

# Astrocytoma, ependymoma, and oligodendroglioma classification with deep convolutional neural network

Romi Fadillah Rahmat<sup>1</sup>, Mhd Faris Pratama<sup>1</sup>, Sarah Purnamawati<sup>1</sup>, Sharfina Faza<sup>2</sup>,  
Arif Ridho Lubis<sup>2</sup>, Al-Khowarizmi<sup>3</sup>, Muharman Lubis<sup>4</sup>

<sup>1</sup>Department of Information Technology, Universitas Sumatera Utara, Medan, Indonesia

<sup>2</sup>Department of Computer Engineering and Informatic, Politeknik Negeri Medan, Medan, Indonesia

<sup>3</sup>Department of Information Technology, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

<sup>4</sup>School of Industrial Engineering, Telkom University, Medan, Indonesia

## Article Info

### Article history:

Received Sep 30, 2021

Revised Jul 1, 2022

Accepted Jul 30, 2022

### Keywords:

Brain tumor

Deep convolutional neural  
network

Glioma classification

## ABSTRACT

Glioma as one of the most common types of brain tumor in the world has three different classes based on its cell types. They are astrocytoma, ependymoma, oligodendroglioma, each has different characteristics depending on the location and malignance level. Radiological examination by medical personnel is still carried out manually using magnetic resonance imaging (MRI) medical imaging. Brain structure, size, and various forms of tumors increase the level of difficulty in classifying gliomas. It is advisable to apply a method that can conduct gliomas classification through medical images. The proposed methods were proposed for this study using deep convolutional neural network (DCNN) for classification with k-means segmentation and contrast enhancement. The results show the effectiveness of the proposed methods with an accuracy of 95.5%.

*This is an open access article under the [CC BY-SA](#) license.*



## Corresponding Author:

Romi Fadillah Rahmat

Department of Information Technology, Universitas Sumatera Utara

Medan, 20143, Sumatera Utara, Indonesia

Email: romi.fadillah@usu.ac.id

## 1. INTRODUCTION

Glioma is one of the primary brain tumors [1]. This tumor originated from glial cells, which are neuron-supporting cells that function to help specify synaptic contact and preserve the signaling capability of neurons [2]. If we see statistics 15,000 peoples with brain cancer annually are glioma patients. There are 4 types of glioma classification based on cell types, they are astrocytoma, ependymoma, oligodendroglioma, and non-glioma [3]. Each type of tumor has certain property which varied depending on the location and the level of tumor malignancy [4]. Early diagnosis of astrocytoma and oligodendroglioma can be performed through clinical and radiological examination [5], [6]. Clinical examination is carried out to ascertain whether the patient has a brain disorder based on the symptoms experienced by the patient. In contrast, the radiological examination can be conducted using magnetic resonance imaging (MRI) imaging [7]. As we know that brain structure, size, and various forms of tumors increase the complexity level in classifying gliomas [8], [9]. In addition, there is the possibility of diagnostic errors that cause errors in handling, so a method is recommended to facilitate the classification of gliomas through medical images.

As a branch of artificial intelligence (AI), machine learning (ML) is covering the design and development of algorithms on computer systems to enable computers to perform and develop behaviors based on empirical data such as sensor data or databases [10]. While there is the latest powerful methodology such as deep learning (DL) which solve weakness in ML. DL approach is an breakthrough algorithm in

solving human life problem nowadays [11]. Supported by overwhelming computational power and big - complex dataset problem solver, DL has proven to be superior in image processing and preferably in medical imaging [12].

Numerous research have been carried out in classifying medical image using ML approach. Some of the researches use k-means clustering to find tumors in medical images [13]. In this study, k-means generated an excellent result in segmenting the brain tumor. K-means method was modified so it could segment the data on a large scale. Several combination methods also applied such as support vector machine (SVM) with k-means [14].

In the area of deep neural network (DNN), we also can get several previous studies. In 2017, Esteva *et al.* [15] applied DNN method for skin lesion images classification. The test was conducted by validating the system with real dermatologist to diagnose the clinical images with two classification use cases. This study succeeded in achieving an accuracy of 93.3% [15]. Cho *et al.* [16] did research to classify gliomas into low-level and high-level glioma using multi-modal image radiomics features. The test was implemented using 220 images of high level glioma and 54 images of low level glioma. We can see 88% of accuracy from this study [16]. Other research shown great brain tumor segmentation using deep neural network [17], convolutional neural network (CNN) with 2-phase training and brain tumor image segmentation benchmark (BRATS) test data [18], CNN with 3x3 kernel [19] and many novel-adjusted CNN [20], [21], also CNN implementation in 3-dimentional MRI Scan [22], [23] and also extreme learning machine (ELM) [24].

Identifying and classify astrocytoma, ependymoma, oligodendroglioma, and unspecified glioma has become problem nowadays. Using a normal DL approach will sure give a better result than common methods such as feed forward neural network or SVM. However, we need to get the differences between existed methodology and proposed methodology. Based on previous study we can see that implementing DL on glioma tumor is already exist, however we spot several differences. The differences are in glioma dataset used in this study, we can say it has not been implemented with deep convolutional neural network (DCNN), especially DCNN with k-means segmentation which will be described in our proposed method later.

## 2. RESEARCH METHOD

The classification process of glioma brain tumor consists of several steps. These steps are pre-processing of training data and testing data, training process, and classification process using DCNN method, and the system testing process. Figure 1 shows the general architecture of this research.

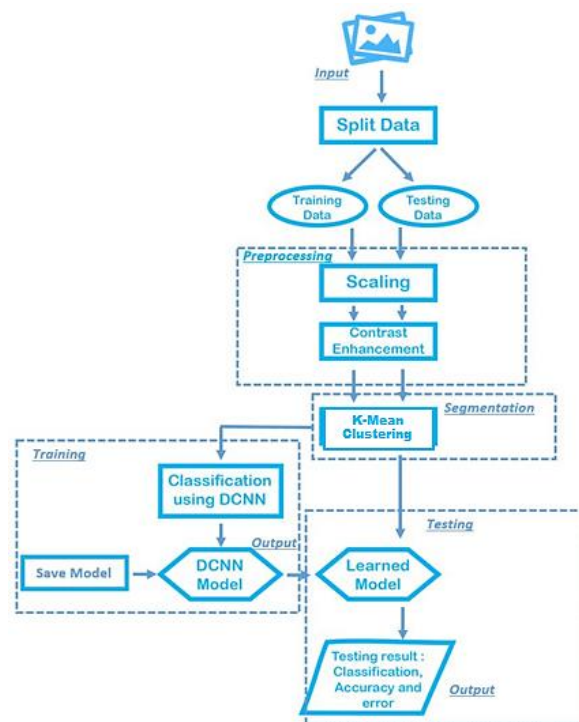


Figure 1. General architecture

The proposed method for this study consists of several stages. The stages are as follows: acquisition of glioma-detected images. All preprocessed image will be divided into training and testing data to be train in DCNN environment and to be tested later using testing data. The image preprocessing process span from grayscale image conversion; image cropping to minimize the black background of the image; contrast enhancement to get more precise data and including segmentation and thresholding to get specify area of glioma. Then we applied segmentation using k-means clustering. Then we do train process which will create DCNN model. This DCNN model then will be used as classification engine to get the results.

## 2.1. Data

We are using benchmark data in this research, where the data is in the form of images with joint photographic expert group (JPEG) format, that has been used for the training process and are available the cancer imaging archive (TCIA) [25]. These images consist of three categories, namely astrocytoma, ependymoma, and oligodendroglioma, several image samples shown in Figure 2. In addition to the above categories, this study also includes the images of brains without gliomas, with a total of four types of gliomas. Using data from the same source, researchers categorized brains that did not have gliomas as brains without tumors. The total data carried out in this study were 350 data, in the form of MRI images, that have been converted to images in JPG format. The distribution of datasets in this study can be seen in Table 1.

Table 1. Research data

Dataset	Astrocytoma	Ependymoma	Oligodendroglioma	Normal	Total
Training data	80	80	80	20	260
Testing data	25	25	25	15	90
Total	105	105	105	45	350

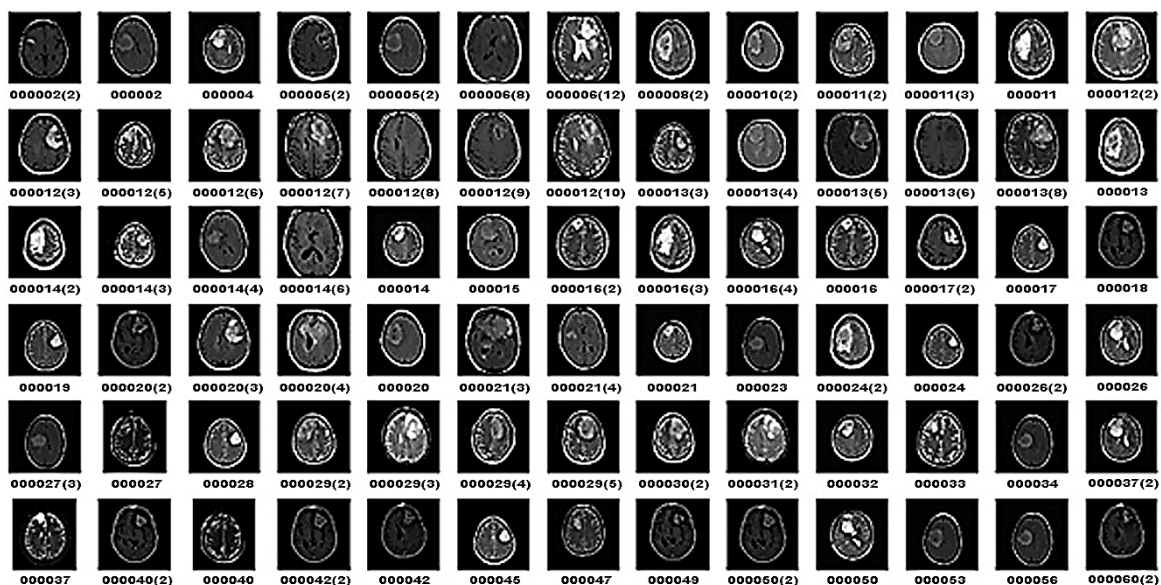


Figure 2. Images samples of astrocytoma, ependymoma, and oligodendroglioma glioma tumor [25]

## 2.2. K-means segmentation process

The segmentation process using k-means begins with reading the input data which is the result of the file from the gaussian filter in the previous stage. Then, the width and height of the image will be obtained, followed by the pixel value of the image. After that the pixel that will be used as the center of the cluster (centroid) is determined. The number of centroids has the same number as the specified number of clusters (k).

Then the calculation and sorting (sort) of the Euclidean distance values contained in the cluster will be carried out followed by partitioning (grouping) in each cluster. The grouping is done based on the calculation of the smallest Euclidean distance value generated between pixels and each centroid. The clustering is performed on all pixels so that the pixels are divided into groups according to the number of

clusters determined in the previous step. After pixels, the calculation of the average pixel value in each cluster is carried out. The resulting average pixel value will be the new centroid. After the new centroid value is obtained, the calculation and measurement of the Euclidean distance value, partition, and the calculation of the average pixel value will be carried out again until the number of iterations specified in the first step is successfully achieved or the convergence condition is successfully met. Convergence is a condition where the calculated mean value is equal to the centroid. When the convergence is met or the specified iteration has been carried out, the process will stop and the resulting output is a pixel that is in accordance with the desired number of clusters.

### 2.3. Deep convolutional neural network

DCNN is a DL based neural network approach used to analyze visual images, detect and recognize objects in an image, which are high-dimensional vectors, that will involve many parameters to characterize the network. Basically, DCNN is an advance version or deeper version of CNN. Bionic CNN was proposed to decrease the phase of pre-processing and alter the model architecture specifically for visual tasks. The CNN is usually arranged by a set of layers that can be grouped according to their function. CNN consists of neurons that have a weight, bias, and activation function with advance architecture and consist of two main parts such as feature extraction layer and fully-connected layer.

The process of classifying glioma brain tumors is carried out through several steps. These steps are training and testing data processing, and classification process using DCNN method, and lastly testing process by the system. In the input section data will be divided into 2, training and testing data. The data will go through a preprocessing stage or called pre-processing, while the pre-processing process is consisted of scaling, contrast enhancement and thresholding. If the incoming data is the training data, it will pass through the training stage as well as classify the data according to its classification. The incoming testing data will be directly processed by the model that has been created through the previous training process. Then the data will come out and issue the results of the processed classification.

### 2.4. Parameter setup in environment of deep convolutional neural network

In order to get the optimal number of hidden layers, we do several parameter tunings which varies from 2 until 10 hidden layers, then we get the best result based on optimal hidden layer shown in testing. For number of neurons inside hidden layer we prefer using 12 neurons in a fully-connected layer. To produce the output from each layer are calculated by the parameter from input, weight, bias and activation function. In terms of activation function, we use the ReLU activation function for convolutional layer and the Softmax on the output layer to get categorical data results. While for optimizer to determine optimization inside DCNN, in this study we use Root Mean Square Propagation (RMSProp). Thus, in terms of batch size that can determine the number of observations made before weight changing based on computer specification and setup, we use default batch size, which is 32. For the epoch or number of repeated learning process. The greater the number of epochs, the higher the level of learning outcomes until our models reach convergence. The number of epochs used in this study is 100.

## 3. RESULTS AND DISCUSSION

Here in this section, we describe our application, we can see in the Figure 3 that the system is built in web-based application as a frontend with the DL Python in the backend. Figure 3 is the main user interface of our proposed system. Table 2 is the result of the process through which the proposed method goes. It can be seen the stages of image processing starting from the original image, scaling, contrast enhancement and thresholding processes. In this example we can see tumor region which will become input for DCNN model in our system.

The standard used in evaluating the system performance using testing data is the gold standard, which is based on true false-positive negative. True positive (TP) is when the desired output and actual output generate the same result of tumor area detected. False positive (FP) is a state where the desired output should be an undetected tumor area while the actual output generates a detected tumor area. True negative (TN) is a condition where the desired output and actual output produce an undetected tumor area. Then, false negative (FN) is a state where the desired output is a detected tumor area while the actual output generates the opposite. The test was performed using a dataset of 90 data of patients with brain tumors of oligodendroglioma, astrocytoma and glioblastoma, and healthy brains. The result of data testing can be seen in Table 3.

The amount of data used in this study amounted to 350 images for three classes of glioma brain cancer classified plus one type of brain classification that does not have a tumor/glioma. Due to the insufficient amount of data, the classification process has not yet perfected. Apart from the lack of data, some

data have similar characteristics and locations, which produced an error in the classification error. It is mandatory to have the correct feature extraction to process the image before the classification process.

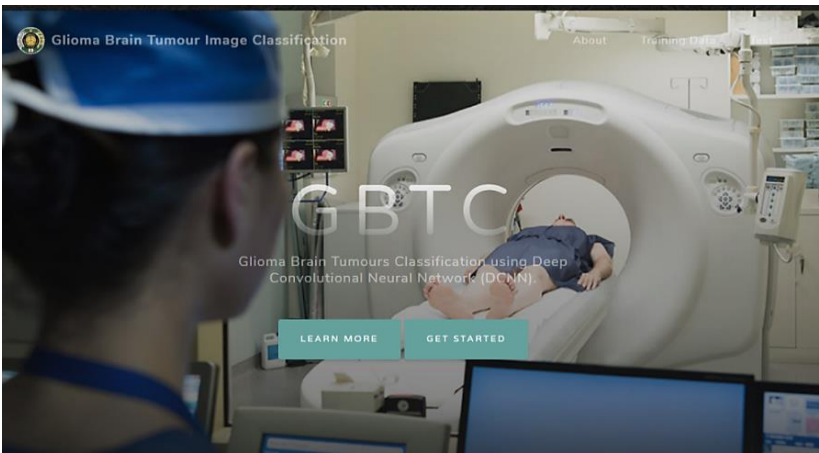
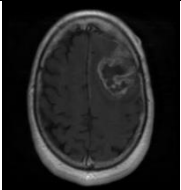
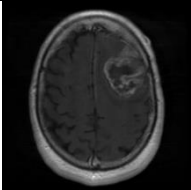
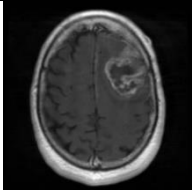
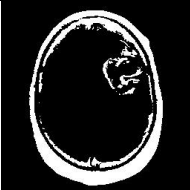
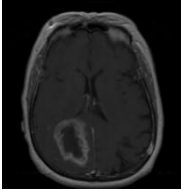
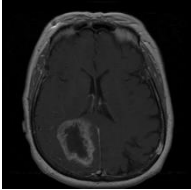
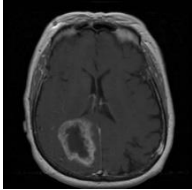
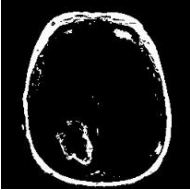
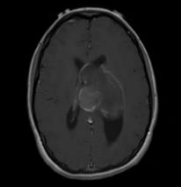
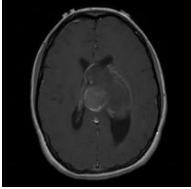
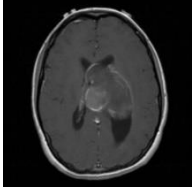
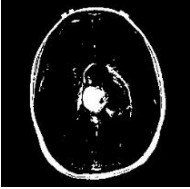
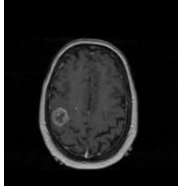
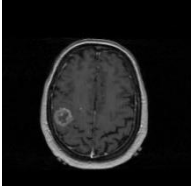
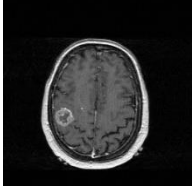

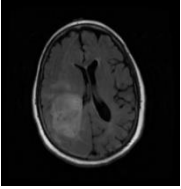
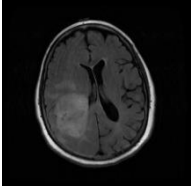
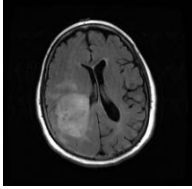
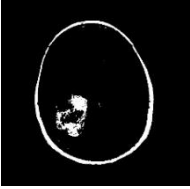


Figure 3. Glioma brain tumor classification GUI

Table 2. Image pre-processing results			
Original Image	Scaling	Contrast Enhancement	Thresholding
			
			
			
			
			

Based on the testing result of our proposed glioma brain tumor classification system, we get an average of training accuracy value of 90%. While for the speed of training process is in 184 seconds/epoch, thus it takes around 3 hours 7 minutes to complete training process up to 100 epochs. The test results of the glioma classification system using the DCNN, the system achieved 95.5% accuracy.

$$Accuracy = \frac{23 + 24 + 24 + 15}{23 + 1 + 1 + 24 + 1 + 1 + 24 + 15} = \frac{86}{90} \times 100\% = 95.5\%$$

As we can see there are only 4 images that categorize as false classified. An example of an image that failed to be classified by the system can be seen in Figure 4(a). This image is an image of a type of astrocytoma that was misclassified by the system as an ependymoma. This is because the type of ependymoma is located in the ventricular part of the brain which is located adjacent to the tumor, therefore the image processing used is not quite precise because it has not been able to detect the exact location of the ventricular and part of the tumor. Another thing that causes the classification of this system to be disturbed is that the image processing section used is not yet capable of eliminating unnecessary parts, and incorrectly selecting tumors in the data. One of the data that failed to be classified is in Figure 4(b). In Figure 4(b), this image is an image of a type of astrocytoma that was misclassified by the system as oligodendroglioma. This is because the system has not been able to remove unnecessary parts as shown in Figure 4(c) and has not been able to select the correct tumor section as shown in Figure 4(d).

Table 3. Testing data results

	Astrocytoma	Ependymoma	Oligodendroglioma	No Tumor
Astrocytoma	23	0	0	0
Ependymoma	1	24	1	0
Oligodendroglioma	1	1	24	0
No Tumor	0	0	0	15

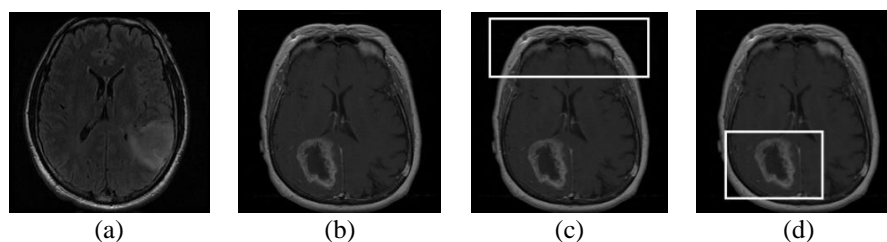


Figure 4. False classification images by the system as (a) ependymoma, (b) oligodendroglioma, (c) unnecessary parts and (d) the part that the system failed to pick up

#### 4. CONCLUSION

DCNN method is capable of performing glioma classification using MRI images with an accuracy value of 90% using a dataset of 270 images for the training process. The accuracy was affected by the consistent training data in one class of glioma. For the testing process, the system obtained an accuracy of 95.5% with a total of 90 testing data. For further research, it is strongly encouraged to have more data to increase the accuracy, as well as more methods in data image pre-processing and to perform data visualization to obtain the best data for the research.

#### ACKNOWLEDGEMENTS

We would like to thank you Universitas Sumatera Utara, for supporting this research through university grant no 448/UN5.2.3.1/PPM/SPP-TALENTA USU/2020.

#### REFERENCES




- [1] A. Drevelegas, *Imaging of brain tumors with histological correlations*, 2nd ed. Springer Berlin Heidelberg, 2011.
- [2] D. Purves *et al.*, *Neuroscience*, 2nd ed. Sunderland (MA): Sinauer Associates, 2001.
- [3] A. Perkins and G. Liu, "Primary brain tumors in adults: diagnosis and treatment," *American Family Physician*, vol. 93, no. 3, pp. 211–217, Feb. 2016.






- [4] T. Schneider, C. Mawrin, C. Scherlach, M. Skalej, and R. Firsching, "Gliomas in adults," *Deutsches Ärzteblatt international*, vol. 107, no. 45, pp. 799–808, Nov. 2010, doi: 10.3238/arztebl.2010.0799.
- [5] M. Smits, "Imaging of oligodendroglioma," *The British Journal of Radiology*, vol. 89, no. 1060, p. 20150857, Apr. 2016, doi: 10.1259/bjr.20150857.
- [6] M. Kapoor and V. Gupta, *Astrocytoma*. Treasure Island (FL): StatPearls Publishing, 2022.
- [7] K. K. Koeller and E. J. Rushing, "Oligodendroglioma and its variants: radiologic-pathologic correlation," *RadioGraphics*, vol. 25, no. 6, pp. 1669–1688, Nov. 2005, doi: 10.1148/rg.256055137.
- [8] M. M. Kounelakis *et al.*, "Strengths and weaknesses of 1.5T and 3T MRS data in brain glioma classification," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 4, pp. 647–654, Jul. 2011, doi: 10.1109/titb.2011.2131146.
- [9] P. M. Patel, B. N. Shah, and V. Shah, "Image segmentation using k-mean clustering for finding tumor in medical application," *International Journal of Computer Trends and Technology (IJCTT)*, vol. 4, no. 5, pp. 1239–1242, 2013.
- [10] T. G. Dietterich, "Machine learning," pp. 1056–1059, Jan. 2003, doi: 10.1002/0470018860.s00036.
- [11] S. Krishnan, P. Magalingam, and R. Ibrahim, "Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, pp. 5467–5476, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5467-5476.
- [12] M. Al-Smadi, M. Hammad, Q. B. Baker, S. K. Tawalbeh, and S. A. Al-Zboon, "Transfer deep learning approach for detecting coronavirus disease in X-ray images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, pp. 4999–5008, Dec. 2021, doi: 10.11591/ijece.v11i6.pp4999-5008.
- [13] V. S. A. Kumar and T. C. S. Rao, "Brain tumor extraction by k-means clustering based on morphological image processing," *V. Sagar Anil Kumar (PG Scholar), T. Chandra Sekhar Rao*, vol. 3, no. 10, pp. 8539–8546, Oct. 2014.
- [14] L. K. Abdulgafoor and A. George, "Detection of brain tumor using modified k means algorithm and SVM," in *National Conference on Recent Trends in Computer Applications NCRTCA*, 2013, pp. 28–31.
- [15] A. Esteve *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, Jan. 2017, doi: 10.1038/nature21056.
- [16] H. Cho and H. Park, "Classification of low-grade and high-grade glioma using multi-modal image radiomics features," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2017, pp. 3081–3084, doi: 10.1109/embc.2017.8037508.
- [17] M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, Jan. 2017, doi: 10.1016/j.media.2016.05.004.
- [18] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, May 2016, doi: 10.1109/tmi.2016.2538465.
- [19] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, May 2016, doi: 10.1109/tmi.2016.2548501.
- [20] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: deep convolutional neural network for segmenting neuroanatomy," *NeuroImage*, vol. 170, pp. 434–445, Apr. 2018, doi: 10.1016/j.neuroimage.2017.02.035.
- [21] J. Dolz, C. Desrosiers, and I. Ben Ayed, "3D fully convolutional networks for subcortical segmentation in MRI: a large-scale study," *NeuroImage*, vol. 170, pp. 456–470, Apr. 2018, doi: 10.1016/j.neuroimage.2017.04.039.
- [22] K. Kamnitsas *et al.*, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, Feb. 2017, doi: 10.1016/j.media.2016.10.004.
- [23] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, pp. 221–248, Jun. 2017, doi: 10.1146/annurev-bioeng-071516-044442.
- [24] R. F. Rahmat, R. M. Wijaya, S. Faza, E. B. Nababan, F. Nadi, and Al-Khowarizmi, "Classification of primary and secondary brain tumor using extreme learning machine," Nov. 2021, doi: 10.1109/databia53375.2021.9650272.
- [25] K. Clark *et al.*, "The cancer imaging archive (TCIA): maintaining and operating a public information repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, Jul. 2013, doi: 10.1007/s10278-013-9622-7.

## BIOGRAPHIES OF AUTHORS






**Romi Fadillah Rahmat**    is a researcher and lecturer in Universitas Sumatera Utara. Currently, he is a Vice Dean in Faculty of Computer Science and Information Technology. He finished his undergraduate and master program from Universiti Sains Malaysia and pursuing his Ph.D. program in Universitas Sumatera Utara. He is also an associate professor in Universitas Sumatera Utara since 2020. His research interest is in image processing, deep learning, brain-like computing, internet of things and mobile computing. He can be contacted at email: romi.fadillah@usu.ac.id.






**Mhd Faris Pramata**    was a researcher in Universitas Sumatera Utara. Currently, he is an information technology operations analysis in PT. Bank Negara Indonesia. He finished his undergraduate from Universitas Sumatera Utara in 2019. His research interest is in image processing, deep learning, and information technology. He can be contacted at email: farispratamaa@gmail.com, LinkedIn: <https://id.linkedin.com/in/mhd-faris-pratama-9ba4ab15b>.






**Sarah Purnamawati**    is a researcher and lecturer in Universitas Sumatera Utara. Currently, she is a Head of Information Technology in Faculty of Computer Science and Information Technology. She finished her master program from Universiti Sains Malaysia. Her research interest is in Database and Image Processing. She can be contacted at email: sarah\_purnamawati@usu.ac.id.






**Sharfina Faza**    is a researcher and lecturer in Politeknik Negeri Medan. She received her undergraduate and master program from Universitas Sumatera Utara. She is currently the Secretary of Research Program in Politeknik Negeri Medan. Her research interests are in the area of image processing, deep learning, and information technology. She can be contacted at email: sharfinafaza@polmed.ac.id.






**Arif Ridho Lubis**    he got master from University Utara Malaysia in 2012 and graduate from University Utara Malaysia in 2011, both information technology. He is a lecturer in Department of Computer Engineering and Informatics, Politeknik Negeri Medan in 2015. His research interest includes computer science, network, science and project management. He can be contacted at email: arifridholubis@polmed.ac.id.



**Al-Khowarizmi**    was born in Medan, Indonesia, in 1992. He is a Dean in Faculty of Computer Science and Information Technology at Universitas Muhammadiyah Sumatera. He got master from University of Sumatera Utara in 2017. His main research interest is data science, big data, machine learning, neural network, artificial intelligence and business intelligence. He can be contacted at email: alkhwarizmi@umsu.ac.id.



**Muharman Lubis**    has finished his Doctoral degree recently in Information Technology at 2017 in International Islamic University Malaysia, he also received his Master degree from same university at 2011 and Bachelor degree from University Utara Malaysia at 2008, both in Information Technology. He joined as a Lecturer in the School of Industrial Engineering, Telkom University, in 2017. His research interests include privacy protection, information security awareness, knowledge management and project management. He can be contacted at email: muharmanlubis@telkomuniversity.ac.id.