

Gingivitis Identification via Grey-level Cooccurrence Matrix and Extreme Learning Machine

Wen Li^{1,a,#}, Yiyang Chen^{2,b,#}, Leiying Miao^{1,c,*},
Mackenzie Brown^{3,d}, Weibin Sun^{4,e} and Xuan Zhang^{4,f}

¹ Department of Endodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China

² School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing, Jiangsu 210023, China

³ School of Engineering, Edith Cowan University, Joondalup WA 6027, Australia

⁴ Department of Periodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China

^a18844501367@163.com, ^bcyy15950539675@126.com, ^cmiaoleiying80@163.com

^dmackbrown@ieee.org, ^ewbsun@nju.edu.cn, ^fzhxuan2015@163.com

Those two authors contributed equally to this paper

* The corresponding author

Keywords: Gingivitis; Graylevel Cooccurrence Matrix; Extreme Learning Machine

Abstract. The diagnosis of gingivitis often occurs years later by using a series of conventional oral examination, and they depended a lot on dental records which are physically and mentally laborious task for dentists. In this study, our research presented a new method to diagnose gingivitis, which is based on gray-level cooccurrence matrix (GLCM) and extreme learning machine (ELM). The experiments demonstrate that this method is more accurate and sensitive than two state-of-the-art approaches: naïve Bayes classifier and wavelet energy.

Introduction

The tooth is one of the most important structures in the human mouth. There are a lot of diseases with the tooth and chronic gingivitis is a gum tissue inflammatory disease caused by long-term bacterial infection [1]. The main manifestations are redness, bad breath and bleeding which leads to major damage to the periodontium [2]. Plaque is the main irritant of periodontal disease and it can be reformed in the tooth surface after a few hours. Therefore, it is necessary to brush effectively every day to remove plaque and prevent periodontal disease.

The diagnosis of gingivitis often occurs years later by using a series of conventional oral examination. Recent studies have been shown that dental records play an important role in the accurate detection of these gingivitis. However, most dentists are inexperienced at recording the “dental chart” [3] for gum loci, and it is a physically and mentally laborious task. For overcoming these drawbacks, studies have proposed automatically obtaining dental information from the dental images. Each tooth was first isolated from its neighbors and the “tooth contour” was extracted on the basis of intensity.

In this study, we investigated an automated method for classifying tooth types on dental images using gray-level cooccurrence matrix (GLCM) [4] and extreme learning machine (ELM) [5-8]. Unlike the conventional methods, the proposed method is advantageous in obtaining high classification accuracy without the need for precise tooth segmentation. The deep learning methods, such as autoencoders [9-11] and convolutional neural networks [12-15], are not used, since our 52-image dataset is relatively small.

The aim of this study is to propose a robust framework for segmentation and classification of different anatomical gingival soft tissue structure from images in both accuracy and efficiency.

Method

Three gingivitis patients in Nanjing Stomatological hospital were selected to observe the gingival inflammation. Selected several teeth of each patient randomly and take images. The images used in this study were obtained using two dental Digital Single Lens Reflex (DSLR), namely A and B, which were used to acquire 33 and 19 images, respectively. Marked mesial, distal and middle three regions in mages. The images were obtained from Nanjing Stomatological Hospital, Nanjing, China. The diameter of the field of view ranged from 51 to 200 mm, and the voxel resolution ranged from 0.1 to 0.39 mm.

Therefore, in this study, the length and width of region-of-interests were manually adjusted to have an appearance similar to the model image, in which the dental region was clearly visualized. The average length was 701 ± 461 and the average width was 1338 ± 973 for 12-bit images. The gray level was then reduced to 8 bits prior to following training and testing. Figure 1 shows two examples of our dataset. Figure 1(a) shows one gingivitis image, and Figure 1(b) shows one healthy gum image.



(a) Gingivitis (b) Healthy gum
Figure 1. Finite Examples of our dataset

Grey Level Cooccurrence Matrix (GLCM) is also called as Grey Tone Spatial Dependency Matrix. It is used to find out the texture feature of an image [16]. Except GLCM, there are other commonly-used features, for example, the wavelet features [17-22]. Nevertheless, we used GLCM in this study because it can provide spatial-related texture information, which can used to identify gingivitis.

GLCM is a two-dimensional matrix which is computed using a displacement vector d , and orientation θ . The d values ranging from 1 to 10 and every pixel has eight neighboring pixels with θ value $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ$, or 315° respectively.

The gray-level co-occurrence matrix is adopted to measure the characteristics of the edge pixel pairs. We use co-occurrence matrix because sometimes information provided by histograms is not good enough for selecting a proper threshold, and the gray-level co-occurrence matrix can be an aid to histograms. GLCM calculates how often a pixel with gray-level value occurs horizontally adjacent to a pixel with the value j in an image I given an $M \times N$ neighborhood of an input image containing G gray levels from 0 to $G - 1$, the elements of the GLCM are given by equation.

$$P(i, j|d, \theta) = WQ(i, j, d|d, \theta) \quad (1)$$

where

$$W = \frac{1}{(M-d)(N-\theta)} \quad (2)$$

$$Q(i, j|d, \theta) = \sum_{n=1}^{N-\theta} \sum_{m=1}^{M-d} A \quad (3)$$

$$A = \begin{cases} 1 & \text{if } f(m, n) = i \text{ and } f(m + d, n + \theta) = j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

and $f(m, n)$ is the intensity at sample m and line n of the neighbourhood.

Extreme learning machine (ELM) is an simple and easy-to-implement classifier [23]. It can give

better performance with less computation cost than traditional classifiers, such as feedforward neural network [24-26] and support vector machine [27-29]. Given a set of N training datasets (x_i, y_i) , where $i = 1, 2, \dots, N$, $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$ and $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T \in R^m$, x_i is an n -dimensional input vector and y_i is the expected output. The output function of ELM with L hidden nodes is represented as follows:

$$\sum_{i=1}^L \beta_i g(w_i \cdot x_j + b_i) = o_j, \quad j = 1, 2, \dots, N \quad (5)$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T \in R^n$ is the weight vector of input nodes to hidden nodes and b_i is the bias of i th hidden node, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}] \in R^m$ is the weight vector between hidden nodes and the output nodes, $g(x)$ is the activation function of hidden layer, and o_j is the output vector.

If the SLFNs with L hidden nodes can approximate the N samples with zero error, we know that (5) can be converted to the following formula [29]:

$$\sum_{i=1}^L \beta_i g(w_i \cdot x_1 + b_i) = y_1 \quad (6)$$

$$\sum_{i=1}^L \beta_i g(w_i \cdot x_N + b_i) = y_N \quad (7)$$

The above equations can be written as

$$H\beta = Y, \quad (8)$$

where

$$H(w_1, \dots, w_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \quad (9)$$

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}_{L \times m} \quad (10)$$

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}_{N \times m} \quad (11)$$

Hence, training the SLFNs corresponds to finding the norm least-squares solution $\hat{\beta}$, which can be shown as follows:

$$\hat{\beta} = H^+ Y \quad (12)$$

where H^+ is the Moore-Penrose generalized inverse of hidden layer output matrix H . Then, according to KKT theorem, (12) can be expressed as

$$\beta = \begin{cases} \left(\frac{I}{\lambda} + H^T H \right)^{-1} H^T Y & N > L \\ H^T \left(\frac{I}{\lambda} + H H^T \right)^{-1} Y & N < L \end{cases} \quad (13)$$

where I is the unit matrix and λ is the regularization coefficient. Thus, the learning steps of the ELM can be summarized as in Table 1.

Table 1 Pseudocode of ELM

Step 1. Determine the structure of neural networks and set random values to the input weights w_i and the hidden layer
Step2. Calculate the hidden layer output matrix H according to (9).
Step3. Calculate the output weight vector β according to (12).

Experiments and Results

From the 52 images, 5 images each from the two imaging systems were randomly selected and used as the evaluation dataset, and the remaining were used as the training dataset, as shown in Table 2. Cost-sensitive learning was used to balance the size of gingivitis images and healthy gum images in the training set.

Table 2 Dataset division

	DSLR A	DSLR B
Training	28	14
Test	5	5
Total	33	19

To mitigate the randomness, the dataset division was randomly sampled 10 times, and the mean and standard deviation (SD) of final classification performances were reported. The results were shown in Table 3. Here TP represents gingivitis correctly identified as gingivitis, TN healthy gum correctly identified as healthy gum. FN represents gingivitis wrongly identified as healthy gum, and FP healthy gum wrongly identified as gingivitis. The average sensitivity, specificity, and accuracy of our method is 72%, 70%, and 71%, respectively.

Table 3 Statistical Results of 10 runs

Run	Indicator				Performance (%)		
	TP	FN	TN	FP	Sensitivity	Specificity	Accuracy
1	3	2	4	1	60	80	70
2	5	0	2	3	100	40	70
3	3	2	4	1	60	80	70
4	2	3	2	3	40	40	40
5	3	2	5	0	60	100	80
6	5	0	3	2	100	60	80
7	4	1	4	1	80	80	80
8	3	2	5	0	60	100	80
9	4	1	3	2	80	60	70
10	4	1	3	2	80	60	70
Average					72	70	71

Finally, we compared this proposed “GLCM+ELM” method with state-of-the-art approaches, including naïve Bayes classifier (NBC) [30] and wavelet energy (WE) [31]. The results are shown in Table 4 and Figure 2.

Table 4 Method Comparison

Method	Sensitivity	Specificity	Accuracy
NBC [30]	66	62	64
WE [31]	62	68	65
GLCM+ELM (Our)	72	70	71

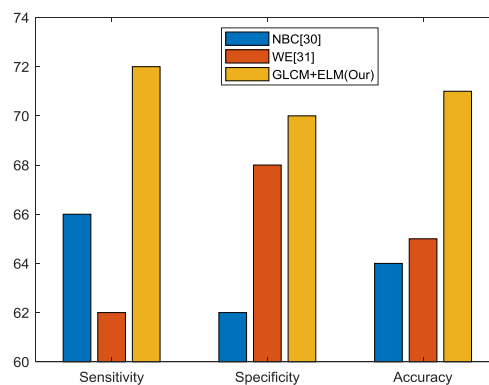


Figure 2. Finite Method Comparison

From Figure 2 we can observe that this proposed “GLCM+ELM” method gives better performance than the two basis methods, NBC [30] and WE [31], in terms of sensitivity, specificity, and accuracy. This demonstrates the effectiveness of GLCM.

Conclusions

The combination of grey-level cooccurrence matrix (GLCM) and extreme learning machine (ELM) is an efficient and accurate method which we investigated to classify tooth types and diagnose the gingivitis. Unlike the conventional methods, we don’t need to do precisely tooth segmentation before classification. After experimental processing and analysis, our research is more accurate and sensitive than state-of-the-art approaches.

However, as we can find from the experiment, some samples were identified falsely, which should be improved in the following study and experiment. The database and the number of the samples is relatively small, it will cause overfitting, which will be improved in the following research as well. This study will help the dentists from the laborious task.

References

- [1] Supranoto, S.C., et al., *The effect of chlorhexidine dentifrice or gel versus chlorhexidine mouthwash on plaque, gingivitis, bleeding and tooth discoloration: a systematic review*. International Journal of Dental Hygiene, 2015. **13**(2): p. 83-92
- [2] Sangeetha, K.M., *Effect of Triclosan Containing Tooth Paste and Conventional Fluoride Tooth Paste on Plaque and Gingivitis: A Randomized Clinical Trial*. International Journal of Scientific Study, 2017. **4**(10): p. 58-61
- [3] Ohira, H., et al., *Transmission of dental data on a microdotted paper-based dental chart from a digital pen - trial operations at disaster exercise sites and at a landslide disaster at Hiroshima City*. Canadian Society of Forensic Science Journal, 2016. **49**(4): p. 189-202
- [4] Djuricic, G.J., et al., *Fractal and Gray Level Cooccurrence Matrix Computational Analysis of Primary Osteosarcoma Magnetic Resonance Images Predicts the Chemotherapy Response*. Frontiers in Oncology, 2017. **7**: p. 8: Article ID. 246
- [5] Lu, S., *A Pathological Brain Detection System based on Extreme Learning Machine Optimized by Bat Algorithm*. CNS & Neurological Disorders - Drug Targets, 2017. **16**(1): p. 23-29
- [6] Zhao, G., *Smart Pathological brain detection by Synthetic Minority Oversampling Technique, Extreme Learning Machine, and Jaya Algorithm*. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-5023-0.
- [7] Wu, Y., *Extreme learning machine used for focal liver lesion identification*. Journal of gastroenterology and hepatology, 2017. **32**(S3): p. 168
- [8] Muhammad, K., *Ductal carcinoma in situ detection in breast thermography by extreme learning machine and combination of statistical measure and fractal dimension*. Journal of Ambient

- Intelligence and Humanized Computing, 2017, DOI: 10.1007/s12652-017-0639-5.
- [9] Hou, X.-X., *Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed*. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-4554-8.
- [10] Jia, W., *Five-category classification of pathological brain images based on deep stacked sparse autoencoder*. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-5174-z.
- [11] Jia, W., *Three-Category Classification of Magnetic Resonance Hearing Loss Images Based on Deep Autoencoder*. Journal of Medical Systems, 2017. **41**(10): Article ID. 165
- [12] Zhao, G., *Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units*. Journal of Real-Time Image Processing, 2017, DOI: 10.1007/s11554-017-0717-0.
- [13] Tang, C., *Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification on GPU platform*. Multimedia Tools and Applications, 2018, DOI: 10.1007/s11042-018-5765-3.
- [14] Lv, Y.-D., *Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling*. Journal of Medical Systems, 2018. **42**(1): Article ID. 2
- [15] Muhammad, K., *Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation*. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-5243-3.
- [16] Zainal, Z., R. Ramli, and M.M. Mustafa, *Grey-Level Cooccurrence Matrix Performance Evaluation for Heading Angle Estimation of Moveable Vision System in Static Environment*. Journal of Sensors, 2013: p. 6: Article ID. 624670
- [17] Phillips, P., *Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm*. Neurocomputing, 2018. **272**: p. 668-676
- [18] Han, L., *Identification of Alcoholism based on wavelet Renyi entropy and three-segment encoded Jaya algorithm*. Complexity, 2018. **2018**: Article ID. 3198184
- [19] Nayak, D.R., *Detection of unilateral hearing loss by Stationary Wavelet Entropy*. CNS & Neurological Disorders - Drug Targets, 2017. **16**(2): p. 15-24
- [20] Li, Y., *Detection of Dendritic Spines using Wavelet Packet Entropy and Fuzzy Support Vector Machine*. CNS & Neurological Disorders - Drug Targets, 2016. **15**, DOI: 10.2174/1871527315666161111123638
- [21] Li, P. and G. Liu, *Pathological Brain Detection via Wavelet Packet Tsallis Entropy and Real-Coded Biogeography-based Optimization*. Fundamenta Informaticae, 2017. **151**(1-4): p. 275-291
- [22] Yang, J.Q., et al., *A Novel Compressed Sensing Method for Magnetic Resonance Imaging: Exponential Wavelet Iterative Shrinkage-Thresholding Algorithm with Random Shift*. International Journal Of Biomedical Imaging, 2016: Article ID. 9416435
- [23] Pacheco, A.G.C., R.A. Krohling, and C.A.S. da Silva, *Restricted Boltzmann machine to determine the input weights for extreme learning machines*. Expert Systems with Applications, 2018. **96**: p. 77-85
- [24] Ji, G., *Fruit classification using computer vision and feedforward neural network*. Journal of Food Engineering, 2014. **143**: p. 167-177
- [25] Wu, J., *Fruit classification by biogeography-based optimization and feedforward neural network*. Expert Systems, 2016. **33**(3): p. 239-253
- [26] Wei, L., *Fruit classification by wavelet-entropy and feedforward neural network trained by fitness-scaled chaotic ABC and biogeography-based optimization*. Entropy, 2015. **17**(8): p. 5711-5728
- [27] Wu, L., *Classification of fruits using computer vision and a multiclass support vector machine*. Sensors, 2012. **12**(9): p. 12489-12505
- [28] Dong, Z., *Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree*. Progress In Electromagnetics Research, 2014. **144**: p. 171-184

- [29]Liu, A., *Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine*. Journal of Medical Imaging and Health Informatics, 2015. **5**(7): p. 1395-1403
- [30]Zhou, X. *Detection of pathological brain in MRI scanning based on wavelet-entropy and naive Bayes classifier*. in *International Conference on Bioinformatics and Biomedical Engineering (IWBBIO)*. 2015. Granada, Spain: Springer International Publishing. p. 201-209
- [31]Feng, C., et al., *Automated Classification of Brain MR Images using Wavelet-Energy and Support Vector Machines*, in *International Conference on Mechatronics, Electronic, Industrial and Control Engineering*, C. Liu, G. Chang, and Z. Luo, Editors. 2015, Atlantis Press: USA. p. 683-686.