Detection of Caries in Panoramic Dental X-ray Images using Back-Propagation Neural Network

Ainass A. Alhahbah Hazem M. El-Bakry Sameh Abd-Elghany
Department of Information Systems, Department of Information Systems, Department of Information Systems,
Faculty of Computer and Information Faculty of Computer and Information Faculty of Computer and Information
Sciences Sciences Sciences
Mansoura University, Mansoura University, Mansoura University,
Mansoura, Egypt Mansoura, Egypt Mansoura, Egypt

Abstract – Recently, artificial neural network (ANNs) has been adopted widely for solving many complex problems in different fields due to its high performance and its ability to generalize. One of these fields is the medical image processing for diagnosing purposes. In this paper, tooth caries detection strategy is introduced based on a back propagation (BP) neural network for analyzing the dental X-ray images. The neural network used the inter-pixel autocorrelation as input features. The accuracy of classification is satisfactory where the tooth caries detection is clearly improved when compared to the diagnosing process performed by a rule-based computer assisted program and a group of doctors.

Keywords – Component; Tooth Decay; ANNs; Histograms of Oriented Gradients (HOG).

I. INTRODUCTION

In clinical dentistry, the early detection of tooth decay (caries) is important. This is increasingly because recent dental research has shown that caries which is confined to the dental enamel may be repaired biologically by topical fluoride. The two clinical methods of caries detection in use at present are radiography, and clinical examination by mirror and probe. Radiography cannot reliably distinguish caries which is confined to the enamel. Clinical examination is subjective and varies according to extend factors such as lighting and the experience of the dentist. In addition, if the probe is pushed too hard, it may damage the relatively softer surface layer of the early carious lesion and prevent its future repair [1]. Thus, present clinical methods of caries detection are not satisfactory at the detection of very early carious enamel.

During the past century, the nature of dental decay or dental caries has been changed markedly due to the introduction of fluoride to the drinking water, the use of fluoride dentifrices and rinses, application of fluoride topical in the dental office and improved dental hygiene [1]. In spite of these advances, dental decay continues to be the leading cause of tooth loss. The nature of the caries problem has changed dramatically with the majority of newly discovered caries lesions being highly localized to the occlusal pits and fissures of the posterior dentition and the interproximal contact sites between teeth as shown in Fig.1 [2].

These early carious lesions are often obscured or “hidden” in the complex and convoluted topography of the pits and fissures or are concealed by debris that frequently accumulates in those regions of the posterior teeth. Moreover, such lesions are difficult to detect in the early stages of development [3].

By definition, early caries lesions are those lesions confined to the enamel and have not yet penetrated into the inner dentin. In the caries process demineralization occurs as organic acids generated by bacterial plaque diffuse through the porous enamel of the tooth dissolving the mineral. If the decay process is not arrested, the demineralization spreads through the enamel and reaches the dentin where it rapidly accelerates due to the markedly higher solubility and permeability of dentin. The lesion spreads throughout the underlying dentin to encompass a large area, resulting in loss of integrity of the tissue and cavitation [4].

Caries lesions are usually not detected until after the lesions have progressed to the point which surgical intervention and restoration are necessary, often resulting in the loss of healthy tissue structure and weakening of the tooth. The caries screening and treatment paradigms that were developed in the past, based on radiography, for example, are adequate for large, cavitated lesions; however they do not have sufficient sensitivity or specificity for the detection of early non-cavitated caries, particularly in the early stages. Therefore, new imaging and detection technologies are needed for the early detection of such lesions. Moreover, the treatment for early dental decay or caries is shifting away from aggressive cavity preparations that attempt to completely remove demineralized tooth structure toward non-surgical or minimally invasive restorative techniques. In non-surgical therapy, a clinician prescribes antibacterial rinses, fluoride treatments, and dietary changes to arrest demineralization and facilitate remineralization of the caries lesion before it becomes irreversible [5, 6]. The success of this type of therapy is contingent on early caries detection and also requires imaging modalities that do not require ionizing radiation that can be used safely and accurately to monitor the success of such treatment [7].

A. Pre-processing

In this phase segmentation of the Jaws area is done after resizing to extract the region of interest and avoid extraction of noisy and background areas and increasing the performance evaluation of feature extraction. Threshold is applied to the image by assigning the Jaws pixels to “1” for decayed and all the remaining pixels to “0” so the resulting binary image is suitable for feature extraction.
B. Processing
Extraction features is done by Histograms of Oriented Gradients (HOG) descriptors provide excellent performance relative to other existing feature sets including wavelets.

C. Post Processing
Classifying image of jaws to be decay or not decayed using ANNs labeled by “0” to be decayed or “1” otherwise.

II. BACKGROUND

In the following subsections the necessary concepts and background are introduced briefly.

Adaptive Binarization
Jaws Image Binarization transforms the 8-bit Gray image to a 1-bit image with 0-value for decayed teeth and 1-value otherwise. A locally adaptive binarization method is performed to binarize the jaw image. This method is done by transforming a pixel value to 1 if the value is larger than the threshold value to labeled ad not decayed and transforming a pixel value to 0 if the value is smaller than the threshold value to labeled ad decayed [8].

Segmentation and Extraction the Region of Interest
Extracting the Region of Interest (ROI) is useful to be accurate and save time for each jaw image. First discard the image area without effective information and as it only holds background information. The Crop Image tool is used here. It is a moveable, resizable rectangle that can be position over the image and perform the crop operation interactively [9].

Neural Network
ANNs is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a weight connection between each other. In most cases ANNs is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.

Neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. [10].

One of the most popular NN algorithms is back propagation algorithm [18-21]. As claimed in [11] that BP algorithm could be broken down to four main steps. After choosing the weights of the network randomly, the back propagation algorithm is used to compute the necessary corrections. The algorithm can be decomposed in the following four steps [11]:

i. Feed-forward computation.
ii. Back propagation to the output layer.
iii. Back propagation to the hidden layer.
iv. Weight updates.

The algorithm is stopped when the value of the error function has become sufficiently small.

This is very rough and basic formula for BP algorithm. There are some variations proposed by other scientist but Rojas definition seem to be quite accurate and easy to follow. The last step, weight updates is happening throughout the algorithm [11].

Histogram of Oriented Gradient
The histogram of oriented gradients (HOG) is a component descriptor utilized as a part of PC vision and picture handling with the end goal of item identification. The strategy includes events of slope introduction restricted parts of a picture. This technique is like that of edge introduction histograms, scale-invariant component change descriptors, and shape connections, however varies in that it is processed on a thick network of consistently separated cells and utilizes covering nearby contrast standardization for enhanced exactness [12].

The basic idea is that local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions, even without precise knowledge of the corresponding gradient or edge a measure of nearby histogram "vitality" over to some degree bigger spatial areas (squares) and utilizing the outcomes to standardize the greater part of the cells in the piece. We will allude to the standardized descriptor hinderers as Histogram of Oriented Gradient (HOG) descriptors. Tiling the identification window with a thick (truth be told, overlapping) framework of HOG descriptors and utilizing the joined element vector as a part of an ordinary SVM based window classifier gives our human discovery chain [13].

The Swine descriptor has a couple key focal points over different descriptors. Since it works on nearby cells, it is invariant to geometric and photometric changes, aside from item introduction. Such changes would just show up in bigger spatial locales. In addition, as Dalal and Triggs found, coarse spatial inspecting, fine introduction testing, and solid neighborhood photometric standardization allows the individual body development of people on foot to be disregarded insofar as they keep up a generally upright position. The HOG descriptor is consequently especially suited for human identification in pictures. Swine comprise of four stages as shown in Fig.2.

The first step of calculation in many feature detectors in image pre-processing is to ensure normalized color and gamma values (Gradient computation). The most common method is to apply the 1-D centered, point discrete derivative mask in one or both of the horizontal and vertical directions [12, 13]:

The second step of estimation is making the cell histograms (Orientation binning). Every pixel inside of the cell makes a weighted choice for an introduction construct histogram channel based with respect to the qualities found in the inclination calculation. The cells themselves can either be rectangular or outspread fit as a fiddle, and the histogram channels are uniformly spread more than 0 to 180 degrees or 0 to 360 degrees, contingent upon whether the slope is "unsigned" or "signed". Dalal and Triggs found that unsigned inclinations utilized as a part of conjunction with 9 histogram directs performed best in their human recognition tests. Concerning the vote weight, pixel commitment can either be the inclination extent itself, or
some capacity of the greatness. In tests, the inclination
greatness itself for the most part delivers the best results.
Different alternatives for the vote weight could
incorporate the square root or square of the slope
greatness, or some cut form of the size [14, 15].
The third step is to represent changes in light and
contrast (Descriptor obstructs), the slope qualities must be
privately standardized, which requires gathering the
phones together into bigger, spatially joined squares. The
HOG descriptor is then the linked vector of the parts of
the standardized cell histograms from the greater part of
the square areas. These pieces commonly cover, implying
that every cell contributes more than once to the last
descriptor. Two fundamental piece geometries exist:
rectangular R-HOG squares and round C-HOG squares.
R-HOG pieces are for the most part square networks,
spoke to by three parameters: the quantity of cells per
hinder, the quantity of pixels per cell, and the quantity of
channels per cell histogram.
The fourth step is block normalization. There are
different methods for block normalization. Let \( v \) be non-
normalized vector containing all histogram in given block,
\( v_k \) be its k-norm for k= 1, 2 and e be some small constant
(whose value will not influence the results). Then the
normalization is applied [12, 16].

### III. THE PROPOSED METHOD

The purpose of the proposed model is to achieve higher
performance that may not be possible from previous researcher
to detect tooth decay (caries). The phases of the
proposed system are shown in Fig.3. Here proposed
method start to resize the medical image of tooth to be
appropriate for enhancing and extracting feature in simple
and good way next figure explain the overall process in
the proposed method.

Segmentation and region of interest this step is execute
the upper and lower jaw and tooth, the input images of
this stage are the images jaw after resize. Where the
problem of the input Images is that they only contains jaw
and tooth but in the majority of the cases all images
contains extra background corresponding.

Cropping method is applied here to extract region of
interest. Now after converting image to RGB, threshold is
applied to the image by assigning the Jaws pixels to “1”
and all the remaining pixels to “0” so the resulting binary
image is suitable for feature extraction. Preprocessing
increased the contrast of the image, highlighting the dark
areas and the bright areas of the original image. In other
case the preprocessing in teeth that were present in the
original image that already don’t contrasts with the rest of
the image tends to disappear with the implementation of
preprocessing. This occurs only in some due to the poor
quality of the input image that disables the good
effectiveness of the preprocessing. After binaries the image is to extract feature of tooth decay (caries) histogram of oriented gradient (HOG) have been used because Local object appearance and shape can often be
categorized rather well by the distribution of local
intensity gradients or edge directions, better invariance to
illumination and shading. It is useful to contrast-
normalize the local responses and Accumulate local
histogram “energy” over a larger regions (“blocks”).

A three layers network topology was adopted with
sigmoid function as its transfer function. The input layer
consists of a group of processing units that are responsible
for acceptance of data imported to the network. Input data
are normalized autocorrelation coefficients that well
reflect the inter-pixel correlation within a decayed tooth
image. A hidden layer connects the input layer to the
output layer. There is one nodal point in the output layer
telling decayed or not labeled by 0 or 1.

### IV. EXPERIMENTAL RESULTS

The proposed framework was implanted and developed
using Matlab 2014b x64 bit on Computer system with
specific features. These features were classified as
Hardware including Intel Processor, 2.0 GHz and 64 bit
architecture rather than Windows 7 Home Edition as
Software platform. The proposed system is based on using
the dataset.

ANNs experiments have the following configuration
shown Table I that describes total number of instances
used for evaluation purpose. Where, error percent is
defined as fraction of samples which are misclassified. In
which, 0 value represents no misclassification and 100
represents maximum misclassification.

Precision is another criterion used to measure the
performance of our system. It is a description of a level of
measurement that yields consistent results when repeated.
It is associated with the concept of "random error", a form
of observational error that leads to measurable values
being inconsistent when repeated [9]. The precision
values of different phases are shown in Table II and Fig.4
sequentially.

A receiver operating characteristic (ROC), or ROC
curve, is a graphical plot that illustrates the performance
of a binary classifier system as its discrimination
threshold is varied. The curve is created in Fig.5 by
plotting the true positive rate (TPR) against the false
positive rate (FPR) at various threshold settings. The true-
positive rate is also known as sensitivity,
or recall in machine learning. The false-positive rate is
also known as the fall-out and can be calculated as (1-
specificity) [8].

Another used measure is the cross-entropy (CE) method
is shown in Fig.6. It is a new generic approach to
combinatorial and multi-external optimization and rare
event simulation according CE is computed [17].

### V. CONCLUSION

In this paper we proposed an efficient neural network
tooth decay diagnostic system in which normalized
autocorrelation coefficients were used to differentiate
decayed and normal teeth. The back propagation
algorithm was used to construct the weights for the neural
network classifier in this system. The experimental results
indicated that decayed and normal teeth could be
distinguished accurately by our diagnostic model. The fact that the performances of neural network tooth decay diagnosis system is better than the rule-based computer assisted system means that for the problem as complicated as tooth decay detection “rules” cannot work as well as we expect. Teeth of different patients are quite different depend on teeth position, patient’s age, health condition, living environment and genes. It is hard to find all the factors affecting X-ray image decay behaviors and how these factors affect the behaviors. And it is even harder to conclude to a set of “rules”, which can correctly charge all these situations to make correct diagnostic decision. Artificial neural network’s ability of handling complex situation did a good job here.

REFERENCES


Fig. 1. New caries lesions are found in either contact sites between teeth, interproximal, or in the pits and fissures of occlusal surfaces.
Fig. 2. The different stages of swine descriptor.

Fig. 3. The structure of the proposed model.
Fig. 4. The confusion matrices at the different phases of ANNs

Fig. 5. ROC curves at the different phases of ANNs
Fig. 6. ROC curves at the different phases of ANNs

Table I: The number of samples and the error percent during the different phases of ANNs

<table>
<thead>
<tr>
<th>Phase</th>
<th>#Samples</th>
<th>Error Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>39</td>
<td>17.95%</td>
</tr>
<tr>
<td>Testing</td>
<td>9</td>
<td>33.33%</td>
</tr>
<tr>
<td>Validation</td>
<td>9</td>
<td>55.56%</td>
</tr>
</tbody>
</table>

Table II: The values of precision during the different phases of ANNs

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Training precision</td>
<td>82.1%</td>
</tr>
<tr>
<td>Validation precision</td>
<td>66.7%</td>
</tr>
<tr>
<td>Testing precision</td>
<td>44.4%</td>
</tr>
<tr>
<td>Overall precision</td>
<td>73.7%</td>
</tr>
</tbody>
</table>

Table III: The values of accuracy and specificity

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>64.91</td>
</tr>
<tr>
<td>specificity</td>
<td>64.91</td>
</tr>
</tbody>
</table>