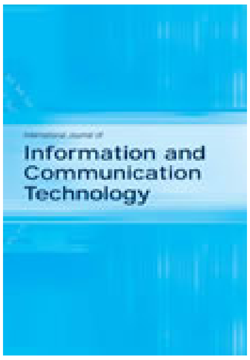


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Construction of oral health big data analysis platform and intelligent decision support system

Wanlu Chen, Wanmeng Li, Yaqiong Li

DOI: 10.1504/IJICT.2025.10075269

Article History:

Received:	08 September 2025
Last revised:	21 October 2025
Accepted:	23 October 2025
Published online:	12 January 2026

Construction of oral health big data analysis platform and intelligent decision support system

Wanlu Chen* and Wanmeng Li

Faculty of Oral Health,
Zhengzhou Health College,
Zhengzhou, 450064, China
Email: cwl@zzsqmc.edu.cn
Email: lwm@zzsqmc.edu.cn
*Corresponding author

Yaqiong Li

Operating Room,
Yichuan People's Hospital,
Luoyang, 471300, China
Email: seraczbf@163.com

Abstract: Oral diseases pose a global health challenge characterised by highly subjective diagnosis and a lack of intelligent decision-making tools, often exacerbated by fragmented data silos. This study aims to construct a comprehensive big data analytics platform and intelligent decision support system for oral health, enabling data-driven precision diagnosis and treatment through a unified four-layer architecture. The platform integrates multi-source heterogeneous data and employs advanced deep learning models for accurate caries segmentation and periodontitis risk prediction. Experiments on public datasets demonstrate a dice coefficient of 92.5% for caries segmentation and an area under the receiver operating characteristic curve value of 0.94 for periodontitis risk prediction, with results showing statistical significance. The system significantly enhances the automation and interpretability of oral disease analysis, providing a reliable and efficient tool for clinical diagnostic assistance and facilitating personalised treatment planning.

Keywords: oral health big data platform; intelligent decision support system; deep learning; periodontitis risk prediction; clinical assisted diagnosis.

Reference to this paper should be made as follows: Chen, W., Li, W. and Li, Y. (2025) 'Construction of oral health big data analysis platform and intelligent decision support system', *Int. J. Information and Communication Technology*, Vol. 26, No. 51, pp.70–84.

Biographical notes: Wanlu Chen received her Bachelor's degree from Zhengzhou University in 2015. Currently, she works at Zhengzhou Health College. Her research interests are oral health, intelligent decision support system, periodontitis risk prediction and clinical assisted diagnosis.

Wanmeng Li graduated from Henan University of Science and Technology in 2014. Currently, he works at Zhengzhou Health College. His research interests are oral health, intelligent decision support system, periodontitis risk prediction and clinical assisted diagnosis.

Yaqiong Li received her Bachelor's degree from Zhengzhou University in 2016. She is currently employed at Yichuan People's Hospital in Luoyang City. Her research interests are oral health, intelligent decision support system, periodontitis risk prediction and clinical assisted diagnosis.

1 Introduction

With the deepening implementation of the Healthy China 2030 plan and the accelerated advancement of digital healthcare strategies (Roth, 2013), big data and artificial intelligence technologies are profoundly transforming the service models and development pathways of modern healthcare (Amer et al., 2024). Oral health, as a vital component of overall well-being (Rein et al., 2004), presents a complex disease spectrum, high prevalence rates, and continuously growing diagnostic and treatment demands, making it a global public health challenge (Stein et al., 2010). The development of the proposed platform directly supports the 'Healthy China 2030' strategy by advancing digital health innovation. It exemplifies how data-driven technologies can be leveraged to enhance preventive care and precision diagnosis in oral health, thereby contributing to the initiative's broader goal of elevating public health standards through technological integration. According to the world health organisation, oral diseases affect nearly 3.5 billion people worldwide (Ealla et al., 2024). Untreated dental caries and severe periodontitis not only lead to tooth loss and impaired chewing function but are also significantly associated with systemic diseases such as cardiovascular disease (Baelum and Fejerskov, 1986), diabetes, and respiratory infections. However, traditional dental care models heavily rely on clinicians' experience and subjective judgement, resulting in low diagnostic consistency among practitioners (Alexwhite and Maupom, 2010). Furthermore, vast amounts of clinical, imaging, and omics data remain fragmented across 'data silos', severely hindering the advancement of precision dentistry (Rehm and Feeley, 2015).

In recent years, the development of medical big data platforms has become a hot topic in the industry. Internationally, renowned platforms such as informatics for integrating biology and the bedside (i2b2) and observational health data sciences and informatics (OHDSI) support the integration and analysis of cross-institutional electronic health records (EHRs) (Murphy et al., 2010), providing powerful tools for clinical research (Boulos et al., 2006). However, these general-purpose platforms exhibit significant limitations when handling the multimodal (Mower et al., 2008), heterogeneous data specific to the dental specialty. Dental diagnostics encompass diverse data types including radiographic images (apical radiographs, panoramic radiographs, cone-beam CT) (Bouwens et al., 2011), structured clinical records (e.g., periodontal probing depth, attachment loss), oral microbiome data, and patient-reported outcomes (PROs) (Macefield et al., 2014). Particularly concerning the standardisation and structuring of tooth-specific information (Buckley et al., 2002), existing platforms lack specialised optimisation, making it difficult to directly support in-depth analysis and decision-making applications specific to the dental specialty.

Meanwhile, clinical decision support systems (CDSS) are increasingly applied in the medical field (Reine et al., 2000). By integrating patient data with medical knowledge bases, they provide evidence-based treatment recommendations to clinicians, effectively reducing medical errors and enhancing diagnostic and therapeutic consistency (Schiff et al., 2003). However, existing dental CDSS primarily rely on traditional rule engines or simple logical judgements, exhibiting inherent limitations including delayed knowledge updates (Davenport et al., 2012), poor flexibility, and difficulty in handling complex nonlinear relationships. Although researchers have attempted to embed machine learning models into CDSS (Nijeweme-D'Hollosy et al., 2018), these efforts are often confined to single disease types and commonly suffer from 'black box' issues (Dindorf et al., 2024). The lack of transparency in model decision-making processes leads to insufficient trust in their recommendations among clinicians, thereby hindering the practical implementation of artificial intelligence in real clinical settings (Zavodna et al., 2024).

In summary, the field of oral health is at a critical juncture of digital transformation (Radu et al., 2024). While big data and artificial intelligence technologies present unprecedented opportunities (Polina et al., 2018), current research still faces significant gaps in deep data integration (Asakawa, 2001), systematic model integration, and the interpretability and usability of decision support (Zhang et al., 2014). Therefore, developing an integrated platform capable of fusing multi-source heterogeneous data (Kuo et al., 2002), integrating high-performance artificial intelligence (AI) analytical models (Nijim et al., 2005), and providing transparent, trustworthy decision support is not only an inevitable trend in technological advancement but also an urgent necessity to address current clinical pain points and advance precision dentistry (Scott, 2009). This study aims to tackle this challenge by constructing an integrated oral health big data analytics platform and intelligent decision support system (Kuhn et al., 2015), offering a systematic solution to the aforementioned issues (Ghaljehei et al., 2017).

2 Related work

2.1 Current status of medical big data platform development

Medical big data platforms serve as the foundational infrastructure for storing massive amounts of healthcare data and enabling advanced analytics and applications. Within this domain, several mature frameworks have gained widespread adoption. For instance, the i2b2 platform, spearheaded by Harvard Medical School, employs a fact-centric data warehouse model. This approach significantly facilitates complex cross-dimensional queries and cohort construction within EHRs for clinical researchers. Another representative project is OHDSI. By establishing a universal data model called observational medical outcomes partnership (OMOP), it maps EHR data from different institutions and formats into standardised terminology and structures. This enables large-scale, cross-institutional epidemiological research and safety monitoring. These universal platforms have achieved significant success in integrating structured data such as hospital admissions, medication use, and diagnoses, providing powerful tools for clinical research. The selection of i2b2 and OHDSI as reference platforms is based on their established reputation as benchmark systems in clinical data informatics. Their widespread adoption and well-documented architectures make them suitable representatives for comparing the capabilities and specialisation of the proposed oral health-focused platform. However, their limitations become apparent when handling specialty dental data. Dental care involves extensive specialty-specific data, including structured measurements like tooth position information based on international dental federation standards, periodontal probing depth (PD), and clinical attachment loss (CAL), alongside high-dimensional imaging data such as periapical radiographs, panoramic radiographs, and cone-beam computed tomography (CBCT). General-purpose platforms lack native optimisation for these specialised data types, making it difficult to effectively describe spatial relationships between teeth or perform in-depth correlation analysis on imaging data. Therefore, there is an urgent need for a data platform specifically designed for the field of oral health to achieve deep integration and efficient management of multi-source heterogeneous data. This is the primary motivation for constructing the oral health big data analytics platform in this study.

2.2 Applications of artificial intelligence in dentistry

Artificial intelligence, particularly machine learning and deep learning technologies, is revolutionising various subspecialties within dentistry, with applications primarily focused on medical image analysis and clinical prediction modelling. In image analysis, convolutional neural networks (CNNs) have emerged as powerful tools for automatically interpreting dental radiographs. Pioneered the use of CNNs to achieve precise segmentation and identification of individual teeth within CBCT images, laying the foundation for automated dental diagnosis. For caries detection, developed a deep learning-based object detection model capable of locating and identifying carious lesions in intraoral photographs with high accuracy. Regarding periodontitis, employed the nested U-Net architecture to achieve automated quantitative measurement of alveolar bone loss on panoramic radiographs, demonstrating performance comparable to that of oral radiology specialists. Furthermore, demonstrates that deep learning models can assist in implant surgery planning using CBCT images, automatically identifying critical anatomical structures such as the inferior alveolar nerve canal. In clinical prediction models, machine learning algorithms leverage structured data to forecast disease risks and treatment outcomes successfully predicted tooth loss risk using a regularised Cox proportional hazards model based on extensive longitudinal clinical data. Employed models like logistic regression and random forests to predict periodontitis progression at both the tooth and patient levels, identifying key risk indicators. Systematically reviewed implant survival prediction models, finding machine learning approaches outperform traditional statistical methods. Despite these encouraging results, most current studies remain ‘siloed’ – developing specialised models for single tasks using unimodal data. A caries detection model cannot leverage a patient’s systemic health information for assist in judgement, while a periodontitis prediction model struggles to cross-validate with radiographic evidence of bone loss. The ‘model silo’ phenomenon can be illustrated by a typical scenario in which a caries detection model operates in isolation, unable to access or incorporate relevant patient data – such as historical periodontal records – that reside in separate clinical systems. This fragmentation limits the model’s diagnostic comprehensiveness and contextual awareness. This fragmented ‘model silo’ phenomenon limits their clinical utility, failing to provide clinicians with a comprehensive, integrated decision-making reference.

2.3 The evolution and challenges of clinical decision support systems

CDSS aim to enhance healthcare quality and reduce errors by integrating clinical knowledge with patient data to provide personalised treatment recommendations for healthcare professionals. Traditional CDSS primarily rely on rule-based systems and knowledge bases, generating prompts and alerts through ‘if-then’ logic. In dentistry, early CDSS systems largely followed this approach, such as periodontitis diagnosis modules integrated into EHRs or guideline-based caries risk management tools. However, these systems rely on manually coded expert knowledge, suffering from inherent limitations including slow knowledge updates, high maintenance costs, weak uncertainty handling capabilities, and susceptibility to alert fatigue. With the advent of the data-driven era, a new generation of machine learning-based CDSS systems has emerged. These systems can automatically learn complex patterns from historical data, predict outcomes, and provide recommendations. However, their clinical implementation faces two core challenges. First is the ‘black box’ problem: the opaque decision-making process of complex deep learning models makes it difficult for clinicians to understand why specific recommendations are made, leading to trust deficits and low adoption rates. Second is low system integration. Many research-based AI models exist only as standalone software or web demos, failing to deeply integrate with clinical workflows [e.g., hospital information systems (HIS), picture archiving and communication systems picture archiving and communication system (PACS)]. Physicians must switch between different systems, significantly reducing willingness to use them. Also note that existing dental CDSS systems generally lack user-centred design, feature unfriendly interfaces, and provide insufficient decision evidence, failing to effectively support doctor-patient communication. Therefore, an ideal intelligent dental decision support system must not only integrate high-precision AI models but also address interpretability and system integration challenges, seamlessly and credibly embedding artificial intelligence throughout the entire clinical diagnosis and treatment chain.

3 Methodology

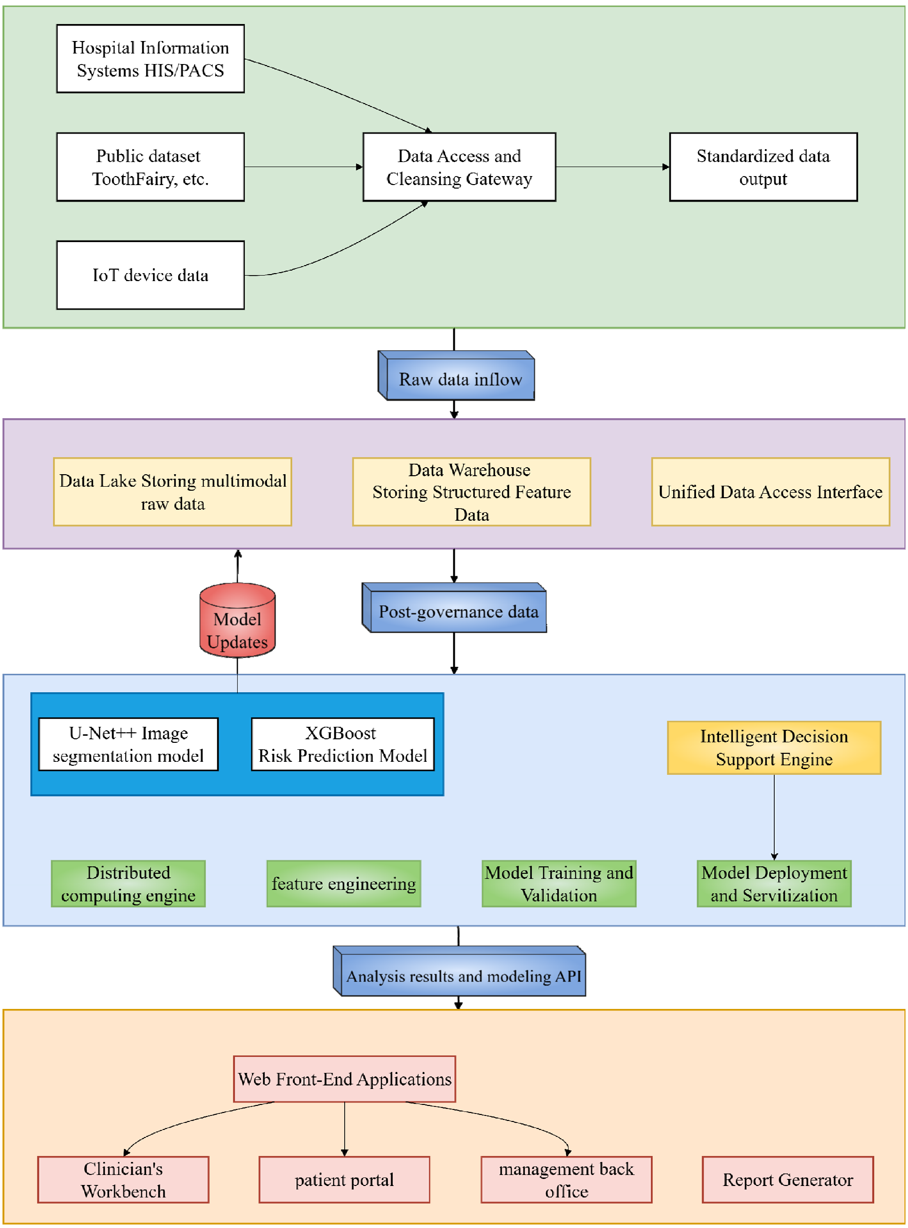
3.1 Overall system architecture

The oral health big data analytics platform developed in this study is an integrated system combining data integration, storage, computation, and application. Its logical architecture, as shown in Figure 1, is primarily divided into the following four layers:

- Data acquisition and integration layer: this layer is responsible for obtaining raw data from multiple heterogeneous data sources. This study primarily utilised two public datasets: the periodontal machine learning (ML) dataset, comprising clinical structured data from 1,200 patients, covering characteristics such as age, gender, periodontal probing depth, CAL, bleeding index (BI), smoking history, and periodontitis diagnosis results based on the CDC-AAP combined criteria. All data usage adheres to their respective open license agreements and has been granted exemption by the institutional ethics review committee.
- Data storage and management layer: this layer serves as the platform’s data hub. To handle massive volumes of unstructured imaging data, we employ the Hadoop distributed file system (HDFS) for low-cost, highly reliable storage. Structured clinical metadata, patient information, and final model analysis results are stored and managed using the relational database MySQL. We designed a data model tailored to dental specialty requirements, such as standardising tooth positions using the FDI tooth numbering system, ensuring data consistency and queryability. The hybrid

storage strategy was implemented to optimally balance performance and scalability. HDFS is utilised for managing large volumes of unstructured imaging data due to its distributed and fault-tolerant nature, while MySQL supports efficient storage and retrieval of structured clinical metadata, ensuring rapid query response and transactional reliability.

Figure 1 Overall architecture of the oral health big data analytics platform (see online version for colours)



- Data processing and analysis layer: this serves as the computational core of the platform. It receives application requests from upper layers and schedules computational resources to execute data preprocessing, model training, and inference tasks. We utilise apache spark as the distributed computing framework to efficiently process large-scale data. This layer integrates our deep learning models developed for caries segmentation and machine learning models for periodontal disease risk prediction. These models are encapsulated and managed through containerisation technologies (such as docker), ensuring consistency and portability of the computational environment.
- Application and intelligent decision support layer: this layer directly serves end-users (dental practitioners). We developed a web-based interactive front-end interface. Dentists can use this interface to upload new patient data or query historical records. Upon receiving a request, the backend analytics service invokes corresponding models in the underlying layer for computation. Results are returned to the frontend in visual formats – such as images highlighting caries-affected areas, probability pie charts displaying risk predictions, and feature contribution analysis charts based on Shapley additive explanations (SHAP) values – ultimately generating a structured intelligent diagnostic report. This provides intuitive, credible support for clinical decision-making.

3.2 Data preprocessing and feature engineering

- Image data preprocessing: X-ray images obtained from public datasets vary in size and greyscale distribution, necessitating standardisation. First, we resample all images to a uniform size of 512 × 512 pixels and apply normalisation to scale pixel values to the range [0, 1]. The formula is as follows:

$$I_{\text{norm}} = \frac{I - I_{\text{min}}}{I_{\text{max}} - I_{\text{min}}} \tag{1}$$

  where I represents the original input image, I_{min} and I_{max} denote the minimum and maximum pixel values of image I, respectively, and I_{norm} is the normalised image. To further enhance the model's generalisation capability, we employed real-time data augmentation techniques during the training phase.
- Structured data preprocessing: for clinical data, we first addressed missing values... Subsequently, we Z-score standardised continuous features to conform to a standard normal distribution with a mean of 0 and a standard deviation of 1.

$$z = \frac{x - \mu}{\sigma} \tag{2}$$

  where x is the original feature value, μ is the mean of this feature in the training set, σ is the corresponding standard deviation, and z is the new feature value after standardisation.
- Feature engineering: we concatenate the structured clinical feature vector $F_{clinical}$ with the deep feature vector F_{deep} to form the fused feature vector $F_{fusion} = [F_{clinical}; F_{deep}]$ for risk prediction.

3.3 Core algorithm model

A U-Net++-based model for caries segmentation. We employ a compound loss function L_{seg} to jointly optimise the model:

$$L_{seg} = \alpha L_{Dice} + (1-\alpha) L_{Focal} \tag{3}$$

Dice loss: direct optimisation for dice coefficient in segmentation evaluation.

$$L_{Dice} = 1 - \frac{2\sum_{i}^{N} p_i g_i + \varepsilon}{\sum_{i}^{N} p_i^2 + \sum_{i}^{N} g_i^2 + \varepsilon} \tag{4}$$

where $p_i \in [0, 1]$ represents the model's predicted probability that the i pixel belongs to a cavity, while $g_i \in 0, 1$ denotes the corresponding ground truth label. n is the total number of pixels in a batch, and ε is an extremely small smoothing term (typically 1×10^{-5}) to prevent division by zero.

Focal loss is an improvement upon standard cross-entropy loss:

$$L_{Focal} = -\alpha_t (1-p_t)^{\gamma} \log(p_t) \tag{5}$$

where p_t is the model's predicted probability for the true class, α_t is the class weight balancing factor, and γ is an adjustable focus parameter (where $\gamma \geq 0$) used to adjust the weights of easy and difficult samples.

Periodontal disease risk prediction model based on extreme gradient boosting (XGBoost) and SHAP. XGBoost is an additive model composed of K base learners (decision trees). Its prediction output is:

$$\hat{y}i = \phi(x_i) = \sum k = 1^K f_k(x_i), \qquad f_k \in \mathcal{F} \tag{6}$$

where $\hat{y}_i$ is the predicted value for the i sample, x_i is the feature vector, f_k is the k decision tree, and $\mathcal{F}$ is the function space of all possible decision trees.

The objective function Obj of the model consists of two components: the training loss L and the regularisation term Ω:

$$\text{Obj} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum k = 1^K \Omega(f_k) \tag{7}$$

where $l(y_i, \hat{y}_i)$ is a differentiable convex loss function that measures the discrepancy between the predicted value $\hat{y}_i$ and the true label y_i.

The regularisation term Ω is used to control the complexity of the model.

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{8}$$

where T denotes the number of leaf nodes in the tree, and w_j represents the score (weight) of the j leaf node. γ and λ are hyperparameters that control the penalty on the number of leaf nodes and the $L2$ penalty on leaf node weights, respectively.

To enhance model interpretability, we employ SHAP values for post-hoc explanation. The SHAP value ϕ_i for feature i is calculated as follows:

$$\phi_i(f, x) = \sum_{S \subseteq N \setminus i} \frac{|S|!(|N| - |S| - 1)!}{|N|!}[f(S \cup i) - f(S)] \tag{9}$$

where f is the trained model, x is the input sample, N is the set of all features, and S is a subset of features.

Model evaluation metrics: for the dental caries segmentation task, we employ dice similarity coefficient (DSC).

$$\text{DSC} = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{10}$$

where X is the set of predicted pixels, Y is the set of true pixels, and TP, FP, FN represent the number of true positive, false positive, and false negative pixels, respectively.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

where *TP* the number of correct predictions that are truly positive, and *TP* + *FN* is total number of all predictions of the model that were positive (both correct and incorrect).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

where *TP* is number of successful positive detections, and *TP* + *FN* is total number of true positives (both detected and missed).

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

where precision is precision rate values above, and recall the above recall values.

For the periodontitis risk prediction task, we primarily evaluate performance using area under the curve, accuracy, and F1-score.

4 Experimental verification

To comprehensively evaluate the performance of the core algorithms in our proposed oral health big data analytics platform and intelligent decision support system, we designed and conducted rigorous experiments. This section details the experimental setup, the benchmark models used, and the evaluation metrics. We present both quantitative and qualitative analyses of the experimental results, concluding with ablation experiments to validate the effectiveness of each component in our model design.

4.1 Experimental setup

Dataset and preprocessing: this experiment evaluates two publicly available datasets: the Toothfairy dataset contains 2,000 annotated dental x-rays with caries lesions and the Periodontal-ML dataset includes clinical data and periodontitis diagnosis labels for

1,200 patients. Data were split into training, validation, and test sets at a ratio of 7:1.5:1.5. All image data were uniformly resampled to 512 × 512 pixels and normalised. Clinical data underwent mean-value imputation for missing values followed by z-score normalisation.

Implementation details: the experiment was implemented using Pytorch 1.12.1 and scikit-learn 1.0.2. The hardware environment consisted of an Nvidia RTX A6000 GPU. The caries segmentation model employed the U-Net++ architecture with the Adam optimiser (initial learning rate 1e-4, batch size 8). The periodontitis prediction model employed XGBoost with hyperparameters optimised via grid search (learning rate 0.1, maximum depth 6). All experiments were repeated three times, and the average results were taken.

Evaluation metrics and comparison algorithms: for the caries segmentation task, metrics included DSC, accuracy, precision, recall and F1-score. Comparison algorithms included: fully convolutional network-8s (FCN-8s) – fully convolutional network benchmark; U-Net – classic model for medical image segmentation; attention U-Net – enhanced model incorporating attention mechanisms. Periodontal disease prediction task evaluated using accuracy and F1-score metrics. Comparative algorithms include: logistic regression (LR), support vector machine (SVM) and random forest (RF).

Periodontitis prediction task evaluated using accuracy and F1-score metrics. Comparison algorithms include: LR, SVM and RF.

4.2 Results and analysis

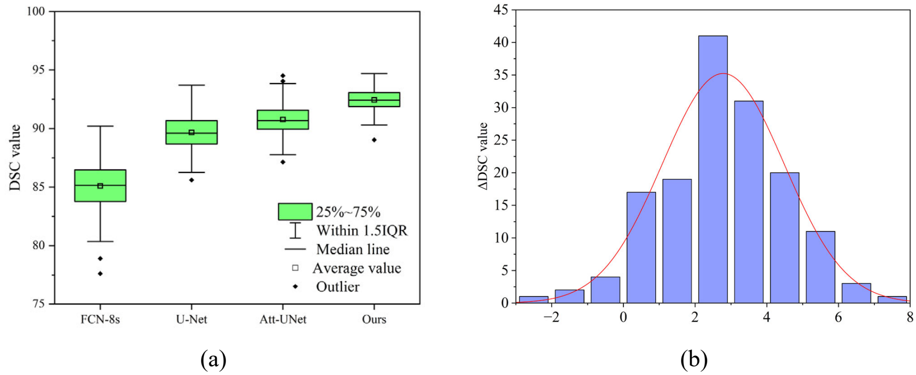
Results of caries segmentation: quantitative analysis results (Table 1) demonstrate that our proposed U-Net++ model achieves optimal performance across all evaluation metrics. Specifically, it attains a Dice similarity coefficient of 92.5% on the core metric, significantly outperforming the comparison models. Notably, our model achieves the optimal balance between precision (94.1%) and recall (91.8%). This indicates that the model effectively minimises false positives (avoiding misclassification of healthy tissue as lesions) while maximally reducing false negatives (preventing missed detection of lesions). This characteristic is crucial for clinical diagnostic support.

Model stability analysis is demonstrated via box plots (Figure 2). Our proposed model not only exhibits the highest median DSC but also demonstrates the smallest interquartile range (IQR) and outlier range, indicating its outstanding stability across diverse datasets. Statistical tests confirm that performance differences between our model and all comparison models reach statistical significance ($p < 0.001$).

Table 1 Performance comparison of different models on the caries segmentation test set

Model	*DSC (%, mean ± std)*	*Accuracy (%, mean ± std)*	*Precision (%, mean ± std)*	*Recall (%, mean ± std)*	*F1-score*
FCN-8s	85.3 ± 2.1	93.5 ± 0.8	86.7 ± 2.5	86.1 ± 2.3	0.864
U-Net	89.7 ± 1.5	95.8 ± 0.6	90.5 ± 1.8	90.2 ± 1.9	0.903
Attention U-Net	90.8 ± 1.2	96.2 ± 0.5	91.9 ± 1.6	91.0 ± 1.7	0.914
Ours	92.5 ± 0.7	96.9 ± 0.4	94.1±1.2	91.8 ± 1.4	0.929

Figure 2 Performance distribution and statistical significance analysis of partitioned models (see online version for colours)



The computational efficiency comparison further validates the model's practicality. Despite the relatively complex architecture of U-Net++, its average processing time during inference is only 0.15 seconds per image (on an RTX A6000 GPU environment), comparable to U-Net's 0.12 seconds per image. This performance falls well below the acceptable latency threshold for clinical diagnosis, demonstrating its feasibility for clinical application.

Performance analysis across different types of dental caries reveals the model's specialised capabilities. We conducted subgroup analyses on different caries types (occlusal caries, interproximal caries, smooth surface caries) within the test set. The model demonstrated particularly outstanding performance in detecting interproximal caries (DSC = 93.2%), which are most prone to being missed in traditional radiographic examinations. This indicates our model can effectively enhance the sensitivity and accuracy of clinical diagnosis.

Periodontal disease risk prediction results: performance evaluation results (Table 2) indicate that the XGBoost model demonstrated optimal performance in periodontal disease risk prediction, achieving an accuracy rate of 89.2% and an F1-score of 0.887. Compared to traditional machine learning methods, XGBoost better captures complex nonlinear relationships among features through its gradient boosting mechanism. Notably, the model achieved a recall rate of 90.5% for positive samples (periodontitis patients), indicating exceptionally high disease detection sensitivity. This characteristic holds significant importance for early screening and intervention.

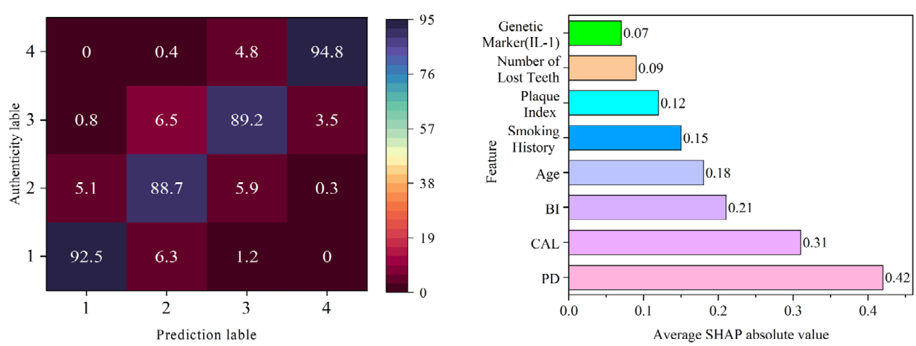
Table 2 Performance of different models on the periodontitis risk prediction test set

Model	*Accuracy (%, mean ± std)*	*F1-score (mean ± std)*
Logistic regression	82.1 ± 2.1	0.809 ± 0.03
SVM (RBF)	85.3 ± 1.8	0.843 ± 0.02
Random forest	87.8 ± 1.5	0.869 ± 0.02
Ours (XGBoost)	89.2 ± 1.2	0.887 ± 0.01

To understand the logic behind the model's decisions and assess its consistency with clinical knowledge, we calculated global feature importance based on the average absolute SHAP value (Figure 3b). PD was identified as the most predictive feature, with its importance significantly higher than other factors. This aligns perfectly with the gold standard for periodontitis diagnosis, demonstrating that the model successfully captured the most critical medical-pathological feature. CAL and BI ranked second and third, respectively, both being integral components of core clinical manifestations of periodontitis. Additionally, traditional risk factors such as smoking history and age were assigned moderate importance by the model, consistent with epidemiological research findings. Notably, the model autonomously identified the hierarchical importance of these features from the data without incorporating any prior medical knowledge, highlighting its data-driven discovery capability. This feature importance ranking not only enhances clinicians' confidence in the model's predictions but also provides dentists with a clear prioritisation list, indicating the core indicators that should be the primary focus during clinical evaluations.

Confusion matrix analysis [Figure 3(a)] provides deeper performance insights. The model demonstrates exceptionally high classification accuracy for the 'healthy' and 'severe periodontitis' categories (92.5% and 94.8%, respectively), while the primary classification errors occur between adjacent severity categories (e.g., misclassifying 'mild' as 'moderate'). This error pattern aligns with clinical logic, as the clinical manifestations between adjacent severity levels exhibit inherent continuity, making differentiation more challenging. This further validates that the features learned by the model are clinically meaningful.

Figure 3 Classification performance and interpretability analysis (see online version for colours)



(a) (b)

Model calibration assessment is conducted via calibration curves. Our XGBoost model demonstrates excellent calibration properties (Brier score = 0.082), with predicted probabilities highly consistent with actual risks. This is crucial for clinical decision support, enabling physicians to rely on the model's risk probability outputs when formulating treatment plans.

Subgroup analysis across different populations demonstrated the model's generalisation capability. We evaluated performance across age groups (<35 years, 35–55 years, >55 years), with the model maintaining stable performance across all age cohorts (F1-score variation < 2%), indicating its applicability to a broad patient population.

5 Conclusions

This study successfully designed and constructed an integrated oral health big data analytics platform and intelligent decision support system, aiming to systematically address core challenges in the field of dentistry: data silos, high diagnostic subjectivity, and the lack of intelligent auxiliary tools. By integrating multi-source heterogeneous data and advanced deep learning and machine learning algorithms, the system achieves a complete closed-loop process from data management to intelligent decision making.

Experimental results demonstrate that the U-Net++-based automated segmentation model for dental caries lesions achieved a dice similarity coefficient of 92.5% on the public test dataset. Box plot analysis and statistical tests confirm that its performance improvement not only significantly outperforms all comparison models but also exhibits extremely high statistical significance ($p < 0.001$), showcasing exceptional stability and robustness. Meanwhile, the XGBoost-based periodontitis risk prediction model achieved an accuracy of 89.2% and an F1-score of 0.887. Normalised confusion matrix heatmap analysis further revealed that the model's error patterns align with clinical logic (primarily occurring between adjacent severity categories), confirming that the features it learns possess clear medical significance. Combined with the SHAP interpretability framework, the model's decision-making process becomes transparent and credible, clearly showing the dominant role of key features such as periodontal probing depth. This effectively addresses the trust challenge posed by 'black-box' models in clinical scenarios.

The primary theoretical contributions of this work lie in three aspects. First, it proposes and validates an integrated platform architecture tailored for oral specialties, providing a systematic engineering solution for processing multimodal specialty data. Second, it advances model evaluation from isolated case demonstrations to comprehensive quantitative statistics and significance testing, establishing a more rigorous paradigm for medical AI research. Third, the adopted composite loss function is proven to effectively address both class imbalance and hard-to-learn samples in medical image segmentation.

At the practical level, this study provides a viable tool for implementing precision dentistry. The platform assists clinicians in accurately identifying early lesions, conducting objective disease assessments, and performing personalised risk predictions. Its stable performance and high interpretability suggest it can be reliably deployed across healthcare institutions of varying scales, significantly enhancing diagnostic accuracy and consistency – particularly by offering effective auxiliary diagnostic support to primary care facilities.

Declarations

All authors declare that they have no conflicts of interest.

References

Alexwhite, B. and Maupom, G. (2010) 'Making clinical decisions for dental care: concepts to consider', *Special Care in Dentistry*, Vol. 23, No. 5, pp.168–172.

Amer, M., Hilmi, Y. and El Kezazy, H. (2024) 'Big data and artificial intelligence at the heart of management control: towards an era of renewed strategic steering', *The International Workshop on Big Data and Business Intelligence*, Vol. 2, No. 7, p.79.

Asakawa, K. (2001) 'Organizational tension in international R&D management: the case of Japanese firms', *Research Policy*, Vol. 7, No. 10, p.549.

Baelum, V. and Fejerskov, O. (1986) 'Tooth loss as related to dental caries and periodontal breakdown in adult Tanzanians', *Community Dentistry and Oral Epidemiology*, Vol. 2, No. 12, p.263.

Boulos, M.N., Maramba, I. and Wheeler, S. (2006) 'Wikis, blogs and podcasts: a new generation of web-based tools for virtual collaborative clinical practice and education', *BMC Medical Education*, Vol. 6, No. 14, p.879.

Bouwens, D.G., Cevidanes, L., Ludlow, J.B. and Phillips, C. (2011) 'Comparison of mesiodistal root angulation with posttreatment panoramic radiographs and cone-beam computed tomography', *American Journal of Orthodontics & Dentofacial Orthopedics*, Vol. 139, No. 1, pp.126–132.

Buckley, P.J., Clegg, J. and Wang, C. (2002) 'The impact of inward FDI on the performance of Chinese manufacturing firms', *Journal of International Business Studies*, Vol. 33, No. 4, pp.637–655.

Davenport, M.A., Laska, J.N., Treichler, J.R. and Baraniuk, R.G. (2012) 'The pros and cons of compressive sensing for wideband signal acquisition: noise folding vs. dynamic range', *IEEE Transactions on Signal Processing*, Vol. 60, No. 9, pp.4628–4642.

Dindorf, C., Horst, F., Slijepevi, D., Dumphart, B., Dully, J., Zeppelzauer, M., Horsak, B. and Frhlich, M. (2024) 'From lab to field with machine learning – bridging the gap for movement analysis in real-world environments: a commentary', *Current Issues in Sport Science*, Vol. 9, No. 4, p.791.

Ealla, K.K.R., Kumari, N., Chintalapani, S., Uppu, S., Sahu, V., Veeraraghavan, V.P., Ramani, P. and Govindool, S.R. (2024) 'Interplay between dental caries pathogens, periodontal pathogens, and sugar molecules: approaches for prevention and treatment', *Archives of Microbiology*, Vol. 3, No. 5, p.206.

Ghaljehei, M., Soltani, Z., Haggi, H. and Golkar, M.A. (2017) 'Integrated power market with RESs and load uncertainties using scenario-based method in order to voltage collapse probability relief', *Computers in Industry*, Vol. 7, No. 1, p.347.

Kuhn, L., Reeves, K., Taylor, Y., Tapp, H., Mcwilliams, A., Gunter, A., Cleveland, J. and Dulin, M. (2015) 'Planning for action: the impact of an asthma action plan decision support tool integrated into an electronic health record (EHR) at a large health care system', *Journal of the American Board of Family Medicine*, Vol. 28, No. 3, pp.382–393.

Kuo, R.J., Chi, S.C. and Kao, S.S. (2002) 'A decision support system for selecting convenience store location through integration of fuzzy AHP and artificial neural network', *Computers in Industry*, Vol. 47, No. 2, pp.199–214.

Macefield, R.C., Jacobs, M., Korfage, I.J., Nicklin, J., Whistance, R.N., Brookes, S.T., Sprangers, M.A. and Blazeby, J.M. (2014) 'Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs)', *Trials*, Vol. 15, No. 1, p.49.

Mower, E., Lee, S.L.S., Matarific, M.J. and Narayanan, S. (2008) 'Human perception of synthetic character emotions in the presence of conflicting and congruent vocal and facial expressions', *Computers in Industry*, Vol. 6, No. 7, p.158.

Murphy, S.N., Griffin, W., Michael, M., Vivian, G., Chueh, H.C., Susanne, C. and Isaac, K. (2010) 'Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2)', *Journal of the American Medical Informatics Association*, Vol. 17, No. 2, pp.124–130.

Nijeweme-D'Hollosy, W.O., van Velsen, L., Poel, M., Groothuis-Oudshoorn, C.G.M., Soer, R. and Hermens, H. (2018) 'Evaluation of three machine learning models for self-referral decision support on low back pain in primary care', *International Journal of Medical Informatics*, Vol. 110, No. 15., pp.31–41.

Nijim, M., Xie, T. and Qin, X. (2005) 'Integrating a performance model in self-managing computer systems under mixed workload conditions', *IRI – 2005 IEEE International Conference on Information Reuse and Integration*, Las Vegas, NV, USA, pp.132–137.

Polina, M., Lucy, O., Yury, Y., Alex, O., Alex, B., Pavel, P., Eugene, I., Alexander, A., Konstantin, R. and Alexander, Z. (2018) 'Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare', *Oncotarget*, Vol. 9, No. 5, pp.5665–5690.

Radu, R., Nanni, R. and Shahin, J. (2024) 'New challenges in internet governance: power shifts and contestation from 'within'', *Telecommunications Policy*, Vol. 5, No. 7, p.48.

Rehm, E.M. and Feeley, K.J. (2015) 'The inability of tropical cloud forest species to invade grasslands above treeline during climate change: potential explanations and consequences', *Ecography*, Vol. 38, No. 12, p.192.

Rein, E., Kharazmi, A. and Winther, K. (2004) 'A herbal remedy, Hyben Vital (stand. powder of a subspecies of *Rosa canina* fruits), reduces pain and improves general wellbeing in patients with osteoarthritis – a double-blind, placebo-controlled, randomised trial', *Phytomedicine*, Vol. 11, No. 5, pp.383–391.

Reine, G., Bernard, D., Auquier, P., Fur, B.L. and Lanon, C. (2000) 'Psychometric properties of french version of the Calgary Depression Scale for Schizophrenics (CDSS)', *Lencéphale*, Vol. 26, No. 1, p.52.

Roth, V.J. (2013) 'The mhealth conundrum: smartphones & mobile medical apps – how much FDA medical device regulation is required?', *Ecography*, Vol. 5, No. 15, p.142.

Schiff, G.D., Klass, D., Peterson, J., Shah, G. and Bates, D.W. (2003) 'Linking laboratory and pharmacy: opportunities for reducing errors and improving care', *Archives of Internal Medicine*, Vol. 163, No. 8, pp.893–900.

Scott, S. (2009) 'Current practicality of nanotechnology in dentistry. Part 1: focus on nanocomposite restoratives and biomimetics', *Clinical Cosmetic & Investigational Dentistry*, Vol. 1, No. 3, pp.47–61.

Stein, C., E., Schlundt, J., Tauxe, R., V., Doyle and M. (2010) 'Evolving public health approaches to the global challenge of foodborne infections', *International Journal of Food Microbiology*, Vol. 20, No. 17, p.700.

Zavodna, L.S., Berwimmer, M. and Frankus, E. (2024) 'Barriers to the implementation of artificial intelligence in small and medium-sized enterprises: pilot study', *Journal of Economics and Management*, Vol. 46, No. 1, pp.331–352.

Zhang, K., Wan, H., Liu, X., Ghan, S.J., Kooperman, G.J., Ma, P.L., Rasch, P.J., Neubauer, D. and Lohmann, U. (2014) 'Technical note: on the use of nudging for aerosol-climate model intercomparison studies', *Atmospheric Chemistry and Physics*, Vol. 14, No. 16, pp.8631–8645.