TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 19, No. 4, August 2021, pp. 1284~1290
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v19i4.18874

Lung cancer classification based on CT scan image by applying FCM segmentation and neural network technique

Ahmad Zoebad Foeady, Siti Ria Riqmawatin, Dian Candra Rini Novitasari
Department of Mathematics, UIN Sunan Ampel Surabaya, Indonesia

Article Info

Article history:

Received Sep 7, 2020
Revised Jan 8, 2021
Accepted Jan 21, 2021

Keywords:

ANFIS
ELM
FCM
GLCM
KELM
Lung cancer
RNN

ABSTRACT

The number of people with lung cancer has reached approximately 2.09 million people worldwide. Out of 9.06 million cases of death, 1.76 million people die due to lung cancer. Lung cancer can be automatically identified using a computer-aided diagnosis system (CAD) such as image processing. The steps taken for early detection are pre-processing feature extraction, and classification. Pre-processing is carried out in several stages, namely grayscale images, noise removal, and contrast limited adaptive histogram equalization. This image feature extracted using gray level co-occurrence matrix (GLCM) and classified using 2 method of neural network which is feed forward neural network (FFNN) dan feed backward neural network (FBNN). This research aims to obtain the best neural network model to classify lung cancer a. Based on training time and accuracy, the best method of FFNN is kernel extreme learning machine (KELM), with a training time of 12 seconds and an accuracy of 93.45%, while the best method of FBNN is Backpropagation with a training time of 18 minutes 04 seconds and an accuracy of 97.5%.

Corresponding Author:

Dian Candra Rini Novitasari
Department of Mathematics
UIN Sunan Ampel Surabaya
Ahmad Yani No. 117 Wonocolo Surabaya 60237, Indonesia
Email: diancrini@uinsby.ac.id

1. INTRODUCTION

Lung cancer is a malignant tumor that irregularly grows in one or both lungs with an abnormal and irregular direction of spread. Cancer cells spread from the lungs through the bloodstream or lymph fluid that surrounds the lung tissue. Cancer cells can also spread to an other organ, which is called the metastatic process [1]. There are four types of lung cancer that is squamous cell carcinoma, small cell carcinoma, large cell carcinoma, and adenocarcinoma [2]. According to the World Health Organization (WHO) in its report on cancer which stated that in 2018 people with lung cancer accounted for around 2.09 million people worldwide, and 1.76 million people died of lung cancer out of 9.06 million cases of death [3]. The American Cancer Society estimates that by 2020 America has 228,820 new cases of lung cancer and 135,720 deaths. This report is based on estimates for one year in 2020 [4]. This disease is characterized by pain accompanied by shortness of breath because the cancer cells fill the space in the lungs, and the capacity of the lungs for air storage is getting narrower [5]. This disease is also a frequent cause of other cancers because it spreads rapidly in the body and in the lungs [6]. Lung cancer is often called primary cancer because it is the beginning of the formation of other cancer cells in the body [7]. To identify this disease can be done by doing a computerized tomography (CT) scan or magnetic resonance imaging (MRI). However, most patients prefer to use CT scans because of the low cost and relatively accurate results. CT scan is a compilation of x-ray images taken from different angles with compilation processing to take pictures of particular body parts. In general, a CT scan image is black and white [8]. There are 4 stages in lung cancer that is stage I, II, III, and IV. Stage I, II, and III can be predicted by an axial CT scan of the lungs. While Stage IV patients can feel the symptoms because the cancer has spread throughout the body and cannot be identified from the CT scan [9]. Stage I lung cancer is said to be benign because it does not kill the patient. However, when the cancer reaches stage II onwards, it is classified as malignant because it has a mortality rate of 60-70%, and is difficult to cure. Therefore, early detection of lung cancer is required to determine whether a cancer is malignant or not so that treatment can be done as soon as possible [10].

This lung cancer early detection system uses a computer-aided diagnosis system (CAD) such as image processing. The steps taken for early detection are pre-processing feature extraction, and classification [11]. Pre-processing is carried out in several stages, namely, grayscale images, noise removal, histogram equalization, and Fuzzy C-Means (FCM) segmentation. FCM is the best cluster method, and all parameters must be pre-determined [12]. Balafar et al. stated that medical images had a lot of noise and inhomogeneity. Based on the paper he wrote, it proved that FCM was the best segmentation method for medical image data such as CT scans and MRI because it could reduce noise and maximize the feature and background selection process [13]. Huang *et al.* once applied FCM segmentation to CT brain images to detect brain cancer. In this research, it was explained that FCM could extract the mass feature on CT images well so that it was suitable for use in cancer detection [14]. Apart from CT scan images of the brain, FCM segmentation has also been applied to identify lung cancer, as was done by Dhaware *et al*. In this research, the FCM segmentation was combined with the gray level co-occurrence matrix (GLCM) method, which is suitable for extracting the texture of an image. The application of this method has successfully identified the mass po sition [15].

Image data will be easier to process when it is numeric data. The application of feature extraction is required to create numerical data from the image. In this case, the method used in feature extraction is the GLCM [16]. The next stage is the classification. This stage is used to determine the stage of lung cancer. The result of feature extraction from the GLCM will be the initial input of this process. In the classification process it self, the data will be divided into two parts, namely training data and testing data. In this research, several neural network methods are used to form the best model of a system. Neural network has been applied to CT scan data by Thanammal and Sudha [17]. This research used one of the neural network methods that is Backpropagation. This research showed that the application of segmentation could apply accuracy to the neural network method. The accuracy obtained in this research was 95% [17]. The application of neural network for lung cancer CT scan data has been conducted by Arulmurugan *et al.* [18] and Shaukat *et al.* [19]. The research explained that the neural network could classify a CT scan image well with an average accuracy of 94.5%. However, this research demonstrated that Backpropagation had a slow training time [18], [19]. Based on this case, lung cancer will be identified based on CT scan data to determine the best neural network model that can be used. This research will conduct trials on two neural network models, namely feed-forward (FFNN) and feed backward (FBNN). FFNN is a neural network that sends data or input in one direction, that is through the input node and out at the output node [20]. FBNN is a neural network that sends data or input in two directions that is through the input node to the output node and back again to the node input [21]. Based on several research reviews above, automatic detection of lung cancer will be carried out using both methods to maximize the results obtained. This research aims to achieve the best neural network model to classify lung cancer.

2. PRELIMINARIES

2.1. The fuzzy C-means segmentation algorithm (FCM)

Fuzzy C-means (FCM) is a data clustering technique in which the existence of each data point in a cluster is determined by the degree of membership. FCM segmentation is the separation of the background with features by clustering the image matrix [22]. The initial step required is the initialization of the initial FCM inputs such as iterations, multiple clusters, errors, and weights. Then, the membership matrix (μ_{ik}) is initialized randomly, where V_{kj} is the center of the cluster X_{ij} is the input matrix, w is the initial weight with the default value of 2, and k is a cluster [23]. The center of the cluster is calculated using (1).

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^{w} * X_{ij})}{\sum_{i=1}^{n}(\mu_{ik})^{w}} \tag{1}$$

After obtaining the cluster center, then the objective function is calculated using the formula shown in (2) and the change in the membership value matrix (u_{ik}) is calculated using (3). The iteration is said to stop when the minimum error or maximum iteration is reached [24].

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right](\mu_{ik})^{w}\right) \tag{2}$$

$$u_{ik} = \frac{[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2]^{\frac{-1}{w-1}}} \quad (3)$$

2.2. Neural network

Neural networks are a set of algorithms which is human brain modeled like and designed to recognize patterns [25]. Like the brain, an artificial neural network is a collection of connected units that can also be called neurons. Connections between neurons in an artificial neural network can carry signals in the form of real values that determine the weight or strength of the signal [26]. Input data or information sent by neurons can be single or multiple. Where a is the neuron output, w is the signal weight, p is the neuron input, and b is the bias [27]. There are several neural networks implemented based on mathematical operations, and a set of parameters is required to determine the output [25], [27].

- Feed forward neural network (FFNN)

FFNN is a neural network that sends data or input in one direction, that is through the input node and out at the output node. There are several feed-forward neural network methods, that is radial basis function NN (RBFNN), extreme learning machine (ELM), kernel extreme learning machine (KELM), and perceptron [20]. Figure 1 shows the design of the FFNN algorithm with circles in a network that forms neurons in an artificial neural network [28].

- Feed backward neural network (FBNN)

FBNN is a neural network that sends data or input in two directions that is through the input node to the output node and back again to the node input. There are several backward neural network methods that is Backpropagation, recurrent neural network (RNN), adaptive neuro-fuzzy inference system (ANFIS), and self-organizing map (SOM). Figure 2 shows the design of an FBNN in a network, in which there are nodes (circles) connected by edges that form neurons in a dummy network. The movement in FBNN occurs in feedback because the input data is sent in two directions [21].

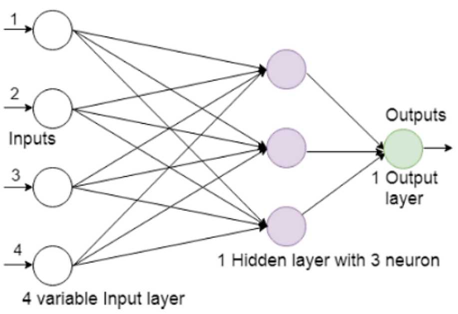


Figure 1. Feed-forward neural network algorithm [25]

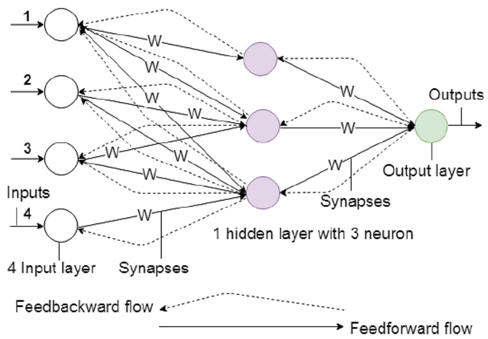


Figure 2. Feed backward neural network algorithm [25]

3. RESEARCH METHOD

3.1. Research type

This research is descriptive quantitative research because it involves a lot of calculations to find out the results, and the data processing must be analyzed in each stage. Calculations are carried out in all processes, while data analysis is carried out when the data has been processed and obtains the results. This research have 3 main process which is pre-processing, feature extraction, and classification.

3.2. Data collection and analysis

The data was obtained from a cancer imaging archive of 351 images consisting of several stages that is 72 data of stage I, 77 data of stage II, and 202 data of stage III. Furthermore, the data is processed into pre-processing, feature extraction, and then classification is carried out to determine whether the cancer is malignant or benign. Testing the evaluation data begins with the feature extraction process consisting of grayscale images, noise removal, histogram equalization, and FCM segmentation. Then the features are taken using the GLCM method, and finally, the data from the feature extraction is used as input to the neural network classification. The entire series of research stages can be seen in the flow chart in Figure 3. After the feature extraction results are obtained, then the training data and testing data are distributed to be included in the neural network method. The neural network methods used next are divided into two that is FFNN and FBNN. In

FFNN, the methods used are ELM, KELM, perceptron, and RBFNN, while in FBNN, the methods used are Backpropagation, RNN, ANFIS, and SOM. The training data is used to form the neural network method, and the testing data is used to test the system accuracy level. Finally, the data on lung cancer is divided into three classes that is the stage I, stage II, and stage III.

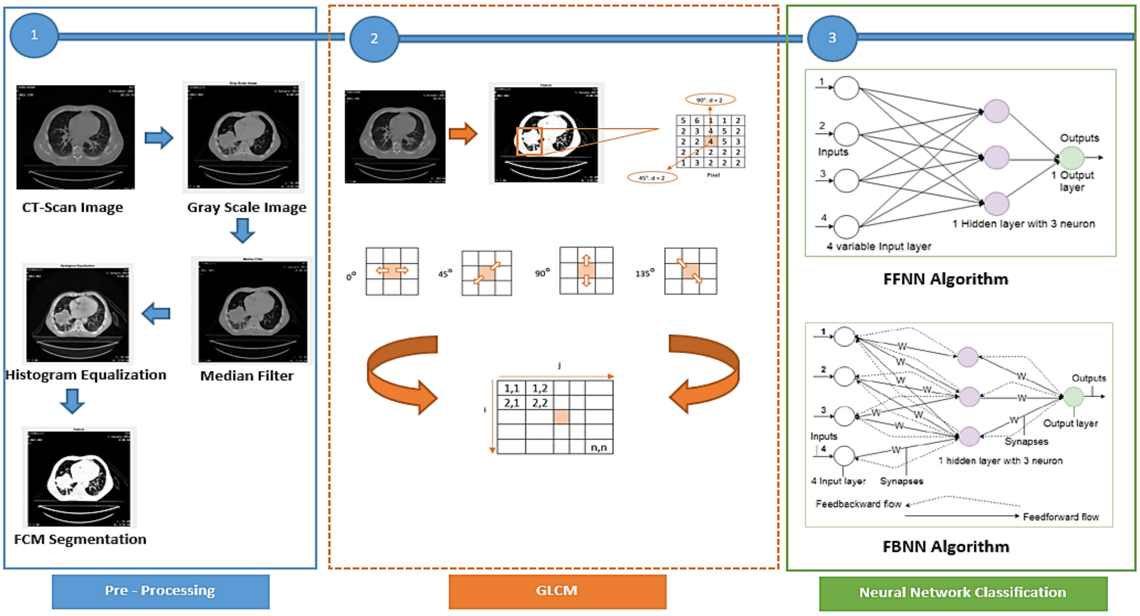
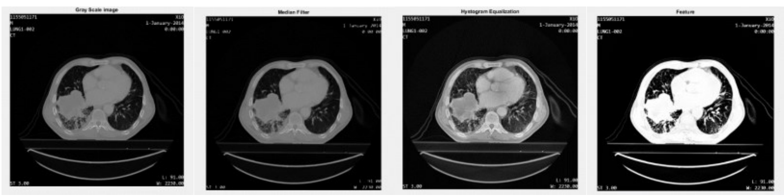


Figure 3. Graphical abstract of lung cancer classification based on CT scan image by applying FCM segmentation and neural network technique

4. RESULT AND DISCUSSION

Several processes are required to detect the level of malignancy of lung cancer. The results of pre-processing can be seen in Figure 4. Figure 4 (a) is a grayscale image and the $3x3$ median filter result shown in Figure 4 (b). The next step is the histogram equalization shown in Figure 4 (c). The last stage of pre-processing is the FCM segmentation whose result shown in Figure 4 (d), where the existing image has no background, and only features of the image are taken. Determination of clusters in the FCM segmentation in this research using a trial system. The trial system aims to determine the best feature and background separation. This research uses a trial of the number of clusters 2, 3, and 5. The best results achieved in this study are the number of clusters 3 which can be seen in Figure 4 (d) because it has high contrast on features and disguises the background image. Next, feature extraction using for parameters that is energy, correlation, homogeneity, and contrast. Sample data from feature extraction can be seen in Table 1.



(a) (b) (c) (d)

Figure 4. (a) Grayscale, (b) Median Filter, (c) Histogram Equalization, (d) FCM Segmentation

The next process is classification divided into two systems, namely FFNN and FBNN. FFNN uses ELM, KELM, RBFNN, and perceptron methods, while FBNN uses Backpropagation, SOM, ANFIS, and RNN methods. The data from the feature extraction is divided into 2 that is training data and testing data. Data sharing uses the K-fold cross validation method with $k = 5$.

As seen from Table 2 that each method has different accuracy values. The parameters used for FFNN and FBNN methods are the same that is hidden nodes of 100. The results obtained are KELM is the best method to classify lung cancer from FFNN. This is because KELM has implemented a kernel algorithm so that it can map data to a higher dimension which makes it easier for the system to recognize data patterns. While the best FBNN method is Backpropagation. This is because Backpropagation is known as weight updating, which allows the system to recognize data patterns better in each iteration. The worst accuracy among the methods is RNN. This is because RNN is a method based on data sequencing and it is more suitable for use in predictions that have a correlation between data. The results of the comparison of the FFNN and FBNN methods can be seen in Figures 5 and 6, respectively.

Table 1. Sample of GLCM results

Contrast	Correlation	Energy	Homogeneity	Label
0.5544	0.7681	0.9014	0.9762	Stage I
0.4881	0.7683	0.9292	0.9844	
1.1003	0.7120	0.8376	0.9544	
0.5425	0.8093	0.8943	0.9775	
0.4750	0.6018	0.9396	0.9832	Stage II
0.4971	0.7358	0.9264	0.9818	
0.4301	0.7483	0.9312	0.9839	
0.6083	0.6911	0.9175	0.9776	
0.5124	0.7490	0.9236	0.9815	Stage III
0.5156	0.6715	0.9343	0.9822	
0.4513	0.8942	0.9342	0.9846	
0.6142	0.6024	0.8535	0.9703	

Table 2. Classification result of FFNN and FBNN with 100 hidden nodes

FFNN			FBNN		
Method	Fold	Accuracy	Method	Fold	Accuracy
ELM	Fold 1	88.65%	SOM	Fold 1	91.50%
	Fold 2	88.00%		Fold 2	92.00%
	Fold 3	92.00%		Fold 3	75.45%
	Fold 4	90.50%		Fold 4	77.45%
	Fold 5	82.45%		Fold 5	80.00%
Max		92.00%	Max		92.00%
KELM	Fold 1	90.50%	RNN	Fold 1	77.50%
	Fold 2	90.00%		Fold 2	79.50%
	Fold 3	92.50%		Fold 3	82.75%
	Fold 4	92.25%		Fold 4	81.00%
	Fold 5	88.45%		Fold 5	77.25%
Max		92.50%	Max		82.75%
RBFNN	Fold 1	85.45%	ANFIS	Fold 1	88.65%
	Fold 2	74.45%		Fold 2	85.85%
	Fold 3	82.15%		Fold 3	93.45%
	Fold 4	84.00%		Fold 4	89.45%
	Fold 5	77.65%		Fold 5	88.65%
Max		85.45%	Max		93.45%
Perceptron	Fold 1	89.50%	Backpropagation	Fold 1	94.50%
	Fold 2	88.65%		Fold 2	95.75%
	Fold 3	90.00%		Fold 3	92.50%
	Fold 4	75.65%		Fold 4	95.45%
	Fold 5	82.50%		Fold 5	96.00%
Max		90.00%	Max		96.00%

As shown in the graph in Figure 6, overall, the accuracy levels of the four methods rise slightly and stabily which proves that FFNN with these four methods has a high match to recognize the patterns of lung cancer data. Then, the entire FFNN methods reach their highest point of accuracy at the hidden nodes of 100-250. Therefore, the training experiment was carried out using 100 hidden nodes because from several NN methods, the average has the best accuracy at 100 hodeen nodes. After that the accuracy levels of the four methods slowly decline. This is because the increasing number of hidden nodes results in overlapping cases so that there is data that is not included in any class. Out of the four FFNN methods, KELM has the highest accuracy of 93.45% at hidden nodes of 250. As seen in the graph in Figure 6 on the results of FBNN that RNN is less capable of classifying lung cancer. The results of RNN have a fairly far range from the other FBNN methods. RNN has the highest accuracy of 82.75% at the hidden nodes of 100. The ANFIS and SOM methods have a high accuracy at hidden node 100 with almost the same accuracy of 93.45% for ANFIS and 92% for SOM. In this research, the highest accuracy of FBNN was achieved using the backpropagation method with the highest accuracy of 97.5%. In addition to accuracy, it is necessary to pay attention to the performance of a model based on the required training time. The results on training time can be seen in Table 3. These results are also compared with systems that do not implement FCM segmentation.

Table 3 shows that the application of FCM segmentation can affect the results of the accuracy of the system being created. This is shown by comparing systems that use FCM segmentation and systems that use thresholding. The overall NN method has better accuracy if there is FCM segmentation. Table 3 also reveals that the FFNN method does not require long training time, while the FBNN method requires long training time. This is because the FBNN method requires iteration in the training process, while the FFNN method does not. Based on training time and accuracy, the best FFNN method is KELM with a training time of 12 seconds and an accuracy of 93.45%, while the best FBNN method is Backpropagation with a training time of 18 minutes 04 seconds and an accuracy of 97.5%. The results of a system can also be seen from the error. Error can be defined as the highest accuracy which is $f(x) = 100\% - accuracy$ and it can be concluded that the best FFNN method, KELM, has an error of 6.55%, while the best FBNN method, namely Backpropagation, has an error of 2.5%.

In this research, researches still lacked in pre-processing stage because the data used were not balanced. Therefore, data augmentation is needed to overcome the imbalance of the data. One of the methods that can be used for data augmentation is conditional generative adversial networks (GANs) conducted by Shanqing Gu. Manisha Pednekar, and Robert Slater [29].

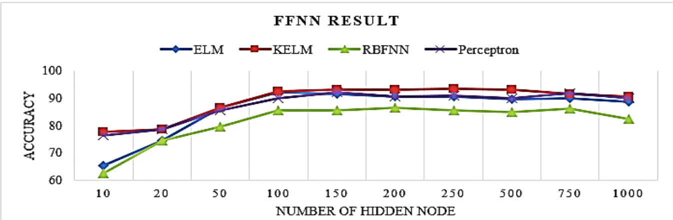


Figure 5. Graph of FFNN result

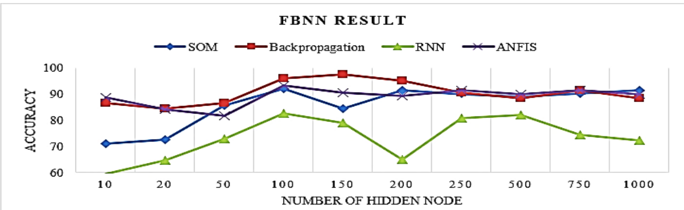


Figure 6. Graph of FBNN result

Table 3. Training time and best accuracy comparison

Model	Method	Number of Hidden Node	FCM Segmentation		Thresholding	
			Accuracy	Training Time	Accuracy	Training Time
FFNN	ELM	100	92.00%	5 second	89.60%	5 second
	KELM	250	93.45%	12 second	90.50%	10 second
	RBFNN	200	86.50%	8 second	82.00%	8 second
	Perceptron	150	92.15%	11 second	89.75%	10 second
FBNN	SOM	200	91.45%	14 minutes 32 second	85.15%	13 minutes 12 second
	RNN	500	82.00%	28 minutes 12 second	77.50%	27 minutes 48 second
	ANFIS	100	93.45%	12 minutes 41 second	91.00%	10 minutes 53 second
	Backpropagation	150	97.50%	18 minutes 04 second	96.50%	17 minutes 15 second

5. CONCLUSION

This research aims to obtain the best neural network model to classify lung cancer. The FCM segmentation has succeeded in increasing the accuracy of the lung cancer classification system. This explains that the FCM segmentation is able to identify the mass in lung cancer so that the classifier can recognize the pattern well. Based on training time and accuracy, the best FFNN method is KELM with a training time of 12 seconds and an accuracy of 93.45%, while the best FBNN method is Backpropagation with a training time of 18 minutes 04 seconds and an accuracy of 97.5%.

REFERENCES

[1] P. Kalavathi and A. Dhavapandiammal, "Segmentation of Lung Tumor in CT Scan Images using FA-FCM Algorithms," vol. 18, no. 5, pp. 74-79, 2016, doi: 10.9790/0661-1805047479.

[2] K. Inamura, "Lung cancer: understanding its molecular pathology and the 2015 WHO classification," *Front. Oncol.*, vol. 7, p. 193, 2017, doi: 10.3389/fonc.2017.00193.

[3] A. A. Adjei, "Lung cancer worldwide," *J. Thorac. Oncol.*, vol. 14, no. 6, p. 956, 2019.

[4] The American Cancer Society medical and editorial content team, "Key Statistics for Lung Cancer," *American Cancer Society*, 2020. https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html (accessed Aug. 09, 2020).

[5] M. Mufti, S. Ching, S. Farjami, S. Shahangian, and S. Sobnosky, "A case series of two patients presenting with pericardial effusion as first manifestation of non-small cell lung cancer with BRAF mutation and expression of PD-L1," *World J. Oncol.*, vol. 9, no. 2, p. 56, 2018, doi: 10.14740/wjon1092w.

[6] V. W. Q. Lou *et al.*, "Respiratory symptoms, sleep, and quality of life in patients with advanced lung cancer," *J. Pain Symptom Manage.*, vol. 53, no. 2, pp. 250–256, 2017, doi: 10.1016/j.jpainsymman.2016.09.006.

[7] M. P. Yavropoulou *et al.*, "Distant lung metastases caused by a histologically benign phosphaturic mesenchymal tumor," *Endocrinol. diabetes Metab. case reports*, vol. 2018, no. 1, 2018, doi: 10.1530/EDM-18-0023.

[8] S. Makaju, P. W. C. Prasad, A. Alsadoon, A. K. Singh, and A. Elchouemi, "Lung cancer detection using CT scan images," *Procedia Comput. Sci.*, vol. 125, pp. 107-114, 2018, doi: 10.1016/j.procs.2017.12.016.

[9] S. Park *et al.*, "Volume doubling times of lung adenocarcinomas: correlation with predominant histologic subtypes and prognosis," *Radiology*, vol. 295, no. 3, pp. 703-712, 2020, doi: 10.1148/radiol.2020191835.

[10] L. Ni *et al.*, "Heart V5 predicts cardiac events in unresectable lung cancer patients undergoing chemoradiation," *J. Thorac. Dis.*, vol. 11, no. 6, p. 2229, 2019, doi: 10.21037/jtd.2019.06.29.

[11] M. Thohir, A. Z. Foeady, D. C. R. Novitasari, A. Z. Arifin, B. Y. Phiadelvira, and A. H. Asyhar, "Classification of Colposcopy Data Using GLCM-SVM on Cervical Cancer," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2020, pp. 373-378.

[12] A. Kapoor and A. Singhal, "A comparative study of K-Means, K-Means++ and Fuzzy C-Means clustering algorithms," in *2017 3rd international conference on computational intelligence & communication technology (CICT)*, 2017, pp. 1-6, doi: 10.1109/CIACT.2017.7977272.

[13] M. A. Balafar, "Fuzzy C-mean based brain MRI segmentation algorithms," *Artif. Intell. Rev.*, vol. 41, no. 3, pp. 441-449, 2014, doi: 10.1007/s10462-012-9318-2.

[14] H. Huang, F. Meng, S. Zhou, F. Jiang, and G. Manogaran, "Brain image segmentation based on FCM clustering algorithm and rough set," *IEEE Access*, vol. 7, pp. 12386-12396, 2019, doi: 10.1109/ACCESS.2019.2893063.

[15] B. U. Dhaware and A. C. Pise, "Lung cancer detection using bayasein classifier and FCM segmentation," in *2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)*, 2016, pp. 170-174, doi: 10.1109/ICACDOT.2016.7877572.

[16] A. Z. Foeady, D. C. R. Novitasari, A. H. Asyhar, and M. Firmansjah, "Automated Diagnosis System of Diabetic Retinopathy Using GLCM Method and SVM Classifier," *2018 5th Int. Conf. Electr. Eng. Comput. Sci. Informatics*, pp. 154-160, 2018, doi: 10.1109/eecsi.2018.8752726.

[17] K. K. Thanammal and J. S. J. Sudha, "Enhancement of fissure using back propagation neural network and segmentation of lobes in CT scan image," *Int. J. Biomed. Eng. Technol.*, vol. 20, no. 1, pp. 1-11, 2016.

[18] R. Arulmurugan and H. Anandakumar, "Early detection of lung cancer using wavelet feature descriptor and feed forward back propagation neural networks classifier," in *Computational Vision and Bio Inspired Computing*, Springer, 2018, pp. 103-110.

[19] F. Shaukat, G. Raja, R. Ashraf, S. Khalid, M. Ahmad, and A. Ali, "Artificial neural network based classification of lung nodules in CT images using intensity, shape and texture features," *J. Ambient Intell. Humaniz. Comput.*, vol. 10, no. 10, pp. 4135-4149, 2019, doi: 10.1007/s12652-019-01173-w.

[20] A. Lomuscio and L. Maganti, "An approach to reachability analysis for feed-forward relu neural networks," *arXiv Prepr. arXiv1706.07351*, 2017.

[21] L. N. M. Tawfiq and O. M. Salih, "Design Suitable Feed Forward Neural Network to Solve Troesch's Problem," Sci. Int.(Lahore), vol. 31, pp. 41-48, 2019.

[22] D. C. R. Novitasari *et al.*, "Identify Education Quality Based on Islamic Senior High School Data in Kompetisi Sains Madrasah Using Fuzzy C-Means Clustering," in *Smart Trends in Computing and Communications: Proceedings of SmartCom 2020*, Springer, 2020, pp. 487-495.

[23] Y. K. Dubey and M. M. Mushrif, "FCM clustering algorithms for segmentation of brain MR images," *Adv. Fuzzy Syst.*, vol. 2016, 2016, doi: 10.1155/2016/3406406.

[24] B. Thamaraichelvi and G. Yamuna, "Gaussian kernel-based FCM segmentation of brain MRI with BPNN classification," *Int. J. Biomed. Eng. Technol.*, vol. 20, no. 2, pp. 116-131, 2016, doi: 10.1504/IJBET.2016.074198.

[25] O. I. Abiodun *et al.*, "Comprehensive review of artificial neural network applications to pattern recognition," *IEEE Access*, vol. 7, pp. 158820-158846, 2019, doi: 10.1109/ACCESS.2019.2945545.

[26] Y. Xu, H. Yao, and K. Lin, "An overview of neural networks for drug discovery and the inputs used," *Expert Opin. Drug Discov.*, vol. 13, no. 12, pp. 1091-1102, 2018, doi: 10.1080/17460441.2018.1547278.

[27] M. Islam, G. Chen, and S. Jin, "An Overview of Neural Network," *Am. J. Neural Networks Appl.*, vol. 5, no. 1, pp. 7-11, 2019, doi: 10.11648/j.ajnna.20190501.12.

[28] L. D. Medus, T. Iakymchuk, J. V. Frances-Villora, M. Bataller-Mompeán, and A. Rosado-Muñoz, "A Novel Systolic Parallel Hardware Architecture for the FPGA Acceleration of Feedforward Neural Networks," *IEEE Access*, vol. 7, pp. 76084-76103, 2019, doi: 10.1109/ACCESS.2019.2920885.

[29] S. Gu, M. Pednekar, and R. Slater, "Improve Image Classification Using Data Augmentation and Neural Networks," *SMU Data Sci. Rev.*, vol. 2, no. 2, p. 1, 2019.