TELKOMNIKA, Vol.17, No.3, June 2019, pp.1492~1500
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v17i3.12240

Improving of classification accuracy of cyst and tumor using local polynomial estimator

Nur Chamidah*[1], Kinanti Hanugera Gusti[2], Eko Tjahjono[3], Budi Lestari[4]
[1,2,3]Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, Indonesia
[4]Department of Mathematics, FMIPA, The University of Jember, Kampus Tegal Boto, Indonesia
*Corresponding author, e-mail: nur-c@fst.unair.ac.id[1], kinantihanugera@gmail.com[2], eko-t@fst.unair.ac.id[3], budi.fmipa@unej.ac.id[4]

Abstract

Cyst and tumor in oral cavity are seriously noticed by health experts along with increasing death cases of oral cancer in developing country. Early detection of cyst and tumor using dental panoramic image is needed since its initial growth does not cause any complaints. Image processing is done by mean for distinguishing the classification of cyst and tumor. The results in previous studies about classification of cyst and tumor were done by using a mathematical computation approach namely supports vector machine method that have still not satisfied and have not been validated. Therefore, in this study we propose a method, i.e., nonparametric regression model based on local polynomial estimator that can be improve the classification accuracy of cyst and tumor on human dental panoramic image. By using the proposed method, we get the classification accuracy of cyst and tumor, i.e., 90.91% which is greater than those by using the support vector machine method, i.e., 76.67%. Also, in validation process we obtain that the nonparametric regression model approach gives a significant Press's Q statistical testing value. So, we conclude that the nonparametric regression model approach improves the classification accuracy and gives better outcome to classify cyst and tumor using dental panoramic image than the support vector machine method.

Keywords*: accuracy, classification, cyst, tumor, local polynomial*

1. Introduction

Statistical modeling plays important role in data analysis, for examples, predicting, forecasting, classifying and interpreting of the nature phenomenon based on data. There are two approaches for building the statistical model, i.e., parametric regression model and nonparametric regression model approaches. We can use parametric regression model when researchers assume the form of regression function is known, for examples, linear, quadratic, cubic, etc. Yet, if the form of regression function is unknown, we can use nonparametric regression model. There are many estimators in the nonparametric regression model, for example, spline estimator, kernel, local linear, wavelet estimator, etc. In the real cases, we are often faced to the problem in which response variable is observed at several values of the predictor variables. Spline estimator has been studied by [1-6]. Designing growth charts of children up to five years old based on local polynomial modeling, kernel and local linear, respectively have been done by [7-10].

Tumor is categorized as an alarming disease as being positioned number seven cause of death in Indonesia with 5.7% death rate of total population [11]. Currently, the case of tumor is increasing because it is often newly detected in an advanced stage after the emergence of clinical symptoms felt by the patient. According to [12], the diseases that occur in the oral cavity have become a serious concern by experts along with the increasing cases of death caused by cancer of the oral cavity due to nearly 75% cases occur in developing countries.

In addition, tumor and cyst are found in the oral cavity which early detection of them is still difficult due their initial symptoms do not cause complaint [13]. A panoramic image facilitates health experts to see abnormalities of them. However, the determination of the abnormality has not been optimum because it is difficult to distinguish cells based on naked eye and depend on the knowledge/experience of doctors. Besides that in developing countries like Indonesia, the panoramic machines available have low resolution which affect on resulting image [14].

Received June 4, 2018; Revised January 9, 2019; Accepted February 7, 2019

Whereas, the certainly of those abnormalities can be known after the biopsy process (body tissues taking for laboratory examination) with expensive cost and take a long time, so that it requires a computer programming to help classification of human teeth panoramic images results which produce high dimension data in the form of $(n \times p)$-matrix with p (number of predictor variables) is more than n (number of observations).

Programming which done by researcher is the image processing to improve quality of image and facilitates them for giving information based on image [15-18]. Studies related to the classification of cyst and tumor have been done by several researchers, i.e., [19-21] who used a mathematical calculation approach namely the support vector machine (SVM) method. The study by [19] used preprocessing like histogram equalization, normalization and manual segmentation with doctor's help and got the result of classification with accuracy of 76.67%. Preprocessing based on texture analysis which supports the use of extraction features (features of histogram and gray level co-occurrence matrix (GLCM)) has been proposed by [20] that produced 63.33% classification accuracy, and the study by [21] used image processing like Gaussian filter, normalization, and active contour segmentation that gave 79.85% of classification accuracy. Other studies such as [22] who used statistical calculation approach namely wavelet transformation method and local linear estimator in detection of brain tumors produced 100% classification accuracy. The statistical approach requires $n > p$, so the dimension reduction is required for further analyzing data.

Researchers [22-25] used wavelet transform for reduction dimension, but in statistical modelling, the method as a dimension reducer required a principal component analysis to overcome the existence of multi-collinear in dimension reduction results. The research requires two reducing stages, so in this study the statistical method to be used is GEFA to omit multi-collinear cases in the reduction results and shorten the reduction stage [26]. The result of image data by using GEFA dimension reducer is analyzed further by using statistical approach to get the image classification. The statistical approach used is a parametric regression model that has good and simple properties for simply statistical inference, such as easy interpretation, and linear, efficient, consistent estimators, also best linear unbiased estimator (BLUE) [27]. While nonparametric regression analysis is flexible in modelling data patterns [28], and the estimator used is a local polynomial estimator that can reduce asymptotic bias and produces a good estimation [29].

The data set used in this study are categorized into two categories namely cyst image $(Y = 0)$ and tumor image $(Y = 1)$, so that the response variable on the dental panoramic image data follows the Bernoulli distribution. One of the statistical methods that can be used is binary logistic regression. Logistic regression is a statistical analysis method that is used to describe relationships between variables, and hypotheses testing of categorical response to one or more predictor variables whether continuous or categorical uses logit link function [30]. In statistical modelling, there are two approaches, i.e., parametric approach by using GLM and nonparametric approach by using GAM. In GLM and GAM, we estimate the model by estimating its parameters and its regression functions, respectively. In GAM, we use local scoring algorithm to estimate its regression functions to accommodate nonparametric additive regression models whose distribution of response variables is included in an exponential family [31] whereas Bernoulli distribution included in the exponential distribution family.

In [19, 20] the used of mathematical computation approaches gave the accuracy of classifications for cyst and tumor on human dental panoramic images, i.e., 76.67% according to [19] and 63.3 % according to [20], that have still not satisfied. Therefore, in this paper, we propose a new method by using statistical modelling approach to classify cyst and tumor on human dental panoramic images. The statistical model with logit link function used are parametric model (GLM) and nonparametric model (GAM) approaches, then we compare the accuracy classifications between GLM and GAM approaches. In addition, to validate the accuracy classifications based the best model, we use Press's Q statistical testing.

2. Research Method

The analysis steps of this research are as follows:

1. Steps of image processing on 22 human dental panoramic images are given as Figure 1:

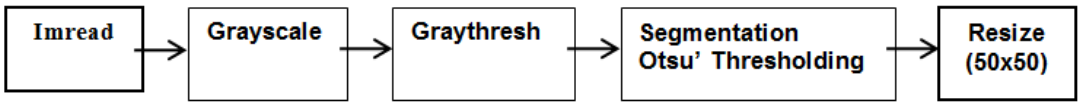

Figure 1. Steps of image processing on 22 human dental panoramic

Resizing of image size which is represented by a matrix where predictors are as columns of matrix and observations are as rows of matrix.

2. Reducing dimension of image processing human dental panoramic image by using GEFA method, we create R-code with the following algorithm:
   a. Obtain **Z**, a matrix of predictors with dimension (22×2500) from step 1.
   b. Determine the number of observations (n), number of predictors (p), and the number of predictors after reducing (k) under terms $n < p$, $k < p$, and $k < n$.
   c. Construct **F**, a reduction matrix of **Z** with dimension $(n\times k)$, and **U**, a matrix of random number generated from singular value decomposition on each element with dimension $(n\times p)$.
   d. Determine the value of old function, $f_{old} = \|Z\|_F^2$
   e. Minimize the equation of $f(F,L,T,\psi) = \|Z - FL^T - T\psi\|_F^2$ under iteration procedure as follows:
      i. calculate the first loading factor $L = Z^T F$ and initial specific variety $\psi = diag(T^T Z)$
      ii. make block matrix $B = [F \quad U]$ and $A = [L \quad \psi]$
      iii. determine f value by using formula $f = \|Z\|_F^2 + tr(B^T BA^T A) - 2tr(B^T ZA)$
      iv. calculate value of **B*** by using formula $B^* = UV^T$ where **U** and **V** are singular value decomposition of $A^T Z = UDV^T$
      v. obtain matrix $F^* = B^*[1 \quad k]$ and $U^* = B^*[k+1 \quad p]$, so that matrices **L*** and ψ^* are obtained
      vi. repeat steps (i) to (v) until the convergence is completed under term $|f_{old} - f| < 10^{-6}$
   f. Get the reduction matrix **F*** that ready will be used as the predictors in the model.
3. Making the classification of cyst and tumor from the result of dimension reduction data (in step 2) by using parametric logistic regression (GLM) approach through the following steps:
   a. Estimate binary logistic regression with logit link function.
   b. Estimate probability value on each observation by using formula $\pi_i = \dfrac{\exp(X_i\beta)}{1+\exp(X_i\beta)}$.
   c. Classify π_i value based of cut off probability value for 0.5.
   d. Estimate the accuracy of classification based on apparent error rate (*APPER*).
4. Making the classification of cyst and tumor from the result of dimension reduction data (in step 2) by using nonparametric logistic regression (GAM) approach and by creating R-code through on the following algorithm:
   a. Define response variable (Y) and predictors variable (X) obtained from step 2
   b. Determine orde and optimum bandwidth (*h*) based on minimum cross validation (CV) criterion for each predictor by using local polynomial estimator.
   c. Estimate GAM based on local polynomial estimator by using local scoring algorithm as follows:

i. estimate function $\mu(\mathbf{x}) = \frac{\exp\left(m_{h_1}(x_1) + m_{h_2}(x_2) + m_{h_3}(x_3) + m_{h_4}(x_4)\right)}{1 + \exp\left(m_{h_1}(x_1) + m_{h_2}(x_2) + m_{h_3}(x_3) + m_{h_4}(x_4)\right)}$

where $\hat{m}_{h_j}(x_i) = \boldsymbol{X}(x_0)\left(\boldsymbol{X}^T(x_0)\boldsymbol{K}_{h_j}(x_0)\boldsymbol{X}(x_0)\right)^{-1}\boldsymbol{X}^T(x_0)\boldsymbol{K}_{h_j}(x_0)\boldsymbol{y}$, $\mathrm{j} = 1,2,3,4$ and the orde and optimum bandwidth of each predictor (h_j) obtained from step 4 (b).

ii. determine the initial value which is used in initial iteration $r = 0$

iii. iterate the scoring iteration (outer loop) for $(r = 0,1,2,...)$ with these following steps:

(a). Iterate weighted back fitting (inner loop) for $(s = 0,1,2,...)$ as follows:

(i) Define $\hat{m}_j^{(s)}(X_j) = \hat{m}_j^{(r)}(X_j)$ and $Avg(RSS)^{(s)} = Avg(RSS)^{(r)}$ while initial iteration $(s = 0)$

(ii) Define residual partial value

$$R_j^{(s+1)} = z - \sum_{k=1}^{j-1} m_k^{(s+1)}(X_k) - \sum_{j=k+1}^{p} m_k^{(s+1)}(X_k)$$

(iii) Determine the finer functions in model $m_j^{(s+1)}(X_j) = A(h_j)R_j^{(s+1)}$

(iv) Determine the average of weighted sum square error ($Avg(RSS)^{(s+1)}$) by using formula $Avg(RSS)^{(s+1)} = \frac{1}{n}\left\{(z^{(s+1)} - \eta^{(s+1)})^T B^{(s+1)}(z^{(s+1)} - \eta^{(s+1)})\right\}$

(v) Repeat step (ii) to (iv) until the value of difference between $Avg(RSS)^{(s+1)}$ and $Avg(RSS)^{s}$ is convergent

(b). Determine matrix $B^{(r+1)}$ and vector $z^{(r+1)}$ based on μ_i and η_i values which obtained after the convergence of $Avg(RSS)$ is fulfilled

(c). Calculate the average of deviance

$$Avg(D(y_i \quad \mu_i))^{(r+1)} \approx \frac{-2}{n}\sum_{i=1}^{n}\left\{y_i \log \mu_i^{(r+1)} + (1 - y_i)\log(1 - \mu_i^{(r+1)})\right\}$$

(d). Repeat step (iv) to (v) until the value of difference between $Avg(Dev)^{(s+1)}$ and $Avg(Dev)^{s}$ is convergent.

d. Determine the cut of probability value (threshold value) based on the highest classification accuracy score.

e. Define the initial incidence of cyst and tumor and classify the estimation of μ_i value

f. Calculate and compare the accuracy classification values of cyst and tumor between parametric and nonparametric models approaches.

3. Results and Analysis

3.1. Image Processing

The purpose of image processing is to improve the image quality for the retrieval of existing information in the image can be processed to the next stage [18]. Stages of image processing in this study is done by using a software with several stages, starting from reading process of panoramic image data file, histogram equalization process, otsu' thresholding segmentation, and image resizing process. Results are given in Figure 2.

3.2. Reducing Dimensions of Images Processing Result Using GEFA Method

Dimensions reducing is necessary because the results obtained in image processing have very high dimensions and do not meet the criterion of the data analyzed by statistical

approach. This criterion must be met such that the estimation and inference process can be applied. Dimension reducing method used in this study is GEFA which is selected for the results of the reduction given in new variables that are mutually independent between each other [26].

In this study, the data set obtained by image processing has number of predictor variable of 2500 and then reduced to 4 predictors. The GEFA method uses the convergence criterion to perform the iteration process and it will stop if the converging criterion is met. The iteration process on the human dental panoramic image data stops at the 26th-iteration with the value of $|f_{old} - f_{new}| = 6.16\times10^{-7}$ which has met converging criterion at 10^{-6}. Matrix $\mathbf{F}^*$ obtained from OSS-R as reduced data is used as a predictor data for classifying of cyst and tumor in human dental panoramic image.

Figure 2. Cyst image–histogram equalization–otsu' thresholding segmentation

3.3. Classifying Cyst and Tumor by Using GLM and GAM Approaches

In this section we determine classification cyst and tumor by using a GLM approach, i.e., parametric logistic regression model, and by using a GAM approach, i.e., nonparametric logistic regression model based on local polynomial estimator. From 22 insample observations, we reduce dimensions by using GEFA such that produces 4 predictor variables of 2500 predictors which later we will model them to determine the classification of cyst and tumor in the image. Firstly we determine classification cyst and tumor by using parametric logistic regression model approach with link function, logit. In this step, we estimate parameters, do simultaneously or individually significant testing, and calculate classification accuracy in each observation. The estimating of parameters results are given in Table 1.

Table 1. Parameter Model Estimation for GLM

Predictor	Coeff	SE Coeff	Z	P
Constant	24.4	18.3234	1.33	0.184
X1	89.4	66.9229	1.34	0.182
X2	57.8	43.6072	-1.33	0.185
X3	75.6	57.5643	1.31	0.189
X4	25.3	19.3332	1.31	0.190

Based on the analysis by helping a statistical software we got G=2.105 and p-value = 0.716. Next, we test for significance simultaneously and individually at 5% level, then there is no simultaneous effect between the predictor variables and the response. So, there is no individual variable predictor that able to explain the effect of it on the response variable. We restrict to no considering the test of significance of the parameters either simultaneously or individually, so that the obtained model can be used to calculate the accuracy of cyst or tumor image classification. In this study, each observation is estimated the probability value π_i. For example, on the 2nd observation with $x_1 = -0.07, x_2 = 0.33, x_3 = -0.03, x_4 = 0.11$ we obtain the estimated probability:

$$\pi_2 = \frac{e^{(24.4+89.4(-0.07)-57.8(0.33)+75.6(-0.03)+25.3(0.11))}}{1+e^{(24.4+89.4(-0.07)-57.8(0.33)+75.6(-0.03)+25.3(0.11))}} = 0.38$$

The estimated π_2 is less than 0.5, so that it is classified as cyst image (Y=0). Next, we determine the probability value at each observation by using the nonparametric regression model approach based local polynomial estimator. Based on step 4 (a), we obtain optimum bandwidth for each predictor as given in Table 2.

Table 2. Optimum Bandwidth Value for each Predictor Variable

No	Predictor Variable	Order	Minimum CV	Optimum Bandwidth
1	X_1	1	0.2498356	1.499949
		2	0.2699279	1.499953
2	X_2	1	0.2496131	1.499953
		2	0.2480063	0.4011054
3	X_3	1	0.2520231	1.499943
		2	0.2440422	1.499951
4	X_4	1	0.2525861	1.499942
		2	0.2612995	1.499952

Based on the optimum bandwidth values and the obtained order in Table 2, we determine the initial value $m_{h_j}(x_{ij})$ for every predictor by using the local scoring algorithm stage (step 4 (b) part (ii)). In this study, every observation has an initial value and the expected value. So, for example, we will discuss for the 2nd observation only. The 2nd observation by defining cyst image (Y=0) has initial values estimate as follows:

$$m_{h_1}(x_{12}) = \text{-}270.12+1011.83(x_1 + 0.37)$$

$$m_{h_2}(x_{22}) = 19.41\text{- } 655.69(x_2 + 0.01) + 11.04(x_2 + 0.01)^2$$

$$m_{h_3}(x_{32}) = 294.10+938.22(x_3 - 0.07) + 194.25(x_3 - 0.07)^2$$

$$m_{h_4}(x_{42}) = \text{-}45.47 + 291.72(x_4 \text{ - } 0.10)$$

Next, by using nonparametric logistic regression we get the following estimated values:

$$\hat{\eta}_2 = \text{-}270.12+19.41+294.10 \text{ - } 45.47 = \text{-}2.08$$

$$\hat{\mu}_2 = \frac{e^{-2.08}}{1+e^{-2.08}} = 0.11$$

Then, we classify categories $Y = 0$ for cyst, and $Y = 1$ for tumor. This is done by determining the threshold value in step 4(d) that will be used as a comparison or cut off probability in the classification of cyst and tumor as given in Table 3.

Table 3. Threshold Values and Accuracy Classification

No	*Threshold*	Accuracy Classification	No	*Threshold*	Accuracy Classification
1	0	50.00	8	0.40	81.8
2	0.10	72.27	9	0.50	77.27
3	0.15	86.36	10	0.80	77,27
4	0.32	90.9	11	0.92	72.72
5	0.33	90.9	12	0.97	68.18
6	0.36	90.9	13	0.99	59.09
7	0.37	86.36	14	1	50.00

Threshold that will be used as a reference for cut-off category 0 or category 1 is determined by looking at the highest classification accuracy score and highest threshold value which has the highest classification accuracy. So that the selected threshold value obtained previously equals to 0.11 that is smaller than 0.36. If value $\hat{\mu}_i$ is greater than threshold value, then it will be classified as tumor image $Y=1$, vice versa. Value $\hat{\mu}_2$ obtained based on this calculation will be classified as cyst image. The classification result based on value $\hat{\mu}_i$ in the 1st to the 22nd observations which have been analyzed by using parametric and nonparametric regression models approaches are given in Table 4.

Table 4. Classification Accuracy Using Parametric (left) and Nonparametric (right) Approaches

Observation	Prediction		Total
	Cyst	Tumor	
Cyst	7	4	11
Tumor	4	7	11
Total	11	11	22

Observation	Prediction		Total
	Cyst	Tumor	
Cyst	10	1	11
Tumor	1	10	11
Total	11	11	22

next step, we calculate the probability of errors by apparent error rate (APPER) calculation based on parametric regression model approach (left) and nonparametric regression model approach (right) as follows:

$$APPER = \frac{4+4}{7+4+4+7} \times 100\% = 36.36\% \qquad APPER = \frac{1+1}{10+1+1+10} \times 100\% = 9.09\%$$

based on APPER values, we obtain the classification accuracy values of cyst and tumor for in sample data are 63.64% (for parametric regression model approach) and 90.91% (nonparametric regression model approach). It means that the nonparametric logistic regression model approach based on local polynomial estimator is better than the parametric logistic regression model approach based on logit link function, for determining classification accuracy value of cyst and tumor in human dental panoramic images. Next, we validate the classification accuracy of nonparametric regression model approach. In this step we calculate the following Press's Q value [32]:

$$\text{Press'Q} = \frac{(\text{N-(nK)})^2}{\text{N(K-1)}} = \frac{(22-(20)(2))^2}{22(2-1)} = 14.72$$

the Press's Q is compared with $\chi^2_{(0.05,1)} = 3.841$. Because of the Press's Q =14.72 is greater than $\chi^2_{(0.05,1)} = 3.841$, then the nonparametric regression model approach based on local polynomial estimator is significantly appropriate for classifying cyst and tumor on human dental panoramic images. Finally, based on the threshold values and accuracy classification in Table 3, and APPER value and Press's Q value we conclude that by using the local polynomial estimator, the classification accuracy of cyst and tumor, i.e., 90.91% is greater than those by using the support vector machine method proposed by [19], i.e., 76.67%.

4. Conclusion

To determine the classification accuracy of cyst and tumor on human dental panoramic images, the use of a statistical modeling, i.e., nonparametric regression model approach based on local polynomial estimator is better than the use of mathematical calculation approach namely the support vector machine (SVM) methods that were proposed by [19, 20]. The Use of local polynomial estimator can improve the classification accuracy of cyst and tumor on human dental panoramic images from 76.67% to 90.91%.

References

[1] Wang Y, Guo W, Brown MB. Spline Smoothing for Bivariate Data with Applications to Association between Hormones. *Statistica Sinica*. 2000; 10(2): 377-397.

[2] Lestari B, Budiantara IN, Sunaryo S, Mashuri M. Spline Estimator in Multi-Response Nonparametric Regression Model with Unequal Correlation of Errors. *Journal of Math. & Stat.* 2010; 6 (3): 327-332.

[3] Lestari B, Budiantara IN, Sunaryo S, Mashuri M. Spline Smoothing for Multi-Response Nonparametric Regression Model in Case of Heteroscedasticity of Variance. *Journal of Math. & Stat.* 2012; 8(3): 377-384.

[4] Chamidah N, Lestari B. Spline Estimator in Homoscedastic Multi-Response Nonparametric Regression Model in Case of Unbalanced Number of Observations. *Far East Journal of Mathematical Sciences* (*FJMS*), 2016; 100(9): 1433-1453.

[5] Lestari B, Fatmawati, Budiantara, IN, Chamidah N. Estimation of Regression Function in Multi-Response Nonparametric Regression Model Using Smoothing Spline and Kernel Estimators. *Journal of Physics: Conference Series*, 2018; 1097 012091.

[6] Lestari B, Anggraeni D, Saifudin T. Estimation of Covariance Matrix Based on Spline Estimator in Homoscedastic Multi-Responses Nonparametric Regression Model in Case of Unbalanced Number of Observations. *Far East Journal of Mathematical Sciences* (*FJMS*), 2018; 108(2): 341-355.

[7] Chamidah N, Budiantara IN, Sony S, Zain I. Designing of Child Growth Chart Based on Multi-Response Local Polynomial Modeling. *J. Math. & Stat.* 2012; 8(3): 342-347.

[8] Chamidah N, Saifudin T. Estimation of Children Growth Curve Based on Kernel Smoothing in Multi-Response Nonparametric Regression. *Applied Mathematical Sciences*. 2013; 7(37): 1839-1847.

[9] Chamidah N,, Tjahjono E,, Fadilah AR,, Lestari B. Standard Growth Charts for Weight of Children in East Java Using Local Linear Estimator. *Journal of Physics: Conference Series*, 2018; 1097 012092.

[10] Chamidah N, Rifada M. Local Linear Estimation in Bi-Response Semiparametric Regression Model for Estimating Median Growth Charts of Children, *Far East Journal of Mathematical Sciences* (*FJMS*). 2016; 99(8): 1233-1244.

[11] Anonymous. Basic Health Research, Health Research and Development Agency, Ministry of Republic of Indonesia, Jakarta (in Indonesia: Riset Kesehatan Dasar, Badan Penelitian dan Pengembangan Kesehatan Kementrian RI, Jakarta). 2007.

[12] Hasibuan S. Early Detection Procedure and Diagnosis of Oral Cavity Cancer, University of North Sumatra, Medan (in Indonesia: Prosedur Deteksi Dini dan Diagnosis Kanker Rongga Mulut, Universitas Sumatera Utara, Medan). 2004.

[13] Lynch HT, Aaltonen LA, Peltom$\bar{a}$ ki P, Mecklin JP, Jarvinen H, Jass JR, Green JS, Watson P, Tallqvist G, Juhola M, Sistonen P, Hamilton SR, Kinzler K, Vogelstein B, Chapelle A. Replication Errors in Benign and Malignant Tumor from Hereditary Nonpolyposis Colorectal Cancer Patients. *American Association for Cancer*. 1994; 54(7): 1645-1648.

[14] Supriyanti R, Setiadi AS, Ramadhani Y, Widodo HB. Point Processing Method for Improving Dental Radiology Image Quality. *International Journal of Electrical and Computer Engineering (IJECE)*. 2016; 6(4): 1587-1594.

[15] Na'am J. Edge Detection on Objects of Medical Images with Enchancement multiple Morphological Gradient (EmMG). *IEEE 4th Proc. EECSI*. 2017; 61-67.

[16] Gunawan W, Arifin AZ, Indraswari R, Navastara DA. Fuzzy Region Merging Using Fuzzy Similarity Measurement on Image Segmentation. *International Journal of Electrical and Computer Engineering (IJECE)*. 2017; 7(6): 3402-3410.

[17] Na'am J, Harlan J, Madenda S, Santony J, Suharinto C. Detection of Proximal Caries at the Molar Teeth Using Edge Enhancement Algorithm. *International Journal of Electrical and Computer Engineering (IJECE)*. 2018; 8(5): 3259-3266.

[18] Na'am J, Harlan J, Nercahyo GW, Arlis S, Sahari, Mardison, Rani LN. Detection of Infiltrate on Infant Chest X-Ray, *TELKOMNIKA Telecommunication, Computing, Electronics and Control* . 2017; 15(4): 1937-1945.

[19] Angkoso CV. Classification of Tumors & Cysts in Human Dental Panoramic Images Using the Support Vector Machine (SVM) Method, Thesis, Surabaya: ITS Post Graduate. (in Indonesia: Klasifikasi Tumor & Kista pada Citra Panoramik Gigi Manusia Menggunakan Metode Support Vector Machine (SVM), *Tesis*, Surabaya: Post Graduate ITS). 2011.

[20] Angkoso CV, Nurtanio I, Purnama IKE, Purnomo MH. Texture Analysis for Cyst and Tumor Classification on Dental Panoramic Images, *Seminar on Intelligent Technology and Its Applications (SITIA)*. 2011; 15–19.

[21] Nurtanio I. Classification of Cysts and Tumors in Human Dental Panoramic Images. Ph.D. Thesis. Surabaya: ITS Post Graduate; 2013 (in Indonesia: Klasifikasi Kista dan Tumor pada Citra Panoramik Gigi Manusia. Ph.D. Thesis. Surabaya: Post Graduate ITS; 2013).

[22] Kubo, M., Koshinaka, H., Muramoto, K-I. *Extraction of Clouds in the Antarctic Using Wavelet Transforms*, Proceedings of IEEE Geoscience and Remote Sensing Symposium, 2000: 2170-2172.

[23] Zheng J, Regentova E. *Wavelet based feature reduction method for effective classification of hyperspectral data*. Proceedings ITCC 2003. International Conference on Information Technology: Coding and Computing, 2003. DOI: 10.1109/ITCC.2003.1197489
[24] Phinyomark A, Nuidod A, Phukpattaranont P, Limsakul C. Feature Extraction and Reduction of Wavelet Transform Coefficients for EMG Pattern Classification, *Electronics and Electrical Engineering*, 2012; 6(122): 27-32.
[25] Huang R, Heng F, Hu B, Peng H, Zhao Q, Shi Q, Han J. Artifacts Reduction Method in EEG Signals with Wavelet Transform and Adaptive Filter. In: Slezak D,Tan A-H, Peters JF, Schwabe L.Editors. Brain Informatics and Health. Springer Cham; 2014:122-131.
[26] Zulfikar F. Generalized Exploratory Factor Analysis (GEFA) Approach to Reduce Data Dimensions in Calibration Modeling, Thesis, Bogor Agricultural Institute, Bogor (in Indonesia: Pendekatan Generalized Exploratory Factor Analysis (GEFA) Untuk Mereduksi Dimensi Data dalam Pemodelan Kalibrasi, Tesis, Institut Pertanian Bogor, Bogor). 2013.
[27] Sunaryo S, Retnaningsih SM. Application of Calibration Model with Discrete-Partial Least Square (TWD-PLS) Wavelet Transformation on Gingerol Data (in Indonesia: Penerapan Model Kalibrasi dengan Transformasi Wavelet Diskrit–Partial Least Square (TWD-PLS) pada Data Gingerol). *Jurnal ILMU DASAR*. 2008; 9(1): 56-61
[28] Eubank, R. L. *Nonparametric Regression and Spline Smoothing Second Edition*, Marcel Dekker, New York.1999.
[29] Welsh AH, Yee TW. Local regression for vector responses, *Journal of Statistical Planning and Inference*. 2006; 136(9): 3007-3031.
[30] Peng CJ, Lee KL, Ingersoll GM. An Introduction to Logistic Regression Analysis and Reporting, *the Journal of Educational Research*. 2002; 96(1): 3-14
[31] Hastie TJ, Tibshirani RJ. *Generalized Additive Models*, Chapman & Hall, London. 1990.
[32] Meyers LS, Glenn G, Guarino AJ. *Applied Multivariate Research*, 3 edition. SAGE Publication. United States of America. 2016.