Brain Tumor Detection and Segmentation Using Mask R-CNN Algorithm

Samoua Alsamoua, No:422705

Software Engineering Department, Karadeniz Teknik University, Trabzon, Turkey.
Email: samoua.alsamoua@gmail.com

ABSTRACT: brain tumor disease affecting millions of people around the world. If this disease is not diagnosed early, the survival rate of patients is very low. Therefore, the diagnosis of brain tumor must be fast and accurate. Mask-RCNN is a recently proposed state-of-the-art algorithm for object detection, object localization, and object sample segmentation of MRI images. In this study, using a dataset consisting of 1500 MR images with tumor and 1500 MR images without a tumor, and the location and size of the tumors were detected with Mask R-CNN, an accuracy rate of 88%-94% was obtained.

KEYWORDS: *Convolutional Neural Networks, Brain Tumor, Mask-RCNN, Tumor Detection.*

INTRODUCTION:

Computer vision tasks include methods for acquiring, processing, analyzing, and understanding digital images, and extraction of high-dimensional data from the real world [1]. Real-time computer vision can be performed using the open-source computer vision (OpenCV) programming library. OpenCV has vast application areas such as facial recognition systems, human-computer interaction, object identification, mobile robotics, motion tracking, and augmented reality. R-CNNs (Region Based Convolutional Neural Networks) are a kind of machine learning model used in computer vision, specifically object detection. As a result, Mask R-CNN is a natural and intuitive concept. However, the additional mask output differs from the class and box outputs, necessitating the extraction of a much finer spatial layout of an object. For each candidate item, Mask R-CNN produces two outputs: a class label and a bounding-box offset; to this, we add a third branch that produces the object mask [2]. As a result, Mask R-CNN is a natural and intuitive concept. However, the additional mask output differs from the class and box outputs, necessitating the extraction of a much finer spatial arrangement of an item. This is accomplished by adding a branch for anticipating an object mask alongside the existing branch for bounding box recognition as shown in Figure 1.

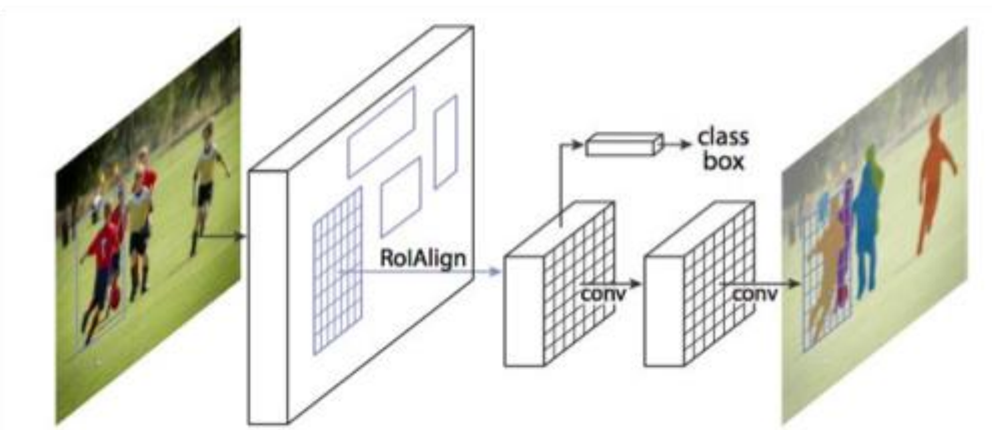

Figure 1. The General Structure of Mask R-CNN

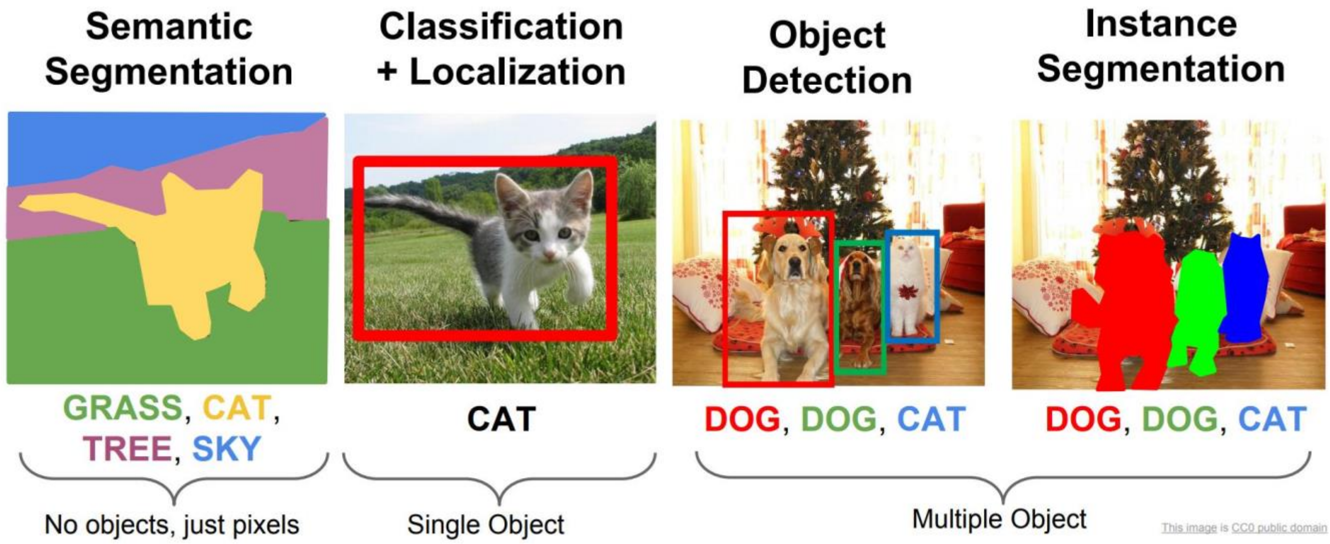

Figure 2. Types of Computer Vision Tasks

A Brain Tumor is an abnormal growth of brain tissue that can be life-threatening if not diagnosed early and treated appropriately. Typically, Magnetic resonance imaging (MRI) and Computed Tomography (CT) scans are used by medical personnel to obtain detailed images of the brain for initial analysis in invasive procedures such as tissue biopsies. In addition, the use of computer-based image analysis in collaboration with medical knowledge can contribute significantly to aiding early diagnosis [1]. Several different types of algorithms have been studied for many years in the fields of image classification and segmentation, image processing for supervised and unsupervised feature extraction, and computer vision. Therefore, an increasing number of existing and new computer-based image classification and segmentation algorithms are being implemented and validated by many researchers in this field of study [2][3]. Misdiagnosis of a brain tumor can lead to a serious problem and reduce the patient's chances of survival. To overcome the disadvantages of manual diagnostics, there is an increased interest in designing automated image processing systems. [4]

Related Work:

Mohammed et al. [8] compared the performance of K-Means Clustering and Fuzzy C-Mean Clustering techniques in three different data sets in tumor detection from MR images. They achieved an average accuracy of 92.27% with the K- Mean Clustering technique and 96.66% with the Fuzzy C-Mean Clustering technique.

Naghsh et al. [9] used the ROI method for tumor detection from MR brain images. The presented method consists of different image processing techniques such as morphological operations, low- pass filtering and thresholding. Tumor segmentation was performed with an average of 98.48% success from ten different data sets. Sajjad et al. In their study, tumor region was selected by segmentation from MRIs for classification of brain tumors. Then, the classification process was carried out with the proposed Convolutional Neural Network (CNN) model. The success in classification is 94.58% [7]. In the study by İbrahim et al., MR images were classified using CNN. Data from the CIPR database were used as training data. The size of each image used in the developed model is 3x58. With the results obtained, the classification accuracy has been demonstrated as 96.33% [8].

Ghosh et al using the fuzzy k-means clustering algorithm with MRI images of patients, classified different tumor types in the brain and other brain-related areas with 89.2% accuracy.[10]

Mohammad et al. proposed a method they call Caps Net to classify brain tumors. With the proposed method, they aimed for a higher accuracy score with more data. For this purpose, they used 64 features obtained from a single convolution layer. They achieved an accuracy rate of 86.56% in the\ classification of brain tumors with the method.[11]

Ergin et al. proposed a method for detecting brain tumors on MR images, including deep learning and K mean segmentation. As a result of the study, they detected the brain tumor with an accuracy rate of 84.45% and a sensitivity of 95.04%.[12]

Cheng et al. [13] proposed a method to improve brain tumor classification performance by magnifying the tumor region through image expansion and then subdividing it into subregions. They used three approaches to extract the features; The density histogram achieved the best accuracy of 91.28% using the Gray Level Co-occurrence Matrix (GLCM) and Bag of Words (BOW) and finally the ring form section in addition to tumor region magnification. In this article, an experimental study was carried out on the localization of tumors using the Mask R-CNN algorithm from the CNN Family, one of the supervised machine learning algorithms in computer science for the detection of tumors on MR images. The findings show that Mask R-CNN performs well in determining the location of the tumors.

Our goal in this work is to develop a comparably enabling framework for instance segmentation. First, let's review the Faster R-CNN. The Faster R-CNN consists of two phases: (i) The Region Proposal Network (RPN) and the basic Fast R-CNN model. (ii) RPN is employed in the generation of candidate regions.

THE PROPOSED METHOD:

Mask R-CNN is an object recognition algorithm. It was released in 2018 by Kaiming He et al. It is the fourth member of the R-CNN family. Mask R-CNN is a branch of the Faster R-CNN architecture to predict the mask of an object in addition to drawing a bounding box with the existing architecture of the Faster R-CNN algorithm. It differs from classical object detection models such as Mask R-CNN, Faster R-CNN. Here, in addition to defining the class and the bounding box location, the pixels in the bounding box corresponding to that class are also defined. This pixel information may be necessary, for example, to detect the path in autonomous vehicles and the objects they want to pick up in robots.

a. **Preprocessing:** 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.
b. **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 (Fig. 3).
c. **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) backbone to extract more discriminating and reliable features. 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 region proposal network (RPN).
d. **Region Proposal Network, RPN:** The feature map is obtained by passing the input image shown in Figure 3 through the convolution layers and this feature map is used as the input of the RPN. RPN, on the other hand, detects different sized identification frames at different locations, using the anchors of different sizes and aspect ratios, and detects areas that may be objects with the sliding window method. In this way, anchors that can contain objects are selected. If several anchors overlap too much, the one with the highest Intersection over Union (IoU) value is selected. This process is called non-max suppression. Thanks to this process, intersection regions (Region of Interest or called RoI)

are obtained. 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 back- ground (FG/BG). The bounding box regression (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.

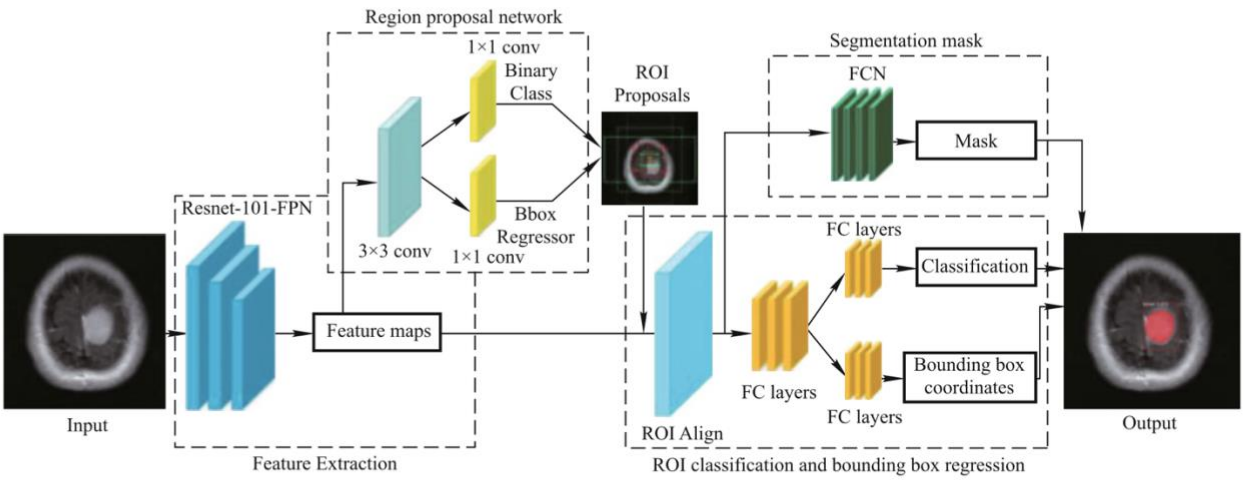

Figure 3. Architecture of the proposed framework

e. **ROI** This network takes the proposed ROI and feature map as input (Figure 3). Unlike RPN, this network is deeper and categorizes ROIs according to a specific class such as tumor/non-tumor, further improving the size of the bounding box. BBR aims to improve the position and size of the bounding box to fully encapsulate the tumor site. Because the feature map is down sampled k times the size of the original image, usually the boundaries of the ROI do not coincide with the detail level of the feature map. To resize feature maps, the RoIAlign layer is applied to extract fixed-length feature vectors for candidate regions of arbitrary size.

f. **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 in- formation 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.

g. **Mask R-CNN Loss Function** it is calculated as follows:

$$LOSS = L\text{class} + L\text{loc} + L\text{mask} \qquad (1)$$

where Lclass is the loss function of each classification result, $Lreg$ is the loss function of the regression process used to determine the bounding box, showing the sum of the squares of the difference between the binary image and the label defined in each image, Lmask corresponds to the loss function of the segmentation result. We can select the functions as follow:

$Lclass$ is the log loss of the two classes(tumor/non-tumor), defined as:

$$Lclass(ti, t_i^*) = -\log[ti t_i^* + (1 - ti)(1 - t_i^*)], \qquad (2)$$

where t_i is the target prediction probability that candidate anchor i contains a tumor and t_i^* is the ground-truth GT label, which is 1 for the positive anchor, otherwise 0. The bounding-box regression loss function is shown in Eq. (3).

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

where

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

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:

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

where x_{ij} represents the value of a pixel (i, j) in a GT mask of size $n \times n$ and P^k_{ij} is the predicted value of the same pixel in the mask learned for class k ($k = 1$ for tumor and 0 for non-tumor).

Dataset: The dataset is taken from the **Kaggle site** and consists of MR images of tumor patients, and it constitutes a total of 1500 MR images with tumor and 1500 MR images without a tumor. As the suggested method, the Mask R- CNN algorithm is suitable for the data set, and the algorithm tries to locate the tumor by scanning pixels on the images, and in case of correct detection of the tumor, the location of the tumor is determined in a red area. The presented approach is evaluated on two online available datasets, Brain Tumor Fig share (BTF) Dataset [6] and Brain Tumor Kaggle (BTK) Dataset [7] that are diverse in terms of structural complexity, acquisition angle, devices, noise, and bias field-effect, etc.

As we said before applying mask R-CNN to MR images, a data preprocessing step was performed to accurately predict the location of the tumors on these images. In this way, the data was prepared for RPN. If the presence of the object is detected, it takes the accuracy value of 1 and switches to the regression layer in the visual RPN. In the regression layer, a bounding box is drawn around the object. The next step is to switch to the RoIAlign layer. In the RoIAlign layer, each object is additionally masked according to its spatial arrangement. While using RESNET101 as the basic architecture, a pre-trained model with the **Kaggle** dataset is used for training.

For the Mask R-CNN algorithm to produce correct results, it is very important to determine the epoch number correctly. where the epoch number denotes the cumulative iteration number of the training process applied on the dataset. In the proposed method, the number of epochs is used as 15, and the accuracy rate is between 88% and 94%.

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 **Kaggle** 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.

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. 4 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, the method can precisely segment the brain tumor by overcoming the challenges of variations in location, shape, and size.

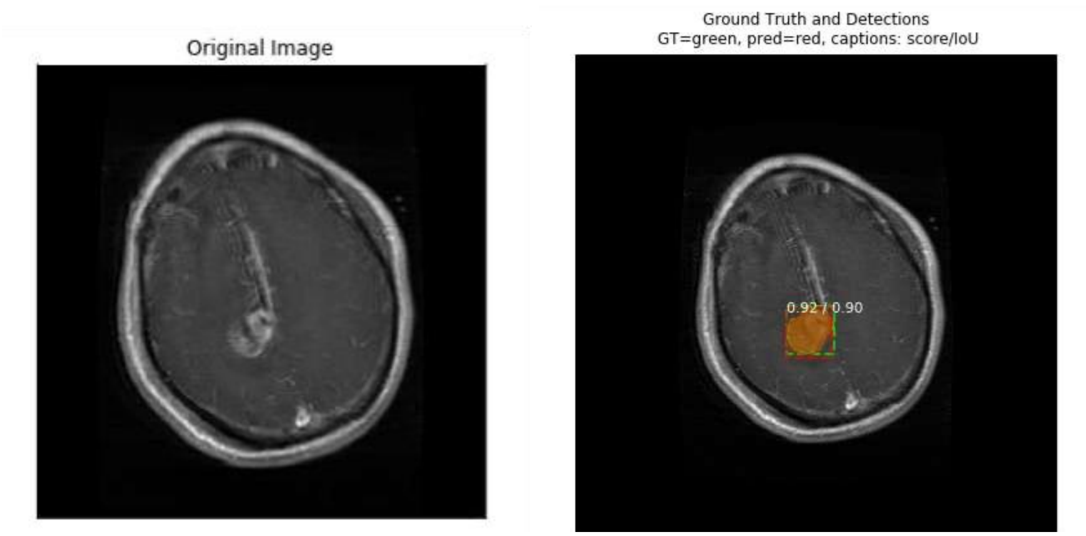

Figure 4. Tumor detected by Mask R-CNN

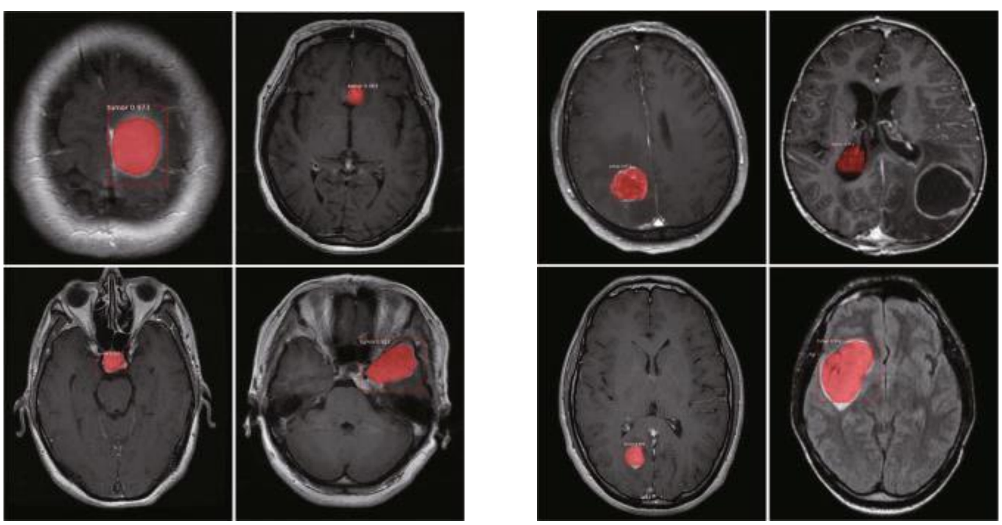

Figure 5. More results of Tumor detected by Mask R-CNN

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. The results illustrate that the proposed method precisely delineates the tumor region and serves as an effective automated tool for diagnostic purposes.

References

1. Qasem S N, Nazar A, Attia Qamar S. A learning based brain tumor detec- tion system. CMC-Computers Materials & Continua, 2019, 59(3):713– 727
2. 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
3. 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. 2018, arXiv preprint arXiv:1809.07786
4. 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
5. He K, Gkioxari G, Dollár P, Girshick R B. Mask R-CNN. In: Proceed- ings of the IEEE International Conference on Computer Vision.. 2017, 2980–2988
6. Cheng J, Huang W, Cao S, Yang R, Yang W, Yun Z, Wang Z, Feng Q. En- hanced performance of brain tumor classification via tumor region aug- mentation and partition. PLoS ONE, 2015, 10(10): e0140381
7. Çinar A, Yıldırım M. Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Medical Hypotheses,2020, 139: 109684
8. Mohammed, E., Eman, A.M., Rashid A.A., Brain tumor segmentation based on a hybrid clustering technique, Egyptian Informatics Journal, 16 (1), 71–81, 2015.
9. Naghsh, A.,Abu-Bakar, Mousavi, S.M. S.A.R., A heuristic automatic and robust ROI detection method for medical image warermarking, Journal of Digital Imaging, 28 (4), 417–427,2015.
10. Ghosh, D., & Bandyopadhyay, S. K., Brain tumor detection from MRI image: An approach. IJAR. 3 (6): 1152-1159, 2017.
11. Mohammadi A, Afshar P, Plataniotis K. N. Brain Tumor Type Classification Via Capsule Networks. In: 25th IEEE International Conference on Image Processing (ICIP);7-10 October 2018; Athens, Greece: IEEE. pp. 3129–3133.
12. Tas M O, Ergin S. Detection of the Brain Tumor Existance Using a Traditional Deep Learning Technique and Determination of Exact Tumor Locations Using K-Means Segmentation from MR Images. İleri Mühendislik Çalışmaları ve Teknolojileri Dergisi 2020;1: 91–97.
13. J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, „„Enhanced performance of brain tumor classification via tumor region augmentation and partition,"" PloS ONE, vol. 10, no. 10, Oct. 2015, Art. no. e0140381.