Original Research Paper

Volume - 12 | Issue - 10 | October - 2022 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Dental Science

AN ASSOCIATIVE CLASSIFICATION BASED APPROACH TOWARDS ANALYSIS OF DENTAL CARIES X-RAY IMAGES

Sovamayee	HiTech Dental College and Hospital, Bhubaneswar, Odisha, India.
Debasmita Das	HiTech Dental College and Hospital, Bhubaneswar, Odisha, India.
Raghunath Dey	GITAM (Deemed to be University), Visakhapatnam, India.
Rakesh Chandra Balabantaray	IIIT Bhubaneswar, Odisha, India.

ABSTRACT The most common disease on the planet is dental caries, also known as cavities. Almost everyone has had this condition at some point in their lives. Early identification of dental caries can considerably reduce the risk of serious damage to teeth in people who have dental disease. As medical imaging becomes more efficient and faster to use, clinical applications are having a greater impact on patient care. Recently, there has been a lot of interest in machine learning approaches for categorizing and analyzing image data. In this study, we describe a new strategy for locating and identifying dental caries from X-ray photos as a dataset and using associative classification as a classification method. This technique incorporates both classification and correlation. The numerical discrimination approach is also used in the strategy. This is the first study to employ association-based classifications to determine dental cavities and root canal treatment positions. This method was tested on real data from hundreds of patients and found to be very good at finding unexpected damage to teeth.

KEYWORDS : Radiographs, Dental X-rays, association based classification, discritization

INTRODUCTION:

MRI (Magnetic Resonance Imaging), CT (Computerized Tomography), x-rays, and other medical imaging procedures are used to determine what's wrong with a person. Many problems, including segmentation, glaucoma detection, and others, can be solved using machine learning approaches. CNN is excellent at detecting errors, but it takes a long time. In the research literature, several classification systems have been employed to discover and evaluate dental caries, but no studies have used associative classification. This study employed associative classification to figure out where cavities are in teeth and their appropriate positions for root canals.

Radiographs are needed to get an accurate diagnosis and find problems that would be hard to see otherwise. A panoramic view is a type of dental x-ray that is often used in orthodontics. It gives a clear view of the teeth and jaws, as well as screening and diagnostic information for a number of illnesses and abnormalities that can be found in the oral section.

The paper has the following structure: Some of the recent work is in the second section. The third section talks about our methods and shows how an association-based classifier can be used to train and classify samples of teeth. In the fourth section, the results of several tests will be shown. This study ends with Section 5.

Related Work:

A large amount of studies existing in the literature related to the proposed work. In this section these studies are collected and presented in a compact and concise manner to easily grasp the idea and knowledge of the work. The presentation of the study has been done in two segments. One highlights the different approaches used to analyze the dental caries, whereas the second section gives knowledge on the application of associative classifications.

Analysis Of Dental Caries:

Dental caries is one of the most common dental issues worldwide. It means a cavity, which is another name for tooth decay. Dental caries progresses through various stages, but the goal here is to classify the condition.

This study [1] implies digital radiography's texture factor can detect dental cavities. The dental diagnosis system uses the Laplacian filter for image sharpening, adaptive threshold and morphological procedures for image segmentation, and SVM for image classification. The Classifier uses segmented picture texture signals to determine if an image has caries. Textural features like (Grey Level Co-occurrence Matrix) and GLDM (Grey Level Difference Method) are better by several parameters, like accuracy, sensitivity, specificity, and precision. An Analysis of Variance (ANOVA) was conducted at a 5% significance level.

To extract crucial x-ray features, a capsule network uses transfer learning to draw prediction results and advantages from several pretrained deep learning models **[2]**. On a test set of 470 panoramic shots and 240 annotated images, the model scored 86.05%. As long as authentic patients' panoramic x-rays are used, the resulting score demonstrates satisfactory detection performance and an increase in caries detection speed. In the test set, the model got recall scores of 69.44% for mild caries lesions and 90.52% for severe caries lesions. This suggests that severe caries spots are easier to find and that moderate caries spots need a larger dataset to find.

Automatically classify dental problems from panoramic x-rays that were taken at three different dental clinics and analysed to show 14 different tooth problems **[3]**. Using the annotated data, a CNN was trained to learn about semantic segmentation. Each tooth in the area of interest was given a label, and the problem that was causing it was decided by a majority vote based on the histogram. Several ways were used to judge the system, such as intersection over union for semantic segmentation and accuracy, precision, recall, and F1-score for bounding box detection.

This paper **[4]** suggests a way to classify oral malignant disorders and the standard region using preprocessing techniques like the Deriche–Canny edge detector and the circular Hough transform (CHT); a textural analysis approach based on the gray-level co-occurrence matrix (GLCM); and a feature selection algorithm (linear discriminant analysis (LDA) followed by k-nearest neighbour) (KNN). When it comes to telling the difference between normal and abnormal oral cavity regions, accuracy, sensitivity, and specificity are 83%, 85%, and 84%, respectively. To measure performance, graphs were made of the receiver operating characteristics of diagnosing periodontitis with and without the HPIL system.

The study **[5]** discusses how to discover cavities in x-ray images by converting RGB to Gray, creating a binary image, choosing the region of interest, removing the backdrop, identifying regions, breaking the image into many blocks, and lastly locating the cavities.

Using Associative Classification

The associative classification (AC) approach was used to figure out what's wrong with centrifugal chillers **[6]**. To begin with, association rules with high confidence values and significant support are identified. To speed up the process of mining for association rules, FP-growth is used instead of the Apriori method. Second, only class association rules (CARs) with fault class consequences are kept. Third, pruned CARs are formed by ranking CARs and deleting extraneous rules with the assumption that "higher rank" is preferable. Fourth, the AC algorithm is used to pick a small number of rules from the trimmed CARs to make an associative classifier.

This work **[7]** presents a way to classify things based on models that can be understood and explained. A framework has been presented for building rule-based classifiers using class association rules. Two real-world datasets have been investigated: one that was collected using wearable sensors (accelerometers) in the field of personalized health, and another that was collected using smart phone sensors to detect activity.

This paper **[8]** introduces the context-aware personalized human activity recognition (CAPHAR) framework, which uses contextual information to generate class association rules between low-level actions, sensor activations, and high-level activities. Personalization in CAPHAR reduces the activity management challenge by taking advantage of how each person performs using a similarity score.

This research **[9]** presents a new artificial immune system model for associative classification with competitive breast cancer detection performance. The proposed model is based on the immune system of the human body. The immune system's ability to recognize antigens is used as a model for finding the correct antigens. Other well-known categorization models have shown that the quality of the suggested model is comparable to theirs. The model also has a low computational cost, which is beneficial. This model did well on classification challenges, demonstrating that swarm intelligence can be used for problems other than optimization.

Single-neuron recordings of epileptic patients performing an associative long-term memory test before surgery were used **[10].** During encoding and retrieval, auditory beat stimuli improved memory. Neurons encode familiarity, recall, connection to source memory, and overlap of neuron groups. It has been shown that MTL neurons respond to stimuli with understanding even when familiar material is presented ("old/new effect"). Familiarity-coding neurons and retrieval-coding neurons are comparable. Neurons that stored information about a source were different. Neurons coding for source memory and item memory must not intersect. Several neural groups govern other functions.

From the above mentioned literature, it has been observed that associative classification has been successfully used and implemented in several studies but never towards dental cavity classification and recognition. This motivates to do implement associative classification on some dental cavity classification task.

Classification Based On Association Methodology

Association and classification are two of the most commonly used data mining techniques. It is known as "association rule mining" when a set of rules has a particular level of support and confidence. A technique for finding the fewest possible classifier-applicable rules is called "classification rule mining". Here the approach combines the idea of association and classification, hence called associative classification (AC). Association rule mining and categorization rule mining are necessary for such a practice. The essential steps are shown in Fig. 1. The class is the only determining objective. AC is done out in this situation in three stages **[11]**:

1. Discretization of continuous data,
2. Rule production, and
3. Making classifier.

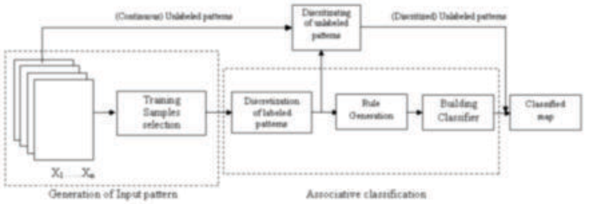

Fig 1. Block Diagram Of The Proposed Dental X-ray Image Classification Using Associative Clasification

Discretization Of Continuous Attributes

In general, the range of a continuous-valued attribute is divided into two intervals multiple times to make it discrete. A threshold value of 'thresh' is discovered for the continuous-valued data. Values less than *'thresh*' should be kept to the left, and more than 'thresh' should be kept to the right for every possible scenario. The *'thresh'* represents the cut point. The best cut point $thresh_f$ for any continuous-valued feature, 'F', is to select a value from its range.

This research applies the entropy approach presented to discretize continuous feature values. Due to the continuous nature of the attribute values, the ranges were transformed into discrete values using the entropy-based discretization approach **[12]**. With its assistance, the classification-based on association technique is simple to apply.

Rule Generation And Building Classifier

The classification-based association (CBA) rule generation technique detects all rule items by iteratively reviewing the data. In the first phase, the number of people who agree with each rule item is counted, and then it is determined whether or not that rule item is frequently used. It begins with the rule items discovered to be prominent in the first pass and continues with subsequent passes. This set of potential rule items generates new rule items that will most likely be used frequently. The actual support for these candidate rule items is determined during the data pass. The end of the iteration determines which candidate rule items are standard. It constructs class association rules using this collection of frequently used rule objects. The final step in the process of modifying these rules is rule pruning. C4.5's pruneRules function employs a pessimistic error rate-based pruning technique.

Rule items that frequently appear must be employed in order to meet the minSup. The rules that match MinConf are the most accurate. Rule items that meet the same requirements and the highest level of confidence are considered potential rules (PRs). Class associative rules (CAR_S) are both frequent and accurate. To find the most common k-rules, look at the ruleitems with ruleSup>= minSup. The details can be studied from the article **[13]**. The *ruleitems* are the set of conditions of antecedents and the class labels.

The Apriori algorithm **[14]** is used in the generation of candidates. It must increase both the condset and ruleitem support counts separately, whereas the algorithm apriori affects only one count. This support count allows us to assess the confidence of the rule item.

To determine the most frequent-1 ruleitems, the rule generation technique counts the number of occurrences of each item and class. This set of rules constitutes the CAR set known as CAR1. CAR1 is cut using a pruning process. The C4.5 **[15]** prune Rules function implements a pessimistic error rate-based pruning strategy. In the (k-1)th pass, the candidate rule items are produced from the often discovered rule items F_{k1}. C_k. searches the database and modifies the totals of support for each contender. Then, new frequently recurring rule items were discovered, making F_k outdated. The CAR_k rules are then generated by the algorithm. Finally, these rules are submitted to rule cutting.

Building Classifier

The best classifier must be able to generate rules from all available data. It would have to be evaluated using every feasible subset of the training data. This section demonstrates how to construct a classifier with CARs using the CBA-CB (Classifier Building) **[16]** technique based on pruning CARs—choosing the subset with the minimal number of errors and the correct rule sequence. Before proceeding with the algorithm, one needs to understand the order of the rules.

Sorting in precedence order: - There are two rules: R_I comes first if confidence in R_I is more significant than in R_J. If both are equal, R_I comes first, but support for R_I is more significant than support for R_J. Even though both confidence and support are similar, R_I arrived first. Using these criteria, the rules are sorted in order of importance. The detailed procedure can be obtained from article **[17]**.

The classifier is of the following format

$< R_1, R_2, ..., R_N, \text{DefaultClass} >$

Experimental Evaluation

In this section, the experiments carried out to compare and evaluate the built classification system are described. For collecting the training sample manually, 50 numbers of pixels from all the four class groups including the background has been chosen.

Data Set Used For Classification

The model is designed to predict the labels on an x-ray image of teeth. In total, there are 89 samples present in the dataset. Of those, only 24 samples are chosen randomly, of which there are four classes. The significance of the three classes are shown in Fig 2. (regular teeth, hole in teeth, and Background). The fourth class is the wire, which is optional and always present at the top portion of the samples, as shown in one sample in Fig 3. All the samples do not have wires. There are several bands through which x-ray images can be captured from a patient's teeth, such as Bitewing X-Ray, Periapical X-Ray, Occlusal X-Ray, Panoramic X-Ray, Cephalometric Projection, and Cone Beam X-ray. Out of these, only four bands of x-ray images are present in the collected dataset, and one set of samples is shown in Fig 4. The selected 24 samples out of the whole 89samples are current in four different bands.

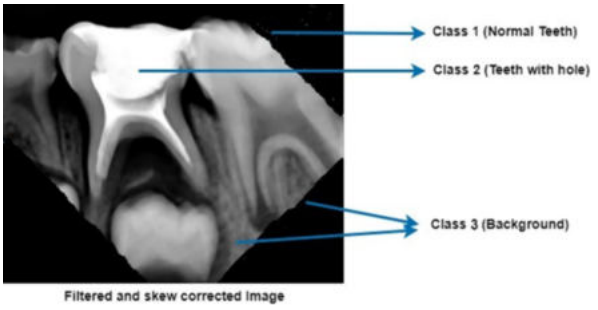

Fig 2. Three Classes Showing Normal Teeth, Hole Or Cavity And The Background Itself.

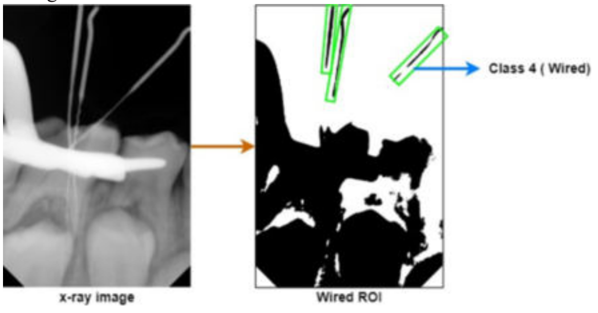

Fig 3. The Fourth Class (wire) Present At The Top Of The Teeth.

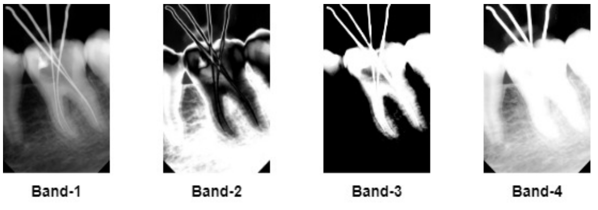

Fig 4. The Samples From Same Teeth Taken Through Different Four Bands.

Classification

Initially, the samples of the datasets are resized into 32x64 dimensions irrespective of the patients and the mode of taken x-ray images. Then the required percentages of training samples are chosen from both object and background randomly. Here the objects are four, one for the normal teeth, second for the cavity or hole present, the third one is the background itself, and fourth is the penetrated wire for the root canal. First few coordinates are selected from the four set of images of four bands. It means a total of 16 images are chosen for preparing the classification rules. The grey label intensity values of corresponding coordinate points from four different regions of those selected 16 images are being extracted based on each of the band of images. The intensity values are appended based on the image band. It means for each of image we have 4x32x64 number of intensity values. Similarly the dataset would have four columns with 4x32x64 number of intensity values with respect to their four bands.

These training samples are consisting of the gray label intensity values of the different images of different bands and with their respective class values (three kinds of objects and one background). Therefore, the classification-based on association technique will have more features if there are more bands. Using an entropy-based discretization technique similar to, these intensity values are converted into ranges. Class-based association **[16]** was utilised to create class association rules (CARs). As was already said, sorting happens after CAR creation. These CARs are used to construct the necessary classifier. Rule set pruning **[15]** is a different method for reducing classifier size. During this, each band in the image is separated using discretizing labels that were taken from training data. These discretized unlabeled patterns are then applied to the classifier. By the help of this classifier the whole image is classified into available maximum four regions including the background.

RESULT DISCUSSION

Out of the 24 randomly chosen samples, only four samples are used for the generation of class association rules to develop the classifier. There after remaining 20 samples are used as unknown samples for testing. The Table 1 describes the evaluation report for band 1. Similarly, for all the remaining three bands of images have been evaluated. However, only the results of the band 1 are presented. Here the T-Class stands for the number of points of true class present inside the samples and P-class stands for number of correctly classified predicted class by the proposed classifier.

After collecting the results for all the 20 samples for each of four the bands, their accuracy has been computed individually. There after the average accuracy has been computed, this is 92.22%. Although the result is not so high, it is still promising**.**

Samples	T-Class1	P-Class1	T-Class2	P-Class2	T-Class3	P-Class3	T-Class4	P-Class4
1	369	360	834	796	420	411	425	418
2	372	368	417	412	837	835	422	418
3	360	354	815	813	432	430	441	437
4	276	271	917	915	440	436	415	410
5	548	541	376	372	679	675	445	439
6	445	442	468	465	677	675	458	455
7	703	701	463	461	367	366	615	612
8	1023	1011	413	412	105	103	507	505
9	399	396	786	785	492	491	371	368
10	297	295	839	834	763	760	149	145
11	655	650	399	396	298	297	696	691
12	729	728	492	490	658	650	169	166
13	476	471	428	426	387	385	757	755
14	398	396	289	288	762	761	599	592
15	656	650	376	375	268	264	748	744
16	971	971	452	450	376	371	249	246
17	444	441	517	512	481	478	606	602
18	485	482	617	613	510	507	436	433
19	696	692	217	210	843	841	292	292
20	1004	998	378	375	428	426	238	237

CONCLUSION

Cavities, sometimes called dental caries, are the most prevalent disease worldwide. Almost everyone has had this ailment at some point in their lives. As medical imaging technology grows more effective and quick to use, clinical applications increasingly impact patient care. In this work, we propose a novel approach for localizing and identifying dental caries on X-ray images as a dataset and associative classification as a classification method. In this approach, classification rules and associative rules are both combined. Additionally, the numerical discretization method is used in the procedure. The study is the first attempt to identify the locations of dental cavities and root canal treatment using association-based classifications. This method was examined using actual patient data from hundreds of patients and was found to be highly successful at identifying unforeseen tooth damage. The chosen samples have four different class groups. The number of pixels in each group was computed and compared with the result produced by the proposed classifier. The proposed classifier has achieved an overall average accuracy of 92.22%. In the future, the model needs to be tested on a large scale, and further modifications are required to enhance its recognition accuracy.

REFERENCES

[1] Geetha, V., & Aprameya, K. S. (2019). Textural analysis based classification of digital X-ray images for dental caries diagnosis. *Int J Eng Manuf (IJEM)*, *9*(3), 44-5.

[2] Haghanifar, A., Majdabadi, M. M., & Ko, S. B. (2020). Paxnet: Dental caries detection in panoramic x-ray using ensemble transfer learning and capsule classifier. *arXiv preprint arXiv:2012.13666*.

[3] M. P. Muresan, A. R. Barbura and S. Nedevschi, "Teeth Detection and Dental Problem Classification in Panoramic X-Ray Images using Deep Learning and Image Processing Techniques," *2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2020, pp. 457-463, doi: 10.1109/ICCP51029.2020.9266244.

[4] Awais, M., Ghayvat, H., Krishnan Pandarathodiyil, A., Nabillah Ghani, W. M., Ramanathan, A., Pandya, S., ... & Faye, I. (2020). Healthcare professional in the loop (HPIL): classification of standard and oral cancer-causing anomalous regions of oral cavity using textural analysis technique in autofluorescence imaging. Sensors, 20(20), 5780.

[5] Gonsalves, P. (2017). Diagnosis of dental cavities using image processing.

[6] Huang, R., Liu, J., Chen, H., Li, Z., Liu, J., Li, G., ... & Wang, J. (2018). An effective fault

diagnosis method for centrifugal chillers using associative classification. *Applied Thermal Engineering*, *136*, 633-642.

[7] Atzmueller, M., Hayat, N., Trojahn, M., & Kroll, D. (2018, January). Explicative human activity recognition using adaptive association rule-based classification. In *2018 IEEE International Conference on Future IoT Technologies (Future IoT)* (pp. 1-6). IEEE.

[8] Khowaja, S.A., Yahya, B.N. & Lee, SL. CAPHAR: context-aware personalized human activity recognition using associative learning in smart environments. *Hum. Cent. Comput. Inf. Sci.* **10,** 35 (2020). https://doi.org/10.1186/s13673-020-00240-y

[9] González-Patiño, D., Villuendas-Rey, Y., Argüelles-Cruz, A. J., Camacho-Nieto, O., & Yáñez-Márquez, C. (2020). AISAC: An artificial immune system for associative classification applied to breast cancer detection. *Applied Sciences*, *10*(2), 515.

[10] Derner, M., Dehnen, G., Chaieb, L., Reber, T. P., Borger, V., Surges, R., ... & Fell, J. (2020). Patterns of single-neuron activity during associative recognition memory in the human medial temporal lobe. *Neuroimage*, *221*, 117214.

[11] Dey, R. (2012). *Remotely-Sensed Image Classification and Change Detection using Associative Classification* (Doctoral dissertation).

[12] Lee, C. H. (1997). Discretization of Continuous-Valued Attributes for Classification Learning. *The Transactions of the Korea Information Processing Society*, *4*(6), 1541-1549.

[13] Frank, E., & Witten, I. H. (1998). Generating accurate rule sets without global optimization.

[14] Han, J., & Kamber, M. (2000). Data Mining: Concepts and Techniques, The Morgan Kaufinann Series in Data Management Systems, Jim Gray, Series Editor Morgan Kaufmann Publishers. *Retrieved on April*, *10*, 2020.

[15] Salzberg, S. L. (1994). Book Review: C4. 5: Programs for Machine Learning Morgan Kaufmann Publishers by J. Ross Quinlan. Inc., 1993. *Machine Learning,© KluwerAcademic Publishers, Boston. Manufactured In The Netherlands*, *16*, 235-240.

[16] Ma, B. L. W. H. Y., Liu, B., & Hsu, Y. (1998, August). Integrating classification and association rule mining. In *Proceedings of the fourth international conference on knowledge discovery and data mining* (Vol. 50, pp. 55-00).

[17] Frank, E., & Witten, I. H. (1998). Generating accurate rule sets without global optimization.