Available *online* at : http://jnte.ft.unand.ac.id/

Jurnal Nasional Teknik Elektro

| ISSN (Print) 2302-2949 | ISSN (Online) 2407-7267 |

Model Design of the Image Recognition of Lung CT scan for COVID-19 Detection Using Artificial Neural Network

Maulana Akbar Dwijaya [1], Umar Ali Ahmad [1], Rudi Purwo Wijayanto [2], Ratna Astuti Nugrahaeni [1]

[1] *Department of Computer Engineering, School of Electrical Engineering, Telkom University, Bandung, Indonesia*
[2] *Agency for the Assessment and Application of Technology (BPPT), PUSPIPTEK, South Tangerang, Indonesia*

ARTICLE INFORMATION

Received: November 17, 2021
Revised: February 27, 2022
Available online: March 29, 2022

KEYWORDS

COVID-19, CT scan of the lungs, digital image processing, artificial neural networks

CORRESPONDENCE

Phone: +62 811-2341-066
E-mail: umar@telkomuniversity.ac.id

A B S T R A C T

COVID-19 has become a pandemic and is a big problem that needs to be checked out immediately. CT scan images can explain the lung conditions of COVID-19 patients and have the potential to be a clinical diagnostic tool. In this research, we classify COVID-19 by recognizing images on a computer tomography scan (CT scan) of the lungs using digital image processing and GLCM feature extraction techniques to obtain grayscale level values in CT images, followed by the creation of an artificial neural network model. So that the model can classify CT scan images, the results in this research obtained the most optimal model for COVID-19 classification performance with 90% accuracy, 88% precision, 91% recall, and 90% F1 score. This research can be a useful tool for clinical practitioners and radiologists to assist them in the diagnosis, quantification, and follow-up of COVID-19 cases.

INTRODUCTION

By following the times, research using digital image processing (image processing) is widely applied in various fields, one of which is in the health or biomedical field. In health or biomedical, image processing is usually used to detect a disease from a person. COVID-19 (Coronavirus Disease 2019) has caused a pandemic and has become a Public Health Emergency of international concern (PHEIC). This virus was identified in Wuhan, China, in December 2019 with the emergence of a group of patients with pneumonia of unknown cause [1].

In the COVID-19 pandemic, there is an urgent need to detect accurate and quick classification of COVID patients. Currently, the most used method is an invasive method for patients, such as RT-PCR with a nasal swab [2]. In digital image processing technology in medicine, computed tomography (CT) is the non-invasive technology of choice because internal porosity detection and characterization is considered a promising technique [3]. If the patient has symptoms of COVID-19 such as fever, cough, and wheezing, a chest X-ray or CT scan may be done to diagnose COVID-19 infection with high sensitivity.

Computed tomography (CT) scan technology tests serve as a screening and diagnostic function to detect people with COVID-19. CT chest imaging of the lungs showed that numerous small plaques and interstitial changes occurred, then developed into multiple opacity and infiltrated both lungs, lung consolidation, or even "white lung" in some severe cases [4]. CT scans' most common abnormal findings are bilateral surrounding ground-glass opacities (GGO), consolidation with disease progression, and crazy-paving patterns [5].

Ground glass opacity (GGO) sign is scattered throughout the lungs. They represent alveoli are filled with fluid and turn gray on a CT scan. In more advanced infections, the appearance of glassy opacity results in solid white consolidation as more fluid accumulates in the lung lobes. Crazy paving stone pattern due to nodule of the interstitial space along the lung wall.

The digital imaging method of computed tomography (CT scan) for those infected with COVID-19 can be recognized from the CT scan image features of the COVID-19 lungs. In this research, the author designed and analyzed an image recognition model that can detect COVID-19 using a CT scan image of a patient's lungs. The image recognition used in this research uses an artificial neural network algorithm using different activation functions on

https://doi.org/10.25077/jnte.v11n1.984.2022

the hidden layer parameters to find the best activation function performance.

Gray Level Co-Occurate Matrix

To extract secondary features of texture from an image, GLCM (Gray Level Co-occurrence Matrix) can be used to show adjacency between pixels in the image at different orientations θ and at a spatial distance d. In doing texture analysis, a matrix is known to be powerful [6]. The GLCM matrix in the image a [y, z] is a double matrix [y, z], and the probability of appearing at a distance d and intensity levels y and z at a particular angle θ is represented by each element of the matrix. If the dimension of the GLCM matrix is [L+1] × [L+1], the maximum intensity value of the image is L. Generally, four directions are used to create a GLCM matrix. That is, for the direction of the angle θ = 0°, 45°, 90°, and 135°, a GLCM matrix exists for each value selected from the distance d and the angle θ [7].

In using GLCM in feature extraction from an image, the first thing to do is to change image type to grayscale, do quantization to create a GLCM matrix by converting a grayscale value (0 to 255), followed by co-occurrence which is the number of occurrences a level quantization with the same intensity value as adjacent pixels between two pixels values with one other pixel intensity level within a certain distanced and at the GLCM angle orientation θ. After obtaining the co-occurrence matrix, change the matrix into a symmetrical matrix, and normalize the symmetrical matrix. A way to normalize a GLCM matrix consists of dividing each element of the matrix by the sum of all its components, using the result of normalization of the GLCM matrix to perform a second-order extraction. The feature extraction of the GLCM process can be seen in a flowchart in Figure 1.

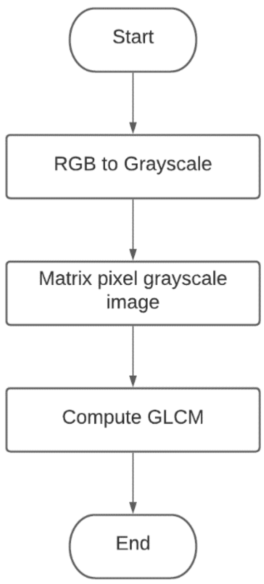


Figure 1. Feature Extraction Process of GLCM

Figure 1 shows the feature extraction process flow for the GLCM method, which starts by converting the RGB image to a grayscale image, and then creating a matrix with the provisions of the GLCM rules such as finding the value of the intensity occurrence between 2 pixels with GLCM distance and angle. The value of the intensity occurrence between the 2 pixels of the normalized GLCM matrix is calculated in the pre-processing stage as the contrast, entropy, homogeneity, and energy values. The work area values in the GLCM features are reference pixel values in the GLCM matrix (i), neighboring pixel values in the GLCM matrix (j), and values in rows i and j of the GLCM matrix (P), which can be seen in Figure 2.

j \ i	j_1	j_2	j_3	j_4
i_1	$P_{i_1 j_1}$	$P_{i_1 j_2}$	$P_{i_1 j_3}$	$P_{i_1 j_4}$
i_2	$P_{i_2 j_1}$	$P_{i_2 j_2}$	$P_{i_2 j_3}$	$P_{i_2 j_4}$
i_3	$P_{i_3 j_1}$	$P_{i_3 j_2}$	$P_{i_3 j_3}$	$P_{i_3 j_4}$
i_4	$P_{i_4 j_1}$	$P_{i_4 j_2}$	$P_{i_4 j_3}$	$P_{i_4 j_4}$

Figure 2. The Work Area of GLCM Features

A. *Contrast*

In an area of an image, there are grayscale level values in which the difference can be measured using a feature value named contrast. The distribution of the image intensity value can be seen in contrast value in the form of magnitude. The lighting level strongly influences the contrast value in an image in image capture [7]. The formula in Eq. (1) can be used to calculate contrast in the GLCM matrix.

$$Contrast = \sum_i \sum_j (i-j)^2 \; P_{(i,j)} \tag{1}$$

B. *Entropy*

A feature value called entropy can be calculated to show the level of regularity of structure and texture in the image. Value of entropy value is large if the image has both regular structure and texture. Otherwise, the entropy will be small for images with irregular structure and texture [7]. Entropy in the GLCM matrix can be calculated using Eq. (2).

$$Entropy = \sum_i \sum_j P_{(i,j)} \; log \, P_{(i,j)} \tag{2}$$

C. *Homogeneity*

Measurements on similarity can be done using Homogeneity. It is also known as Inverse Difference Moment (IDM), which means that measurements of the distribution of the GLCM elements compared to the diagonal GLCM will give the same intensity value in the image. Homogeneity will be of high value if all pixels have a uniform value [7]. Eq. (3) is a formula for calculating homogeneity in the GLCM matrix.

$$Homogeneity = \sum_i \sum_j \frac{P_{(i,j)}}{1+|i-j|} \tag{3}$$

https://doi.org/10.25077/jnte.v11n1.984.2022

D. Energy

A measure of uniformity in the image is represented by energy. Higher energy value means the higher the similarity in the image [8]. Energy is also called Angular Second Moment (ASM). If the pixel values are very similar, the energy value will be high [7]. Homogeneity will be of high value if all pixels have a uniform value. Eq. (4) is a formula for calculating energy in the GLCM matrix.

$$Energy = \sum_i \sum_j P_{(i,j)}{}^2 \quad (4)$$

Artificial Neural Network (ANN)

One of the areas of computer science concerned with creating intelligent systems that mimic human behavior is the notion of artificial neural network (ANN). ANN methods are increasingly attractive, efficient, effective, and successful in use cases for pattern recognitions in various problems [9][10]. It is a nonlinear computational system inspired by the brain's biological structure, behavior, and learning skills [11] in an artificial neural network, adaptive to the given input, which is carried out by the learning process to obtain knowledge with the relationship between elements.

Artificial neural networks have emerged as a better variant for solving complex human problems in a variety of disciplines. Therefore, a neural network is made up of three layers: input layer, hidden layer, and output layer [12]. The input layer receives the input feature from the user, and the hidden layer does all the intermediate calculations to get the required output from the received input. Finally, the output layer produces the result. For applications such as pattern recognition and classification, the efficiency depends primarily on the architectural algorithms neural networks adopted. The neuron is the basic component forming artificial neural network. Artificial neural networks include input layers, hidden and output layers, activation functions, learning algorithms, and weights. All layers are made up of neurons that need to be interconnected to create a network.

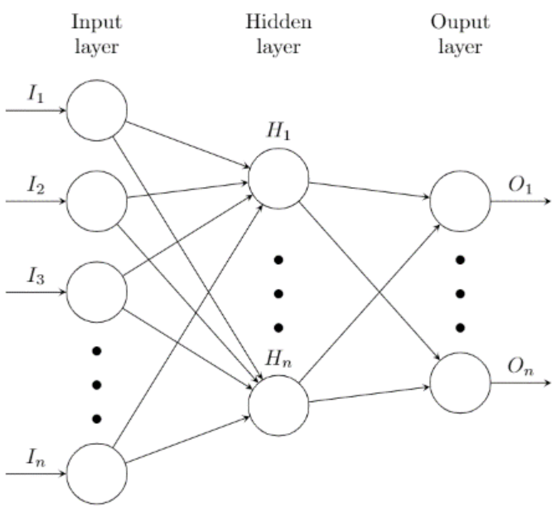


Figure 3. Neural Network

A. Activation function

The activation function has a main role in artificial neural networks that function to activate or deactivate neurons. The weighted input count is applied in the trigger ability to receive outgoing feedback. The activation function is an abstract function that represents the discharge rate in a cell. Neurons in the same layer must use the same activation function.

Sigmoid function is referred to as a logistic function in some publications [13]. Sigmoid is a nonlinear activation function used primarily in feedforward neural networks [14]. A real value will be taken as input, and the function outputs a finely derived value between 0 and 1, defined for a real input value, with a positive derivative. Eq. (5) is a formula for the sigmoid function.

$$f(x) = \left(\frac{1}{1+e^{-x}}\right) \quad (5)$$

Rectified Linear Unit (ReLU) function is the most widely used ReLU function for deep learning applications with the most recent research work. ReLU function proposed by Nair and Hinton 2010, is the most widely used for deep learning applications with the latest research work [15]. Because the representation of this function is like a linear function, it maintains the attributes of the linear function and can be easily optimized using the gradient-descent method [16]. For each input element with a value less than 0 or a negative value, this function will perform a threshold operation and set the value into 0. Eq. (6) is a formula for the ReLU function.

$$f(x) = max(0, x) = \begin{cases} x_i, & if\ x_i \geq 0 \\ 0, & if\ x_i < 0 \end{cases} \quad (6)$$

Leaky Rectifier Linear Unit (Leaky ReLU/LReLU) function proposed a slight negative slope activation function for ReLU in 2013. This is done to continuously update the weights throughout the propagation process [17]. The problem with the normalized ReLU activation function is that some neurons die during training. That is, the release of non-zero neurons stop. Neurons die when their weight changes significantly. Therefore, the total weight entry for all sets is negative. LReLU calculates gradients with minimal constant values for negative gradients in the range 0.01 [14]. LReLU function is created to solve the ReLU function problem. Eq. (7) is a formula for the LReLU function.

$$f(x) = ax + x = \begin{cases} x, & if\ x > 0 \\ ax, & if\ x \leq 0 \end{cases} \quad (7)$$

Exponential Linear Units (ELU) function proposed in 2015 by Clevert et al. This activation function is used to accelerate training in deep neural networks. ELU adjusts the slope of the negative part of the function. Since this function has a negative value, it can activate units with an average close to 0, reducing costs and increasing learning rates [18]. ELU pushes moderate unit activation towards zero during training, which reduces bias and makes ELU a good alternative to ReLU. Eq. (8) is a formula for ELU function.

$$f(x) = \begin{pmatrix} x, & if\ x > 0 \\ a\,(e^x - 1), & if\ x \leq 0 \end{pmatrix} \quad (8)$$

B. Adam optimization algorithm

Adam stands for Adaptive moment estimation. This method can be used for efficient stochastic optimization requiring only first-order gradients with minimal memory usage. For different parameters of the gradient's first and second moment estimates, individual adaptive learning rates are used for computation using this method [19]. Adam Optimization Algorithm is a popular optimization algorithm in artificial neural networks because it achieves good results in a short amount of time.

C. Dropout layer

In neural networks, dropout can overcome overfitting problems and speed up the learning process. Overfitting is a condition where a good percentage is reached for almost all data that has gone through the training process. However, a discrepancy in the prediction process or data validation which results in a value gap between the two processes, indicates the model results are not optimal. In its working system, the dropout layer is the method of dropping a neuron that is on the hidden layer in the artificial neural network with a single parameter named dropout rate. The neurons to be dropped are randomly selected [20].

METHOD

Figure 4 shows the proposed method in this research, which uses GLCM feature extraction and an artificial neural network algorithm to classify the result. First, the lung CT scan images will be pre-processed by converting RGB image to grayscale image and resizing the image size. Then some calculations will be done, such as the contrast in (1), entropy in (2), homogeneity in (3), and energy in (4) for each of the four directions of the angle θ GLCM, from each lung CT scan image. The results from GLCM extraction will be processed for classification using an artificial neural network algorithm between the COVID-19 label and non-COVID-19 label. The result of artificial neural network classification will be tested for acquiring the performance.

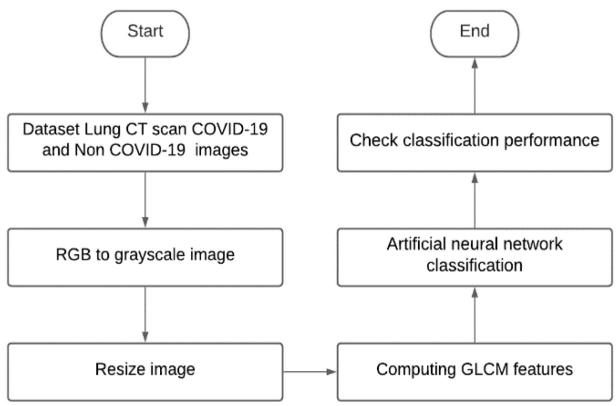


Figure 4. Detection of Lung COVID-19

Figure 5 is an architectural design of our model artificial neural network. The architecture consists of an input layer, four hidden layers, and an output layer. The input layer has the number of inputs according to the results of the 16 GLCM features. In the first hidden layer, there are 32 neurons, the second hidden layer has 512 neurons, the third hidden layer has 256 neurons, and the fourth hidden layer has 128 neurons. The dropout layer is in the hidden layer model. The output layer has one neuron.

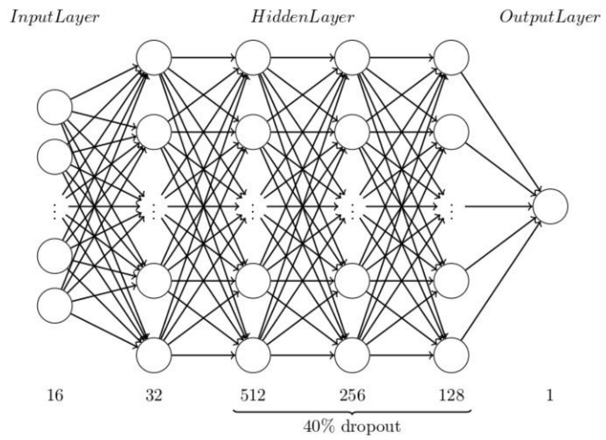


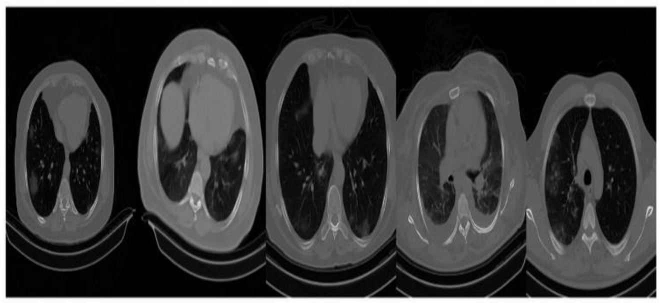
Figure 5. Architecture Our Model

The architecture of this artificial neural network model will be used to classify CT images with gray level feature values using GLCM as model input, and the model will compute the values according to the connected design model nodes to produce one value that will be worth a CT image between the COVID-19 label and non-COVID-19 label.

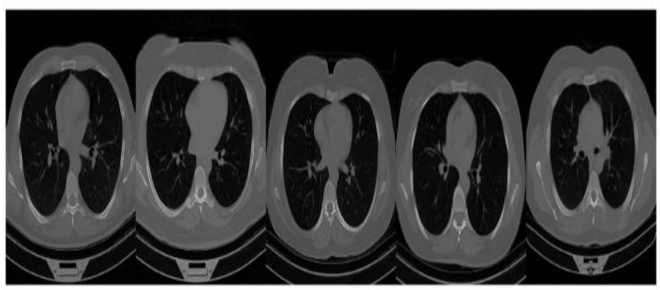
Data Collection and Pre-processing Images

A. Dataset

This experiment used 1000 CT scan images of the lungs consisting of 500 samples for COVID-19 and 500 samples for normal, sourced from research by M. Maftouni et al. [21].



(a)



(b)

Figure 6. CT Images; (a) COVID-19, and (b) Normal

https://doi.org/10.25077/jnte.v11n1.984.2022

B. *Pre-processing*

The pre-processing process includes two stages, namely converting RGB to grayscale image and resizing image into 128×128 pixels. The purpose is to make calculations possible during the feature extraction process.

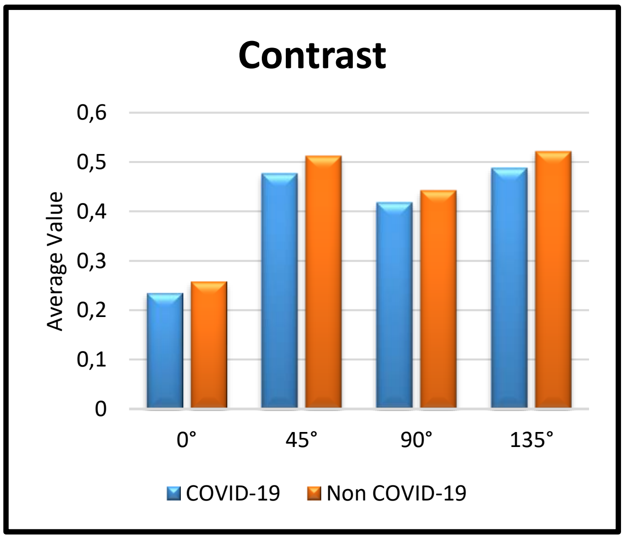
C. *Feature extraction*

Image feature extraction is the step of extracting feature information from the object you want to identify to distinguish it from other objects. After extraction, this attribute is used as an input parameter to distinguish objects in the classification stage. At this stage, all images, both COVID-19 and normal CT scans, each test will use four GLCM features extraction with each feature having four angles (0°, 45°, 90°, 135°). So, there is a total of 16 GLCM features for the classification process. GLCM feature extraction values are standardized with decimal scaling technique by dividing each column into the data by a number to the power of base 10. Eq. (9) is a formula for the decimal scaling technique. In Eq. (9), X is feature data values, j is the number of digits in the largest value of each feature value, and X_i is new feature data values.

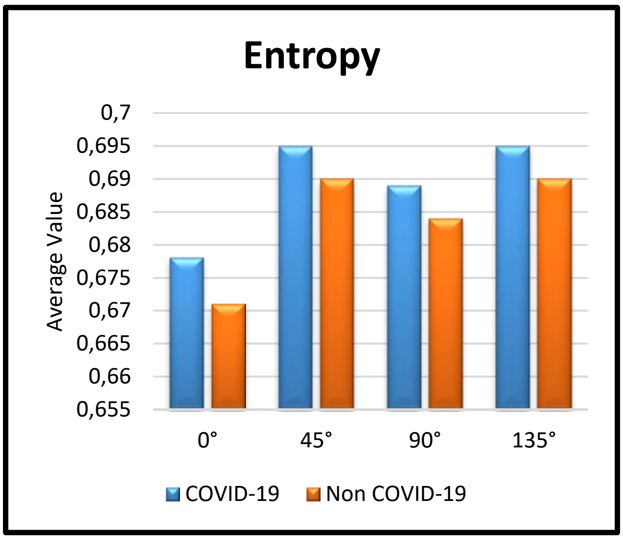
$$X_i = \frac{X}{10^j} \tag{9}$$

RESULTS AND DISCUSSION

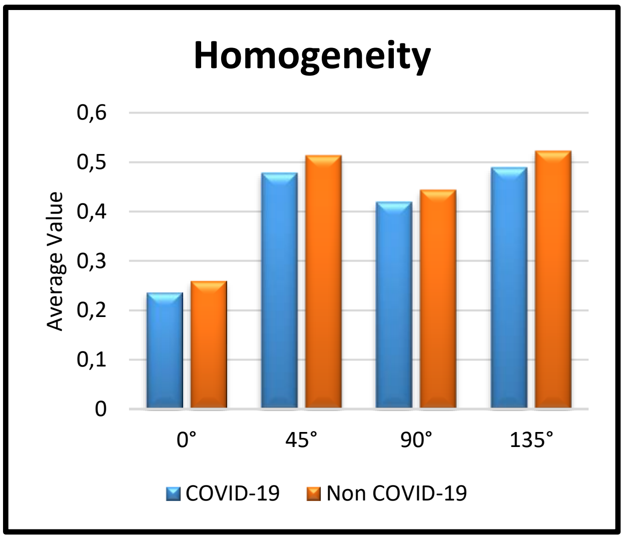
Experimentation is a process conducted to assess whether the designed model is as expected. This research also experimented with setting the model output to obtain optimum accuracy results. After images get the feature extraction values, Figure 7 compares the average GLCM feature values (contrast, entropy, homogeneity, and energy) between Covid-19 and non-Covid-19 images. These GLCM feature values will be used for further processing as input data in the artificial neural network to pattern recognition in image classification that will determine the image of Covid-19 and non-Covid-19 patients.



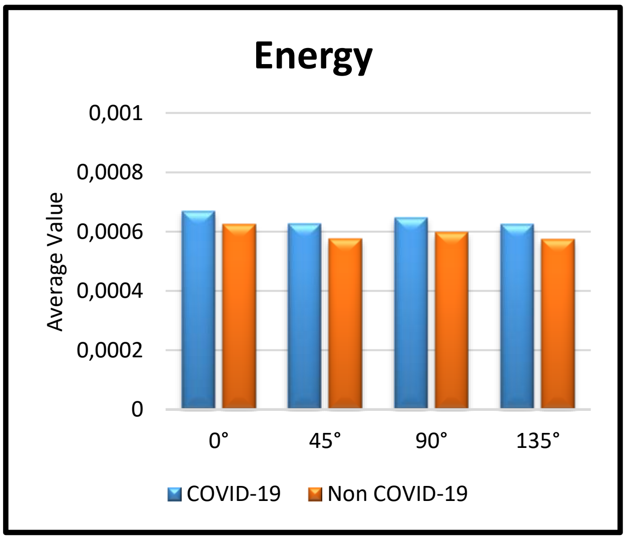
(a)



(b)



(c)



(d)

Figure 7. Result Features Values; (a) Contrast, (b) Entropy, (c) Homogeneity, and (d) Energy

Classifying Artificial Neural Network

The results feature extraction data will be split into two parts, training and testing. The classification in this research used 800 training data and 200 testing data, containing 98 images for COVID-19 and 102 for non-COVID-19 images. The results of the GLCM characteristic extraction will be classified using an artificial neural network with the model architectures in Figure 4 and the hyperparameter in Table 1. Trainable parameters are the number of elements/neurons that can be trained on the artificial neural network model architecture. Data samples assigned to the neural network model in batch, with the amount of data is based on batch size value. While to compute the weight correction value over time training process, a learning rate is needed. As for epoch, it is meant to train a neural network using all the training data in one cycle.

Table 1. Hyperparameter And Architecture Model

Factors	Values
Trainable parameters	181793
Maximum epochs	3000, 4000, 5000
Batch size	32
Optimizer	Adam
Learning rate	0.001
AF on hidden layer	ReLU, LReLU,ELU
AF on output layer	Sigmoid
Dropout rate	0.4

Table 2 shows the classification results using an artificial neural network to various epochs and activation function (AF) on the hidden layer. Here, TP is True Positive (the model detects COVID-19 is correct), TN is True Negative (the model detects COVID-19 is wrong), FN is False Negative (the model detects Non-COVID-19 is correct), and FP is False Positive (the model detects Non-COVID-19 is wrong).

From the overall data, to describe how accurate the model is in classifying the data correctly, accuracy can be used as it gives the proportion of correct predictions (positive and negative) to the overall data. In other words, it shows the level of closeness of the predicted value to the actual value. Eq. (10) shows the general formula to get accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

The level of accuracy can be determined by calculating precision, which shows the level of accuracy between the requested data and the prediction provided by the model. Comparing the ratio of positive true predictions and overall predicted results will give precision value of every positive class been predicted correctly. The general formula of precision is given in Eq. (11).

$$Precision = \frac{TP}{(TP+FP)} \tag{11}$$

Recall portrays the positive achievement of the model in retrieving information. Comparing the ratio of true positive predictions with all data that are true positive will give recall value. The general formula of recall is given in Eq. (12).

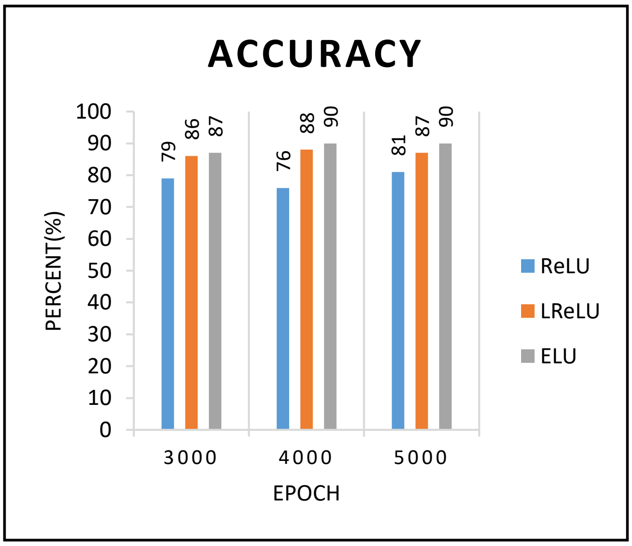
$$Recall = \frac{TP}{(TP+FN)} \tag{12}$$

The general formula of F1-score is given in Eq. (13), which can give a weighted comparison of given recall and average precision.

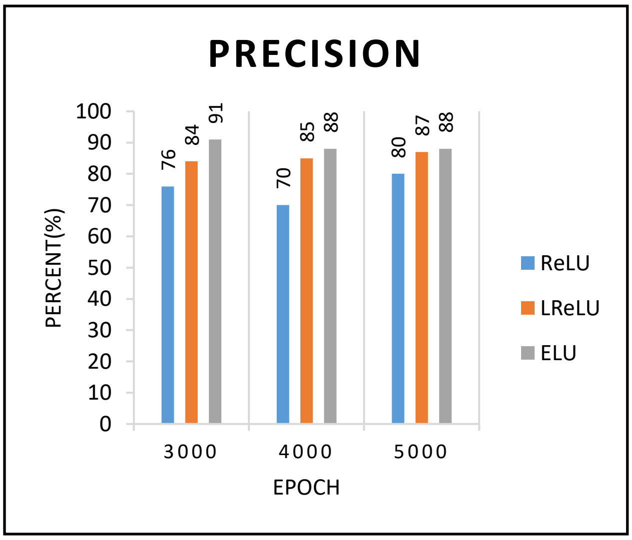
$$F1\text{–}score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

Table 2. Classification Results Of Various Epochs And Activation Function (AF) On Hidden Layer

Epoch	AF on hidden layer	TP	FP	TN	FN
3000	ReLU	74	24	84	18
	LReLU	82	16	90	12
	ELU	89	9	85	17
4000	ReLU	69	29	83	19
	LReLU	83	15	92	10
	ELU	86	12	93	9
5000	ReLU	77	21	98	4
	LReLU	85	13	89	13
	ELU	86	12	94	8

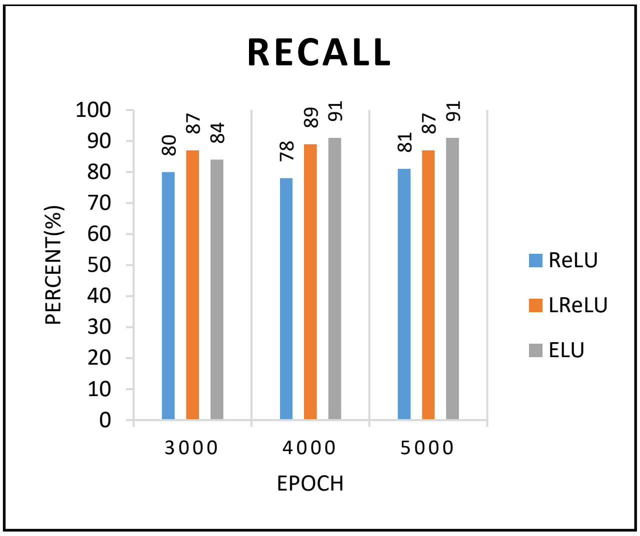


(a)

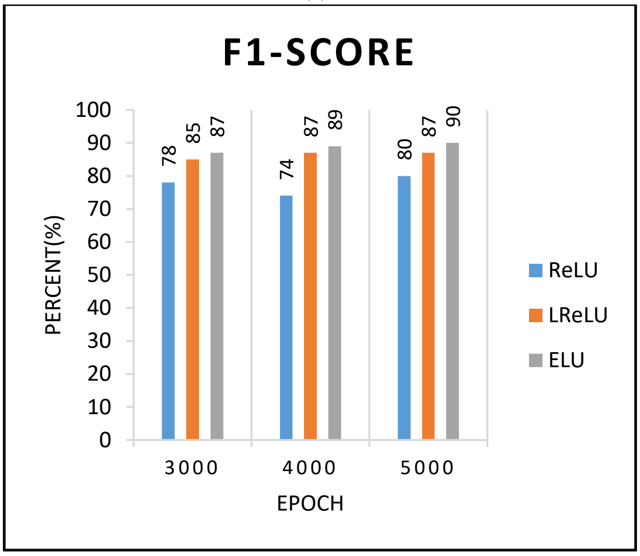


(b)

https://doi.org/10.25077/jnte.v11n1.984.2022



(c)



(d)

Figure 8. Result Performance; (a) Accuracy, (b) Precision, (c) Recall, and (d) F1-score

The testing has been carried out by 1000 sets of lung CT scan image data (800 training data and 200 test data). The result can be seen in Figure 8, namely by testing the dataset with various epochs and activation function (AF) on the hidden layer. The most optimum performance model is obtained with 5000 epochs and using the ELU activation function in the hidden layer with 90% accuracy, 88% precision, 91% recall, and 90% F1-score.

CONCLUSIONS

This article introduces an approach to modeling an artificial neural network by constructing our own model architecture using image processing features with the GLCM method to classify COVID-19 seen on CT scan images of the lungs. The artificial neural network consists of an input layer, four hidden layers, and an output layer. The optimizer is also used to tune the hyperparameters. This consists of a dropout layer followed by a fully connected sigmoid layer, predicting class performance.

Using a dataset of 1000 CT scan images of the lungs with balanced data between classes, the innovative approach achieves high accuracy. In addition, hyperparameter tuning was used to achieve 90% accuracy, 88% precision, 91% recall, and 90% F1-score values with the Adam optimizer, using a value of 5000 epochs, 32 batch sizes, and 0.4 dropout rate.

Based on the results obtained in this research, using the GLCM method to get the value of image features that can take the value of the grayscale level of images data, the model can be made with the results in our research that the most optimum model has an accuracy of 90%. The artificial neural network that has been conducted could classify COVID-19 from CT images. This contribution proves the prospect of improving the diagnosis of COVID-19 images data. In future research, it is possible to increase the number of CT images of patients from various image data sources and use various features extraction methods so that better models can be developed so can be a useful tool for clinical practitioners and radiologists to assist them in the diagnosis, quantification, and follow-up of COVID-19 cases.

ACKNOWLEDGMENT

This research was supported by the Department of Computer Engineering, School of Electrical Engineering, Telkom University, Indonesia.

REFERENCES

[1] N. Zhu *et al.*, “A Novel Coronavirus from Patients with Pneumonia in China, 2019,” *N. Engl. J. Med.*, vol. 382, no. 8, pp. 727–733, 2020, doi: 10.1056/nejmoa2001017.

[2] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu, “Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing,” *J. Clin. Microbiol.*, vol. 58, no. April, pp. 1–5, 2020.

[3] P. Hermanek and S. Carmignato, “Porosity measurements by X-ray computed tomography: Accuracy evaluation using a calibrated object,” *Precis. Eng.*, 2017, doi: 10.1016/j.precisioneng.2017.03.007.

[4] Z. Wang, W. Qiang, and H. Ke, “A Handbook of 2019-nCoV Pneumonia Control and Prevention,” *Hubei Sci. Technol. Press*, 2020.

[5] T. F. Ali, M. A. Tawab, and M. A. ElHariri, “CT chest of COVID-19 patients: what should a radiologist know?,” *Egypt. J. Radiol. Nucl. Med.*, vol. 51, no. 1, 2020, doi: 10.1186/s43055-020-00245-8.

[6] T. Sutojo, D. R. I. M. Setiadi, P. S. Tirajani, C. A. Sari, and E. H. Rachmawanto, “CBIR for classification of cow types using GLCM and color features extraction,” *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017*, vol. 2018-Janua, pp. 182–187, 2018, doi: 10.1109/ICITISEE.2017.8285491.

[7] C. Irawan, E. N. Ardyastiti, D. R. I. M. Setiadi, E. H. Rachmawanto, and C. A. Sari, “A survey: Effect of the number of GLCM features on classification accuracy of lasem batik images using K-nearest neighbor,” *2018 Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2018*, pp. 33–38, 2018, doi: 10.1109/ISRITI.2018.8864443.

[8] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, “Tomatoes classification using K-NN based on GLCM and HSV color space,” *Proc. - 2017 Int. Conf. Innov. Creat. Inf. Technol. Comput. Intell. IoT, ICITech 2017*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/INNOCIT.2017.8319133.

[9] O. I. Abiodun *et al.*, “Comprehensive Review of

Artificial Neural Network Applications to Pattern Recognition," *IEEE Access*, vol. 7, no. February 2017, pp. 158820–158846, 2019, doi: 10.1109/ACCESS.2019.2945545.

[10] R. Bala and D. Kumar, "Classification Using ANN: A Review," *Int. J. Comput. Intell. Res.*, vol. 13, no. 7, pp. 1811–1820, 2017, [Online]. Available: http://www.ripublication.com.

[11] R. Dharwal, "Applications of Artificial Neural Networks: A Review," *Indian J. Sci. Technol.*, vol. 9, no. 1, pp. 1–8, 2016, doi: 10.17485/ijst/2016/v9i47/106807.

[12] L. Kaur, M. Sharma, R. Dharwal, and A. Bakshi, "Lung Cancer Detection Using CT Scan with Artificial Neural Netwok," *ICRIEECE 2018*, pp. 1624–1629, 2018, doi: 10.1109/ICRIEECE44171.2018.9009244.

[13] J. Turian, J. Bergstra, and Y. Bengio, "Quadratic features and deep architectures for chunking," *Proc. Hum. Lang. Technol. 2009 Annu. Conf. North Am. Chapter Assoc. Comput. Linguist.*, no. June, pp. 245--248, 2009, [Online]. Available: https://aclanthology.org/N09-2062.

[14] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning," pp. 1–20, 2018, [Online]. Available: http://arxiv.org/abs/1811.03378.

[15] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," *ICML*, 2010.

[16] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, [Online]. Available: https: //doi.org/10.1038/nature14539.

[17] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," *ICML Work. Deep Learn. Audio, Speech Lang. Process.*, vol. 28, 2013.

[18] D. A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)," *4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc.*, pp. 1–14, 2016, [Online]. Available: http://arxiv.org/abs/1511.07289%0A[67].

[19] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.

[20] A. Özgür and Fatih Nar, "Effect of Dropout layer on Classical Regression Problems," *2020 28th Signal Process. Commun. Appl. Conf.*, pp. 1–4, 2020, doi: 10.1109/SIU49456.2020.9302054.

[21] M. Maftouni, A. Chung, C. Law, B. Shen, and Z. J. Kong, "A Robust Ensemble-Deep Learning Model for COVID-19 Diagnosis based on an Integrated CT Scan Images Database," *Proc. 2021 IISE Annu. Conf.*, 2021.