Turk Neurosurg, 2021

Original Investigation

DOI: 10.5137/1019-5149.JTN.29217-20.2

Received: 21.01.2020
Accepted: 20.07.2020

Published Online: 16.09.2021

Utilizing Deep Convolutional Generative Adversarial Networks for Automatic Segmentation of Gliomas: An Artificial Intelligence Study

Ebru AYDOGAN DUMAN[1], Seref SAGIROGLU[1], Pinar CELTIKCI[2], Mustafa Umut DEMIREZEN[3], Alp Ozgun BORCEK[4], Hakan EMMEZ[5], Emrah CELTIKCI[5]

[1]Gazi University Faculty of Engineering, Department of Computer Engineering, Ankara, Turkey
[2]Baskent University Faculty of Medicine, Department of Radiology, Ankara, Turkey
[3]The Digital Transformation Office of The Presidency of The Republic of Turkey, The Artificial Intelligence and Big Data Unit, Ankara, Turkey
[4]Gazi University Faculty of Medicine, Department of Neurosurgery, Division of Pediatric Neurosurgery, Ankara, Turkey
[5]Gazi University Faculty of Medicine, Department of Neurosurgery, Ankara, Turkey

Corresponding author: Emrah CELTIKCI ✉ drceltikci@gmail.com

ABSTRACT

AIM: To describe a deep convolutional generative adversarial networks (DCGAN) model which learns normal brain MRI from normal subjects than finds distortions such as a glioma from a test subject while performing a segmentation at the same time.

MATERIAL and METHODS: MRIs of 300 healthy subjects were employed as training set. Additionally, test data were consisting anonymized T2-weigted MRIs of 27 healthy subjects and 27 HGG patients. Consecutive axial T2-weigted MRI slices of every subject were extracted and resized to 364x448 pixel resolution. The generative model produced random normal synthetic images and used these images for calculating residual loss to measure visual similarity between input MRIs and generated MRIs.

RESULTS: The model correctly detected anomalies on 24 of 27 HGG patients' MRIs and marked them as abnormal. Besides, 25 of 27 healthy subjects' MRIs in the test dataset detected correctly as healthy MRI. The accuracy, precision, recall, and AUC were 0.907, 0.892, 0.923, and 0.907, respectively.

CONCLUSION: Our proposed model demonstrates acceptable results can be achieved only by training with normal subject MRIs via using DCGAN model. This model is unique because it learns only from normal MRIs and it is able to find any abnormality which is different than the normal pattern.

KEYWORDS: Artificial intelligence, Deep learning, Glioma, Machine learning, Segmentation

INTRODUCTION

Medical data which is produced continuously is unstructured. Conventional statistical methods remain inadequate for understanding intricate relationships. Knowledge gained from the evaluation of medical big data by artificial intelligence (AI) is growing exponentially and keeps changing paradigms (3,4,7,10,15,16,18). Algorithms are capable of perceive patterns including predicting molecular markers from magnetic resonance imaging (MRI) of gliomas which is not possible with human perception (1,6,9,12).

Analyzing the MRIs of gliomas via AI is multidirectional. Classifying normal and abnormal images would require no previous segmentation for training however, in most studies, automatic segmentation is acquired by performing segmentation in the training step. Segmentation by human intervention is the best way to gather a ground truth for the algorithm but it is also

Ebru AYDOGAN DUMAN : 0000-0001-8231-6022
Seref SAGIROGLU : 0000-0003-0805-5818
Pinar CELTIKCI : 0000-0002-1655-6957
Mustafa Umut DEMIREZEN : 0000-0002-9045-4238
Alp Ozgun BORCEK : 0000-0002-6222-382X
Hakan EMMEZ : 0000-0002-3290-179X
Emrah CELTIKCI : 0000-0001-5733-7542

time consuming. Need for automatic glioma segmentation became great clinical significance and there are several methods reported for this task (8,11,14,21,23,24).

In this paper we present automatic segmentation of gliomas using a deep convolutional generative adversarial networks (DCGAN) model from T2-weighted images of high-grade glioma (HGG) patients. This study is one of the pilot studies of the Turkish Brain Project (TBP) under the supervision of The Artificial Intelligence and Big Data Unit, The Digital Transformation Office of The Presidency of The Republic of Turkey.

MATERIAL and METHODS

Data Acquisition

All data were acquired using a 3-Tesla Magnetom Verio® (Siemens, Erlangen, Germany). Fully anonymized T2-weighted MRIs of 300 healthy subjects were employed as training set. Additionally, test data were consisting anonymized T2-weigted MRIs of 27 healthy subjects and 27 HGG patients. Consecutive axial T2-weigted MRI slices of every subject were extracted and resized to 364x448 pixel resolution. Ethics committee approval and authorization of the Digital Transformation Office of The Presidency of The Republic of Turkey were obtained.

Unsupervised Generative Model

We propose a framework to train and test the Generative Model as shown in Figures 1, 2. Aim of this framework is training the model for randomly producing a normal synthetic T2-weighted image. The difference between the input and output can then be used to calculate anomaly score and to classify normal versus abnormal images. The normalization of images is common for both training and testing stages. During MR image acquisition, different environmental or patient specific conditions may result in intensity variations. For this reason, we applied intensity normalization process to the images before training step in our framework. The normalization method is using the Gaussian method. The

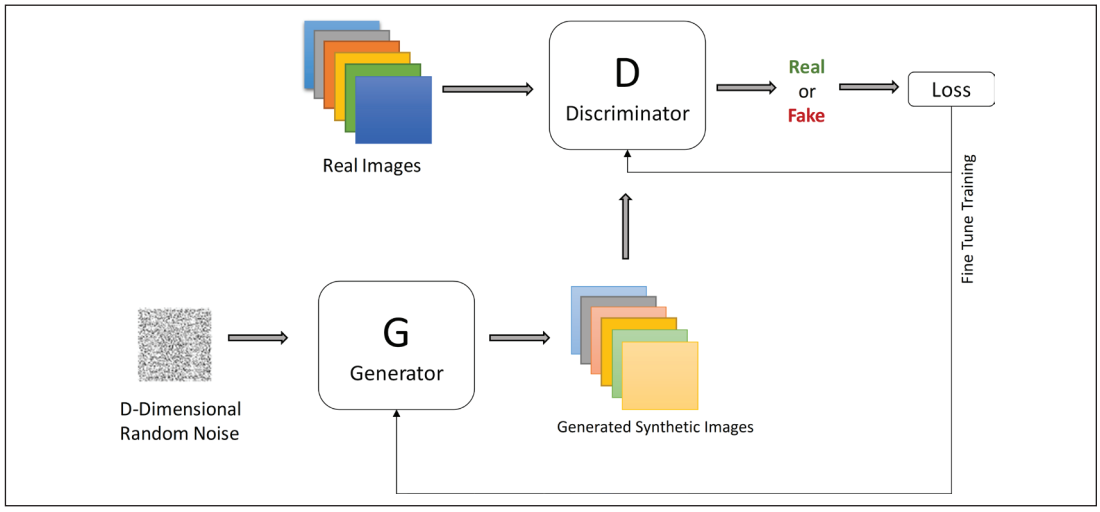

Figure 1: Diagram summarizing structure of Generative Adversarial Networks.

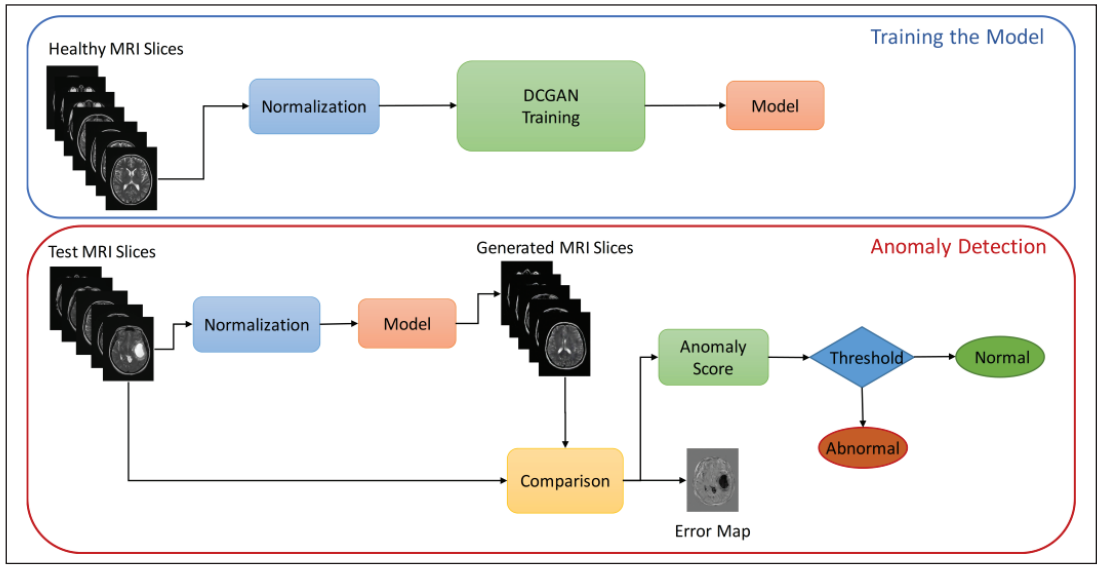

Figure 2: Diagram summarizing structure of proposed anomaly detection framework.

normalization method rescales the intensities by $I_{new} = I / SD$. Where I is the intensity and SD is the standard deviation of the whole image. The principle of the method is that each scan has the same intensity distribution.

We are given M MR images I_m and K image patches are extracted $x_{k,m} \varepsilon X$, during training phase model is trained with I_m to learn manifold X. For testing phase, a set of N MR images, only given to the system during testing, $I_n.L_n$ gets 0 or 1 and it is an array of binary image-wise ground-truth labels. The labels are used only during testing to accurately evaluate the model efficiency.

The generator is able to generate synthetic image z. In order to generate a realistic image, loss function is used in an iterative process with backpropagation steps. The residual loss is used to measure visual similarity between input MRI and generated MRI $G(z)$ and defined as: $L_r(z_\gamma) = \sum |x - G(z_\gamma)|$ Where x is the target image and $G(z_\gamma)$ is the generated image. In case of the assumption of generator performs without errors, image x and $G(z_\gamma)$ are the loss function is zero. The discriminator loss is defined as follow: $L_d(z_\gamma) = \sum \left| f(x) - f(G(z_\gamma)) \right|$ where f is a discriminator layer. Statistical closeness of target image f(x) and generated image $f(G(z_\gamma))$ in the latent space are calculated. During lesion detection in MR images, each patient MRI is evaluated as a normal or abnormal. In the method anomaly score can obtained as the equation: $A(x) = (1-\lambda) \cdot R(x) + \lambda \cdot D(x)$, R is residual score and D(x) is discrimination score mentioned above. The anomaly score values $A(x)$ are in the range of 0 to 1.

Performance Metrics

Brain anomaly detection problem is considered as a two-class classification problem and we consider the tumorous brain as positive and normal brain as negative. We used accuracy, precision, recall, Receiver Operating Characteristic (ROC) curve and area under the curve (AUC) metrics to quantitatively evaluate the detection performance of our model. Besides, confusion matrix is presented to compare the detection of tumorous and normal brain. The confusion matrix elements obtained for a two-class classifier are given in Table I.

RESULTS

The model trained for 150 epoch with MRI images of healthy subjects, using a learning rate 0.002. Training phase sample is shown in Figure 3. The first epochs are mostly blur until

Table I: Confusion Matrix for Two-Class Classifier

		Predicted Label	
		Negative	**Positive**
Actual Label	**Negative**	True Negative (TN): number of negative samples correctly predicted.	False Positive (FP): number of negative samples wrongly predicted as positive
	Positive	False Negative (FN): number of positive samples wrongly predicted as negative	True Positive (TP): number of positive samples correctly predicted

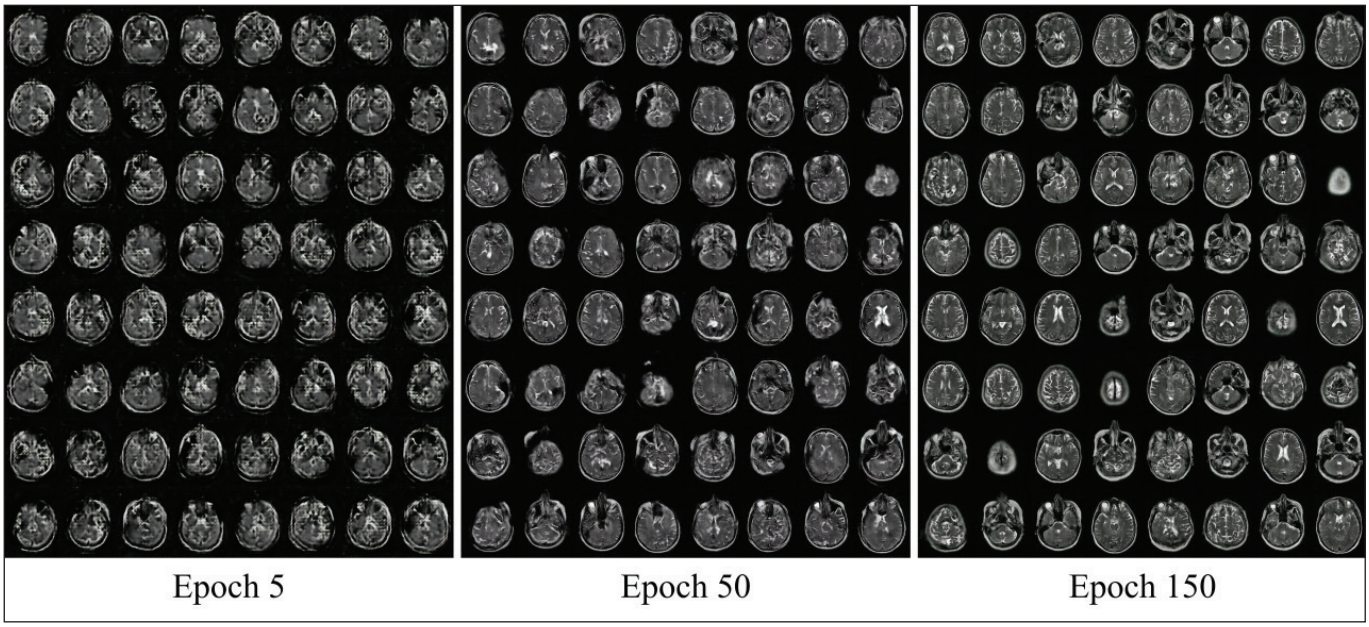

Figure 3: Samples from generator during training at different epochs. As seen on the figure, during epoch 150 generator is able to reproduce synthetic MRIs which has slight differences from real ones.

reaching epoch 20. After training phase for 150 epochs, the model learns healthy brain anatomy for each MRI slices. In this way, the model is not capable of reconstruction brain lesions or any kind of distortions such as geometric or intensity distortions.

After the model learns the distribution of MRI images of healthy subjects, it is expected that the model is able to detect MRI with lesions or any kind of distortions not presented in the training dataset. For assessment of the model, test database was employed. The model generated synthetic MRI slices for each test MRI. In order to determine a normal MRI or an abnormal MRI, an anomaly score was obtained by comparing whole MRI slices with corresponding generated MRI slices. An MRI consists of slices and the anomaly score of MRI is the average anomaly score of its slices. For the location of errors on MRI, residual images were computed as pixel level intensity difference between an input image and a corresponding generated image. Test results for a healthy subject is given in Figure 4A-C and for a HGG patient is given in Figure 5A-C. The model correctly detected anomalies on 24 of 27 HGG patients' MRIs and marked them as abnormal. Besides, 25 of 27 healthy subjects' MRIs in the test dataset detected correctly as healthy MRI. The accuracy, precision, recall, and AUC were 0.907, 0.892, 0.923, and 0.907, respectively. When misdiagnosed subjects and patients were investigated it was observed that algorithm diagnosed ischemic changes in the white matter as a tumor. This is an expectable error hence algorithm was trained with T2-weighted images only. Additionally, when we investigate misdiagnosed HGG patients' MRIs we found that algorithm did labelled those images. This misdiagnosis was probably due to the limited sample size of the training set.

DISCUSSION

In this study, we demonstrated our results for visual inspection of the model. Additionally, we provided quantitative results using the metric mentioned in the performance metric section. Hence, our model learned the normal anatomy from healthy subjects' MRI scans, the model always produces normal synthetic MRI slices even input MRI slice has a big lesion causing geometric distortion. In literature there are mainly 3 classes of automatic segmentation algorithms which are supervised machine learning methods, deep learning methods with unsupervised learning and atlas based segmentation algorithms (2,13,17,19,20,22). Different than previously described methods, we proposed a new method by using DCGAN which compares the MRI slices of a patient with a randomly reproduced, fully synthetic T2-weighted images and segmenting the tumor as an error.

Manual segmentation of MRIs is time consuming, showing inter- or intra-operator differences and prone to human derived errors (5). Manual segmentation of brain tumors is a daily practice for neurosurgeons especially for the ones who performs radiosurgery. As a result, automatic segmentation of brain tumors became a desired feature. However, automatic segmentation of brain tumors is challenging because the anatomical structures in a random brain MRI includes both normal and abnormal structures in greyscales.

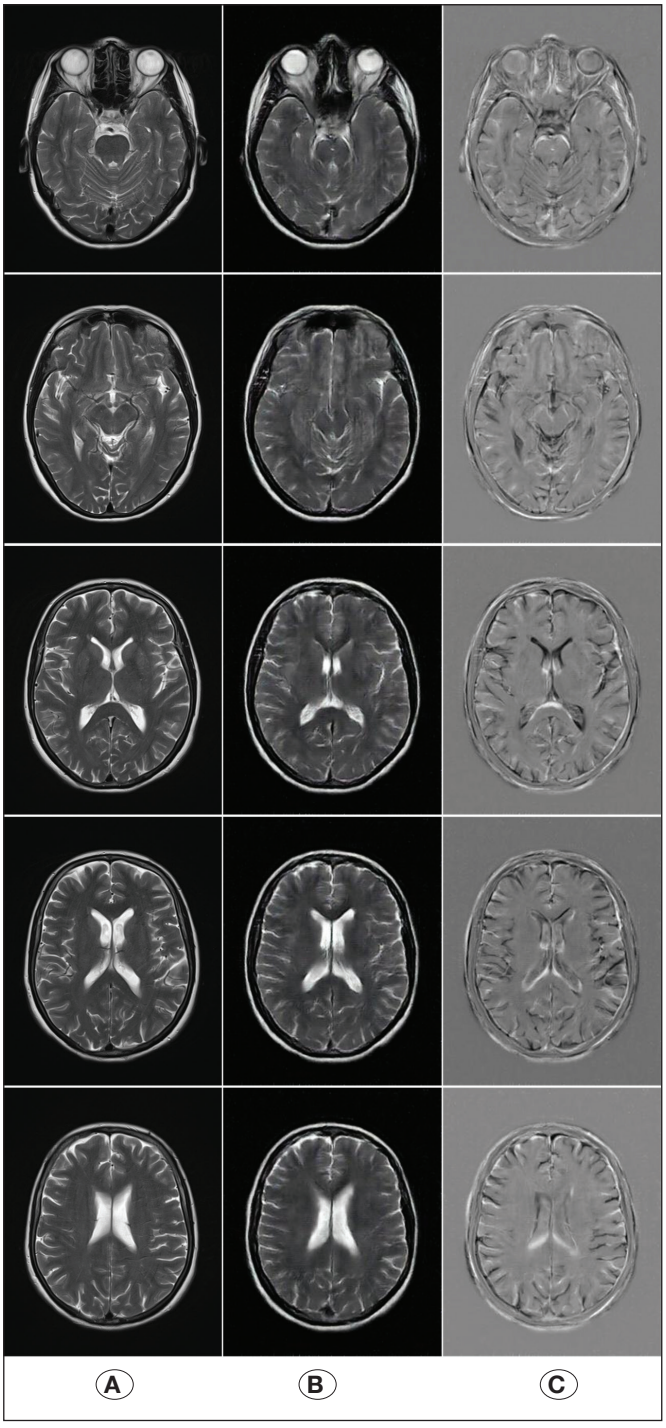

Figure 4: A) Selected T2 weighted axial original input slices which belongs to a healthy subject; **B)** generated MR slices **C)** difference of input MR slices and generated MR slices.

Another key feature of our study is our generalized adversarial network does not need training with HGG MRIs. Training of the algorithm is provided with only normal subject T2-weighted MRIs. Even though we tested our algorithm only with T2-weighted MRIs of HGG patients, it is applicable for any lesion which causes a distortion different than a normal

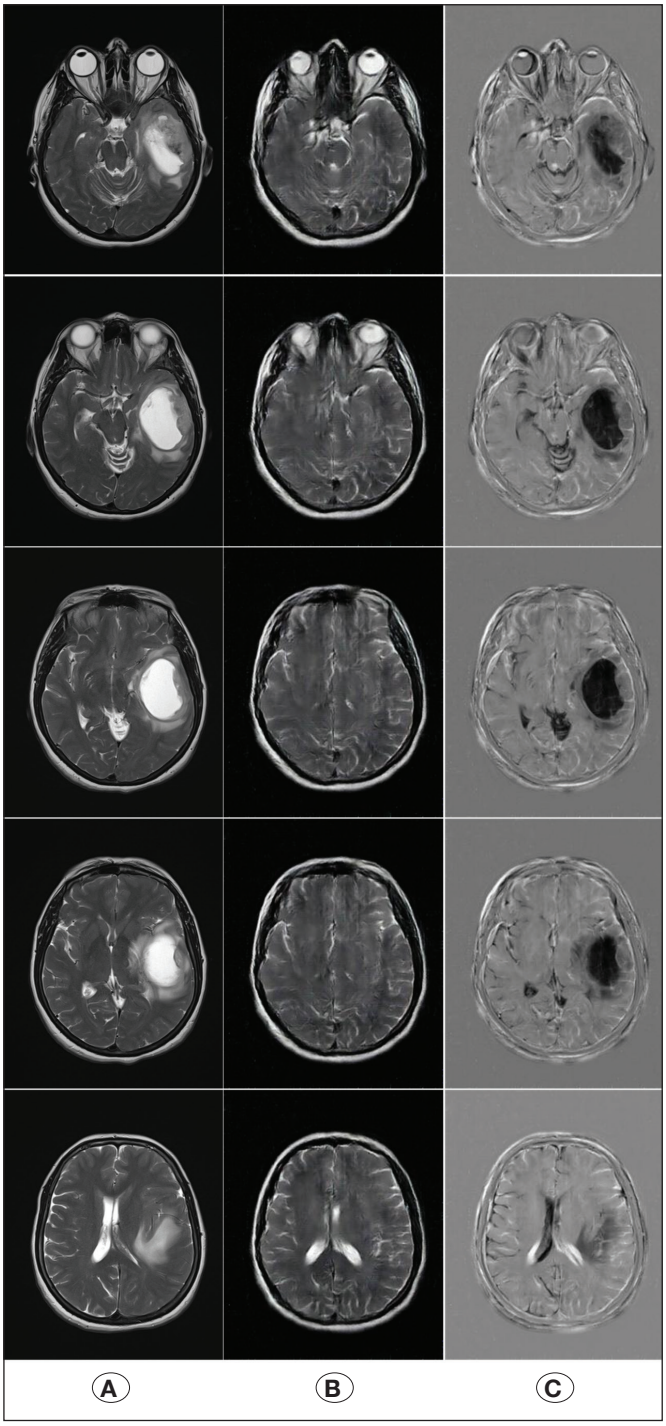

Figure 5: A) Selected T2 weighted axial original input slices which belongs to a HGG patient; **B)** generated MR slices **C)** difference of input MR slices and generated MR slices. As seen on the right column, tumor and edema causes a distortion which makes a difference from normal (black area).

MRI. This provides high efficiency and low training time for our proposed method.

Although our proposed method has acceptable accuracy, precision and recall rates, it has limitations. First, hence this is study is one of the pilot studies under the supervision of The Artificial Intelligence and Big Data Unit, The Digital Transformation Office of The Presidency of The Republic of Turkey, study comprise only T2-weighted MRIs. Second, number of normal subjects for training is limited and as we mentioned in the results section this particular limitation is the main reason for lower accuracy. Third, test subjects are comprised only HGG patients and normal subjects. This is a major limitation because proposed model needs further testing on different pathologies. All the limitations mentioned above are being investigated and subjects of further studies.

CONCLUSION

Automatic segmentation of brain tumors by using artificial intelligence methods are getting more popular as the daily neurosurgical practice evolves into a new technology era. Deep convolutional generative adversarial networks provides training with only normal MRIs and requires no MRIs with abnormal features. Future studies are warranted in order to increase effectiveness of artificial intelligence methods on the field.

REFERENCES

1. Acharya UR, Hagiwara Y, Sudarshan VK, Chan WY, Ng KH: Towards precision medicine: From quantitative imaging to radiomics. J Zhejiang Univ Sci B 19:6-24, 2018
2. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D: Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. NeuroImage 46:726-738, 2009
3. Atici MA, Sagiroglu S, Celtikci P, Ucar M, Borcek AO, Emmez H, Celtikci E: A novel deep learning algorithm for the automatic detection of high-grade gliomas on T2-weighted magnetic resonance images: A preliminary machine learning study. Turk Neurosurg 30(2):199-205, 2020
4. Celtikci E: A systematic review on machine learning in neurosurgery: The future of decision-making in patient care. Turk Neurosurg 28:167-173, 2018
5. Chen H, Dou Q, Yu L, Qin J, Heng PA: VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. Neuroimage 170:446-455, 2018
6. Gevaert O, Mitchell LA, Achrol AS, Xu J, Echegaray S, Steinberg GK, Cheshier SH, Napel S, Zaharchuk G, Plevritis SK: Glioblastoma multiforme: Exploratory radiogenomic analysis by using quantitative image features. Radiology 273: 168-174, 2014
7. Hassanipour S, Ghaem H, Arab-Zozani M, Seif M, Fararouei M, Abdzadeh E, Sabetian G, Paydar S: Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: A systematic review and meta-analysis. Injury 50:244-250, 2019
8. Jones TL, Byrnes TJ, Yang G, Howe FA, Bell BA, Barrick TR: Brain tumor classification using the diffusion tensor image segmentation (D-SEG) technique. Neuro Oncol 17:466-476, 2015

9. Kickingereder P, Bonekamp D, Nowosielski M, Kratz A, Sill M, Burth S, Wick A, Eidel O, Schlemmer HP, Radbruch A, Debus J, Herold-Mende C, Unterberg A, Jones D, Pfister S, Wick W, von Deimling A, Bendszus M, Capper D: Radiogenomics of glioblastoma: Machine learning-based classification of molecular characteristics by using multiparametric and multiregional MR imaging features. Radiology 281:907-918, 2016
10. Knoll MA, Oermann EK, Yang AI, Paydar I, Steinberger J, Collins B, Collins S, Ewend M, Kondziolka D: Survival of patients with multiple intracranial metastases treated with stereotactic radiosurgery: Does the number of tumors matter? Am J Clin Oncol 41:425-431, 2018
11. Li Z, Wang Y, Yu J, Shi Z, Guo Y, Chen L, Mao Y: Low-grade glioma segmentation based on CNN with fully connected CRF. J Healthc Eng 2017:9283480, 2017
12. Mazurowski M: Radiogenomics: What it is and why it is important. J Am Coll Radiol 8:862-866, 2015
13. Moeskops P, Benders MJNL, Chiţ SM, Kersbergen KJ, Groenendaal F, de Vries LS, Viergever MA, Išgum I: Automatic segmentation of MR brain images of preterm infants using supervised classification. Neuroimage 118:628-641, 2015
14. Odland A, Server A, Saxhaug C, Breivik B, Groote R, Vardal J, Larsson C, Bjørnerud A: Volumetric glioma quantification: Comparison of manual and semi-automatic tumor segmentation for the quantification of tumor growth. Acta Radiol 56:1396-1403, 2015
15. Oermann EK, Kress M-AS, Collins BT, Collins SP, Morris D, Ahalt SC, Ewend MG: Predicting survival in patients with brain metastases treated with radiosurgery using artificial neural networks. Neurosurgery 72:944-951; discussion 952, 2013
16. Oermann EK, Rubinsteyn A, Ding D, Mascitelli J, Starke RM, Bederson JB, Kano H, Lunsford LD, Sheehan JP, Hammerbacher J, Kondziolka D: Using a machine learning approach to predict outcomes after radiosurgery for cerebral arteriovenous malformations. Sci Rep 6:21161, 2016
17. Pereira S, Pinto A, Oliveira J, Mendrik AM, Correia JH, Silva CA: Automatic brain tissue segmentation in MR images using random forests and conditional random fields. J Neurosci Methods 270:111-123, 2016
18. Sarkiss CA, Germano IM: Machine learning in neuro-oncology: Can data analysis from 5,346 patients change decision making paradigms? World Neurosurg 2019 (Online ahead of print)
19. Shi F, Fan Y, Tang S, Gilmore JH, Lin W, Shen D: Neonatal brain image segmentation in longitudinal MRI studies. NeuroImage 49:391-400, 2010
20. Stollenga MF, Byeon W, Liwicki M, Schmidhuber J: Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. In: Cortes C, Lawrence ND, Lee DD, Sugiyama M, Garnett R (eds), Advances in Neural Information Processing Systems 28. Curran Associates, Inc, 2015
21. Sun R, Wang K, Guo L, Yang C, Chen J, Ti Y, Sa Y: A potential field segmentation based method for tumor segmentation on multi-parametric MRI of glioma cancer patients. BMC Med Imaging 19:48, 2019
22. Wang L, Gao Y, Shi F, Li G, Gilmore JH, Lin W, Shen D: LINKS: Learning-based multi-source integration framework for segmentation of infant brain images. NeuroImage 108:160-172, 2015
23. Wang Y, Li C, Zhu T, Zhang J: Multimodal brain tumor image segmentation using WRN-PPNet. Comput Med Imaging Graph 75:56-65, 2019
24. Wu Y, Zhao Z, Wu W, Lin Y, Wang M: Automatic glioma segmentation based on adaptive superpixel. BMC Med Imaging 19:73, 2019