VOL. 13, NO. 2, JANUARY 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

INTEGRATED IMAGE PROCESSING ANALYSIS AND NAÏVE BAYES CLASSIFIER METHOD FOR LUNGS X-RAY IMAGE CLASSIFICATION

M. Arief Bustomi[1], Anifatul Faricha[2], Alfiana Ramdhan[1] and Faridawati[1]
[1]Department of Physics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
E-Mail: bustomi@physics.its.ac.id

ABSTRACT

To diagnose the lungs condition, commonly, the radiologists analyze the lungs image purely based on the X-ray image result with the naked eye. Hence, this method leads the subjectivity issue. In this study, the combination of image processing analysis and Naïve Bayes Classifier (NBC) are expected to overcome the issue. In the image processing analysis, we used the median filter and adaptive histogram equalization to enhance the lungs X-ray image quality. The five image features i.e., the feature mean, the feature SD, the feature kurtosis, the feature skewness, and the feature entropy were determined to obtain the characteristic of each lungs condition i.e., the normal lungs, the pleural effusion, and the lung cancer. In the NBC method, the five image features were used as the predictors to determine the lungs class i.e., the normal lungs class, the pleural effusion class, and the lung cancer class. The classification using NBC method consisted of two processes i.e., the training process and the validation process. The training process included the total numbers of 90 lungs X-ray images, whereas the validation process used the total numbers of 60 lungs X-ray images. According to the numerical calculation in the validation process, the performance of NBC method has 70% accuracy.

Keywords: lungs X-ray image, image processing, image features, naïve bayes classifier.

INTRODUCTION

In the medical area, the lungs which part of the vital organs are difficult to diagnose. Basically, to diagnose the lungs condition, the X-ray technology (*Rontgen*) is used to capture the lungs image [1]. The X-ray gives particular image result for normal lungs or abnormal lungs which shown in Figure-1. As we can see in Figure-1, the X-ray technology produces specific image result for each lungs condition. For example, at the normal lungs, the color of the lungs X-ray image result is completely opaque black. Whereas, the lungs X-ray image result from the abnormal lungs is not fully opaque black [2], [3].

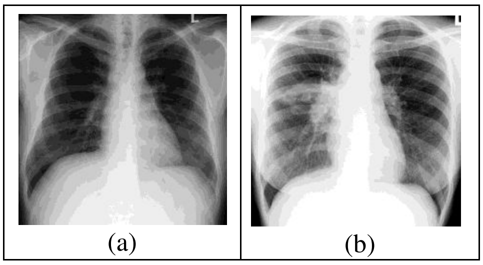

Figure-1. The lungs X-ray image results: (a) Normal lungs, (b) Abnormal lungs.

To assess the lungs condition, generally, the radiologists only use the visual observation to analyze the lungs X-ray image [4]. Hence, the diagnosis results are non-objective and lead the subjectivity issue (i.e., based on the radiologist's personal experience). Therefore, a digital tool which is able to classify the lungs condition is completely advantageous to assist the radiologist in analyzing the lungs condition.

In these current days, the development of the image processing technology has been increased extensively, particularly in the medical area. For example, reported by Anant Madabhushi and George Lee, the using of image processing improved the modeling of the disease appearance which difficult to be observed by a pathologist [5]. According to Shivangi Jain *et al.,* and B. N. Kumar *et al*., the image processing tools could also be used to detect the presence of the skin cancer and the Glaucoma (the most lethal eye disease) [6], [7]. In the image processing analysis, various techniques have been implemented to enhance the image quality such as the median filter and the adaptive histogram equalization [8]. The median filter technique offers the noise reduction of the image, whereas the adaptive histogram equalization is used to improve the image contrast [9].

The Naïve Bayes Classifier (NBC) is one of the classification methods which based on the statistics. It is broadly reported that the pros of using NBC method are easy to build with no complicated iterative parameter. Besides its simplicity, the NBC method often outperforms than other sophisticated classification methods [10], [11]. The previous research showed that the NBC methods successfully implemented in the several cases i.e., as a tool for public health surveillance from large administrative databases, forecasting the hospitalization and complications for European children, and as a tool for diagnosing the B-Chronic Lymphocytic Leukemia [12]-[14].

This study investigated the implementation of the image processing analysis at the X-ray image results and proposed the Naïve Bayes Classifier (NBC) method to classify the lungs condition. In the section of materials and methods, we describe the image processing analysis and the Naïve Bayes Classifier. Furthermore, in the next section, we demonstrate the results and verification. Finally, we present our conclusion in the last section.

MATERIALS AND METHODS

Materials

In this study, the lungs X-ray images were obtained from Airlangga University Hospital. The X-ray images are grouped into several lungs conditions i.e., the normal lungs, the pleural effusion, and the lung cancer shown in Figure-2a, 2b, and 2c sequentially. The total numbers of X-ray image samples used in this study were 150, which consisted of 50 normal lungs X-ray images, 50 pleural effusion X-ray images, and 50 lung cancer X-ray images. The amounts of 90 X-ray image samples were used for the training process. In the training process, we used the total numbers of 30 X-ray images samples from each group. Whereas, the total numbers of 60 X-ray image samples were used as the validation process, which included 20 normal lungs X-ray images, 20 pleural effusion X-ray images, and 20 lung cancer X-ray images.

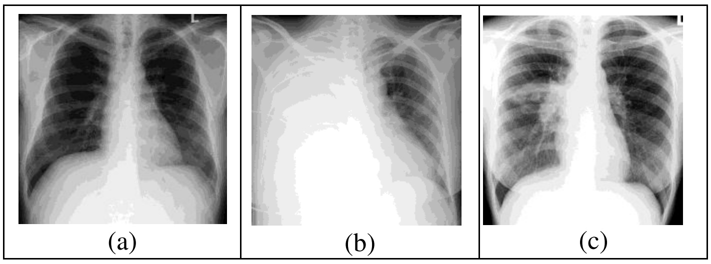

Figure-2. The X-ray image results from several lungs conditions: (a) Normal lungs, (b) Pleural effusion, (c) Lung cancer.

METHODS

Feature extraction at the image processing analysis

In this study, we applied the image processing to enhance the quality of X-ray image samples. Firstly, the median filter technique was used to filter out the noise from the lungs X-ray image sample [15]. Furthermore, to improve the contrast at the lungs X-ray image sample, we used the adaptive histogram equalization [16]. Figure-3 describes the image processing analysis implemented to the lung cancer X-ray image sample. As we can see in Figure 3b, after applyingthe median filter technique, the noise at the X-ray image sample vanishes but it has the low contrast level. As a result, the detail of X-ray image result is still unclear. Therefore, the adaptive histogram equalization is implemented to enhance the contrast level at the X-ray image sample as shown in Figure 3c.

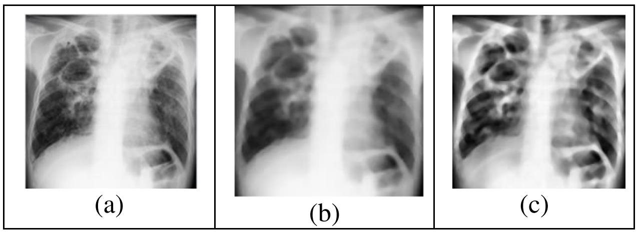

Figure-3. The image processing analysis applied at the X-ray image result from the lung cancer case: (a) The initial image sample, (b) Applying the median filter technique, (c) Applying the adaptive histogram equalization.

The feature extraction at the image processing analysis is used to reduce the dimension size and to obtain the characteristic of the X-ray image samples from each group (i.e., the normal lungs, the pleural effusion, and the lung cancer). In this study, we used the five image features i.e., the feature mean (A_1), the feature SD (A_2), the feature kurtosis (A_3), the feature skewness (A_4), and the feature entropy (A_5) [17]-[19]. The definition and the formula of each feature are as follow:

***a.* The feature mean**

The feature mean (A_1) shows the average of grayscale intensity value from the total pixels of an image.

$$A_1 = \sum_{n=1}^{N} f_n p(f_n). \tag{1}$$

***b.* The feature SD**

The feature SD (A_2) determines the amount of grayscale intensity variation from each pixel against the feature mean (A_1) of an image.

$$A_2 = \sqrt{\sum_{n=1}^{N} (f_n - A_1)^2 p(f_n)}. \tag{2}$$

***c.* The feature skewness**

The feature skewness (A_3) calculates the asymmetry level from the pixels distribution around the feature mean (A_1) of an image.

$$A_3 = \frac{1}{(A_2)^3} \sum_{n=1}^{N} (f_n - A_1)^3 p(f_n). \tag{3}$$

***d.* The feature kurtosis**

The feature kurtosis (A_4) is a measure of the normal distribution level at the intensity distribution in the pixels of an image.

VOL. 13, NO. 2, JANUARY 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

$$A_4 = \frac{1}{(A_2)^4}\sum_{n=1}^{N}(f_n - A_1)^4\, p(f_n) - 3. \tag{4}$$

***e.* The feature entropy**

The feature entropy (A_5) explains the randomness of the intensity value of an image (cline, *et al.*, 1990:1037-1045).

$$A_5 = -\sum_{n=1}^{N} p(f_n)\log_2 p(f_n). \tag{5}$$

Where:

n	=	Number of pixels of an image, n=1, 2, ….., N.
f_n	=	The grayscale intensity value.
$p(f_n)$	=	The histogram value of an image.
A_1	=	The feature mean of an image.
A_2	=	The feature SD of an image.
A_3	=	The feature skewness of an image.
A_4	=	The feature kurtosis of an image.
A_5	=	The feature entropy of an image.

Naïve Bayes classifier (NBC)

The Naïve Bayes Classifier (NBC) is used for the classification method which follows the Bayes' theorem i.e., having the strong independent assumption between the predictors [20]. In this study, the classification using the NBC method included two processes i.e., the training process and the validation process. The total numbers of 90 lungs X-ray image samples were included for the training process, which consisted of 30 normal lungs X-ray images, 30 pleural effusion X-ray images, and 30 lung cancer X-ray images.

In the training process, the lungs class was divided into three classes i.e., the normal lungs class, the pleural effusion class, and the lung cancer class which denoted the lungs conditions i.e., the normal lungs, the pleural effusion, and the lung cancer sequentially. Furthermore, the five image features which produced by the feature extraction from the image processing analysis were used as the input parameters (predictors) in the NBC method. The five image features are the feature mean (A_1), the feature SD (A_2), the feature kurtosis (A_3), the feature skewness (A_4), and the feature entropy (A_5). Each lungs class has the specific value of the image feature. Finally, Equation 6 determines the basic model of NBC method used in this study [21], [22].

$$P\left(C / A_1,...,A_5\right) = \frac{P\left(C\right)P(A_1,\ldots,A_5\left|C\right.)}{P(A_1,\ldots,A_5)} \tag{6}$$

Where:

C	=	The lungs class i.e., normal lungs class, pleural effusion class, and lung cancer class.
A	=	The predictors i.e., the five image features.
$P(C\|A_1,\ldots,A_5)$	=	The posterior probability of the class which given by the predictors.
$P(C)$	=	The prior probability of the class.
$P(A_1,\ldots,A_5\|C)$	=	The probability of the predictors which given by the class.
$P(A_1,\ldots,A_5)$	=	The prior probability of predictors.

It is broadly reported that the classification process using the NBC methods requires the two known parameters i.e., the mean of a class, and the standard deviation of a class [23]. In this study, we divided the lungs class into three classes i.e., normal lungs class, pleural effusion class, and lung cancer class. Each class is represented by the five image features i.e., A_1, A_2, A_3, A_4, and A_5. However, in this study, we used the total numbers of 30 lungs X-ray images for each class in the training process. Hence, each class has many image features. For example, the normal lungs class has the total number of 30 feature mean (A_1), 30 feature SD (A_2), 30 feature kurtosis (A_3), 30 feature skewness (A_4), and 30 feature entropy (A_5). Therefore, in this study, we have the total numbers of five mean per class i.e., mean of A_1, mean of A_2, mean of A_3, mean of A_4, and mean of A_5. Finally, the mean of each image feature in a class is determined by Equation 7, whereas the Equation 8 calculates the standard deviation of each image feature in a class.

$$\mu_i = \frac{1}{M}\sum_{j=1}^{M=30} A_{i(j)} \tag{7}$$

$$\sigma_i = \sqrt{\frac{1}{M-1}\sum_{j=1}^{M=30}(A_{i(j)} - \mu_i)^2} \tag{8}$$

Where:

μ_i	=	The mean of an image feature in a class (μ_1, μ_2, μ_3, μ_4, and μ_5 define the mean of the image feature A_1, A_2, A_3, A_4, and A_5 sequentially).
j	=	The total numbers of images used in the training process in a class (j=1, 2, ….., 30).
A_i	=	The image features (A_1, A_2, A_3, A_4, and A_5).
σ_i	=	The standard deviation of an image feature in a class (σ_1, σ_2, σ_3, σ_4, and σ_5 determine the standard deviation of the image feature A_1, A_2, A_3, A_4, and A_5 sequentially).

In this study, the validation process includes the total numbers of 60 lungs X-ray image samples i.e., 20 normal lungs X-ray images, 20 pleural effusion X-ray images, and 20 lung cancer X-ray images. The amounts of 60 lungs X-ray image generates the specific image features which to be classified into the particular lungs class (i.e., normal lungs class, pleural effusion class, and lung cancer class). The validation process using NBC method follows the Gaussian distribution probability formula, which derived in Equation 9 [24].

$$P(A_i \mid C) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left[-\frac{1}{2}\left(\frac{(A_i-\mu_i)^2}{\sigma_i^2}\right)\right] \qquad (9)$$

Where:

μ_i	=	The mean of an image feature in a class obtained from training process.
σ_i	=	The standard deviation of an image feature in a class obtained from training process.
A_i	=	The image features (A_1, A_2, A_3, A_4, and A_5).
C	=	The lungs class.
$P(A_i \mid C)$	=	The Gaussian distribution probability of the image feature given by the class.

To describe the performance of the NBC method, the 3x3 confusion matrix table was applied for this study, which shown in Table-1 [25], [26]. In the confusion matrix, the actual result is the result based on the observation (reality), whereas the predicted result is assessed by the NBC method. Both of actual result and predicted result have three lungs classes i.e., the normal lungs class, the pleural effusion class, and the lung cancer class denote the normal lungs condition, the pleural effusion condition, and the lung cancer condition sequentially. Then the case was divided into nine cases: *TN*, *FN1*, *FN2*, *FP1*, *TP*, *FP2*, *FL1*, *FL2*, and *TL*. For example, the *TN* case is when the actual result and the predicted result give the same conclusion i.e., the normal lungs class. The *FN1* case is when the actual result has the pleural effusion result, but the predicted result gives the normal lungs result (For more detail of each case, see Table-1). Finally, the accuracy (*AC*) of the NBC method in the validation process is given by Equation 10[27], [28].

Table-1. The 3x3 confusion matrix table.

		Actual Result		
		Normal lungs	**Pleural effusion**	**Lung cancer**
Predicted Result	Normal lungs	*TN*	*FN1*	*FN2*
	Pleural effusion	*FP1*	*TP*	*FP2*
	Lung cancer	*FL1*	*FL2*	*TL*

Where:

TN	=	Correctly classified normal lungs class.
TP	=	Correctly classified pleural effusion class.
TL	=	Correctly classified lung cancer class.
FN 1	=	Classified pleural effusion class into normal lungs class.
FN 2	=	Classified lung cancer class into normal lungs class.
FP1	=	Classified normal lungs class into pleural effusion class.
FP2	=	Classified lung cancer class into pleural effusion class.
FL1	=	Classified normal lungs class into lung cancer class.
FL2	=	Classified pleural effusion class into lung cancer class.

$T = TN + TP + TL$;

$$AC = \frac{T}{T + FN1 + FN2 + FP1 + FP2 + FL1 + FL2} \times 100\%. \qquad (10)$$

RESULTS AND VERIFICATION

In this study, the lungs X-ray images obtained from Airlangga University Hospital were applied to analyze the lungs condition, which divided into three groups i.e., the normal lungs, the pleural effusion, and the lung cancer. To improve the lungs X-ray images quality, we implemented the image processing analysis. In the image processing analysis, each lungs X-ray image was determined to obtain the five image features i.e., the feature mean (A_1), the feature SD (A_2), the feature kurtosis (A_3), the feature skewness (A_4), and the feature entropy (A_5). Then the five image features would be used for the input parameter in the classification process.

The Naïve Bayes Classifier was selected as the classification method in this study. The classification process using NBC method was divided into two processes i.e., the training process and the validation process. The training process included the total numbers of 90 lungs X-ray image samples, whereas the validation process used the total numbers of 60 lungs X-ray image samples. In the training process, the lungs class as the output was divided into three classes i.e., the normal lungs class, the pleural effusion class, and the lung cancer class, which denoted the normal lungs condition, the pleural effusion condition, and the lung cancer condition sequentially. Whereas, the five image features (i.e., A_1, A_2, A_3, A_4, A_5) were employed as the input parameter (predictors).

The classification process using NBC methods needs two known parameters i.e., the mean of a class and the standard deviation of a class. In this study, each lungs X-ray image was represented by the five image features i.e., A_1, A_2, A_3, A_4, A_5. Therefore, the total numbers of mean per class were five, i.e., the mean of A_1, the mean of A_2, the mean of A_3, the mean of A_4, and the mean of A_5, which denoted as μ_1, μ_2, μ_3, μ_4, and μ_5 respectively. In addition, the total numbers of standard deviation per class were also five, i.e., the standard deviation of A_1, the standard deviation of A_2, the standard deviation of A_3, the standard deviation of A_4, and the standard deviation of A_5, which symbolized by σ_1, σ_2, σ_3, σ_4, and σ_5 respectively. Table-2 and Table-3 present the mean value and the standard deviation from several lungs class i.e., normal lungs class, pleural effusion class, lung cancer class. As we can see from Tables 2 and 3, every lungs class has the specific value of mean and standard deviation which shows the characteristic of each lungs condition.

Table-2. The mean value of each image features from the various lungs conditions.

Lungs class	Mean				
	μ_1	μ_2	μ_3	μ_4	μ_5
Normal Lungs	66.13	62.23	9.094	6.643	118.8
Pleural Effusion	74.47	67.83	6.753	2.890	77.72
Lung Cancer	78.78	71.45	6.694	3.772	117.9

Table-3. The standard deviation of each image features from the various lungs conditions.

Lungs class	Standard deviation				
	σ_1	σ_2	σ_3	σ_4	σ_5
Normal Lungs	12.57	5.063	0.551	2.652	85.93
Pleural Effusion	9.690	3.503	0.322	1.938	76.29
Lung Cancer	10.94	3.020	0.267	2.918	102.2

In the validation process, the Gaussian distribution probability formula was used to classify the image features generated by 60 lungs X-ray image samples as the testing, which consisted of 20 normal lungs X-ray images, 20 pleural effusion X-ray images, and 20 lung cancer X-ray images. Table-4 shows the confusion matrix result between the actual result and the predicted result. The numerical results of case *TN*, *FN1*, *FN2*, *FP1*, *TP*, *FP2*, *FL1*, *FL2*, and *TL* are 16, 2, 2, 11, 9, 0, 0, 3, and 17 sequentially. Furthermore, to assess the performance of NBC method, the accuracy (*AC*) was determined in this study. According to the numerical calculation, the validation process using NBC method has the accuracy (*AC*) of 70%, which means from the total numbers of 60 lungs X-ray image samples as testing; 42 correctly classified and 18 incorrectly classified.

Table-4. The 3x3 confusion matrix result.

		Actual result		
		Normal lungs	Pleural effusion	Lung cancer
Predicted Result	Normal lungs	16	2	2
	Pleural effusion	11	9	0
	Lung cancer	0	3	17

CONCLUSIONS

This study successfully demonstrated the integrated image processing analysis and Naïve Bayes Classifier to classify the lungs condition based on the lungs X-ray image. The lungs conditions were divided into three groups i.e., the normal lungs, the pleural effusion, and the lung cancer. In the image processing analysis, the median filter and adaptive histogram equalization were used to enhance the quality of the lungs X-ray image. Each lungs X-ray image was analyzed to produce the five image features i.e., the feature mean (A_1), the feature SD (A_2), the feature kurtosis (A_3), the feature skewness (A_4), and the feature entropy (A_5). Furthermore, the five image features were used for input parameters in the classification process. In this study, the classification process using Naïve Bayes Classifier was divided into two processes: the training process and the validation process. According to the numerical result, the Naïve Bayes Classifier has the accuracy (*AC*) of 70%. These findings indicate that the combination of the image processing analysis and the Naïve Bayes Classifier can be applied to classify the lungs condition and to assist the radiologist in diagnosing.

ACKNOWLEDGEMENT

The authors would like to acknowledge Airlangga University Hospital to provide the lungs X-ray image and LPPM ITS for the intense discussion.

REFERENCES

[1] Raziee H.; Hope A.; Faruqi S.; Yap ML, Roberts H.; Kandel S.; *et a*l. 2015. Classification and Reporting of Late Radiographic Changes after Lung Stereotactic Body Radiotherapy: Proposing a New System. Clinical lung cancer. 16(6): 245-51.

[2] J. Budna; Ł. Spychalski; M. Kaczmarek; M. Frydrychowicz; J. Goździk-Spychalska; H. Batura-Gabryel; J. Sikora. 2017. Regulatory T cells in malignant pleural effusions subsequent to lung carcinoma and their impact on the course of the disease. Immunobiology. 222(3): 499-505.

[3] Elaine The, Elizabeth Belcher. 2014. Lung cancer: diagnosis, staging and treatment. Surgery (Oxford). 32(5): 242-248.

[4] Ma X.; Phi Van; Kimm; M. A.; Prakash J.; Kessler H.; Kosanke; Wildgruber M. 2017. Integrin-Targeted Hybrid Fluorescence Molecular Tomography/X-ray Computed Tomography for Imaging Tumor Progression and Early Response in Non-Small Cell Lung Cancer. Neoplasia (New York, N.Y.). 19(1): 8-16.

[5] AnantMadabhushi; George Lee. 2016. Image analysis and machine learning in digital pathology: Challenges

and opportunities. Medical Image Analysis. 33, pp. 170-175.

[6] Shivangi Jain; Vandanajagtap; NitinPise. 2015. Computer Aided Melanoma Skin Cancer Detection Using Image Processing. Procedia Computer Science. 48, pp. 735-740.

[7] B. N. Kumar; R. P. Chauhan; N. Dahiya. 2016. Detection of Glaucoma using image processing techniques: A review. International Conference on Microelectronics, Computing and Communications (MicroCom), Durgapur. pp. 1-6.

[8] KesariVerma; Bikesh Kumar Singh; A.S. Thoke. 2015. An Enhancement in Adaptive Median Filter for Edge Preservation. Procedia Computer Science. 48, pp. 29-36.

[9] Jayprakash Upadhyay; Ayushi Jaiswal. 2015. A joint implementation of adaptive histogram equalization and interpolation. Optik - International Journal for Light and Electron Optics. 126(24): 5936-5940.

[10] Ahmed I. Saleh; Arwa E. Abulwafa; Mohammed F. Al Rahmawy. 2017. A web page distillation strategy for efficient focused crawling based on optimized Naïve bayes (ONB) classifier. Applied Soft Computing. 53, pp. 181-204.

[11] R. Vedala; B. R. Kumar. 2012. An application of Naive Bayes classification for credit scoring in e-lending platform. International Conference on Data Science & Engineering (ICDSE), Cochin, Kerala. pp. 81-84.

[12] Marucci-Wellman HR; Lehto MR; Corns HL. 2015. A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Naive Bayes algorithms. Accident; analysis and prevention. 84, pp. 165-76.

[13] Paola Berchialla; Francesca Foltran; Dario Gregori. 2013. Naïve Bayes classifiers with feature selection to predict hospitalization and complications due to objects swallowing and ingestion among European children. Safety Science. 51, pp. 1–5.

[14] Lakoumentas J.; Drakos J.; Karakantza M.; Sakellaropoulos G.; Megalooikonomou V.; Nikiforidis G. 2012. Optimizations of the naive-Bayes classifier for the prognosis of B-Chronic Lymphocytic Leukemia incorporating flow cytometry data. Computer methods and programs in biomedicine. 108(1): 158-67.

[15] Dyah Ekashanti Octorina Dewi; Ahmad Athif Mohd. Faudzi; Tati Latifah Mengko; Koichi Suzumori. 2017. Recursive Gauss-Seidel median filter for CT lung image denoising. Proc. SPIE 10225, Eighth International Conference on Graphic and Image Processing.

[16] H. Kaur; J. Rani. 2016. MRI brain image enhancement using Histogram Equalization techniques. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai. pp. 770-773.

[17] Shofwatul 'Uyun; Sri Hartati; AgusHarjoko; Subanar. 2013. Selection Mammogram Texture Descriptors Based on Statistics Properties Backpropagation Structure. International Journal of Computer Sciences and Information Security. 11(5).

[18] S. bintiYusoff; Y. Bee Wah. 2012. Comparison of conventional measures of skewness and kurtosis for small sample size. International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi. pp. 1-6.

[19] G. M. Georgiou; K. Voigt. 2015. Stochastic computation of moments, mean, variance, skewness and kurtosis. in Electronics Letters. 51(9): 673-674.

[20] GherardoVarando; Concha Bielza and Pedro Larrañaga. 2016. Decision functions for chain classifiers based on Bayesian networks for multi-label classification. International Journal of Approximate Reasoning. 68, pp. 164-178.

[21] SergiosTheoridis and KonstantinosKoutroumbas. 2009. Pattern Recognition. Elsevier: San Diego.

[22] Devesh Kumar; Ravinder Singh; Abhishek Kumar; Nagesh Sharma. 2015. An Adaptive Method of PCA for Minimization of Classification Error Using Naïve Bayes Classifier. Procedia Computer Science. 70, pp. 9-15.

[23] F. Gumus; C. O. Sakar; Z. Erdem; O. Kursun. 2014. Online Naive Bayes classification for network intrusion detection. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), Beijing. pp. 670-674.

[24] X. Deng; J. Guo, Y. Chen; X. Liu. 2012. A Method for Detecting Document Orientation by Using NaÏve Bayes Classifier. International Conference on Industrial Control and Electronics Engineering, Xi'an. pp. 429-432.

[25] Ghani; N. L. A.; Abidin; S. Z. Z.; Khalid N. E. A. 2015. Accuracy assessment of urban growth pattern classification methods using confusion matrix and ROC analysis. In Communications in Computer and Information Science. 545, pp. 255-264.

[26] A. Leijon; G. E. Henter; M. Dahlquist. 2016. Bayesian Analysis of Phoneme Confusion Matrices. In IEEE/ACM Transactions on Audio, Speech, and Language Processing. 24 (3), pp. 469-482.

[27] Bhuvaneswari T; Prabaharan S.; Subramaniyaswamy V. 2015. An Effective Prediction Analysis Using J48. ARPN Journal of Engineering and Applied Sciences. 10 (8).

[28] Wicaksono; Danang Triantoro Murdiansyah; Fhira Nhita. 2016. Optimization of fuzzy inference system's parameter using evolutionary programming algorithm for teacher certification in Indonesia. 4th International Conference on Information and Communication Technology (ICoICT). pp. 1-4.