Comparison of Convolution Neural Network Architecture for Colon Cancer Classification

https://doi.org/10.3991/ijoe.v18i03.27777

Indrarini Dyah Irawati[1], Ivosierra Andrea Larasaty[2], Sugondo Hadiyoso[1](✉)
[1]School of Applied Science, Telkom University, Bandung, Indonesia
[2]School of Electrical Engineering, Telkom University, Bandung, Indonesia
sugondo@telkomuniversity.ac.id

Abstract—In 2021, colon cancer is the second most common cause of death for this type of cancer. Therefore, in this study, a colon cancer classification system was developed to help medical staff classify 2 types of cancer colon adenocarcinomas and benign colonic tissues. The classification method uses the Convolution Neural Network (CNN) with the architecture VGG16, VGG19, ResNet101, ResNet152, MobileNetV2, DenseNet201 and InceptionV3. We used 10.000 image datasets that divided into 7200 training data, 1800 validation data and 1000 test data. Pre-trained models are used to extract new features and training data. The best performance parameter based on accuracy, precision, recall and f1-score and confusion matrix are obtained in 3 architectures, namely VGG19, ResNet101 and ResNet152. These architectures can identify and classify both types of colon cancer with 100% accuracy.

Keywords—classification, colon cancer, CNN, pre-trained network

1 Introduction

According to WHO, cancer is a non-contagious disease that causes death worldwide [1], [2]. Generally cancer cells can multiply out of control [3]. One of the cancers with a high prevalence of around 10% is colon cancer or colorectal cancer (CRC) of all cases in the world [4]. Colon cancer is the second most deadly disease [5]. The proliferation of cancer cells is so fast that it requires a health system that can detect cancer early, and get treatment as early as possible [6]. In diagnosing colorectal cancer, clinical pathologists visually examine colon surface tissue samples using a microscope. To support analysis, tissue samples were resected, fixed and stained using Hematoxylin and Eosin (H&E) [7], [8]. From previous studies, it has been proven that the degree of malignancy can also be analyzed accurately and effectively using this technique. Thus, early detection can also be done. However, visual evaluation requires a lot of effort and is time consuming. The expertise and experience of the clinician also greatly determines the accuracy of detection. Therefore, it is important to develop a decision support system for automatic classification of colon cancer from tissue images.

In recent years, research on the detection and classification of cancer tissue based on clinical pathology, has received great attention. Classification based on genetic profile is reported in [9]–[11]. However, detection techniques using genetic analysis tend to be high cost. Some other studies use traditional machine learning by performing feature extraction to get a feature vector which is then trained and validated as reported in [12]–[14]. Nevertheless, traditional machine learning will be time consuming if applied to large datasets, besides variations in image background, image source, size and color of images can reduce accuracy. Another issue about traditional machine learning techniques that needs to be reduced is human intervention [15]. Thus, an alternative method is needed to tackle this problem. One method that has high potential in image classification is convolutional neural network (CNN) and deep learning [15], [16].

In most cases of classification and detection, convolutional neural network shows significant performance. Recently, CNN classification of lung cancer based on computer tomography (CT) image has been reported in [17]–[20]. Study by Saleh [17], combined CNN-SVM in CT image-based lung cancer detection and succeeded in generating 97.91% accuracy. Meanwhile, similar classification cases have been reported in [18], [19], generate 93.54% and 100% accuracy respectively. A lung cancer identification system based on CT images that produces high accuracy is also proposed in [20], which generates accuracy up to 98.52%. Another significant study on the use of CNN in the identification and classification of Alzheimer's in MRI images also demonstrated high performance as reported in [21]–[23]. The application of CNN in the detection of brain tumors on MRI images is reported in [24]. By considering CNN's performance in several literature reviews for various image sources, it gives an idea that CNN has the potential to classify colon cancer. Therefore, in this study, CNN-based system architecture is proposed for the classification of colon cancer. The developed system was tested on various CNN architectures to get the best performance.

As a reminder, this paper consists of material and method are described in Section 2. The detail of results and followed by discussion will be described in Section 3. The study will be concluded in Section 4 with implication and future works.

2 Material and methods

2.1 Dataset

This study uses a dataset taken from [25]. The dataset is an image of colon cancer with dimensions of 768×768 pixels in jpeg format. There are 10.000 images of colon cancer, which are categorized into two classes, such as 5.000 images of adenocarcinoma (ACA) and 5.000 images of benign colonic tissue (N). The image dataset is divided into 7200 training data, 1800 validation data and 1000 test data.

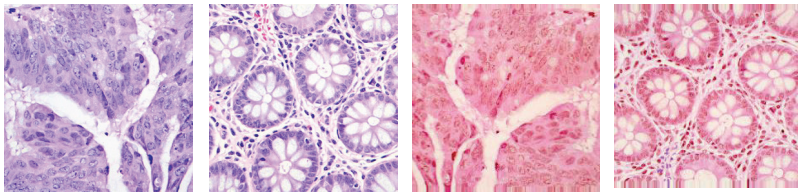
In the preprocessing step, the image dimensions are changed to 224×224 pixels. Then followed by the shuffle technique to randomize the order of the training images. The train test split technique is used to sort the images, as well as adding a static count feature to maintain the number of output test during testing. The training and test image data types were also changed from uint8 type to float32 with a range of 0 to 1. The test data labels which were previously in the form of class name labels were changed to

categorical numeric using the default library from Keras TensorFlow. The dataset uses the augmentation system which parameters shown in Table 1.

Table 1. Image augmentation parameter

Parameter	Value
Horizontal & Vertical flip	0.5
Zoom Range	0.1
Brightness	Brightness 0.4 Probability 0.75
Image Size	224 × 224
Color mode	RGB
Class mode	Categorical
Seed & Random State	Seed = 42 Random State = 42
Validation Split	0,2

Figure 1 shows the original and an augmented images of colon cancer with the specified augmentation parameters.



a) Original ACA **b) Original N** **c) Augmented ACA** **d) Augmented N**

Fig. 1. Original and augmented colon image

2.2 CNN architectures

The general approach of computer vision to build accurate models in a time efficient manner is by CNN transfer learning model [26]. Transfer learning is a solution to the problem of large data collection, which the collection process does not start with learning from the beginning, but simply uses the previously learned model. In this research, we use 7 transfer learning models such as: VGG16, VGG19, ResNet101, ResNet152, MobileNetV2, DenseNet201 and InceptionV3, were pre-trained on the ImageNet. Testing several architectures aims to test the robustness of the proposed model.

2.3 Performance parameter

The performance parameters to evaluate the CNN classification model are accuracy, precision, recall and F1-score, and confusion matrix. The formula of the parameters shown as follows [27]:

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

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

$$F1 - score = \frac{2 * Precison * Recall}{Precison + Recall} \tag{3}$$

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

True Positive (TP) is a prediction of the complementary class as a positive class number. True Negative (TN) is a negative class prediction as a hostile class number. False Positive (FP) is a negative class prediction as a positive class number. False Negative (FN) is a prediction of the complimentary class as a hostile class number.

3 Results and discussion

The simulation of the system is run on google colab pro using TPU v2 with 64 GB High Bandwidth Memory (HBM). To avoid overfitting, several parameters were added including a learning rate of 0.00001, with 755 decay steps, 0.9 decay rate and using Adam's optimization optimizers. Then for the modeling made using the ReLU and Softmax activation functions, then for the dense layer using the value 512. Next for the training process using batch size 32 with epochs of 10.

Table 2. Performance parameter

Architecture	Performance Parameter			
	Prec.	Recall	F1-score	Acc.
VGG16	0.99	0.99	0.99	0.99
VGG19	1.00	1.00	1.00	1.00
ResNet101	1.00	1.00	1.00	1.00
ResNet152	1.00	1.00	1.00	1.00
MobileNetV2	0.96	0.96	0.96	0.96
DenseNet201	0.93	0.93	0.92	0.93
InceptionV3	0.89	0.89	0.89	0.89

Performance parameters of precision, recall, F1-score, and accuracy are reported in Table 2. This test was carried out using seven CNN architectures such as VGG16, VGG19, ResNet101, ResNet152, MobileNetV2, DenseNet201, IceptionV3. Each

architecture produces adequate results with almost the same value. For all parameters architecture, VGG19, ResNet101 and ResNet152 have the best performance with a perfect value of 100%. Followed by the closest is VGG16 then MobileNetV2 and DenseNet201. The worst performance is InceptionV3 with values for all performance parameters in below 0.89. Figure 2 expresses confusion matrix of each pre-trained architecture. The value in the row represents the actual class, the column represents the predicted class. Based on these results, the architecture with the best performance in determining the classification class are VGG19, ResNet101 and ResNet152.

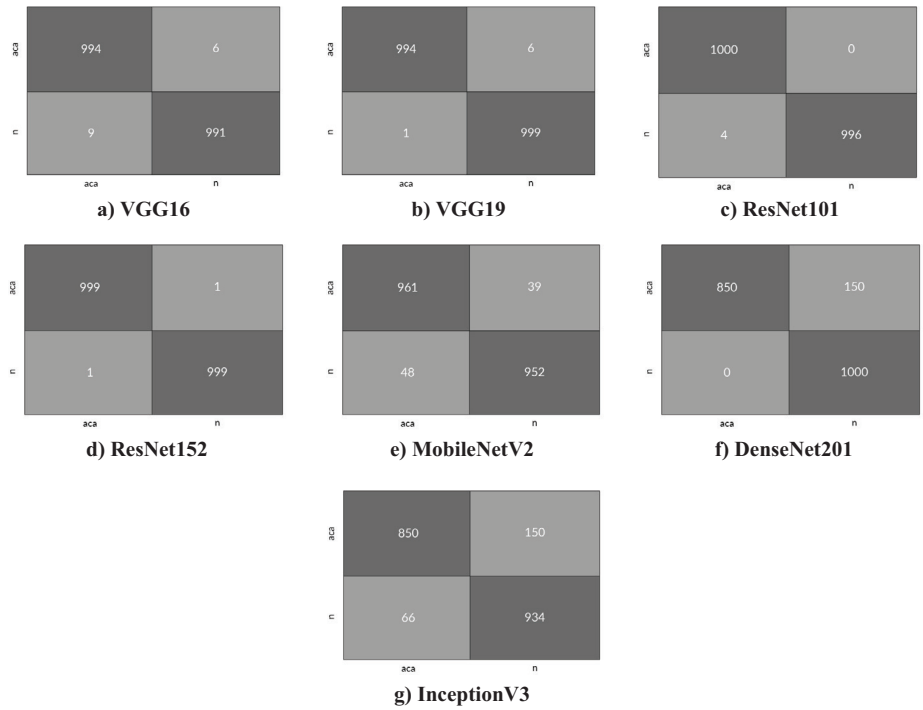
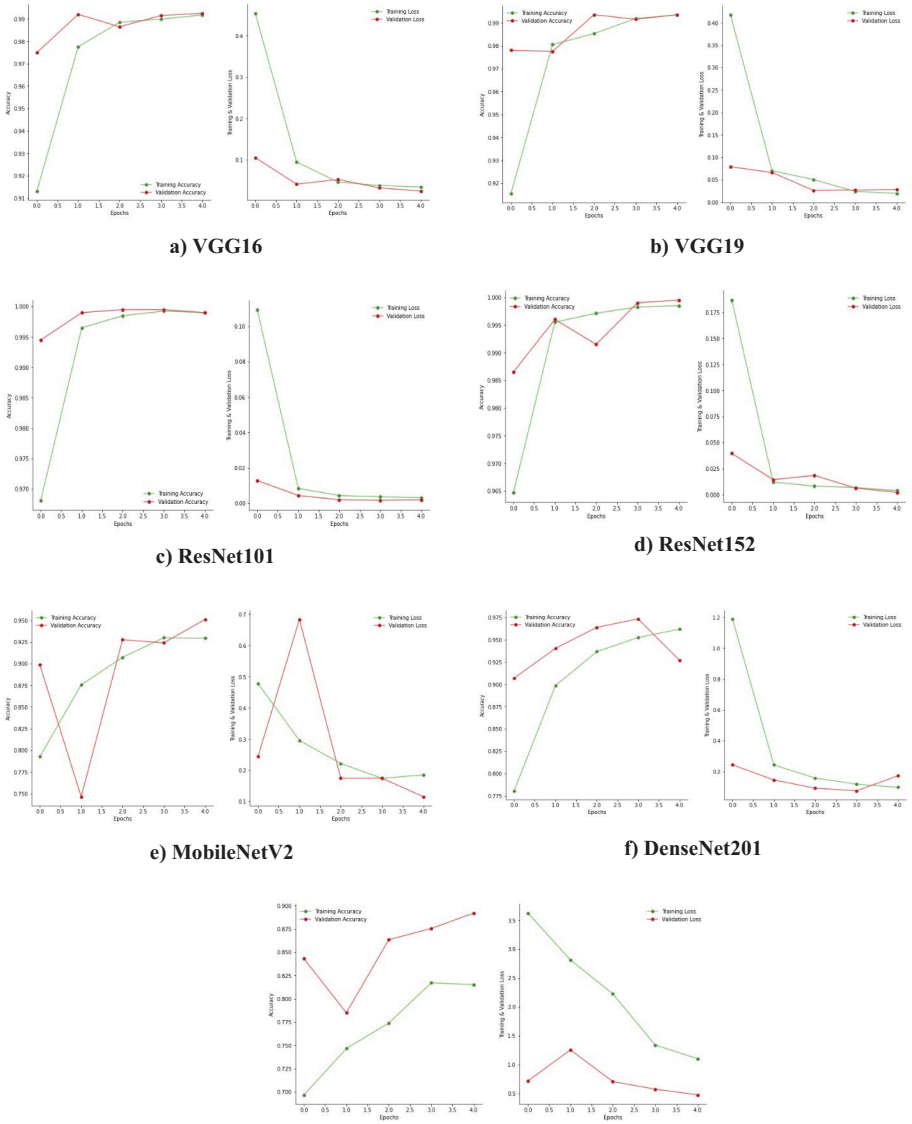


Fig. 2. Confusion matrix

Simulation performance is also shown in the learning curve which is one of the parameters in deep learning. The learning curve displays performance during training by always updating during the training process. The model is evaluated with both training data sets and validation data. Figure 3 shows the training and validation results of each pre-trained CNN architecture. The figure shows the learning curve and validation curve of the accuracy function against the epoch which determines the number of learning algorithms that will work to process the entire training dataset. In all pre-trained models, the higher the epoch, the greater the accuracy. Meanwhile, the curve between the loss function and the epoch aims to describe the prediction error of the model's performance. Based on the figure, the more epochs, the loss curve tends towards zero. In epoch 4, the model VGG19, ResNet101, ResNet152 result the highest accuracy and

the lowest loss. Accuracy obtained reaches 100%. The proposed method outperform compared to our previous study using the KNN technique [28].



a) VGG16 **b) VGG19**

c) ResNet101 **d) ResNet152**

e) MobileNetV2 **f) DenseNet201**

g) InceptionV3

Fig. 3. Training and validation results

4 Conclusion

Colon cancer is one type of cancer with a high prevalence and the leading cause of death. Early detection is the best way to arrange the right therapy. Early detection can be done by examining clinical pathology on microscopic tissue images. However, visual observation is often difficult and tends to be subjective depending on the expertise of the clinician. Therefore, computer-based analysis is needed to support the diagnosis. In this study, we have successfully demonstrated an automatic classification system for detecting colon cancer. The system developed is based on CNN with various architectures including VGG16, VGG19, ResNet101, ResNet152, MobileNetV2, DenseNet201 and InceptionV3 to get the best performance. A total of 1800 validation data and 1000 test data were used to test the robustness of the proposed system. The implementation results show that the proposed system can generate an average accuracy of up to 100%. The best performance is achieved using the VGG19, ResNet101 and ResNet152 architectures. With this proposed system, it is hoped that it can support clinical diagnosis in addition to visual observation. This system can also be used for screening large populations. Future work is to apply this system to assess cancer severity in a larger and more varied source dataset.

5 References

[1] A. Damasceno, *Noncommunicable Disease*. 2016. https://doi.org/10.1002/9781119097136.part5

[2] A. Boutayeb and S. Boutayeb, "The burden of non communicable diseases in developing countries," *Int. J. Equity Health*, vol. 4, pp. 1–8, 2005. https://doi.org/10.1186/1475-9276-4-2

[3] S. F. Ameer, Z. T. Nayyef, Z. H. Fahad, and I. R. N. AlRubee, "Using morphological operation and watershed techniques for breast cancer detection," *Int. J. online Biomed. Eng.*, vol. 16, no. 5, pp. 140–149, 2020. https://doi.org/10.3991/ijoe.v16i05.12999

[4] M. Egeblad, E. S. Nakasone, and Z. Werb, "Tumors as organs: complex tissues that interface with the entire organism," *Dev Cell.*, vol. 18, no. 6, pp. 884–901, 2011. https://doi.org/10.1016/j.devcel.2010.05.012

[5] Recio-Boiles and Cagir, *Colon Cancer*. 2021.

[6] M. Hamim, I. El Moudden, M. D. Pant, H. Moutachaouik, and M. Hain, "A hybrid gene selection strategy based on fisher and ant colony optimization algorithm for breast cancer classification," *Int. J. online Biomed. Eng.*, vol. 17, no. 2, pp. 148–163, 2021. https://doi.org/10.3991/ijoe.v17i02.19889

[7] J. P. Hinton et al., "A method to reuse archived H&E stained histology slides for a multiplex protein biomarker analysis," *Methods Protoc.*, vol. 2, no. 4, pp. 1–15, 2019. https://doi.org/10.3390/mps2040086

[8] B. S. Restall et al., "Virtual hematoxylin and eosin histopathology using simultaneous photoacoustic remote sensing and scattering microscopy," *Opt. Express*, vol. 29, no. 9, pp. 1–12, 2021. https://doi.org/10.1364/OE.423740

[9] M. L. R. AbdElNabi, M. Jasim, H. M. EL-Bakry, M. H. N. Taha, N. E. M. Khalifa, and 1, "Breast and colon cancer classification from gene expression profiles using data mining techniques," *Symmetry (Basel).*, vol. 12, pp. 1–15, 2020. https://doi.org/10.3390/sym12030408

[10] M. A. Fahami, M. Roshanzamir, N. H. Izadi, V. Keyvani, and R. Alizadehsani, "Detection of effective genes in colon cancer: A machine learning approach," *Informatics Med. Unlocked*, vol. 24, no. April, pp. 1–11, 2021. https://doi.org/10.1016/j.imu.2021.100605

[11] M. Al-Rajab, J. Lu, and Q. Xu, "A framework model using multifilter feature selection to enhance colon cancer classification," *PLoS One*, vol. 16, no. 4 April, pp. 1–26, 2021. https://doi.org/10.1371/journal.pone.0249094

[12] F. Keskin, A. Suhre, K. Kose, T. Ersahin, A. E. Cetin, and R. Cetin-Atalay, "Image classification of human carcinoma cells using complex wavelet-based covariance descriptors," *PLoS One*, vol. 8, no. 1, pp. 1–10, 2013. https://doi.org/10.1371/journal.pone.0052807

[13] C. Sindhu, S. Subhashini, T. Swathi, and G. S. S, "colorectal cancer detection using image processing techniques: A knowledge transfer perspective," *AJAST*, vol. 2, no. 2, pp. 1–9, 2018.

[14] A. M. Sarhan, "A novel lung cancer detection method using wavelet decomposition and convolutional neural network," *J. Biomed. Sci. Eng.*, vol. 13, no. 05, pp. 81–92, 2020. https://doi.org/10.4236/jbise.2020.135008

[15] M. J. Tsai and Y. H. Tao, "Deep learning techniques for the classification of colorectal cancer tissue," *Electron.*, vol. 10, no. 14, pp. 1–26, 2021. https://doi.org/10.3390/electronics10141662

[16] A. Elnakib, H. M. Amer, and F. E. Z. Abou-Chadi, "Early lung cancer detection using deep learning optimization," *Int. J. online Biomed. Eng.*, vol. 16, no. 6, pp. 82–94, 2020. https://doi.org/10.3991/ijoe.v16i06.13657

[17] A. Y. Saleh, C. K. Chin, V. Penshie, and H. R. H. Al-Absi, "Lung cancer medical images classification using hybrid cnn-svm," *Int. J. Adv. Intell. Informatics*, vol. 7, no. 2, pp. 151–162, 2021. https://doi.org/10.26555/ijain.v7i2.317

[18] A. Khan and Z. Ansari, "identification of lung cancer using convolutional neural networks based classification," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 10, pp. 192–203, 2021.

[19] H. F. Al-Yasriy, M. S. Al-Husieny, F. Y. Mohsen, E. A. Khalil, and Z. S. Hassan, "Diagnosis of lung cancer based on CT scans using CNN," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 928, no. 2, pp. 1–10, 2020. https://doi.org/10.1088/1757-899X/928/2/022035

[20] S. H. M. Mohammed and A. Çinar, "Lung cancer classification with convolutional neural network architectures," *Qubahan Acad. J.*, vol. 1, no. 1, pp. 33–39, 2021. https://doi.org/10.48161/qaj.v1n1a33

[21] J. Bin Bae et al., "Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging," *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, 2020. https://doi.org/10.1038/s41598-020-79243-9

[22] W. Lin et al., "Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment," *Front. Neurosci.*, vol. 12, no. NOV, pp. 1–13, 2018. https://doi.org/10.3389/fnins.2018.00777

[23] F. E. K. Al-Khuzaie, O. Bayat, and A. D. Duru, "Diagnosis of Alzheimer disease using 2D MRI slices by convolutional neural network," *Appl. Bionics Biomech.*, vol. 2021, no. ii, pp. 1–9, 2021. https://doi.org/10.1155/2021/6690539

[24] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Sci. Rep.*, vol. 11, no. 1, pp. 1–17, 2021. https://doi.org/10.1038/s41598-021-90428-8

[25] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A DeLand, S. M. Mastorides. "Lung and colon cancer histopathological image dataset," (LC25000). arXiv:1912.12142v1 [eess.IV], 2019

[26] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng*., vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191. https://doi.org/10.1109/TKDE.2009.191

[27] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 01–11, 2015. https://doi.org/10.5121/ijdkp.2015.5201

[28] I. D. Irawati, S. Hadiyoso, and A. Fahmi, "Compressive sensing in lung cancer images for telemedicine application," in 2021 International Conference on Electronics, Communications and Control Engineering, ICECC, 2021. https://doi.org/10.1145/3462676.3462685

6 Authors

Indrarini Dyah Irawati obtained a doctoral degree in the School of Electrical and Information Engineering, Institute of Technology Bandung. She joined Telkom Applied Science School, Telkom University as an Instructor (2007–2019), Associate Professor (2019-present). Her main research interests are in the areas of compressive sensing, watermarking, signal processing, and computer network. She is currently a member of the Association for Computing Machinery (ACM) and the International Association of Engineers (IAENG) (email: indrarini@telkomuniversity.ac.id).

Ivosierra Andrea Larasaty is currently a Telecommunication Engineering undergraduate student in Telkom University, Bandung, Indonesia. She is active as a research assistant at the Digital Signal Processing Research Laboratory at Telkom University, as a programmer for Maltab and python as well as in writing various scientific journals. Her primary research interests are in the field of signal processing, networking and IoT (email: andrea.ivosierra@gmail.com).

Sugondo Hadiyoso received the Master.in Electrical-Telecommunication Engineering from Telkom University, Bandung, Indonesia in March 2012. His research interests are wireless sensor network, embedded system, logic design on FPGA and biomedical engineering. In 2018-present, he became a doctoral student in electrical engineering at the Bandung Institute of Technology. The focus of his doctoral research is signal processing and analysis of EEG waves (email: sugondo@telkomuniversity.ac.id).

Article submitted 2021-10-20. Resubmitted 2021-11-19. Final acceptance 2021-11-23. Final version published as submitted by the authors.