



RESEARCH

Open Access



Machine learning in the detection of dental cyst, tumor, and abscess lesions

Vyshiali Sivaram Kumar^{1†}, Pradeep R. Kumar^{1†}, Pradeep Kumar Yadalam^{2*}, Raghavendra Vamsi Anegundi², Deepti Shrivastava^{3*} , Ahmed Ata Alfurhud^{4,5}, Ibrahem T. Almaktoom⁵, Sultan Abdulkareem Ali Alftaikhah⁵, Ahmed Hamoud L Alsharari⁵ and Kumar Chandan Srivastava^{6,7*} 

Abstract

Background and Objective Dental panoramic radiographs are utilized in computer-aided image analysis, which detects abnormal tissue masses by analyzing the produced image capacity to recognize patterns of intensity fluctuations. This is done to reduce the need for invasive biopsies for arriving to a diagnosis. The aim of the current study was to examine and compare the accuracy of several texture analysis techniques, such as Grey Level Run Length Matrix (GLRLM), Grey Level Co-occurrence Matrix (GLCM), and wavelet analysis in recognizing dental cyst, tumor, and abscess lesions.

Materials & Methods The current retrospective study retrieved a total of 172 dental panoramic radiographs with lesion including dental cysts, tumors, or abscess. Radiographs that failed to meet technical criteria for diagnostic quality (such as significant overlap of teeth, a diffuse image, or distortion) were excluded from the sample. The methodology adopted in the study comprised of five stages. At first, the radiographs are improved, and the area of interest was segmented manually. A variety of feature extraction techniques, such as GLCM, GLRLM, and the wavelet analysis were used to gather information from the area of interest. Later, the lesions were classified as a cyst, tumor, abscess, or using a support vector machine (SVM) classifier. Eventually, the data was transferred into a Microsoft Excel spreadsheet and statistical package for social sciences (SPSS) (version 21) was used to conduct the statistical analysis. Initially descriptive statistics were computed. For inferential analysis, statistical significance was determined by a p value < 0.05 . The sensitivity, specificity, and accuracy were used to find the significant difference between assessed and actual diagnosis.

Results The findings demonstrate that 98% accuracy was achieved using GLCM, 91% accuracy using Wavelet analysis & 95% accuracy using GLRLM in distinguishing between dental cyst, tumor, and abscess lesions. The area under curve (AUC) number indicates that GLCM achieves a high degree of accuracy. The results achieved excellent accuracy (98%) using GLCM.

Conclusion The GLCM features can be used for further research. After improving the performance and training, it can support routine histological diagnosis and can assist the clinicians in arriving at accurate and spontaneous treatment plans.

[†]Vyshiali Sivaram Kumar and Pradeep R. Kumar have contributed equally and should be considered as first author.

*Correspondence:

Pradeep Kumar Yadalam
pradeepkumar.sdc@saveetha.com
Deepti Shrivastava
sdeepti20@gmail.com
Kumar Chandan Srivastava
drkcs.omr@gmail.com

Full list of author information is available at the end of the article



Keywords Dental Images, Artificial Intelligence, Digital Image Processing, GLRLM, GLCM

Introduction

The most traditional way to confirm a dental cyst, tumor, or abscess lesion is by performing an incisional or excisional biopsy [1]. A dental practitioner or a specialist may carry this out this invasive procedure. The sample is then sent to an oral pathologist who examines the tissue under a microscope to check for abnormality. The success of this method depends on how accurately the needle is placed, skill and experience of the oral pathologist [2, 3]. An unsuitable, unrepresentative sample is unproductive, and the patient might have to undergo repeated procedures. A second, unnecessary biopsy may be required if the patient receives inadequate medical care at any stage. It is often time-consuming and is subjected to unacceptable inter and even intra-observer variations [4, 5]. There are various other ways to determine the presence or extent of oral cancer. Radiographs can be especially helpful in assessing the presence and extent of oral cancer as they provide a non-invasive and comprehensive view of the affected areas, aiding in early detection and precise treatment planning. One of the most valuable, standard, and routinely used diagnostic tools are radiographs, which are an adjunct to clinical examination in diagnosing dental diseases [6, 7]. It is a painless and quick method that is the most preferred and reliable [8].

Effective diagnosis and early intervention are crucial in preventing dental complications, with various radiographic modalities playing a key role in this process. Dental abscesses and cysts are common periapical conditions that can affect the maxillary and mandibular jaws [9], and their timely detection and treatment can save patients from considerable discomfort and expense. Clinical classifications help differentiate between these lesions, each with its own distinct causes [10, 11]. Detecting and treating dental issues before they progress to more significant complications may save time, money, and pain. Various dental radiographic modalities such as intraoral periapical, panoramic, and cone beam computed tomographic radiographs play a pivotal role in diagnosis and treatment planning [12–15]. Oral panoramic images play an important role in identification of lesions which related to not only to the oral diseases but also to various systemic disorders. This eventually assist the clinician to plan, treat and make referral for systemic conditions [16, 17]. By analyzing the resulting radiographic images with distinct patterns of intensity fluctuations, computer-aided image analysis may be able to identify aberrant tissue masses. Automatic segmentation

will simplify lesion analysis, making it easier to spot and identify lesions in the clinic. Oral panoramic pictures cannot reliably distinguish between lesions due to their uniformly smooth, spherical or oval borders [18, 19]. To help pathologists in the detection of cysts and tumors, researchers have devised quantitative methodologies for computer-aided diagnosis utilizing dental panoramic radiographs [20, 21].

Oral cancer segmentation isolates oral cancer in a medical image. This can be done manually or automatically with computer vision. Automatic segmentation is more accurate but computationally expensive than manual segmentation, which is time-consuming and imprecise [22–24]. The process begins with the segmentation of the lesion. Rough sets, fuzzy logic & neural networks are examples of soft computation methods that may be used to arrive at solutions that are more intelligent, affordable, and understandable than those reached by more conventional means [25]. By efficiently analyzing and predicting illnesses based on data stored in the database, artificial neural networks provide a useful tool for clinicians. Implementing neural networks on a Field Programmable Gate Array (FPGA) enables more adaptability in software-defined hardware [26]. FPGAs are ideal for real-time applications because of their compact size and fast processing rates [27, 28].

GLCM (Gray Level Co-occurrence Matrix) is a statistical approach that analyses the spatial relationship between image pixels. It builds a matrix showing the frequency of each grey level combination at a given distance and direction. GLRLM (Gray Level Run Length Matrix) is a statistical method that measures the length of consecutive pixels with the same grey level in a picture. It builds a matrix of run length frequencies for each grey level. GPGA (Gray Level Peak Grouping Analysis) is a statistical method that counts and sizes peaks in an image's grey level histogram. It builds a matrix of peak size frequencies for each grey level. Wavelet analysis is a mathematical method that breaks down a signal into wavelets. Each signal wavelet has a particular frequency and scale. Analyzing wavelet coefficients can extract characteristics from photos.

Abnormalities were detected from non-enhanced oral panoramic radiographs using Wavelet Image processing. Image characteristics that could identify tumors, abscesses, and cystic lesions were analyzed using texture analysis. Features of the image were derived from the GLCM, comprising correlation, contrast, homogeneity, & energy [29]. When a wavelet transform was

used on the pictures, the coefficients were calculated [30]. The GLCM & wavelet transform image characteristics were used as input vectors for the neural network training process.

Images were sorted into "normal" and "abnormal" categories based on the network's conclusions. Segmentation based on in-tensity was used to locate the lesion area in the irregular pictures. For more precise diagnosis of oral tumor, cyst, or abscess, an artificial neural network [31] was created on FPGA with the help of a system generator & Xilinx ISE 14.3. The objective of this study was to compare the diagnostic accuracy in detecting dental cysts, abscesses as well as tumor lesions via image processing on panoramic dental images to the actual diagnosis.

Materials and Methods

Ethics

Before starting the study, ethical clearance was obtained from the institutional ethics committee of Saveetha University (IHEC/SDC/FACULTY/22/PERIO/104).

Study characteristics

The purpose of the study was to detect dental cysts, abscesses as well as tumor lesions via image processing based on panoramic dental radiographs. The Department of public health dentistry designed an in vitro, retrospective study with the help from Department of Oral Medicine and Radiology, Saveetha dental college, and the Hospital.

Null Hypothesis (H0): There is no significant difference in the accuracy of texture analysis techniques, including Grey Level Run Length Matrix (GLRLM), Grey Level Co-occurrence Matrix (GLCM), and wavelet analysis, in recognizing dental cyst, tumor, and abscess lesions.

Alternate Hypothesis (Ha): There is a significant difference in the accuracy of texture analysis techniques, including Grey Level Run Length Matrix (GLRLM), Grey Level Co-occurrence Matrix (GLCM), and wavelet analysis, in recognizing dental cyst, tumor, and abscess lesions.

Sample characteristics

Digital panoramic radiographs with cystic, tumor or abscess lesions were retrieved. Radiographs of poor diagnostic quality (significant overlap of teeth, diffused image, or distortion) were excluded. The sample size was determined based on the re-search conducted by Ingrid et al.; 2003, N=172 (calculated using G-power) (at 95% power and 5% α -error) [32].

Study protocol

A total of 172 digital panoramic radiographs were selected based on the pre-determined inclusion and exclusion criteria. Tumor lesions in the selected

orthopantomograph (OPGs) include ameloblastoma, odontoma, ameloblastic fibroma, adenomatoid odontogenic tumour, hemangioma, enostosis, exostosis, cementoblastoma, torus mandibularis, torus palatinus, myxoma, osteoma, and osteoid osteoma. In addition to above-mentioned tumor lesion, dental cyst and abscess lesions were also selected. These measurements were taken with a digital orthopantomograph (PlanmeccaProOne; Helsinki, FINLAND) at the Department of Oral Medicine and Radiology, Saveetha Dental College and Hospital. Patient's identification data including age, gender, and the date of radiograph date were all recorded for each patient.

In the present study, the characteristics of dental panoramic radiographs were used to create methods for differentiating cyst and tumor lesions. The correlation between the evaluated diagnosis and the actual diagnosis followed (Fig. 1).

- A. *Data acquisition:* A set of dental panoramic images were collected from Saveetha Dental College and Hospital. Ground truth image was considered as a reference (Fig. 2).
- B. *Preprocessing and Enhancement:* Pre-processing is required to enhance the picture quality of dental panoramic radiographs since they are difficult to interpret. This step facilitates subsequent stages of feature extraction that are easier and more trustworthy. The radiographs were cleaned up to remove any

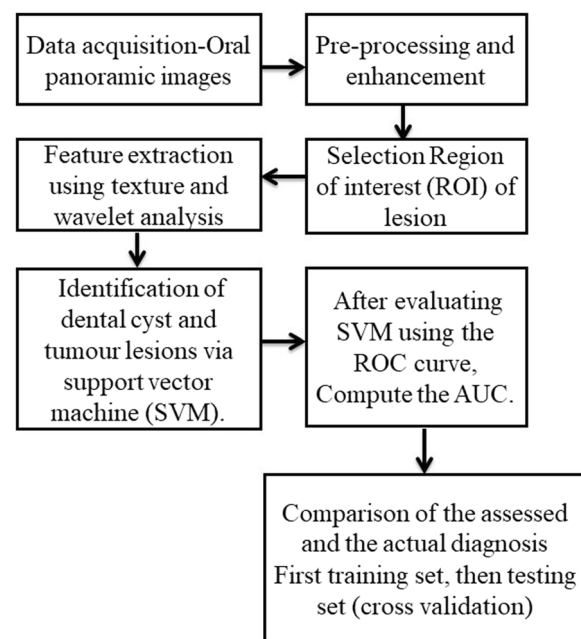


Fig. 1 A decomposition at level 1 B decomposition at level3



Fig. 2 Schematic representation of methodology



Fig. 3 Data Acquisition-Oral panoramic image of ameloblastoma (dental tumor)—Data Acquisition-Oral panoramic image of ameloblastoma (dental tumor). Increased radiolucency with ill-defined margins is observed in the jaw’s lower right posterior tooth region, with related tooth structures missing

artefacts or background noise during the preprocessing procedure. Noise was reduced and smoothed down using the gaussian filter, allowing for better contrast in dental panoramic radiographs. Noise was reduced by using a 3-by-3-square median filter. In order to get rid of the outliers without reducing the image’s sharpness, a median filter is used since it is less sensitive to extreme values. When the median filter is applied, abnormalities become sharper in the homogenous background. The area around the lesion has to be amplified. The purpose of Imadjust is to enhance images. The intensity of a grayscale picture is modified by imadjust to a new value that is 1% of the data at the extremes of brightness. This improves the image’s contrast (Fig. 3).

- C. *Selection Region of interest (ROI) of a lesion:* Dental panoramic radiographs with cyst and tumor lesions served as the regions of interest for this study. A trained operator manually identified all lesions in the radiographs, and these findings were then verified by a radiologist. With the mass placed in the middle of the window, a 40×40-pixel region of interest was extracted. The masses were then split in two: the training group and the evaluation group. The training set featured 40% cyst pictures and 40% tumor images, whereas the testing set contained 20% cyst images and 20% tumor images using one-three-hold-out cross-validation (Fig. 4).
- D. *Feature extraction using texture and wavelet analysis* (Fig. 5):

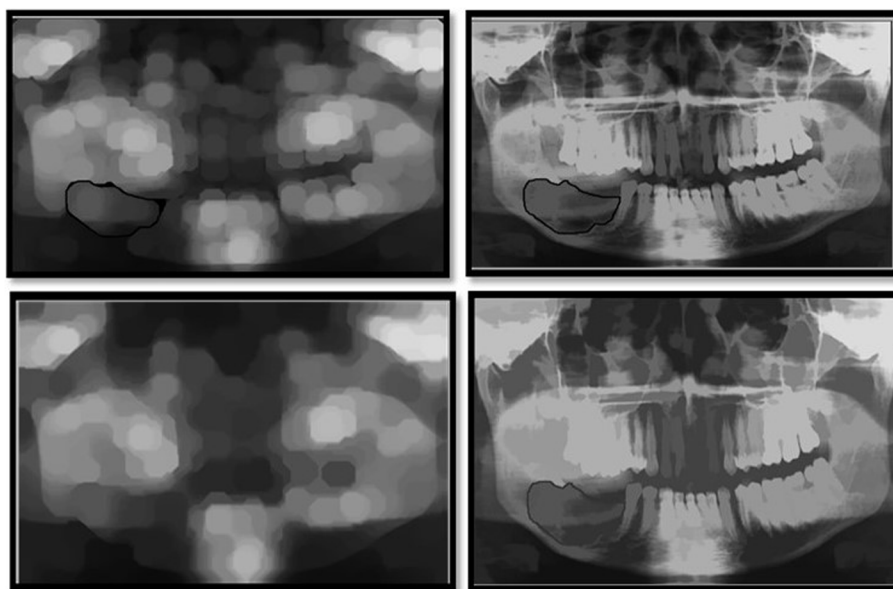


Fig. 4 Preprocessing and enhancement of the obtained OPG image—It shows the Preprocessing and enhancement of the obtained OPG image. In this step, the removal of artifacts and noise from the Image is carried out

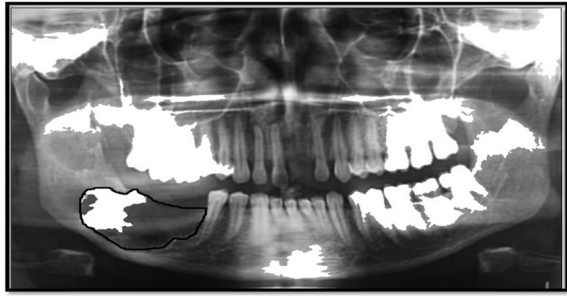


Fig. 5 Morphological processing for extracting Region of Interest – It shows the Morphological processing for extracting the Region of Interest. In this step, the lesion is associated with the nearby anatomical landmark and is separated from the OPG image

a. Feature extraction using texture analysis: The Image’s texture is the local variations in the image’s intensity. Using a co-occurrence matrix, statistical texture analysis is possible. This matrix is produced by estimating the pairwise statistics of image pixel intensity. The Grey level Co-occurrence Matrix [GLCM] was developed based on the premise that identical gray-level configurations are repeated in a texture. The co-occurrence matrix is denoted by $P(i, j | d, \theta)$, in which i and j are grey level readings separated by a distance d and an angle. With GLCM, it is possible to extract second-order statistical features. The GLCM includes the same number of rows and columns (Fig. 6). The following notations are used in GLCM.

- μ - Mean value of P
- μ_x - Mean value of Px
- μ_y - Mean value of Py
- σ_x - Standard deviation of Px
- σ_y - standard deviation of Py
- G- Size of the co-occurrence matrix

The characteristics of GLCM that were utilized in this study included energy, correlation, contrast, and homogeneity.

i. Energy: It provides the uniformity of an image. The energy of a constant image is 1.

$$Energy = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j)^2$$

ii. Correlation: measures level dependence among pixels at the specified relative position to one another. For an image with multiple linear structures, correlation will be high.

$$Correlation = \frac{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i, j\} \times P(i, j) - (\mu_x \times \mu_y)}{\sigma_x \times \sigma_y}$$

iii. Contrast: refers to the total amount of local differences present in an image. This indicates the degree to which $p(i, j)$ from the diagonal element $i = j$ influences the degree of contrast. The greater the image variation, the more diagonal $P(i, j)$ values will be concatenated, resulting in a higher contrast.

$$contrast = \sum_{n=0}^{G-1} n^2 \left\{ \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j) \right\}, |i - j| = n$$

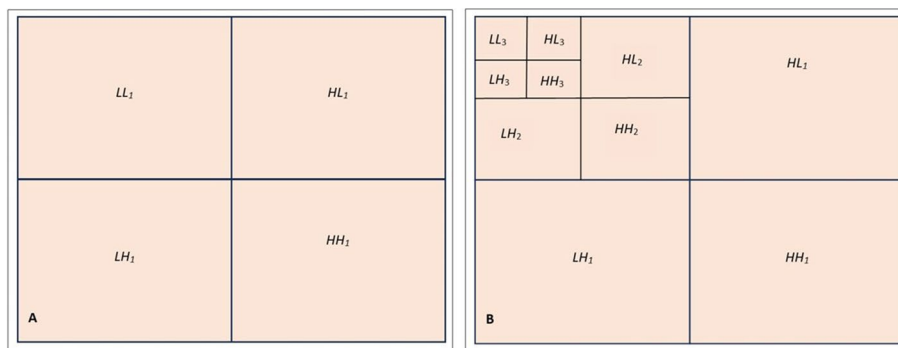


Fig. 6 It shows the regional maxima superimposed on the original for enhancing the Region of Interest (ROI). Here, the lesion is segmented and undergoes wavelet and texture analysis (GLCM, GLRLM)

Homogeneity: It measures the nearness of the components in GLCM with the diagonal of the GLCM matrix and is used to evaluate the homogeneity of the GLCM. Homogeneity is 1 for the diagonal elements. The Co-occurrence matrix of a heterogeneous images provide an even distribution of P (i, j) values, whereas homogeneous images include a combination of high as well as low P (i, j) values.

An Input vector to the neural network and the wavelet transform coefficients were given.

z. *Feature extraction utilizing Wavelet Analysis:* To approximate a picture, we can utilize matrix A with the pixel intensities at each grey level P(i,j) as its elements [10]. A is a square with dimensions of 2n×2n, where n is A integer. Here is how the wavelet transform works: The rows of A are processed using filters G and H. The resultant matrices are HrA and LrA of size 2n2n-1. Next, the HrA and LrA column matrix is processed once again with the G and H filters, yielding four matrices HcLrA, HcHrA, LcLrA, and LcHrA of size 2n-1×2n-1. The LcLrA matrix is an average matrix, whereas the other matrices are more detailed matrices. It's also possible to execute many degrees of decomposition (Fig. 7).

$$Homogeneity = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{P(i, j)}{1 + |i - j|}$$

E. Identifying dental cysts or tumor lesions using the support vector machine (SVM) method:

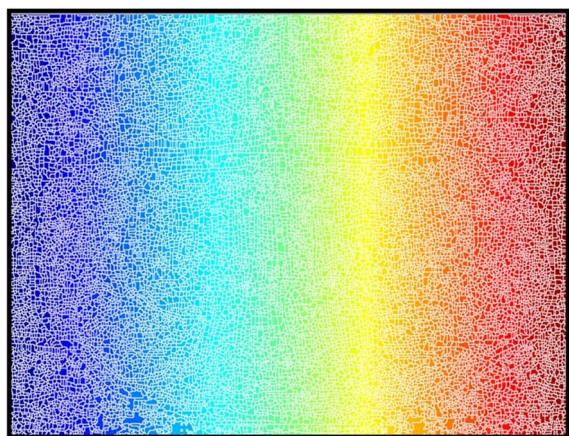


Fig. 7 shows the Wavelet analysis for features extraction from wavelet coefficients. After this, the Image is subjected to a support vector machine classifier (SVM)

Following feature extraction and selection, these data were fed into a classifier that categorize each picture to one of two classes: cysts or tumors. In order to classify these lesions, we used Support Vector Machine [SVM] technique.

The ROC curve was used in this study because of its capacity to provide a comprehensive and objective assessment. The ROC curve is the scatter plot of True Positive Fraction (TPF) as the function of False Positive Fraction (FPF). The area that is under the receiver operating characteristic curve (AUC) can be used as a criterion [33, 34].

F. Evaluate SVM using the ROC curve and compute the AUC:

The ROC curve was used in this study since it is both comprehensive and objective in its assessment [32]. The ROC curve is the scatter plot of True Positive Fraction (TPF) as the function of False Positive Fraction (FPF). The area that is under the receiver operating characteristic curve (AUC) can be used as a criterion [33, 34].

Statistical analysis

The gathered data was transferred into a Microsoft Excel spreadsheet and SPSS (version 21) was used to conduct the statistical analysis. Initially descriptive statistics were computed. For inferential analysis, statistical significance was determined by a p-value of 0.05.

Measures of performance evaluation

The effectiveness of the system is evaluated using a variety of criteria. Classification Accuracy (AC) and Mathews Correlation Coefficient (MCC) are the measures used. The Confusion Matrix is used for calculating these values. The results of a classification system's actual classifications and predictions are summarized in a confusion matrix (Kohavi and Provost, 1998) [35]. The matrix information is often used to assess the effectiveness of such systems. The confusion matrix for a two-class classifier is shown in the tabular column below.

-
- TP (True Positive) – Correct prediction of abnormal
 - TN (True Negative) – Correct Prediction as normal
 - FP (False Positive) – Incorrect prediction of abnormal
 - FN (False Negative) – Incorrect prediction of normal
-

From the confusion matrix, accuracy (AC) can be obtained:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

The Matthews correlation coefficient (MCC): In machine learning, the Matthews coefficient of correlation (MCC) measures the accuracy of two-class(binary) classifications.

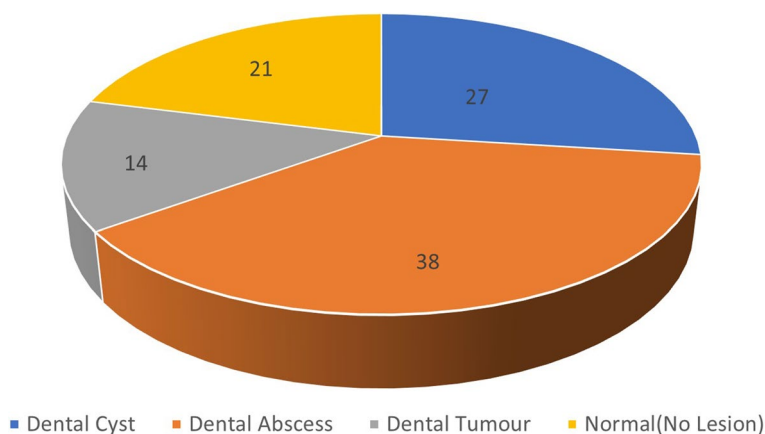


Fig. 8 Distribution of OPG images based on dental cysts, tumors, and abscess lesions

The MCC, which yields a value between -1 and +1, is essentially a correlation coefficient that exists between observed as well as predicted binary classifications. A coefficient of +1 indicates a perfect prediction, coefficient 0 is no better than a random prediction, and a coefficient of -1 denotes complete disagreement between the prediction and the observation. The MCC is estimated by using:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Clinical tests are evaluated based on their sensitivity and specificity.

- Sensitivity of a clinical test is defined as the ability of the test to identify the percentage of patients with disease, Calculated as: $TP / (TP + FN)$
- Specificity is defined as the ability of test to appropriately identify healthy individuals, Calculated as: $TN / (TN + FP)$
- Positive Predictive Value (PPV) or precision rate is the proportion of positive test results that are the true positives (accurate diagnosis). The PPV is denoted by the formula, $PPV = TP / (TP + FP)$

Results

Figure 8 shows the distribution of 172 OPG images in which 73 (38%) were dental abscesses, 52 (27%) were dental cysts, 27 (14%) were dental tumors, and 40 (21%) images were normal without any lesions.

Table 1 depicts the performance of GLCM by identifying the true negative (39), false negative (1), false positive (0), and true positive (40) by comparing the actual and predicted classifications done by confusion matrix for diagnosing dental cyst, tumor, and abscess lesions using

Table 1 Confusion matrix for GLCM features

		Predicted	
		Negative(-)	Positive(+)
Actual	Negative(-)	39(TN)	1(FN)
	Positive(+)	0(FP)	40(TP)

TN True Negative, FN False Negative; FP False Positive; TP True Positive; GLCM Grey Level Co-occurrence Matrix

80 OPG images in which 40 images had the lesion and 40 images were without any lesion.

Table 2 depicts the performance of GLCM by identifying the true negative (40), false negative (2), false positive (2), and true positive (36) by comparing the actual and predicted classifications done by confusion matrix for diagnosing dental cyst, tumor, and abscess lesions using 80 OPG images in which 40 images had the lesion and 40 images were without any lesion.

Table 3 depicts the performance of GLCM by identifying the true negative (30), false negative (4), false positive (3), and true positive (43) by comparing the actual and predicted classifications done by confusion matrix for diagnosing dental cyst, tumor, and abscess lesions using 80 OPG images in which 40 images had the lesion and 40 images were without any lesion.

Table 2 Confusion matrix for GLRLM features

		Predicted	
		Negative(-)	Positive(+)
Actual	Negative(-)	40(TN)	2(FN)
	Positive(+)	2(FP)	36(TP)

TN True Negative, FN False Negative; FP False Positive; TP True Positive; GLRLM Grey Level Run Length Matrix

Table 3 Confusion matrix for wavelet analysis

		Predicted	
		Negative(-)	Positive(+)
Actual	Negative(-)	30(TN)	4(FN)
	Positive(+)	3(FP)	43(TP)

TN True Negative, FN False Negative; FP False Positive; TP True Positive

Table 4 Matrix for all three techniques

Value	Wavelet analysis	GLCM	GLRLM
TP	43	40	36
FP	3	0	2
FN	4	1	2
TN	30	39	40

TN True Negative, FN False Negative; FP False Positive; TP True Positive; GLRLM Grey Level Run Length Matrix; GLCM Grey Level Co-occurrence Matrix

Table 5 Evaluation results

Measure	Wavelet analysis	GLCLM	GLRLM
AC	91%	98%	95%
MCC	0.82	0.97	0.89
SN	90%	97%	94%
SP	90%	100%	95%
PPV	93%	100%	94%

TN True Negative, FN False Negative; FP False Positive; TP True Positive; GLRLM Grey Level Run Length Matrix; GLCM Grey Level Co-occurrence Matrix

Table 4 depicts the performance of GLCM, GLRLM, and wavelet analysis by identifying the true positive, true negative, false positive, and false negative by comparing the actual and predicted classifications done by confusion matrix for diagnosing dental cyst, tumor, and abscess lesions using 80 OPG images in which 40

images had the lesion and 40 images were without any lesion.

Table 5 demonstrates the calculated accuracy, specificity, sensitivity, Mathews Correlation Coefficient, and Precision Rate calculated using values from Table 4. Various models are used to extract texture features, and the results are inconsistent (Fig. 9).

Figure 10 shows the receiver operator characteristic (ROC) curve evaluating the performance of GLCM, GLRLM, and wavelet analysis by plotting sensitivity along the y-axis and Specificity along the x-axis. The area under the ROC curve (AUC) is a criterion.

Discussion

Based on their morphology and protein expression, histology and immunohistochemistry currently diagnose dental tumors and cysts. A biopsy is a painful procedure that cannot be recommended every time [36]. Therefore, the computer-aided diagnosis of a dental cyst, tumor, and abscess lesions using oral panoramic images is attempted in this study [37]. This will avoid exposure to unnecessary biopsies. An intelligent, low-cost, and interpretable solution may be achieved via the use of soft computing methods such as texture analysis and neural networks for extracting the features than traditional techniques [38].

Today texture-based medical imaging has a vast number of research being done on the early detection of stroke, Automatic Diagnostic Systems for CT Liver and CT (computed tomography) images are widely utilized to diagnose liver disease and develop a computer-assisted classification system for detecting cancer using digital mammograms [39–45].

This study approaches the problem by running texture (GLCM and GLRM) and wavelet analysis to extract the features in the Image. We calibrated values by feeding the recognition and differentiation of dental cysts, abscesses, and tumour lesions using the abovementioned methods.

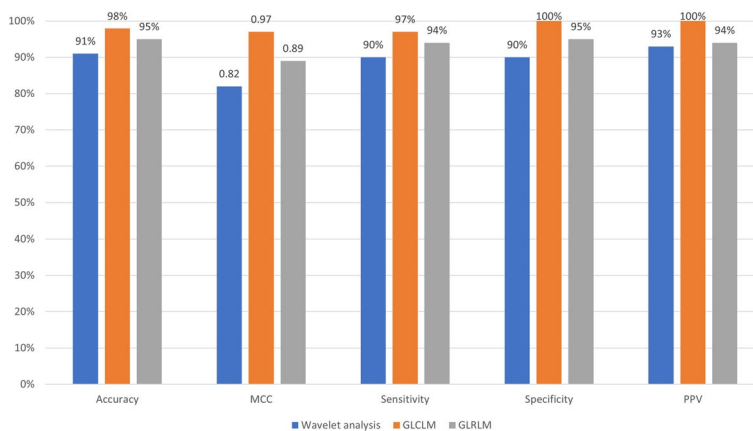


Fig. 9 Analysis of the Performance of wavelet analysis, GLCM, and GLRLM

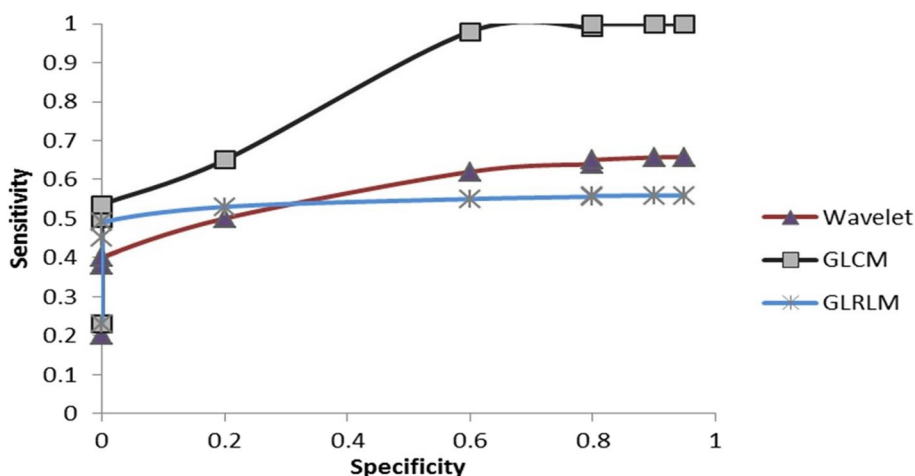


Fig. 10 Receiver operator characteristic curve

All the methods analysed the change in the expression of each lesion. Later sensitivity, Specificity, accuracy, MMC, and PPV of the classification were calculated. All the experiments were performed using MatlabVer 7.1 on a PC Intel Pentium Centrino with 1 GB of RAM. A total of 53 dental abscesses, 17 tumors, and 52 dental cyst lesions measuring 40×40 pixels were transformed into wavelet analysis, GLCM, and GLRLM. The minimum and maximum values of all three lesions are different. We acquired up to 91% accuracy with wavelet analysis, 98% accuracy with GLCM, and 95% accuracy with GLRLM; utilizing the ROC curve along with computing the AUC, we acquired up to 0.9444 AUC. with FO, we estimated the AUC for GLCM up to 0.9361 and for GLRLM up to 0.8722.

The feature values from both classes overlap, but all the three lesions have different minimum and maximum values. Due to the overlapping value, this conclusion suggests that the categorization process is not straightforward (non-linear). The non-linear classification is made feasible by the maximum and minimum feature values within each class. Lesion types such as cysts have median values between 13.7650 and 213.2888, whereas tumors have mean values between 85.0038 and 252.3044. This indicates that SVM may be used to differentiate cyst lesions, abscesses, and tumors.

The texture traits were shown to be useful in identifying cysts and tumor lesions in this research. When attempting to locate a region of interest (ROI) in an image, texture is an important factor to consider. Wavelet analysis was used to extract 5 texture features, 13 texture features were extracted based on GLCM, and 7 texture features were extracted based on GLRLM. Wavelet analysis provides a set of five features: central tendency, dispersion, kurtosis, skewness, and dispersion. GLCM utilized the

following measures: autocorrelation, contrast, correlation, cluster prominence and cluster shade, energy, dissimilarity, Homogeneity, information measure, maximum probability, inverse difference, inverse difference normalized, and inverse difference moment normalized. GLRLM utilized the following measures: Low Gray Level Run Emphasis, High Gray Level Run Emphasis, Short Runs Emphasis, Long Runs Emphasis, Run per-centage, Gray Level Non-uniformity, Run Length Non-uniformity.

The ROC curve was used in this study because of its capacity to provide a com-prehensive and fair evaluation ability [43]. A ROC curve is the plotting of true positive fraction (TPF) as the function of false positive fraction (FPF) [46, 47]. The area under the ROC curve (AUC) can be used as a criterion.

A strength of this method is that it is simple, non-invasive, and quick. This method holds significant promise for cancer patients by potentially enabling more personalized and targeted treatment approaches based on the specific genetic signatures identified, ultimately enhancing the effectiveness of cancer therapies. The primary objective of the research was to optimize the classification and prediction models for these cancers, which led to the identification of 96 genes that are extremely relevant to these malignancies. Nonetheless, we anticipate that the above list can be extended by employing more detailed arrays and larger training sample sets.

Although we achieved a high degree of specificity and sensitivity for diagnostic classification, we expect that the sensitivity of these models can be improved for diagnostic purposes in clinical practice with larger arrays and more samples. Research into the staging and biological behavior of malignancies, with the goal of predicting prognosis and guiding treatment, is a potential future use of these technologies. Furthermore, future

research ought to be performed to improve the precision of GLRLM as well as wavelet analysis texture characteristics. Since we believe that this approach provides a robust and alternative means for identifying dental tumours, cysts, and abscesses, and given that the results indicate enhanced system performance with a larger set of training images, we recommend this method as a valuable tool in clinical practice.

Conclusions

The accuracy of the results achieved through GLCM analysis yielded an impressive 98 percent. These GLCM features demonstrate strong potential for further exploration. Consequently, there is a pressing need for future investigations aimed at refining the precision of GLRLM and wavelet analysis texture characteristics. This supports the prospective use of these techniques as an adjunct to standard histological diagnosis. This research will aid dentists in establishing accurate diagnoses and prompt treatment plans for the benefit of their patients.

Acknowledgements

None.

Authors' contributions

"Conceptualization, V.S.K.; PK.Y.; D.S.; and K.C.S.; methodology, V.S.K.; PK.R.; PK.Y.; R.V.A.; D.S.; and K.C.S.; software, PK.R.; A.A.F.; I.T.A.; S.A.F.; A.H.S.; validation, PK.R.; A.A.F.; I.T.A.; S.A.F.; A.H.S.; formal analysis, PK.Y. and K.C.S.; investigation, V.S.K.; PK.R.; A.A.F.; I.T.A.; S.A.F.; A.H.S.; K.C.S.; writing—original draft preparation, PK.Y.; D.S.; and K.C.S.; writing—review and editing, V.S.K.; PK.R.; PK.Y.; R.V.A.; D.S.; A.A.F.; I.T.A.; S.A.F.; A.H.S.; K.C.S.; supervision, K.C.S.; project administration, PK.Y.; funding acquisition, K.C.S. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Availability of data and materials

The data will be available on reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate.

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Saveetha Dental college and Hospitals, Saveetha Institute of Medical and Technical Sciences (IHEC/SDC/FACULTY/22/PERIO/104). Being a retrospective study, the informed consent from the participants was waived by the approval of local ethics committee of Saveetha Dental college and Hospitals, Saveetha Institute of Medical and Technical Sciences.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Public Health Dentistry, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. ²Department of Periodontics, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Saveetha University,

Chennai, Tamil Nadu, India. ³Department of Preventive Dentistry, College of Dentistry, Jouf University, 72345 Sakaka, Saudi Arabia. ⁴Oral Surgery Department, Institute of Dentistry, Queen Mary University of London, London E1 2AD, UK. ⁵Present Address: College of Dentistry, Jouf University, 72345 Sakaka, Saudi Arabia. ⁶Department of Oral & Maxillofacial Surgery & Diagnostic Sciences, College of Dentistry, Jouf University, 72345 Sakaka, Saudi Arabia. ⁷Department of Oral Medicine and Radiology, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu 602105, India.

Received: 24 June 2023 Accepted: 23 October 2023

Published online: 06 November 2023

References

- Chen S, Forman M, Sadow PM, August M. The Diagnostic Accuracy of Incisional Biopsy in the Oral Cavity. *J Oral Maxillofac Surg*. 2016;74(5):959–64. <https://doi.org/10.1016/j.joms.2015.11.006>.
- Romeo U, et al. Biopsy of Different Oral Soft Tissues Lesions by KTP and Diode Laser: Histological Evaluation. *Scientific World Journal*. 2014;2014:1–6. <https://doi.org/10.1155/2014/761704>.
- Narayan M, Augustine D, Jacob T, Shrivastava K, Shrivastava D, Narayan S. Does Clinical Staging and Histological Grading Show Parallelism In Oral Submucous Fibrosis? A Retrospective Study from an Indian City. *J Interdiscip Histopathol*. 2014;2(3):145. <https://doi.org/10.5455/jihp.20140406010651>.
- Diamanti N, Duxbury AJ, Ariyaratnam S, Macfarlane TV. Attitudes to biopsy procedures in general dental practice. *Br Dent J*. 2002;192(10):588–92. <https://doi.org/10.1038/sj.bdj.4801434>.
- Bänkfalvi A, Piffkò J. Prognostic and predictive factors in oral cancer: the role of the invasive tumour front. *J Oral Pathol Med*. 2000;29(7):291–8. <https://doi.org/10.1034/j.1600-0714.2000.290701.x>.
- D. Srivastava, D. Shrivastava, and D. Austin, "Journey Towards The 3d Dental Imaging- The Milestones In The Advancement Of Dental Imaging," *Int J Adv Res (Indore)*, vol. 4, no. 12, pp. 377–382, Dec. 2016, doi: <https://doi.org/10.21474/IJAR01/2405>.
- A. Sushanth. A et al., "Recommendations, Practices and Infrastructural Model for the Dental Radiology Set-up in Clinical and Academic Institutions in the COVID-19 Era," *Biology (Basel)*, vol. 9, no. 10, p. 334, Oct. 2020, doi: <https://doi.org/10.3390/biology9100334>.
- K. Chandan Srivastava et al., "Prevalence and Characteristics of Idiopathic Osteosclerosis in Saudi Arabian Population Using Digital Panoramic Radiographs: A Retrospective Study," 2020.
- A. M. do Couto et al., "Chronic inflammatory periapical diseases: a Brazilian multicenter study of 10,381 cases and literature review," *Braz Oral Res*, vol. 35, pp. 1–12, 2021, doi: <https://doi.org/10.1590/1807-3107BOR-2021.VOL35.0033>.
- Kumar J, Vanagundi R, Manchanda A, Mohanty S, Meher R. Radiolucent Jaw Lesions: Imaging Approach. *Indian J Radiol Imaging*. 2021;31(1):224–36. <https://doi.org/10.1055/s-0041-1729769>.
- M. Soluk-tekkesin and J. M. Wright, "The world health organization classification of odontogenic lesions: a summary of the changes of the 2022 (5th) edition," *Turkish Journal of Pathology*, 2022, doi: <https://doi.org/10.5146/tjpath.2022.01573>.
- Sghaireen MG, et al. A CBCT Based Three-Dimensional Assessment of Mandibular Posterior Region for Evaluating the Possibility of Bypassing the Inferior Alveolar Nerve While Placing Dental Implants. *Diagnostics*. 2020;10(6):406. <https://doi.org/10.3390/diagnostics10060406>.
- Sghaireen MG, et al. Comparing the Diagnostic Accuracy of CBCT Gray-scale Values with DXA Values for the Detection of Osteoporosis. *Appl Sci*. 2020;10(13):4584. <https://doi.org/10.3390/app10134584>.
- Iqbal A, et al. Evaluation of root canal morphology in permanent maxillary and mandibular anterior teeth in Saudi subpopulation using two classification systems: a CBCT study. *BMC Oral Health*. 2022;22(1):171. <https://doi.org/10.1186/s12903-022-02187-1>.
- K. C. Srivastava, "A CBCT aided assessment for the location of mental foramen and the emergence pattern of mental nerve in different dentition status of the Saudi Arabian population," *Braz Dent Sci*, vol. 24, no. 1, Dec. 2020, doi: <https://doi.org/10.14295/bds.2021.v24i1.2372>.
- Assiri KI, Sandeepa NC, Asiri RS, Mulawi SA, Najimi SM, Srivastava KC. Assessment of Oral-Systemic Disease Association amongst Dental Patients: A Retrospective Panoramic Radiographic Study. *J Contemp Dent Pract*. 2020;21(7):748–55.

17. Srivastava KC, et al. Assessing the Prevalence and Association of Pulp Stones with Cardiovascular Diseases and Diabetes Mellitus in the Saudi Arabian Population—A CBCT Based Study. *Int J Environ Res Public Health*. 2020;17(24):9293. <https://doi.org/10.3390/ijerph17249293>.
18. Özen T, Kamburoğlu K, Cebeci ARI, Yüksel SP, Paksoy CS. Interpretation of chemically created periapical lesions using 2 different dental cone-beam computerized tomography units, an intraoral digital sensor, and conventional film. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*. 2009;107(3):426–32. <https://doi.org/10.1016/j.tripl.2008.08.017>.
19. S. Gamanagatti, "Pictorial Essay Singapore Med Radiographical approach to jaw lesions," 2008.
20. MA. Bittencourt et al., "Accuracy of computer-aided image analysis in the diagnosis of odontogenic cysts: A systematic review," *Med Oral Patol Oral Cir Bucal*, pp. e368–e378, 2021, doi: <https://doi.org/10.4317/medoral.24238>.
21. Nakamoto T, Hatsuta S, Yagi S, Verdonshot RG, Taguchi A, Kakimoto N. Computer-aided diagnosis system for osteoporosis based on quantitative evaluation of mandibular lower border porosity using panoramic radiographs. *Dentomaxillofacial Radiology*. 2020;49(4):20190481. <https://doi.org/10.1259/dmfr.20190481>.
22. Paderno A, et al. Deep Learning for Automatic Segmentation of Oral and Oropharyngeal Cancer Using Narrow Band Imaging: Preliminary Experience in a Clinical Perspective. *Front Oncol*. 2021;11: 626602. <https://doi.org/10.3389/fonc.2021.626602>.
23. Dildar M, et al. Skin Cancer Detection: A Review Using Deep Learning Techniques. *Int J Environ Res Public Health*. 2021;18(10):5479. <https://doi.org/10.3390/ijerph18105479>.
24. Tong N, Gou S, Yang S, Ruan D, Sheng K. Fully automatic multi-organ segmentation for head and neck cancer radiotherapy using shape representation model constrained fully convolutional neural networks. *Med Phys*. 2018;45(10):4558–67. <https://doi.org/10.1002/mp.13147>.
25. Shiu SCK, Pal SK. Case-Based Reasoning: Concepts, Features and Soft Computing. *Appl Intell*. 2004;21(3):233–8. <https://doi.org/10.1023/B:APIN.0000043556.29968.81>.
26. P. Škoda, T. Lipić, A. Srp, B. Medved Rogina, K. Skala, and F. Vajda, "Implementation framework for Artificial Neural Networks on FPGA," 2011. [Online]. Available: <https://www.researchgate.net/publication/224250544>
27. R. Joost and R. Salomon, "Advantages of FPGA-based multiprocessor systems in industrial applications," in *31st Annual Conference of IEEE Industrial Electronics Society, 2005. IECON 2005.*, IEEE, 2005, p. 6 pp. doi: <https://doi.org/10.1109/IECON.2005.1568946>.
28. García G, Jara C, Pomares J, Alabdo A, Poggi L, Torres F. A Survey on FPGA-Based Sensor Systems: Towards Intelligent and Reconfigurable Low-Power Sensors for Computer Vision, Control and Signal Processing. *Sensors*. 2014;14(4):6247–78. <https://doi.org/10.3390/s140406247>.
29. Hope TMH, Seghier ML, Leff AP, Price CJ. Predicting outcome and recovery after stroke with lesions extracted from MRI images. *Neuroimage Clin*. 2013;2:424–33. <https://doi.org/10.1016/j.nicl.2013.03.005>.
30. H. B. Kekre and S. Gharge, "Segmentation of MRI Images Using Probability and Entropy as Statistical Parameters for Texture Analysis," 2009. [Online]. Available: <http://www.ripublication.com/acst.htm>
31. A. Procházka, A. Gavlasová, and O. Vyšata, "Texture Segmentation and Classification in Biomedical Image Processing."
32. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175–91. <https://doi.org/10.3758/BF03193146>.
33. C. M. Florkowski, "Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests," *Clin Biochem Rev*, vol. 29 Suppl 1, no. Suppl 1, pp. S83–7, Aug. 2008.
34. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*. 2007;115(5):654–7. <https://doi.org/10.1161/CIRCULATIONAHA.105.594929>.
35. Regezi JA. Periapical diseases: spectrum and differentiating features. *J Calif Dent Assoc*. 1999;27(4):285–9.
36. U. Nischal, Nischal Kc, and U. Khopkar, "Techniques of skin biopsy and practical considerations," *J Cutan Aesthet Surg*, vol. 1, no. 2, pp. 107–11, Jul. 2008, doi: <https://doi.org/10.4103/0974-2077.44174>.
37. Lee J, Kim D, Jeong S. Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. *Oral Dis*. 2020;26(1):152–8. <https://doi.org/10.1111/odi.13223>.
38. M. Avanzo et al., "Machine and deep learning methods for radiomics," *Med Phys*, vol. 47, no. 5, May 2020, doi: <https://doi.org/10.1002/mp.13678>.
39. J. Luts, F. Ojeda, R. Van de Plas Raf, B. De Moor, S. Van Huffel, and J. A. K. Suykens, "A tutorial on support vector machine-based methods for classification problems in chemometrics," *Anal Chim Acta*, vol. 665, no. 2, pp. 129–145, Apr. 2010, doi: <https://doi.org/10.1016/j.aca.2010.03.030>.
40. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Inform*. 2005;38(5):404–15. <https://doi.org/10.1016/j.jbi.2005.02.008>.
41. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett*. 2006;27(8):861–74. <https://doi.org/10.1016/j.patrec.2005.10.010>.
42. "IJCS_39_2_08".
43. H. Wibawanto, A. Susanto, T. S. Widodo, and S. M. Tjokronegoro, "Discriminating Cystic and Non Cystic Mass using GLCM and GLRLM-based Texture Features," 2010. [Online]. Available: <http://www.ripublication.com/ijeer.htm>
44. Sghaireen MG, et al. Machine Learning Approach for Metabolic Syndrome Diagnosis Using Explainable Data-Augmentation-Based Classification. *Diagnosics*. 2022;12(12):3117. <https://doi.org/10.3390/diagnostics12123117>.
45. Sghaireen MG, et al. Vitamin D, Cholesterol, and DXA Value Relationship with Bimaxillary Cone Beam CT Values. *J Clin Med*. 2023;12(7):2678. <https://doi.org/10.3390/jcm12072678>.
46. Reddy TK, Kumaravel N. Wavelet based texture analysis and classification of bone lesions from dental CT. *Int J Med Eng Inform*. 2010;2(3):319. <https://doi.org/10.1504/IJMEI.2010.035223>.
47. T. K. Reddy, T. Kaliraja, B. Ramesh, and G. Gurumurthy, "Computer Aided Diagnosis of Dental CT images for Bone Quality Assessment," *Int J Biosci Biochem Bioinforma*, pp. 226–233, 2011, doi: <https://doi.org/10.7763/IJBBB.2011.V1.42>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

