DENTAL DISEASE DETECTION BY USING GABOR FILTER AND K-NEAREST NEIGHBORS BASED ON IMAGE TEXTURES

Setiawan Sudrajat[1], Kania Evita Dewi[2].

[1,2] Teknik Informatika – Universitas Komputer Indonesia
Jl.Dipatiukur 112-116 Bandung
E-mail : setiawanuyab@gmail.com[1], kania.evita.dewi@email.unikom.ac.id[2]

ABSTRACT

The disease that experienced mostly by the Indonesian people is a dental disease. Periodontal disease composed by gingivitis and periodontitis. Two second disease have difference where if gingivitis caused by bacterial plaque stick teeth but teeth still strong, then periodontitis is advanced of gingivitis disease because bacterial infection of the teeth that can ruined teeth. The difference of second disease cause periodontal disease difficult to detection by people even can become fatal problem if not handled well.

That cans detection disease of teeth, either using image processing application. Image processing application using gabor filter method to texture extraction of image and algorithm K-NN to classification class. By using image processing application can detection dental disease since early.

The test result detection dental disease with gabor filter and KNN get accuracy with each K=3 amount 86%, K=5 amount 82%, and K=7 amount 80%. With the test result, image processing application using gabor filter and K-NN can detection second disease that well so this application can used by people to detection dental disease since early.

Keyword: Gabor Filter Method, K-NN Algorithm, Dental Disease, Image Processing Application, Periodontal.

1. INTRODUCTION

1.1 Background

The disease that experienced mostly by the Indonesian people is a dental disease [1]. Accord result *Survey Kesehatan Rumah Tangga* (SKIRT) in 2004, dental disease is mostly by experienced half Indonesian is a periodontal disease [1]. Periodontitis by composed two diseases that gingivitis and periodontitis. Two diseases have same that disease caused by bacterial plaque of teeth. Difference two diseases that if gingivitis caused by bacterial plaque stick of teeth but still not broken so no until damage other network. While, periodontitis disease cause infection teeth effect bacterial plaque can damage network bone teeth and gum so cause damage teeth [2]. The difference two dental disease, especially periodontal cause people hard detection disease even if handled not well then fatal problem. For solution diseases that needed by detection dental disease that using development technology current that is image processing. Image processing application that expected can detection dental disease.

For detection image that needed by extraction texture. Extraction texture is taking feature value of object can give information to image. One of extraction feature to implementation by method gabor filter. Gabor filter have excellence that can give image information into the orientation and frequency so can get value on detail. Accord research using by Helsi Rosyida, gabor filter method can detection fundus eye well in process segmentation of retina vessels. [3]

Classification process is grouping object based same class. One of classification method to using by KNN algorithm. KNN Algorithm is algorithm technique of grouping class to object with find the closest Euclidean distance. Accord research using by Elva Budiantara [4], KNN method can classification image with accuracy each among 88.75% to testing without background and 73.375% to testing with using background. Based result research and background then create image processing application with gabor filter and KNN method.

1.2 Identification of Problem

Identification of problem exposed to background that has not measured accuracy in detecting dental disease using by filter gabor and KNN Algorithm..

1.3 Intention and Purpose

Intention from research that detection dental disease with implementation filter gabor method and KNN.

Purpose from research that measured accuracy filter gabor and KNN algorithm in detection dental disease.

1.4 Research Methods

Research methods in this case that

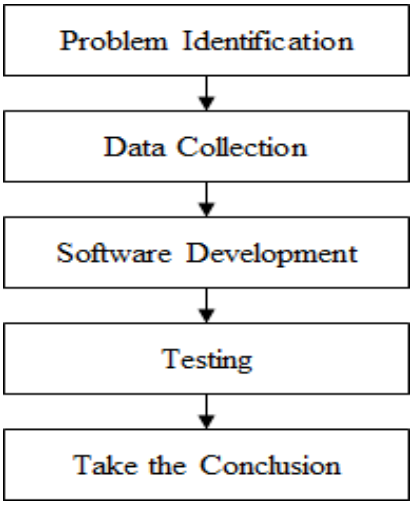


Image 1. Research Method

Stage research methods that

1. Make identification of problem about research to be made.
2. Collection data to research. For research data of in this case composed by study of literature and study of working. Study of literature to get journal, article, and book as reference in this case while study of working to get other data for research.
3. Make software development for build application in this research. The development methods used is agile method. Using testing to get accuracy. Stage method by composed two stage First test training data to set feasible or not to use. Second test testing data to get accuracy. As for functional testing used is blackbox and equivalent-class to test input data.
4. Make conclusion result research.

2. CONTENT

2.1 Application Overview.

This Research has proper for build image processing application so that can detection dental disease with using gabor filter and KNN. As for dental disease detected is periodontitis and gingivitis. Process application overview can be seen on image 2 as follows.

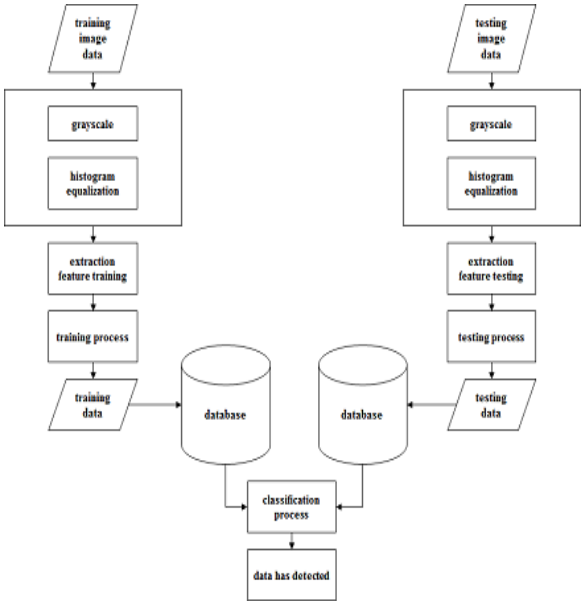


Image 2. Application overview of dental disease detection process

2.2 Input Data

Input data to taken in this research is image of periodontitis and gingivitis. Insert image is size 125x125 pixel and formatted png. After insert image next step is using preprocessing.

2.3 Preprocessing

Preprocessing is early stage for start process data before do extraction feature and classification. Process implemented in thesis is grayscale and histogram equalization.

Grayscale is process for change color image become color grayscale. Following calculation grayscale value to every pixel image [5].

$$Grayscale = (R \times 0.299) + (B \times 0.587s) + (G \times 0.114) \quad (1)$$

Where

Grayscale = Grayscale value
R = Red Value
B = Blue Value
G = Green Value

As for plot grayscale can be seen on image 3.

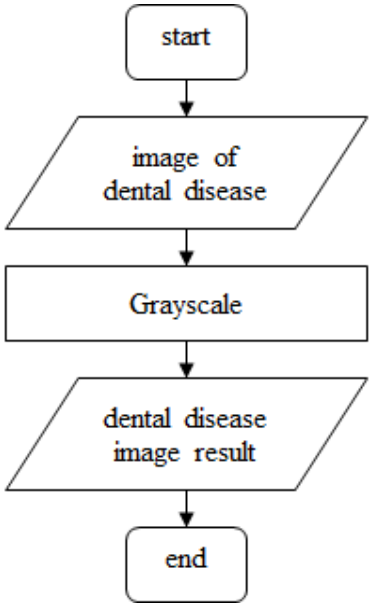


Image 3. Flowchart Grayscale Process

After is calculation grayscale process. Next step is *histogram equalization* process. Histogram equalization is process to smooth grayscale value evenly forming new histogram. Histogram equalization intend for sharpening gray equality of image. Histogram Equalization started with calculation gray intensity value (rk) where gray value (k) divided by gray degree (L). Following is equation gray intensity (rk).

$$r_k = \frac{k}{L-1} \quad (2)$$

Where

rk = gray intensity value
k = gray value
L = gray degree value

Calculation every rk from beginning to finish with gray value (k) from 0 to 255 and gray degree around 255 [6].

In the histogram equalization, 0 is black value and 1 to white value in gray scale. After determined gray intensity value (rk), then calculation probability appearance new pixel for new histogram. Probability appearance (pr) pixel for gray intense value defined by Pr(rk). Then probability (pr) of every pixel is calculated, where amount gray pixel value that appears at a certain intensity divided by total pixel value. Following is equation probability intensity value (rk). [6].

$$P_r(r_k) = \frac{n_k}{n} \quad (3)$$

Where

Pr(rk) = Probability appearance new pixel
nk = image pixels value that appears at a certain intensity.
n = total pixel value.

For begin histogram equalization, image must already be processed in format grasycale. Next step, histogram equalization used for smooth intensity value in the image to equal with approach cumulative frequency distribution.

After calculation probability appears new pixels, next step distribute cumulative probability value to new pixel color gray (sk). Cumulative frequency distribution aim to calculate probability value with equal from initial pixels to final pixels so that get new gray value (sk) with smooth. Following is equation every new gray value (sk) [6].

$$s_k = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n} = \sum_{j=0}^{k} P_r(r_j) \quad (4)$$

Where

sk = cumulative distribution of gray value.
nj = image pixels value that appears at a certain intensity.
n = total pixel value.
T = rk transformation value

Plot histogram equalization can be seen on Image 4.

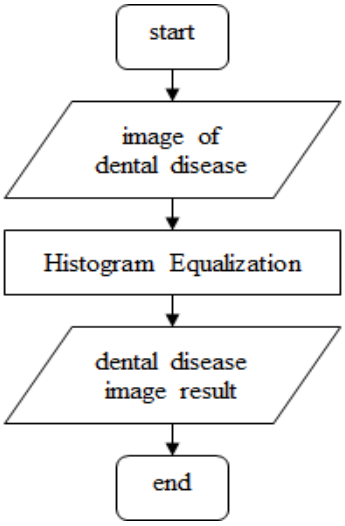


Image 4. Flowchart Histogram Equalization Process

After do this step, next step using extraction feature.

2.4 Extraction feature Gabor Filter

Gabor filter is feature extraction method with calculate equation sinusoid function and Gaussian envelope function. Step use extraction feature can be seen on image 5.

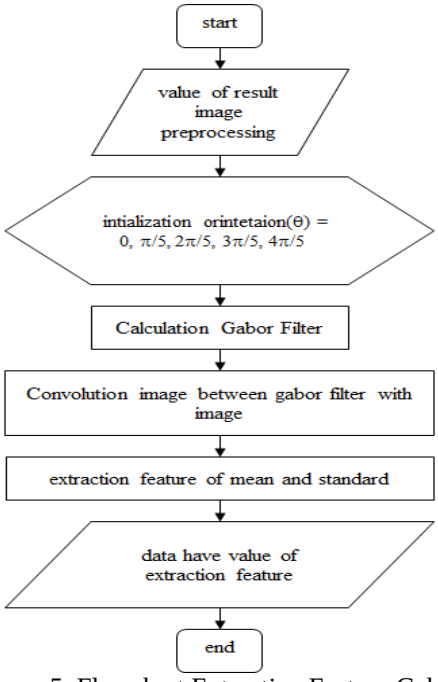


Image 5. Flowchart Extraction Feature Gabor Filter

First step, insert result preprocessing value from grayscale and histogram equalization. The second, initialization orientation degree value each amount 5 values, that is value 0, π/5, 2π/5, 3π/5, and 4π/5 [8].

After initialization orientation degree value, next step calculate gabor filter value. Following is calculation *filter gabor* value [9].

$$g_\sigma(x,y) = \frac{1}{2\pi\,\sigma_x\,\sigma_y} * \exp -\frac{1}{2}\left(\frac{x'}{\sigma_x} + \frac{y'}{\sigma_y}\right) * \cos(2\pi\omega_{x0}x + 2\pi\omega_{y0}y) \quad (5)$$

Where

ω_{x0} = angular frequency dimension x value

ω_{y0} = angular frequency dimension x value

θn = orientation degree to parallel line gabor function

σx = standard deviation Gausian Envelope dimension x

σy = standard deviation Gausian Envelope dimension y

x = coordinate point x

y = coordinate point y

With value

$x_{\theta n}$ = x cos (θn) + y sin (θn) (6)

$y_{\theta n}$ = - x cos (θn) + y sin (θn) (7)

After calculate gabor filter, next step convolution process. Convolution is one of technique change original image with kernel. In the case of this research, result gabor filter to be convolution by original image produce image result convolution. Convolution equation operation is [6].

$$h(x,y) = f(x,y) * g(x,y) \quad (8)$$

Where

h = convolution value

f = original image value

g = kernel gabor value

After using convolution, next step calculate extraction feature value composed to two value, that is mean and standard deviation value. Following is equation mean and standard deviation [10].

1. Mean value

$$\bar{x} = \frac{\sum x_i}{n} \quad (9)$$

Where

$\bar{x}$ = mean value

xi = amount data

n = many data value

2. Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n}(xi - \bar{x})^2}{n-1}} \quad (10)$$

Where

s = standard deviation

xi = amount data

n = many data value

Calculation extraction feature is calculated as much 5 times with value 0, π/5, 2π/5, 3π/5, and 4π/5 so that produce extraction mean value as many 5 values and extraction standard deviation as many 5 values..

2.5 Training Process

After extraction feature process, next steps training process. Training process intend to get training data to be made sample for testing data. Training data that include data already process at stage *preprocessing* and get extraction feature value. Training data also inserted to database. Plot training process in save data training can be seen on image 6.

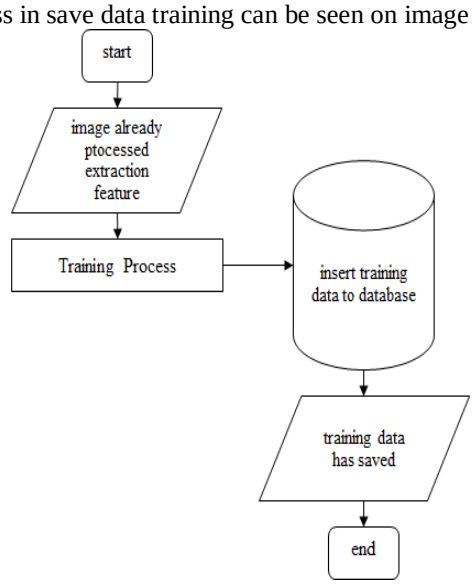


Image 6. Training Process

Training data to be compared by testing data whether data is accurate or not. Training data used in this research composed of 20 training data with amount 10 data periodontitis and amount 10 data gingivitis.

2.6 Testing Process

Testing process in the case of this research is used to get testing data. Testing process in the case of this research is used to get testing data. Testing data that include data already process at stage *preprocessing* and get extraction feature value. Testing data also inserted to database. Plot testing process in save data testing can be seen on image 7.

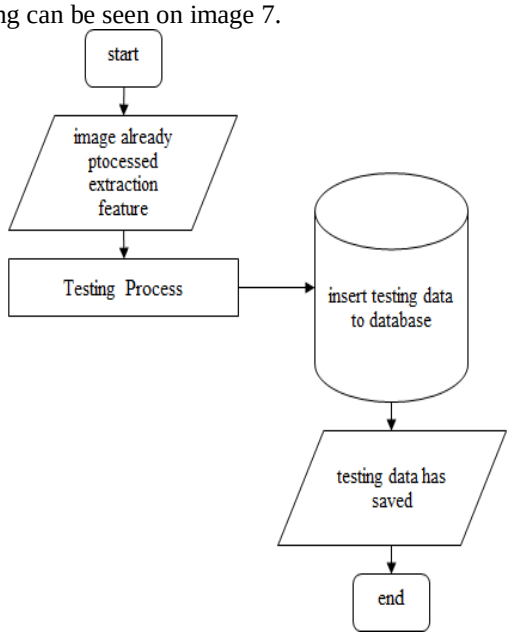


Image 7. Testing Process

Testing data already insert into database to compare with training data so that get accuracy. Testing data is used many 100 data for in the case of this research.

2.7 KNN Classification

After process extraction feature, Next step use classification KNN with calculate Euclidean distance, where with the mean and standard deviation training data to compared by mean and standard deviation value. Plot process classification KNN can be seen on image 8 [11].

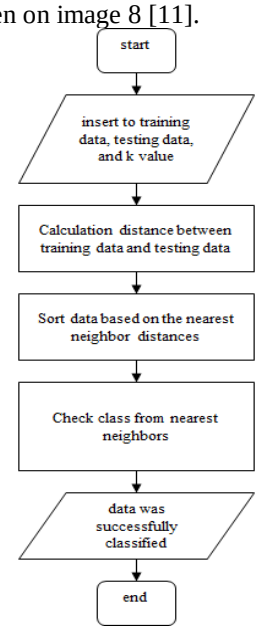


Image 8. KNN Classification Process

2.8 Testing

Testing used to measure accuracy. Testing composed to 2 stages, which are test training data with same data and test training data with testing data. K value used in the test is value K=3, K=5, and K=7. Training data used as much 20 data with 10 data periodontitis and 10 data gingivitis while, testing data used as much 100 data. Result test training data with same data can be seen on table 1 and result test training data with testing data can be seen on table 2.

Table 1. Test training data with same data

description	K=3	K=5	K=7
Training data of accurate	100%	100%	100%

Of the test result, concluded that training data with value K=3, K=5, and K=7 have accuracy each 100%. Therefore, training data worth test to detection testing data. After test training data process, next step test testing data process. Testing data used as much 100 data with difference data for get accuracy. Result test testing data can be seen on table 2.

Table 2. Testing test data with training data

Description		K=3	K=5	K=7
Periodontitis	accuracy	41	36	32
	not	9	14	18
Gingivitis	accuracy	45	44	46
	not	5	6	4
Total accuracy		86	80	78
Total		100	100	100
$\frac{\sum jumlah\ akurasi}{\sum total} * 100\%$		86%	80%	78%

Of the test result, conclude form 100 testing data detected as successful get accuracy with each 86 data for value K=3, 80 data for value K=5, and 78 data for value K=7. From data accuracy between periodontitis disease detection many 41 data for K=3, 36 for K=5, and 32 for K=7. As for gingivitis detection 45 data for K=3, 44 for K=5, and 46 for K=7. Therefore, accuracy percentage on test testing data to K=3, K=5, and K=7 each amount 86%, 80%, and 78%.

3. CLOSING

3.1 Conclusion

The Conclusion from the test show periodontal disease detection suitable amount 86% for value K=3, value K=5 amount 80%, and value K=7 amount 78%. From value K, value K=3 has the highest accuracy in identification dental disease because can detection data well. Influencing factor in detecting dental disease is difference extraction value standard deviation on periodontitis and gingivitis. Along majority value K appears on testing accuracy can affect accuracy in detection dental

disease. Result research also show increase of value K, then accuracy decreased [12].

3.2 Recommendation

Recommendation to given by writer is append other extraction feature because on this research. Extraction feature mean value has same but standard deviation value different so that happen inconsistency between one images with another image.

4. BIBLIOGRAPHY

[1] Departemen Kesehatan RI. *Kebijakan Pelayanan Kedokteran Gigi Keluarga*. Jakarta: Depkes, 2005.

[2] Dokter Sehat. *Periodontitis Peradangan Pada Jaringan Pendukung Gigi*. 2016. Diakses 19 Februari 2016 <http://doktersehat.com/periodontitis-peradangan-pada-jaringan-pendukung-gigi/>.

[3] Mandasari, Helsi Rosyida, Handayani Tjandrasa dan Arya Yudhi Wijaya. *Segmentasi Pembuluh Darah Retina pada Citra Fundus Mata dengan 2D-Gabor Filter*. Skripsi. Surabaya: Institut Teknologi Sepuluh Nopember (ITS), 2011.

[4] Budianita, Elvia, Jasril dan Lestari Handayani. "Implementasi Pengolahan Citra dan Klasifikasi K-Nearest Neighbor Untuk Membangun Aplikasi Pembeda Daging Sapi dan Babi." *Jurnal Sains, Teknologi dan Industri, Vol. 12 No.2* (2015): 242-247.

[5] Gurusinga, Pertumpun dan Kelvin Permana Arbi. "Rancangan Aplikasi Pengenalan Wajah Dengan Algoritma Eigenface." *Jurnal Matematika Statistik, Vol. 13 No. 2* (2013): 132-139.

[6] Munir, Rinaldi. *Pengolahan Citra Digital*. Bandung: Penerbit Informatika Bandung, 2002.

[7] Putra, Darma. *Pengolahan Citra Digital*. Yogyakarta: Andi, 2010.

[8] Khatun, Amina dan Md. Al-Amin Bhuiyan. "Neural Network based Face Recognition with Gabor Filter." *IJCSNS Internatioal Journal of Computer Science and Network Security, Vol. 11 No.1* (2011): 71-76.

[9] Derpanis, Konstantinos G. "Gabor Filters." *York University kosta@cs.yorku.ca Version1.3* 23 April 2007: 1-4.

[10] Rumus Statistik. *Rata-Rata, Varian, Dan Standar Deviasi*. 2013. 21 Februari 2018 <https://www.rumusstatistik.com/2013/07/varian-dan-standar-deviasi-simpangan.html>.

[11] Sukma, Alfian, et al. *K-Nearest Neighbor Information Retrieval (Sistem Temu Kembali Informasi)*. Skripsi. Surabaya: Universitas Airlangga, 2014.

[12] Karo, Edis Siswanto Karo dan Pulung Nurtantio Andono. *Klasifikasi Gambar Gerak Pemain Bola Voli Dengan Menggunakan Metode K-Nearest Neighbor (K-NN)*. Skripsi. Semarang: Universitas Dian Nuswantoro, 2011.