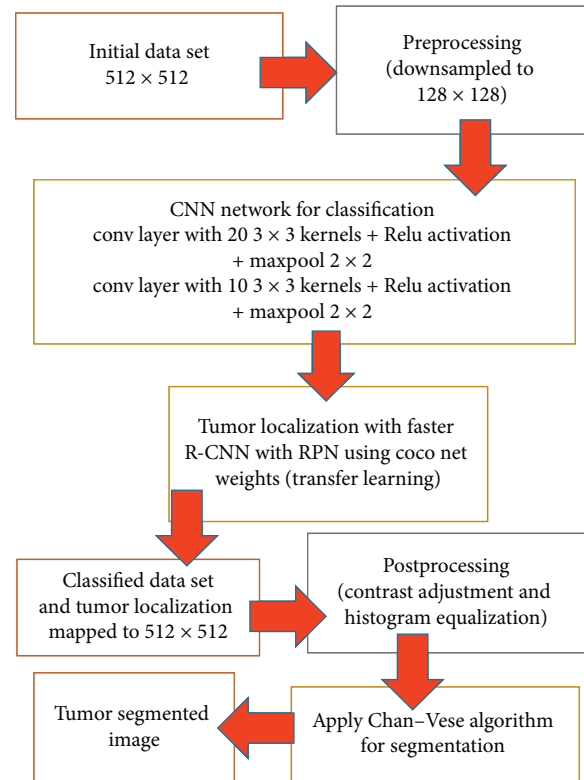

Figure 1: Proposed algorithm.

to an accurate boundary. Hence, we were able to obtain a much precious boundary for tumor segmentation within a low computational time. The outputs obtained at steps 2 and 3 are presented in Figure 2.

3. Background Works

3.1. Basic Operation of CNN and R-CNN. CNN is a class of layered deep neural network architecture built using convolution, activation, pooling, and fully connected layers to analyse visual imagery. The convolutional layer uses a set of learnable filters with different sizes to extract various feature maps to learn the correlation between neighbouring pixels, while drastically reducing the number of weighted parameters. The pooling layer introduces nonlinear downsampling to the system architecture while the activation increases the nonlinear properties of the decision function of the overall network independent of the convolution layer. Followed by several combinations of convolution, pooling, and activation, CNN has the fully connected layers, where high-level decision-making takes place. At the final stage of the design, the dense layer or loss layer maps the trained outcome with the predefined output class. In a fully connected CNN architecture, these operations are executed forward and backward, through forward learning and backpropagation as a designed architecture fine-tune, that is, training cycle, to optimize the decision-making capacity of the CNN architecture.

The R-CNN is an object detection and localization mechanism evolved from CNN architecture. It is a region-based segmentation method which follows segmentation

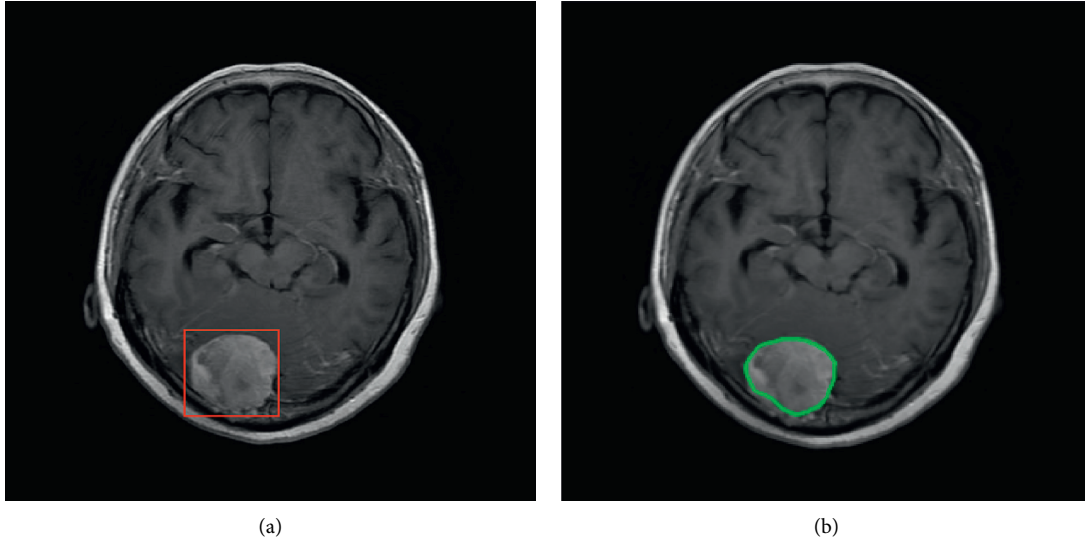(a)   (b)

FIGURE 2: (a) Brain tumor localization bounding box obtained at step 2 (after Faster RCNN). (b) Tumor segmentation output obtained at step 3 (after Chan–Vese).

using a recognition approach. It first extracts the free-form regions of interest from the input image and then conducts region-based classification on the extracted region of interest (ROI). The faster R-CNN consists of two main subnetworks, R-CNN and RPN [27]. RPN itself narrows down the number of search regions in the image by generating anchors as in Figure 3 and works as a classifier that trains CNNs to classify these selected ROIs, called hereafter "region proposals," into object categories. At first, R-CNN takes an input image and segments it into many subimages called regions with different dimensions. Next, each region is treated as an isolated image, and this isolated image is classified into a predefined set of object labels. Finally, a greedy algorithm is used to recursively combine subimages with similar regions to generate the region proposals with the predicted object labels.

R-CNN uses selective search algorithms to select these ROIs that lead to a huge computational cost and slow response rate as it initially generates over 2000 regions for each input image. Hence, RPN based bounding box detection algorithm was introduced into Faster R-CNN as the cost of generating region proposals is quite smaller in RPN compared to the selective search algorithm [27]. The significant difference between the two techniques is that the R-CNN uses the region proposals at the pixel level as input, whereas Faster R-CNN uses the region proposals at feature map level as its input. In general, RPN generates 9 anchors using the input image as in Figure 4 and predicts the probability of an anchor being in the background or foreground. Based on two significant factors, positive or negative labels were assigned to these anchors. It is observed that anchors which have higher intersection-over-union (IOU), correspond with the ground truth box. Hence, if an anchor and ground truth's IOU overlap is over 0.7, the anchor target gets a positive label, and if it is less than 0.3, the area is given a negative label [27]. The anchors where IOU lies between 0.3 and 0.7 $(0.3 < \text{IOU} < 0.7)$ are not followed through for

learning. The training phase of the RPN network is based on the loss function in (1), which is defined using the values assigned to the anchors:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*),$$

(1)

where i is the index of an anchor in a minibatch, p_i is the predicted probability of an anchor i being an object, t_i is the vector representing the 4 parameterized coordinates of the predicted bounding box, L_{cls} is the classification loss (log loss over two classes), L_{reg} is the regression loss, N_{cls} is the minibatch size, and N_{reg} is the number of the anchor locations.

It should be noted that when defining the loss function of the RPN for training purposes, a binary class label was assigned to each anchor. If the desired object is inside the anchor, the algorithm assigns 1 for p_i^* to indicate a positive anchor, whereas 0 is assigned to indicate a negative anchor. t_i^* is that of the ground truth box associated with a positive anchor. The regression loss, L_{reg}, of the loss function, adopted in faster R-CNN is defined as

$$L_{\text{reg}}(t_i, t_i^*) = R(t_i - t_i^*),$$

(2)

where R represents robust loss function. It should be noted that the regression loss is activated only for positive anchors $(p_i^* = 1)$ and is disabled otherwise $(p_i^* = 0)$, due to $p_i^* L_{\text{reg}}$ term in (1).

A trained RPN generates different sizes of feature maps as its output. As it is not easy to work with different sizes of feature maps, ROI pooling splits the input feature map into equal size regions and applies max pooling to every region. It is worth noting that, the output of the ROI pooling is always independent of the input size.
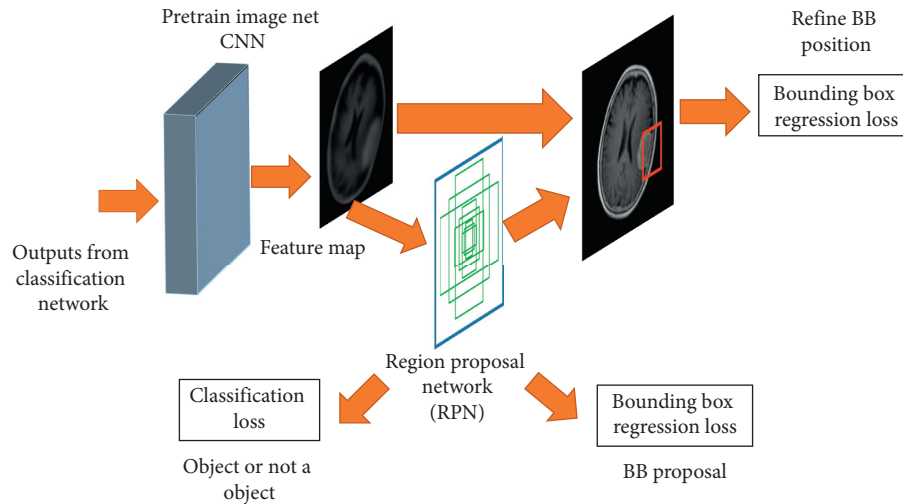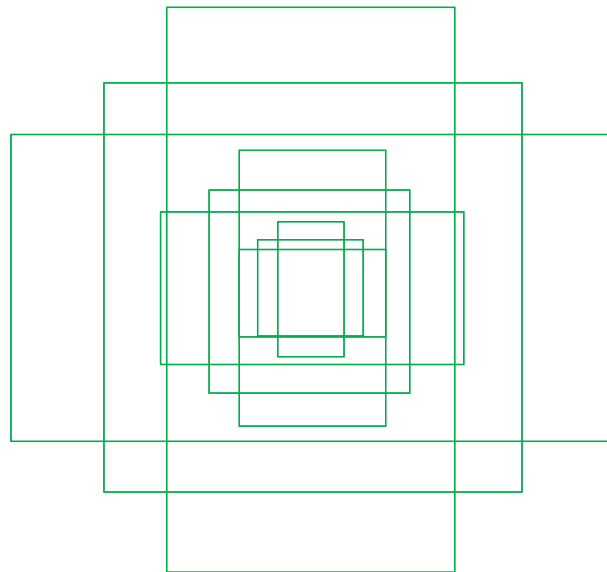
FIGURE 3: Region Proposal Network.

FIGURE 4: 9 anchors used by RPN at (320, 320).

3.2. Chan–Vese Segmentation. In image processing, many edge detection techniques based on the gradient of intensity, such as Sobel, Prewitt, and Roberts, are used for object boundary detection and segmentation [23]. In MRI, it is challenging to get accurate boundary detection using these operators as MRI itself contains Rician noise, which causes irregularities in the edge estimation. As both Rician noise and the edges of the image contain high-frequency components, it is challenging to get accurate results with these edge detection operators which depend on the gradient of the intensities. In [28], although Gaussian smoothing filters were used to reduce the Rician noise, it resulted in blurred and distorted MRI, which may lead to erroneous diagnosis. Therefore it is worthwhile to exploit the edge detection algorithms which are not defined based on the gradient of the intensity for MRIs.

As an alternative to the edge-based segmentation, active contour or Snake models [29] were developed for image segmentation which is governed by the gradient variation of the pixels. It starts with an initial estimation of the boundary curve plotted around the object of interest. With iterations, the estimated boundary moves towards its interior and stops on the true boundary of the object based on an energy minimizing model [30]. Even though active contouring performs better than the operators that depend on the gradient of intensity, the final output is sensitive to the initial condition of the algorithm. Hence, it is critical to set the correct boundary box at the beginning of the algorithm. Furthermore, the difficulties associated with topological changes, like merging and splitting of the evaluation curve, contribute to the active contour algorithm being a less popular choice for the segmentation problem.

Although there exist many algorithms with improved segmentation methodologies based on active contour algorithm, the level set method performs better with the noisy images [31, 32]. A level set method is a powerful tool for

contour evaluation which easily reins the topological changes like merging and splitting, which is difficult to tackle with classical active contour models [30]. In literature, the level set function is denoted as $\varnothing(i, j, t)$, where i, j are coordinates of the image and t is the artificial time. Segmentation is defined as two regions where $\{\varnothing < 0\}$ belongs to region 1 and $\{\varnothing \geq 0\}$ belongs to region 2. The output of the level set function, edge contour, is defined in the literature using $\{i, j; \varnothing(i, j, t) = 0\}$. This approach is used in Chan–Vese algorithm to form the level set function [30].

The Chan–Vese model, which is based on the Mumford–Shah functional for segmentation [14] is used to detect objects whose boundaries are not necessarily detected by the gradient [30, 33]. Mumford–Shah model is based on energy function and it finds a pair of (u, C) for a given input image u_0. Here, C denotes a smooth and closed segmentation curve of the object and u represents a nearly piecewise smooth approximation of the given image. The Chan–Vese algorithm presents an alternative solution to the Mumford–Shah model. It refers to the fitting energy function and achieves minimization through solving the following equation:

$$E^{CV}(c1, c2, C) = \mu \cdot \text{Length}(C) + \lambda_1$$
$$\cdot \int_{\text{inside}(C)} |u_0(x, y) - c_1|^2 \, dx \, dy \qquad (3)$$
$$+ \lambda_2 \cdot \int_{\text{outside}(C)} |u_0(x, y) - c_2|^2 \, dx,$$

where μ, λ_1, and λ_2 are positive constants. Typically, $\lambda_1 = \lambda_2 = 1, c_1$ and c_2 are intensity average of the given image μ_0 inside C and outside C, respectively. This minimization problem can be solved by replacing the curve C with the level set function $\phi(x, y)$. If $\phi(x, y) > 0$, the point (x, y) is inside the C, and if $\phi(x, y) < 0$, the point (x, y) is outside C. In addition, (x, y) is on the curve if $\phi(x, y) = 0$. For further details on the Chan–Vese model, the reader is directed to [30, 33–35]. The advantage of this method is that even if the image is very noisy and initial conditions are not well defined, still the locations of the boundaries are accurately estimated by the model.

4. Materials and Methods

4.1. Preparation of the Dataset. T1-weighed MRI brain tumor dataset presented at [13, 36] is used in this research. It is a collection of MRI data from Nanfang Hospital, Guangzhou, China, and General Hospital, Tianjing Medical University, China from 2005 to 2010. It was first published online in 2015, and the last updated version was released in 2017. It has been extensively used in MRI tumor analysis research recently [16, 18, 37]. We have employed the updated version (in 2017) of the dataset for training, testing, and validating of the proposed system. This dataset presents MRIs of coronal, sagittal, and axial plan, of 233 patients with 3 types of brain tumors, namely, meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices). The total images in the dataset are 3064 MRIs. Since this research only addressed the classification between meningioma and glioma, the pituitary tumor images were discarded at the

preprocessing stage. Then, the authors use only the axial MRI of meningioma (119 slices) and glioma (138 slices) tumors for the segmentation task. Hence, this research evaluates the performance of faster R-CNN under a smaller annotated dataset environment.

4.2. Preprocessing and the Parameter Setting. At the preprocessing stage, the input images were resized into 128×128 as it was not feasible to train the neural network with the original size of 512^*512. Yet, some finer features of the input, present at original resolution, could be lost during the downsampling process, reducing the sensitivity of the output by a small fraction. Nevertheless, we have downsampled the input image to 128×128 to reduce the complexity of the network with the objective of reducing the training of the network. Initially, the dataset is randomly split into 3 sets as training, testing, and validation with the ratio of $0.70 : 0.15 : 0.15$, and 5-fold cross-validation is applied to the training set with the scikit-learn library of python. Batch normalization was applied to input image to rearrange the input intensities to the scale 0-1.

4.3. Applying Chan–Vese Segmentation. After extracting the bounding box through the proposed Faster R-CNN based model, a segmentation algorithm is applied to obtain a precise tumor boundary definition. As the first step of the segmentation process, the bounding box obtained at faster R-CNN from the downsampled 128^*128 image was mapped to the original image size 512^*512. Then, general image preprocessing techniques such as contrast adjustment and histogram equalization were applied to adjust the contrast levels, brightness level, and sharpness of the images, to reduce the noise levels while enhancing the details of the image. After applying the preprocessing techniques, the Morphological Active Contours (Morphological Chan–Vese) technique was applied to identify the precise tumor region. In this research, the Chan–Vese algorithm uses a square-level set function, instead of general circle-level set functions, because the input boundary boxes are defined as squares at the faster R-CNN. Furthermore, parameters λ_1 and λ_2 in (3) were set to 1 following [30]. During the simulations, 100 iterations were followed to obtain the best convergence of the contour lines, and at each iteration, smoothing operators are applied 8 times. The values for this number of iterations and smoothing operators were selected by adopting the trial and error method. After the segmentation, the output of the Chan–Vese algorithm is compared against segmentation output of Prewitt edge detection technique and ground truth demarcations provided by neurologists, to justify the performance of our proposed system.

4.4. Statistical Performance Analysis. For the performance evaluation of the overall cascaded MRI tumor segmentation system proposed, the ground truth demarcations provided by experts in the field (neurologists) were compared against the mask obtained from the prediction process of the designed system. Dice Score, Rand Index (RI), Variation of

Information (VOI), Global Consistency Error (GCE), Boundary Displacement Error (BDE), Peak Signal to Noise Ratio (PSNR), and Mean Absolute Error (MAE) were calculated as the objective performance measure parameters [38].

Dice Score (F1 score) is a statistical parameter used to evaluate the similarity of two samples. Dice Score lies between 0 and 1, with 1 signifying the highest similarity between predicted and truth. F1 score is calculated using

$$\text{Dice score}\,(F1\,\text{score}) = \frac{2TP}{2TP + FP + FN}, \quad (4)$$

where TP is true positive, FP is false positive, FN is false negative, and TN is true negatives.

Rand Index (RI) counts the fraction of pairs of pixels whose labelling is consistent between the computed segmentation and the ground truth image. RI lies between 0 and 1 and if the two images are identical, the RI should be equal to 1. The RI value is calculated using

$$\text{RI} = \frac{TP + TN}{TP + FP + FN + TN}. \quad (5)$$

VOI computes the measure of information content in each of the segmentations and how much information one segmentation gives about the other; that is, it measures the information distance between the two segmentation. VOI is defined using the entropy and mutual information as

$$\text{VOI}\left(S_g, S_t\right) = H\left(S_g\right) + H\left(S_t\right) - 2MI\left(S_g, S_t\right), \quad (6)$$

where S_g and S_t are the fuzzy segmentations of the image, $H(S)$ is the marginal entropy, $H(S_g, S_t)$ is the joint entropy between two images, and $MI(S_g, S_t)$ is the mutual information between two images [39].

The GCE measures the extent to which a particular segmentation can be viewed as a refinement of the other. Segmentations that are related are considered to be consistent since they could represent the same image segmented at a different scale. The mathematical expression for GCE can be written as

$$\text{GCE}\,(S_1, S_2) = \frac{1}{n}\min\left(\sum_i E\,(S_1, S_2, p_i),\ \sum_i E\,(S_2, S_1, p_i)\right), \quad (7)$$

where S_1 and S_2 are two segmentations and p_i is a pixel position.

The Boundary Displacement Error measures the average displacement error between the boundary pixels in the predicted segmentation and the corresponding closes boundary pixels in the ground truth segmentation as follows:

$$\mu_{LA}\,(u, v) = \begin{cases} \dfrac{u - v}{L - 1}; & 0 < u - v. \end{cases} \quad (8)$$

Mean Absolute Error (MAE) is the average of the difference between the predicted and the actual values in all test cases; that is, it is the average prediction error. It is a quantity used to measure how close forecasts or predictions are to the eventual outcomes and the mathematical representation is given as

$$\text{MAE} = \frac{1}{N}\sum_i \sum_j \left|E_{ij} - O_{ij}\right|, \quad (9)$$

where N is the size of the image, E is the edge image, and O is the original image.

5. Experimental Results

The experimental outcome of the implemented architecture presented in Figure 1 is analysed at two stages, namely, first after the classification stage and second after the segmentation stage. Both the training and validation accuracies of the classifier are presented using the confusion matrices in Figure 5. The ROC curves of the training and validation stages of the classification model are presented in Figure 6. Also, the performance of the segmentation algorithm is illustrated visually in Figure 7. Table 1 summarises the statistical performance evaluations of the complete model for selected MRIs, whereas Table 2 presents the overall performance summary of the segmentation performance at the final stage of the proposed architecture.

5.1. Results of the Classification Model. We present a confusion matrix to illustrate the performance of the classification model against the ground truth. The confusion matrices for the training dataset and the test samples are shown in Figure 5. In each confusion matrix, green squares represent TP and TN values, light orange squares represent FP and FN values, and blue squares were used to represents positive predictive value (PPV), negative predictive values (NPV), specificity (Sp), and sensitivity (Sn), respectively, in clockwise, from the top right to bottom left. Overall correct classification rate (accuracy) was given in the purple square. The classification error for the training and testing set is equal to 7.69% and 6.42%. Overall summery of the classification process is tabulated in Table 3.

As the dataset is more biased towards glioma, we have used Cohen's kappa statistic [40] as it is a very good measure of the performance of a classifier against the imbalanced class problem. It estimates the designed classifier performance against a random guessing classifier based on the frequency of the class occurrence. In general, Cohen's kappa value above 0.81 is an indication of a perfect classifier while a value less than 0 indicates a nonperforming classification outcome. Cohen's kappa statistics value for the proposed model is 0.843 in training and 0.872 in testing which indicates a good agreement in the classification process.

Area Under the Curve-Receiver Operating Characteristic (AUC-ROC) is another powerful metric used to evaluate the performance of machine learning algorithms with an imbalance dataset. ROC curve illustrates TP rate versus FP rate at various threshold values and is commonly used in medical statistics. The ROC curve for the training model and testing are shown in Figure 6 and AUC values are at 0.93 and 0.94, respectively.
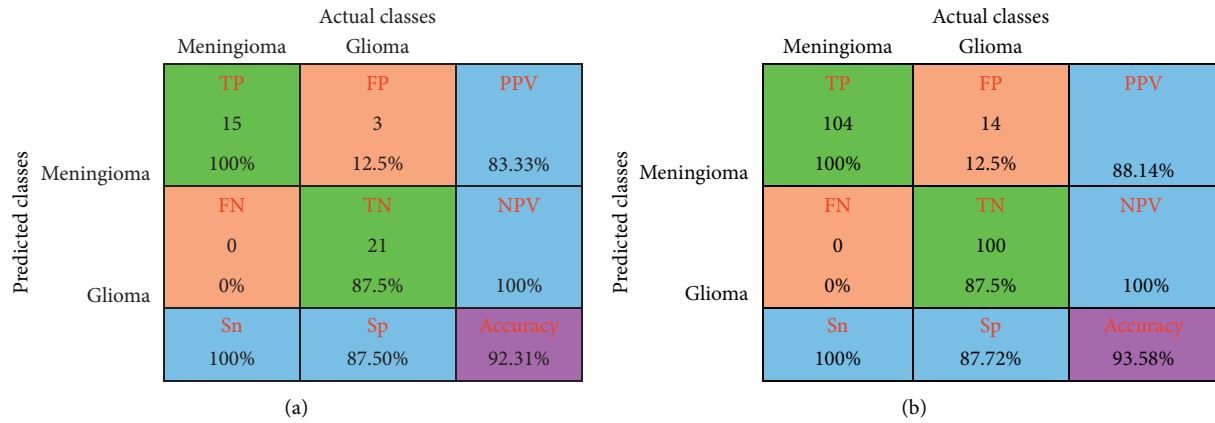
		Actual classes	
		Meningioma	Glioma	
Predicted classes	Meningioma	TP 15 100%	FP 3 12.5%	PPV 83.33%
	Glioma	FN 0 0%	TN 21 87.5%	NPV 100%
		Sn 100%	Sp 87.50%	Accuracy 92.31%

(a)

		Actual classes	
		Meningioma	Glioma	
Predicted classes	Meningioma	TP 104 100%	FP 14 12.5%	PPV 88.14%
	Glioma	FN 0 0%	TN 100 87.5%	NPV 100%
		Sn 100%	Sp 87.72%	Accuracy 93.58%

(b)

FIGURE 5: Confusion matrices. (a) Confusion matrix for testing. (b) Confusion matrix for training.

(a)

(b)

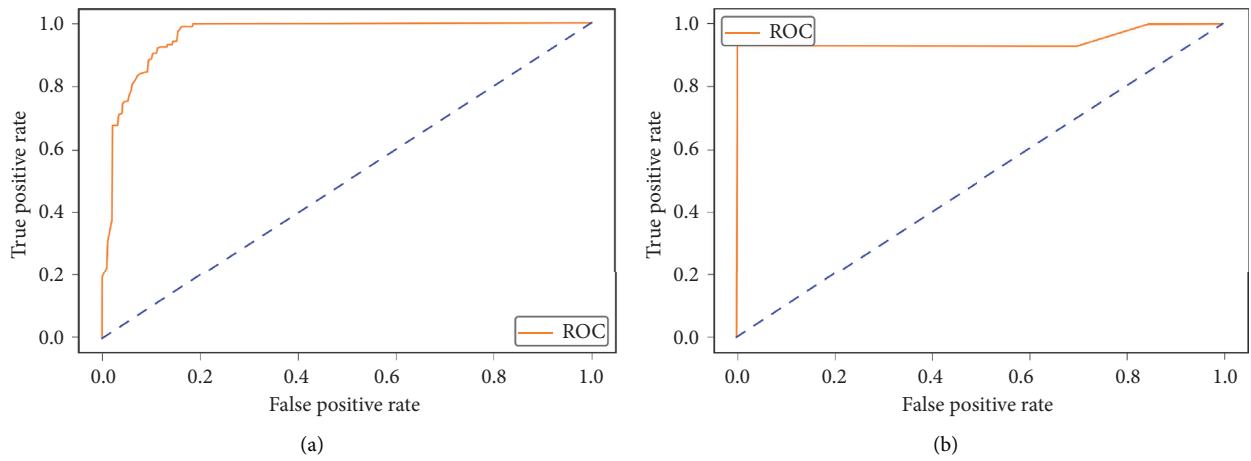FIGURE 6: Receiver operation characteristic (ROC) curve. (a) Training. (b) Testing.

(a)   (b)   (c)
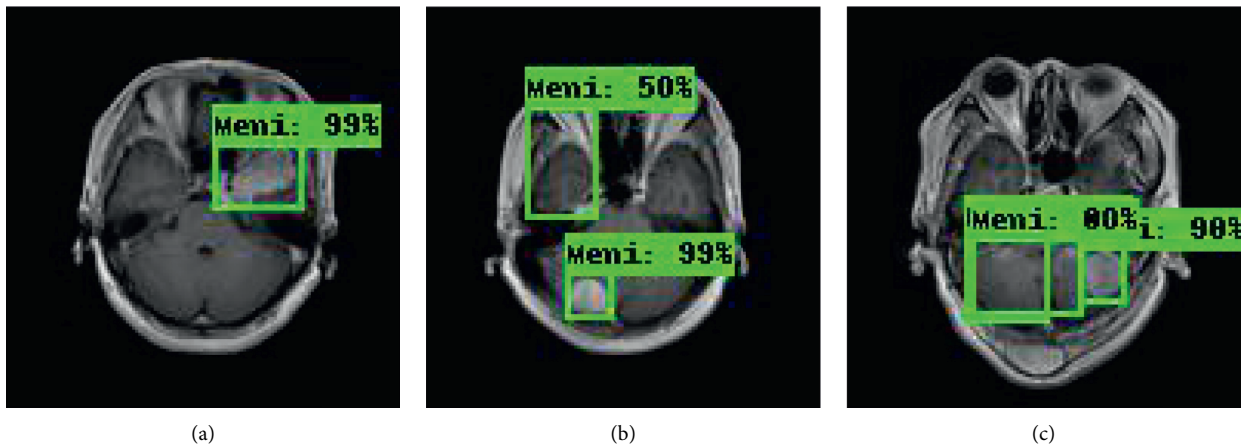
FIGURE 7: Random sample of tumor localization of the faster RCNN.

5.2. Results of the Segmentation Model.

In our proposed model, the Faster R-CNN extracts the bounding box of the tumor and it is followed by the Chan–Vese algorithm to obtain a precise tumor boundary for an accurate segmentation of the tumors. The Faster R-CNN model was able to generate the boundary boxes with 93.6% confidence interval and 99.81% accuracy. Figure 7 illustrates random sample outputs of the faster R-CNN. As in Figure 7(a), the tumor is localized correctly with 99% confidence interval. There are two possible outcomes predicted in Figure 7(b). Once closely

TABLE 1: Performance comparison between Chan–Vese (CV) and Prewitt for selected tumor MRIs.

| Sample MRIs	Tumor 1		Tumor 2		Tumor 3		Tumor 4		Tumor 5	
Confidence level of detection	0.99		0.98		0.98		0.99		Correct detection with 80%, erroneous diagnosis 98%.	
Boundary outlining method	CV	prewitt	CV	prewitt	CV	prewitt	CV	prewitt	CV	prewitt
Dice Score	0.9256	0.9025	0.9125	0.8545	0.9237	0.8723	0.9164	0.8845	0.6256	0.5941
Accuracy	0.9427	0.9254	0.9485	0.9123	0.9486	0.9406	0.9503	0.9372	0.7831	0.7232
RI	0.9872	0.9542	0.9953	0.9648	0.9967	0.9865	0.9955	0.9947	0.8263	0.7639
VOI	0.0112	0.1626	0.0252	0.2683	0.0281	0.0965	0.0376	0.0752	0.873	1.94
GCE	0.0034	0.1795	0.005	0.068	0.0031	0.046	0.0044	0.0089	0.7361	0.927
BDE	1.83	2.57	2.0298	3.325	1.3171	2.56	1.7234	2.78	5.8263	7.7836
PSNR	30.5287	13.18	27.2865	15.26	27.8105	9.56	26.499	16.25	13.578	9.354
MAE	58.54	143.23	51.28	121.56	65.23	128.26	45.6	99.57	86.6521	145.764
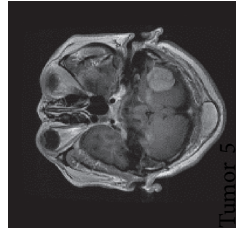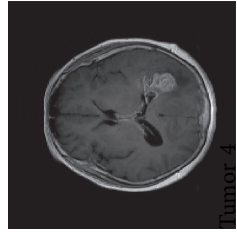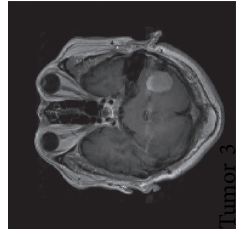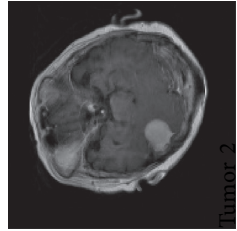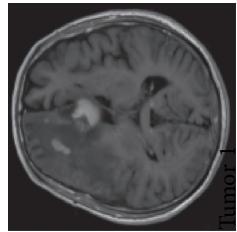
Table 2: Overall performance of the system.

Boundary estimation method	Chan–Vese	Prewitt
Dice Score	0.92	0.90
Accuracy	0.9457	0.9125
RI	0.9936	0.9642
VOI	0.0301	0.1542
GCE	0.004	0.083
BDE	2.099	3.581
PSNR	77.076	13.43
MAE	52.946	126.47

Table 3: Performance of the classification network.

Brain tumor class	Glioma	Meningioma
No if images taken for transfer learning	114	104
Training accuracy (%)	88.5	100
Testing accuracy (%)	87.5	100
No if images taken for testing	24	15
No if correctly predicted images	21	15

examined, it is evident that the false positive area, which comprises brain matter, is predicted as tumour only with a 50% confidence interval. The correct tumor area in Figure 7(b) is identified with a 99% confidence interval.

However, in certain slices of axial MRIs where the image consists of complex anatomical structures such as skull and the eye socket, a false detection could happen with a high confidence interval (example: tumor 5 of Table 1). In such scenarios, the system fails to completely recover from the erroneous detection. Yet, a close examination of the results presented in Table 1 yields that the most accurate detections have a confidence level greater than 90%, compared to the false detections less than 80%. Another false detection is presented in Figure 7(c), where the false detection carries 80% of confidence level, while all the other positive detections carry a confidence level of 98%.

To determine the overall performance of the proposed Chan–Vese algorithm-based system, we compare the output of Chan–Vese algorithm against the ground truth, gold standard, as well as against the output of a simple gradient-based edge detection technique. Prewitt is adopted in the experiment as the gradient-based edge detection method. The statistical quality measures obtained for the segmented output of both the Chan–Vese algorithm and the Prewitt edge detection are tabulated in Table 1. Columns with Tumors 1–4 present positive detections, while column with Tumor 5 presents a false detection. The value of GCE, VOI, and BDE must be low, whereas the RI should be high for a good segmented image. It is observed from the results that the Chan–Vese algorithm exhibits a superior performance than the gradient-based edge detection algorithm, Prewitt. This is also evident from the visual inspections shown in Figure 8. Although the RI for both Chan–Vese and Prewitt algorithms have a significantly higher score for all the test images, the RI values of the Chan–Vese algorithm have a relatively higher value than the Prewitt. Also, the MAE value, which should be low for better prediction, is the lowest for Chan–Vese according to Table 1. Hence, we can conclude

that the tumor boundary detected by Chan–Vese is superior to that of Prewitt.

However, in some critical cases, as illustrated in Figure 7(c), the system may not perform as desired. Furthermore, the performance measurements for such a scenario are presented in the last column of Table 1, Tumor 5. It is evident that some of the performance measures for false detection, specifically GCE, VOI, and BDE, show a deviation from that of the positive detection.

The main reason behind false detection is the missprediction and classification of brain matters as tumors. These inaccuracies in the model can be easily rectified by further tuning the proposed model using a larger dataset.

The objective of the system presented in this manuscript is to obtain a finer segmentation of brain tumors. To achieve this objective, we firstly employed Faster R-CNN to obtain the initial boundary region and secondly used the Chan–Vese active contour to refine the initial boundary assessment to obtain an exact boundary extraction. Table 2 summarises the overall performance of the proposed architecture. According to the objective measurements, the overall system achieved a Dice Score of 0.92, accuracy of 0.946, RI of 0.99, and PSNR of 77.1, against the gold standard, pointing to excellent precision of the proposed segmentation method.

6. Discussion

This study presents an automated MRI tumor classification and segmentation algorithm based on deep learning techniques and active contours. Results obtained from the experiments demonstrate remarkable performance at brain tumor segmentation with a Dice Score of 0.92, accuracy of 0.9457, RI of 0.9936, VOI of 0.0301, GCE of 0.004, BDE of 2.099, PSNR of 77.076, and MAE of 52.946.

The model presented in this manuscript provides a 0.915 Dice Score for glioma segmentation with the Figshare data set [11, 13], with the faster R-CNN and Chan–Vese algorithms. In comparison, the authors in [41] have developed a
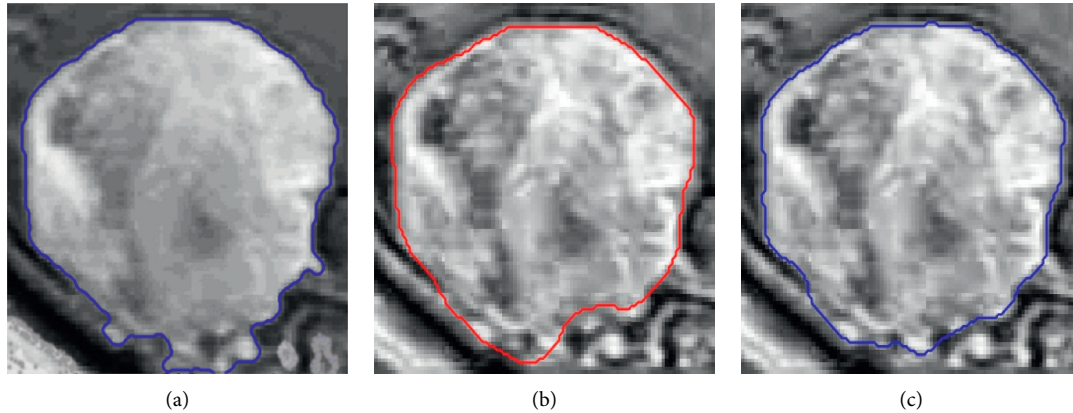
Figure 8: Segmented images versus ground truth. (a) Ground truth. (b) Prewitt. (c) Chan–Vese.

CNN based glioma segmentation algorithm and achieved a 0.87 Dice Score on BRATS 2013 and 2015 datasets. Experimental results were presented in [20] for an accurate glioma segmentation algorithm which obtained 0.897 Dice Score. An automatic semantic segmentation model was developed on the BRATS 2013 dataset by the authors in [42] and the Dice Score was around 0.80.

In our study, 0.926 Dice Score was obtained only for the meningioma segmentation which is comparably high among similar research works. The authors in [43] obtained dice coefficient of around 0.81 for 56 meningioma cases by using deep learning on a routine multiparametric model. One recent study [44] achieved dice coefficients ranging between 0.82 and 0.91 by employing an algorithm based on a 3D deep convolutional neural network.

Furthermore, the authors in [45] proposed a CNN-based algorithm on the same Figshare dataset [13] and achieved a Dice Score of 0.71 for both meningioma and glioma together with axial MR images. The authors in [37] were able to increase the segmentation accuracy using Cascaded Deep Neural Networks and obtained around 0.8 in Dice Score in both Meningioma and Glioma segmentation. In our study, we achieved an average Dice Score of 0.92 for both meningioma and glioma segmentation using the same dataset.

Hence, a comparison with the comparable state-of-the-art methods shows that the proposed methodology exhibits a remarkable improvement in brain tumor segmentation. Nevertheless, all these researchers proved that deep neural networks are capable of performing significantly accurate brain tumor segmentation in MR images. We show that the segmentation output can be further improved by using active contour algorithms along with deep learning architectures.

7. Conclusions

In the research presented, we have proposed R-CNN and Chan–Vese algorithms based model for meningioma and glioma brain tumor classification and segmentation. The proposed model is validated using Figshare dataset with 5-fold cross-validation and objective quality metrics Dice Score, RI, VOI, GCE, BDE, PSNR, and MAE are calculated to analyse the performance of the proposed architecture. We have used R-CNN to obtain the initial tumor boundary box which is followed by active contouring to obtain the exact tumor outline. We adopt level set functions based Chan–Vese algorithm which is independent of Rician noise, for both meningioma and glioma brain tumor segmentation and we compare the performance of the proposed segmentation method against that of the typical-gradient based edge detection algorithm Prewitt. We were able to achieve a much more accurate segmentation result through the Chan–Vese algorithm with an average Dice Score of 0.92 for both tumor types. In addition, experts in the field have cross-checked our segmentation output and have validated it with a high confidence interval. In the research presented, we have shown with evidence that active contour algorithms along with localized outputs of deep learning architectures such as R-CNN, are capable of improving the segmentation accuracy and the precision in MRI tumor segmentation applications. Hence, we can conclude that the model present can be used as a reliable aid for brain tumor classification and segmentation in low human resource, expertise, environments.

Appendix

Proposed Algorithm

The first stage of the proposed system is the classification, where input MRI images are analysed to detect the presence and absence of the tumours. To achieve this objective, a simple CNN architecture with a total of 5 layers was designed. The CNN designed consists of two 2D convolutional (2DConv) layers with 20-3x3 kernels and 10-3x3 kernels, respectively. Both of the 2DConv layers adopt nonlinearity through Relu activation function and each layer is followed by maxpooling layers with 2x2 window size. The 2D feature map generated through convolution and maxpooling is converted into 1D feature vectors by flattening and these 1D feature vectors are fed into the fully connected layer for the classification task. The output layer classifies the inputs as a meningioma or glioma using the softmax function. Adam optimizer was used as the optimization function of the convolution layers and the learning rate was

set to 0.02 in the training process. The proposed classification architecture was trained and validated, before testing for classification accuracy. The validation process was carried out by using 5-fold cross-validation with 500 epochs at each validation fold. At the end of the first stage, the system is capable of classifying the input MRI according to the tumor type.

Next, the classified images were fed into the second stage of the proposed architecture, which is the Region Proposal Network (RPN) based faster R-CNN for tumor localization. Due to the limitation of annotated images in the dataset, the transfer learning method with pretrained faster R-CNN weights [46] was utilized in our architecture. The faster R-CNN was able to localize the tumor with a bounding box and produce a confidence interval for each decision.

As the last part of the proposed architecture, the Chan–Vese active contour algorithm was used to segment the tumor in the Axial MRI. In general, active contour algorithms are utilized to obtain finer and precise object boundaries, given a user-defined initial guessed boundary. Thus, in MRI brain tumor segmentation using Chan–Vese algorithm, it is essential to provide a reasonably accurate initial region of interest search area for successful segmentation. We have used the bounding box obtained after faster R-CNN as the initial boundary for the Chan–Vese algorithm to automate the entire brain tumor segmentation process.

Data Availability

All MRI data are publicly available in the brain tumor imaging archive (https://figshare.com/articles/brain_tumor_dataset/1512427).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] A. Kabir Anaraki, M. Ayati, and F. Kazemi, "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 63–74, 2019.

[2] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018.

[3] S. Khan and S.-P. Yong, "A deep learning architecture for classifying medical images of anatomy object," in *Proceedings of the 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1661–1668, Kuala Lumpur, Malaysia, December 2017.

[4] M. W. Nadeem, M. A. Al Ghamdi, M. Hussain et al., "Brain tumor analysis empowered with deep learning: a review, taxonomy, and future challenges," *Brain Sciences*, vol. 10, no. 2, pp. 1–33, 2020.

[5] "Health in the European Union–facts and figures-Statistics Explained," 2019, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Health_in_the_European_Union_–_facts_and_figures.

[6] Bureau of Labor Statistics, "Radiologic and MRI Technologists: Occupational Outlook Handbook:: U.S. Bureau of Labor Statistics," United States Department of Labor, 2018, https://www.bls.gov/ooh/healthcare/radiologic-technologists.htm.

[7] A. Işin, C. Direkoǧlu, and M. Şah, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Procedia Computer Science*, vol. 102, pp. 317–324, 2016.

[8] P. Mlynarski, H. Delingette, A. Criminisi, and N. Ayache, "Deep learning with mixed supervision for brain tumor segmentation," *Journal of Medical Imaging*, vol. 6, no. 3, p. 1, 2019.

[9] C. G. Madamombe, "Deep learning techniques to classify and analyze medical imaging data," *International Journal of Computational Science and Engineering*, vol. 7, no. 4, pp. 109–113, 2018.

[10] M. Soltaninejad, G. Yang, T. Lambrou et al., "Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 69–84, 2018.

[11] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-net based fully convolutional networks," *Communications in Computer and Information Science*, vol. 1, pp. 506–517, 2017.

[12] M. Soltaninejad, G. Yang, T. Lambrou et al., "Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in flair MRI," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 2, pp. 183–203, 2017.

[13] J. Cheng, "Brain Tumor Dataset", figshare. Dataset, 2017, https://doi.org/10.6084/m9.%20figshare.1512427.v5.

[14] J. Cheng, W. Huang, S. Cao et al., "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS One*, vol. 10, no. 10, Article ID e0140381, 2015.

[15] Z. N. K. Swati, Q. Zhao, M. Kabir et al., "Content-based brain tumor retrieval for MR images using transfer learning," *IEEE Access*, vol. 7, pp. 17809–17822, 2019.

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

[17] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, "A deep learning model based on concatenation approach for the diagnosis of brain tumor," *IEEE Access*, vol. 8, pp. 55135–55144, 2020.

[18] S. Tripathi, A. Verma, and N. Sharma, "Automatic segmentation of brain tumour in MR images using an enhanced deep learning approach," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp. 1–10, 2020.

[19] A. M. Alqudah, H. Alquraan, I. A. Qasmieh, A. Alqudah, and W. Al-Sharu, "Brain tumor classification using deep learning technique - a comparison between cropped, uncropped, and segmented lesion images with different sizes," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3684–3691, 2019.

[20] X. Chen, M Zeng, Y. Tong et al., "Automatic prediction of MGMT status in glioblastoma via deep learning-based MR image analysis," vol. 2020, Article ID 9258649, 9 pages, 2020.

[21] H. N. T. K. Kaldera, S. R. Gunasekara, and M. B. DIssanayake, "Brain tumor classification and segmentation using faster R-CNN," in *Proceedings of the 2019 Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, UAE, March 2019.

[22] H. N. T. K. Kaldera, S. R. Gunasekara, and M. B. Dissanayake, "MRI based Glioma segmentation using Deep Learning algorithms," in *Proceedings of the International Research Conference on Smart Computing and Systems Engineering (SCSE)*, pp. 51–56, Colombo, Sri Lanka, 2019.

[23] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[24] T. R. Shultz, S. E. Fahlman, G. I. Webb, and C. Sammut, "Clustering," in *Encyclopedia of Machine Learning and Data Mining*, Springer, Berlin, Germany, 2017.

[25] S. Yang and G. Berdine, "The receiver operating characteristic (ROC) curve," *The Southwest Respiratory and Critical Care Chronicles*, vol. 5, no. 19, p. 34, 2017.

[26] T. F. Chan, B. Y. Sandberg, and L. A. Vese, "Active contours without edges for vector-valued images," *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 130–141, 2000.

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," 2015, https://arxiv.org/abs/1506.01497.

[28] V. G. Narendra and K. S. Hareesh, "Study and comparison of various image edge detection techniques used in quality inspection and evaluation of agricultural and food products by computer vision," *International Journal of Agricultural and Biological Engineering*, vol. 4, no. 2, 2011.

[29] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[30] X.-F. Wang, D.-S. Huang, and H. Xu, "An efficient local Chan-Vese model for image segmentation," *Pattern Recognition*, vol. 43, no. 3, pp. 603–618, 2010.

[31] S. Z. Lin Sun, X. Meng, and J. Xu, "An image segmentation method based on improved regularized level set model," *Applied Sciences*, vol. 8, no. 12, 2018.

[32] X. Jiang, R. Zhang, and S. Nie, "Image segmentation based on level set method," *Physics Procedia*, vol. 33, pp. 840–845, 2012.

[33] P. Getreuer, "Chan–vese segmentation simplified mumford – shah model level set functions," *Image Processing On Line*, vol. 2, pp. 214–224, 2012.

[34] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[35] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[36] L. Lhotska, L. Sukupova, I. Lacković, and G. S. Ibbott, *World Congress on Medical Physics and Biomedical Engineering 2018*, Vol. 1, Springer, Berlin, Germany, 2018.

[37] Z. Sobhaninia, S. Rezaei, N. Karimi, A. Emami, and S. Samavi, "Brain tumor segmentation by cascaded deep neural networks using multiple image scales," in *Proceedings of the 2020 28th Iranian Conference on Electrical Engineering (ICEE)*, pp. 11–14, Tabriz, Iran, August 2020.

[38] R. Kumar and A. M. Arthanariee, "Performance Evaluation and comparative Analysis of Proposed Image Segmentation Algorithm," *Indian Journal of Science and Technology*, vol. 7, no. 1, pp. 39–47, 2014.

[39] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, 2015.

[40] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 2020.

[41] S. Hussain, S. M. Anwar, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," *Neurocomputing*, vol. 282, pp. 248–261, 2018.

[42] S. Cui, L. Mao, J. Jiang, C. Liu, and S. Xiong, "Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network," *Journal of Healthcare Engineering*, vol. 2018, Article ID 4940593, 14 pages, 2018.

[43] K. R. Laukamp, F. Thiele, G. Shakirin et al., "Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI," *European Radiology*, vol. 29, no. 1, pp. 124–132, 2019.

[44] K. R. Laukamp, L. Pennig, F. Thiele et al., "Automated meningioma segmentation in multiparametric MRI: comparable effectiveness of a deep learning model and manual segmentation," *Clinical Neuroradiology*, 2020.

[45] Z. Sobhaninia, "Brain Tumor segmentation using deep learning by type specific sorting of images," 2018, https://arxiv.org/abs/1809.07786.

[46] "models/research/object_detection/samples/configs at master·tensorflow/models," 2020, https://github.com/tensorflow/models/tree/master/research/object_detection/samples/configs.