Automatic Approach for Cervical Cancer Detection Based on Deep Belief Network (DBN) Using Colposcopy Data

Dian Candra Rini Novitasari
Department of Mathematics
UIN Sunan Ampel Surabaya
Surabaya, Indonesia
diancrini@uinsby.ac.id

Ahmad Zoebad Foeady
Department of Mathematics
UIN Sunan Ampel Surabaya
Surabaya, Indonesia
deddyzoebad711@gmail.ac.id

Muhammad Thohir
Department of Arabic Language Education
UIN Sunan Ampel Surabaya
Surabaya, Indonesia
muhammadthohir@uinsby.ac.id

Ahmad Zaenal Arifin
Department of Mathematics
Universitas PGRI Ronggolawe
Tuban, Indonesia
az_arifin@unirow.ac.id

Khoirun Niam
Department of Islamic Education
UIN Sunan Ampel Surabaya
Surabaya, Indonesia
khoirunniam@uinsby.ac.id

Ahmad Hanif Asyhar
Department of Mathematics
UIN Sunan Ampel Surabaya
Surabaya, Indonesia
hanif@uinsby.ac.id

***Abstract*— Cervical cancer is one of the diseases with the highest mortality rate. In the world, cervical cancer is ranked as the fourth most dangerous disease. Based on these problems, this paper can be an alternative to help medical authorities in detecting cervical cancer with the help of the Computer-Aided Diagnosis (CAD) System. CAD System used has two processes, such as preprocessing and classification. Preprocessing is useful to improve the image so that it is easier to do the process of identifying features. Preprocessing used is greyscale, histogram equalization, and median filter. Preprocessing results will be formed into a vector matrix using the reshaping process. The final step is the process of classifying data using the Deep Belief Network method. The best accuracy results obtained from the identification process of cervical cancer using the DBN method is 84%. Based on the results of accuracy, is expected to help reduce the number of deaths from cervical cancer with early detection.**

Keywords—Cervical Cancer, Deep Learning, Deep Belief Network, Colposcopy

I. INTRODUCTION

Cancer is the main cause of death worldwide. Every year, around 7.4 million people are dying because of cancer. Based on records, cancer patients have increased every year, so it is predicted that cancer patients in 2030 will reach 11.5 million deaths [1]. Among men, the most common cancer is lung cancer that caused the smoking habit, and among women, the most common cancers are breast cancer and cervical cancer [2].

Cervical cancer is one of a disease with a high mortality rate, especially for women. In the world, cervical cancer is ranked 4th as the most dangerous disease suffered. Every year, around 530,000 cases of cervical cancer appear, and around 270,000 patients die are 48-year-old women [3]. There are about 75% of women in developing and developed countries suffer from cervical cancer [4], [5]. This cancer is caused by several factors, namely environmental factors and gene factors. Cervical cancer is also formed through the spread of viruses [6].

The main virus that causes cervical cancer is the Human papillomavirus (HPV). Until now, 80 types of HPV are known, and around 40 species can be transmitted through genes. There are 11 of the 40 types of HPV that have a high chance of death [7]. The main reason for death is the lack of knowledge in the early identification of cervical cancer due to symptoms that are difficult to recognize by sufferers. Most cervical cancer sufferers feel symptoms when they have reached a high stage, so it is almost impossible for doctors to deal with the problem. Deaths caused by cervical cancer can be reduced by identifying it early, and it can be overcome before reaching a higher stage [8]. Based on this problem, this paper can be an alternative to help medical authorities in detecting cervical cancer with the help of the Computer-Aided Diagnosis (CAD) System [9]. The CAD system is an intelligence system that can help in diagnosing or detecting something like classification. The CAD process has several steps that must be done, and it can provide a decision in detecting something [10]. CAD systems tend to be used to detect things based on numerical data, signals, or in the form of images. In the case of cervical cancer, data that can be processed using a CAD system is a colposcopy image [11].

In general, data from colposcopy will be used as data processing with functional approach techniques or called soft computing. The soft computing technique is a model approach using computers that works to imitate human reason. Some stages in soft computing techniques include image processing (preprocessing), feature extraction, and classification [12]. If the process uses a deep learning method, the soft computing process is carried out without using the feature extraction process so that the process obtained is preprocessing and classification [13].

The first process carried out in this research is preprocessing. Preprocessing is useful for improving the image, and it is easier for the process of identifying features. In preprocessing is always accompanied by a grayscale process, because in the grayscale level the image contains more information. The preprocessing stage to identify cervical cancer has been investigated like the research conducted by Liang et al. [14] in 2013 on how to identify cervical cancer based on colposcopy images guided by classification using SVM for learning new feature input vectors and very clear low-level features for recognition of image accuracy. In another research of techniques for identifying cervical cancer cells carried out by Gupta et. al. [15] in 2014 based on statistical values including sensitivity, specificity, positive predictive value (PPV), and negative predictive value (not) to see the comparison of the accuracy of colposcopy tests on lesion size, large or small. The obtained statistical values are 85.85% for sensitivity, 87.65% for specificity, 75.83% for positive predictive value (PPV), and 95.38% for negative predictive value (NPV). The results of the preprocessing will be made a feature for the classification process.

The classification process consists of training and testing, so the results of preprocessing will be divided into training data and testing data. Training data used as modeling or pattern that will later be used in classification. Testing data used to test models that have been formed from the training process. Research conducted by Kavitha et al. [16] to classifies leukemia using the SVM method. The results obtained are good enough, with an accuracy of 65%. The SVM method is quite efficient as a classification method due to the speed of the process, but in this modern era, there is a deep learning method that emphasizes the learning process and has advantages over methods that are not deep learning. One of the deep learning methods that are quite popular is the Deep Belief Network (DBN) [17].

DBN has been widely used in research such as that conducted by Nassih Bouchra et al. [18] in 2018 about face identification using DBN. The results obtained from the paper are arguably satisfactory, with an accuracy of 98.86%. Prasanna Tamilselvan et al. [19] in 2012 also used DBN to diagnose someone's health. In this paper comparing DBN with SVM, which results in better accuracy when using DBN with a difference of approximately 1-3%. Maryam Sabzevari et al. [20] in 2010 also researched DBN to recognize facial patterns and the research was successfully applied and recognized faces well. Kamada et al. [21] in 2018 also compared the DBN method with other deep learning methods, namely CNN, with the best accuracy using the DBN method. Based on research conducted by Prasanna, Nassih, and Kamada [18], [19], [21], produce conclusions that the DBN method compared with other methods, the DBN is still better based on its accuracy. Research on DBN is also applied because it can properly recognize patterns, and it can be able to classify well. [20]. Based on several studies about DBN, this research aims can get good accuracy, and it can help the medical authorities in handling cervical cancer sufferers early.

II. PRELIMINARIES

A. *Cervical Cancer*

Cervical Cancer is the second most common cancer suffered by women. This is caused by the low level of medical services that can overcome cervical cancer [22]. The most effective method to date for identifying cervical cancer is to use a pap smear test or use a colposcopy image.

Fig.1 Colposcopy Image

Colposcopy is a series of tests to determine cervical cancer sufferers by taking pictures of a woman's cervix. The results of colposcopy images can be distinguished based on the shape and size of the cancer cells [23]. The results from colposcopy images can be seen in Figure 1.

B. *Deep Belief Network*

Deep Neural Network (DBN) generally requires large amounts of data labeled balanced. Deep Belief Network is a type of probabilistic generative model that is not monitored and is built by stacking the Restricted Boltzmann Machine (RBM) method. The difference between DBN and Deep Neural Network (DNN) is that DBN uses contrastive divergence to determine the weight which is then entered into backpropagation while DNN requires a backpropagation process to determine the weight of each process [20].

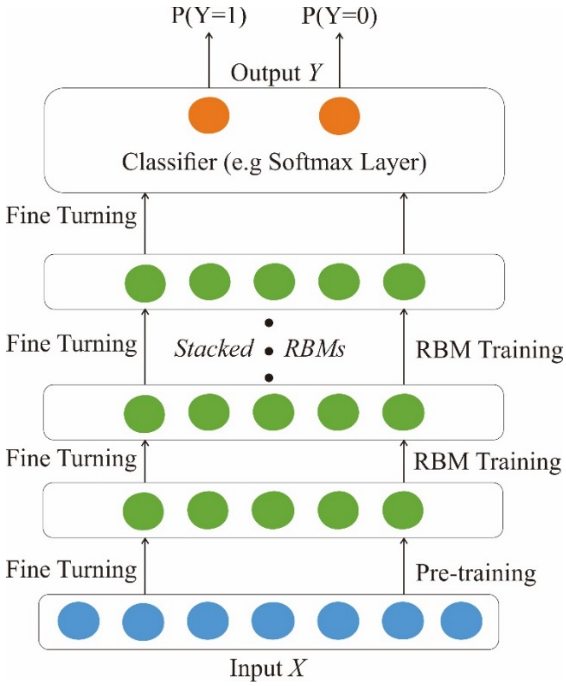


Fig.2 DBN Architecture

The parameters for RBM are trained by the Contrastive Divergence (CD) algorithm. CD is learning without supervision and labeled data is not needed at this stage. In the second stage, pre-trained networks will be adapted to supervised learning models such as softmax/logistic regression or backpropagation. However, the Deep Belief Network parameters were almost fixed after the CD and the second stage only set the model parameters [17]. DBN Architecture can be seen in Figure 2,

a. *Pre-training*

Pre-training is applied to the RBM stack. The hidden layer from the previous RBM will be the visible layer in the next RBM. Next is the stage of the pre-training process by taking X as an input matrix [24]:

1. Train the first layer of RBM on X to get a weight matrix using the Contrastive Divergence algorithm.
2. Transform X with a weight matrix to reconstruct new data X '.
3. Use X ' as the new input for the second layer.
4. Repeat steps 1 through 3 to the last two layers of tissue.

b. RBM *Training*

RBM is a two-layer stochastic model consisting of hidden layers and visible layers. The visible layer is denoted as v and hidden layer denoted as h. The two layers are connected, but the same layer is independent of each other. In RBM, the joint probability distribution (v, h) give in Equation 9 [25].

$$p(v,h) = \frac{1}{Z} e^{-E(v,h)} \tag{9}$$

Energy function $E(v, h)$ defined in Equation 10 [26].

$$E(v,h) = -\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} h_i v_j - \sum_{j=1}^{m} a_j v_j - \sum_{i=1}^{n} b_i h_i \tag{10}$$

The marginal probability of v can be calculated by counting the number of values $e^{-E(v,h)}$ from h [25].

$$p(v) = \frac{1}{Z} \sum_h e^{-E(v,h)} \tag{11}$$

Where Z is the normalization factor which is the sum of all combinations (v, h).

$$Z = \sum_{v,h} e^{-E(v,h)} \tag{12}$$

In the general case from binary conditions (where v_j and $h_i \epsilon \{0,1\}$), the probability with $vj = 1$ can be calculated by Equation 13, with σ is the sigmoid function.

$$p(v_j = 1|h) = \sigma(a_j + \sum_i h_i w_{ij}) \tag{13}$$

Same as probability with $hi = 1$ can be calculated by Equation 14.

$$p(h_i = 1|v) = \sigma(b_i + \sum_i v_j w_{ij}) \tag{14}$$

The main process for RBM training is to learn the weight matrix $W = \{w_{ij}\}$ that maximizes log Likelihood $\log p(v)$. From Equation 10 and 11, is obtained

$$\frac{\partial E(v,h)}{\partial w_{ij}} = -v_j h_i$$

$$\frac{\partial \log p(v)}{\partial w_{ij}} = \langle v_j h_i \rangle^0 - \langle v_j h_i \rangle^\infty$$

$$\Delta \partial w_{ij} = \epsilon(\langle v_j h_i \rangle^0 - \langle v_j h_i \rangle^\infty) \tag{15}$$

With $\langle v_j h_i \rangle^t, t = 0, \dots, \infty$ shows the expectations of random variables $v_j h_i$ in the sample sampling step t and ϵ is the learning rate to update weights.

c. Fine Tuning

In this research, backpropagation and softmax regression are used in the fine-tuning of DBN.

1. *Backpropagation*

Backpropagation is used to adjust the weights to the principle of the chain in the error model. Errors will be propagated from the last layer to the first layer. The two main parameters here are the batch size and number of epochs. In fine-tuning, the training sample is divided into groups of the same size [27]. The number of epochs related to fine-tuning iteration. In general, the network will fine-tune fine-tuning with smaller batch sizes or a greater number of epochs [28].

The derivative for the final layer can be determined as follows.

$$\frac{\partial L}{\partial w_{ij}} = \delta_j \frac{\partial \mu_j}{\partial w_{ij}} + \tau_j \frac{\partial p^2{}_j}{\partial w_{ij}} \tag{16}$$

$$\delta_j = \frac{\partial L}{\partial \mu_j} = \frac{\partial L}{\partial o_j} \frac{\partial o_j}{\partial n_j} \frac{\partial n_j}{\partial n_j} \tag{17}$$

$$\tau_j = \frac{\partial L}{\partial p^2{}_j} = \frac{\partial L}{\partial o_j} \frac{\partial o_j}{\partial n_j} \frac{\partial n_j}{\partial p^2{}_j} \tag{18}$$

Where L is a loss function, o_j is an output of j nodes. n_j is the input of the activation function node j-th, and i, j identifies the input and output node. In the proposed algorithm, the input of the activation function on Equation 19.

$$n = \frac{\mu}{\sqrt{1 + \rho^2 \pi / 8}} \tag{19}$$

For the hidden layer, the derivative can be written as follows

$$\frac{\partial L}{\partial w_{ij}} = \delta_j \frac{\partial \mu_j}{\partial w_{ij}} + \tau_j \frac{\partial p^2{}_j}{\partial w_{ij}} \tag{20}$$

$$\delta_j = \alpha_j \frac{\partial n_j}{\partial \mu_j} \tag{21}$$

$$\tau_j = \alpha_j \frac{\partial n_j}{\partial p^2{}_j} \tag{22}$$

$$\alpha_j = \left[\sum_k \left\{ \delta_k \frac{\partial \mu_k}{\partial o_j} + \tau_k \frac{\partial p^2{}_k}{\partial o_j} \right\} \right] \frac{\partial o_j}{\partial n_j} \tag{23}$$

Where δ_k, τ_k is a derivative propagated from the previous layer and k identify the nodes from the previous layer.

2. Softmax Regression

Softmax Regression [17] is an extension of logistic regression, which can be applied to more than two classes. Probability of $Y = l$ given RBM output X ', the coefficient matrix C and interception d are shown in Equation 24. The final prediction is the class with the greatest probability in Equation 25.

$$P(Y = l|X', C, d) = \frac{e^{C_l{}^T x' + d_l}}{\sum_k e^{C_k{}^T x' + d} \; k} \tag{24}$$

$$\text{Prediction} = max_{l \epsilon L}\{P(Y = l|X', C, d)\} \tag{25}$$

C. Confusion Matrix

The confusion matrix is a method to find out information that contains actual and predictive data from the classification system results. In classification, it is expected to be able to classify data appropriately, and it has good results with small errors. Therefore, this method exists to help find out how successful a classification [29]. There are three results which are the main results of the confusion matrix, and these results are accuracy, specifications, and sensitivity. To get this result, the things that need to be identified are True Positive (TP), Positive False (FP), False Negative (FN), and True Negative (TN) [30][31]. The confusion matrix table can be seen in Table 1.

TABLE I. CONFUSION MATRIX

Actual	Classification	
	+	-
+	True Positive (TP)	False Negative (FN)
_	False Positive (FP)	True Negative (TN)

After getting the parameters TP, FN, FP, and TN, the calculation aims to find accuracy, specifications, and sensitivity that can be seen in Equations 8, 9, and 10.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (26)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (27)$$

$$Spesificity = \frac{TN}{TN + FP} \quad (28)$$

III. Research Method

A. *Types of Research*

Research on the detection of cervical cancer using the Deep Belief Network (DBN) method is a quantitative descriptive study because in the process requires mathematical calculations and the results are in the form of an analysis of the detection of cervical cancer in the hope that later it can help the medical party in early identification of this disease.

B. *Data Collection and Analysis*

The data used is colposcopy image data obtained from [32]. The data obtained has been carried out the augmentation process to 500 data. This data will be divided into testing data and training data with 400 and 100, respectively. This division of data will then be forwarded to the preprocessing and classification process. Pre-processing contains the steps to get the uterine wall features. Furthermore, the results of preprocessing will be used as input from the classification process. The classification method used is the Deep Belief Network (DBN) which was originally a stacked RBM method and then a backpropagation process or other linear classification method is carried out.

C. *Testing Data Evaluation*

The first stage for conducting this research is the preprocessing process. This process is used to improve images to simplify the classification process. The work steps can be seen through the flowchart in Figure 2.

a. *Preprocessing*

The preprocessing stage is needed in a CAD process to assist in the process of improving data. In image data, most preprocessing will be used to improve or improve image quality that can make the identification process has better results than other research. In this research, the preprocessing stage used is grayscale, histogram equalization, and median filter. Grayscale is useful for reducing noise caused by RGB images [14]. Furthermore, the grayscale image will be improved in light intensity. This process is used to flatten the intensity of the light to make images that are too dark or too bright will get a sharper light intensity. The last step is to use a median filter [15]. The median filter will help to reduce the small noise left behind due to errors in shooting so the image will look clearer.

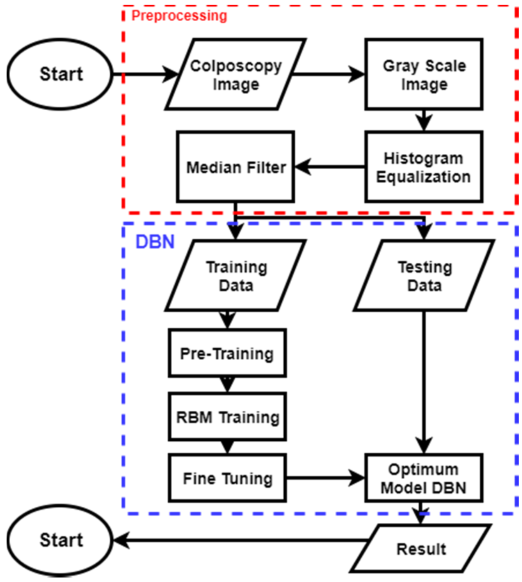


Fig.2 Flowchart of Cervical Cancer Identification

b. *Deep Belief Network*

The results of the image from the pre-processing then carried out the classification process using the DBN method. The initial step of the DBN method is to pre-train with input in the form of an image, and the output is on Hidden layer 1. Then the output from Hidden Layer 1 will be used as input for the hidden Layer 2 output until the last layer. This stage is called RBM training [17]. This process is carried out until getting the desired final output value. The next step is the backward pass process using the supervised classifier method. The method used in the backward pass is backward propagation. After all the processes are fulfilled, the results of the RBM training will be in the form of the DBN model, which will then be tested using data testing [24].

IV. Result And Discussion

The results of this research are classification accuracy which aims to measure the success of identification. The greater the accuracy produced, the greater the success of identification. The identification process requires two stages, such as preprocessing and classification.

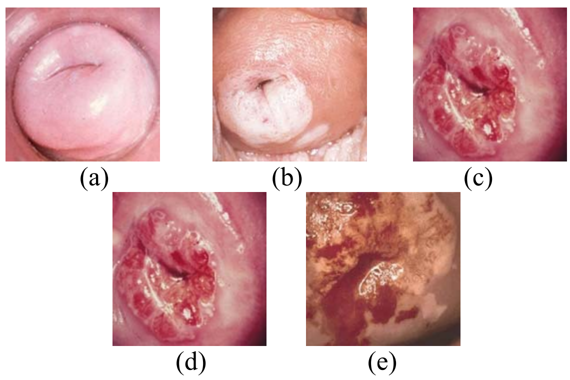
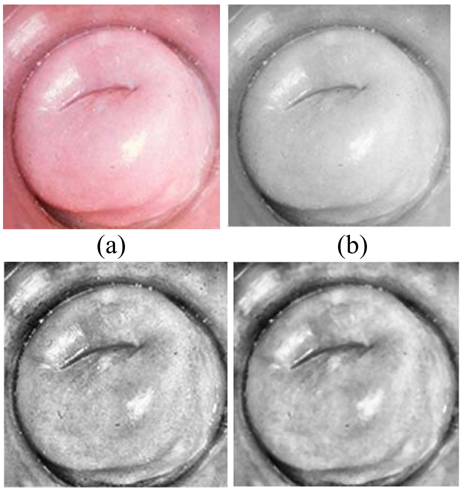


Fig.3 Data Sample (a) Normal (b) Stage 1 (c) Stage 2 (d) Stage 3 (e) Stage 4

a. *Preprocessing*

Preprocessing aims to improve the quality of the image so that it can calcify properly between normal data, stage 1, stage 2, stage 3, and stage 4. Data samples in this research can be seen in Figure 3. The data will be processed using grayscale, histogram equalization, and median filter. The process is only intended so the image has more information than reducing unnecessary information. The results of this process can be seen in Figure 4.

In Figure 4.b that the image that has been carried out the grayscale process has a lot of information that is reduced so that the rest of the important information based on cervical cancer features will be taken. Colposcopy images have varying light intensities. These problems can be overcome through the process of equalization of light intensity. Figure 4.c histogram equalization process is performed to smooth the intensity of light so that the image looks sharper and has more specific information. On the results of histogram equalization, the image still has some noise that must be removed. A median filter process is performed to clean up the remaining noise to solve this problem. The results of the median filter process can be seen in Figure 4.d.



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

Fig.4 (a) Original Image (b) Gray Scale Image (c) Histogram Equalization (d) Median Filter

b. *Classification*

Before the classification process is carried out, the preprocessing data will be formed into a matrix-vector. The results of preprocessing have a size of 200 x 200 pixels, and it converted to a vector matrix will have a size of 1 x 40000. Changing the matrix to this vector is called the process of reshaping data.

$$x_{200\,x\,200} = \begin{bmatrix} 23 & \cdots & 154 \\ \vdots & \ddots & \vdots \\ 211 & \cdots & 70 \end{bmatrix}$$
$$= [x_{1,1}, \dots, x_{1,200}, \dots, x_{200,1}, \dots, x_{200,200}]$$
$$= [23, \dots, 154, \dots, 211, \dots, 70]$$

Due to the amount of data used as much as 500 data and the result of preprocessing output is a matrix with a size of 500 x 40000.

After reshaping the data, the data will be classified into four classes, the first-class shows normal cervix, the second class shows the results of stage 1 cervical cancer, while the third class, fourth class, and fifth class show the results indicated stage 2 cervical cancer, 3, and 4. The data classification process is carried out using the Deep Belief Network method. Implementation of the Deep Belief Network method for classification requires the number of hidden nodes specified by the user. Determination of the number of hidden nodes will determine the results of the classification. In this research, the classification was carried out using four different hidden nodes.

TABLE II. DBN TRIAL TEST RESULTS

Trial	Accuracy	Hidden nodes
1	74%	50
2	80%	100
3	84%	200
4	62%	500
5	52%	1000

The first classification experiment was carried out with a number of hidden layers of 50. The second, third, and fourth experiments were carried out with hidden nodes of 100, 200, and 500. The accuracy of each experiment produced different values. From Table 2 it can be seen that the best accuracy results obtained in the 3rd experiment with the number of hidden nodes were 200. Based on trials that have been carried out using nodes as many as 50, 100, 200, and 500 have the results of succession are 74%, 80%, 84%, and 62%. In this trial, the best accuracy results were obtained at 200 nodes with an accuracy of 84%. Table 2 also shows that more hidden layers cannot influence the value of accuracy. This is evidenced in the fourth experiment that uses as many as 500 hidden nodes has an accuracy of 62%. In the fourth experiment has more hidden layers than the third experiment, but the results of its accuracy prove that the third experiment has a greater value. this case occurs because the hidden nodes above 200 have overfitting results and from that case get poor accuracy. The accuracy value is obtained from the confusion matrix value. Confusion matrix results from the 4th experiment can be seen in Table 3.

TABLE III. CONFUSION MATRIX WITH NODES 200

Actual Data	**Predicted Data**				
	Normal	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**
Normal	16	1	3	0	0
Stage 1	1	17	0	0	2
Stage 2	0	1	17	2	0
Stage 3	1	0	2	17	0
Stage 4	1	0	2	0	17

In Table 3, it can be seen that there are four normal wrong data identified. This affects the accuracy value. But more important is the result of cancer detection, which is detected as normal. If there are normal data identified by cancer, the results of the identification can be used as advice to the patient despite a misdiagnosis, but the advantage is that the patient can examine before cancer occurs. Conversely, when there is identified normal cancer data, there will be a big mistake because cancer patients will not examine because they believe that he is normal. Based on these problems, trials should continue using better methods to avoid these errors.

In this research still uses the basic DBN method. This research can be improved by optimizing the initial weight determination or trying to use the dropout method to selection the hidden nodes and the expectation for this evaluation will be much better.

V. Conclusion

The results of the implementation of the DBN method to identify cervical cancer with a grayscale image preprocessing process, histogram equalization, and median filter are a system that can classify cervical cancer accurately. The preprocessing process succeeded in getting a cervical feature which will later be used in the identification process and the classification process using the DBN method. In this method, the hidden nodes 50, 100, 200, and 500 have hidden comparisons with 74%, 80%, 84%, and 62% accuracy respectively. The best results obtained are 84%. By using 200 nodes. These results will be used as a reference for researchers if they want to detect cervical cancer using DBN. These results can also help medical authorities detect early by applying this method in the Detection Machine System.

References

[1] E. Weisshaar, "Cancers," *Pruritus Second Ed.*, pp. 283–287, 2016.

[2] World Health Organization, "Cancer control - Knowledge into action," 2008.

[3] W. Small *et al.*, "Cervical cancer: A global health crisis," *Cancer*, vol. 123, no. 13, pp. 2404–2412, 2017.

[4] H. I. Hyacinth, O. A. Adekeye, J. N. Ibeh, and T. Osoba, "Cervical Cancer and Pap Smear Awareness and Utilization of Pap Smear Test among Federal Civil Servants in North Central Nigeria," *PLoS One*, vol. 7, no. 10, pp. 1–8, 2012.

[5] World Health Organization, "Global burden of disease report: Causes of death in 2004. Global Burden of Disease Report.," Guneva, 2004.

[6] World Health Organization, "Human Papillomavirus," pp. 1–6, 2018.

[7] N. Muñoz *et al.*, "Epidemiologic classification of human papillomavirus types associated with cervical cancer," *N. Engl. J. Med.*, vol. 348, no. 6, pp. 518–527, 2003.

[8] World Health Organization, "Accelerating cervical cancer elimination Report by the Director-General Cervical Cancer: A Global Public Health Priority," vol. 94, no. November 2018, 2020.

[9] P. Elayaraja and M. Suganthi, "Automatic approach for cervical cancer detection and segmentation using neural network classifier," *Asian Pacific J. Cancer Prev.*, vol. 19, no. 12, pp. 3571–3580, 2018.

[10] Z. Stjepanovič, "Computer-aided processes in garment production Features of CAD/CAM hardware," *Int. J. Cloth. Sci. Technol.*, vol. 7, no. 2–3, pp. 81–88, 1995.

[11] J. W. Sellors and R. Sankaranarayanan, "Colposcopy and Treatment of Cervical Intraepithelial Neoplasia. A Beginner's Manual," *IARC Press*, 2013.

[12] D. K. Chaturvedi, "Soft Computing Techniques and Their Applications," pp. 247–256, 2015.

[13] S. R. Kannan, "A new segmentation system for brain MR images based on fuzzy techniques," *Appl. Soft Comput. J.*, vol. 8, no. 4, pp. 1599–1606, 2008.

[14] M. Liang, G. Zheng, X. Huang, G. Milledge, and A. Tokuta, "Identification of abnormal cervical regions from colposcopy image sequences," *21st Int. Conf. Cent. Eur. Comput. Graph. Vis. Comput. Vision, WSCG 2013 - Commun. Pap. Proc.*, pp. 130–136, 2013.

[15] V. Gupta, A. Tandon, A. Nanda, A. Sharma, and N. Bansal, "Colposcopic evaluation of cervical lesions: a prospective study," *Int. J. Clin. Trials*, vol. 1, no. 3, p. 110, 2014.

[16] K. R. Kavitha, A. Gopinath, and M. Gopi, "Applying Improved SVM Classifier for Leukemia Cancer Classification using FCBF," pp. 61–66, 2017.

[17] W. Feng and K. Kunkle, "A Deep Belief Network Based Machine Learning System for Risky Host Detection."

[18] N. Bouchra, A. Aouatif, N. Mohammed, and H. Nabil, "Deep Belief Network and Auto-Encoder for Face Classification," vol. 5, pp. 22–29, 2018.

[19] P. Tamilselvan and P. Wang, "Structural Health Diagnosis Using Deep Belief Network Based State Classification," *Collect. Tech. Pap. - AIAA/ASME/ASCE/AHS/ASC Struct. Struct. Dyn. Mater. Conf.*, 2012.

[20] M. Sabzevari, S. R. Quchani, S. Toosizadeh, and V. Abrishami, "A fast and accurate facial expression synthesis system for color face images using face graph and deep belief network," *ICEIE 2010 - 2010 Int. Conf. Electron. Inf. Eng. Proc.*, vol. 2, no. ICEIE, pp. 354–358, 2010.

[21] S. Kamada and T. Ichimura, "An Adaptive Learning Method of Deep Belief Network by Layer Generation Algorithm.," pp. 2–5, 2018.

[22] World Health Organization, "WHO Guidance Note. Comprehensive Cervical Cancer Prevention And Control : A Healthier Future For Girls And Women http://www.who.int/cancer/detection/en/," 2013.

[23] P. A. Cohen, A. Jhingran, A. Oaknin, and L. Denny, "Cervical cancer," *Lancet*, vol. 393, no. 10167, pp. 169–182, 2019.

[24] P. Saikia, P. Vij, and R. D. Baruah, "Unsupervised Pre-training on Improving the Performance of Neural Network in Regression," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2018-July, no. July, pp. 1–6, 2018.

[25] A. Fischer, "Training Restricted Boltzmann Machines," *KI - Künstliche Intelligenz*, vol. 29, no. 4, pp. 441–444, 2015.

[26] G. E. Hinton, "A practical guide to training restricted boltzmann machines," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7700 LECTU, pp. 599–619, 2012.

[27] B. Gopika, K. Srilaxmi, D. Alekhya, B. B. Rao, and B. Rama Mohan, "Face Recognition by Using Gabor Feature Extraction and Neural Networks," *IOSR J. Electron. Commun. Eng. Ver. II*, vol. 10, no. 2, pp. 2278–2834, 2015.

[28] B. Soediono, "Backpropagation," *J. Chem. Inf. Model.*, vol. 53, p. 160, 1989.

[29] S. Visa, B. Ramsay, A. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection," *CEUR Workshop Proc.*, vol. 710, pp. 120–127, 2011.

[30] D. A. Adyanti *et al.*, "Analisis Citra Dental Panoramic Radiograph (Dpr) Pada Tulang Mandibula Untuk Deteksi Osteoporosis Menggunakan Metode GLCM – SVM Multiclass (Gray Level Co-Occurrence Matrix – Support Vector Machine Multiclass)," 2018.

[31] 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, 2019.

[32] Guneva Foundation, "Geneva Foundation for Medical Education and Research," 2018. [Online]. Available: www.gfmer.ch. [Accessed: 20-Sep-2019].

[33] M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, "Breast Cancer Classification from Histopathological Images with Inception Recurrent Residual Convolutional Neural Network," *J. Digit. Imaging*, pp. 1–13, 2019.