Research Article

Predicting Tooth Surface Loss Using Genetic Algorithms-Optimized Artificial Neural Networks

Ali Al Haidan, Osama Abu-Hammad, and Najla Dar-Odeh

College of Dentistry, Taibah University, Al Madina Al Munawara, Saudi Arabia

Correspondence should be addressed to Najla Dar-Odeh; najla_dar_odeh@yahoo.com

Our aim was to predict tooth surface loss in individuals without the need to conduct clinical examinations. Artificial neural networks (ANNs) were used to construct a mathematical model. Input data consisted of age, smoker status, type of tooth brush, brushing, and consumption of pickled food, fizzy drinks, orange, apple, lemon, and dried seeds. Output data were the sum of tooth surface loss scores for selected teeth. The optimized constructed ANN consisted of 2-layer network with 15 neurons in the first layer and one neuron in the second layer. The data of 46 subjects were used to build the model, while the data of 15 subjects were used to test the model. Accepting an error of ±5 scores for all chosen teeth, the accuracy of the network becomes more than 80%. In conclusion, this study shows that modeling tooth surface loss using ANNs is possible and can be achieved with a high degree of accuracy.

1. Introduction

Tooth surface loss (TSL) is a universal problem that involves an irreversible, multifactorial, noncarious, physiologic, pathologic, or functional loss of dental hard tissues [1].

It is difficult to estimate the precise prevalence of tooth surface loss, because of the differences in assessment criteria. Also, the data on the prevalence are inconsistent [2, 3]. Oginni and Olusile reported that 64.28% of the patients attending a dental hospital had attrition, 15.87% had abrasion, 7.14% had erosion, and 12.69% had attrition and abrasion combined [4]. Al-Zarea reported the prevalence of tooth surface loss among a sample of 400 patients [2]. The results showed that 300 (75%) had attrition, 360 (90%) had erosion, 60 (15%) had abrasion, and about 95% had more than one type of tooth surface loss. Wang et al. carried out a survey on a sample of 12-13-year-old children from 10 schools and found that 27.3% of his sample had a tooth surface loss [5]. The most affected teeth were the upper and lower central incisors [5].

Many studies reported that tooth surface loss severity increases with age. Van'T Spijker et al. reported that tooth surface loss in adults increases from 3% of 20 years old to 17% of 70 years old [6]. Also, a study by Dugmore and Rock found that 59.7% of 12-year-old children had tooth surface loss [7].

In the same sample, 2.7% had exposed dentine but in a sample of 14-year-old children this percentage increased to 8.9%.

TSL is known to be multifactorial. Diet, biting on foreign objects, bruxism, parafunctional activity, environment, occupation, medicaments, gastrointestinal problems, and acid regurgitation are among the aetiological factors that lead to TSL [2]. Al-Zarea attributed TSL mainly to diet where 78% of TSL was caused by acidic food and drinks, followed by parafunctional habits (70%) and unilateral chewing (50%). Occupational risks and intake of acidic medications were the least common factors (0.5% each) [2]. Tantbirojn et al. reported that people with gastroesophageal reflux disease had higher scores for tooth surface loss [8].

There are many indices to measure TSL, and these can be either quantitative or qualitative. Quantitative indices identify the physical measurements, for example, depth of groove, area of facet, or height of crown. Qualitative indices consider the clinical descriptions (mild, moderate, and severe) [9]. Eccles index, a qualitative index, originally classified lesions as early, small, and advanced. Later, it was expanded to score the severity and site of erosion [9].

Smith and Knight Index was the first index that scored multifactorial tooth loss (0-No loss of enamel surface, 1-loss of enamel surface, 2-loss of enamel exposing dentine for less

than one-third of surface, 3-loss of enamel exposing dentine for more than one-third of surface, and 4-complete enamel loss-pulp exposure) [9].

Artificial neural network (ANN) is a data-driven tool for analyzing and modeling complex relationships. ANN-based models are where the data do not have to fit predefined conditions (e.g., linearity and normal distribution); instead, available data are used in a training phase to develop experimental models [10]. This method has been used in many applications where the expert knowledge is not clearly defined, and it is based on the adjustment of weights between the neurons for any input-output function approximation [11].

ANN was originally used in medicine to investigate the hidden relationship of a number of factors to different diseases and it was found to have relatively high accuracy [12, 13].

The aim of this study was to construct an ANN mathematical model capable of predicting TSL scores in individuals without the need to conduct clinical examination.

2. Subjects and Methods

This study was conducted at the Dental Clinics of Taibah University, Saudi Arabia. A convenient sample of TSL male patients was recruited by simply asking patients attending one intern doctor clinic to participate in the study. Patients accepting were asked to sign a consent form.

Sample size was decided to be more than 30 cases for the buildup of the network and not less than 10 for the testing of the network.

Ethical approval was obtained from the Scientific Research Ethics Committee at the College of Dentistry, Taibah University.

Before data collection, examiner (AA) explained to the participants the aim of the study and guarantee of confidentiality, and consent was obtained before being recruited into the study.

All residents of Al Madinah patients who are having a full set of natural teeth (except for the wisdom teeth) were included into the study regardless of age. Participants who received treatment for TSL were excluded from the study. Patients were then asked to complete a questionnaire and then they were examined for the severity of TSL.

2.1. The Questionnaire. A questionnaire was prepared to collect the data from the study sample. It included the following:

(i) personal data of name, gender, age, education, and occupation;

(ii) review of medical history including gastrointestinal problems and history of vomiting;

(iii) review of parafunctional habits such as clenching, foreign body biting, and bruxism;

(iv) oral hygiene habits like tooth brushing;

(v) information on diet such as fizzy drinks, fruits (orange, apple, or lemon), pickles, and dried seeds consumption;

(vi) history of tobacco smoking.

2.2. Clinical Examination. Clinical examination was carried out to note TSL score for each exposed surface of teeth of the dentition up to the second molar teeth in both jaws.

In this study, the tooth surface loss index used was that of Smith and Knight (0-No loss of enamel surface, 1-loss of enamel surface, 2-loss of enamel exposing dentine for less than one-third of surface, 3-loss of enamel exposing dentine for more than one-third of surface, and 4-complete enamel loss-pulp exposure) [14].

Prior to the conduction of this study, the first ten patients were reinterviewed and examined one week after their first evaluation. The results showed 100% coincidence between first and second evaluations.

2.3. Data Selected for Analysis. Scores for TSL were variable for different teeth and different patients. TSL score for each patient was calculated as the sum of recorded TSL scores of the different surfaces of his dentition.

Data that were chosen as input data were 11: age, smoking status, type of tooth brush, and frequency of the following: tooth brushing, bruxism, eating oranges, apples, lemon, pickles, or dried seeds, and drinking carbonated fizzy drinks.

Date for the age, and all other data, were fed into Pythia as such. The software has the capability to normalize data. After normalization, the smallest age was referred to as 0 and the maximal age was 1. All other ages were expressed as the corresponding fraction of the maximum. Smoker status was referred to as 0 when not smoking, whereas a smoker was recorded as 1.

Frequency of tooth brushing, bruxism, eating or drinking oranges, apples, lemon, pickles, dried-roasted seeds, and fizzy drinks was calculated as the number of occurrences per month, that is, discrete data, except for dried seeds which was calculated as the approximate number of minutes of eating this type of foodstuff per month (i.e., continuous scale data). In this study there was only one output: TSL scores. Output was fed as the sum of TSL scores for all examined teeth for a given subject (which could take any discrete value between 0 and 336).

It was also decided to use the data of the first 80 subjects to build up and train the ANN and keep the remainder of the data (remaining 15 subjects) for testing the network.

Responses with more than 2 nominal pieces of data (such as occupation and place of birth) could not be used as input data for ANNs. These data were discarded.

2.4. Buildup of the Networks. The ANN software was run on a PC using x64-based Intel Core i7-2630QM CPU @ 2.00 GHz processor, with 8 GB of RAM, running 32-bit Windows 8.1 pro © 2013 Microsoft Corporation. The ANN software (Pythia ©, Runtime Software, 2000) used input and output data for the first 81 subjects to construct and optimize an appropriate neural network that reproduced
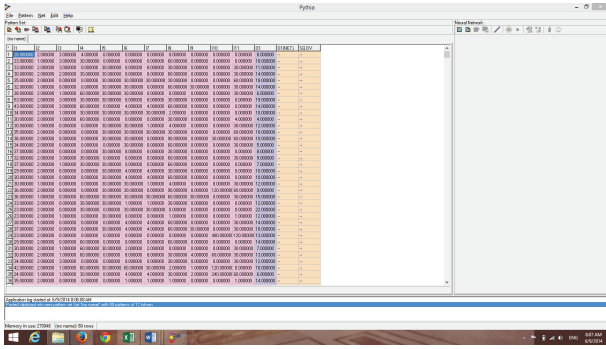
FIGURE 1: The graphical user interface (GUI) of the software after importing input and output data for the first group of patients (80 patients) prior to construction of the ANN.
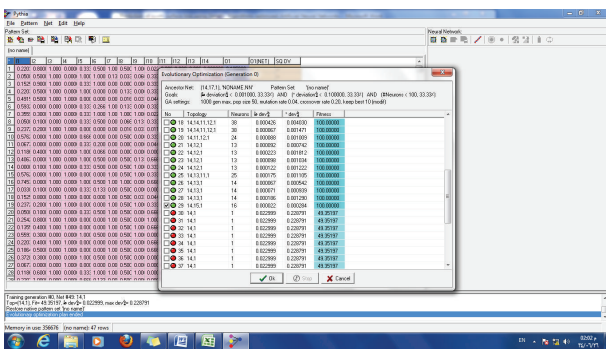

FIGURE 2: The optimizer while trying to find the appropriate ANN. All green marks point to network configurations that are capable of reproducing data with 100% accuracy.

FIGURE 3: Actual output (O1) and net output (O1 (NET)) with square deviations (SQ DV). The network is evident at the right-hand side of the GUI of the software.

FIGURE 4: The network was a two-layer network with 12 neurons in the first layer and only one neuron in the second layer.

output data (scores of TSL) with high degree of accuracy approaching 100%. The network was further trained to yield higher accuracy. Figures 1–6 show the process of network buildup step by step.

Figure 1 shows the graphical user interface (GUI) of the software after importing the data for the first 80 patients (data not normalized view).

2.5. The Use of Pythia to Construct an ANN. Pythia has the capability to employ genetic algorithms to choose the characteristics of an ANN that would be able to reproduce the output data for the first group of subjects that was used to build it, that is, to choose an ANN with best fit to given data.

Figure 2 shows the genetic algorithms optimizer while trying to find appropriate ANNs that are capable of reproducing output data of TSL scores.

The network with the least reported deviations was chosen. This network was a two-layer network with 12 neurons in the first layer and only one neuron in the second layer.

After the network was constructed, input data were reproduced by the network. The deviations were small, but not negligible. Figure 3 shows the output data, predicted output data, and square deviations. It is noted that the deviations are small.
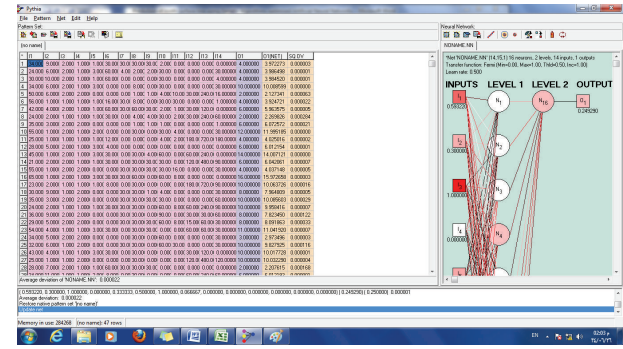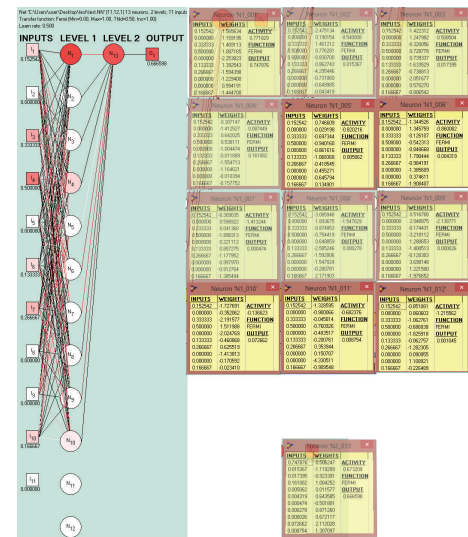
As differences were not negligible, the network was trained to further bring down the deviations of predicted TSL scores. The network was trained in 1000 cycles for 16 times (16000 training cycles in total). Each time the training finishes, the net output was noted.

Training the ANN will result in more accurate predictions of TSL scores (net output) and fewer deviations.

The network was trained a total of 50000 times.

Figure 4 shows the constructed network with its synaptic weights. The network was a two-layer network with 12 neurons in the first layer and only one neuron in the second layer.

2.6. Reproduction of the Output Using the Selected ANN before Training and after Training. As differences were not negligible, the network was trained to further bring down the deviations or predicted TSL scores.

2.7. Testing the ANN. The optimized trained network was ready at this stage for testing.
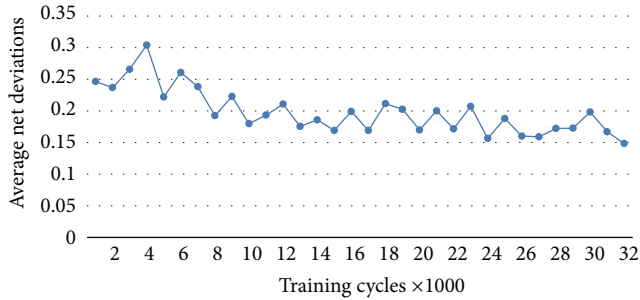
FIGURE 5: Sums of deviations: before training, then at 1000 consecutive training cycles till 30000, and then at 50000 training cycles. Deviations decline with more training.

The data for the last 15 patients were used to test the network by exposing to the network input data for a given subject; then, the network would predict the TSL score for that subject.

2.8. Using Neural Networks Function of SPSS Statistical Analysis Software to Categorize Importance of Input Variables. As Pythia could not provide information on the importance of the different variables in contributing the predictions, SPSS (IBM (International Business Machines Corp.) SPSS Statistics, version 20.0.0) was used for this purpose.

All data of the 96 subjects were fed into the software and a multilayer perceptron (network) was designed with the same topology of that created in Pythia in the initial phase of the analysis. All default settings were accepted as such in dialogue tables of the analysis and the analysis was conducted.

3. Results

The total number of subjects agreeing to participate into this study was 96 subjects. Twelve patients did not want to sign a consent form and were dropped from the study. All patients were males with a mean age of 35 and an age range of 16–80. Forty-three patients indicated they were smokers while 53 stated they were not. Thirty-one subjects did not brush their teeth, 42 used Miswak (wooden sticks from Silvadora Persica plant), and 5 used soft tooth brush, while 18 used hard tooth brush. Frequency of brushing teeth ranged from 0 to twice daily. While many patients did not report bruxism ($n = 52$), some patients complained of it, and some of them indicated that they experience it daily ($n = 43$). Twenty-one patients did not drink or eat oranges, but some of them ($n = 3$) indicated that they eat 2 oranges/day. Forty-four patients did not eat apples

Data for the first 81 subjects were used to construct and optimize (train) the network; data of the other 15 subjects were used to test the network.

The genetic algorithms optimizer included in the software chose a network with a buildup that could reproduce the data of the first 46 subjects with almost perfect manner. Training the ANN, however, has resulted in more accurate predictions of the already known TSL scores (net output). This is shown in Table 1.

TABLE 1: Original outputs (sum of TSL scores) for the first 80 subjects, network output before training. Deviations of network output and original output are also displayed before and after training.

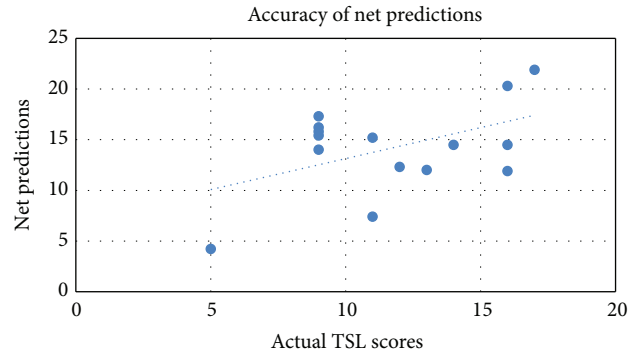| Subject	Sum of actual outputs in subjects	Sum of net outputs before training	Sum of deviations before training	Sum of deviations after 50000 cycles
80	935	935.078	0.247	0.148

FIGURE 6: Comparison between net predictions with actual TSL scores for the 15 test subjects and a projected trend line for the predictions.

Figure 5 also shows how deviations were reduced with more training of the network.

The network was trained a total of 16000 times.

Predicted TSL scores for all 15 subjects are shown in Table 2 and Figure 6.

Testing the ANN showed that the network was capable of predicting unknown scores with near 100% degree of accuracy for 2 subjects only out of the 15 (13.3% of subjects).

However, accepting an error of ±5 scores for TSL gives the network an accuracy of 73.3%. Moreover, accepting an error of ±6.3 scores for the TSL produces a network accuracy of 100%.

As for each subject TSL scores were registered for 3 surfaces of his 28 teeth with the maximum possible score being 4, maximal TSL scores on this scale for any given subject allow for 336. Considering this, accepting results of ±5 scores (i.e., error of 10 scores with an accuracy of 97% of the TSL score value) will make 73.3% of the predictions accurate. And when accepting results of ±10 scores (i.e., error of 20 scores with an accuracy of 94% of the TSL score value), it will make 100% of the predictions accurate.

The use of SPSS neural networks allowed for the classification of input variables according to importance. Table 3 shows this classification.

4. Discussion

Tooth surface change can be measured in a variety of ways; however, no single technique provides a comprehensive

TABLE 2: Actual TSL scores and ANN predictions for those scores of the test (number = 15) subjects.

Subject number	TSL score	Net prediction	Difference (residual)	Accurate prediction rounded to nearest integer	Accurate predictions rounded to nearest 5	Accurate predictions when rounded to nearest 8.3
81	13	12.01	0.99		Yes	Yes
82	16	11.9	4.1		Yes	Yes
83	11	7.4	3.6		Yes	Yes
84	5	4.2	0.8		Yes	Yes
85	9	15.4	−6.4			Yes
86	9	17.3	−8.3			Yes
87	16	20.3	−4.3		Yes	Yes
88	9	16.2	−7.2			Yes
89	14	14.49	−0.49	Yes	Yes	Yes
90	16	14.49	1.51		Yes	Yes
91	9	14	−5		Yes	Yes
92	12	12.3	−0.3	Yes	Yes	Yes
93	9	15.8	−6.8			Yes
94	17	21.9	−4.9		Yes	Yes
95	11	15.2	−4.2		Yes	Yes
Overall accuracy of predictions				13.33%	73.33%	100%

TABLE 3: Independent variable importance as calculated by SPSS neural networks module.

Independent variable importance		
	Importance (or weight in the ANN)	Normalized importance
Age	0.160	100.0%
Smoker status	0.029	18.0%
Type of brush	0.074	46.2%
Frequency of brushing/month	0.105	65.5%
Frequency of bruxism/month	0.027	16.7%
Frequency of oranges eating/month	0.127	79.1%
Frequency of apples eating/month	0.078	48.7%
Frequency of lemon eating/month	0.073	45.8%
Frequency of pickle eating/month	0.136	84.8%
Minutes of seeds eating/month	0.075	46.7%
Frequency of fizzy drinks/month	0.118	73.8%

assessment of the remaining tooth surface, and each technique suffers from its own limitations [15].

This study is the first to utilize ANN for the prediction of this rather prevalent problem of TSL. The software used in this study is Pythia, and it is software that is capable of optimizing artificial neural networks using genetic algorithms, building artificial neural networks, and operating them. The software allows the creation of the ANN and its testing using other data.

It has been noticed that the more the network is trained on the supplied set of data, the more accurate it becomes [16]. As far as the prediction of unknown data is concerned, the accuracy of the ANNs reached was 73.3%. This in itself has a clinical value. Another advantage of this method besides accuracy is saving time for dentists since evaluating TSL may be a time-consuming practice.

Sample size determination in ANN-based research is not consistent with epidemiologic or similar research. Sample size was decided to be 46 cases for the buildup of the network and 15 for the testing of it. It is well established that these numbers are within what is acceptable for the buildup and testing of the ANNs.

But still the numbers for subjects (total $n = 61$) for network construction and testing remain a limitation. It is also well established that increasing these numbers will yield a more accurate network and more reliable predictions. Another limitation was that all cases were of the male gender. We recruited our patients from the dental hospital, and this was exclusively a male hospital. The female section of the college was not included in the study as it was at the other side of the town. Regardless of the level of accuracy of the network, it is applicable only to males not females. Females
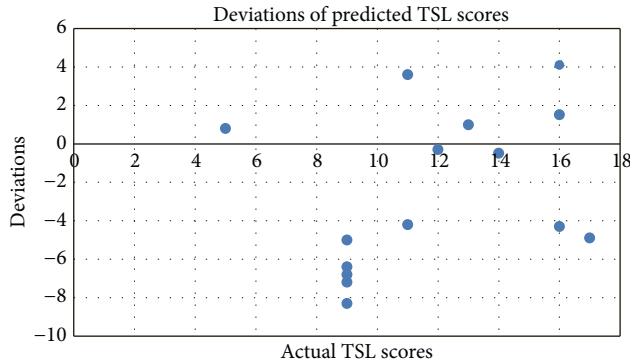
FIGURE 7: Deviations of the predicted TSL score values for the 15 test subjects.

have different factors that can influence TSL such as the number of previous pregnancies and the associated morning sickness and vomiting episodes. Vomiting was found to be a significant factor in TSL in many other studies. In this study, however, none of our patients complained of vomiting or gastric reflux; hence, it was excluded as a possible factor.

In a previous study that used conventional statistics, factors influencing TSL were concluded. The same factors were used in this study [17].

Results show that about 73.3% of the predicted TSL scores of the testing sample deviate from the actual scores by ±5 scores. Accepting this ±5 error in TSL scores is so small considering the maximal possible TSL scores of "336" for any given subject (i.e., the error margin being calculated as 10/336 × 100% = 3%). Moreover, the results show that all predicted TSL scores lie within ±8.3 TSL scores of the actual recordings (i.e., error here becomes 16.6/336 × 100% = 4.9%), Figure 7.

Although 73.3% is an acceptable level of accuracy, other studies predicting other outcomes in other biologic systems using ANNs achieved 90% accuracy [16]. Perhaps increasing the number of subjects with representative TSL will bring predictions closer to this value or perhaps such a high margin of accuracy could be achieved as there were better associations between predicting variables and the output.

In a previous study [17], a number of factors were found to be of significant importance to TSL while others were found to have less effect. Vomiting, brushing, bruxism, and consumption of dried-roasted seeds were found to exert a significant effect on TSL, whereas fizzy drinks, orange consumption, grape eating, apple consumption, pickle consumption, alcohol, lemon eating, and smoking status were found to have an effect that was not found to be significant. The idea of using TSL was to find patterns of association between given variables and an outcome or outcomes. So it was decided to include as much as possible of these variables. In this study vomiting was excluded as no one reported vomiting.

Table 3 displayed input factors according to importance. Age was the most important factor associated with TSL, followed, respectively, by pickles, oranges, fizzy drinks, frequency of brushing, apples, dried seeds, type of brush, and lemons. Smoker status and finally bruxism were the least

important variables. This bit of information displayed in Table 3 (i.e., the classification of input variables according to importance) is not directly related to the theme and aim of the study which is the buildup of a mathematical model that is capable of predicting TSL. However, some researchers, particularly epidemiologists, will be more interested in it.

5. Conclusions

Construction of mathematical models using artificial neural networks to predict TSL is possible and can result in reasonable degree of accuracy. The most important factors for the prediction of TSL were found to be age, pickle consumption, orange consumption, drinking fizzy drinks, and brushing frequency. The least important factors were found to be smoking status and bruxism.

6. Recommendations

It is recommended to use larger number of patients for the construction phase of these studies in order to yield higher accuracy in predictions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. Kelleher and K. Bishop, "Tooth surface loss: an overview," *British Dental Journal*, vol. 186, no. 2, pp. 61–66, 1999.

[2] B. K. Al-Zarea, "Tooth surface loss and associated risk factors in northern Saudi Arabia," *ISRN Dentistry*, vol. 2012, Article ID 161565, 5 pages, 2012.

[3] C. M. Kreulen, A. Van't Spijker, J. M. Rodriguez, E. M. Bronkhorst, N. H. J. Creugers, and D. W. Bartlett, "Systematic review of the prevalence of tooth wear in children and adolescents," *Caries Research*, vol. 44, no. 2, pp. 151–159, 2010.

[4] O. Oginni and A. O. Olusile, "The prevalence, aetiology and clinical appearance of tooth wear: the Nigerian experience," *International Dental Journal*, vol. 52, no. 4, pp. 268–272, 2002.

[5] P. Wang, H. C. Lin, J. H. Chen, and H. Y. Liang, "The prevalence of dental erosion and associated risk factors in 12-13-year-old school children in Southern China," *BMC Public Health*, vol. 10, article 478, 2010.

[6] A. Van'T Spijker, J. M. Rodriguez, C. M. Kreulen, E. M. Bronkhorst, D. W. Bartlett, and N. H. J. Creugers, "Prevalence of tooth wear in adults," *The International Journal of Prosthodontics*, vol. 22, no. 1, pp. 35–42, 2009.

[7] C. R. Dugmore and W. P. Rock, "The prevalence of tooth erosion in 12-year-old children," *British Dental Journal*, vol. 196, no. 5, pp. 279–282, 2004.

[8] D. Tantbirojn, M. R. Pintado, A. Versluis, C. Dunn, and R. Delong, "Quantitative analysis of tooth surface loss associated with gastroesophageal reflux disease: a longitudinal clinical study," *Journal of the American Dental Association*, vol. 143, no. 3, pp. 278–285, 2012.

[9] P. F. Bardsley, "The evolution of tooth wear indices," *Clinical Oral Investigations*, vol. 12, no. 1, Supplement, pp. 15–19, 2008.

[10] C. Winter, J. P. Payet, and C. A. Suttle, "Modeling the winter-to-summer transition of prokaryotic and viral abundance in the arctic ocean," *PLoS ONE*, vol. 7, no. 12, Article ID e52794, 2012.

[11] N. Gargouri, A. D. Masmoudi, D. S. Masmoudi, and R. Abid, "A new GLLD operator for mass detection in digital mammograms," *International Journal of Biomedical Imaging*, vol. 2012, Article ID 765649, 2012.

[12] J. S. Allison, J. Heo, and A. E. Iskandrian, "Artificial neural network modeling of stress single-photon emission computed tomographic imaging for detecting extensive coronary artery disease," *The American Journal of Cardiology*, vol. 95, no. 2, pp. 178–181, 2005.

[13] A. Bartosch-Härlid, B. Andersson, U. Aho, J. Nilsson, and R. Andersson, "Artificial neural networks in pancreatic disease," *The British Journal of Surgery*, vol. 95, no. 7, pp. 817–826, 2008.

[14] B. G. Smith and J. K. Knight, "An index for measuring the wear of teeth," *British Dental Journal*, vol. 156, no. 12, pp. 435–438, 1984.

[15] J. Field, P. Waterhouse, and M. German, "Quantifying and qualifying surface changes on dental hard tissues in vitro," *Journal of Dentistry*, vol. 38, no. 3, pp. 182–190, 2010.

[16] N. S. Dar-Odeh, O. M. Alsmadi, F. Bakri et al., "Predicting recurrent aphthous ulceration using genetic algorithms-optimized neural networks," *Advances and Applications in Bioinformatics and Chemistry*, vol. 3, no. 1, pp. 7–13, 2010.

[17] R. A. B. M. Al-Habahbeh, S. Al-Jutaily, N. Dar-Odeh, and O. Abu-Hammad, "Effect of dried roasted seeds on tooth surface loss," *Pakistan Oral & Dental Journal*, vol. 31, no. 2, pp. 439–442, 2011.