# Classifying Cyst and Tumor Lesion on Dental Panoramic Images Using Support Vector Machine Based on GLRLM Texture Features.

Ingrid Nurtanio[1], Eha Renwi Astuti[2] and Christoforus Yohannes[1]
[1]Department of Electrical Engineering, University of Hasanuddin, Makassar, Indonesia
(Tel : +62-815-252-2716; E-mail: ingrid_unhas@yahoo.com)
[2]Department of Dentistry, Airlangga University, Surabaya, Indonesia
(Tel : +62-816-536-030; E-mail: e_renwi_a@yahoo.com)

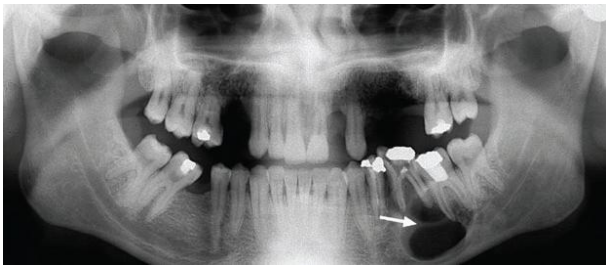
**Abstract- Many lesions that occur in the jaw have a similar radiographical appearance and it is often difficult to differentiate among them. In this research, an approach is proposed to develop a computer-aided classification system for cyst and tumor lesion from dental panoramic images. The proposed system consist of feature extraction based texture using GLRLM (Gray Level Run Length Matrix). We used a set of 7 GLRLM features. The 7 selected features are: SRE (Short Runs Emphasis), LRE (Long Runs Emphasis), GLN (Gray Level Non-uniformity), RP (Run Percentage), RLN (Run Length Non-uniformity), LGRE (Low Gray Level Run Emphasis), HGRE (High Gray Level Run Emphasis) and the 7 features extracted from grey level run-length matrix could distinguishing cyst and tumor lesion with accuracy up to 76,92% using Support Vector Machine Method. Further analysis showed that Area Under the Receiver Operating Characteristic Curve (AUC) was 0.8722. Based on the AUC value, the level of accuracy of this prediction could be categorized as 'Good'.**

**Keyword: Area Under the Receiver Operating Characteristic Curve, cyst and tumor lesion, feature extraction, Gray-level Run Length Matrix, Support Vector Machine.**

## I. INTRODUCTION

A variety of disorders can be found in human jawbone. These disorders consist of various types of cyst and tumor lesions that have been clinically classified [1]–[6]. Many lesions that occur in the jaw have a similar radiographical appearance and it is often difficult to differentiate among them [6]. Both types of lesions, typically have a smooth, round or oval periphery [4], [6] and are not easily differentiated, see Fig. 1. This situation makes a dentist unable to determine exactly whether it is a tumor or a cyst. Thus, we proposed to develop a computer aided classification system for cyst and tumor lesion from dental panoramic images. The proposed system consist of feature extraction based texture using GLRLM (Gray Level Run Length Matrix) and SVM (Support Vector Machine) as classification method. To evaluate the performance of SVM, we compute Area Under Receiver Operating Characteristic (ROC) Curve (AUC).

This paper is organized as follows : In section 2 we give the materials and methods concerning the GLRLM, SVM, ROC and AUC. In section 3, we show the experimental result about the feature extraction and classification. After that, compute AUC from ROC curve to measure SVMs classifier performance. Section 4, discusses the result and section 5 contains the conclusions.



(a)

(b)

Fig. 1. The cyst and tumor lesion on dental panoramic images. (a) cyst lesion (arrow), (b) tumor lesion (arrow).

## II. MATERIALS AND METHODS

### A. Materials

We have considered a dataset of 133 dental panoramic images including various of cyst lesion (radicular cyst, dentigerous cyst, buccal bifurcation cyst, keratocyst, calcifying odontogenic cyst, nasopalatine cyst, simple bone cyst) and various of tumor lesion (ameloblastoma, ameloblastic, fibroma, adenomatoid, odontogenic tumor, odontoma, cementoblastoma, torus palatinus, torus mandibularis, exostosis, enostosis, myxoma, osteoma, hemangioma, osteoid osteoma, osteo blastoma).from Oral Radiology [4]. All images were already in digital forms. The total images of the cyst lesions were 53

images and those for the tumor lesions were 80 images. The position of the cyst and tumor regions was provided by an experienced radiologist acting as the co-author of this paper.

B. Methods

In this paper, we develop method to classify cyst or tumor lesion using the properties of dental panoramic images. The steps of our method are as follow :

1. Preprocessing : a color image is first transformed into gray image by normalizing the values of its pixels with respect to the length of the gray scale. Using Gaussian filter to smooth the images. The region of interest (ROI) for each lesion was manually outlined by a well-trained operator and further confirmed by an experienced radiologist. The ROI of size 40 x 40 pixels is extracted with mass centered in the window, and divided into two sets : the learning set and the testing set.
2. Feature Extraction : using GLRLM to extract the features of cyst and tumor lesions. We used a set of 7 features. The 7 selected features are: SRE (Short Runs Emphasis), LRE (Long Runs Emphasis), GLN (Gray Level Non-uniformity), RP (Run Percentage), RLN (Run Length Non-uniformity), LGRE (Low Gray Level Run Emphasis) and HGRE (High Gray Level Run Emphasis).
3. Classification : using SVM method to classify the result of features from cyst and tumor lesion.
4. Evaluation : using AUC to evaluate the performance of SVM.

C. GLRLM (Gray Level Run Length Matrix)

Gray level run length matrix is a matrix from which the texture features can be extracted for texture analysis. Texture is understood as a pattern of gray intensity pixel in a particular direction from the reference pixels. Run length is the number of adjacent pixels that have the same gray intensity in a particular direction. Gray level run length matrix is a two dimensional matrix where each element $p(i,j/\theta)$ is the number of elements j with the intensity i, in the direction θ. For example, figure 2 below shows a matrix of size 4x4 pixel image with 4 gray levels. Figure 3 is a representation matrix GLRL (Gray Level Run Length) in the direction of $0^0$ [$p(i,j/\theta=0^o)$]

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 1 | 3 | 4 | 4 |
| 3 | 2 | 2 | 2 |
| 4 | 1 | 4 | 1 |

| Gray level | Run length (j) | | | |
|---|---|---|---|---|
| i | 1 | 2 | 3 | 4 |
| 1 | 4 | 0 | 0 | 0 |
| 2 | 1 | 0 | 1 | 0 |
| 3 | 3 | 0 | 0 | 0 |
| 4 | 3 | 1 | 0 | 0 |

Fig. 2. Matrix of image 4X4 pixels Fig. 3. GLRL Matrix

In addition to the $0^o$ direction, GLRL matrix can also be formed in the other direction, i.e. $45^o$, $90^o$ or $135^o$. Some texture features can be extracted from the GLRL matrix, such as Short Runs Emphasis (SRE), Long Runs Emphasis (LRE), Gray Level Non-uniformity (GLN), Run Percentage (RP), Run Length Non-uniformity (RLN), Low Gray Level Run Emphasis (LGRE), High Gray Level Run Emphasis (HGRE).

The feature can be defined as follows in eq. (1-7) :

$$SRE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)}{j^2} \quad (1)$$

$$LRE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)*j^2 \quad (2)$$

$$GLN = \frac{1}{n_r}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}p(i,j)\right)^2 \quad (3)$$

$$RP = \frac{n_r}{p(i,j)*j} \quad (4)$$

$$RLN = \frac{1}{n_r}\sum_{j=1}^{N}\left(\sum_{i=1}^{M}p(i,j)\right)^2 \quad (5)$$

$$LGRE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)}{i^2} \quad (6)$$

$$HGRE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)*i^2 \quad (7)$$

D. Classification

After the features have been extracted and selected, they are input into classifier to categorize the images into cyst or tumor lesion. We use Support Vector Machine (SVM) method to categorize these lesions.

D.1. Support Vector Machine (SVM) Classifier

SVM is a state-of-the-art classification method introduced in 1992 by Boser, Guyon & Vapnik [7] for binary classification. The key concept of SVMs, which were originally first developed for binary classification problems, is the use of hyperplanes to define decision boundaries separating between data points of different classes. SVMs are able to handle both simple, linear, classification tasks, as well as more complex, i.e. nonlinear, classification problems. Both separable and nonseparable problems are handled by SVMs in the linear and nonlinear case. The idea behind SVMs is to map the original data points from the input space to a high dimensional, or even infinite-dimensional, feature space such that the classification problem becomes simpler in the feature space. The mapping is done by a suitable choice of a kernel function. Kernel functions are used to map the input data into a higher dimension space where the data are supposed to have a better distribution, and then an optimal separating hyperplane in the high-dimensional feature space is chosen [8].

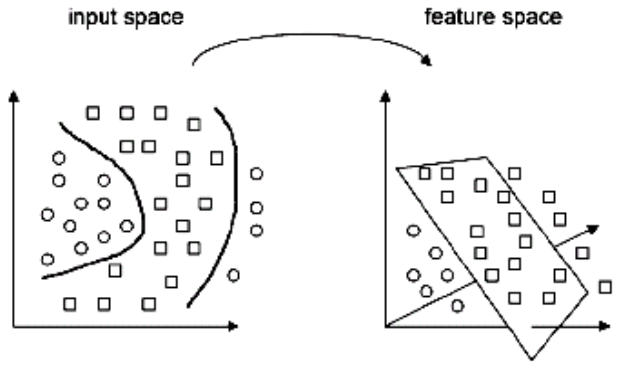



Fig. 4. SVMs allow mapping of the data from the input space to a high-dimensional feature space [9].

Consider a training data set $\{x_i,y_i\}$, with $x_i \in \Re^d$ being the input vectors and $y_i \in \{-1,+1\}$ the class labels. SVMs map the d-dimensional input vector x from the input space to the d1-dimensional feature space using a (non)linear function $\phi(.) : \Re^d \rightarrow \Re^{d1}$ . The separating hyperplane in the feature space is then defined as $W^T \phi(x) + b = 0$, with $b \in \Re$ and $W$ an unknown vector with the same dimension as $\phi(x)$. A data point $x$ is assigned to the first class if $f(x) = \text{sign}(W^T \phi(x) + b)$ equals +1 or to the second class if $f(x)$ equals -1.

However, in our study, data of both classes are overlapping, which makes a perfect linear separation impossible. Therefore, a restricted number of misclassification should be tolerated around the margin. The resulting optimization problem for SVMs, where the violation of the constraints is penalized, is written as

$$\min_{W,b,\xi} \frac{1}{2} W^T W + C \sum_{i=1}^{N} \xi_i$$

$$\text{subject to} \quad y_i (W^T \phi(x_i) + b) \geq 1 - \xi_i \qquad (8)$$

$$\xi_i \geq 0, \qquad i = 1,\ldots\ldots\ldots, N,$$

where C is a positive regularization constant. The regularization constant in the cost function defines the trade-off between a large margin and misclassifation error.

For non-separable data, the soft-margin SVM uses the slack variable ($\xi$) to control an upper bound of the misclassification error. The value of $\xi_i$ indicates the distance of $x_i$ with respect to the decision boundary. Equivalently, the optimization problem for SVMs can be written in the dual space us ing the Lagrangian with Lagrange multipliers $\alpha_i \geq 0$ for the first set of constraints. The solution for the Lagrange multipliers is obtained by solving a quadratic programming problem. Finally, the SVM classifier takes the form :

$$f(x) = sign\left( \sum_{i=1}^{\#SV} \alpha_i y_i K(x, x_i) + b \right) \qquad (9)$$

where #SV represents the number of support vectors and the kernel function K(.,.) is positive definite.

Furthermore, $K(x,x_i) \equiv \phi(x)^T \phi(x)$ is called the kernel function. In the optimization problem only K(.,.) is used which is related to $\phi(.)$. This enables SVMs to work in a high-dimensional (or infnite-dimensional) feature space, without actually performing calculations in this space. We use Gaussian Kernel in this study with kernel function is :

$$K(x, x_i) = \exp\left( -\frac{\|x - x_i\|^2}{2\sigma^2} \right) \qquad (10)$$

where $\sigma$ is the kernel parameter.

E. Evaluation

A receiver operating characteristic (ROC) curve is most frequently used because of its comprehensive and fair evaluation ability [8] . A ROC curve is a plotting of true positive fraction (TPF) as a function of false positive fraction (FPF) [10], [11]. The area under the ROC curve (AUC) can be used as a criterion. Table I shows the classifying level of accuracy based on AUC [12].

Other frequently used criteria are [12]–[14] :

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (11)$$

$$specificity = \frac{TN}{TN + FP} \qquad (12)$$

$$sensitivity = \frac{TP}{TP + FN} \qquad (13)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives.

TABLE I

CLASSIFYING LEVEL OF ACCURACY BASED ON AUC [12]

| AUC value | Classified as |
|---|---|
| 0.90 – 1.00 | Excellent |
| 0.80 – 0.90 | Good |
| 0.70 - 0.80 | Fair |
| 0.60 – 0.70 | Poor |
| 0.50 – 0.60 | Fail |

## III. EXPERIMENTS AND RESULTS

All the experiments were conducted in Matlab Ver 7.1 by using a PC Intel-Pentium Centrino with RAM 1 GB.

A total of 53 cyst lesions and 80 tumor lesions measuring 40x40 pixels are transformed into GLRLM texture. Based on GLRLM, 7 texture features was extracted.

Table II shows the features' values extracted from GLRLM both lesions. It can be seen that the value of the features both classes are overlapping, which makes a perfect linear separation impossible but the minimum and maximum value for both classes are different. This shows that classification process cannot be easily done (not linear) because the overlapping value. But the difference in each class maximum and minimum

features' value makes classification still possible to conduct (non linear classification).

TABLE II
TABULATED MINIMUM AND MAXIMUM FEATURES VALUE EXTRACTED FROM GLRLM AT DIRECTION= $0^0$

| Features | Cyst Lesion | | Tumor Lesion | |
|---|---|---|---|---|
| | Min. | Max. | Min. | Max. |
| SRE | 0.2193 | 0.8226 | 0.3392 | 0.7997 |
| LRE | 1.9790 | 109.3407 | 2.8501 | 182.9587 |
| GLN | 39.3656 | 273.6864 | 37.0880 | 278.0305 |
| RP | 0.1431 | 0.7725 | 0.1513 | 0.6838 |
| RLN | 22.8428 | 745.6715 | 54.7440 | 641.1950 |
| LGRE | 0.0261 | 0.3054 | 0.0264 | 0.1718 |
| HGRE | 9.1975 | 46.5710 | 12.5282 | 63.9416 |

For example, the SRE values of cyst lesion are from 0.2193 until 0.8226 while the SRE values of tumor lesion are from 0.3392 until 0.7997. This problem is handled by SVMs in the nonlinear case. SVMs maps the original data points from the input space to a high dimensional feature space (see Fig. 4). We use one third hold out cross validation of 133 data images by randomly selecting 94 images referring to each class as data training, while the rest (39 images) was used for the data test. The experiments were conducted 20 times. Using SVM with kernel Gaussian and parameter value C = 10000, $\alpha$ = 1e-7 and $\gamma$ = 4000, we get accuracy up to 76.92 %. Using ROC curve (see Fig. 5) and computing the AUC, we get the AUC up to 0.8722. This means that the cyst lesions and tumor lesions can be distinguished. We have achieved the GOOD classifying level of accuracy. (see Table I)

## IV. CONCLUSION

Based on experimental results, it concludes that texture features based on GLRLM can be used to distinguish between cyst and tumor lesion on dental panoramic image, with accuracy levels up to 76.92 %, and AUC value of 0.8722.

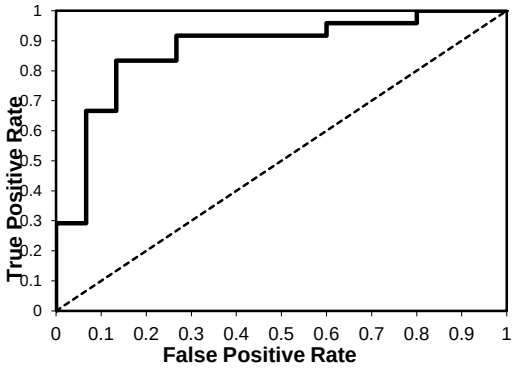



Fig. 5. ROC Curve with AUC value of 0.8722

## REFERENCES


[1] S. J. Theodorou, D. J. Theodorou, and D. J. Sartoris, "Imaging characteristics of neoplasms and other lesions of the jawbones: Part 1. Odontogenic tumors and tumorlike lesions," *Clin. Imaging*, vol. 31, no. 2, pp. 114–119, Mar. 2007.

[2] S. J. Theodorou, D. J. Theodorou, and D. J. Sartoris, "Imaging characteristics of neoplasms and other lesions of the jawbones: Part 2. Odontogenic tumor-mimickers and tumor-like lesions," *Clin. Imaging*, vol. 31, no. 2, pp. 120–126, Mar. 2007.

[3] D. J. Theodorou, S. J. Theodorou, and D. J. Sartoris, "Primary non-odontogenic tumors of the jawbones: An overview of essential radiographic findings," *Clin. Imaging*, vol. 27, no. 1, pp. 59–70, Jan. 2003.

[4] S. C. White and M. J. Pharoah, *Oral Radiology: Principles and Interpretation*. Elsevier Health Sciences, 2008.

[5] F. Pasler and H. Visser, *Pocket Atlas of Dental Radiology*. Thieme, 2007.

[6] Z. Neyaz, A. Gadodia, S. Gamanagatti, and S. Mukhopadhyay, "Radiographical approach to jaw lesions," *Singapore Med. J.*, vol. 49, no. 2, pp. 165–176; quiz 177, Feb. 2008.

[7] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, 1992, pp. 144–152.

[8] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognit.*, vol. 43, no. 1, pp. 299–317, Jan. 2010.

[9] J. Luts, F. Ojeda, R. Van de Plas, B. De Moor, S. Van Huffel, and J. A. K. Suykens, "A tutorial on support vector machine-based methods for classification problems in chemometrics," *Anal. Chim. Acta*, vol. 665, no. 2, pp. 129–145, Apr. 2010.

[10] T. A. Lasko, J. G. Bhagwat, K. H. Zou, and L. Ohno-Machado, "The use of receiver operating characteristic curves in biomedical informatics," *J. Biomed. Inform.*, vol. 38, no. 5, pp. 404–415, Oct. 2005.

[11] T. Fawcett, "An introduction to ROC analysis," *ROC Anal. Pattern Recognit.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

[12] A. K. Mohanty, S. Beberta, and S. K. Lenka, "Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammogram," *Int. J. Eng. Res. Appl. IJERA*, vol. 1, no. 3, pp. 687–693, 2011.

[13] H. Wibawanto, A. Susanto, T. S. Widodo, and S. M. Tjokronegoro, "Discriminating Cystic and Non Cystic Mass using GLCM and GLRLM based Texture Features," *Int. J. Electron. Eng. Res.*, vol. 2, no. 4, pp. 569–580, 2010.

[14] M. Çınar, M. Engin, E. Z. Engin, and Y. Ziya Ateşçi, "Early prostate cancer diagnosis by using artificial neural networks and support vector machines," *Expert Syst. Appl.*, vol. 36, no. 3, Part 2, pp. 6357–6361, Apr. 2009.