Recursive Journal of Informatics
Vol. 3, No. 2, September 2025

p-ISSN 2963-5551 https://journal.unnes.ac.id/journals/index.php/rji e-ISSN 2986-6588

Brain Tumor Classification on Magnetic Resonance Imaging Images using Convolutional Neural Network with Cycle Generative Adversarial Network and Extreme Gradient Boosting

Akbar Lintang Aji[1], Endang Sugiharti[2]
[1,2]Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang, Indonesia

Abstract. With the current advancement in technology, image classification process can be carried out through computer processing. This can also be applied to various fields, one of which is the health sector. The health sector is known for its high complexity in pattern recognizing of diseases. One of the diseases that is difficult to classify is brain tumors.
Purpose: This study aims to improve the accuracy of classification in brain MRI images, which are known to have a small and unbalanced sample. This limitation poses challenges in developing an effective classification model. The classification model is highly dependent on the quantity of data used for training. Therefore, data augmentation techniques play a crucial role in influencing the model's performance.
Methods/Study design/approach: In this study, CNN model using VGG-19 architecture was used to learn feature of brain tumor in brain MRI images. Additionally, CycleGAN is used to augment and balance the data, addressing issues related to data scarcity and imbalance, thus improving diversity of the dataset. And then, XGBoost is applied to classify the feature learned by the CNN model.
Result/Findings: CycleGAN has the ability to generate new image by transferring characteristics between images with different classes, making it a suitable to replace traditional data augmentation techniques in CNN. Additionally, XGBoost can be used to improve the classification results by classifying the features learned by CNN model during the training process. The proposed combination method achieves a highest accuracy of 97.37%.
Novelty/Originality/Value: CNN combined with CycleGAN and XGBoost successfully improved the accuracy of the model and addresed data scarcity and imbalances in the dataset used. This combined method can improve the accuracy of the classification models. This is proven by an accuracy increase of 0.36% when compared to previous research.

Keywords: Image Classification, Convolutional Neural Network, VGG-19, Cycle Generative Adversarial Network, Extreme Gradient Boosting

Received January 2024 / **Revised** September 2025 / **Accepted** September 2025

INTRODUCTION

In recent years, the use of technology has become an integral part of everyday life. Machine Learning (ML), which is a subfield of Artificial Intelligence (AI), is a technology capable of doing classification for complex tasks such as decision making [1]. ML can be applied to various computational fields, such as face detection, traffic prediction, product recommendations, and medical diagnosis [2]. Image processing is also within its capabilities, making image recognition one of the extensively developed technologies in recent times. Healthcare sector had a significant improvement in implementing image processing to address various issues [3]. The detection of brain tumors is one of the issues that can be addressed, where the tumors can be located in random positions with varying sizes and shapes [4].

One of the algorithms that exhibits strong performance in image processing is Convolutional Neural Network (CNN). CNN is a neural network designed for processing grid-structured data that utilizes convolution as a key operation in its layer. This operation involves a linear algebra operation that multiplies the filter matrix with the image for processing [5]. The widespread use of the CNN algorithm is attributed to its capability to simultaneously perform feature extraction and classification [6]. Despite its known

[1]*Corresponding author.
Email addresses: akbarlintang@students.unnes.ac.id (Aji)
DOI: 10.15294/rji.v3i2.486

effectiveness, CNN algorithms have weaknesses in terms of the amount of data and time required to train the classification models.

One way to address the training time of a CNN model training is by implementing Transfer Learning, where a pre-trained model on a specific task is reused to perform a different task. One of the pre-trained CNN models is VGG-19, which is a deep learning neural network with 16 convolutional layers and 3 fully connected layers [7]. In addition to reducing training time, Transfer Learning can also yield better performance compared to models trained from scratch [8]. Then, to overcome the issue of data scarcity, data augmentation techniques can be used to increase the training samples, which improves the diversity of the training data [9]. However, the use of traditional data augmentation techniques within CNNs may fall short in providing sufficient diversity in the training data.

To overcome this issue, Cycle Generative Adversarial Network (CycleGAN) method can be used as an alternative data augmentation technique to increase the number of training data. CycleGAN consists of two mirror-symmetric GANs arranged in a ring network that leverages cycle consistency loss to optimize the frame structure [10]. CycleGAN structure involves the sharing of two generators and two discriminators between both GANs. CycleGAN stands out as one of the commonly used approach in synthetic medical image applications [11]. Following this, XGBoost can be used to classify the features learned through CNN. XGBoost is a machine learning algorithm that employs a gradient boosting structure and is rooted in the principles of decision trees [12]. The combination of CNN, CycleGAN, and XGBoost has the potential to enhance the accuracy of brain tumor classification in brain MRI images.

METHODS

In this research, the first step involves preparing the dataset to be used. Subsequently, several stages of data preprocessing are conducted, including cropping, resizing, and data splitting. Afterward, the CycleGAN method is used to augment and balance the training data, replacing traditional data augmentation techniques. The data generated by the CycleGAN's generator is then combined with the original training data, resulting in a combined training dataset, which then are used as input data for the training of CNN model using VGG-19 architecture. The next step involves extracting the dense layer from the trained CNN model to predict input data classification probability. These extracted data are then used as input for the XGBoost model, which performs classification based on the data learned by the CNN. Flowchart of the combined method used in this research is as shown in Figure 1.

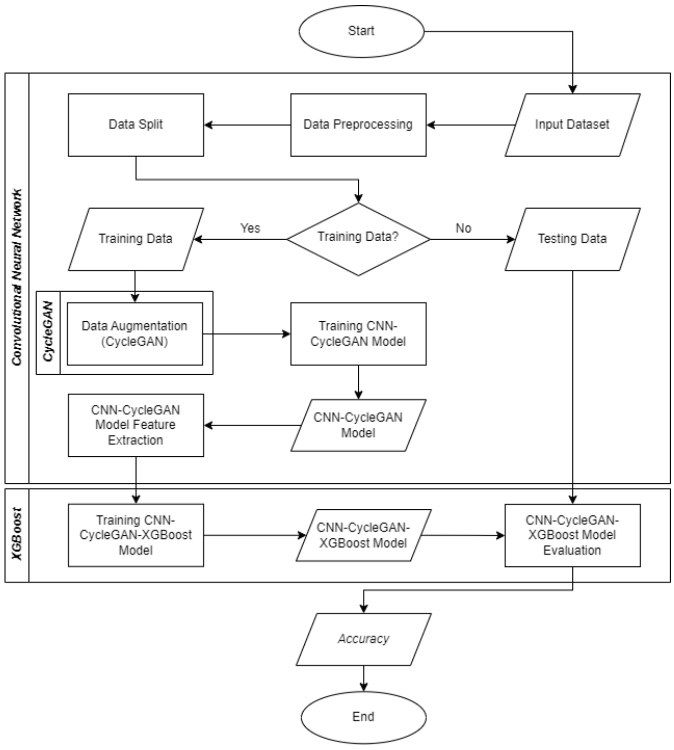


Figure 1. Research design

Data Collection

The data used in this research are Brain MRI Images for Brain Tumor Detection obtained from Kaggle. This dataset contains 253 brain MRI images which are divided into 2 classes, where 155 images are labeled as "yes" and 98 images are labeled as "no". Brain MRI images provide detailed information about the brain structure, facilitating the identification of anomalies in brain tissue [13].

Data Preprocessing

Data preprocessing is required so that images can be processed in the next stage. The steps of data preprocessing conducted are as follows.

Crop

Cropping is used to eliminate unnecessary parts of the image so that the model training process can focus on the main parts of the brain.

Resize

Resizing is used to standardize the dimensions of the images, as the dataset used contains images of various sizes. In this study, all images are resized to a dimension of 256 x 256 pixels, as it yields sufficiently high accuracy results with a relatively short training time for the CNN model [14].

Data Split

Data split is used to separate training data, validation data, and testing data. The training data will be used to train the classification model, while the validation data is used to validate the classification results during the training process. And the testing data is used to evaluate the accuracy of the trained combined model. In this study, the ratio used for splitting is 70% for training data, 15% for validation data, and 15% for testing data.

Data Augmentation using CycleGAN

In this study, CycleGAN is utilized to augment and balance the number of data between classes. CycleGAN enables the transformation of images from one domain to another domain with unpaired images [15]. Data labeled 'yes' will have its characteristics transformed into 'no', and vice versa for data labeled 'no,' transforming its characteristics into 'yes'. Subsequently, the data with altered characteristics is combined with the original training data, ensuring a balanced number of training data between classes.

Feature Learning using CNN

The combined training data from the previous stage is used as input data to train the CNN model using VGG-19 architecture. The feature extraction component of the CNN typically comprises convolutional and pooling layers, while the classification segment involves fully connected and classification layers [16]. In this study, the dense layer or the fully connected layer of the trained model is extracted to predict the classification probabilities of the input data, represented as a decimal number between 0 and 1. Thus, feature extraction from the CNN classification model is obtained and subsequently used as input data to train the XGBoost model.

Classification using XGBoost

The features extracted in the previous stage are used as input data to train the XGBoost model. The use of XGBoost as the final-level classifier replaces the role of the fully connected layer in CNN, thereby reducing the complexity of the model and the number of parameters [17].

Model Evaluation

The evaluation stage is conducted to test the combined model trained in the previous stages. A confusion matrix is used to evaluate the model's performance by calculating some metrics such as accuracy, precision, recall, and F1-score [18]. The assessment table based on the confusion matrix is as shown in Table 1

Table 1. Confusion matrix

		Predicted	
		0	1
Actual	0	True Positive (TP)	False Positive (FP)
	1	False Negative (FN)	True Negative (TN)

Based on Table 1, the precision, recall, F1-score, and accuracy values can be formulated as in Equation 1, Equation 2, Equation 3, and Equation 4, respectively.

$$precision = \frac{TP}{TP + FP} \tag{1}$$

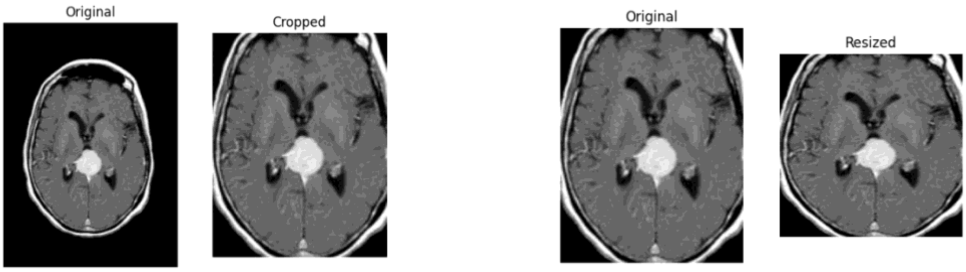
$$recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2\,(precision * recall)}{precision + recall} \tag{3}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

RESULT AND DISCUSSION

The Brain MRI Images for Brain Tumor Detection dataset obtained from Kaggle, consisting of 253 brain MRI images, is used as input data in data preprocessing. The data preprocessing stage includes cropping, resizing, and data splitting. The results of the initial data preprocessing can be seen in Figure 2(a) for the crop results and 2(b) for the resize results.



(a) (b)

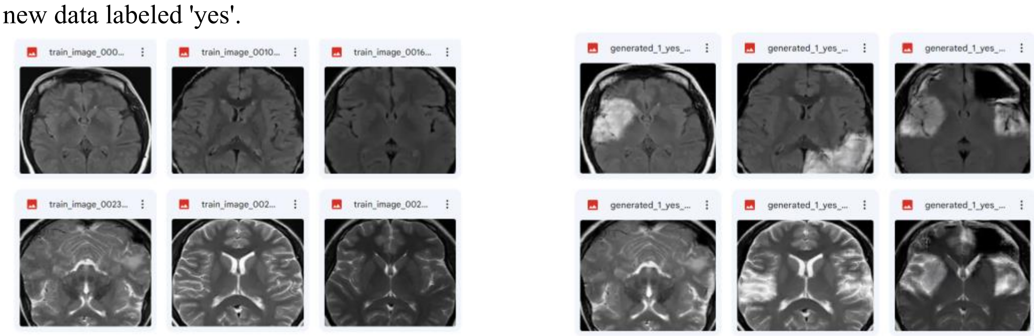
Figure 2. (a) Crop result, (b) Resize result

After undergoing the crop and resize processes, the data is then divided into three groups: training data, validation data, and testing data. The data splitting ratio used in this study is as shown in Table 2.

Table 2. Data split

Set	Percentage	Total
Training Data	70%	177
Validation Data	15%	38
Testing Data	15%	38

Subsequently, the training data obtained from the series of previous processes is used to train the CycleGAN model. CycleGAN has a hyperparameter called epoch, where each epoch signifies one complete round of training the model on the input data. In this study, several scenarios are conducted with different numbers of epochs used to train CycleGAN, namely with 50 epochs, 100 epochs, 150 epochs, and 200 epochs. This is done to determine the number of epochs that can yield the highest accuracy for the CNN-XGBoost classification model later on. An example of CycleGAN results can be seen in Figure 3(a) for input data labeled 'no' and Figure 3(b) for the CycleGAN generator's output from input data labeled 'no' transformed into new data labeled 'yes'.

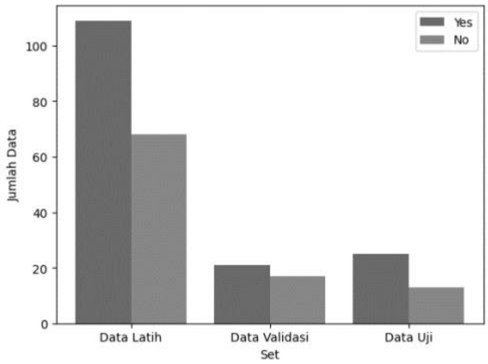


(a) (b)

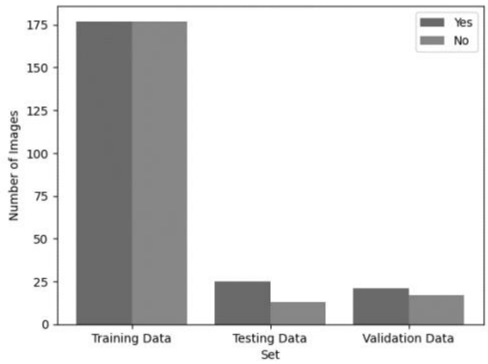
Figure 3. (a) Input data label no, (b) Generated data label yes

The next step is to combine the training data with the data generated by the CycleGAN. Training data labeled 'yes' is fed into the 'no' generator, while training data labeled 'no' is fed into the 'yes' generator. As a result, the combined training data has a balanced quantity. The data on each set before data augmentation is as shown in Figure 4(a), while the data on each set after data augmentation with CycleGAN is as shown in Figure 4(b).

.



(a)



(b)

Figure 4. (a) Data on each set before augmentation, (b) Data on each set after augmentation

The combined training data is then used as input data to train the CNN model using VGG-19 architecture. The training of the CNN model is conducted according to each testing scenario. After the CNN model in each testing scenario is trained, the next step is to perform feature extraction by taking the last layer of the CNN model and making predictions on the input data.

The feature extraction results from each testing scenario of the CNN are then used as input data to train the XGBoost model. This results in the combined CNN-CycleGAN-XGBoost model, which then the model's performance can be evaluated using the testing data and confusion matrix.

After undergoing various tests scenario, the best classification model with the highest accuracy was obtained. The results of the accuracy comparison from each evaluation test are displayed in Table 3.

Table 3. Comparison of model accuracy

Test	Data augmentation	CNN epoch	CNN accuracy	CNN-XGboost accuracy
1	Traditional	10	89.47%	92.11%
2	CycleGAN 50 epoch	25	81.58%	89.47%
3	CycleGAN 100 epoch	25	89.47%	84.21%
4	CycleGAN 150 epoch	5	94.74%	94.74%
5	CycleGAN 200 epoch	10	89.47%	97.37%

Based on Table 3, the test that yielded the highest accuracy is Test 5 with CycleGAN data augmentation at 200 epochs and CNN at 10 epochs. Performance evaluation of the model was then conducted using a confusion matrix, as shown in Figure 5.

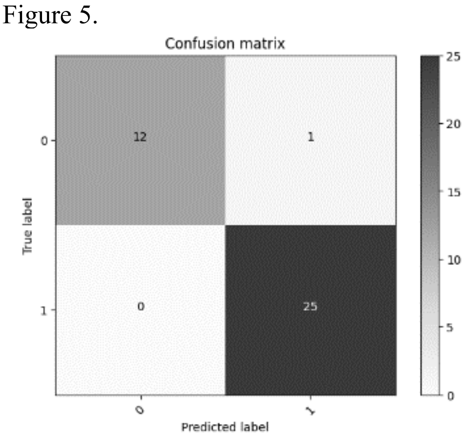


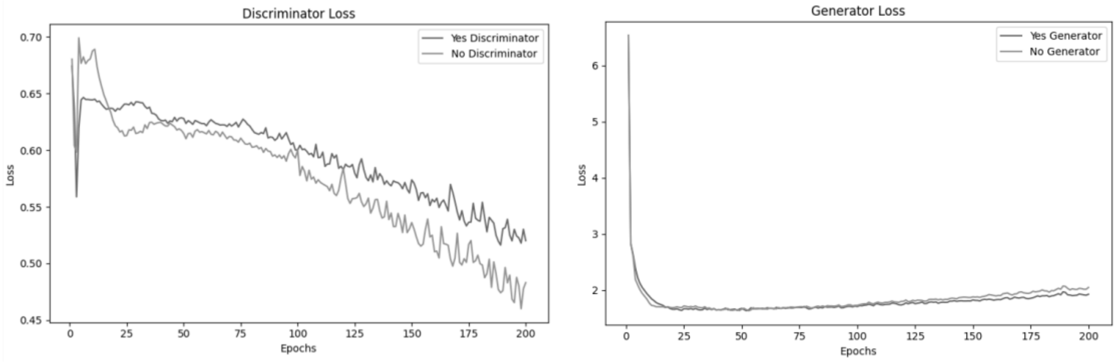
Figure 5. Confusion matrix of test 5

Based on Figure 5, evaluation metrics of the model from test 5 can be calculated using the formulation of Equation 1, Equation 2, Equation 3, and Equation 4. The results of confusion matrix during test 5 are presented in Table 4.

Table 4. Confustion matrix results of test 5

Evaluation metrics	Value
Precision	0.962
Recall	1.000
F1-score	0.980
Accuracy	0.9737

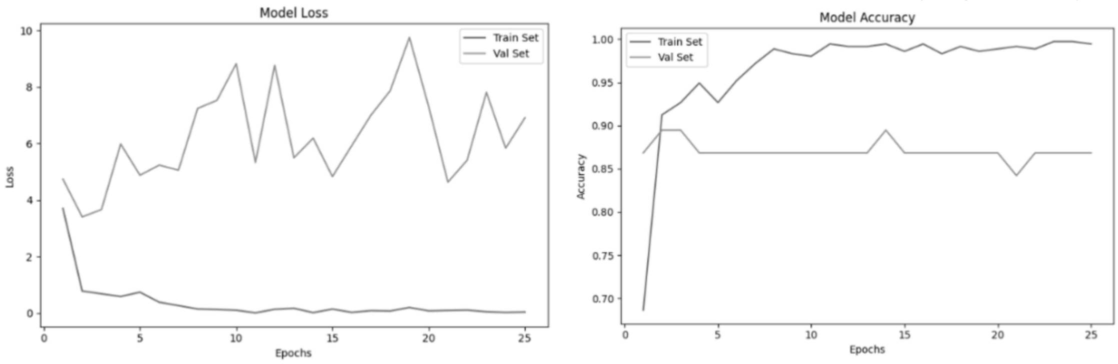
In test 5, the epoch used for training CycleGAN model was 200 epoch, which is the highest among all the test scenario. The progression of the CycleGAN model is as shown by the graph in Figure 6(a), which depicts a decrease in discriminators loss and Figure 6(b) depicts an increase in generators loss. This indicates that the CycleGAN model improved over time, because the generator is getting better at producing images that are more difficult for the discriminator to distinguish from real ones. Nevertheless, this also indicates that the generator might produces distorted image if continued, which could harm the discriminator and decreasing the overall quality.



(a) (b)

Figure 6. (a) Cyclegan discriminator loss, (b) Cyclegan generator loss

In test 5, an accuracy loss of 0.0978 and a validation loss of 8.8180 were obtained from the CNN model training. This is shown by the graph in Figure 7(a), which depicts an increase in validation loss with the increasing number of epochs. However, the progression of validation accuracy, as shown in Figure 7(b), can be considered relatively stable, indicating that the model still maintains a reasonably high accuracy.



(a) (b)

Figure 7. (a) Training and validation loss, (b) Training and validation accuracy

Based on the results of the series of tests above, test 5, which exhibits the best model performance, is utilized for comparison with previous studies. The proposed method, which combines CNN with CycleGAN and XGBoost, achieves a superior accuracy of 97.37%. In comparison, Sofian & Laluma [19] employed KNN with Image Threshold and GLCM, resulting in an accuracy of 83.33%. Wahid et al. [20] utilized ELM, achieving an accuracy of 86%. Meanwhile, Siddique et al. [21] and Cinar & Yildirim [22], both employing CNN with different architectures, achieved accuracies of 96% and 97.01%, respectively. The accuracy comparison between this study and previous studies is presented in Table 5.

Table 5. Comparison of accuracy based on previous research

Researcher	Research title	Research method	Research result
Sofian & Laluma (2020)	Klasifikasi Hasil Citra MRI Otak Untuk Memprediksi Jenis Tumor Otak dengan Metode Image Threshold dan GLCM Menggunakan Algoritma K-NN (Nearest Neighbor) Classifier Berbasis Web	KNN + Image Threshold + GLCM	83.33%
Wahid et al. (2020)	Implementasi Metode Extreme Learning Machine untuk Klasifikasi Tumor Otak pada Citra Magnetic Resonance Imaging	Extreme Learning Machine	86%
Siddique et al. (2020)	Deep Convolutional Neural Networks Model-based Brain Tumor Detection in Brain MRI Images	Deep CNN (VGG-16)	96%
Cinar & Yildirim (2020)	Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture	Hybrid CNN (Resnet-50)	97.01%
Proposed Method	Brain Tumor Classification on Magnetic Resonance Imaging Images using Convolutional Neural Network with Cycle Generative Adversarial Network and Extreme Gradient Boosting	CNN (VGG-19) + CycleGAN epoch 200 + XGBoost	97.37%

CONCLUSION

Implementation of CNN using VGG-19 architecture combined with CycleGAN and XGBoost can yield promising results. The use of CycleGAN to replace traditional CNN data augmentation techniques has a significant impact on the accuracy achieved by the CNN model. The implementation of CycleGAN can increase and balance the quantity of training data, indicating its ability to increase diversity in the dataset, thereby aiding in the training of the CNN model. Additionally, the use of XGBoost can improve the accuracy of the CNN classification model in several tests. However, there are instances where the implementation of the XGBoost method did not improve and even lowered the accuracy of CNN model. Therefore, careful configuration of certain CNN parameters is necessary to generate a well-performing combined model. In this study, a combined classification model with an accuracy of 97.37% was achieved with the configuration of CycleGAN at 200 epochs and CNN at 10 epochs.

REFERENCES

[1] C. Alexopoulos, Z. Lachana, A. Androutsopoulou, V. Diamantopoulou, Y. Charalabidis, and M. Avgerinos Loutsaris. "How Machine Learning is Changing e-Government," in *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance (ICEGOV '19), New York, NY, USA, April 2019*, pp. 354–363.

[2] J. Alzubi, A. Nayyar, & A. Kumar, "Machine Learning from Theory to Algorithms: An Overview," *J. Phys.: Conf. Ser.*, vol. 1142, December 2018

[3] S. Bhattacharya, P. K. R. Maddikunta, Q. V. Pham, T. R. Gadekallu, S. R. Krishnan, C. L. Chowdhary, M. Alazab, M. J. Piran, "Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey," *Sustain. Cities Soc.*, vol. 65, November 2020.

[4] H. U. Rehman, S. Masood, and M. A. Khan, "Brain Tumor Classification using Deep Learning Methods," *Int. J. Inf. Syst. Comput. Technol.*, vol. 1, no. 1, pp. 38-53, January 2022.

[5] D. C. Febrianto, I. Soesanti, and H. A. Nugroho, "Convolutional Neural Network for Brain Tumor Detection," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 771, March 2020

[6] K. Amalia, R. Magdalena, and S.Saidah, "Klasifikasi Penyakit Tumor Otak Pada Citra MRI Menggunakan Metode CNN Dengan Arsitektur Alexnet," *E-Proc. Eng.*, vol. 8, no. 6, pp. 3247-3254, December 2022.

[7] T. H. Nguyen, T. N. Nguyen, and B. V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," *AgriEngineering*, vol. 4, no. 4, pp. 871-887, October 2022.

[8] M. Shaha and M. Pawar, "Transfer learning for image classification," in *2018 second international conference on electronics communication and aerospace technology (ICECA), Coimbatore, India,* March 29-31, 2018. pp. 656-660.

[9] K. Yasaka, H. Akai, A. Kunimatsu, S. Kiryu, and O. Abe, "Deep learning with convolutional neural network in radiology," *Jpn. J. Radiol.*, vol. 6, no. 4, pp. 257-272, April 2018.

[10] Y. Chen, Y. Zhu, and Y. Chang, "CycleGAN Based Data Augmentation for Melanoma images Classification," in *Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition (AIPR '20), New York, NY, USA, June 26-28, 2020*, pp. 115-119.

[11] M. Hammami, D. Friboulet, and R. Kechichian, “Cycle GAN-Based Data Augmentation for Multi-Organ Detection in CT Images Via Yolo,” in *2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, October 25-28, 2020*, pp. 390-393

[12] E. D. Carvalho, E. D. Carvalho, A. O. de Carvalho Filho, F. H. D. de Araújo, and R. d. Andrade Lira Rabêlo, “Diagnosis of COVID-19 in CT image using CNN and XGBoost,” in *2020 IEEE Symposium on Computers and Communications (ISCC), Rennes, France, July 7-10, 2020*, pp. 1-6.

[13] J. Seetha and S. S. Raja, “Brain Tumor Classification Using Convolutional Neural Networks,” *Biomed. Pharmacol. J.*, vol. 11, no. 3, pp. 1457-1461, September 2018.

[14] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, “Deep learning for brain tumor classification,” in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging 2017, Orlando, Florida, USA*, February 11-16, 2017.

[15] V. Sandfort, K. Yan, P. J. Pickhardt, and R. M. Summers, “Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks,” *Sci. Rep.*, vol. 9, no. 1, November 2019.

[16] E. Irmak, “Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework,” *Iran. J. Sci. Technol. Trans. Electr. Eng.*, vol. 45, pp. 1015-1036, April 2021.

[17] S. Thongsuwan, S. Jaiyen, A. Padcharoen, P. Agarwal, “ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost,” *Nucl. Eng. Technol.*, vol. 53, no. 2, pp. 522-531, February 2021.

[18] A. Septiarini, R. Saputra, A. Tejawati, M. Wati, Hamdani, and N. Puspitasari, “Analysis of Color and Texture Features for Samarinda Sarong Classification” in *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia,* December 16, 2021, pp. 102-107.

[19] J. Sofian and R. H. Laluma, “Klasifikasi Hasil Citra MRI Otak untuk Memprediksi Jenis Tumor Otak Dengan Metode Image Threshold dan GLCM Menggunakan Algoritma K-NN (Nearest Neighbor) Classifier Berbasis Web,” *Infotronik J. Teknol. Inf. dan Elektron.*, vol. 4, no. 2, pp. 51-56, December 2019.

[20] R. R. Wahid, F. T. Anggraeni, B. Nugroho, “Implementasi Metode Extreme Learning Machine untuk Klasifikasi Tumor Otak pada Citra Magnetic Resonance Imaging,” *Prosiding Seminar Nasional Informatika Bela Negara*, vol. 1, pp 16-20, November 2020.

[21] M. A. B. Siddique, S. Sakib, M. M. R. Khan, A. K. Tanzeem, M. Chowdhury, and N. Yasmin, “Deep Convolutional Neural Networks Model-based Brain Tumor Detection in Brain MRI Images,” in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India*, October 7-9, 2020, pp. 909-914.

[22] A. Çinar and M. Yildirim, “Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture,” *Med. Hypotheses*, vol. 139, June 2020.