Gingivitis identification via multichannel gray-level co-occurrence matrix and particle swarm optimization neural network

Wen Li[1,#,*], Xianwei Jiang[2,4,#], Weibin Sun[3], Yu-Dong Zhang[4,#,*], Shui-Hua Wang[5,*], Xuan Zhang[3], Leiying Miao[1,#,*],

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

2 Nanjing Normal University of Special Education, Nanjing 210038, China

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

4 Department of Informatics, University of Leicester, Leicester, LE1 7RH, UK

5 School of Architecture Building and Civil engineering, Loughborough University, Loughborough, LE11 3TU, UK

# Those authors contributed equally to this paper

* Email: L Miao (miaoleiying80@163.com), S H Wang (shuihuawang@ieee.org), Y D Zhang (yudongzhang@ieee.org), W Li (18844501367@163.com)

Abstract: The oral maintenance of patients with periodontal disease mainly depends on clinical examination. However, insufficient number of medical workers cannot carry out detailed oral health education for a large number of patients within limited time and provide these patients with proper and effective oral health nursing methods under the guidance of doctors. In this study, our research presented a new artificial intelligence method to diagnose chronic gingivitis, which is based on contrast limited adaptive histogram equalization (CLAHE) and multichannel gray-level co-occurrence matrix (MGLCM) and particle swarm optimization neural network (PSONN). Meanwhile, different training algorithms have been used as comparison groups. The dataset contains 800 images: 400 chronic gingivitis images and 400 healthy gingiva images. The results certify that the sensitivity, specificity, precision, accuracy and F1 Score of our MGLCM (PSONN as a classifier) method is 78.17%, 78.23%, 78.24% ,78.20% and 78.17%, respectively. The combination of CLAHE and MGLCM and PSONN is more efficient and accurate than state-of-the-art approaches: NBC, WN+SVM, ELM and CLAHE+ELM.

Keywords: multichannel gray-level co-occurrence matrix; contrast limited adaptive histogram equalization; artificial neural network; particle swarm optimization; gingivitis identification; pattern recognition

1 Introduction

Gingivitis, of which the main manifestations are redness, bleeding and bad breath, is a common chronic oral infectious disease of gingival tissue induced by bacterial infection [1]. The continuous progression of gingivitis without prompt treatment will further destroy the periodontal tissues, such as the gingiva, alveolar bone and pericementum around the teeth [2], which eventually results in Looseness, and

detachment of teeth. It has been reported that periodontosis was the most widespread cause of teeth loss among adults [3]. Therefore, chronic gingivitis not only affects the masticatory function, facial aesthetics and pronunciation function of patients, but also further endangers patients' physical and mental health and quality of life [4, 5].

Clinical examination is conventionally and widely used to diagnose and predict gingivitis. However, it costs a large number of human labor and has problems such as bias, delay of diagnosis. Therefore, it tends to use the artificial intelligence auxiliary means of clinical diagnosis and prediction. In this study, we presented a new Artificial Intelligence (AI) based method to identify chronic gingivitis, which is based on multichannel gray-level co-occurrence matrix (MGLCM) and particle swarm optimization neural network (PSONN). The gray level co-occurrence matrix (GLCM) as we all known was a spatial correlation feature extraction method used in texture analysis for big data [6]. To our knowledge, the GLCM is a powerful method, which states the relationship between a pixel of gray level *i* and a pixel of gray level *j* appears in a specific spatial. However, GLCM method is concerned with multi/hyperspectral images and is not always suitable for texture feature extraction of multispectral images. Therefore, Lucieer, Stein (2005) [7] produced a multivariate local binary pattern (MLBP) for segmentation of remotely sensed images, which proved the necessity of multichannel textural extraction. Palm and Lehmann (2002) [8] also proposed a Gabor filtering in RGB color space. It has been approved that the color textures can achieve better results than the grayscale features. Some studies have been reported for other common algorithms on image feature extraction. For instance, Zhou (2015) [9] used naïve Bayesian classifier (NBC) to classify pathological brains. Feng, Zhang (2015) [10] used wavelet energy (WN) and support vector machine (SVM). Brown (2018) [11] employed extreme learning machine (ELM) to identify gingivitis. Li (2019) [12] used contrast-limited adaptive histogram equalization (CLAHE) and ELM, and their accuracy procured 74%. Multichannel Level Co-Occurrence Matrix (MGLCM) is an improved method of the GLCM and showing the Spectroscopic information of the multispectral photography. Whether the MGLCM method can produce valid features for multispectral image classification is still worth testing.

Compared with gradient-based algorithms, Particle swarm Optimization (PSO) is a swarm intelligence method for global optimization and does not need any gradient information [13]. Also, it is clearly demonstrated that PSO can produce better results in a cheaper, faster way compared with many other methods. Therefore, in this study, the purpose was to evaluate the potential accuracy and effectiveness of computer-vision and image-processing method CLAHE, MGLCM and PSO algorithm for diagnosing and predicting gingivitis.

The contribution of this study are two folds: (i) We introduced a multichannel GLCM (MGLCM) method to replace the traditional single GLCM method. (ii) We used the particle swarm optimization neural network (PSONN) as a stable and reliable classifier.

The paper is structured as follows: Section 2 shows the methodology of this paper. Section 3 shows the experiments and results. Section 4 discusses our results and section 5 concludes this paper.

2 Method

2.1 Subjects and Dataset

First, we selected twenty gingivitis patients and people with healthy gums in Department of Periodontics, Nanjing Stomatological Hospital between January 2018 and December 2018 randomly. Some

exclusion and inclusion criteria should be formulated in order to ensure the accuracy of the experiment. Inclusion criteria: According to the criteria of World Health Organization-standardized for the diagnosis of chronic gingivitis, pink and pliable gingiva with a clinical attachment level (CAL) of<3 mm in a periodontal probe examination were diagnosed as healthy gingiva. Teeth was probed for bleeding and there was no loss of attachment or alveolar resorption were diagnosed as gingivitis. Exclusion criteria: The teeth with severe cervical caries, severe periodontitis with evident gingival atrophy and residual root and crown.

Via randomly selecting several teeth with inflammatory gums or healthy gums of each candidate and using Digital Single Lens Reflex (DSLR) for periodontal disease to gain images. 400 gingivitis images and 400 healthy gum images were acquired to build the training dataset. Then marked the inflammation areas of gingiva in each image to identify easily. The diameter of the field of view ranged from 26 to 100 mm, and the voxel resolution ranged from 0.2 to 0.41 mm. Adjusted the length and width of region of interests (ROI) to have similar size to the virtual image. All the authors had access to information that could identify individual participants during or after data collection.

2.2 CLAHE

Plain histogram equalization (HE) can adjust the histogram of an image globally. Suppose a discrete image [x], the probability of an occurrence of a pixel of level k in the image is:

$$p_x(k) = p(x = k) = \frac{n_k}{n} \tag{1}$$

where n_k is the number of occurrences of graylevel values k.

The cumulative distribution function (CDF) is defined as

$$CDF_x(k) = \sum_{m=0}^{k} p_x(m) \tag{2}$$

The aim of HE is to create a new transformation y = T(x) to produce a new image [y] which has a flat histogram as shown in **Figure 1**.

$$CDF_y(k) = ck \tag{3}$$

where c is some constant.

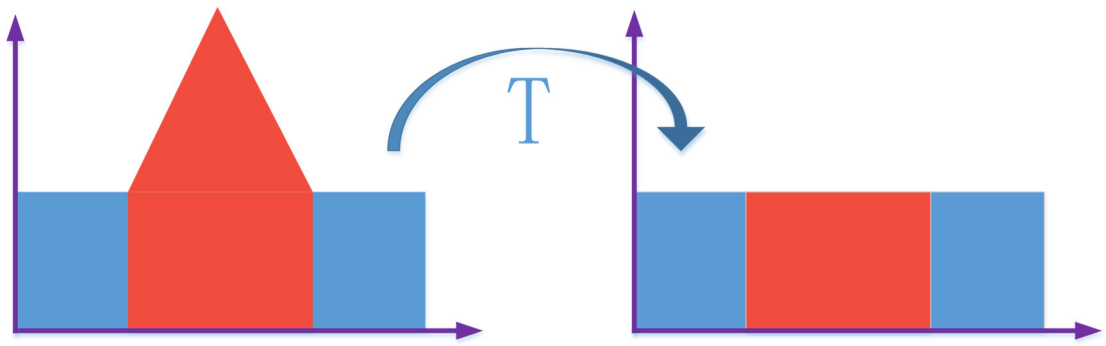


Figure 1 Illustration of histogram equalization

Contrast-limited adaptive histogram equalization (CLAHE) is an advanced HE. CLAHE employs bilinear interpolation for combining neighboring tiles without artifacts. Compared to standard HE, the CLAHE has three pros: (a) It depresses the noises especially in the homogeneous areas, (b) It uses the histogram of different tiles to re-distribute the lightness value of the image, (c) It improves the local

contrast and enhancing the definitions of edges.

2.3 Multichannel gray-level co-occurrence matrix

The method of extracting texture features based on gray level co-occurrence matrix (GLCM) is extremely classical in statistical analysis field, which was first proposed by Haralick in 1973. It gradually turned into a commonly used method and measurement technology in dealing with texture feature [14]. GLCM gives the number of times between two related pixels with grayscale intensity value i and j in the image, which is actually a statistical form of the joint distribution of such two pixels. We can use the following formula to represent GLCM.

$$P(\mathrm{I},\mathrm{j},\mathrm{D},\theta) = \{[(\mathrm{x},\mathrm{y}),(\mathrm{x}+\Delta\mathrm{x},\mathrm{y}+\Delta\mathrm{y})]|\mathrm{f}(\mathrm{x},\mathrm{y}) = \mathrm{I},\mathrm{f}(\mathrm{x}+\Delta\mathrm{x},\mathrm{y}+\Delta\mathrm{y}) = \mathrm{j}; \quad \mathrm{x} = 0,1,2,\cdots,\mathrm{N_x}-1;\mathrm{y} = 0,1,2,\cdots\mathrm{N_y}-1; i,j = 0,1,\ldots,L-1\} \tag{4}$$

where L is the number of gray levels of the image, the values of i and j range from 0 to L-1, x and y are the pixel coordinates in the image, $\Delta\mathrm{x}$ and $\Delta\mathrm{y}$ represent the offset, N_x and N_y are the number of rows and columns of the image. θ indicates the direction and the general calculation process will be carried out in several different directions such as horizontal 0°, vertical 90°, and 45° and 135°. Hence, the gray level co-occurrence matrix calculates the probability $P(\mathrm{I},\mathrm{j},\mathrm{D},\theta)$ from the pixel whose gray level is i and coordinates is (x, y) to another pixel whose distance is D and the gray level is j appearing at the same time [15].

For the sake of simplicity, we generally use the features, that is, energy (ENE), contrast (CON), entropy (ENT), and correlation (COR) to extract the texture features of the image. Their calculation formulas are given below.

$$\mathrm{ENE} = \sum_i \sum_j P(i,j)^2 \tag{5}$$

$$\mathrm{CON} = \sum_i \sum_j (i-j)^2 P(i,j) \tag{6}$$

$$\mathrm{ENT} = -\sum_i \sum_j P(i,j)\log P(i,j) \tag{7}$$

$$\mathrm{COR} = \frac{[\sum_i \sum_j ((i,j)P(i,j) - \mu_x\mu_y]}{\sigma_x\sigma_y} \tag{8}$$

Here ENE is a measure of the uniformity of grayscale changes in image texture. CON measures how the values of the matrix are distributed and determines how much local changes in the image. The randomness of image texture is measured by ENT. COR gives how similar the spatial gray level co-occurrence matrix elements are in row or column direction.

Although GLCM is very efficient and powerful, in practice we often need to extract texture-based features on multi-channel images. Therefore, multichannel GLCM is introduced in this study. An effective method is to perform GLCM on each RGB channel separately. This laid the foundation for the creation of multichannel gray level co-occurrence matrices (MGLCM).

Nevertheless, there are other texture feature extraction techniques, such as: edge detection, scale-invariant feature transform (SIFT), template matching, Hough transform, wavelet transform, local binary pattern, etc. We chose MGLCM due to its simplicity and effectiveness.

2.4 Artificial Neural Network

The artificial neural network (ANN) abstracts the human brain neural network based on the information processing perspective, establishes a simple model, and forms different networks according to different connection methods [16]. ANN is the classifier model we used for gingivitis identification.

Hence, we add the description here in methodology, in order for the readers understanding our methodology.

A neural network is an operational model consisting of a large number of nodes (or neurons) connected to each other. Each node indicates a specific output function called an activation function. The connection between every two nodes represents a weighting value for passing the connection signal, called weight, which is equivalent to the memory of the artificial neural network. The output of the network varies depending on the connection method of the network, the weight value and the excitation function. The mathematical representation of the output of a single neuron is given below.

$$t = f(wA' + b) \tag{9}$$

The function of a neuron is to obtain a scalar result via a nonlinear transfer function after finding the inner product of the input vector and the weight vector [17]. Where t is the neuron output, f is the transfer function, W is the weight vector, A is the input vector, A' is the transpose of the vector A, and b is the bias vector. The following **Figure 2** shows a common multilayer feedforward network consisting of three parts: input layer, hidden layer, and output layer.

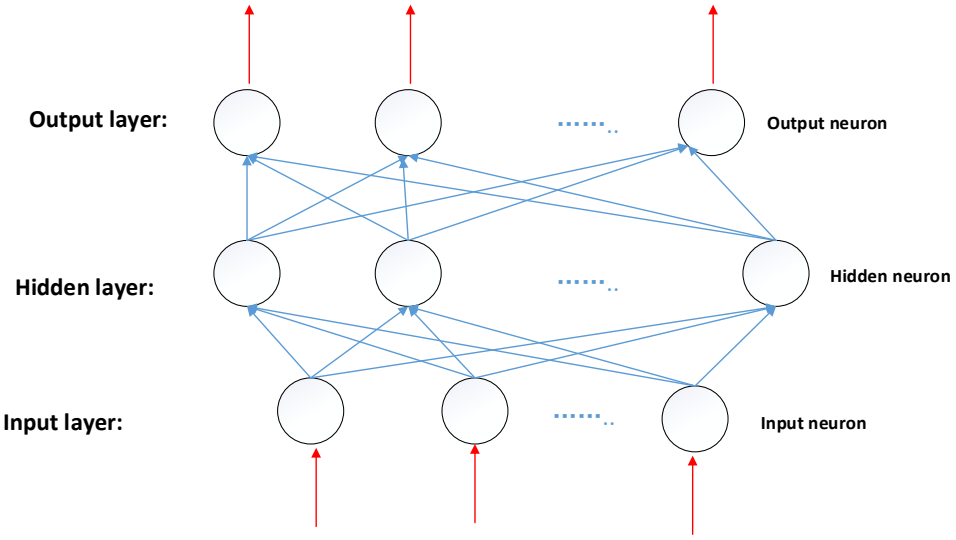


Figure 2 Diagram of artificial neural network

Here, at the input level, many neurons receive a large number of non-linear input messages. The hidden layer is the layer of many neurons and links between the input layer and the output layer, which can have one or more layers. The number of hidden layers determines the robustness of the neural network. Information are transmitted, analyzed, weighed in neuron links, and the output is formed and finally given at the output layer.

The training network is terminated until the training error reaches the required condition (mean squared error, MSE). The MSE is defined as follows.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(p_i - e_i)^2 \tag{10}$$

where n indicates the number of samples of data set, p_i is the predicted value, and e_i represents the realistic output of samples.

On the other hand, deep learning is currently widely used as a common classifier. The reason that we do not choose deep learning in this study is because the size of the dataset is only 800, relatively small to train a complicate model such as deep neural network.

2.5 Particle Swarm Optimization

There are many traditional neural network training methods, e.g., back-propagation (BP) [18], genetic algorithm (GA) [19], multi-agent GA (MAGA) [20], and simulated annealing (SA) [21] approaches. Nevertheless, those methods still may get stuck in local optimal solutions.

Particle Swarm Optimization (PSO) is a random search algorithm based on swarm co-operation developed by imitating bird foraging behavior, which is developed by J. Kennedy et al. in 1995. Its basic core is to make use of the individual's sharing of information in the group so that the whole group's motion can be transformed from disorder , so that gain the optimal solution of the problem [22]. In PSO, the solution to each optimization problem is equal to a bird in the search space, which is called particles. All particles have the associated fitness values which are determined by the fitness function, and each particle has a velocity that determines the direction and distance they fly [23]. The particles then follow the current optimal particle to search in the solution space. **Figure 3** shows the flowchart of particle swarm optimization.

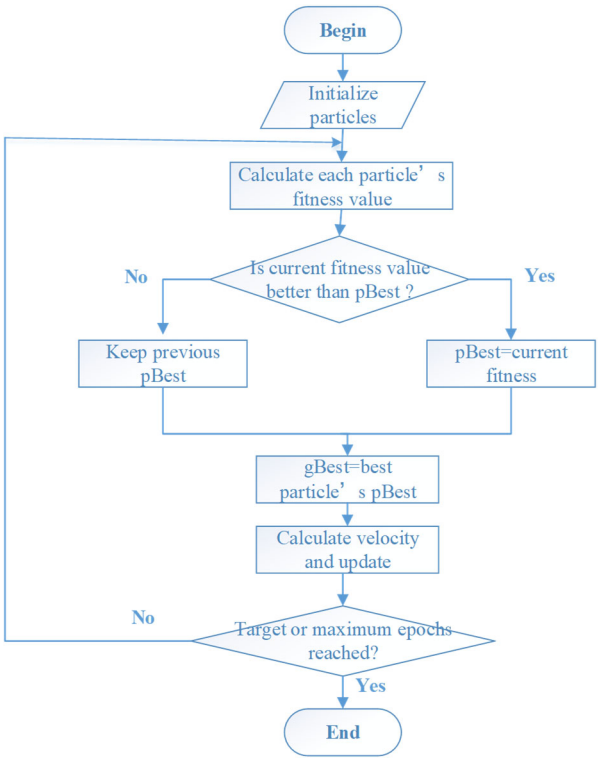


Figure 3 Flowchart of particle swarm optimization

The PSO algorithm sets a population of random particles initially and reiteratively approaches the optimum solution [24]. In each iteration, the particles update themselves by tracking two "extreme values". The first is the optimal solution found by the particle itself. This solution is called the individual extremum *pBest*, and the other extremum is the optimal solution found by the entire population. This extremum is the global extremum *gBest*. The position of the particle is updated as in the equation below.

$$x_i(t) = x_i(t-1) + v_i(t) \tag{11}$$

$$v_i(t) = wv_i(t-1) + C_1r_1[pbest - x_i(t)] + C_2r_2[gbest - x_i(t)] \tag{12}$$

where w is the inertia weight, and C_1 and C_2 are the acceleration constants. v_i indicates the velocity of the particle, x_i presents the position of the current particle. The algorithm will not stop until the specified number of iterations or error values is reached.

2.6 Implementation

Figure 4 shows the pipeline of proposed method. Here the input teeth images were firstly enhanced by CLAHE algorithm [25]. Second, each channel of R, G, B are sent to calculate the corresponding GLCM matrix. Then, the energy (ENE), contrast (CON), entropy (ENT), and correlation (COR) are calculated for each GLCM matrix. Those features are concatenated to a row vector as a representation for the input teeth image. Finally, the matrix was split via k-fold cross validation, and the model was trained on the basis of artificial neural network with training method of particle swarm optimization.

Ten runs of 10-fold cross validation were carried out. That means each fold will contains 80 images, 40 of which are gingivitis and the rest healthy. Nine folds were used for training, and the rest fold for test. The ideal confusion matrix on the test set is

$$C = \begin{bmatrix} 40 & 0 \\ 0 & 40 \end{bmatrix} \tag{13}$$

After single run of 10-fold cross validation (we have 10 test sets), the ideal confusion matrix is the summation of 10 trials, which is:

$$C = \begin{bmatrix} 400 & 0 \\ 0 & 400 \end{bmatrix} \tag{14}$$

Finally, notice that we run this 10-fold cross validation 10 times, hence, the ideal confusion matrix will be 10 times of the previous confusion matrix. That is,

$$C = \begin{bmatrix} 4000 & 0 \\ 0 & 4000 \end{bmatrix} \tag{15}$$

The measures include the sensitivity, specificity, precision, accuracy, and F1-score over all ten runs in the format of mean plus standard deviation. Suppose TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative, respectively, then above measures are defined as below:

$$SEN = \frac{TP}{TP+FN} \tag{16}$$

$$Spc = \frac{TN}{TN+FP} \tag{17}$$

$$Prc = \frac{TP}{TP+FP} \tag{18}$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \tag{19}$$

$$F1 = 2TP/(2TP + FP + FN) \tag{20}$$

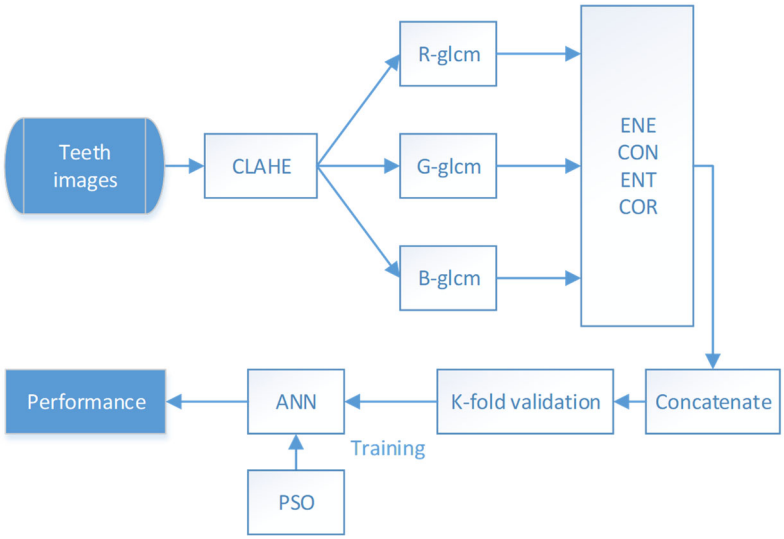
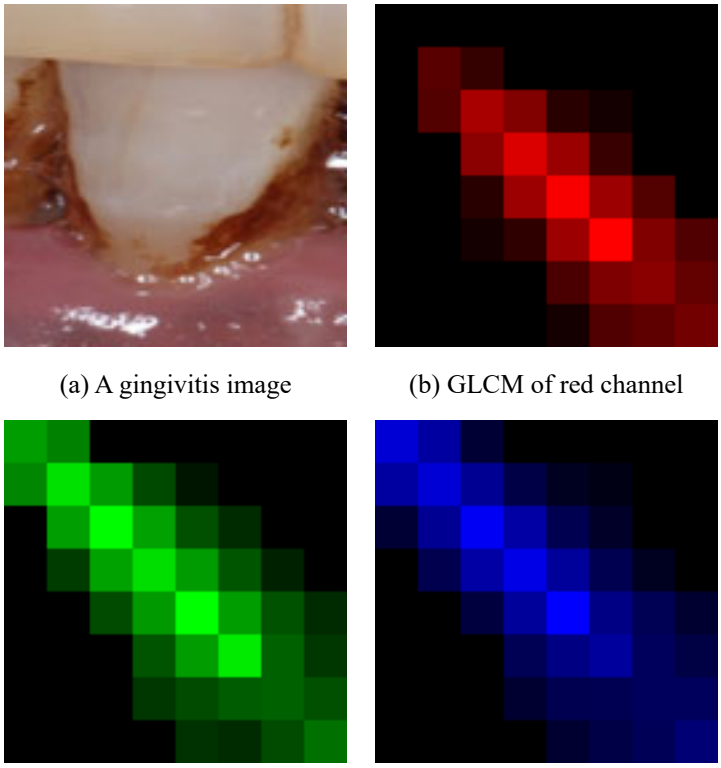


Figure 4 Pipeline of proposed method

3 Experiment and Results

3.1 MGLCM Results

Figure 5(a) shows an original gingivitis image. Figure 5(b-d) show the GLCM matrix of red, green, and blue channels. For each GLCM, we calculate the energy, contrast, entropy, and correlation measures. The greatest values lie along the diagonal line, since the adjacent values are similar most of the time.



(a) A gingivitis image (b) GLCM of red channel

(c) GLCM of green channel (d) GLCM of blue channel

Figure 5 Multichannel GLCM (The pixel values in the pictures are log scaled for clear vision)

3.2 Classification Results

The feature matrix of all samples was sent to ANN classifier trained via PSO. The classification results are shown in **Table 1**. We can observe that the overall sensitivity, overall specificity, overall precision, overall accuracy, and overall F1-score are 78.17± 2.53%, 78.23± 1.05%, 78.24± 0.66%, 78.20± 1.04%, and 78.17± 1.36%, respectively.

Table 1 Results of 10x10-fold cross validation of our method

Run	Sensitivity	Specificity	Precision	Accuracy	F1 Score
1	73.75	79.50	78.27	76.63	75.91
2	79.50	79.75	79.72	79.62	79.60
3	73.25	79.50	78.13	76.38	75.60
4	78.00	77.25	77.46	77.63	77.69
5	79.00	76.75	77.28	77.88	78.11
6	78.25	79.00	78.86	78.63	78.54
7	79.75	77.50	78.02	78.63	78.85
8	81.75	77.50	78.42	79.62	80.05
9	79.25	78.00	78.38	78.63	78.75
10	79.25	77.50	77.88	78.38	78.55
Mean±SD	78.17± 2.53	78.23± 1.05	78.24± 0.66	78.20± 1.04	78.17± 1.36

3.3 Training algorithm comparison

In this experiment, we compared PSO algorithm with traditional training methods, e.g., back-propagation (BP) [18], genetic algorithm (GA) [19], multi-agent genetic algorithm (MAGA) [20], and simulated annealing (SA) [21] approaches. The comparison results are shown in **Table 2** and **Figure 6**.

Table 2 Training algorithm comparison

Training method	Sensitivity	Specificity	Precision	Accuracy	F1 Score
BP [18]	71.98± 3.12	72.00± 3.52	72.05± 3.14	71.99± 2.98	71.98± 2.93
GA [19]	76.33± 2.62	76.35± 2.76	76.41± 2.46	76.34± 2.30	76.33± 2.31
MAGA [20]	77.38± 2.17	77.35± 1.47	77.39± 1.49	77.36± 1.66	77.36± 1.74
SA [21]	71.40± 4.72	71.50± 4.00	71.48± 4.07	71.45± 4.18	71.41± 4.31
PSO (Ours)	**78.17± 2.53**	**78.23± 1.05**	**78.24± 0.66**	**78.20± 1.04**	**78.17± 1.36**

(Bold means the best)

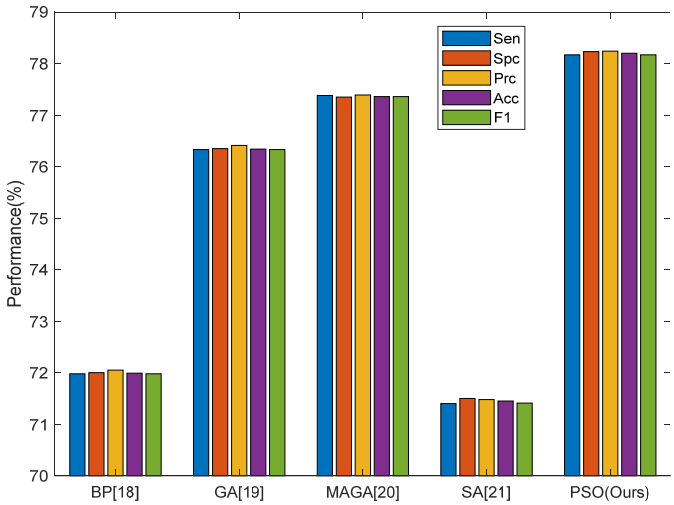


Figure 6 Training method comparison

3.4 Multichannel versus single channel

We first need to compare the multichannel GLCM against single channel GLCM. We replaced the multichannel GCLM in Figure 4 with single-channel GLCM, and re-do the experiment. All the other components in the methodology are not changed. The results of using single-channel GLCM are shown in **Table 3**.

Table 3 Results of 10x10-fold cross validation of single-channel GLCM method

Run	Sensitivity	Specificity	Precision	Accuracy	F1 Score
1	75.25	76.00	75.82	75.63	75.51
2	80.25	73.75	75.37	77.00	77.71
3	74.50	73.00	73.45	73.75	73.95
4	76.00	75.50	75.63	75.75	75.81
5	76.50	75.00	75.41	75.75	75.93
6	73.75	73.75	73.78	73.75	73.74
7	75.25	73.50	73.98	74.38	74.58
8	75.00	79.75	78.78	77.38	76.82
9	76.75	79.50	78.95	78.13	77.81
10	71.25	74.75	73.85	73.00	72.51
Mean± SD	75.45± 2.19	75.45± 2.27	75.50± 1.87	75.45± 1.63	75.44± 1.66

The comparison of single-channel GLCM with multichannel GLCM is listed in **Table 4**. The error bar is shown in **Figure 7**. As is seen, using multichannel GLCM can get better performance. The reason is multichannel GLCM capitalize all the color information, while sing-channel GLCM get rid of color information. Hence, the former one can procure better results than the latter one.

Table 4 Comparison of single-channel GLCM with multichannel GLCM (PSONN as classifier)

Method	Sensitivity	Specificity	Precision	Accuracy	F1 Score

Single-channel GLCM	75.45± 2.19	75.45± 2.27	75.50± 1.87	75.45± 1.63	75.44± 1.66
Multichannel GLCM	78.17± 2.53	78.23± 1.05	78.24± 0.66	78.20± 1.04	78.17± 1.36

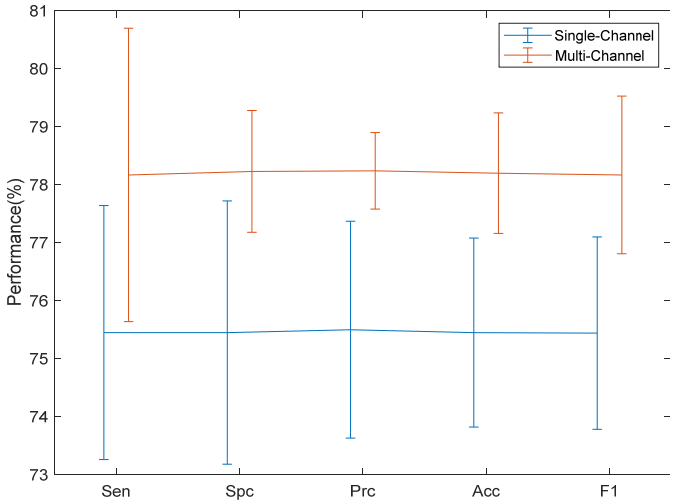


Figure 7 Error bar of single-channel and multi-channel methods

3.5 Comparison of State-of-the-art Approaches

In this experiment, we compared our proposed "CLAHE+MGLCM+PSONN" method with state-of-the-art approaches: NBC [9], WN+SVM [10], ELM [11], CLAHE+ELM [12]. The comparison results were listed below in

Table 5 and **Figure 8**. We could observe that our method gained the best accuracy among all five methods.

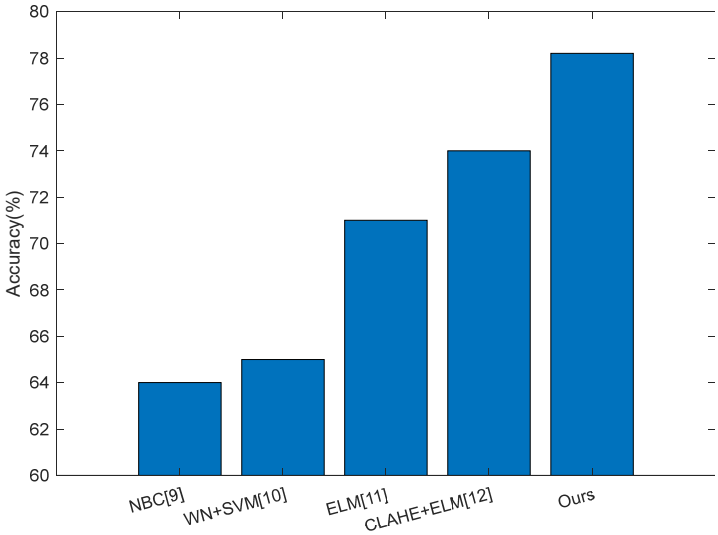


Figure 8 Bar plot of state-of-the-art algorithm comparison

Table 5 Algorithm comparison

Approach	Accuracy
NBC [9]	64%

WN+SVM [10]	65%
ELM [11]	71%
CLAHE+ELM [12]	74%
CLAHE+MGLCM+PSONN (Ours)	**78.20%**

(Bold means the best)

4 Discussions

Chronic gingivitis as one of the most common oral diseases can affect the health of residents in our country, in recent 10 years, the national oral health epidemiological survey results showed that the periodontal health status and oral hygiene of the elderly is obviously decreased. With dramatic increasing of chronic gingivitis patients, insufficient number of medical workers cannot carry out detailed oral health education for a large number of patients within limited time and provide these patients with proper and effective oral health nursing methods under the guidance of doctors. Recently, some researches have demonstrated the feasibility and accuracy of artificial intelligence methods in processing clinical photographs, cone beam computed tomography (CBCT) and magnetic resonance imaging (MRI) images [26, 27]. Particularly, both X-rays and CT play the vital role in the auxiliary diagnosis of chronic gingivitis according to the study. Although imaging is valuable in diagnosing chronic gingivitis, it is cumbersome and time-consuming. Consequently, this research proposed a new computer-vision and image-processing method—CLAHE+MGLCM+PSONN—to diagnose and predict chronic gingivitis.

We used CLAHE to enhance the contrast and edges of input images, so as to reduce the noises and improves the local contrast at the first phase. As we all know, GLCM is a method of feature extraction of texture images [28, 29], which can be an aid to threshold selection, because information provided by histograms is not good enough for selecting a proper threshold. In this research, we propose to calculate the MGLCM of each image and seek out the correlation coefficient between frames. Compared to traditional single channel GLCM, our methods MGLCM can provide exact results for multi/hyper-channel images. From **Table 1**, we use MGLCM method to identify Chronic gingivitis and the results showed that the overall sensitivity, overall specificity, overall precision, overall accuracy, and overall F1-score are 78.17± 2.53%, 78.23± 1.05%, 78.24± 0.66%, 78.20± 1.04%, and 78.17± 1.36%, respectively. However, from **Table 3**, we replaced the multichannel GCLM with single-channel GLCM, the results showed that the sensitivity, specificity, precision, accuracy, and F1-score are 75.45± 2.19%, 75.45± 2.27%, 75.50± 1.87%, 75.45± 1.63% and 75.44± 1.66%, respectively. As is seen from **Table 4**, using multichannel GLCM can get better performance. The reason is multichannel GLCM capitalize all the color information, while sing-channel GLCM get rid of color information. Hence, the former one can procure better results than the latter one.

The PSO is an population-based stochastic optimization algorithm, which was proposed by Eberhart and Kennedy in 1995 [30, 31]. It generally relies on the foraging model of birds to find the optimal value. Each particle within the particle swarm continuously changes its search mode by learning the experience of other particles and records the best local position easier and faster they have discovered. Because of the superiority of PSO algorithm, Khatir, Dekemele (2018) [32] used PSO algorithm to forecast the location of damage of a steel cantilever beam. Liao, Liu (2012) [33] adopted an PSO to detect the damage of several kind of framework structures with a high accuracy. Wu, Cole (2016) [34], Pau, Collotta (2017) [35] and other researchers also found the superiority of PSO in dealing with other problems. From **Table 2** and **Figure 6**, we compared PSO algorithm with traditional training methods, back-propagation (BP),

genetic algorithm (GA), multi-agent genetic algorithm (MAGA), and simulated annealing (SA) approaches and the comparison results shown that the sensitivity, specificity, precision, accuracy, and F1-score of c in diagnosing and identifying chronic gingivitis are 75.45± 2.19%, 75.45± 2.27%, 75.50± 1.87%, 75.45± 1.63%, and 75.44± 1.66%, respectively. In those algorithms, PSO algorithm performed best. The Genetic Algorithm (GA) is a search algorithm used efficiently to solve different kind of optimization problems in computational mathematics [36, 37], which was developed from phenomena in evolutionary biology, including heredity, mutation, natural selection, and hybridization.

In GA, evolution occurs from generation to generation and chromosomes shared the information with each other after a iteration, nevertheless in PSO, crossover and mutation is not employed. Instead, PSO uses a simple formula to update the positions of each particle. On the other hand, the Back-Propagation (BP) method is the primary method of artificial neural networks; however, this method is prone to get stuck in the local minima and also experiences slower concentration rate towards the optimum solution [38]. PSO performs better than BP and GA in terms of rate of concentration [39, 40].

According to the above results , we compared our proposed "CLAHE+MGLCM+PSONN" method with state-of-the-art approaches: NBC [9], WN+SVM [10], ELM [11], CLAHE+ELM [12]. the results in

Table 5 and **Figure 8** showed that "CLAHE+MGLCM+PSONN" method performed the best accuracy 78.20% in differentiating chronic gingivitis. As early as before, the Naïve Bayes Classifier is a machine learning algorithm based on Bayes' theorem with independence and normality assumptions among the variables, which has been known to have the ability to work efficiently to develop classification tools in various health domains. An advantage of Naïve Bayes is that it only requires a small amount of training data to estimate the parameters necessary for classification, so the NBC is not applicable to classification of big data. Support Vector Machines (SVMs) are a relatively new form of machine learning that was developed by Vapnik (1995) [41], which is used to refer to both classification and regression methods. Zhang, Li (2012) [42] also used "WN+SVM" methods to Classify the power quality disturbances and performed the best result. However, in this research, the accuracy of "WN+SVM" was lower because SVM algorithm is difficult to implement for large-scale training samples, the same reason for the "CLAHE+ELM".

5 Conclusions

The combination of CLAHE and MGLCM and PSONN is an efficient and accurate method, which we investigated to classify tooth types and diagnose the chronic gingivitis. After experimental processing and analysis, our research is more accurate and sensitive than state-of-the-art approaches: NBC, WN+SVM, ELM and CLAHE+ELM.

However, as we can find from the experiment, the database number of the samples is relatively small and this maybe the reason that some samples were identified falsely, which should be improved in the following study and experiment. In addition, this research provides new ideas with the application of artificial intelligence technology to diagnose periodontal disease and will help the dentists from the laborious task.

Acknowledgement

This work was carried out by Stomatological Hospital of Nanjing University. The paper was

supported by The Project of Invigorating Health Care through Science, Technology and Education Jiangsu Provincial Medical Youth Talent (QNRC2016120), National key research and development plan (2017YFB1103202), Natural Science Foundation of Zhejiang Province (Y18F010018), Henan Key Research and Development Project (182102310629), Open Fund of Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology (17-259-05-011K).

Author Contributions

Sample Collection: Wen Li, Xuan Zhang, Leiying Miao, Weibin Sun
Data Processing: Yu-Dong Zhang, Shui-Hua Wang
Methodology: Wen Li, Xianwei Jiang, Yu-Dong Zhang, Shui-Hua Wang
Writing – original draft: Wen Li, Xianwei Jiang, Yu-Dong Zhang
Writing – review & editing: Weibin Sun, Leiying Miao

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.* Int J Dent Hyg, 2015. **13**(2): p. 83-92
2. JS, L., et al., *National dental policies and socio-demographic factors affecting changes in the incidence of periodontal treatments in Korean: A nationwide population-based retrospective cohort study from 2002-2013.%A Lee JH.* BMC oral health, 2016. **16**(1): p. 118
3. JY, O., et al., *Trends in the incidence of tooth extraction due to periodontal disease: results of a 12-year longitudinal cohort study in South Korea.%A Lee JH.* Journal of periodontal & implant science, 2017. **47**(5): p. 264-272
4. AB, N., et al., *Effect of surgical periodontal treatment associated to antimicrobial photodynamic therapy on chronic periodontitis: A randomized controlled clinical trial.%A Martins SHL.* Journal of clinical periodontology, 2017. **44**(7): p. 717-728
5. D, K., et al., *Nonsurgical and surgical treatment of periodontitis: how many options for one disease?%A Graziani F.* Periodontology 2000, 2017. **75**(1): p. 152-188
6. Ribeiro, C., et al., *A clinical, radiographic, and scanning electron microscopic evaluation of adhesive restorations on carious dentin in primary teeth.* Quintessence Int, 1999. **30**(9): p. 591-9
7. Lucieer, A., et al., *Multivariate texture-based segmentation of remotely sensed imagery for extraction of objects and their uncertainty.* International Journal of Remote Sensing, 2005. **26**(14): p. 2917-2936
8. Palm, C., et al., *Classification of color textures by Gabor filtering.* Machine Graphics & Vision International Journal, 2002. **11**(2/3): p. 195-219
9. 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
10. 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.
11. Brown, M., *Gingivitis identification via grey-level cooccurrence matrix and extreme learning machine.* Advances in Social Science, Education and Humanities Research, 2018. **250**: p. 486-492
12. Li, W., *A gingivitis identification method based on contrast limited adaptive histogram equalization, grey-level cooccurrence matrix, and extreme learning machine.* International Journal of Imaging Systems and Technology, 2019. **29**(1): p. 77-82

13. Marouani, H., et al., *Particle swarm optimization performance for fitting of Levy noise data.* Physica A-Statistical Mechanics and Its Applications, 2019. **514**: p. 708-714
14. Nouri, M., et al., *Non-destructive Evaluation of Bread Staling Using Gray Level Co-occurrence Matrices.* Food Analytical Methods, 2018. **11**(12): p. 3391-3395
15. Yamamoto, M., et al., *Quantitative distinction of neutrophil alkaline phosphatase score through texture analysis using gray level co-occurrence matrix.* International Journal of Laboratory Hematology, 2018. **40**: p. 160-160
16. Kumar, S., et al., *Prediction of jatropha-algae biodiesel blend oil yield with the application of artificial neural networks technique.* Energy Sources Part A-Recovery Utilization and Environmental Effects, 2019. **41**(11): p. 1285-1295
17. Roxas, C.L.C., et al., *An artificial neural network model for the corrosion current density of steel in mortar mixed with seawater.* International Journal of Geomate, 2019. **16**(56): p. 79-84
18. Saeedi, E., et al., *Feed-Forward Back-Propagation Neural Networks in Side-Channel Information Characterization.* Journal of Circuits Systems and Computers, 2019. **28**(1): p. 18: Article ID. 1950003
19. Abdollahi, H., et al., *Prediction and optimization studies for bioleaching of molybdenite concentrate using artificial neural networks and genetic algorithm.* Minerals Engineering, 2019. **130**: p. 24-35
20. Zuo, M.C., et al., *Multi-agent genetic algorithm with controllable mutation probability utilizing back propagation neural network for global optimization of trajectory design.* Engineering Optimization, 2019. **51**(1): p. 120-139
21. Bagherlou, H., et al., *A routing protocol for vehicular ad hoc networks using simulated annealing algorithm and neural networks.* Journal of Supercomputing, 2018. **74**(6): p. 2528-2552
22. Jana, B., et al., *Repository and Mutation based Particle Swarm Optimization (RMPSO): A new PSO variant applied to reconstruction of Gene Regulatory Network.* Applied Soft Computing, 2019. **74**: p. 330-355
23. Prakash, C., et al., *Multi-objective particle swarm optimization of EDM parameters to deposit HA-coating on biodegradable Mg-alloy.* Vacuum, 2018. **158**: p. 180-190
24. Ebadi, Y., et al., *An energy-aware method for data replication in the cloud environments using a Tabu search and particle swarm optimization algorithm.* Concurrency and Computation-Practice & Experience, 2019. **31**(1): p. 10: Article ID. e4757
25. Garg, D., et al., *Underwater image enhancement using blending of CLAHE and percentile methodologies.* Multimedia Tools and Applications, 2018. **77**(20): p. 26545-26561
26. Lehman, C.D., et al., *Diagnostic accuracy of digital screening mammography with and without computer-aided detection.* JAMA internal medicine, 2015. **175**(11): p. 1828-1837
27. Gulshan, V., et al., *Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs.* JAMA, 2016. **316**(22): p. 2402-2410
28. Haralick, R.M., et al., *Textural features for image classification.* IEEE Transactions on systems, man, and cybernetics, 1973(6): p. 610-621
29. Mokji, M., et al. *Gray level co-occurrence matrix computation based on haar wavelet*. in *Computer Graphics, Imaging and Visualisation (CGIV)*. 2007. Bangkok, Thailand: IEEE. p. 38-45
30. Kennedy, J., et al. *Particle swarm optimization*. in *IEEE International Conference on Neural Networks*. 1995. IEEE. p. 1942-1948
31. Eberhart, R., et al. *A new optimizer using particle swarm theory*. in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. 1995. p. 39-43
32. Khatir, S., et al., *Crack identification method in beam-like structures using changes in experimentally measured frequencies and Particle Swarm Optimization.* Comptes Rendus Mécanique, 2018. **346**(2): p. 110-120
33. Liao, J., et al., *Service Composition Based on Niching Particle Swarm Optimization in Service Overlay Networks.* KSII Transactions on Internet & Information Systems, 2012. **6**(4): p. 1106-1127
34. Wu, Q., et al., *Applications of particle swarm optimization in the railway domain.* International Journal of Rail Transportation, 2016. **4**(3): p. 167-190

35. Pau, G., et al., *Bluetooth 5 energy management through a fuzzy-pso solution for mobile devices of internet of things.* Energies, 2017. **10**(7): p. 992
36. Khatir, S., et al., *Damage detection and localization in composite beam structures based on vibration analysis.* Mechanics, 2015. **21**(6): p. 472-479
37. Antonelli, A., et al. *Enhanced PLL system for harmonic analysis through genetic algorithm application*. in *11th International Conference on Environment and Electrical Engineering*. 2012. IEEE. p. 328-333
38. Soodi, H.A., et al., *STATCOM Estimation Using Back-Propagation, PSO, Shuffled Frog Leap Algorithm, and Genetic Algorithm Based Neural Networks.* Computational intelligence and neuroscience, 2018. **2018**: Article ID. 6381610
39. Suresh, A., et al., *Particle swarm optimization over back propagation neural network for length of stay prediction.* Procedia Computer Science, 2015. **46**: p. 268-275
40. Sun, W., et al., *Using a back propagation neural network based on improved particle swarm optimization to study the influential factors of carbon dioxide emissions in Hebei Province, China.* Journal of cleaner production, 2016. **112**: p. 1282-1291
41. Vapnik, V., *The nature of statistical learning theory*. 1995: Springer-Verlag New York, Inc.
42. Zhang, M., et al., *Classification of power quality disturbances using wavelet packet energy and multiclass support vector machine.* COMPEL-The international journal for computation and mathematics in electrical and electronic engineering, 2012. **31**(2): p. 424-442