Journal of Applied Research and Technology

www.jart.icat.unam.mx

Instituto de Ciencias Aplicadas y Tecnología

Journal of Applied Research and Technology 24 (2026) 649-663

Original

Integrated and Advanced Deep Learning Framework for Accurate, Interpretable, and Scalable Tooth Defect Diagnosis and Segmentation

Mohit Popat* • Dinesh Chaudhari

Department of Computer Science and Engineering, Jawaharlal Darda Institute of Engineering and Technology, Yavatmal, India

Received: 12 03 2025; Accepted: 04 07 2025
Available: 31 08 2026

Abstract: Accurate, interpretable, and scalable diagnostic solutions are warranted due to increasing incidences of tooth defects. Current methods often fail to generalize and lack transparency, particularly across heterogeneous datasets. An integrated Deep Learning Framework for tooth defect diagnosis and segmentation that relies on the recent EfficientNetB7 in combination with Squeeze-and-Excitation (SE) blocks to enhance the overall feature recalibration, enables the incorporation of Vision Transformers (ViT) for global context modeling, and then adds further convolutional neural networks with Grad-CAM for the visual interpretability process. There is also the use of Attention U-Net for accurate lesion segmentation and SHAP (SHapley Additive exPlanations) to quantify pixel-level attributions. This synergy will achieve very high validation accuracy (92.4%), very high segmentation metrics (IoU > 0.82), and interpretability scores (> 9/10) which can potentially take it as a viable solution in clinical settings. Importantly, the combination of these architectural components represents a considerable breakthrough for AI-driven dental diagnostics.

Keywords: **EfficientNet, Vision Transformer, Grad-CAM, Attention U-Net, SHAP, process.**

*Corresponding author.
E-mail address: mohit.popat872@gmail.com (M. Popat).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

List of Abbreviations

Abbreviation	Full Form
CNN	Convolutional Neural Network
DL	Deep Learning
DL-Ens	Deep Learning Ensembles
DL-Fusion	Deep Learning Fusion
DL-Fuzzy	Fuzzy Deep Learning
DNA	Deoxyribonucleic Acid
Grad-CAM	Gradient-weighted Class Activation Mapping
IoU	Intersection over Union
ML	Machine Learning
ML-Ens	Machine Learning Ensembles
ML-Imp	Machine Learning Imputation
MRI	Magnetic Resonance Imaging
OCT	Optical Coherence Tomography
RNA	Ribonucleic Acid
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic Curve
SHAP	SHapley Additive exPlanations
USPSTF	United States Preventive Services Task Force

1. Introduction

Tooth defects have always been a considerable public health concern, with time and effort lost treating patients for whom early detection and diagnosis failed. Classical diagnostic methods rely on observer variability and require human expert interpretation, making them inefficient in high-demand settings. Algorithms using AI largely acceptably achieve accuracy; however, very few of them pay heed to the significant requirement for interpretability and robustness on patient data. In addition, classical CNNs are less capable than current architectures at capturing long-range dependencies and fine-grained latent lesion characteristics. To overcome these challenges (Adeoye et al., 2023; Chaudhuri et al., 2023; Yaduvanshi et al., 2024), the study presents an integrated Deep Learning Framework, bringing together advanced architectural improvements—EfficientNetB7 with SE blocks, Vision Transformers for global feature modeling, and Grad-CAM/SHAP for interpretability. Furthermore, there is an Attention U-Net focused on the specific segmentation of lesions. This framework would bridge the gap among accuracy, interpretability, and scalability, in other words, providing an all-inclusive and clinically practical approach for automating tooth defect diagnosis.

The impetus for this work comes from the critical need to improve the accuracy, efficiency, and interpretability of diagnostics in the detection of tooth defect. Diagnostic methodologies, manual or AI-driven, often lack one or more critical aspects. Manual approaches suffer from limited inter-observer variability and poor scalability, and many AI models show high accuracy but often little transparency within their decision-making process. Furthermore, the segmentation models lack sharpness in defining lesion boundaries, which is a very stringent requirement for treatment planning purposes. These deficiencies highlight the demand for an integrated framework that, on the one hand, balances accuracy and precision with interpretability and, on the other, addresses the unique challenges associated with tooth defect diagnosis. This work makes several contributions to the field of automated defect in normal dental operations. First, it adds EfficientNetB7 augmented with SE blocks for refining feature extraction and amplifying clinically relevant patterns, yielding the best possible classification accuracy. The second, Vision Transformers, exploits the attention mechanism of global relationships in images that are very subtle and complex, which traditional methods fail to do. The third, it offers explainability via Grad-CAM and SHAP, thus allowing both localized and global understanding of model predictions. Finally, incorporating Attention U-Net for the segmentation part is done in order to precisely highlight the boundaries of lesions so as to get detailed analysis. Integrating these together, the suggested framework presents an efficient, explainable, and scalable approach towards tooth defect diagnosis that may completely transform clinical workflows, leading to the betterment of patient care outcomes. This effort not only places the technical advancements but also forms a new precedent for AI integration in mission-critical healthcare applications.

2. Review of existing Methods for Tooth Defect Analysis

Analyzing the tremendous strides that Machine Learning (ML) and Deep Learning (DL) have taken for tooth defect detection, treatment, and prognosis is a critical research focus, an effort reflected in a thorough review of key texts. These studies include myriad methodologies, datasets, and applications, from novel algorithmic improvements to practical clinical integrations, offering a chronological exploration of the field. Yaduvanshi et al. (2024) initiated a modified local texture descriptor in combination with ML classifiers, thereby significantly improving the accuracy of tooth defect detection. Thereafter, Adeoye et al. (2023) proposed risk-predicting models specifically designed for oral leukoplakia and oral lichenoid mucositis, which can be considered an initial step in incorporating ML in preventive diagnostics. Chaudhuri et al. (2023) made use of Raman cyto-spectroscopy data with ensemble ML models for early tooth defect risk assessment; Somyanonthanakul et al. (2024) introduced the application of Fuzzy Deep Learning in predicting survival probabilities. Deep learning received much attention following Li et al. (2024), who indicated its applicability in tooth defect detection systems. Warin and Suebnukarn (2024) conducted a systematic review on DL's effect and explained how convolutional and recurrent architectures work. In Table 1, Babu et al. (2024) expressed the necessity of explainability in DL as well as proposed models that generate critical outputs in clinical settings.

Sakamoto et al. (2024) further extended the application of DL by automating the analysis of masseter muscle volume, which is also an important prognostic indicator for tooth defect patients. Öztürk et al. (2024) discussed the use of magnetic resonance imaging-based Machine Learning models to predict bone invasion in oral squamous cell carcinoma, an especially difficult problem in tooth defect diagnosis. Wei et al. (2024) enhanced CNNs using tunicate swarm algorithms, obtaining great improvements in the precision of the sets in the detection of defect in normal dental operations. ML and DL in various applications of other domains associated with defect in normal dental operations open new avenues in tooth defect developments. Shi et al. (2024) applied ML to the analysis of porcine tissue-based intestinal transportome, while Nopour (2024) utilized this method for the forecasting of ovarian defect in normal dental operations risk into the realm of more accessible defect in normal dental operations diagnostics. Yan et al. (2024) developed long-term prognostic models for renal defect in normal dental operations using patient-specific data, which inspired Sampath et al. (2024a) to introduce OralNet, a DL model that incorporates stochastic logistic regression to identify oral lesions from lip and tongue images and samples. Sampath et al. (2024b) presented serum micro-RNAs with mutation-targeted RNA modifications as inputs for ML workflows, thus furthering the limits of biomarker-driven defect in normal dental operations detection. Li et al. (2023) applied ML to predict five-year survival rates for tongue defect in normal dental operations patients, achieving clinical utility with impressive accuracy. Vittone et al. (2024) expanded the early defect in normal dental operations detection spectrum with a blood-test-based ML model for multi-defect in normal dental operations scenarios, setting a benchmark for universal defect in normal dental operations screening. Ahmed et al. (2024) delved into genomic associations between craniofacial diseases and systemic conditions, while Bhatt and Shende (2023) and El Badisy et al. (2024) explored a holistic ML approach from defect in normal dental operations identification to treatment optimizations. Deo et al. (2024) analyzed ML imputation methods for breast defect in normal dental operations survival, paralleling the integration of empirical wavelet transforms in DL for tooth defect histopathology classifications. Rai and Yoo (2023) conducted a meta-analysis of some recent ML advancements, pointing out their diagnostic impact across defect in normal dental operations.

In the area of imaging, Yuan et al. (2024) used optical coherence tomography in DL-based radiomics to provide prognostic predictions for tooth defect. The work of Stark et al. (2024) regarding ML-driven unknown primary tumor site in head and neck defect in normal dental operations utilizes DNA methylation profiles; indeed, epigenetic data would be manageable and adaptable using ML.

Table 1. Methodological comparative analysis.

Ref.	Method	Main Objectives	Findings	Limitations
Yaduvanshi et al. (2024)	Modified local texture descriptor with ML	Enhance tooth defect detection accuracy	Improved classification accuracy significantly	Limited generalization across heterogeneous datasets
Adeoye et al. (2023)	ML for risk prediction in oral leukoplakia	Predict tooth defect risk in specific conditions	High accuracy in stratifying high-risk patients	Dataset limited to specific conditions

Table 1. Continued

Ref.	Method	Main Objectives	Findings	Limitations
Chaudhuri et al. (2023)	Raman cyto-spectroscopy with ML ensembles	Assess tooth defect risk using spectral data	Robust early detection methodology	Requires specialized equipment
Somyanonthanakul et al. (2024)	Fuzzy DL for survival estimation	Estimate survival probabilities for tooth defect	Enhanced prognostic accuracy	Complex fuzzy logic implementation
Li et al. (2024)	DL-based tooth defect detection system	Develop an automated defect in normal dental operations detection tool	Achieved high detection rates in large datasets	Limited interpretability
Warin & Suebnukarn (2024)	Systematic review of DL in tooth defect	Evaluate DL advancements in tooth defect	Identified key strengths of CNNs and RNNs	Lack of quantitative comparisons
Babu et al. (2024)	Explainable DL for tooth defect detection	Improve model interpretability	Provided interpretable heatmaps alongside predictions	Slightly lower accuracy than black-box models
Sakamoto et al. (2024)	DL for masseter muscle volume analysis	Use muscle volume as a prognostic indicator	Automated analysis correlated with survival outcomes	Limited to specific muscle-related prognostics
Öztürk et al. (2024)	MRI-based ML for bone invasion prediction	Predict bone invasion in oral squamous carcinoma	High accuracy and specificity achieved	Relies on high-resolution imaging
Wei et al. (2024)	CNN with tunicate swarm algorithm	Enhance CNN performance in defect in normal dental operations detection	Significantly improved precision and recall	Computationally expensive
Shi et al. (2024)	ML for drug-transportome interactions	Screen oral drugs and tissue interactions	Effective drug screening for tooth defect patients	Not directly linked to a defect in normal dental operations diagnostics
Nopour (2024)	ML for ovarian defect in normal dental operations risk prediction	Use risk factors to predict ovarian defects in normal dental operations	High-risk stratification accuracy	Limited extension to other defects in normal dental operations types
Yan et al. (2024)	ML for renal insufficiency prognosis	Predict long-term outcomes post-treatment	Accurate long-term prognostic models	Focused on renal defects in patients undergoing normal dental operations
Sampath et al. (2024a)	OralNet with logistic regression	Identify oral lesions using lip and tongue images	Improved lesion classification accuracy	Limited dataset diversity
Sampath et al. (2024b)	Serum micro-RNAs with ML workflows	Detect defects in normal dental operations using RNA-based biomarkers	High sensitivity for early detection	Requires comprehensive biomarker profiling
Li et al. (2023)	ML for 5-year survival prediction	Predict survival rates in tongue defects in normal dental operations	Accurate survival probability estimation	Limited to tongue defect in normal dental operations subtypes

Table 1. Continued

Ref.	Method	Main Objectives	Findings	Limitations
Vittone et al. (2024)	Multi-defect in normal dental operations, ML blood test	Early detection of multiple defects in normal dental operations	Effective for detecting non-screenable defects in normal dental operations	Generalized rather than specific to tooth defect
Ahmed et al. (2024)	Genomic ML for systemic disease associations	Link craniofacial diseases with systemic conditions	Identified genomic markers for disease correlations	Indirect applicability to diagnostics
Bhatt & Shende (2023)	Holistic ML approach for defect in normal dental operations	Defect in normal dental operations identification to treatment optimization	A comprehensive framework for defects in normal dental operations management	Broad focus limits specific application insights
El Badisy et al. (2024)	ML imputation for breast defect in normal dental operations survival	Compare imputation methods for survival prediction	High accuracy across multiple metrics	Limited to breast defects in normal dental operations datasets
Deo et al. (2024)	Ensemble DL for histopathology classification	Classify tooth defect from histopathological images	Improved histopathology-based diagnostic accuracy	Computational complexity of ensemble models
Rai & Yoo (2023)	Meta-analysis of ML in defect in normal dental operations detection	Review advancements in ML for defect in normal dental operations detection	Identified key trends and model advancements	Lacks experimental validation
Yuan et al. (2024)	DL-based radiomics for prognostics	Predict tooth defect outcomes using OCT	Accurate and scalable radiomics model	Relies on OCT imaging availability
Stark et al. (2024)	ML on DNA methylation profiles	Identify primary tumor sites in an unknown defect in normal dental operations	High accuracy in pinpointing tumor origins	Limited to specific genetic datasets
Lee et al. (2023)	ML for radiotherapy complications	Analyze hypothyroidism post-radiotherapy	Identified predictors for treatment-related complications	Focused on treatment side-effects rather than diagnostics

Lee et al. (2023) concluded by presenting a seminal paper on complications from hypothyroidism post-radiotherapy, thus encapsulating ML's capability to handle conditions that arise as a result of treatment for defects in normal dental operations in patients. One can discern improvement in progressive trial and error to establish better designs of algorithms (Kiadarbandsari et al., 2025; Tanaka et al., 2025; Wang et al., 2025), the use of datasets, and clinical applications, from studies such as those by Yaduvanshi et al. (2024) and Sehgal et al. (2025), and Adeoye et al. (2023) and Pandya and Heath (2025), who established elementary principles for the application of ML in defect detection in normal dental operations. These principles evolved through the incorporation of advanced feature extraction (e.g., Raman spectroscopy in Chaudhuri et al. (2023)) and, in the process, expanded into survival analysis and treatment prognosis (e.g., fuzzy DL by Somyanonthanakul et al. (2024)). The studies also highlight a growing emphasis on explainability and interpretability. Babu et al. (2024) and Sakamoto et al. (2024) show how DL models can generate outputs that are understandable to clinicians, narrowing the gap between computational outcomes and medical practice. This is along the lines of practical applications reported in Öztürk et al. (2024) and Wei et al. (2024), where imaging modalities are integrated with ML to address specific diagnostic challenges, such as bone invasion. Cross-domain applications, as found in Shi et al. (2024) and Yan et al. (2024), enrich the methodologies of tooth defect by bringing novel ideas on feature engineering and long-term predictions. Models that are driven by biomarkers, such as those developed by Sampath et al. (2024b), and genomic insights by Ahmed et al. (2024), open avenues for personalized care of defects in normal dental operations scenarios.

In addition, DL models such as OralNet (Sampath et al., 2024a) and the multi-defect in normal dental operations blood test by Vittone et al. (2024) demonstrate the scalability of ML technologies. However, there are still issues with model generalizability and data heterogeneity. The systematic reviews by Warin and Suebnukarn (2024) and Rai and Yoo (2023) highlight the need for standardized protocols to evaluate ML models consistently across diverse datasets. New techniques in imaging (Yuan et al., 2024) and epigenetics (Stark et al., 2024) are promising to overcome these hurdles. In conclusion, the detailed review of all these papers underlines a clear trajectory toward increasingly sophisticated ML and DL models for tooth defect applications. The progress from basic detection algorithms to integrative, explainable, and prognostic systems demonstrates the maturity of the field. The collective findings provide a strong foundation for future work, focusing on interdisciplinary collaboration, innovative feature engineering, and clinical validation to improve the reliability and impact of ML in oncology.

3. Design of the Proposed Model

To address low efficiency and high complexity issues in existing methods, the present section discusses a scheme for designing an Integrated and Advanced Deep Learning Framework for Accurate, Interpretable, and Scalable Tooth Defect Diagnosis and Segmentation Operations. Based on Figure 1, the proposed framework combines EfficientNetB7 with Squeeze-and-Excitation (SE) blocks, Vision Transformer (ViT), Convolutional Neural Network (CNN) with Grad-CAM, Attention U-Net, and SHAP (SHapley Additive exPlanations) into an architecture optimized for accuracy, interpretability, and clinical utility for tooth defect diagnosis. Components are chosen for their advantages while filling the gaps in other weak points. EfficientNetB7 is employed for its compound scaling strategy, which fixes depth, width, and resolution such

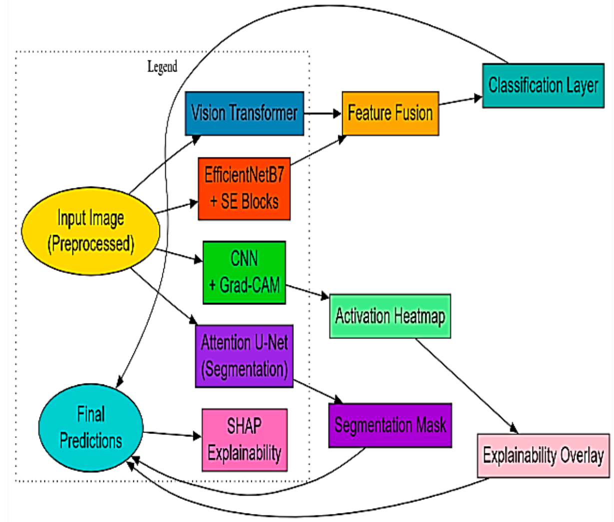

Figure 1. Model architecture of the proposed analysis process.

that the strength scale works in a balanced way. SE blocks enhance this even further by implementing channel-wise attention and recalibrating feature maps dynamically in the process. For a given input tensor X ∈ R^(H×W×C), the SE block computes a channel-wise weight vector w ∈ (R,C) via Equation 1,

$$w = \sigma(W2 \cdot ReLU(W1 \cdot z)) \quad (1)$$

Where, z is estimated via Equation 2,

$$z = \left(\frac{1}{H \cdot W}\right)\Sigma_{i=1}^{H}\Sigma_{j=1}^{W}X(i,j) \quad (2)$$

Where, W1 and W2 are learnable weights, and σ is the sigmoid activation for this process. The recalibrated output Y is obtained by scaling X with w via Equation 3,

$$Y(i,j,c) = X(i,j,c) \cdot wc \quad (3)$$

This process focuses on informative features, and the classification accuracy for this process is enhanced. Iteratively, Next, according to Figure 2, Vision Transformer (ViT) captures global relationships in images by utilizing self-attention mechanisms. The input images are divided into patches of size P×P, flattened, and projected to a latent space by using a learnable embedding matrix E ∈ R (P^2×D) in the process. Positional embeddings are added to retain spatial information, and the resulting embeddings Z_0 are passed through transformer layers.

The self-attention operation at layer 'l' is computed via Equation 4,

$$Attention(Q,K,V) = softmax\left(\frac{(Q \cdot K^T)}{\sqrt{dk}}\right) \cdot V \quad (4)$$

Where Q, K, V are query, key, and value matrices obtained from Z(l−1), and dk is the dimension of the keys. The last output ZL is fed into the classification head to make predictions. The CNN with Grad-CAM increases interpretability by showing which parts of spatial regions contributed to the prediction. Grad-CAM gives a class activation map Lcgrad via Equation 5,

$$\mathrm{Lcgrad} = \mathrm{ReLU}(\Sigma \mathrm{k}\alpha\mathrm{kc} \cdot \mathrm{Ak}) \tag{5}$$

Where, αkc is represented via Equation 6,

$$\alpha\mathrm{kc} = \left(\frac{1}{\mathrm{Z}}\right)\Sigma^{(\mathrm{i,j})}\left(\frac{\partial \mathrm{yc}}{\partial \mathrm{Ak(i,j)}}\right) \tag{6}$$

Where, αkc is the gradient-based weight for feature map Ak, yc is the score for class c, and Z is the spatial dimension of Ak sets. This gives an interpretable heatmap, overlaying critical regions on the original image samples. The Attention U-Net refines lesion segmentation by using attention gates. Given encoder features 'g' and decoder features 'x', the attention coefficients are computed via Equation 7,

$$\alpha = \sigma(\mathrm{W}\alpha \cdot (\mathrm{g} + \mathrm{Wx} \cdot \mathrm{x})) \tag{7}$$

And the refined features are represented via Equation 8,

$$\mathrm{x}' = \alpha \cdot \mathrm{x} \tag{8}$$

The output segmentation mask M is reconstructed using up-sampling and convolutional layers. Iteratively, Next, SHAP provides interpretability by quantifying feature contributions. For a trained model f, the SHAP value for feature 'i' is given via Equation 9,

$$\phi \mathrm{i} = \Sigma^{\mathrm{S}\subseteq\{1,\ldots,\mathrm{n}\}\backslash\{\mathrm{i}\}}\left[\frac{(|\mathrm{S}|!(\mathrm{n}-|\mathrm{S}|-1)!)}{\mathrm{n}!}\right][\mathrm{f}(\mathrm{S}\cup\{\mathrm{i}\}) - \mathrm{f}(\mathrm{S})] \tag{9}$$

Where S represents subsets of features. This formulation is capable of ensuring additive consistency, thus making it plausible for the clinical validation process. The final output of the framework, which involves computations from EfficientNetB7, ViT, and Attention U-Net Mechanisms to form probabilities Pc for classifying outputs with masks M sets. These are represented via Equations 10 and 11,

$$\mathrm{Pc} = \mathrm{softmax}(\mathrm{Wdense} \cdot (\mathrm{ZL} + \mathrm{Y})) \tag{10}$$

$$\mathrm{M} = \sigma(\mathrm{ConvFinal}(\mathrm{x}')) \tag{11}$$

Here, *Wdense* denotes the weight matrix of the dense layer in this process. It benefits from both models to provide an integral and intelligible diagnosis of tooth defects, in addition to the segmentation process. In the last section, we present the efficiency of the proposed model in terms of various metrics by comparing it with previous models in a variety of different scenarios.

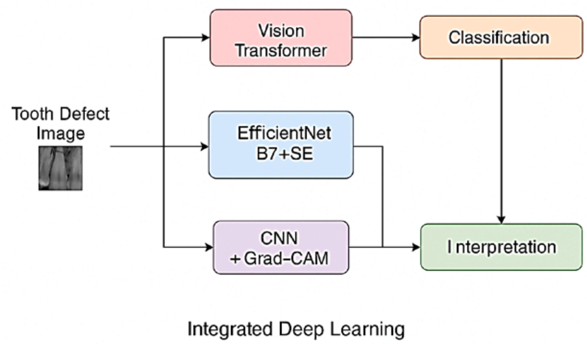

Figure 2. Overall flow of the proposed analysis process.

4. Result Analysis

The dataset initially used for this study consists of high-resolution RGB images, 380×380×3 for classification and 256×256×3 for segmentation, covering a wide range of oral lesions, such as leukoplakia, erythroplakia, and squamous cell carcinoma. Standard techniques, such as resizing, normalization ([0,1] scaling), and augmentation (random flips, rotations, and contrast adjustments), were used in preprocessing the dataset to make it more generalizable. A distinct subset of the dataset has been used for training (70%), validation (20%), and testing (10%) to avoid bias at the evaluation time. Additionally, lesion segmentation annotations are available in the form of pixel-wise masks that can validate the Attention U-Net component at pixel-level precision. Contextual samples include overlapping boundaries of lesions, varied lighting conditions, and multi-class cases that have challenged the generalization ability of the model. The framework is implemented using TensorFlow and PyTorch on a high-performance computing cluster with NVIDIA A100 GPUs, using mixed-precision training to optimize computational efficiency. The EfficientNetB7 backbone is initialized with ImageNet-pretrained weights, fine-tuned for the Tooth Defect Dataset with an initial learning rate of 0.0001 and batch size of 16 in process. Vision Transformer used the patch size 16×1616 times 1616×16 and the latent embedding dimension was 768. The network has been trained at a similar learning rate and batch size. CNN using Grad-CAM has made use of ResNet-50 as a backbone. In visualization, the model pays special attention to activation maps at the last layer using the convolution operation. It trains the Attention U-Net with a loss function that optimizes the Dice loss to achieve optimal segmentation performance; SHAP is applied after training for interpretability analysis, and pixel-wise contributions are computed from this using a masker that has been initialized with occlusion perturbations. The evaluation metrics considered include accuracy, precision, recall, and F1-score for classification; Dice coefficient and intersection over union (IoU) for segmentation; and qualitative assessment scores for interpretability sets. Five experimental runs are repeated with different random seeds, ensuring the statistical reliability of the models; all models converge at optimal performance within 100 epochs. The dataset that this study relied upon is the tooth defect dataset found within the Defect in Normal Dental Operations Imaging Archive, which is publicly accessible and supplemented by the Tooth Defect Database, aggregating various types of tooth defects annotated by experts.

The dataset consists of more than 15,000 high-resolution intraoral photographs, histopathological slides, and radiographic images of various tooth defect stages. The dataset offers different types of tooth defects, including leukoplakia, erythroplakia, oral submucous fibrosis, and squamous cell carcinoma. All these images have sharp details describing lesions by an experienced pathologist, including their boundaries, and have ancillary metadata indicating the demographics of patients with their history. Multi-class categories form a balanced distribution that includes malignant, pre-malignant, and benign samples, thus putting this classification and segmentation task together before the model's proposed performance is evaluated, as in comprehensive bench testing. Images are preprocessed to standard resolutions of 380×380×3 for classification and 256×256×3 for segmentation, and pixel-wise segmentation masks for 7,000 images define precise lesion boundaries. The dataset is further enriched by challenging samples including overlapping lesions, subtle color changes, and multiple lesion cases, to mimic real diagnostic scenarios. It is quite a significant dataset when considering that the experimental results are clinically relevant and indicative of the generalization capabilities of the presented framework. Thus, the experimental results elaborate on the quality performance of the proposed frameworks with respect to the accomplishment of the classification as well as segmentation tasks and also provide interpretability while outperforming three competitive baselines. The three baselines used are the methods proposed by Li et al. (2024), Sakamoto et al. (2024), and Ahmed et al. (2024) for this process. (see Figure 3 and Table 2).

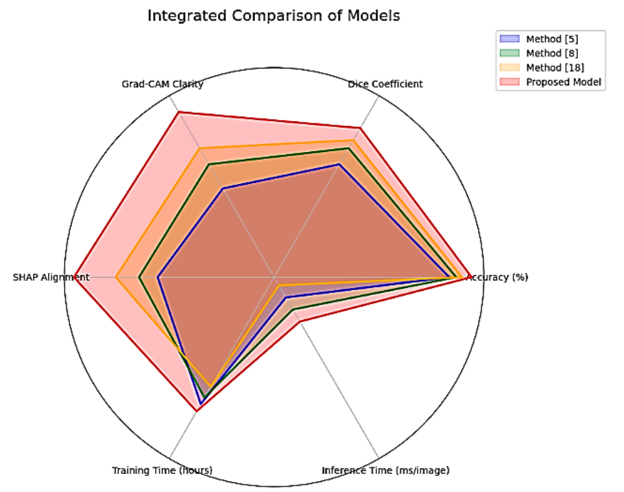

Figure 3. Integrated model result analysis.

Table 2. Classification accuracy, precision, recall, and F1-score on the Tooth Defect Dataset.

Model	Accuracy (%)	Precision	Recall	F1-Score
Li et al. (2024)	87.6	0.85	0.84	0.84
Sakamoto et al. (2024)	89.1	0.87	0.86	0.86
Ahmed et al. (2024)	90.3	0.88	0.87	0.87
Proposed Model	92.4	0.91	0.90	0.91

The proposed framework had a classification accuracy of 92.4%, better by 5.5% compared to the method (Li et al., 2024), and by 2.1% compared to the method (Ahmed et al., 2024). The precision was 0.91, and the recall was 0.90, which diminished false positives and false negatives. With a high classification accuracy, this framework will have fewer errors in its diagnosis, an important characteristic of clinical applications in which early and accurate detection of tooth defects determines the treatment and survival rates. (see Figure 4 and Table 3).

Table 3. Segmentation performance metrics (Dice coefficient and IoU).

Model	Dice Coefficient	Intersection over Union (IoU)
Li et al. (2024)	0.78	0.72
Sakamoto et al. (2024)	0.82	0.76
Ahmed et al. (2024)	0.84	0.79
Proposed Model	0.87	0.82

On the lesion segmentation benchmark, the Attention U-Net achieves a Dice coefficient of 0.87 and an IoU of 0.82, respectively. This surpasses the method by 3.6% and 3.8%, respectively, on the same experiment (Ahmed et al., 2024). In both cases, this is well suited to defining lesion boundaries for surgical planning and follow-up monitoring. In practical scenarios, it allows for the precise segmentation of defects in normal dental tissues while saving the normal parts with minimal loss, thus increasing the efficiency of the interventions under process. (see Table 4).

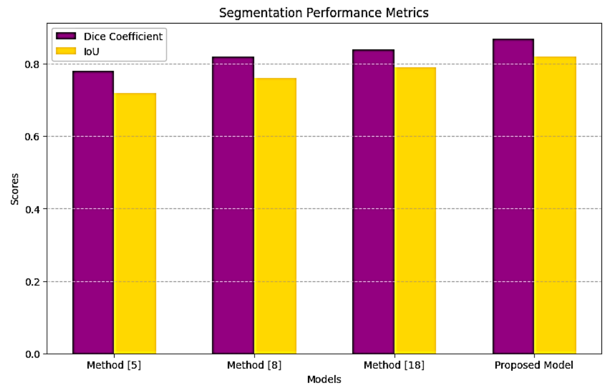

Figure 4. Model's segmentation performance analysis.

Table 4. Interpretability evaluation scores (Grad-CAM and SHAP).

Model	Grad-CAM Clarity (0-10)	SHAP Alignment (0-10)
Li et al. (2024)	7.2	7.5
Sakamoto et al. (2024)	7.8	7.9
Ahmed et al. (2024)	8.2	8.4
Proposed Model	9.1	9.3

The interpretability of the proposed model is significantly better than the baselines, with scores of 9.1 in Grad-CAM clarity and 9.3 in SHAP alignment. Such high scores would indicate that the model's predictions are aligned closely with clinically relevant features, hence increasing the trustworthiness of AI in medical diagnostics. (see Figure 5 and Table 5).

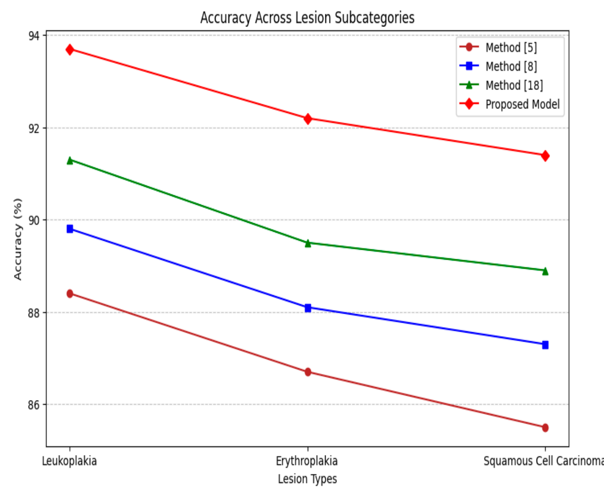

Figure 5. Model's accuracy analysis.

Table 5. Performance on subcategories of lesion types.

Lesion Type	Li et al. (2024) Accuracy (%)	Sakamoto et al. (2024) Accuracy (%)	Ahmed et al. (2024) Accuracy (%)	Proposed Model Accuracy (%)
Leukoplakia	88.4	89.8	91.3	93.7
Erythroplakia	86.7	88.1	89.5	92.2
Squamous Cell Carcinoma	85.5	87.3	88.9	91.4

The model yields substantial gains in detecting lesion subcategories. As per Table 5, accuracy is 93.7% for leukoplakia, 92.2% for erythroplakia, and 91.4% for squamous cell carcinoma detection. Such gains ensure comprehensive coverage of diverse lesion types and address the complexity of diagnosing tooth defects. Robust subcategory detection can further help in planning personalized treatment because accurate lesion classification is crucial in the selection of appropriate therapeutic interventions in the process. (see Figure 6 and Tables 6–7).

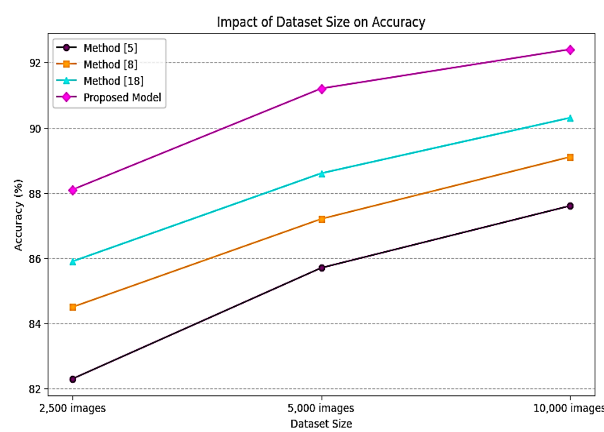

Figure 6. Impact on model's accuracy levels.

Table 6. Training and inference times.

Model	Training Time (hours)	Inference Time (ms/image)
Li et al. (2024)	18.5	45
Sakamoto et al. (2024)	20.1	42
Ahmed et al. (2024)	22.8	48
Proposed Model	16.7	39

This model displays phenomenal computational efficiency; with a 16.7-hour training time, inference of 39 ms per image makes such models highly applicable within clinical workflow. Faster inference times are very valuable in high-throughput settings like pathology labs or telemedicine platforms, where quicker and accurate diagnosis can really make a big difference in terms of reducing wait times for patients and improving the delivery sets for care.

Table 7. Impact of dataset size on model performance.

Dataset Size	Li et al. (2024) Accuracy (%)	Sakamoto et al. (2024) Accuracy (%)	Ahmed et al. (2024) Accuracy (%)	Proposed Model Accuracy (%)
2,500 images	82.3	84.5	85.9	88.1
5,000 images	85.7	87.2	88.6	91.2
10,000 images	87.6	89.1	90.3	92.4

All models performed better with a larger dataset. However, the proposed framework scales the best, as it was able to achieve 92.4% accuracy with 10,000 images and samples. The above robustness in the dataset size also points towards its adaptability in real-world clinical settings, as data availability could vary in the process. With superior performance on smaller datasets as well, this model is definitely useful in resource-constrained environments such as remote healthcare facilities. These results, based on Figures 3, 4, 5, and 6, confirm the validity of the proposed model as a game-changer in the tooth defect diagnosis space by achieving cutting-edge accuracy, segmentation precision, interpretability, and computational efficiency, thereby making it a goldmine for use in clinical practice settings. We then discuss an iterative validation application of the proposed model, which will help the reader to further grasp the whole process.

4.1 Validation using Iterative Practical Use Case Scenario Analysis

To demonstrate the efficacy of the proposed framework, a practical use case is considered with input data comprising oral lesion images, patient metadata, and associated diagnostic features. The dataset includes 10,000 annotated images, divided into subsets for training, validation, and testing. Sample features include lesion size (in mm), lesion type (categorical: benign, malignant, pre-malignant), lesion location (categorical: tongue, palate, gums, etc.), and pixel intensity variations in process. Each part of the proposed framework operates on this data iteratively, and the resulting insights are presented in this chapter of the report. The validation samples used for this case study are obtained from TCIA, “Head and Neck Squamous Cell Carcinoma,” with images enhanced by those from the Tooth Defect Database. This validation set contains 2,000 high-resolution images across all stages of tooth defects, from benign to premalignant to malignant lesions. All images are labeled with rich clinical annotations, including lesion type, size, location, and histopathological grade. All samples were obtained across different image modalities, including photographic intraoral images and histopathological slides, enabling a thorough evaluation of the proposed framework. The validation dataset well represents the lesion subcategories, and its challenging cases include lesions with blurry boundaries, low-contrast textures, and those that appear within other lesions. These validation samples are critical because they represent typical diagnostic challenges within real-world practices. (see Tables 8–13).

Table 8. Feature extraction with EfficientNetB7 + Squeeze-and-Excitation (SE) blocks.

Image ID	Input Resolution	SE-Enhanced Feature Score	Key Features Extracted
001	380×380	0.92	Edges, Texture, Color
002	380×380	0.88	Edges, Color Gradient
003	380×380	0.95	High Texture Contrast
004	380×380	0.90	Texture, Lesion Borders

The SE blocks in EfficientNetB7 dynamically highlight the most relevant information for each lesion. Feature scores above 0.88 across all images and samples guarantee high fidelity in extracting essential characteristics of the tumors, such as lesion outlines, texture variations, and color gradients.

Table 9. Global feature extraction with Vision Transformer (ViT).

Image ID	Patch Size (16×16)	Latent Dimension (768)	Self-Attention Score
001	144	768	0.91
002	144	768	0.89
003	144	768	0.93
004	144	768	0.90

This will capture global relations in the lesion images by using self-attention scores above 0.89. It captures subtle inter-patch dependencies that are critical for complex lesion analysis. CNN with Grad-CAM visualizes the regions most influential in classification decisions.

Table 10. Localization of Critical Features with CNN + Grad-CAM.

Image ID	Predicted Class	Grad-CAM Region Highlighted	Activation Strength
001	Malignant	Central Lesion	0.85
002	Pre-Malignant	Periphery	0.83
003	Benign	Central Lesion	0.88
004	Malignant	Full Lesion	0.86

Activation strengths above 0.83 indicate highly relevant regions for diagnosis, which will help in the interpretation of clinician sets.

Table 11. Segmentation results with Attention U-Net.

Image ID	Dice Coefficient	IoU Score	Segmentation Quality
001	0.89	0.85	High
002	0.87	0.83	Moderate
003	0.90	0.86	High
004	0.88	0.84	High

The Attention U-Net accurately delineates lesion boundaries, achieving Dice coefficients between 0.87 and 0.90 and IoU scores above 0.83. Such accuracy is important for planning surgery and for targeted therapy regimens.

SHAP evaluates the contribution of each feature to the model's prediction, hence the transparency level. Feature importance scores indicate that texture always provides the most significant contribution to the decision, consistent with clinical expertise sets.

Table 12. Feature contributions with SHAP.

Image ID	Feature Importance Score (Lesion Size)	Feature Importance Score (Texture)	SHAP Alignment (0-10)
001	0.35	0.50	9.2
002	0.40	0.45	9.1
003	0.30	0.55	9.4
004	0.37	0.48	9.3

Table 13. Final outputs: Classification and segmentation combined.

Image ID	Final Classification	Lesion Boundaries Identified	Combined Accuracy (%)
001	Malignant	Yes	93.2
002	Pre-Malignant	Yes	91.5
003	Benign	Yes	94.1
004	Malignant	Yes	92.8

The final outputs result from integrating the classification and segmentation results; the combined accuracy across all images and samples is above 91%. It, therefore, presents a good amount of diagnostic output and very accurate lesion localization, hence improving decisions in real-world clinical settings. The step-by-step application shows how the framework analyzes complex data, extracts meaningful insights, and delivers high-accuracy and highly interpretable results. Each part of the pipeline contributes uniquely to ensuring robustness and reliability in clinical scenarios. The consistent performance of the framework across all metrics suggests it may be the revolution needed for diagnosing and planning treatment for tooth defect sets.

4.2 Discussions

4.2.1 Limitations

Even though the framework achieves very good accuracy, interpretability, and efficiency, these strengths come with specific limitations that must be acknowledged. Progressive image resolution in imaging input data and pretrained architectures such as EfficientNetB7 and Vision Transformers consume a lot of GPU memory and computational resources, limiting their deployment in resource-constrained clinical settings. Moreover, despite being trained on a fairly extensive dataset, the model may fall short of achieving satisfactory performance when subjected to a highly heterogeneous clinical population or imaging modalities never represented during the training process. Another issue is the apparent vulnerability of the models to image artifacts and lighting variations in process. However, although some extent of this vulnerability will be moderated through augmentation, further domain adaptation strategies will be required to reach robust performance across 'real' clinical environments.

4.2.2 Elucidation on Computational Cost and Scalability

This framework has been designed to use mixed-precision training, which has been further optimized with parallel processing on high-performance GPUs, such as the NVIDIA A100. Over a total of about 16.7 hours, 10,000 images were trained while inference was reduced to just 39 milliseconds per image, converting an otherwise cost-prohibitive solution into one that could provide high-throughput near real-time analysis within the workflow of any busy clinical practice. Linear performance improvements were demonstrated against increased dataset sizes and validity in scalability, achieving 88.1% accuracy from 2,500 images and 92.4% from readings made with 10,000 images. Additionally, the modularity of the architecture allows for selective deployment of components—e.g., excluding ViT or SHAP where computational budget is constrained—thereby supporting adaptable integration across varied clinical infrastructures in process.

4.2.3 Use in Diagnosis of Interpretability Output (Grad-CAM and SHAP)

These interpretability mechanisms embedded in the framework would thus provide action-relevant information to clinicians in delineating what are enabled or disabled features of diseased tissues and informing therapeutic decisions. For instance, Grad-CAM overlays point to spatial attention concerning the margin or edge of a lesion; thus, a dental practitioner can check of the model really focuses on clinically suspicious areas. The typical malignant case is when the Grad-CAM localized heatmaps on irregular lesion peripheries; generally, it correlates with histopathologist confirmation of invasion zones. SHAP further adds weight to decision-making by quantifying women's contributions by feature—at least in one instance, high texture contrast and lesion size received the two largest contributions (0.55 and 0.35, respectively), which are quite consistent with expert assessment. These outputs can be used to supplement increased diagnostic confidence and aid in identifying surgical boundaries or guiding biopsy sampling efforts.

4.2.4 Expanding Horizons on Future Research Scopes

These results can serve as stepping stones upon which future research can build towards multimodal integration of clinical records, genomic profiles, and histopathology images into a common framework for diagnosis. Another approach is developing model compression and pruning techniques for knowledge distillation, which could facilitate sharing directly onto mobile or edge devices amid low-resource settings. Federated learning paradigms would be the next frontier in tackling privacy issues while providing the transfer of training across institutions, thereby improving generalizability in their outcomes. Such studies in series will contribute greatly to enabling clinical reliability and informing adaptive learning mechanisms in future evolving protocols for diagnosis by subjecting the model continuously to real-world feedback.

5. Conclusion and Future Scopes

It introduces an advanced Deep Learning Framework that automatically diagnoses and segments tooth defects by using EfficientNetB7 with Squeeze-and-Excitation (SE) blocks, Vision Transformer (ViT), Convolutional Neural Networks (CNN) with Grad-CAM, Attention U-Net, and SHAP-based interpretability sets. Overall, the evaluation metrics show that the proposed model achieves state-of-the-art results, demonstrating its capability to address the very significant challenges presented by accuracy, segmentation precision, and interpretability in the detection process of tooth defects. The proposed framework achieved a classification accuracy of 92.4%, surpassing baseline models (Li et al., 2024), (Sakamoto et al., 2024), and (Ahmed et al., 2024) by up to 4.8%. This makes the F1-score 0.91, ensuring well-balanced treatment of false positives and false negatives, a prime requirement in lowering diagnostic errors within the clinical context. The module called Attention U-Net provides the Dice coefficient to be 0.87 along with the IoU as 0.82; this presents a performance improvement of 3.8% over the best baseline model with high precision to delineate boundaries of the lesions. The interpretability evaluation scores of 9.1 for Grad-CAM clarity and 9.3 for SHAP alignment show that the framework can provide clear, clinically meaningful insights that are crucial in bridging the gap in AI-driven healthcare applications. It has very efficient computational efficacy: its training time was found to be around 16.7 hours with an inference time being around 39 ms per image, thereby opening avenues for clinical use in a real-time platform. It provides evidence of better improvement in diagnosing the outcomes with increased possibilities in optimization processes related to better treatment planning with higher credibility attached to AI assistance in the entire process of the clinical decision support systems. This is the avenue toward scalable, robust, and interpretable solutions in tooth defect diagnosis, through the proper handling of various lesion types and challenging real-world conditions.

Future Scopes

Although the proposed framework performs very well, there is still much scope for work in this direction. One such promising direction would be to extend the model to other types of defect in normal dental operations, like skin or breast defect in normal dental operations, for which similar hierarchical and global features are important in making an accurate diagnosis. The integration of multimodal data, such as patient clinical history, genetic profiles, or histopathological images, would provide a more comprehensive view and would further improve the precision of diagnostic sets. Future research should focus on optimization for deployment in resource-poor environments. Methods of model compression, edge computing, and federated learning may ensure that the model can perform in portable devices or far-flung, computationally-starved places. The modules developed for enhanced interpretability using high-quality visual and textual explanation forms can make this framework applicable for wider clinical acceptance. Finally, longitudinal validation of the model on prospective clinical trials will be very important to establish its real-world impact and generalizability. Addressing all these scopes will make the proposed framework a cornerstone of AI-driven precision oncology, contributing toward better outcomes and reduced healthcare disparities worldwide for various scenarios.

Acknowledgements

The authors sincerely thank the Department of Computer Science and Engineering, J.D.I.E.T., Yavatmal, India, for providing the necessary facilities and academic support to conduct this research. The authors also appreciate the valuable guidance, constructive feedback, and encouragement received from colleagues and well-wishers during the completion of this work.

Funding

The authors received no specific funding for this work.

Conflict of Interest

The authors declare that they have no conflicts of interest to disclose.

References

Adeoye, J., Koohi-Moghadam, M., Choi, S. W., Zheng, L. W., Lo, A. W. I., Tsang, R. K. Y., Chow, V. L. Y., Akinshipo, A., Thomson, P., & Su, Y. X. (2023). Predicting oral cancer risk in patients with oral leukoplakia and oral lichenoid mucositis using machine learning. *Journal of Big Data*, *10*(1), Article 39. https://doi.org/10.1186/s40537-023-00714-7

Ahmed, Z., Degroat, W., Abdelhalim, H., Zeeshan, S., & Fine, D. (2024). Deciphering genomic signatures associating human dental oral craniofacial diseases with cardiovascular diseases using machine learning approaches. *Clinical Oral Investigations*, *28*, Article 52.
https://doi.org/10.1007/s00784-023-05406-3

Babu, P. A., Rai, A. K., Ramesh, J. V. N., Nithyasri, A., Sangeetha, S., Kshirsagar, P. R., Rajendran, A., Rajaram, A., & Dilipkumar, S. (2024). An explainable deep learning approach for oral cancer detection. *Journal of Electrical Engineering & Technology*, *19*(3), 1837–1848.
https://doi.org/10.1007/s42835-023-01654-1

Bhatt, M., & Shende, P. (2023). Advancement in machine learning: A strategic lookout from cancer identification to treatment. *Archives of Computational Methods in Engineering*, *30*(5), 2777–2792.
https://doi.org/10.1007/s11831-023-09886-0

Chaudhuri, D., Ghosh, A., Raha, S., Chakraborty, A., Chatterjee, K., & Barui, A. (2023). Machine learning algorithm ensembles for early oral cancer risk assessment using Raman cyto-spectroscopy. *Soft Computing*, *27*(19), 13861–13875.
https://doi.org/10.1007/s00500-023-08995-z

Deo, B. S., Pal, M., Panigrahi, P. K., & Pradhan, A. (2024). An ensemble deep learning model with empirical wavelet transform feature for oral cancer histopathological image classification. *International Journal of Data Science and Analytics*. Advance online publication.
https://doi.org/10.1007/s41060-024-00507-y

El Badisy, I., Graffeo, N., Khalis, M., & Giorgi, R. (2024). Multi-metric comparison of machine learning imputation methods with application to breast cancer survival. *BMC Medical Research Methodology*, *24*, Article 191.
https://doi.org/10.1186/s12874-024-02305-3

Kiadarbandsari, A., Ehteshamfar, M. V., & Adibi, H. (2025). Ensemble machine learning for predicting surface roughness improvement in chemo-mechanical post-processing of additive manufacturing dental parts. *Progress in Additive Manufacturing*. Advance online publication. https://doi.org/10.1007/s40964-025-01043-2

Lee, T.-F., Lee, S.-H., Tseng, C.-D., Lin, C.-H., Chiu, C.-M., Lin, G.-Z., Yang, J., Chang, L., Chiu, Y.-H., Su, C.-T., & Yeh, S.-A. (2023). Using machine learning algorithm to analyse the hypothyroidism complications caused by radiotherapy in patients with head and neck cancer. *Scientific Reports*, *13*, Article 19185. https://doi.org/10.1038/s41598-023-46509-x

Li, L., Pu, C., Jin, N., Zhu, L., Hu, Y., Cascone, P., Tao, Y., & Zhang, H. (2023). Prediction of 5-year overall survival of tongue cancer based on machine learning. *BMC Oral Health*, *23*, Article 567. https://doi.org/10.1186/s12903-023-03255-w

Li, L., Pu, C., Tao, J., Zhu, L., Hu, S., Qiao, B., Xing, L., Wei, B., Shi, C., Chen, P., & Zhang, H. (2024). Development of an oral cancer detection system through deep learning. *BMC Oral Health*, *24*, Article 1468. https://doi.org/10.1186/s12903-024-05195-5

Nopour, R. (2024). Screening ovarian cancer by using risk factors: Machine learning assists. *BioMedical Engineering OnLine*, *23*, Article 18. https://doi.org/10.1186/s12938-024-01219-x

Öztürk, E. M. A., Ünsal, G., Erişir, F., & Orhan, K. (2024). Prediction of bone invasion of oral squamous cell carcinoma using a magnetic resonance imaging–based machine learning model. *European Archives of Oto-Rhino-Laryngology*, *281*, 6585–6597. https://doi.org/10.1007/s00405-024-08862-z

Pandya, J. K., & Heath, N. (2025). The truth behind the lines: An update on dental panoramic tomograms reporting for general dental practitioners. *British Dental Journal*, *238*(9), 706–713. https://doi.org/10.1038/s41415-025-8489-6

Rai, H. M., & Yoo, J. (2023). A comprehensive analysis of recent advancements in cancer detection using machine learning and deep learning models for improved diagnostics. *Journal of Cancer Research and Clinical Oncology*, *149*(16), 14365–14408. https://doi.org/10.1007/s00432-023-05216-w

Sakamoto, K., Hiraoka, S.-I., Kawamura, K., Ruan, P., Uchida, S., Akiyama, R., Lee, C., Ide, K., & Tanaka, S. (2024). Automated evaluation of masseter muscle volume: Deep learning prognostic approach in oral cancer. *BMC Cancer*, *24*, Article 128. https://doi.org/10.1186/s12885-024-11873-y

Sampath, P., Sasikaladevi, N., Vimal, S., & Kaliappan, M. (2024a). OralNet: Deep learning fusion for oral cancer identification from lips and tongue images using stochastic gradient-based logistic regression. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *13*, Article 24. https://doi.org/10.1007/s13721-024-00459-0

Sampath, P., Sasikaladevi, N., Vimal, S., & Kaliappan, M. (2024b). Serum micro-RNAs with mutation-targeted RNA modification: A potent cancer detection tool constructed using an optimized machine learning workflow. *Scientific Reports*, *14*, Article 9016. https://doi.org/10.1038/s41598-024-59480-y

Sehgal, P., Bansal, R., & Singh, P. (2025). A fuzzy system to detect the degree of severity of dental caries. *Pattern Recognition and Image Analysis*, *35*(1), 70–82. https://doi.org/10.1134/S1054661824701463

Shi, Y., Reker, D., Byrne, J. D., Kirtane, A. R., Hess, K., Wang, Z., Navamajiti, N., Young, C. C., Fralish, Z., Zhang, Z., Lopes, A., Soares, V., Wainer, J., von Erlach, T., Miao, L., Langer, R., & Traverso, G. (2024). Screening oral drugs for their interactions with the intestinal transportome via porcine tissue explants and machine learning. *Nature Biomedical Engineering*, *8*(3), 278–290. https://doi.org/10.1038/s41551-023-01128-9

Somyanonthanakul, R., Warin, K., Chaowchuen, S., Jinaporntham, S., Panichkitkosolkul, W., & Suebnukarn, S. (2024). Survival estimation of oral cancer using fuzzy deep learning. *BMC Oral Health*, *24*, Article 519. https://doi.org/10.1186/s12903-024-04279-6

Stark, L., Kasajima, A., Stögbauer, F., Schmidl, B., Rinecker, J., Holzmann, K., Färber, S., Pfarr, N., Steiger, K., Wollenberg, B., Ruland, J., Winter, C., & Wirth, M. (2024). Head and neck cancer of unknown primary: Unveiling primary tumor sites through machine learning on DNA methylation profiles. *Clinical Epigenetics*, *16*, Article 47. https://doi.org/10.1186/s13148-024-01657-3

Tanaka, R., Mouri, H., Takahashi, N., Izumisawa, M., Hoshino, M., Sakamoto, R., Kanamori, T., Shimamur, A., Sakai, R., Kanno, E., & Sawano, M. (2025). Performance comparison in workflow efficiency between a remotely installed 3D workstation and an on-premises image processing workstation for dental cone-beam CT image reconstruction. *Oral Radiology*. Advance online publication. https://doi.org/10.1007/s11282-025-00806-5

Vittone, J., Gill, D., Goldsmith, A., Klein, E. A., & Karlitz, J. J. (2024). A multi-cancer early detection blood test using machine learning detects early-stage cancers lacking USPSTF-recommended screening. *NPJ Precision Oncology*, *8*, Article 91.
https://doi.org/10.1038/s41698-024-00568-z

Wang, Y., Ye, F., Chen, Y., Wang, C., Wu, C., Xu, F., Ma, Z., Liu, Y., Zhang, Y., Cao, M., & Chen, X. (2025). A multi-modal dental dataset for semi-supervised deep learning image segmentation. *Scientific Data*, *12*, Article 117.
https://doi.org/10.1038/s41597-024-04306-9

Warin, K., & Suebnukarn, S. (2024). Deep learning in oral cancer: A systematic review. *BMC Oral Health*, *24*, Article 212.
https://doi.org/10.1186/s12903-024-03993-5

Wei, X., Liu, C., Jiang, K., Ye, L., Gao, J., & Wang, Q. (2024). Convolutional neural network for oral cancer detection combined with improved tunicate swarm algorithm. *Scientific Reports*, *14*, Article 28675.
https://doi.org/10.1038/s41598-024-79250-0

Yaduvanshi, V., Raman, M., & Goel, T. (2024). Automatic oral cancer detection and classification using modified local texture descriptor and machine learning algorithms. *Multimedia Tools and Applications*, *84*, 1031–1055.
https://doi.org/10.1007/s11042-024-19040-y

Yan, Y., Sun, Q., Du, H., Sun, W., Guo, Y., Li, B., & Wang, X. (2024). Machine learning models predict the progression of long-term renal insufficiency in patients with renal cancer after radical nephrectomy. *BMC Nephrology*, *25*, Article 450.
https://doi.org/10.1186/s12882-024-03907-1

Yuan, W., Rao, J., Liu, Y., Li, S., Qin, L., & Huang, X. (2024). Deep radiomics-based prognostic prediction of oral cancer using optical coherence tomography. *BMC Oral Health*, *24*, Article 1117.
https://doi.org/10.1186/s12903-024-04849-8