Enhancing prediction of tooth caries using PCA features and multi-model classifier (#87712)

First submission

Guidance from your Editor

Please submit by **28 Jul 2023** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files
Download and review all files from the materials page.

3 Figure file(s)
2 Latex file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Enhancing prediction of tooth caries using PCA features and multi-model classifier

Shtwai Alsubai Corresp. [1]

[1] Computer Scince, Prince Sattam Bin AbdulAziz University, Al-Kharj, Al-Kharj, Saudi Arabia

Corresponding Author: Shtwai Alsubai
Email address: sa.alsubai@psau.edu.sa

Tooth decay, also known as dental caries, is a prevalent oral health issue that requires timely detection and management to prevent further complications. It is a chronic disease characterized by the demineralization and destruction of the tooth's hard tissues, primarily caused by the interaction between bacteria and dietary sugars. ~~It is a chronic disease characterized by the demineralization and destruction of the tooth's hard tissues, primarily caused by the interaction between bacteria and dietary sugars.~~ Many existing works focus on this research with image-based data but the results are not quite satisfying. This study considers feature-based datasets along with the use of Principle Component Analysis (PCA) and Chi-square (Chi-2) for feature engineering. In the proposed model, features are generated using PCA, utilizing a voting classifier ensemble consisting of Extreme Gradient Boosting (XGB), Random Forest (RF), and Extra Trees Classifier (ETC) algorithms. Extensive experiments are performed in comparison with the proposed approach using Chi-square features and machine learning models to evaluate the efficacy of the proposed approach for tooth caries detection. It is found that the proposed voting classifier using PCA features outperforms and achieves an accuracy, precision, recall, and F1 score of 97.36%, 96.14%, 96.84%, and 96.65% respectively.

Enhancing Prediction of Tooth Caries Using PCA Features and Multi-Model Classifier

Shtwai Alsubai[1]

[1]**Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; Sa.alsubai@psau.edu.sa**

Corresponding author:
Shtwai Alsubai[1]

Email address: Sa.alsubai@psau.edu.sa

ABSTRACT

Tooth decay, also known as dental caries, is a prevalent oral health issue that requires timely detection and management to prevent further complications. It is a chronic disease characterized by the demineralization and destruction of the tooth's hard tissues, primarily caused by the interaction between bacteria and dietary sugars. ~~It is a chronic disease characterized by the demineralization and destruction of the tooth's hard tissues, primarily caused by the interaction between bacteria and dietary sugars.~~ Many existing works focus on this research with image-based data but the results are not quite satisfying. This study considers feature-based datasets along with the use of Principle Component Analysis (PCA) and Chi-square (Chi-2) for feature engineering. In the proposed model, features are generated using PCA, utilizing a voting classifier ensemble consisting of Extreme Gradient Boosting (XGB), Random Forest (RF), and Extra Trees Classifier (ETC) algorithms. Extensive experiments are performed in comparison with the proposed approach using Chi-square features and machine learning models to evaluate the efficacy of the proposed approach for tooth caries detection. It is found that the proposed voting classifier using PCA features outperforms and achieves an accuracy, precision, recall, and F1 score of 97.36%, 96.14%, 96.84%, and 96.65% respectively.

INTRODUCTION

Maintaining good oral health is crucial for overall well-being and enhancing one's quality of life. Oral health encompasses the absence of throat cancer, mouth infections, sores, tooth decay, dental loss, gum problems, and related issues leading to impede activities such as chewing, speaking, smiling, biting, and overall psychosocial well-being. Dental caries are amongst the most common oral health issues, arising from residual food particles adhering to the teeth and leading to calcification. This process results in the teeth becoming porous, hollow, and occasionally fractured. Dental caries affect the enamel, cementum, and dentin, which are the solid tissues of the teeth, manifesting as decayed regions on the teeth. The mineral surface is gradually dissolved leading to the dental caries, eventually progressing inward.

Almost half of the global population is affected by oral diseases in one way or the other, with over 2.2 billion individuals worldwide having dental caries of permanent nature James et al. (2018). Tooth decay (dental caries), is a condition having affected teeth by oral bacteria. The enamel layer of the tooth is directly affected by lactic acid secreted by these bacteria. If left untreated, this can lead to the formation of small gaps between the teeth, resulting in pain, infection, and potential tooth loss Pihlstrom and Tabak (2005); Selwitz et al. (2007); Ogden et al. (2002). It is impossible to identify all caries lesions manually or visually. The detection rate is enhanced by deploying dental imaging techniques Gomez (2015); Metzger et al. (2022); Michou et al. (2022). Early-stage caries can still be missed even after the use of the latest imaging techniques. Hence, the effectiveness and accuracy of the diagnosis depend heavily upon the expertise of the dentists and the equipment used to diagnose the problem Topping and Pitts (2009).

Dental caries, despite being preventable and treatable, are frequently associated with tooth loss and pain. Effective, appropriate, and timely treatment depends upon early detection of these caries. Different equipment is used to identify dental cavities. Dental probes and handheld mirrors are used to visually inspect and probe the visible dental cavities. These orthodox methods prove useful in detecting accessible but partially obscured caries Fejerskov et al. (2015). However, for hidden or inaccessible lesions, X-ray radiography plays an irreplaceable role in diagnosis. Periapical, bitewing, and Panoramic X-rays are commonly utilized radiographs in clinical dental practice. Periapical and Bitewing X-rays are to detail a specific oral area by focusing on it, while panoramic X-rays capture the entire maxillofacial region Kaur et al. (2017). Bitewing radiography is commonly used to detect caries lesions and their depths with its high sensitivity and specificity Schwendicke et al. (2015). However, it cannot provide a complete assessment of all mouth lesions in a single attempt. On the other hand, panoramic imaging, which is taken outside the mouth, offers advantages such as lower infection rates, better patient approval, and reduced radiation exposure Rushton et al. (1999). Due to its cost and efficacy, panoramic imaging is a widely used and accepted radiological tool for diagnosis, clinical dental disease screening, and treatment evaluation.

Various methods have been explored for the detection of dental caries in previous studies. These include the trans-illumination method Datta et al. (2019), the international caries detection and assessment system Sinton et al. (1997), calibrated diaphragm computed tomography Kaur et al. (2017); Schwendicke et al. (2015), and quantitative light-induced fluorescence. Additionally, research has been conducted on panoramic radiology procedures to improve its performance for dental caries diagnosis through image processing techniques Abreu Jr et al. (2001). Computer-aided caries detection using X-rays has also been investigated Neuhaus et al. (2015); Alammari et al. (2013); Tracy et al. (2011). In one study, Tracy et al. utilized the LCD (Logicon Caries Detector) with modified computer-aided design (CAD) software, and the effectiveness of density analysis was confirmed as an auxiliary tool to assist dentists in detecting and organizing caries based on user feedback Alammari et al. (2013). Early dentinal diagnostic tools are employed for caries and this enabled dentists to diagnose the dentinal cavities much earlier than using the orthodox and conventional methods.

In recent decades, there has been a growing interest among scientists in utilizing machine learning techniques for the detection of dental diseases. Traditionally, the evaluation of radiographs and lesion detection is performed physically and subjectively by operators or experts. However, when dealing with large volumes of image data, this task can become tedious and prone to misinterpretations. Computer-aided diagnostics (CAD) using machine learning gaining the attention of many health professionals for the rapid and precise detection of dental diseases. This study is also a step toward detection of the dental disease using CAD systems. this study made the following contributions.

- This research work provides a framework that is contrived to perform tooth caries detection. To improve the effectiveness of tooth caries detection, the proposed model combines a voting classifier with features extracted using PCA.
- In addition, various machine learning models are utilized in this regard like decision tree (DT), extreme gradient boosting (XGB), logistic regression (LR), stochastic gradient descent (SGD), random forest (RF), extra tree classifier (ETC), support vector machine (SVM), gaussian naive bayes (GNB) and a voting classifier combining XGB, RF, and ETC aka VC(XGB+RF+ETC).
- The efficacy of the proposed approach is analyzed with two feature extraction techniques PCA and chi-square. The performance of the Voting classifier in combination with PCA features are analyzed in terms of accuracy (A), precision (P), recall (R), and F1 score (F).

The rest of the sections of this paper are arranged as described as under. Section details the relevant literature related to the current study. In Section , the dataset description, the proposed model, the methodology, and the ML models employed in this study are presented. The outcomes of the paper are presented in Section . Finally, in Section , the study is concluded, and forthcoming research directions are outlined.

RELATED WORK

The combination of data mining and machine learning represents a powerful tool for addressing various challenges. Analyzing potentially large medical data manually is particularly difficult because of its extensive feature vector. Machine learning has emerged as a crucial technique in numerous application domains, including healthcare. It has demonstrated its significance by offering precise and accurate systems for applications related to medicine, even during sensitive data handling in the medical domain.

Similarly, ML models have been successfully utilized in recognizing early-stage hazards associated with conditions such as dental caries.

In a study conducted by ~~Kang et al.~~ Kang et al. (2023), a dental caries detection model was proposed. This approach involved integrating the mRMR and GINI algorithms with the GBDT (Gradient Boosting Decision Tree) classifier. By utilizing only a few clinical test features, this method aimed to save time and cost during dental caries screening. The planned method was matched to other recently recommended dental processes. Among the different classifiers, the highest classification performance was demonstrated by GBDT having a reduced feature set. It achieved precision, impressive accuracy, F1-score, and recall values of 99%, 95%, 93%, and 88% respectively. In their research, Thanh et al. Thanh et al. (2022) put forward a Deep Learning-based system for detecting dental caries by using photos taken by a smartphone from inside the mouth. They employed four deep learning models, namely RetinaNet, YOLOv3, Faster R-CNNs, and Single-Shot Multi-Box Detector (SSD), to identify early-stage cavities and caries lesions. This study indicated that YOLOv3 achieved the highest sensitivity value, specifically 87.4%. This suggests that YOLOv3 exhibited strong performance in accurately detecting dental caries lesions in the intraoral photos captured by smartphones.

Lian et al. Lian et al. (2021) offered a system based on deep learning for the detection and grouping of caries. The dataset in the study was self-annotated. The authors applied a convolutional neural network (nnU-Net) to detect caries lesions and DenseNet121 to classify the lesions based on their depths. The performance of DenseNet121 and nnU-Net models was compared with outcomes from 6 dentists on the test dataset using various evaluation metrics, including Dice coefficient, intersection over union (IoU), recall, accuracy, negative predictive value (NPV), precision, and F1-score. This showed that nnU-Net achieved caries lesion segmentation Dice coefficient and IoU values of 0.663 and 0.785 respectively. The recall and recall rate of nnU-Net were 0.821 and 0.986, respectively. This shows the effectiveness of the nnU-Net model in accurately segmenting caries lesions. Oztekin et al. Oztekin et al. (2023) utilized panoramic radiograph images for dental caries prediction. They employed deep learning models, specifically DenseNet-121, EfficientNet-B0, and ResNet-50. The study's results revealed that the deep learning model ResNet-50 achieved an accuracy score of 92% in predicting dental caries.

Ghamdi et al. AL-Ghamdi et al. (2022) proposed a Neural Search Architecture Network (NASNet) for tooth caries detection. They also equated the performance of NASNet against AlexNet and CNN prototypes. The results of the study indicated that the proposed NASNet model outperformed rest of the deep learning models with respect to accuracy. It detected tooth caries with 96.51% accuracy score, which is by far very impressive. Jader et al. Jader et al. (2018) steered a study using 1500 panoramic X-ray radiographs to develop a tool profile using a convolutional neural network based on mask region. The evaluation of the tool was based on precision, accuracy, recall, F1-score, and specificity as outcome metrics. The results of the investigation showed impressive performance metrics, with recall, F1-score, accuracy, precision, and specificity values of 0.84, 0.88, 0.98, 0.94, and 0.99, respectively.

Muramatsu et al. Muramatsu et al. (2021) utilized a fourfold cross-validation technique with a dataset having 100 dental radiographs of panoramic nature to develop an object detection network. The performance of the network was evaluated based on sensitivity and accuracy for tooth detection. The study reported a sensitivity of 96.4% and an accuracy of 93.2% for tooth detection. Raith et al. Raith et al. (2017) employed a Convolutional Neural Network (CNN) architecture along with the PyBrain package to identify teeth. They achieved a performance score of 0.93, indicating a high level of accuracy in accurately identifying teeth using their proposed approach. Kuhnisch et al. Kühnisch et al. (2022) utilized a model based on deep learning, specifically a Convolutional Neural Network (CNN), for revealing and classification of caries. They compared the diagnostic performance of their model with expert standards in various scenarios. The outcome of the study suggested that the CNN prototype achieved an accuracy score of 92.5%.

The complete summary of the related work is shown in Table 1.

MATERIALS AND METHODS

This section explains the methodology and procedures employed in this study in detail. Figure 1 depicts the architecture of the suggested technique. Starting with data retrieval, the approach follows the generation of text by SMOTE to balance the dataset. Feature extraction is then carried out that involves term frequency (TF), term frequency-inverse document frequency (TF-IDF), and CNN. The data is split for training and testing where the selected machine learning models are utilized for sentiment classification.

Table 1. Summay of the related work.

Reference	Classifiers used	Dataset used	Achieved accuracy
Kang et al. (2023)	GINI, mRMR, GBDT	Korea Centers for Disease Control and Prevention dataset, 2018	95% GBDT
Thanh et al. (2022)	R-CNNs, RetinaNet, Faster YOLOv3, and Single-Shot Multi-Box Detector (SSD)	Self made using mobile camera	87.4% sensitivity using YOLO3
Lian et al. (2021)	nnU-Net, DenseNet121	Stomatology Hospital dataset	0.986, nnU-Net
Oztekin et al. (2023)	EfficientNet-B0, DenseNet-121, and ResNet-50	Firat University Faculty of Dentistry, dataset	92.00%, ResNet-50
AL-Ghamdi et al. (2022)	CNN, AlexNet, NASNet Model	Noor Medical Imaging Center, dataset (kaggle)	NASNet Model 96.51%
Jader et al. (2018)	Mask R-CNN	Silva et al dataset Silva et al. (2018),	98% accuracy using Mask R-CNN
Muramatsu et al. (2021)	CNN with single and multiple data inputs	Asahi University Hospital, dataset	93.2%
Raith et al. (2017)	ANN, CNN with PyBrain package	Ludwig-Maximilians-University Munich dataset	93% CNN
Kühnisch et al. (2022)	CNN (using different amount of data e. g 25%, 50%, 75% and 100% of the dataset)	Self made, using Nikon D7100, D300, or D7200 with a Nikon Micro 105-mm lens	92.5%
Moutselos et al. (2019)	Mask RNN for object detection and DNN for segmentation	Self made	88.9% mask RNN

Tooth caries dataset

The data used in this research work originates from the Oral Health Survey carried out by dentists at the university. Dentists performed oral examinations in each university to conduct this survey and to get accurate results. 10,375 students participated in this survey, and a questionnaire about oral health awareness is chosen for data collection. This questionnaire uses the same attributes that were collected by Korean dentists in their research work Kang et al. (2023). This questionnaire comprised 43 items and a single label, covering various aspects such as place of residence, gender, age, tooth brushing frequency, snack frequency, oral care usage, oral health awareness, smoking experience, and behavior. The label for this dataset is ”act_caries.” It’s worth noting that since the data does not contain any personally identifiable information about patients, it does not require approval from an Institutional Review Board (IRB). For detailed descriptions of the questionnaire items, please refer to the table 2.

Proposed approach

The purpose of this study was to introduce an ML-based oral cavities detection approach utilizing classifiers for more accurate prediction of tooth caries. The diagram 1 illustrating the design of the projected method is presented below.

In this approach, the very crucial and prominent step is the feature selection. Chi-2 and PCA, the two unique techniques used to select the features, are employed for feature fusion. Following this, the dataset is divided into a testing ratio of 30% and training sets ratio of 70%. For classification purposes, an ensemble model is utilized, which combines the XGB, RF, and ETC classifiers using the soft voting criterion. Recall, Accuracy, F1 score, and precision are the performance evaluation metrics of any given model or approach.

Table 2. Dataset attributes details.

Variable	Description
act_caries	It represents the presence of the dental caries (Label)
Sido_No	It shows the area of residence of the respondent of the dental examination.
Region_No	It represents the region of residence of the subject.
Gender	It represents the gender of the respondent.
Prev_caries	It represents the previous history of dental caries.
Calculus	It represents respondent have tartar build-up
Fluorosis	It represents tooth speckle
Bleeding	It show the gingival bleeding
X1	It shows the awareness of the respondent about the dental and gum oral health
X2	It show the respondent dental treatment experience in the last year
X3	It shows the respondent experience of the needing dental treatment but not receiving treatment.
X4_1	It shows the tooth brushed before breakfast
X4_2	It shows the tooth brushed after the breakfast
X4_3	It shows the tooth brushed before lunch
X4_4	It shows the tooth brushed after lunch
X4_5	It shows the tooth brushed before dinner
X4_6	It shows the tooth brushed after dinner
X4_7	It shows the tooth brushed after snack
X4_8	It shows the tooth brushed before going to sleep
X4_9	Tooth not brushed
X5_1	Frequency of the dental floss usage
X5_2	Frequency of handle floss usage
X5_3	Frequency of mouth wash usage
X5_4	Frequency of electric tooth brush usage
X5_5	It represent oral care product usage (if any)
X6	It represent the toothpaste usage
X7	It represents the fluoride tooth paste usage
X8	It represent if any sticky snack eaten today?
X9	It represent if any sticky snack eaten yesterday
X10	It represent the gum bleeding of gum pain while brushing
X11	It shows the pain or discomfort in the tooth in the last 1 year.
X12	It shows parents are smoking or not
X13	It represents any smoking experience?
X14_1	It shows that the respondent living with grandfather
X14_2	It shows that the respondent living with grandmother
X14_3	It shows that the respondent living with father
X14_4	It shows that the respondent living with stepfather
X14_5	It shows that the respondent living with mother
X14_6	It shows that the respondent living with stepmother
X14_7	It shows that the respondent living with older sister/older brother
X14_8	It shows that the respondent living with younger sister/younger brother
X14_9	It shows the not living with the above mentioned family members
X15_1	It represents the house hold economic status
X16	It represents the weekly allowance

Proposed Model XGB+RF+ETC

To evaluate tooth caries detection, this study conducted experiments in three different scenarios: (i) using all features, (ii) using Chi2-selected features, and (iii) using PCA-selected features. The dataset was then split into a 70:30 ratio, with 70% allocated to training the models and the rest of the 30%

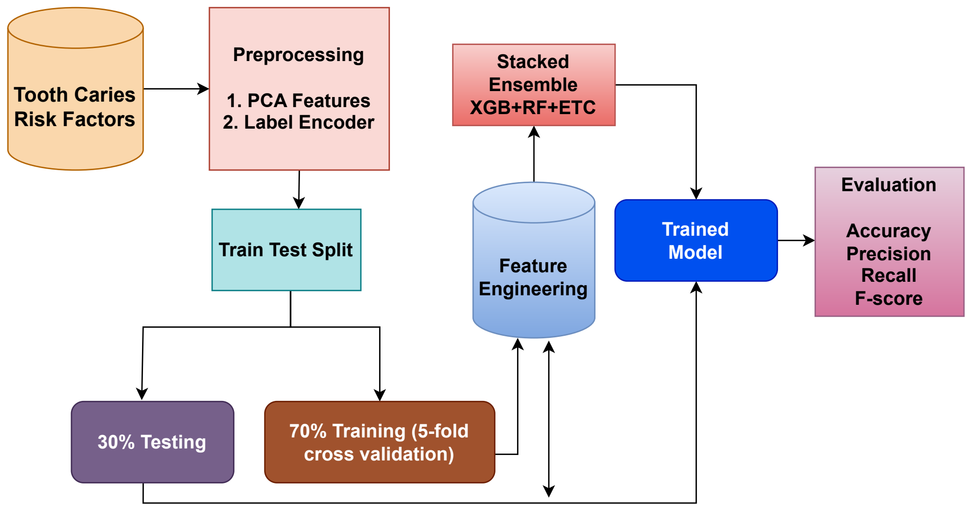

Figure 1. Tooth caries detection workflow methodology.

for prototype testing purposes. For the proposed tooth caries detection system, an ensemble approach called XGB+RF+ETC was utilized. Ensemble models are a powerful technique that combines predictions from multiple models to improve accuracy and robustness. Each model within the ensemble has its own strengths and weaknesses, and by combining them, the overall performance is often enhanced. In the case of tooth caries detection, this study suggests employing an ensemble learning model that integrates three popular algorithms: XGB (eXtreme Gradient Boosting), RF (Random Forest), and ETC (Extra Trees Classifier). By leveraging the strengths of these individual models and their ability to handle different aspects of the data, the ensemble approach aims to achieve better tooth caries detection results. The architecture of the proposed voting system is shown in Figure 2.

The ensemble model has been created by combining the predictions of three different ML algorithms: XGB, RF, and ETC. The general process for constructing an ensemble model involves training multiple models on the same dataset and then merging their predictions. In the case of the XGB+RF+ETC ensemble model, each algorithm is trained individually on the dataset. During training, each model generates predicted probabilities for the different classes of the target variable. These predicted probabilities represent the likelihood of each class for a given observation. To create the final prediction for each observation, the predicted probabilities from each model are combined. One common approach is to calculate a weighted average of the predicted probabilities, where the weights assigned to each model are determined based on their performance on a validation set. Models that demonstrate better predictive performance on the validation set are assigned higher weights in the ensemble combination. By training multiple models and combining their predictions, the ensemble model aims to improve overall performance and reduce the risk of overfitting. This approach leverages the strengths of each individual model and can lead to more accurate and robust predictions for tooth caries detection. The algorithm below details the working of the recommended ensemble model, which can be expressed as:

$$\widehat{p} = argmax\{\sum_{i}^{n} XGB_i, \sum_{i}^{n} RF_i, \sum_{i}^{n} ETC_i\}. \tag{1}$$

In the case of the $\sum_{i}^{n} XGB_i$, $\sum_{i}^{n} RF_i$, and $\sum_{i}^{n} ETC_i$ models, each of them produces prediction probabilities for each test sample. These probabilities represent the likelihood of each class for a given test case. To make a final prediction, the ensemble model utilizes the soft voting criterion. The soft voting criterion involves aggregating the probabilities generated by all three models. This aggregation process combines the predictions from XGB, RF, and ETC to arrive at a consolidated set of probabilities. Once the probabilities from all three models are combined, a final prediction is made based on these combined probabilities. For a test sample, the class having the highest probability is taken as the predicted class. By

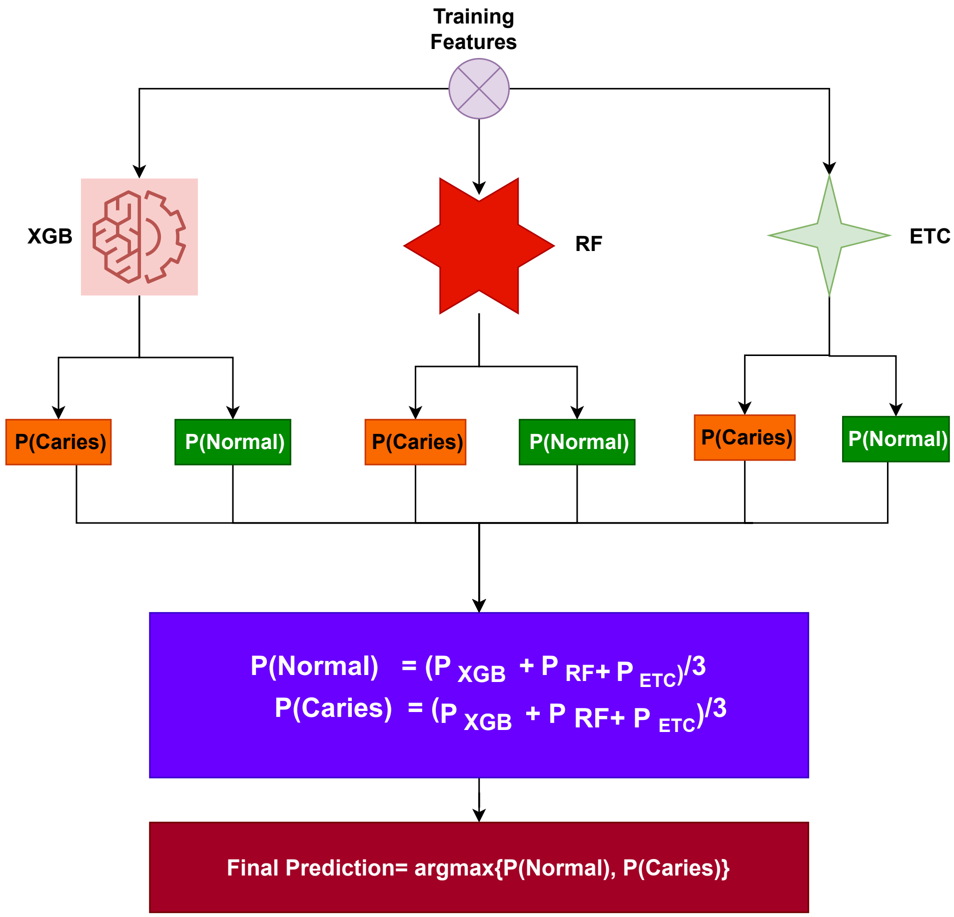

Figure 2. Proposed tooth caries detection framework architectural diagram.

employing the soft voting criterion and combining the predictions from multiple models, the ensemble model benefits from the collective insights and strengths of XGB, RF, and ETC. This approach enhances the robustness and accuracy of the final prediction for each test case in the tooth caries detection task.

In the ensemble model, the final class prediction is determined by selecting the class with the highest average probability score among the combined predictions of the classifiers. The individual classifiers in the ensemble generate probabilities for each class in the target variable. These probabilities are combined, typically through averaging, to obtain a single probability for each class across all the classifiers. After combining the probabilities, the class having the highest average probability score is selected as the final forecast. This means that class that receives the highest collective confidence from the classifiers is selected as the predicted class for a given observation.

Feature selection techniques

To ensure the machine learning model is trained with relevant features, feature selection methods are employed to extract and combine the particular features, resulting in a proficient feature set. The technique of Feature selection plays a crucial role in achieving a well-fitted machine learning model, as each feature holds importance in relation to the object class. So, an approach has been devised to incorporate only the features having a substantial effect on the final class prediction. This approach offers numerous advantages, such as improved interpretability of learning prototypes, reduced model variations, and decreased computational costs and training time. By identifying the ideal subset of features, the complication of the system is reduced, thereby enhancing the accuracy and stability of classification. In this study, chi-2 and PCA, are employed for this drive. Both feature selection techniques, chi-2, and PCA, aim to reduce the feature size while selecting the most relevant features. By doing so, they create a more appropriate feature set for the machine learning model.

Chi-square (Chi2)

The Chi2 feature selection method is widely utilized in machine learning, and it is also employed in the current study to identify the optimal features for model training Narra et al. (2022). The datasets used in the experiment consist of a large number of features, which can introduce complexity in the learning process of the models. To address this, only the best features selected through the Chi2 method are utilized to enhance the performance of the ML models. By focusing on the most informative features, the models can achieve improved performance and potentially overcome challenges associated with the large feature set. Chi2 used the following equation to compute the score:

$$X^2(D,t,c) = \sum_{c_t e[0,1]} \sum_{c_t e[0,1]} \frac{(Ne_t, e_c - Ee_t, e_c)^2}{Ee_t, e_c} \tag{2}$$

In the given context, the variables N and E represent the observed frequency and predicted frequency, respectively. The variable e_t is assigned a value of 1 when the text contains both ”t” and ”0”, and 0 otherwise. On the other hand, the variable e_c takes the value 1 if the document belongs to the ”c” class, and 0 if it belongs to any other class. A high Chi2 score for a feature suggests that the null hypothesis (H0) of independence between the feature and the document class should be rejected. This indicates that the feature and class are interdependent, meaning that the feature has a significant impact on the frequency of occurrence within each class. In the given scenario, if the caries feature exhibits a high Chi2 score, it implies that the feature is closely related to the document class and should be selected for training the model.

PCA (Principal Component Analysis)

PCA (Principal Component Analysis) is a linear feature selection method for the identification of the most relevant features from a given dataset. It is an unsupervised method that utilizes Eigenvector scrutiny to determine the pivotal genuine features of the principal components. The principal components are linear combinations of the observed features, weighted optimally Narra et al. (2022). Consequently, principal components are the outcome of the feature selection method of PCA, which typically has a reduced number of features compared to the genuine dataset. PCA feature selection is beneficial in various problem domains. However, it may not be preferred in cases where there is excessive multicollinearity, which refers to a high correlation among the features. In such situations, the interpretation and effectiveness of PCA may be compromised.

Supervised models for the tooth caries detection

With the increasing popularity of machine learning models, there is a wide range of variations available in the existing literature that can deliver good classification performance. Furthermore, tools like Sci-Kit offer user-friendly functions for implementing these models. In this research, multiple machine learning classifiers are employed to classify tooth caries. The classifiers used include LR (Logistic Regression), DT (Decision Tree), RF (Random Forest), SGD (Stochastic Gradient Descent), ETC (Extra Trees Classifier), XGB (XGBoost), SVC (Support Vector Classifier), and GNB (Gaussian Naive Bayes). For comprehensiveness, each of these models is briefed as under.

Decision Tree

The Decision Tree classifier is a straightforward ML algorithm that constructs relationship rules to identify and guess the particular labels Brijain et al. (2014); Zhang et al. (2014). It falls under the category of Supervised Machine Learning Algorithms. The Decision Tree starts by selecting a root node and then proceeds to traverse the tree, moving from the root node to the leaf nodes in order to make label predictions. The root node in a DT is primarily determined by two methods: Gini Index and Information Gain (IG). These methods assess the quality of potential splits in the data and choose the attribute that maximizes the information gain or minimizes the Gini index as the root node.

EXtreme Gradient Boosting

XGBoost (eXtreme Gradient Boosting) operates similarly to the Gradient Boosting classifier but additionally assigns weights to each sample, like the Adaboost classifier. XGBoost is also a tree-based model that has gained a significant reputation in recent years Zhang et al. (2014). It trains multiple weak learners, such as decision trees, in parallel, contrasting with GB, which does this chronologically. This parallel processing capability of XGBoost provides a speed boost compared to other boosting methods. XGBoost also offers L1 and L2-type regularization techniques, to control overfitting. Both the GB and the Adaboost lack this regularization technique Freund and Schapire (1997). Scalability is also an added benefit of XGBoost which allows it to function on distributed systems and process large datasets efficiently. In terms of the loss function, XGBoost employs the Log Loss function, which aids in minimizing the loss and improving accuracy. The Log Loss function takes into account the probability of false classifications, making it valuable for optimizing the model's performance.

Logistic Regression

LR (Logistic Regression), is a regression adaptive method that constructs predictors using a Boolean combination of binary covariate Sebastiani (2002). The name ”logistic regression” employs the core function deployed in this method, which is a sigmoid function. This function is characterized by an S-shaped curve, capable of turning a real-valued number into a value ranging between 0 and 1. Using the LR is suitable when we have a categorical dependent variable, making it an optimal choice for classification tasks. The logistic function can be computed as:

$$y = \frac{1}{(1 + e^{(-value)})} \tag{3}$$

Overall the logistic regression can be represented as;

$$y = \frac{e^{b_0 + b_1 \times x}}{1 + e^{(b_0 + b_1 \times x)}} \tag{4}$$

Stochastic Gradient Classifier

The SGD (Stochastic Gradient Descent) algorithm incorporates concepts from SVM (Support Vector Machine) and logistic regression, utilizing convex loss functions Zadrozny and Elkan (2002). It is a powerful classifier suitable for multi-class classification problems, employing the One-vs-All (OvA) approach by combining multiple binary classifiers. One of the notable advantages of SGD is its ability to handle large datasets efficiently. It achieves this by using a batch size of 1, processing only a single example per iteration. Additionally, SGD is relatively easy to understand and implement due to its basis in simple regression techniques. However, SGD does have some drawbacks. It can be quite noisy since the example chosen from the batch is random, and accurate results depend on correctly setting the hyperparameters. SGD is also highly sensitive to feature scaling, necessitating careful attention to scaling techniques for optimal performance.

Random Forest

RF (Random Forest) is indeed an alternative term for Random Decision Forest (RDF). It is a versatile algorithm used for various tasks, including grouping, regression, and related tasks constructing multiple DTs Gregorutti et al. (2017). RF is a supervised learning algorithm having an applicability advantage to both regressions as well as classification problems. One of the notable strengths of the RF algorithm is its high accuracy, often outperforming other existing systems. As a result, it has gained significant popularity and is widely adopted as one of the most utilized algorithms in machine learning.

Extra tree classifier:

ETC (Extra Trees Classifier) is an ensemble learning method that consists of randomized trees. It aggregates multiple decorrelated trees within a forest of decision trees to produce a final classification result or output Rustam et al. (2019); Safavian and Landgrebe (1991). The underlying concepts of both the ETC and RF are similar but contradictory in the way that DTs are constructed within a forest while in ETC, K best feature's random samples are used for decision making. The Gini index here is employed as a mathematical criterion for selecting the best feature for splitting the data in each tree. This methodology ensures that the constructed trees in ETC are decorrelated from each other. The selection of features in ETC is based on the Gini feature importance, which determines the relevance and significance of each feature in the classification process. By considering the Gini feature importance, ETC can identify and utilize the most informative features for making decisions within each tree.

Support Vector Classifier

SVC (Support Vector Classifier) is a famous supervised ML prototype utilized for both regression and classification tasks Cortes and Vapnik (1995). It employs a hyperplane to isolate and organize the data points. It aims to locate an ideal hyperplane to maximize the distance between the hyperplane and the sample points, effectively creating a clear separation between different classes. In scenarios where the data is non-linear and cannot be separated by a linear hyperplane, SVC employs a kernel trick. This technique maps the input features into a higher-dimensional space, where the data becomes linearly separable. By transforming the data into a higher-dimensional space, SVC can effectively classify non-linear data by finding an appropriate hyperplane in the transformed feature space.

Gaussian Naive Bayes

GNB (Gaussian Naive Bayes) is a type of the Naive Bayes algorithm, based on Bayes' theorem. GNB predicts the result of an occurrence by using conditional probabilities Ahmed et al. (2023). In GNB, if a sample is classified into k categories, denoted as k = c1, c2, $\cdots$, ck, the resulting output is assigned to one of the classes, denoted as c. The GNB function can be represented as follows, where c represents the class and d represents the sample:

$$P(c|d) = (P(d|c) \times P(c))/P(d) \tag{5}$$

In this equation, the probability of class c given the sample d, is represented by P(c—d), P(d—c) is the probability of the sample d given class c, P(c) denotes the prior probability of class c, and P(d) represents the probability of the sample d. By calculating these probabilities, GNB predicts the most likely class for a given sample based on the principles of Bayes' theorem.

Evaluation parameters

To assess the performance of the ML models, four evaluation parameters are used: accuracy score, recall score, precision score, and F1 score. The accuracy score represents the percentage of correct predictions made by the prototype. To compute the accuracy, we divide the number of correct predictions by the total number of predictions made by the sample. The accuracy score ranges between 0 to 1, with 1 indicating a perfect prediction and 0 indicating no correct predictions. Mathematically,

$$Accuracy(A) = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

The precision score measures the proportion of true positive predictions (correctly identified positive cases) out of all positive predictions (both true positives and false positives). It focuses on the accuracy of positive predictions.

$$Precision(P) = \frac{TP}{TP+FP} \tag{7}$$

Recall score, also known as sensitivity or true positive rate, measures the proportion of true positive predictions out of all actual positive cases in the dataset. It emphasizes the model's ability to identify positive cases.

$$Recall(R) = \frac{TP}{TP+FN} \tag{8}$$

F1 score is the harmonic mean of precision and recall scores, providing a balanced evaluation metric. It considers both precision and recall, which is particularly useful when dealing with imbalanced datasets or when false positives and false negatives have different impacts.

$$F1\ score(F) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{9}$$

These evaluation parameters provide a comprehensive understanding of the model's performance, considering different aspects such as overall accuracy, precision, recall, and balance.

EXPERIMENTAL RESULTS

The results and outcomes of the tooth caries prediction are presented in this section. The machine learning models were implemented using Python 3.8 and executed within a Jupyter Notebook environment. The scikit learn and TensorFlow libraries were utilized for model development and evaluation. The experiments were conducted on a system running 64-bit Windows 10, featuring a 7th-generation Core i7 processor with a clock speed of 2.8 GHz. This information provides context regarding the hardware specifications used for the experiments. To assess the performance of the models, various evaluation metrics were employed. These include accuracy, precision, recall (also known as sensitivity), and F1 score. To evaluate tooth caries detection, this study conducted experiments in three different scenarios: (i) using original features, (ii) using Chi2-selected features, and (iii) using PCA-selected features.

Performance of ML Classifiers Using Original Features

In the first phase of the tests, experiments are performed on the original dataset features. Table 3 shows the results of the ML models using the original features.

Table 3. Results of the machine learning models obtained by original from the dataset

Model	A	**P**	**R**	**F**
LR	85.77	90.54	91.64	90.61
DT	87.24	90.51	90.45	90.27
RF	90.65	91.35	91.75	91.21
SGD	88.59	91.37	90.88	90.66
ETC	90.08	91.35	89.35	90.12
XGB	90.51	90.95	90.89	90.93
SVC	89.45	90.34	90.34	90.62
GNB	82.38	85.44	86.12	85.99
VC(XGB+RF+ETC)	92.03	92.46	92.21	92.33

The results of the experiments indicate that the proposed ensemble model VC(XGB+RF+ETC) outclasses the individual ML classifiers in accuracy and F1-score, achieving values of 92.03% and 92.33% respectively. While the performance of all the ML classifiers is generally close to the VC(XGB+RF+ETC) model, there is a noticeable difference in accuracy and F1-score compared to VC(XGB+RF+ETC). GNB, on the contrary, achieved the lowest accuracy score of 82.38% for the task. It demonstrates that the anticipated VC(XGB+RF+ETC) model exhibits superior performance across all evaluation metrics, including accuracy, precision, recall, and F1 score when compared to the other classifiers employed in the study.

Results Using Chi-2 Features

In the second set of experiments, the ML models were trained and tested using PCA-selected features. In the third set of experiments, the same models were trained and tested using Chi-2 selected features. Table 4 presents the grouping results when Chi-2 features are used to test and train the model. The results indicate that the performance of the ML classifiers improves when features of Chi-2 are used. The proposed ensemble model VC(XGB+RF+ETC) achieves the joint highest accuracy of 91.52%, which is a 0.51% decrease from the accuracy obtained using the original features. Similarly, other machine learning models exhibit improved performance when trained on Chi-2 features

Results Using PCA Features

Table 5 presents the grouping results when the classifiers are accomplished and evaluated on PCA features. The results demonstrate that the performance of the machine learning classifiers is improved when PCA features are used.

The proposed ensemble model VC(XGB+RF+ETC) achieves the top performance with a 97.36% accuracy score, which is a significant improvement of 5.33% compared to the accuracy obtained using the original dataset features and 5.84% compared to the Chi2 features. Additionally, the performance of GNB is also enhanced with PCA features. Moreover, SVC, LR, RF, SGD, XGB, GNB, ETC, and DT models exhibit substantial performance improvements when trained on PCA features. This suggests that PCA

Table 4. Results of the machine learning models obtained by chi-square features from the dataset

Model	A	P	R	F
LR	87.22	87.67	89.64	88.66
DT	88.13	88.35	88.29	88.33
RF	89.21	89.48	90.37	89.55
SGD	89.29	87.17	89.24	88.19
ETC	89.24	89.67	89.21	89.48
XGB	90.25	91.24	87.34	88.29
SVC	85.34	86.65	87.05	86.30
GNB	88.37	88.67	88.34	88.51
VC(XGB+RF+ETC)	91.52	90.25	91.61	90.47

feature selection enhances the performance of the machine learning classifiers, including the proposed ensemble model VC(XGB+RF+ETC), in terms of accuracy.

Table 5. Results of the machine learning models obtained by PCA features from the dataset

Model	A	P	R	F
LR	90.24	92.82	92.71	92.80
DT	92.35	92.62	92.34	92.47
RF	94.28	95.29	95.34	95.32
SGD	91.54	93.78	93.24	93.43
ETC	95.34	95.67	95.19	95.38
XGB	94.08	94.52	94.34	94.43
SVC	93.86	94.45	93.21	93.87
GNB	90.81	91.32	90.36	90.86
VC(XGB+RF+ETC)	97.36	96.14	96.84	96.65

Comparison of Machine Learning Models for all Experiments

To evaluate the efficacy of the suggested system, we performed a comparative evaluation of the performance of various machine-learning models across a range of experiments. The findings revealed a noteworthy improvement in the machine learning models' performance when utilizing the PCA features in the third experiment. Table 6 offers a comprehensive summary of the achieved outcomes by the ML models in all scenarios, facilitating a comprehensive assessment of their performance. This comparison provides a distinct understanding of the influence and advantages of incorporating PCA features in enhancing the predictive abilities of the models. The comparison of all three types of features results is shown in Figure 3.

Table 6. Accuracy Comparison of the machine learning models

Model	With all features	With chi-square features	With PCA features
LR	85.77	87.22	90.24
DT	87.24	88.13	92.35
RF	90.65	89.21	94.28
SGD	88.59	89.29	91.54
ETC	90.08	89.24	95.34
XGB	90.51	90.25	94.08
SVC	89.45	85.34	93.86
GNB	82.38	88.37	90.81
VC(XGB+RF+ETC)	92.03	91.52	97.36

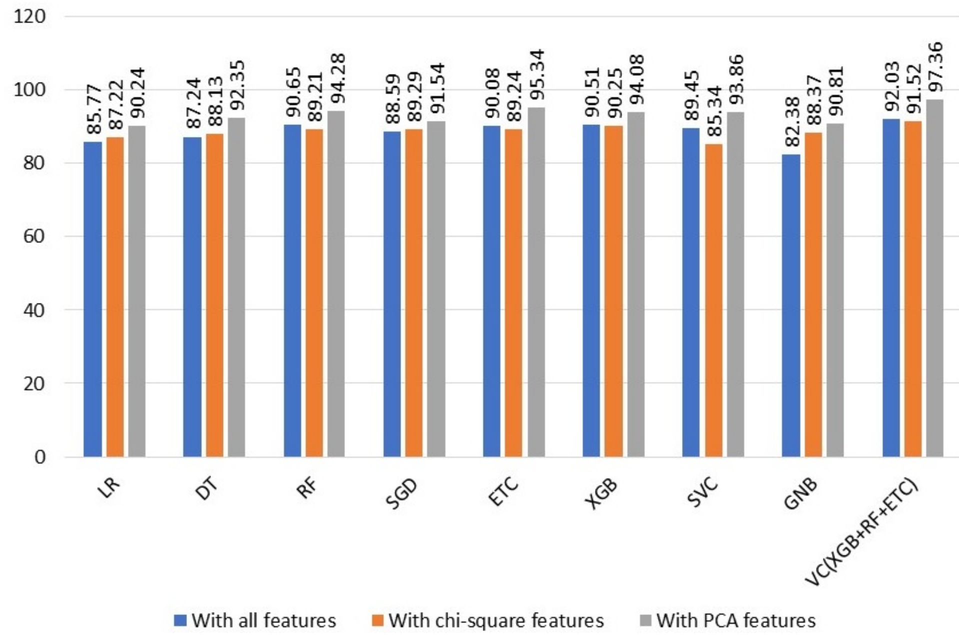

Figure 3. Comparison of all types of features results.

Results of the K-Fold Cross Validation

In order to establish the models' reliability, we utilized K-fold cross-validation. The outcomes of the 5-fold cross-validation, as depicted in Table 7, distinctly demonstrate the superiority of the suggested approach over other models in terms of recall, precision, F1 score, and accuracy. Furthermore, the suggested approach showcases a low standard deviation, underscoring its dependability and consistency. This implies that the proposed approach consistently delivers strong performance across multiple folds, instilling additional assurance in its reliability and resilience.

Table 7. 5- fold cross-validation results for the proposed system

Model	A	P	R	F
1st fold	97.52	96.13	96.61	96.12
2nd fold	97.25	96.34	96.74	96.23
3rd fold	98.64	97.67	98.98	98.81
4th fold	98.08	97.78	97.99	97.85
5th fold	96.98	96.15	96.86	96.33
Average	**97.89**	**96.81**	**96.23**	**96.87**

Comparison with State-of-the-art Approaches

To demonstrate and authenticate the effectiveness of the anticipated ensemble model and PCA feature approach, we conducted a performance comparison between VC(XGB+RF+ETC) and various state-of-the-art methods. The comparison results are presented in the table 8. The findings indicate that the proposed VC(XGB+RF+ETC) surpasses the performance of other modern approaches, achieving an impressive accuracy of 97.36%. This outperforms the previous best accuracy of 96.51% for tooth caries prediction, providing strong evidence of the efficacy and superiority of the proposed approach.

CONCLUSIONS

Prevention and early detection are key in managing dental caries. Maintaining good oral hygiene practices, including regular brushing with fluoride toothpaste, flossing, and using antimicrobial mouthwashes, helps remove plaque and reduce the risk of cavities. Additionally, adopting a balanced diet low in sugary foods

Table 8. Comparison of the proposed approach with best models from the literature

Ref	Classifier	Reported accuracy
Kang et al. (2023)	GINI, mRMR, GBDT	95.05%
Oztekin et al. (2023)	EfficientNet-B0, DenseNet-121, and ResNet-50	92.00%
AL-Ghamdi et al. (2022)	CNN, AlexNet, NASNet Model	96.51%
Muramatsu et al. (2021)	CNN with single and multiple data inputs	93.2%
Raith et al. (2017)	ANN, CNN with PyBrain package	93.29%
Kühnisch et al. (2022)	CNN (using different amount of data e. g 25%, 50%, 75% and 100% of the dataset)	92.50%
Moutselos et al. (2019)	Mask RNN for object detection and DNN for segmentation	88.9%
Proposed model	**VC(XGB+RF+ETC)**	**97.36%**

and beverages can minimize the exposure of teeth to acid-producing bacteria. Regular dental check-ups and professional cleanings are essential for the early detection and management of dental caries. Dentists can identify early signs of decay, provide necessary treatments such as fillings or dental sealants, and offer guidance on oral hygiene practices and dietary modifications. This research work provides a complete framework for automatic tooth caries detection. The first step of the proposed model is based on PCA feature extraction. In the second step, PCA-extracted significant features are fed to the ensemble voting classifier (XGB+RF+ETC). The results obtained show the superiority of the proposed model among all state-of-the-art models by giving an accuracy of 97.36%. The proposed model is also tested with other feature extraction techniques like chi-square and 8 other machine learning models. The future work of this research is to make the combination of machine and deep learning models as an ensemble to get more accurate results. The second future work direction is based on the usage of transfer learning models to augment data and better training.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Abreu Jr, M., Tyndall, D. A., and Ludlow, J. B. (2001). Effect of angular disparity of basis images and projection geometry on caries detection using tuned-aperture computed tomography. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 92(3):353–360.

Ahmed, S., Khan, D. M., Sadiq, S., Umer, M., Shahzad, F., Mahmood, K., Mohsen, H., and Ashraf, I. (2023). Temporal analysis and opinion dynamics of covid-19 vaccination tweets using diverse feature engineering techniques. *PeerJ Computer Science*, 9:e1190.

AL-Ghamdi, A. S., Ragab, M., AlGhamdi, S. A., Asseri, A. H., Mansour, R. F., and Koundal, D. (2022). Detection of dental diseases through x-ray images using neural search architecture network. *Computational Intelligence and Neuroscience*, 2022.

Alammari, M., Smith, P., de Jong, E. d. J., and Higham, S. (2013). Quantitative light-induced fluorescence (qlf): a tool for early occlusal dental caries detection and supporting decision making in vivo. *Journal of dentistry*, 41(2):127–132.

Brijain, M., Patel, R., Kushik, M., and Rana, K. (2014). A survey on decision tree algorithm for classification. *International Journal of Science and Research (IJSR)*.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

Datta, S., Chaki, N., and Modak, B. (2019). A novel technique to detect caries lesion using isophote concepts. *IRBM*, 40(3):174–182.

Fejerskov, O., Nyvad, B., and Kidd, E. (2015). *Dental caries: the disease and its clinical management*. John Wiley & Sons.

Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139.

Gomez, J. (2015). Detection and diagnosis of the early caries lesion. *BMC oral health*, 15(1):1–7.

Gregorutti, B., Michel, B., and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27(3):659–678.

Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., and Oliveira, L. (2018). Deep instance segmentation of teeth in panoramic x-ray images. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 400–407. IEEE.

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., et al. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1789–1858.

Kang, I.-A., Njimbouom, S. N., and Kim, J.-D. (2023). Optimal feature selection-based dental caries prediction model using machine learning for decision support system. *Bioengineering*, 10(2).

Kaur, R., Sandhu, R. S., Gera, A., and Kaur, T. (2017). Edge detection in digital panoramic dental radiograph using improved morphological gradient and matlab. In *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, pages 793–797. IEEE.

Kühnisch, J., Meyer, O., Hesenius, M., Hickel, R., and Gruhn, V. (2022). Caries detection on intraoral images using artificial intelligence. *Journal of dental research*, 101(2):158–165.

Lian, L., Zhu, T., Zhu, F., and Zhu, H. (2021). Deep learning for caries detection and classification. *Diagnostics*, 11(9):1672.

Metzger, Z., Colson, D. G., Bown, P., Weihard, T., Baresel, I., and Nolting, T. (2022). Reflected near-infrared light versus bite-wing radiography for the detection of proximal caries: A multicenter prospective clinical study conducted in private practices. *Journal of Dentistry*, 116:103861.

Michou, S., Vannahme, C., Bakhshandeh, A., Ekstrand, K. R., and Benetti, A. R. (2022). Intraoral scanner featuring transillumination for proximal caries detection. an in vitro validation study on permanent posterior teeth. *Journal of Dentistry*, 116:103841.

Moutselos, K., Berdouses, E., Oulis, C., and Maglogiannis, I. (2019). Recognizing occlusal caries in dental intraoral images using deep learning. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1617–1620. IEEE.

Muramatsu, C., Morishita, T., Takahashi, R., Hayashi, T., Nishiyama, W., Ariji, Y., Zhou, X., Hara, T., Katsumata, A., Ariji, E., et al. (2021). Tooth detection and classification on panoramic radiographs for automatic dental chart filing: improved classification by multi-sized input data. *Oral Radiology*, 37:13–19.

Narra, M., Umer, M., Sadiq, S., Karamti, H., Mohamed, A., Ashraf, I., et al. (2022). Selective feature sets based fake news detection for covid-19 to manage infodemic. *IEEE Access*, 10:98724–98736.

Neuhaus, K., Jost, F., Perrin, P., and Lussi, A. (2015). Impact of different magnification levels on visual caries detection with icdas. *Journal of dentistry*, 43(12):1559–1564.

Ogden, C. L., Kuczmarski, R. J., Flegal, K. M., Mei, Z., Guo, S., Wei, R., Grummer-Strawn, L. M., Curtin, L. R., Roche, A. F., and Johnson, C. L. (2002). Centers for disease control and prevention 2000 growth charts for the united states: improvements to the 1977 national center for health statistics version. *Pediatrics*, 109(1):45–60.

Oztekin, F., Katar, O., Sadak, F., Yildirim, M., Cakar, H., Aydogan, M., Ozpolat, Z., Talo Yildirim, T., Yildirim, O., Faust, O., et al. (2023). An explainable deep learning model to prediction dental caries using panoramic radiograph images. *Diagnostics*, 13(2):226.

Pihlstrom, B. L. and Tabak, L. (2005). The national institute of dental and craniofacial research: research for the practicing dentist. *The Journal of the American Dental Association*, 136(6):728–737.

Raith, S., Vogel, E. P., Anees, N., Keul, C., Güth, J.-F., Edelhoff, D., and Fischer, H. (2017). Artificial neural networks as a powerful numerical tool to classify specific features of a tooth based on 3d scan data. *Computers in biology and medicine*, 80:65–76.

Rushton, V., Horner, K., and Worthington, H. (1999). The quality of panoramic radiographs in a sample of general dental practices. *British dental journal*, 186(12):630–633.

Rustam, F., Ashraf, I., Mehmood, A., Ullah, S., and Choi, G. S. (2019). Tweets classification on the base of sentiments for us airline companies. *Entropy*, 21(11):1078.

Safavian, S. R. and Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674.

Schwendicke, F., Tzschoppe, M., and Paris, S. (2015). Radiographic caries detection: a systematic review and meta-analysis. *Journal of dentistry*, 43(8):924–933.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1):1–47.
Selwitz, R. H., Ismail, A. I., and Pitts, N. B. (2007). Dental caries. *The Lancet*, 369(9555):51–59.
Silva, G., Oliveira, L., and Pithon, M. (2018). Automatic segmenting teeth in x-ray images: Trends, a novel data set, benchmarking and future perspectives. *Expert Systems with Applications*, 107:15–31.
Sinton, J., Wood, R., Pharoah, M., and Lewis, D. (1997). Influence of the addition of restorations on the diagnosis of caries from digitized bitewing radiographs. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 84(4):443–448.
Thanh, M. T. G., Van Toan, N., Ngoc, V. T. N., Tra, N. T., Giap, C. N., and Nguyen, D. M. (2022). Deep learning application in dental caries detection using intraoral photos taken by smartphones. *Applied Sciences*, 12(11):5504.
Topping, G. and Pitts, N. (2009). Clinical visual caries detection. *Detection, assessment, diagnosis and monitoring of caries*, 21:15–41.
Tracy, K. D., Dykstra, B. A., Gakenheimer, D. C., Scheetz, J. P., Lacina, S., Scarfe, W. C., and Farman, A. G. (2011). Utility and effectiveness of computer-aided diagnosis of dental caries. *Gen Dent*, 59(2):136–144.
Zadrozny, B. and Elkan, C. (2002). Transforming classifier scores into accurate multiclass probability estimates. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 694–699.
Zhang, Y., Zhang, H., Cai, J., and Yang, B. (2014). A weighted voting classifier based on differential evolution. In *Abstract and Applied Analysis*, volume 2014. Hindawi.