

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/344302269>

Brain Tumor Localization and Segmentation using Mask RCNN

Article in *Frontiers of Computer Science (electronic)* · September 2020

DOI: 10.1007/s11704-020-0105-y

CITATIONS

47

READS

3,022

5 authors, including:



Momina Masood

35 PUBLICATIONS 2,288 CITATIONS

SEE PROFILE



Tahira Nazir

Riphah International University

61 PUBLICATIONS 2,306 CITATIONS

SEE PROFILE



Marriam Nawaz

University of Engineering and Technology Taxila

67 PUBLICATIONS 3,087 CITATIONS

SEE PROFILE



Ali Javed

University of Engineering and Technology Taxila

186 PUBLICATIONS 4,625 CITATIONS

SEE PROFILE

Brain Tumor Localization and Segmentation using Mask RCNN

Momina MASOOD¹, Tahira NAZIR¹, Marriam NAWAZ¹, Ali JAVED², Munwar IQBAL¹,
Awais MEHMOOD¹

¹ Department of Computer Science, University of Engineering and Technology, Taxila 47050, Pakistan

² Department of Software Engineering, University of Engineering and Technology, Taxila 47050, Pakistan

Front. Comput. Sci., **Just Accepted Manuscript** • 10.1007/s11704-020-0105-y

<http://journal.hep.com.cn> on September 2, 2020

© Higher Education Press 2020

Just Accepted

This is a “Just Accepted” manuscript, which has been examined by the peer-review process and has been accepted for publication. A “Just Accepted” manuscript is published online shortly after its acceptance, which is prior to technical editing and formatting and author proofing. Higher Education Press (HEP) provides “Just Accepted” as an optional and free service which allows authors to make their results available to the research community as soon as possible after acceptance. After a manuscript has been technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an Online First article. Please note that technical editing may introduce minor changes to the manuscript text and/or graphics which may affect the content, and all legal disclaimers that apply to the journal pertain. In no event shall HEP be held responsible for errors or consequences arising from the use of any information contained in these “Just Accepted” manuscripts. To cite this manuscript please use its Digital Object Identifier (DOI(r)), which is identical for all formats of publication.”

Brain Tumor Localization and Segmentation using Mask RCNN

Momina Masood¹, Tahira Nazir^{1*}, Marriam Nawaz¹, Ali Javed², Munwar Iqbal¹, Awais Mehmood¹

¹Department of Computer Science, University of Engineering and Technology, Taxila, 47050, Pakistan

²Department of Software Engineering, University of Engineering and Technology, Taxila, 47050, Pakistan

1. Introduction

Brain tumor is a lethal disease affecting millions of people around the globe and has a high mortality rate. Brain tumor represents a cluster of abnormal cells that grows and multiply uncontrollably within normal brain cells and causes to damage the nervous system that controls different functions in the human body. Early identification and segmentation of brain tumor helps to increase the survival chances of the patient and also saves them from complex surgical processes. Moreover, the precise segmentation of brain tumors facilitates the surgeon for better clinical development and cure. Manual detection and segmentation of brain tumors is slow and suffer from a high error rate. Thus, the demand for effective computer-aided brain tumor segmentation techniques has increased considerably in recent times. However, accurate brain tumor segmentation is still a challenge because of its structural complexity such as variations in location, size, shape, and overlapping tumor boundaries with normal brain tissues, etc. [1].

Existing automated approaches for brain tumor detection can be broadly categorized into handcrafted features and deep learning (DL) based approaches. Qasem et. al [2] used a watershed segmentation algorithm along-with the KNN for brain tumor classification and segmentation. This method performs well on the selected MRI images and unable to accurately segment the tumor regions on challenging images containing tumors with multiple structural complexities. In [3, 4] an encoder-decoder based architecture was proposed to perform pixel-wise segmentation of tumor tissues from the normal brain cells. In FR-MRINet [5], a 33 layer deep model and an encoder with a fully connected decoder was proposed for tumor segmentation. This method [5] provides better segmentation performance but at the expense of increased computational cost. Thus, there exists a need to develop effective brain tumor detection methods that are robust to brain tumors having multiple structural complexities.

In this letter, we propose an automated method to increase the robustness of brain tumor localization and segmentation by employing the Mask RCNN[6] model. Here are the contributions to our work:

- Accurate localization and segmentation of brain tumor using the region proposal network and pixel-to-pixel accurate segmentation.
- A rigorous quantitative and qualitative comparison of the presented and latest techniques was performed on two online available datasets to exhibit the efficacy of our method.
- To the best of our knowledge, it is the first time Mask-RCNN has been employed for brain tumor segmentation.
- The proposed technique is robust to variations in size, shape, location, and overlapping tumor boundaries with normal brain tissues in the presence of MRI artifacts such as noise, bias field effect, and different acquisition angles.

2. Proposed Method

2.1 Preprocessing

The MRI images produced from different MRI machines often include certain artifacts such as noise, bias field, or intensity inhomogeneity, which must be corrected to improve the segmentation results. In the preprocessing step, we applied the level set method for bias field correction and median filter to reduce the noise to get the enhanced image.

2.2 Tumor localization and segmentation using Mask RCNN

For segmentation, our target is to automatically localize and segment the brain tumor from a complex background without requiring any manual intervention. We aim to predict either tumor or non-tumor regions in the given MRI images using the Mask RCNN. The proposed framework (as shown in Fig. 1) consists of the following steps, 1) feature extraction, 2) region proposal network (RPN), 3) region of interest (ROI) classifier and bounding box regressor (BBR), and 4) segmentation. The detailed description of each step is discussed below.

2.2.1 Feature Extraction

The backbone network is utilized to obtain relevant features from the input image. For implementation, we considered ResNet101 with a Feature Pyramid Network (FPN)

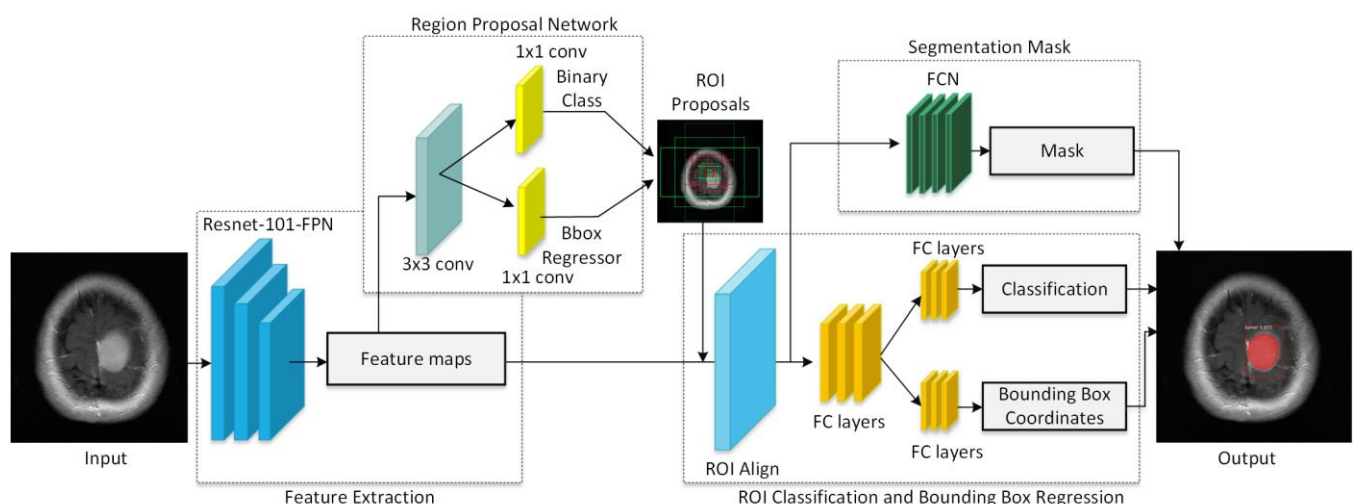


Fig. 1 Architecture of the Proposed Framework.

backbone to extract more discriminating and reliable features. Using ResNet-101 as the backbone provides a gain in both accuracy and processing time [6]. The initial layers extract low-level features i.e., corners and edges, while the deeper layers compute high-level features i.e., texture and color. The resulting feature map is further improved using the FPN that extracts the features with a better representation of the object at different scales for the RPN network.

2.2.2 Region Proposal Network

The feature map computed in the previous step is fed to the RPN network to generate ROIs. A 3×3 convolutional layer is used to scan the image using a sliding window to generate relevant anchors that represent the bounding box with different sizes and distributed over the entire image. There are about 20k anchors of distinct scales and sizes that correspond with each other to cover the image. Binary classification is performed to determine whether an anchor contains the object or background (FG/BG). The BBR generates bounding boxes according to set Intersection-over-Union (IoU) value. More specifically, if an anchor has IoU higher than 0.7 with a ground-truth (GT) box, it is classified as a positive anchor (FG class), otherwise negative. The RPN may produce regions that have significant overlap, the non-maximum suppression algorithm is applied to keep the regions with the highest foreground score while the remaining ones are discarded. The resulting region of interests are passed to the next stage for further classification.

2.2.3 ROI Classification and Bounding Box Regression

This network takes the proposed ROI and feature map as input (Fig. 1). Unlike the RPN, this network is deeper and classifies ROIs to a specific class such as tumor/non-tumor, and further improves the size of the bounding box. The BBR aims to refine the location and size of the bounding box to exactly encapsulate the tumor region. 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. To resize the feature maps, the ROIAlign layer is applied to extract fixed-length feature vectors for arbitrary-size candidate regions. The ROIAlign layer performs the bilinear interpolation to avoid misalignment issues encountered in the ROI pooling layer which uses quantization operation.

2.2.4 Segmentation Mask

The segmentation network takes positive ROI identified by the ROI classifier as input and returns a segmentation mask of 28×28 represented by floating numbers that contains more information over binary masks. The GT masks are scaled down to 28×28 to measure the loss with the predicted mask during the training stage. However, during the inference, the predicted mask is scaled up to match the dimensions of the ROI bounding box and which provides the final output mask.

2.3 Multi-Task loss

The proposed model employs a multi-task loss L on each sampled ROI defined as:

$$L(\text{Mask RCNN}) = L_{\text{class}} + L_{\text{loc}} + L_{\text{mask}} \quad (1)$$

where L_{class} , L_{loc} , and L_{mask} represent the box class labels prediction loss, bounding box refinement loss, and segmentation mask prediction loss respectively.

L_{class} is the log loss of the two classes(tumor/non-tumor), defined as:

$$L_{\text{class}}(t_i, t_i^*) = -\log[t_i t_i^* + (1 - t_i)(1 - t_i^*)] \quad (2)$$

where t_i is the target prediction probability that candidate anchor i contains a tumor and t_i^* is the GT label which is 1 for

the positive anchor, otherwise 0. The bounding-box regression loss function is shown in Eq. 3.

$$L_{\text{loc}}(v_i, v_i^*) = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L1}(v_i - v_i^*) \quad (3)$$

where,

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

Vector v_i 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- $L1$ function is a robust $L1$ loss that is less sensitive to outliers over $L2$ loss. For the training of mask network, the average binary cross-entropy loss is employed that is given as follows:

$$L_{\text{mask}} = -\frac{1}{n^2} \sum_{1 \leq i, j \leq n} [x_{ij} \log P_{ij}^k + (1 - x_{ij}) \log(1 - P_{ij}^k)] \quad (5)$$

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 ($k=1$ for tumor and 0 for non-tumor).

3. Experiments and Results

We have implemented the model using Keras and TensorFlow libraries with ResNet-101, and FPN for feature extraction. We initialized the model using pre-trained weights obtained from the COCO dataset and employed transfer learning to fine-tune the model on MRI datasets for tumor segmentation. For experimentation, we used the 70-30 ratio that is randomly splitted into training (70%) and test (30%) sets.

3.1 Dataset

The presented approach is evaluated on two online available datasets, Brain Tumor Figshare (BTF) Dataset [7] and Brain Tumor Kaggle (BTK) Dataset [8] that are diverse in terms of structural complexity, acquisition angle, devices, noise, and bias field-effect, etc. BTF dataset contains 3064 real MRI images of size 512×512 obtained from 233 subjects. BTK dataset contains 155 MRI images of size 845×845. Both datasets contain T1-weighted MRI images that are contrast-enhanced. The GT mask against each MRI image is created using the VGG Image Annotator tool. We employed the precision, recall, accuracy, dice score (DSC), and IoU to evaluate the proposed method.

3.2 Results and Discussion

This section provides a discussion of the results obtained after performing three different experiments.

In our first experiment, we analyzed the performance of our technique on the BTF dataset and BTK dataset. Fig. 2 shows some of the high-scoring results of the segmented brain tumor obtained by applying the Mask RCNN. The proposed method can accurately localize the brain tumor with an average precision of 0.952 on the BTF dataset and 0.948 on the BTK dataset from the healthy tissues despite discontinuous or blurry boundaries and artifacts in MR images such as noise, bias field effect, and acquisition angle. Moreover, our method can precisely segment the brain tumor by overcoming the challenges of variations in location, shape, and size.

To further understand the performance of our method, we have drawn a boxplot for evaluation metrics on both datasets (Fig. 3). The boxplot represents the spread of results into four quartiles, median, and an outlier. Our method has achieved an average accuracy of 95.1% over the BTF dataset and 94.6% over the BTK dataset. Our method fails to accurately

localize the tumor region in a few images, due to visual similarity with healthy tissues, as shown in Fig. 4.

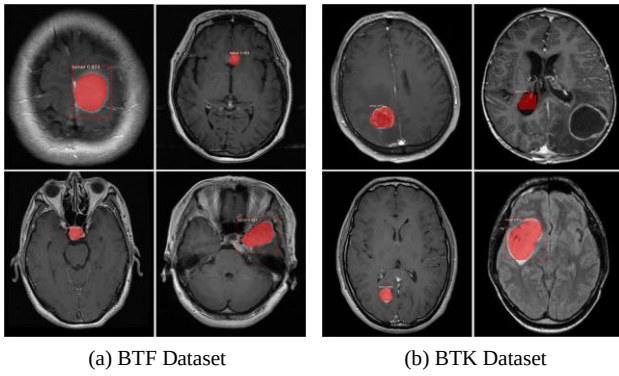


Fig. 2 Visualization of Tumor Segmentation Results in MRI Images

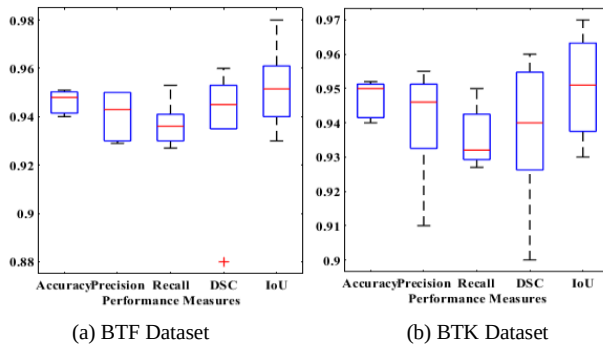


Fig. 3 Performance of the Proposed Method over BTF and BTK Datasets.

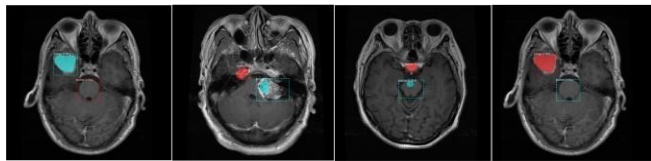


Fig. 4 Example Images of Inaccurately Localized Brain Tumor.

In our next experiment, we have compared the performance of our method with other region-based segmentation methods i.e., RCNN and Faster-RCNN using the BTF dataset, and results are reported in Table 1. The RCNN is computationally complex as it randomly generates region proposals (2000 per image) and uses a selective search algorithm for classification. The Faster RCNN automatically extracts the region proposals using the RPN and shares the convolutional layer among class and bounding box network to reduce the computational cost. The proposed model employs the Mask RCNN that offers an added advantage over Faster-RCNN by providing automated segmentation and localization of the brain tumor. Moreover, our model is easier to train and adds a very small overhead over Faster-RCNN.

Table 1 Comparison of the Proposed Approach with other RCNN Methods.

Methods	accuracy	mAP	dice	sensitivity
RCNN	0.92	0.91	0.87	0.95
Faster-RCNN	0.94	0.94	0.91	0.94
Proposed (Mask RCNN)	0.95	0.94	0.95	0.95

In our last experiment, we have compared the performance of our model with the existing latest approaches using the BTF dataset (Table 2). The proposed technique uses deep features that are more discriminating, reliable, and provide more effective representation of tumor regions over other methods such as [2], which employs the hand-crafted features and unable to better represent the tumor region due to structural complexities. Moreover, in some existing methods [3-5] segmentation is applied directly on the entire image, which results in misclassification due to complex background (i.e., brain tissues overlapping with tumor boundary, MRI artifacts,

etc.), thus reduces the accuracy of the 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 final segmentation result.

Table 2 Comparison of the Proposed Method with Latest Techniques.

Techniques	mean IoU	dice	accuracy
Sultan et. al [2]	-	-	86.0
Arshia et. al [3]	0.76	0.93	93.4
Sobhaninia et. al [4]	-	0.79	-
FR-MRInet [5]	0.94	-	91.4
Proposed method (Mask RCNN)	0.95	0.95	95.1

4. Conclusion

In this letter, we have introduced a Mask RCNN model for the precise segmentation of brain tumor from the MRI images. We showed the significance of Mask RCNN for brain tumor segmentation. Mask RCNN can compute deep features with effective representation of brain tumor regions over existing systems. The results illustrate that the proposed method precisely delineates the tumor region and serves as an effective automated tool for diagnostic purposes. We plan to extend our work by performing the classification of different classes of brain tumor.

Acknowledgements This work was supported and funded by the Directorate ASR&TD of UET-Taxila (UET/ASR&TD/RG-1002).

References

1. 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. *Medical Image Analysis*, 2020: 101692
2. Qasem S N, Nazar A, Attia Qamar S. A Learning Based Brain Tumor Detection System. *CMC-Computers Materials & Continua*, 2019, 59(3): 713-727
3. Rehman A, Naz S, Naseem U, Razzak I, Hameed I A. Deep AutoEncoder-Decoder Framework for Semantic Segmentation of Brain Tumor. *Australian Journal of Intelligent Information Processing Systems*, 2019, 15(4): 53
4. Sobhaninia Z, Rezaei S, Noroozi A, Ahmadi M, Zarrabi H, Karimi N, Emami A, Samavi S. Brain tumor segmentation using deep learning by type specific sorting of images. *arXiv preprint arXiv:1809.07786*, 2018
5. Rayhan F. FR-MRInet: A Deep Convolutional Encoder-Decoder for Brain Tumor Segmentation with Relu-RGB and Sliding-window. *International Journal of Computer Applications*, 2018, 975: 8887
6. He K, Gkioxari G, Dollár P, Girshick R B. Mask R-CNN. *Proceedings of the IEEE international conference on computer vision*, 2017: 2980-2988
7. Brain tumor dataset, https://figshare.com/articles/brain_tumor_dataset/1512427.
8. Brain MRI Images for Brain Tumor Detection, <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>.