*Article*

# A Novel Deep Learning Method for Recognition and Classification of Brain Tumors from MRI Images

Momina Masood [1], Tahira Nazir [1], Marriam Nawaz [1], Awais Mehmood [1], Junaid Rashid [2,*], Hyuk-Yoon Kwon [3,*], Toqeer Mahmood [4] and Amir Hussain [5]

1   Department of Computer Science, University of Engineering and Technology, Taxila 47050, Pakistan; momina.masood@uettaxila.edu.pk (M.M.); tahira.nazir77@gmail.com (T.N.); marriam.nawaz@uettaxila.edu.pk (M.N.); awais.mehmood@uettaxila.edu.pk (A.M.)
2   Department of Computer Science, AIR University Islamabad, Aerospace and Aviation Campus Kamra, Kamra 43570, Pakistan
3   Department of Industrial Engineering, Seoul National University of Science and Technology, Seoul 01811, Korea
4   Department of Computer Science, National Textile University, Faisalabad 37610, Pakistan; toqeer.mahmood@yahoo.com
5   Data Science and Cyber Analytics Research Group, Edinburgh Napier University, Edinburgh EH11 4DY, UK; a.hussain@napier.ac.uk
*   Correspondence: junaidrashid062@gmail.com (J.R.); hyukyoon.kwon@seoultech.ac.kr (H.-Y.K.)

**Abstract:** A brain tumor is an abnormal growth in brain cells that causes damage to various blood vessels and nerves in the human body. An earlier and accurate diagnosis of the brain tumor is of foremost important to avoid future complications. Precise segmentation of brain tumors provides a basis for surgical planning and treatment to doctors. Manual detection using MRI images is computationally complex in cases where the survival of the patient is dependent on timely treatment, and the performance relies on domain expertise. Therefore, computerized detection of tumors is still a challenging task due to significant variations in their location and structure, i.e., irregular shapes and ambiguous boundaries. In this study, we propose a custom Mask Region-based Convolution neural network (Mask RCNN) with a densenet-41 backbone architecture that is trained via transfer learning for precise classification and segmentation of brain tumors. Our method is evaluated on two different benchmark datasets using various quantitative measures. Comparative results show that the custom Mask-RCNN can more precisely detect tumor locations using bounding boxes and return segmentation masks to provide exact tumor regions. Our proposed model achieved an accuracy of 96.3% and 98.34% for segmentation and classification respectively, demonstrating enhanced robustness compared to state-of-the-art approaches.

**Keywords:** MRI; brain tumor; Mask-RCNN; deep learning

## 1. Introduction

A brain tumor is a fatal disease-causing death to thousands of people around the globe. A brain tumor is mainly caused by abnormal growth in brain tissues. As the skull portion of the human body is inflexible and small, any growth inside the brain may affect the functionality of the human organ depending on its origin and position. Moreover, it may also spread in other parts of the body and affect their functionality [1]. Usually, the brain tumor is categorized into two classes, named primary and secondary based on its position. The primary tumor comprises 70% of all brain tumors while the remaining 30% are secondary [2]. A primary brain tumor includes tumors that originate from the brain cells while a secondary brain tumor first originates in another organ and then transfers to the brain through the circulation of the blood [3]. According to an NBTF study, in the USA an estimated 29,000 cases are diagnosed with primary brain tumor, among which,

around 13,000 patients die per year [4]. Similarly, in the UK, more than 42,000 patients with a primary brain tumor die annually.

Among various primary brain tumor types, glioma has the highest mortality and mobility rate [5]. Gliomas usually grow from glial cells of the brain [1] and are classified as low-grade (LG) glioma and high-grade (HG) glioma. The HG glioma is more life-threatening and intense and usually, the victim can survive for two years [6]. A meningioma tumor usually develops in the protective membrane layer which acts as a covering of the human brain and spinal cord. Mostly, meningioma tumors are less threatening and slow-growing [7]. The pituitary tumor starts developing in the pituitary gland, which is located at the base of the brain and is involved in the production of several essential hormones in the body [8]. A pituitary tumor is a benign tumor; however, serious complications may cause hormonal deficiencies or permanent loss of vision because of the overproduction of hormones [1]. Hence, an early-stage detection of brain tumors is critical and is of extreme clinical interest. If it is not diagnosed on time, the disease could become life-threatening or may result in the disability of a person [7].

Depending on the situation and their purpose, numerous medical imaging techniques can be used in clinical practices for tumor diagnosis [9]. Some of those techniques are ultrasonography (US), magnetic resonance imaging (MRI), and computed tomography (CT) [10]. The MRI is the most common non-invasive imaging technique because it uses no damaging ionizing radiations of X-rays during the scan. Moreover, it provides high-quality images of soft tissue without any risk, plus an ability to acquire multiple modalities, e.g., T1, T1c, T2, and FLAIR, using various parameters. Each of these modalities produces noticeably a unique tissue contrast [11].

For timely treatment, the top priority of a neurosurgeon is to mark out tumor regions as precisely as possible, otherwise excessive or insufficient cutting may lead to suffering or a permanent loss. Unfortunately, this manual segmentation process is laborious and time-consuming and yields poor segmentation results [12]. Hence, the use of computer-aided brain tumor segmentation algorithms using MRI to identify and segment brain tumors has received considerable attention from the research community. Therefore, there is a significant need for automated and efficient tumor detection and segmentation technique. Despite recent developments in automatic or semi-automatic techniques for tumor segmentation, it is still a challenging task to segment a tumor accurately because of the following reasons [13]. First, there is a significant change in tumor location, shape, appearance, and size from patient to patient [14]. Second, the tumor boundaries can be discontinuous or blurry as the tumor regions are usually occupied by surrounding healthy tissues [15]. Third, the addition of inadequate signal-to-noise ratio or image distortion usually caused by different factors such as MRI acquisition protocols or variation in imaging devices may further increase the difficulty and influence the precision of final segmentation [16].

The brain tumor detection approaches can be divided into two types, named machine learning (ML)-based [17] and deep learning (DL)-based [18] methods. ML-based techniques mainly include support vector machine [19], conditional random forest [20], decision tree [21], principal component analysis [22], and fuzzy c-means [23]. These techniques require hand-crafted features. The hand-crafted features, here, mean that the features are required to be extracted from training images to start the learning process and perhaps require an expert with extensive knowledge to identify the most important features. Hence, the detection accuracy of the ML-based techniques is dependent on the quality and representation of the extracted features, thus is limited and prone to errors in dealing with large datasets [24]. Meanwhile, DL-based algorithms have shown high performance in various industries including medical imaging [25–27]. The most common or well-known DL model is the convolutional neural network (CNN) that can instinctively learn dense characteristics directly from the training data due to its weight-sharing nature [28]. Based on these advantages, DL-based brain tumor segmentation has grabbed the researcher's attention [29]. Relevant works include patch-based CNN [30], patch-based

multi-scale CNN [31], patch-based DCNN [32], fully convolution-based CNN (FCNN) [33], and U-net based [34] brain tumor segmentation models. The patch-based approaches take a small portion of the image as input to CNN and classify each patch into a different class, which degrades the image content and label correlation. The FCNN, on the other hand, is a modified form of CNN, which predicts probability distribution pixel-wise instead of making patch-wise probability distribution predictions [35]. This improvement enables FCNN to take the full-sized image and perform the prediction for the whole image in just a single forward pass. Despite recent advances, the existing DL-based techniques require several convolution layers (CLs) and kernels, increasing the computational cost resultantly. Hence, an efficient method for accurate tumor identification and segmentation with a less complex network in terms of memory and computing resources usage is still in demand [36].

In this paper, we proposed an automated approach for brain tumor detection and segmentation using MRI images. The proposed technique adopts a DL model using a fully convolution neural network, Mask-RCNN [37] with DenseNet-41 backbone, and utilizes a multitask loss function to achieve an end-to-end training of deep CNN, increasing the detection accuracy. The motivation behind using a custom Mask-RCNN was to achieve a similar level of accuracy with a comparatively simple model, fewer kernels, and two convolutional layers. To show the efficiency of the proposed approach, we evaluated our model on free and online available brain tumor datasets [38,39] using various quantitative measures. The results of brain tumor MRI segmentation are validated through ground truth analysis. The recent works have investigated the effectiveness of a Mask-RCNN model on 3D images of the brain [40] and in other medical fields such as oral disease [41], breast tumor [42], and lung tumor [43] detection and segmentation. The main contributions of the proposed work are as follows:

1.  The proposed method can precisely segment and classify the brain tumors from MRI images under the presence of blurring, noise, and bias field-effect variations in input images.
2.  We have created the annotations which are essential for the training of the proposed model because available datasets do not have a bounding box and mask ground truths (GTs).
3.  The accurate localization and segmentation of tumor regions due to an effective region proposal network of DenseNet-41-based Mask-RCNN as it works in an end-to-end manner.
4.  Extensive experiments are performed using two different datasets to show the robustness of the presented framework and compared obtained results with the existing state-of-the-art methods.

The rest of our paper develops the following structure: the literature review is explained in Section 2, while the brief description of the proposed work is defined in Section 3. In Section 4, the datasets, evaluation parameters used, and experimental results obtained are reported. Finally, a conclusion of this work is described in Section 5.

## 2. Related Work

Due to the high clinical importance and the complex nature of the brain tumor, the advancement of an automatic model is an active research area. This section briefly discusses the relevant works for brain tumor classification and segmentation from MR images. Initially, the ML approaches such as support vector machines vector machines [19], conditional random forests [20], decision forests [21], principal component analysis [22], and fuzzy c-means [23] were presented. A common aspect of these methods is that they classify image voxels based on the pre-defined feature set known as handcrafted features, which require a human expert to figure out the most promising features from training images to start the training process.

In recent years, DL-based approaches have exhibited encouraging results in the automatic segmentation of medical imaging [25]. The most important aspect of DL approaches

is that they can learn complex feature representation automatically from the training data and thus result in a more robust feature vector. Various DL models for the automatic recognition of tumors have been presented, achieving promising results [30–34,36,44]. Pereira et al. [30] trained two different 2D CNN with deeper layers as a sliding window classifier for the segmentation of both LG and HG glioblastomas. Urban et al. [45] proposed a two pathway-based 2D CNN approach on large patches to incorporate both global and local information. A local path focuses on the information in neighbor pixels while a global path captures larger contextual information simultaneously from MRI. Kamnitsas et al. [32] introduced a 3D CNN architecture that considers 3D patches along with global contextual features through down-sampling, as post-processing, a fully connected CRF network was employed. In [46], the extended form of DeepMedic by adding residual connection is proposed for tumor segmentation. These above-mentioned approaches built using CNN for brain tumor segmentation operate at patch-level and consider local regions in MRI images which are then used to classify each patch [47]. Based on the obtained classification results, the central pixel is labeled; thus, it only explores spatially limited contextual information.

Recently, FCNNs have achieved promising results for the segmentation of natural images [35] as well as medical images [36]. In FCNN, convolutional kernels are employed instead of fully connected layers. The original size of the image is restored by using up-sampling and deconvolution layers. Moreover, the model is trained end-to-end and has computational efficiency over patch-level classification approaches. Havaei et al. [31] proposed an architecture, named InputCascadeCNN, which passes pixel-wise probability estimates obtained by first CNN as an additional input to the following one. This two-stage training strategy solves an imbalance of label distribution and captures multi-scale features using a multi-cascaded network. Zhao et al. [33] presented a unified framework by integrating FCNNs with CRFs, such that the results reported after the segmentation of tumors maintain their appearance and spatial consistency. In [28], a multi-cascade convolutional neural network was proposed that takes both local pixel dependencies and more discriminative multi-scale features of 3D MRI images into account. To further refine the obtained results, CRFs are used to smoothen the tumor edges and eliminate false positives. In [36], authors developed the U-Net an encoder–decoder-based mod which consists of a regular FCNN followed by a contracting path capturing contextual features by down-sampling at each layer and an expanding path which raises the image size by up-sampling at each layer, thus enabling precise localization and segmentation, and it is well suited for the segmentation of medical imaging [48]. Don et al. [34] adopted a U-net CNN architecture for the segmentation of a tumor with some minor modifications. They used the technique of data augmentation with dice-based loss function to improve segmentation accuracy.

In [49], a watershed segmentation algorithm was used to segment a brain tumor. However, they employed low dimensional hand-crafted feature vectors as training to the KNN classifier for classification. They achieved an average accuracy of 86%. In [50], an encoder–decoder-based architecture is presented, which performed the pixel-wise segmentation of tumor tissues from normal brain cells. This model used SegNet architecture with an encoder of depth four and VGG16 for generating feature maps, and performed non-linear up-sampling. This approach does not need any post-processing phase and reported an average dice score of 0.931. In FR-MRINet [51], a 33-layer deep model and an encoder with a fully connected decoder instead of the deconvolutional decoder were proposed. However, the approach produces anomaly areas, i.e., small non-tumor areas are predicted to contain tumors and cleaned by using the neighborhood cleaning rule. This work reported a segmentation accuracy of 91.4%.

In this work, we adopt a kind of DL method, Mask-RCNN [37] architecture, that uses end-to-end training for the localization and classification of brain tumors. Instead of applying a threshold-based or boundary-based model for accurate segmentation, this method uses region-based segmentation and generates a mask that achieves improved tumor boundary segmentation accuracy.

## 3. Proposed Methodology

This section illustrates the architecture of our presented method employed for brain tumor detection. Figure 1 shows the workflow of the proposed methodology. The presented DL technique is DenseNet-41-based Mask-RCNN, which aims to perform the accurate localization, segmentation, and classification of brain tumors. Given an MRI image, our aim is to instinctively detect the brain tumor from a complex background without requiring any manual intrusion. First, the input images are preprocessed to remove noise and artifacts added during MRI acquisition. Then, the ground truth segmentation masks are generated that are utilized for model training. Next, custom Mask-RCNN is applied for tumor localization, classification, and segmentation. The custom Mask-RCNN model performs the following steps: (1) keypoints extraction using DenseNet-41, (2) region of interest (RoI) creation, (3) RoI classification and bounding box regression, and (4) segmentation mask acquisition. First, a backbone network based on CLs computes the deep features from preprocessed input MRI images. The obtained features are then used by the region proposal network (RPN)to generate RoIs by mapping each point on the feature map into the original image. Next, the RoIAlign layer is employed to select features from the feature map corresponding to the RoIs obtained from the RPN network and distribute them with correlated layers to segment and classify the RoI.
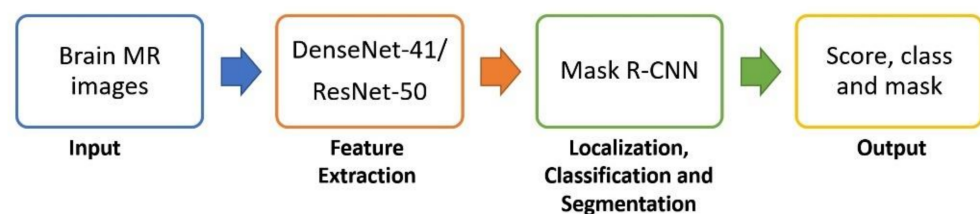


**Figure 1.** Block diagram of the proposed method.

### 3.1. Preprocessing

The MRI images produced from different MRI machines could have a bias field or intensity inhomogeneity, which is an artifact and should be corrected as it affects the segmentation results [52]. In the preprocessing step, we applied the level set method for bias field correction [53]. To obtain the enhanced image, a median filter was applied. The median filter is considered better as compared to linear filtering for eradicating noise [54].

### 3.2. Annotations

The GT mask associated with every MRI image is essential to distinguish the tumor portion for the training procedure. The VGG Image Annotator (VIA) [55] is utilized to annotate the MRI images and then produce a polygon mask for every image. Figure 2 indicates an instance of the original image and the related GT image. The VIA interpretations are saved in a JSON file which comprises the set of polygon points for the tumor region and a value of region attributed 0 or 1. The pixels inside the bounding polygon are related to a tumor region and are given a value of 1 while the rest of the pixels are regarded as background and assigned the value of 0. This file is used to create a mask image corresponding to each MRI image that is later used in the training process.
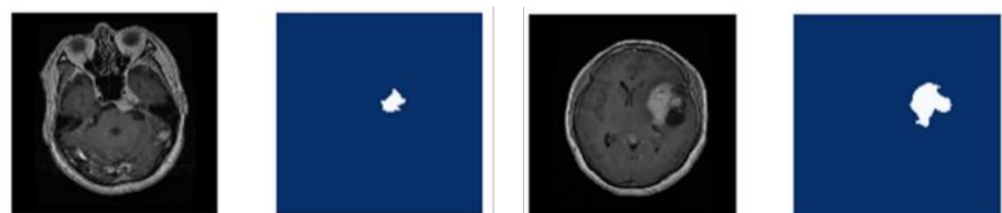


**Figure 2.** Sample original images and corresponding ground truth masks.

### 3.3. Tumor Localization and Segmentation Using Mask-RCNN

Mask-RCNN is the recent DL technique that is used for both object detection and pixel-level segmentation [37]. It is the extension of Faster RCNN [56], which performs segmentation as well along with classification and localization. Furthermore, Mask-RCNN decouples the mask and class prediction. It adds a small network overhead, i.e., an FCN network to perform segmentation. In the presented work, we proposed a custom Mask-RCNN by introducing the DenseNet-41 at the feature computation layer. DenseNet [57,58] is the latest presented approach of CNN, in which the present layer relates to all preceding layers. DenseNet comprises a set of dense blocks that are sequentially interlinked with each other with extra convolutional and pooling layers among successive dense blocks. DenseNet can present the complex transformations which result in improving the issue of the absence of the target's position information for the top-level key points to some degree. DenseNet minimizes the number of parameters, which makes them cost-efficient. Moreover, DenseNet assists the keypoints propagation process and encourages their reuse, which makes them more suitable for brain tumor classification. The structure or flow of the presented approach is illustrated in Figure 3. The custom Mask-RCNN architecture consists of different networks, i.e., a convolutional backbone network, region proposal network, RoI classifier and bounding box regressor, and the segmentation network. A detailed description of each step is discussed below.
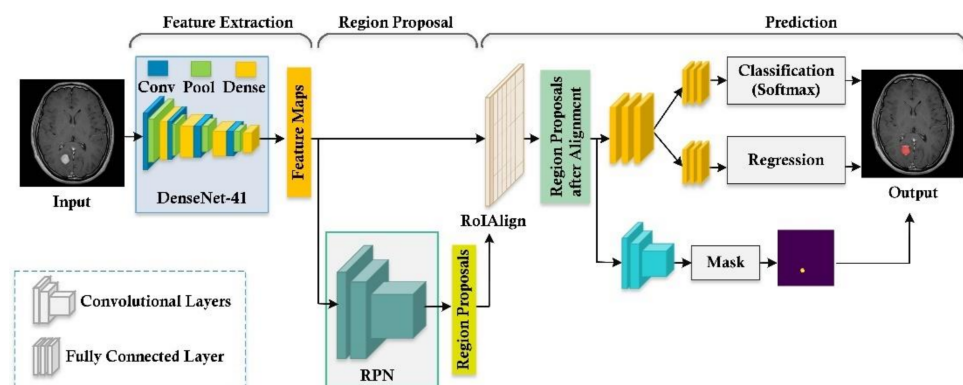


**Figure 3.** The structure of the proposed technique.

#### 3.3.1. Feature Extraction

The backbone network is employed to obtain the relevant feature from the input MR images [59]. This network could be any CNN model intended for image analysis, such as ResNet-50, ResNet-101, and DenseNet. A keypoints extraction network must be intensely adequate with many convolution layers such that it can appropriately learn reliable and more discriminating features. However, the increase in network depth size adds more computational overhead and makes it difficult to optimize the network weights, which may result in an exploding gradients problem. The work in history has employed ResNet for medical image analysis with Mask-RCNN. However, the ResNet model uses skip-connections and comprises many parameters, which eventually results in the vanishing gradient problem. In the presented work, we have implemented Mask-RCNN with both the ResNet and DenseNet framework. More specifically, we have computed features with ResNet-50 and DenseNet-41 frameworks. Since DenseNet contains dense connections, this results in computing a more representative set of image features.

The DenseNet-41 has two potential differences from the traditional DenseNet: (i) DenseNet-41 has fewer parameters from the actual model as instead of 64, it has 24 channels on the first convolution layer and the size of the kernel is $3 \times 3$ instead of $7 \times 7$; (ii) the number of layers within each dense block is attuned to deal with the computational complexity. The dense block is the fundamental part of DenseNet-41, as shown in Figure 4.
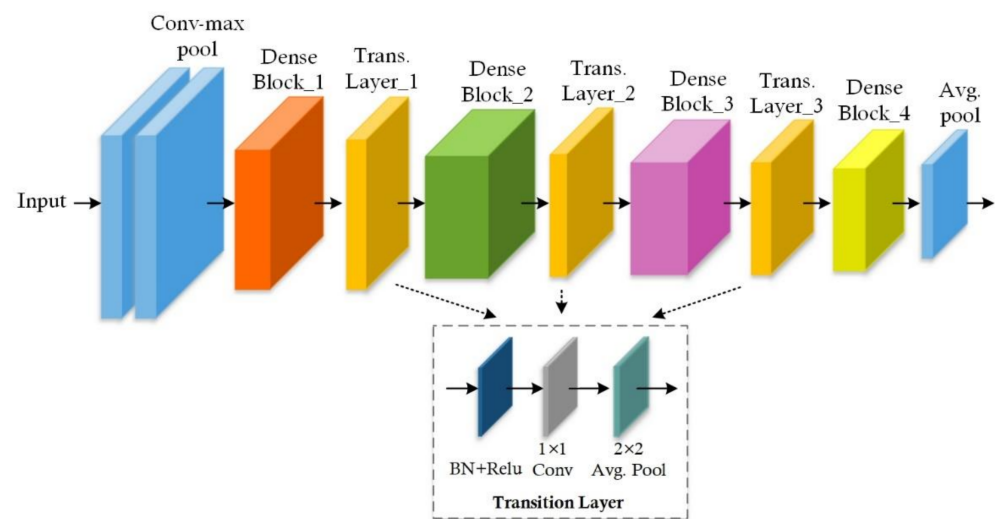
**Figure 4.** DenseNet-41 architecture.

### 3.3.2. Region Proposal Network

In this stage, the feature map obtained by feature extraction is fed to the RPN network to generate RoIs. The RoIs are localized to tumor regions used for the final segmentation and classification. The RPN module uses a 3 × 3 CL to scan the whole image in a sliding window manner to generate relevant anchors. These anchors are the bounding box with different sizes and are distributed over the whole image. Since there are about 20 k anchors of various sizes and scales, they are likely to overlap each other to cover the whole image as much as possible [60]. Using RPN prediction, the top anchors that probably include objects are selected and their position and size are refined using bounding box regression. In the case of the overlapping anchors, the ones with the highest foreground score are kept while the remaining ones are discarded by using non-max suppression. More specifically, if an anchor has an intersection-over-union (IoU) higher than 0.7 with a GT box, it is classified as a positive anchor (fg class), otherwise negative (bg class). This leads to a generation of several RoIs that are passed to the next stage for classification and segmentation.

### 3.3.3. RoI Classification and Bounding Box Regression

This network takes the proposed RoI and feature map as input (Figure 3). Unlike the RPN, which returns two classes such as foreground and background, this network is deeper and classifies proposed RoIs to a specific class, i.e., glioma, meningioma, and pituitary, and further improves the size of the bounding box. This step intends to pool all RoIs on the feature maps to a fixed size. Usually, the boundaries of RoI do not coincide with the granularity of the feature map as the feature map is down-sampled k times from the size of the original image (via convolutions). To resize the feature maps, the RoIAlign layer is utilized to obtain the fixed length of keypoint vectors for arbitrary-size candidate regions and performs the bilinear interpolation to avoid misalignment issues encountered in the RoI pooling layer, which uses quantization operation. These keypoints are fed into categorical classification and regression layers to get the ultimate recognition results.

### 3.3.4. Segmentation Mask Acquisition

In this step, the mask branch works separately from the classification and regression. The segmentation network takes positive regions chosen by the RoI classifier as input and returns a segmentation mask of 28 × 28 resolution. The obtained masks are represented by floating numbers and thus contain more information as compared to binary masks. The GT masks are scaled down to a size of 28 × 28 to measure the loss with the predicted mask during the training stage. However, during the inference phase, the predicted mask is scaled up to match the dimensions of the RoI bounding box and which provides the final output mask.

*3.4. Loss Function*

During training, the Mask-RCNN model [37] employs a multi-task loss L on each sampled RoI, defined as:

$$L(Mask\ RCNN) = L_{class} + L_{bbox} + L_{mask} \tag{1}$$

where $L_{class}$, $L_{bbox}$, and $L_{mask}$ represent the class label prediction loss, bounding box refinement loss, and segmentation mask prediction loss, respectively.

$L_{class}$ is defined as follows:

$$L_{class} = -\log P_u \tag{2}$$

where $P$ is a $(k+1)$ dimensional vector corresponding to the likelihood of a pixel going to the k class or background. For each RoI, $P = P_0, ..., P_k$ and $P_u$ is the probability related to class $u$.

$L_{bbox}$ is defined as follows:

$$L_{bbox}(v_i,\ v_i{}^*) = \sum_{i \in \{x,\ y,w,h\}6} smooth_{L1}(v_i - v_i^*) \tag{3}$$

where

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise,} \end{cases} \tag{4}$$

Vector vi represents four parameters coordinate of the predicted bounding box, and $v_i{}^*$ is the coordinate of the GT relating to the positive anchor. The smooth-$L_1$ function is a robust $L_1$ loss that is less sensitive to outliers over $L_2$ loss. When regression targets are infinite, training with $L_2$ loss can necessitate the careful modification of learning levels to avoid exploding gradients. For the training of the mask network, the average binary cross-entropy loss is employed that is given as follows:

$$L_{mask} = -\frac{1}{n^2} \sum_{1 \le i,\ j \le n} \left[ x_{ij} \cdot log P_{ij}^k + (1 - x_{ij}) \cdot log\left(1 - P_{ij}^k\right) \right] \tag{5}$$

where $x_{ij}$ represents the value of a pixel $(i, j)$ in a GT mask of size $n \times n$ and $P_{ij}^k$ is the predicted value of the same pixel in the mask learned for class $k$.

## 4. Performance Evaluation

*4.1. Experimental Setup*

For implementation, we used the Mask-RCNN model provided by Matterport Inc. [60] released under MIT license based on open-source Tensorflow and Keras Libraries. The Mask-RCNN model with both ResNet-50 and DenseNet-41 frameworks was implemented. In our work, rather than training our model from scratch, we initialized the model using pre-trained weights obtained from MS-COCO and incorporated transfer learning to fine-tune the model on the Brain MRI dataset for tumor segmentation and classification. The motivation to use pre-trained models is that the model has been trained on massive, publicly accessible datasets, such as ImageNet and MS-COCO, and is therefore capable of learning important features. During training, the initial layers learn low-level features; as the network goes up, the layers can learn task-specific patterns. Thus, when the pre-trained models are trained for a new task such as the segmentation and classification of brain tumors, the training speed and accuracy of the new model increase. As the important image features have already been learned, they do not have to be learned again and are transferred to the new task. This process is known as 'transfer learning'. For training, the given data are randomly split into training and test sets containing 70% for training and 30% images for testing. The training parameters for the proposed model using custom Mask-RCNN are displayed in Table 1.

**Table 1.** Training parameters of the presented technique.

| Parameters | Value |
|---|---|
| Epochs | 45 |
| Learning rate | 0.001 |
| IoU Threshold | 0.70 |

### 4.2. Dataset

In this research, we have used two different brain MRI datasets for the evaluation of the proposed technique. The first main dataset used is the 'Figshare brain tumor dataset' obtained from [38] and is one of the largest available datasets for brain tumor detection. It contains a total of 3064 real brain MRI samples, collected from 233 different subjects, among which, samples belonging to the meningioma class are 708, pituitary 930, and glioma 1426. The size of each image is 512 × 512 pixels. The second is the Brain MRI Dataset is obtained from [39] and is relatively small. It contains a total of 253 MRI samples of sizes of 845×845 pixels, among which, 155 MRI samples contain tumors. Both datasets are publicly available. The MRI samples are diverse in terms of structural complexity, acquisition angle, devices, noise, and bias field-effect, etc. The reason for using a T1-weighted MRI image dataset is that it is contrast-enhanced. Thus, it provides a better distinction of the areas affected by the tumor, and they are popular for treatment planning.

### 4.3. Evaluation Metrics

The segmentation results are quantitatively evaluated using parameters such as precision, recall, accuracy (Acc), dice score (DSC), and intersection-over-union (IoU). The equations used for the calculation of Acc, and DSC parameters are given as follows:

$$Acc = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{6}$$

$$DSC = \frac{2 \times TP}{(2 \times TP + FN + FP)} \tag{7}$$

where TP = true positive, TN = true negative, FP = false positive, and FN = false negative cases. Here, note that if the IoU score between the predicted tumor mask and associated GT mask exceeds the threshold value, i.e., 0.7, it will be considered as TP, otherwise it will be considered as FP and it is FN when ground truth mask has no associated predicted tumor mask. Figures 5–7 present the pictorial representation of IOU, precision, and recall parameter, respectively.
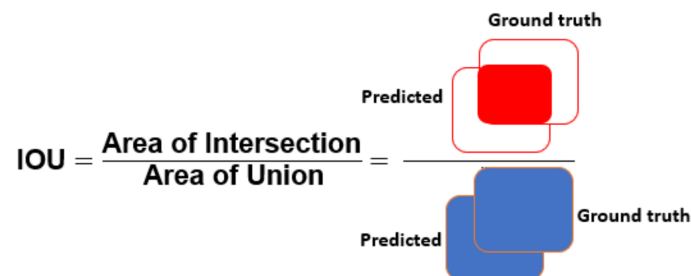


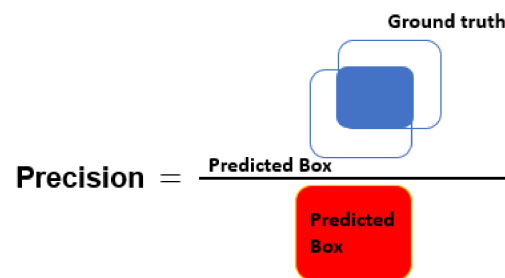**Figure 5.** Pictorial representation of IOU.
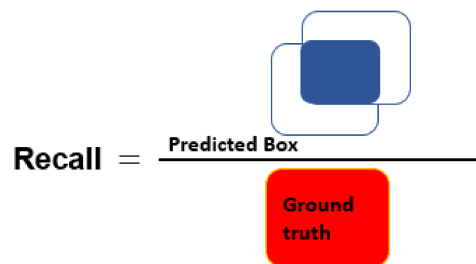
**Figure 6.** Pictorial representation of precision.



**Figure 7.** Pictorial representation of recall.

*4.4. Experimental Results and Discussion*

This section contains a detailed analysis and discussion of the obtained results. We have experimented on real MRI images using two datasets [38,39]. The proposed model uses two different deep neural networks, namely ResNet-50 and DenseNet-41 as backbone networks to learn deep features automatically from the training images. However, we obtained better results with DenseNet-41-based Mask-RCNN due to its ability to compute more robust features as compared to ResNet-50. Some of the visual results for custom Mask-RCNN using both datasets are presented in Figure 8. From the figure, we can observe that the presented approach (with DenseNet-41) can more accurately localize the brain tumor from the healthy tissues despite discontinuous or blurry boundaries and artifacts in MRI samples. Moreover, the custom Mask-RCNN (with DenseNet-41) method can precisely segment the brain tumor by overcoming the challenges of location, shape, and size.

As discussed earlier, the obtained results are evaluated using various quantitative measures such as accuracy, precision, recall, DSC, and IoU. To further understand the accuracy of our method, we have drawn a boxplot for evaluation matrices using both datasets, which is shown in Figure 9. The boxplot represents the spread of results into four quartiles, median, and an outlier for all input training images. Part (a) of the figure shows the results obtained over the Figshare dataset while part (b) presents the results for the MRI dataset with the DenseNet-41 based Mask-RCNN framework. From Figure 9, it can be seen that our approach attains better results for the Figshare database as compared to the MRI dataset. The presented approach has achieved an average accuracy and dice score of 95.9% and 0.955 with ResNet-50 and 96.3% and 0.959 using DenseNet-41 over Figshare dataset. The mean average precision (mAP) of the proposed method to localize the brain tumor region at the regression layer is 0.949. In the case of ResNet-50-based Mask-RCNN, the presented technique fails to accurately localize the brain tumor in a few images due to the visual similarity with healthy tissues, as shown in Figure 10. However, for the DenseNet-41-based network, the system has exhibited more accurate results than ResNet-50.
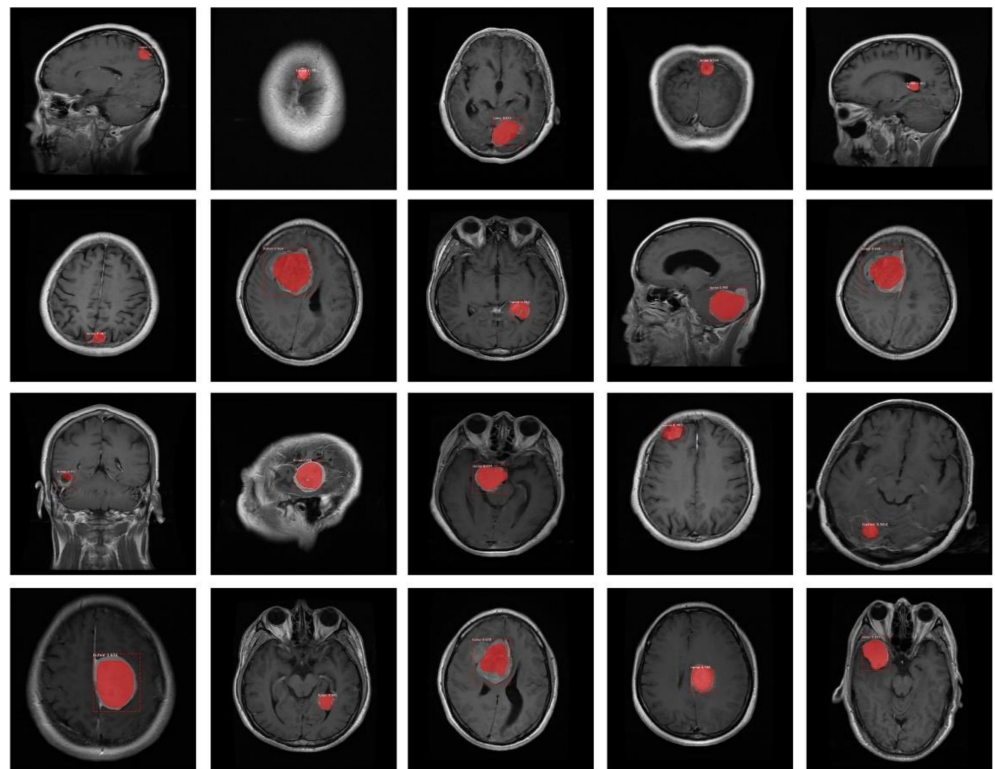
**Figure 8.** Example segmentation results of high-score-obtaining test images using the proposed method. The red contour shows the predicted tumor mask.

Figure 11 shows the confusion matrices that summarize the classification results of the proposed technique against the ground truth for the DenseNet-41-based network using Figshare [38] and Brain MRI [39] datasets. For both datasets, the presented technique attains better results with DenseNet-41-based Mask-RCNN, as reported in Figure 11. Here, part (a) presents the classification accuracy for the Figshare dataset while part (b) shows the classification results for the MRI dataset. The presented approach obtains an overall classification accuracy of 98.34% for the Figshare dataset while it achieves the accuracy value of 97.90% for the Brain MRI dataset. In the case of the Figshare dataset, the attained classification accuracies for glioma, meningioma, and pituitary tumor types over DenseNet-41 are 98.62%, 97.81%, and 98.60%, respectively, while for the Brain MRI dataset, the presented technique attains the class-wise accuracy of 97.74% and 98.06% for tumor and non-tumor classes, respectively.
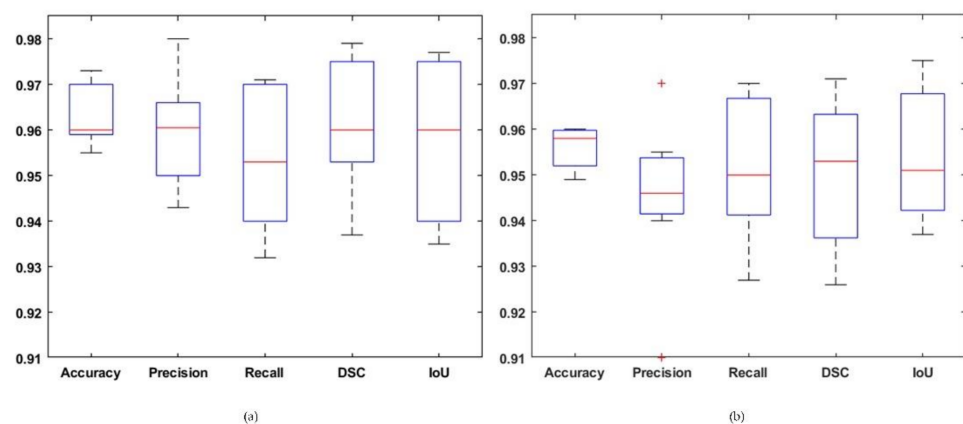


**Figure 9.** Tumor localization results of the proposed approach over datasets using DenseNet-41 (**a**) Figshare, (**b**) Brain MRI dataset. + sign shows the outer value which is larger than the other values.
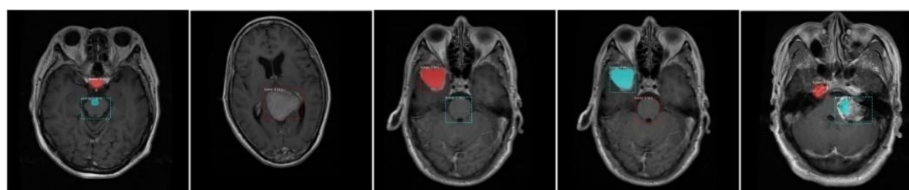
**Figure 10.** Example of inaccurately localized brain tumor images by the proposed method. The red and blue contour shows the predicted tumor region and respective masks.

### 4.4.1. Comparison with RCNN-Based Methods

The performance of our technique compared with other region-based segmentation methods, i.e., RCNN [51] and Faster RCNN [56] using Figshare brain tumor dataset [38], is reported in Table 2. The problem with RCNN methods is that they require significantly too much time to train the model as these techniques randomly generate around 2000 region proposals per image for classification. Additionally, there is no learning process at the region proposal generation step because a fixed selective search algorithm is used, which leads to the generation of false candidate region proposals. Furthermore, the processing time for the test image is approximately 47 s, which is inadequate for obtaining results in real-time. Faster RCNN extracts the region proposals automatically by introducing the region proposal network and shares the CL among class and bounding box network to expedite the process and reduce the computation cost. Faster RCNN and Mask-RCNN give results in real-time, i.e., approximately 0.23 and 0.2 s, respectively, inference time per test image. The advantage of using Mask-RCNN over Faster RCNN is the automated segmentation of the brain tumor along with localization, i.e., it defines the tumor location and draws a high-quality segmentation mask of the tumor region along with classification. Mask-RCNN delineates RoIs automatically and extracts the features from the images layer by layer without providing the features previously. This gives an advantage of detecting the tumor comprehensively besides analyzing the single features. Moreover, it is easier to train and requires a negligible additional overhead as compared to Faster RCNN. Moreover, the presented work with DenseNet-41 is more robust as compared to ResNet-50 due to its dense connections, which result in calculating a more accurate set of image features. Furthermore, the DenseNet-41-based Mask-RCNN is computationally more efficient due to its small number of parameters.
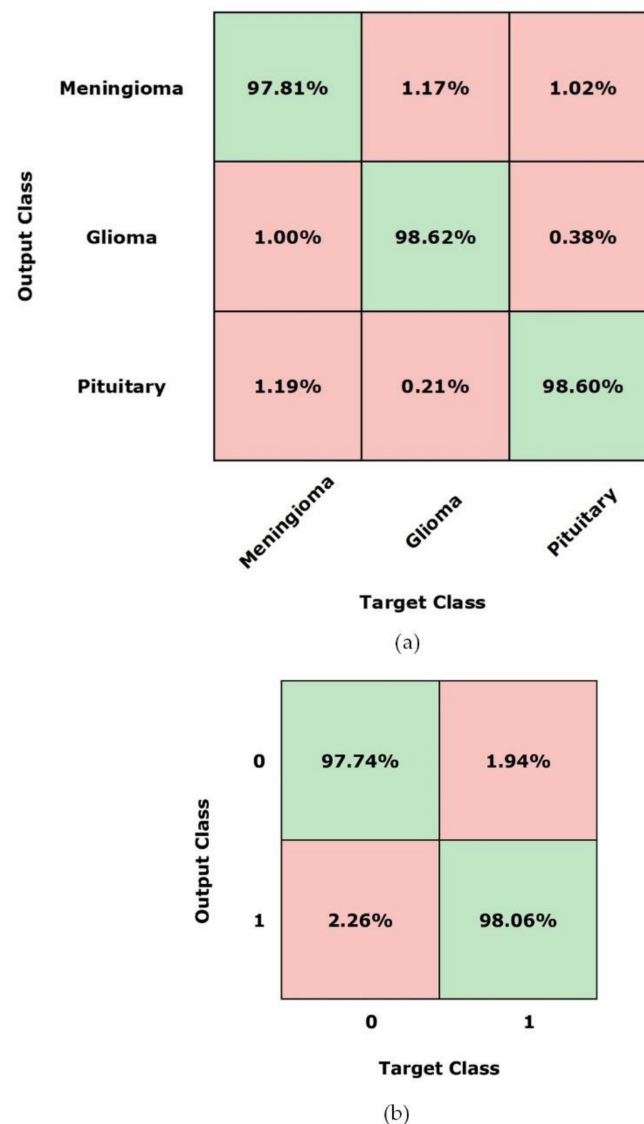
**Figure 11.** Confusion matrix of the presented technique using DenseNet-41. (**a**) Figshare Brain Tumor Dataset, (**b**) Brain MRI Dataset.

**Table 2.** Performance comparison of our technique with other RCNN approaches.

| Method | Evaluation Metrics | | | | |
|---|---|---|---|---|---|
| | Accuracy | mAP | Dice | Sensitivity | Time(s) |
| RCNN [51] | 0.920 | 0.910 | 0.870 | 0.950 | 0.47 |
| Faster RCNN [56] | 0.940 | 0.940 | 0.910 | 0.940 | 0.25 |
| Proposed (Resnet-50) | 0.959 | 0.946 | 0.955 | 0.953 | 0.20 |
| Proposed (Densenet-41) | 0.963 | 0.949 | 0.959 | 0.953 | 0.20 |

### 4.4.2. Comparison with Other Segmentation Techniques

In this section, we compare our proposed model with the other segmentation techniques using the Figshare brain tumor dataset [38] which is one of the largest online available datasets. For performance evaluation, we compared the average highest results of our presented technique with the average results reported in these studies [61–65]. For the presented technique, we have shown the results for the DenseNet-41-based Mask-RCNN framework as we obtained better performance on it as compared to the ResNet-50 framework.

Table 3 shows a quantitative comparison using different performance measurement metrics such as mean IoU, dice score, and accuracy. Sheela et al. [61] proposed a pixel-

based radius contraction and expansion (RCE) technique that uses an active contour model and fuzzy c-means for tumor segmentation. However, the performance depends on the threshold value set for extracting the region of interest. They obtained an average accuracy of 0.91. In [62], the authors proposed a cascaded dual scale LinkNet (CDSL Net), an end-to-end encoder–decoder-based architecture that uses multi-scale inputs for feature concatenation with corresponding layers in the network to perform brain tumor segmentation. They obtained a dice score of 0.8003. Gunasekara et al. [63] employed a faster RCNN model for tumor classification and then a Chan–Vese active contour algorithm for segmentation. They achieved an average dice score of 0.92 and an accuracy of 92.31 for tumor segmentation and classification, respectively. However, the authors considered only the axial MR images of glioma and meningioma brain tumors. In [64], the authors proposed a multi-scale CNN model that processes the input MR image in three different spatial scales using multiple processing pathways. They achieved an average accuracy of 0.973 and a dice score of 0.828 for classification and segmentation, respectively. In [65], we proposed a Mask-RCNN model with ResNet101 as a backbone for brain tumor detection and segmentation only. Our method in [65] achieved a dice score of 0.950; however, this method is computationally expensive due to the large number of parameters.

**Table 3.** Comparison of the presented method with other segmentation techniques.

| Technique | Segmentation Method | Evaluation Metrics | | |
|---|---|---|---|---|
| | | Mean IoU | Dice | Accuracy |
| Sobhaninia et al. [62] | Cascaded CNN | 0.907 | 0.800 | - |
| Gunasekara et al. [63] | Faster RCNN and ChanVese active contour | - | 0.920 | 94.6 |
| Sheela et al. [61] | Active Contour and Fuzzy-C-Means | - | 0.665 | 91.0 |
| Díaz-Pernas et al. [64] | Multi-scale CNN | - | 0.828 | - |
| Masood et al. [65] | Traditional Mask-RCNN | 0.950 | 0.950 | 95.1 |
| Proposed method | Mask-RCNN (ResNet-50) | 0.951 | 0.955 | 95.9 |
| | Mask-RCNN(DenseNet-41) | 0.957 | 0.959 | 96.3 |

As can be seen in Table 3, the proposed work does better than the other state-of-the-art techniques by achieving a dice sore of 0.959, a mean IoU of 0.957, and an overall average accuracy of 96.3% using DenseNet-41 for three different kinds of tumors. The proposed technique uses deep features that are more discriminating, reliable, and provide a more effective representation of tumor regions over other methods such as [61], which employs the hand-crafted features and is unable to better represent the tumor region due to structural complexities. Moreover, in some existing methods [62,64] segmentation is applied directly to the entire image, which results in misclassification due to the complex background (i.e., brain tissues overlapping with tumor boundary, MRI artifacts, etc.), which thus reduces the accuracy of the segmentation. The method in [63] employs a region-based method for tumor localization and requires further processing for tumor segmentation. Unlike these methods, our model performs segmentation on the localized RoIs, which limits the space of segmentation and uses the RoIAlign layer, which ultimately improves the accuracy of the segmentation result. Moreover, our method achieves comparable performance with [65] while using less computational resources.

4.4.3. Comparison with Other Classification Techniques

In this section, we present the comparison of classification results of our approach with results obtained by previous works over the same dataset [38]. Table 4 shows the comparison of tumor classification results with existing approaches in terms of average accuracy. In Table 4, we have presented the highest results of our proposed framework obtained using the DenseNet-41 backbone. In [66], the authors employed a pre-trained GoogleNet model by using transfer learning for feature extraction. The obtained features are classified by using three different classifiers: Softmax, SVM, and KNN, and achieved an accuracy of 97.1%. Swati et al. [67] utilized various DL models: AlexNet, VGG16, and

VGG19. They fine-tuned VGG16 and VGG19 in a block-wise manner and AlexNet using the traditional layer-wise approach. The results showed VGG19 achieved better performance, with an average accuracy of 94.82 for brain tumor classification. Huang et al. [68] proposed a deep CNN model with a modified activation function for brain tumor classification. The CNN model is constructed automatically by a network generator based on three different graph generation algorithms. The activation function is composed of Gaussian error linear units (GeLUs) and rectified linear units (ReLUs). The method achieved an accuracy of 95.49. However, the proposed approach is computationally complex due to network size. In [69], the authors proposed a hybrid feature extraction approach by using a PCA-based normalized GIST descriptor with a regularized extreme learning machine classifier. They obtained an overall accuracy of 94.93%. However, these approaches [66,69] require preprocessed input images, i.e., the normalized pixel values from 0 to 1. In [70], a DL model named BrainMRNet employed three different processing methods such as attention modules, the hypercolumn technique, and residual blocks for brain tumor classification. Besides, the segmentation of the brain using the Otsu method is also performed to determine the lobe region (i.e., left, or right) with more concentrated tumorous cells. They obtained an accuracy of 97.69% for classification.

**Table 4.** Comparison of our method with other classification techniques.

| Technique | Classification Method | Acc (%) |
|---|---|---|
| Deepak et al. [66] | GoogLeNet and SVM | 97.10 |
| Swati et al. [67] | VGG19 | 94.82 |
| Huang et al. [68] | CCN based on complex networks | 95.49 |
| Gumaei et al. [69] | GIST descriptor and ELM | 94.93 |
| BrainMRNet [70] | Attention module, Hypercolumn technique, and Residual block | 97.69 |
| Proposed method | Custom Mask-RCNN | 98.34 |

From Table 4, it can be seen that in comparison to existing techniques, the proposed approach shows an improved overall accuracy of 98.34% for brain tumor type classification. The above-mentioned approaches [66–68] extract features from the whole image which may result in the misclassification of tumor type due to the complex nature of the tumor, i.e., overlapping boundaries and MRI artifacts. In [69], hand-crafted features are employed that are less discriminative and robust. The technique in [70] achieves results comparable to our approach; however, due to the hypercolumn technique, the model leads to overfitting and is computationally expensive, whereas the proposed approach employs deep features that are more discriminative and reliable. Moreover, region-based CNN first localizes the tumor region (RoI) and performs classification that results in improved accuracy.

## 5. Conclusions

In this work, we introduced a DL technique, namely Mask-RCNN with two backbones, ResNet-50 and DenseNet-41, for precise and automated segmentation of brain tumor regions from MRI images. We obtained better segmentation and classification results for DenseNet-41 based Mask-RCNN as compared to the ResNet-50 network, due to its dense connections which result in more robust image feature calculations. Comparative experimental results show that our proposed method more precisely delineates the tumor region and can serve as a new automated tool for diagnostic purposes. Moreover, as compared to state-of-the-art models, our Custom Mask-RCNN can compute deep features with effective representations of brain tumors. In future, we aim to perform classification along with segmentation of brain tumors using more challenging datasets. We also plan to evaluate the robustness of our Custom Mask-RCNN for other medical image analyses applications such as eye disease detection, finger skin recognition, skin cancer, and COVID detection. Furthermore, we aim to increase training samples and optimize hyper-parameters to further improve the accuracy of the model.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing not applicable to this article as authors have used publicly available datasets, whose details are included in the "experimental results and discussions" section of this article. Please contact the authors for further requests.

**Conflicts of Interest:** The authors declare no conflict of interest.

# References

1.  DeAngelis, L.M. Brain tumors. *N. Engl. J. Med.* **2001**, *344*, 114–123. [CrossRef] [PubMed]
2.  Sultan, H.H.; Salem, N.M.; Al-Atabany, W. Multi-classification of Brain Tumor Images using Deep Neural Network. *IEEE Access* **2019**, *7*, 69215–69225. [CrossRef]
3.  Behin, A.; Hoang-Xuan, K.; Carpentier, A.F.; Delattre, J.-Y. Primary brain tumours in adults. *Lancet* **2003**, *361*, 323–331. [CrossRef]
4.  Akil, M.; Saouli, R.; Kachouri, R. Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy. *Med. Image Anal.* **2020**, *63*, 101692.
5.  Maharjan, S.; Alsadoon, A.; Prasad, P.; Al-Dalain, T.; Alsadoon, O.H. A novel enhanced softmax loss function for brain tumour detection using deep learning. *J. Neurosci. Methods* **2020**, *330*, 108520. [CrossRef] [PubMed]
6.  Smoll, N.R.; Schaller, K.; Gautschi, O.P. Long-term survival of patients with glioblastoma multiforme (GBM). *J. Clin. Neurosci.* **2013**, *20*, 670–675. [CrossRef] [PubMed]
7.  Louis, D.N.; Perry, A.; Reifenberger, G.; Von Deimling, A.; Figarella-Branger, D.; Cavenee, W.K.; Ohgaki, H.; Wiestler, O.D.; Kleihues, P.; Ellison, D.W. The 2016 World Health Organization classification of tumors of the central nervous system: A summary. *Acta Neuropathol.* **2016**, *131*, 803–820. [CrossRef] [PubMed]
8.  Nelson, P.B.; Robinson, A.G.; Martinez, J.A. Metastatic tumor of the pituitary gland. *Neurosurgery* **1987**, *21*, 941–944. [CrossRef] [PubMed]
9.  Komninos, J.; Vlassopoulou, V.; Protopapa, D.; Korfias, S.; Kontogeorgos, G.; Sakas, D.E.; Thalassinos, N.C. Tumors metastatic to the pituitary gland: Case report and literature review. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 574–580. [CrossRef]
10. Ullah, M.N.; Park, Y.; Kim, G.B.; Kim, C.; Park, C.; Choi, H.; Yeom, J.-Y. Simultaneous Acquisition of Ultrasound and Gamma Signals with a Single-Channel Readout. *Sensors* **2021**, *21*, 1048. [CrossRef]
11. Bauer, S.; Wiest, R.; Nolte, L.-P.; Reyes, M. A survey of MRI-based medical image analysis for brain tumor studies. *Phys. Med. Biol.* **2013**, *58*, R97. [CrossRef] [PubMed]
12. Olabarriaga, S.D.; Smeulders, A.W. Interaction in the segmentation of medical images: A survey. *Med. Image Anal.* **2001**, *5*, 127–142. [CrossRef]
13. Asa, S.L. *Tumors of the Pituitary Gland*; Amer Registry of Pathology: Washington, DC, USA, 1998.
14. Işın, A.; Direkoğlu, C.; Şah, M. Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Comput. Sci.* **2016**, *102*, 317–324. [CrossRef]
15. Goetz, M.; Weber, C.; Binczyk, F.; Polanska, J.; Tarnawski, R.; Bobek-Billewicz, B.; Koethe, U.; Kleesiek, J.; Stieltjes, B.; Maier-Hein, K.H. DALSA: Domain adaptation for supervised learning from sparsely annotated MR images. *IEEE Trans. Med. Imaging* **2015**, *35*, 184–196. [CrossRef] [PubMed]
16. Yao, J. Image processing in tumor imaging. In *New Techniques in Oncologic Imaging*; Routledge: Abingdon, UK, 2006; pp. 79–102.
17. Ding, Y.; Zhang, C.; Lan, T.; Qin, Z.; Zhang, X.; Wang, W. Classification of Alzheimer's disease based on the combination of morphometric feature and texture feature. In Proceedings of the 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, USA, 9–12 November 2015; pp. 409–412.
18. Ding, Y.; Dong, R.; Lan, T.; Li, X.; Shen, G.; Chen, H.; Qin, Z. Multi-modal brain tumor image segmentation based on SDAE. *Int. J. Imaging Syst. Tech.* **2018**, *28*, 38–47. [CrossRef]
19. Bauer, S.; Nolte, L.-P.; Reyes, M. Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Toronto, ON, Canada, 18–22 September 2011; pp. 354–361.

20. Tustison, N.J.; Shrinidhi, K.; Wintermark, M.; Durst, C.R.; Kandel, B.M.; Gee, J.C.; Grossman, M.C.; Avants, B.B. Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. *Neuroinformatics* **2015**, *13*, 209–225. [CrossRef]

21. Zikic, D.; Glocker, B.; Konukoglu, E.; Criminisi, A.; Demiralp, C.; Shotton, J.; Thomas, O.M.; Das, T.; Jena, R.; Price, S.J. Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nice, France, 1–5 October 2012; pp. 369–376.

22. Kaya, I.E.; Pehlivanlı, A.Ç.; Sekizkardeş, E.G.; Ibrikci, T. PCA based clustering for brain tumor segmentation of T1w MRI images. *Comput. Meth. Prog. Bio.* **2017**, *140*, 19–28. [CrossRef]

23. Hooda, H.; Verma, O.P.; Singhal, T. Brain tumor segmentation: A performance analysis using K-Means, Fuzzy C-Means and Region growing algorithm. In Proceedings of the 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, India, 8–10 May 2014; pp. 1621–1626.

24. Khalid, S.; Khalil, T.; Nasreen, S. A survey of feature selection and feature extraction techniques in machine learning. In Proceedings of the 2014 Science and Information Conference, London, UK, 27–29 August 2014; pp. 372–378.

25. Riaz, H.; Park, J.; Choi, H.; Kim, H.; Kim, J. Deep and Densely Connected Networks for Classification of Diabetic Retinopathy. *Diagnostics* **2020**, *10*, 24. [CrossRef]

26. Mehmood, A.; Iqbal, M.; Mehmood, Z.; Irtaza, A.; Nawaz, M.; Nazir, T.; Masood, M. Prediction of Heart Disease Using Deep Convolutional Neural Networks. *Arab. J. Sci. Eng.* **2021**, *46*, 3409–3422. [CrossRef]

27. Nazir, T.; Irtaza, A.; Javed, A.; Malik, H.; Hussain, D.; Naqvi, R.A. Retinal Image Analysis for Diabetes-Based Eye Disease Detection Using Deep Learning. *Appl. Sci.* **2020**, *10*, 6185. [CrossRef]

28. Hu, K.; Gan, Q.; Zhang, Y.; Deng, S.; Xiao, F.; Huang, W.; Cao, C.; Gao, X. Brain Tumor Segmentation Using Multi-Cascaded Convolutional Neural Networks and Conditional Random Field. *IEEE Access* **2019**, *7*, 92615–92629. [CrossRef]

29. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef] [PubMed]

30. Pereira, S.; Pinto, A.; Alves, V.; Silva, C.A. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1240–1251. [CrossRef]

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

32. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **2017**, *36*, 61–78. [CrossRef]

33. Zhao, X.; Wu, Y.; Song, G.; Li, Z.; Zhang, Y.; Fan, Y. A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. *Med. Image Anal.* **2018**, *43*, 98–111. [CrossRef]

34. Dong, H.; Yang, G.; Liu, F.; Mo, Y.; Guo, Y. Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In Proceedings of the Annual Conference on Medical Image Understanding and Analysis, Edinburgh, UK, 11–13 July 2017; pp. 506–517.

35. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

36. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.

37. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R.B. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.

38. Cheng, J. Brain tumor dataset. Available online: https://figshare.com/articles/brain_tumor_dataset/1512427 (accessed on 21 December 2020).

39. Brain MRI Images for Brain Tumor Detection. Available online: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection (accessed on 1 April 2021).

40. Jeong, J.; Lei, Y.; Kahn, S.; Liu, T.; Curran, W.J.; Shu, H.-K.; Mao, H.; Yang, X. Brain tumor segmentation using 3D Mask R-CNN for dynamic susceptibility contrast enhanced perfusion imaging. *Phys. Med. Biol.* **2020**, *65*, 185009. [CrossRef]

41. Anantharaman, R.; Velazquez, M.; Lee, Y. Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 2197–2204.

42. Chiao, J.-Y.; Chen, K.-Y.; Liao, K.Y.-K.; Hsieh, P.-H.; Zhang, G.; Huang, T.-C. Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine* **2019**, *98*, e15200. [CrossRef]

43. Kopelowitz, E.; Englehard, G. Lung Nodules Detection and Segmentation using 3D Mask-RCNN. *arXiv preprint* **2019**, arXiv:1907.07676.

44. Ismael, S.A.A.; Mohammed, A.; Hefny, H. An enhanced deep learning approach for brain cancer MRI images classification using residual networks. *Artif. Intell. Med.* **2020**, *102*, 101779. [CrossRef]

45. Kleesiek, J.; Biller, A.; Urban, G.; Kothe, U.; Bendszus, M.; Hamprecht, F. Ilastik for multi-modal brain tumor segmentation. In Proceedings of the MICCAI BraTS (Brain Tumor Segmentation Challenge), Boston, MA, USA, 14 September 2014; pp. 12–17.

46. Kamnitsas, K.; Ferrante, E.; Parisot, S.; Ledig, C.; Nori, A.V.; Criminisi, A.; Rueckert, D.; Glocker, B. DeepMedic for brain tumor segmentation. In Proceedings of the International workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Athens, Greece, 17 October 2016; pp. 138–149.

47. Zhang, W.; Li, R.; Deng, H.; Wang, L.; Lin, W.; Ji, S.; Shen, D. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage* **2015**, *108*, 214–224. [CrossRef] [PubMed]

48. Feng, X.; Tustison, N.; Meyer, C. Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features. In Proceedings of the International MICCAI Brainlesion Workshop, Granada, Spain, 16 September 2018; pp. 279–288.

49. Qasem, S.N.; Nazar, A.; Attia Qamar, S. A Learning Based Brain Tumor Detection System. *Comput. Mater. Contin.* **2019**, *59*, 713–727. [CrossRef]

50. Rehman, A.; Naz, S.; Naseem, U.; Razzak, I.; Hameed, I.A. Deep AutoEncoder-Decoder Framework for Semantic Segmentation of Brain Tumor. *Aust. J. Intell. Inf. Process. Syst.* **2019**, *15*, 53.

51. Rayhan, F. FR-MRInet: A Deep Convolutional Encoder-Decoder for Brain Tumor Segmentation with Relu-RGB and Sliding-window. *Int. J. Comput. Appl.* **2018**, *975*, 8887.

52. Gispert, J.D.; Reig, S.; Pascau, J.; Vaquero, J.J.; García-Barreno, P.; Desco, M. Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error. *Hum. Brain Mapp.* **2004**, *22*, 133–144. [CrossRef]

53. Zhan, T.; Zhang, J.; Xiao, L.; Chen, Y.; Wei, Z. An improved variational level set method for MR image segmentation and bias field correction. *Magn. Reson. Imaging* **2013**, *31*, 439–447. [CrossRef]

54. Hwang, H.; Haddad, R.A. Adaptive median filters: New algorithms and results. *IEEE Trans. Image Process.* **1995**, *4*, 499–502. [CrossRef]

55. Dutta, A.; Gupta, A.; Zisserman, A. Vgg Image Annotator (VIA). Available online: http://www.robots.ox.ac.uk/vgg/software/via (accessed on 12 December 2020).

56. Ren, S.; He, K.; Girshick, R.B.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef]

57. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

58. Albahli, S.; Nazir, T.; Irtaza, A.; Javed, A. Recognition and Detection of Diabetic Retinopathy Using Densenet-65 Based Faster-RCNN. *Comput. Mater. Contin.* **2021**, *67*, 1333–1351. [CrossRef]

59. Fedorov, A.; Beichel, R.; Kalpathy-Cramer, J.; Finet, J.; Fillion-Robin, J.-C.; Pujol, S.; Bauer, C.; Jennings, D.; Fennessy, F.; Sonka, M. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **2012**, *30*, 1323–1341. [CrossRef]

60. Mask-RCNN. Available online: https://github.com/matterport/Mask_RCNN (accessed on 12 December 2020).

61. Sheela, C.J.J.; Suganthi, G. Brain tumor segmentation with radius contraction and expansion based initial contour detection for active contour model. *Multimed. Tools. Appl.* **2020**, *79*, 23793–23819. [CrossRef]

62. Sobhaninia, Z.; Rezaei, S.; Karimi, N.; Emami, A.; Samavi, S. Brain Tumor Segmentation by Cascaded Deep Neural Networks Using Multiple Image Scales. In Proceedings of the 2020 28th Iranian Conference on Electrical Engineering (ICEE), Tabriz, Iran, 4–6 August 2020; pp. 1–4.

63. Gunasekara, S.R.; Kaldera, H.; Dissanayake, M.B. A Systematic Approach for MRI Brain Tumor Localization and Segmentation Using Deep Learning and Active Contouring. *J. Healthc. Eng.* **2021**, *2021*. [CrossRef]

64. Díaz-Pernas, F.J.; Martínez-Zarzuela, M.; Antón-Rodríguez, M.; González-Ortega, D. A Deep Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network. *Healthcare* **2021**, *9*, 153. [CrossRef]

65. Masood, M.; Nazir, T.; Nawaz, M.; Javed, A.; Iqbal, M.; Mehmood, A. Brain Tumor Localization and Segmentation using Mask RCNN. *Front. Comput. Sci.* **2020**. (accepted). Preprint version. Available online: https://journal.hep.com.cn/fcs/EN/article/downloadArticleFile.do?attachType=PDF&id=28181 (accessed on 20 April 2021). [CrossRef]

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

67. Swati, Z.N.K.; Zhao, Q.; Kabir, M.; Ali, F.; Ali, Z.; Ahmed, S.; Lu, J. Brain tumor classification for MR images using transfer learning and fine-tuning. *Comput. Med. Imaging Graph.* **2019**, *75*, 34–46. [CrossRef]

68. Huang, Z.; Du, X.; Chen, L.; Li, Y.; Liu, M.; Chou, Y.; Jin, L. Convolutional Neural Network Based on Complex Networks for Brain Tumor Image Classification with a Modified Activation Function. *IEEE Access* **2020**, *8*, 89281–89290. [CrossRef]

69. Gumaei, A.; Hassan, M.M.; Hassan, M.R.; Alelaiwi, A.; Fortino, G. A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. *IEEE Access* **2019**, *7*, 36266–36273. [CrossRef]

70. Toğaçar, M.; Ergen, B.; Cömert, Z. Tumor type detection in brain MR images of the deep model developed using hypercolumn technique, attention modules, and residual blocks. *Med. Biol. Eng. Comput.* **2021**, *59*, 57–70. [CrossRef]