Keywords
Dental radiography, Deep learning, Anomaly detection, Explainable AI, OPG analysis, medical image classification

Authors
Dr. Anurag Rana
Assistant Professor, Yogananda School of AI, Computers and Data Sciences, Faculty of Engineering and Technology, Shoolini University-India (profanuragrana@gmail.com)

Dr. Naeem Akhtar
[2]Research Scholar, Shoolini University India (naeem.akhtar078654@gmail.com)

Address for Correspondence
Dr. Anurag Rana
[1*]Assistant Professor, Yogananda School of AI, Computers and Data Sciences, Faculty of Engineering and Technology, Shoolini University-India (profanuragrana@gmail.com)

Received: 09.12.2025
Accepted: 09.01.2026

DOI: 10.1922/EJPRD_2865Rana22

A Hybrid Deep Learning and Anomaly Detection Model for Interpretable Analysis of Dental OPG

Abstract
Dental panoramic radiography (orthopantomogram or OPG) is a vital tool in diagnosing a range of dental pathologies, yet manual interpretation is time-consuming and subject to variability. This study proposes a hybrid framework that combines deep convolutional neural networks (CNNs) with statistical anomaly detection and explainable artificial intelligence (XAI) to enhance both diagnostic accuracy and clinical interpretability. A fine-tuned ResNet-50 model was trained to extract contextual features from OPG images, which were then fused with point anomaly scores generated by an Isolation Forest algorithm. The system was evaluated on a dataset comprising six diagnostic categories, including rare conditions like fractured teeth and infections. Compared to a baseline CNN, the hybrid model demonstrated higher test accuracy (43.26% vs. 35.12%), macro-F1 score (0.21 vs. 0.10), and macro-AUC (0.70 vs. 0.61). XAI tools-Grad-CAM, SHAP, and saliency maps were employed to visualise decision-critical regions, providing transparent, multi-angle explanations aligned with clinical reasoning. The results confirm that the proposed hybrid approach enhances both performance and trustworthiness, making it a practical solution for AI-assisted dental diagnostics. Future research will explore model generalizability using larger datasets and multi-modal imaging.

1. Introduction
Orthopantomography (OPG), also known as dental panoramic radiography (DPR), is a vital type of imaging in dental diagnostic studies, providing a two-dimensional image of the entire maxillofacial area. It is regularly employed in detecting a wide variety of conditions, including caries, impacted teeth, infections, fractured roots and developmental anomalies. These images, however, are heavily dependent on clinical expertise in their interpretation and are prone to both inter- and intra-observer variability. To address this problem the past few years, have experienced a significant increase in the integration of artificial intelligence (AI) into dental diagnostics, especially due to the popularity of deep learning methods.
CNNs have played a role in the development of automated dental image recognition. Their capability of extracting hierarchical spatial features has made them very applicable in use in dental pathology detection, such as caries detection, tooth segmentation and anomaly classification. Research conducted recently has shown that CNNs are useful in the detection of dental restorations and cavities using panoramic radiographs with promising accuracy and reliability (1,2). Moreover, the latest architectures, such as transformer-based networks, have facilitated the division of complicated anatomical features in OPG images, which provided a better definition of dental components (3).
The systematic reviews verify a rapid development of AI-based tools in the field of dentistry and their increasing potential in clinical adoption (4). Simultaneously, a number of studies investigated frameworks that can use deep learning to identify and classify abnormalities in dental images, as well as rare and subtle lesions (5). Such developments predetermined the more intelligent, automated, and reproducible diagnostic systems.

In addition to the impressive development, there are still several major challenges during the implementation of AI systems in dental radiology.

A majority of deep learning models are black boxes that do not provide much information as to how diagnostic predictions are formed. Such opacity limits their use in clinical practices, where accountability and traceability are vital. Besides, CNNs have difficulties with unbalanced data sets, especially when rare pathologies such as root fractures or infections are to be detected, and they may cause overfitting and weak generalisation (6).

The use of CNNs to use spatial (contextual) features is also the source of another limitation. Although CNNs have the advantage of identifying contextual anomalies like misaligned and affected teeth, they fail to identify point anomalies that do not conform to the statistical norms, but are not characterised by any apparent spatial irregularities. To overcome this, we can present the use of statistical models such as the Isolation Forests that can be used to complement the CNNs in a way that outliers will be labelled according to the distributions of features, hence enhancing sensitivity to rare cases (7).

Moreover, the inability to explain most of the AI models diminishes their trustworthiness and adoption. Gradient-weighted Class Activation Mapping (Grad-CAM) and SHapley Additive exPlanations (SHAP) are methods of interpretation that provide understanding of the decision-making process of the model and show statistically significant regions of impact in the input image (8,9). They need to be incorporated into diagnostic pipelines to be compliant in terms of regulations as well as clinical validation.

The proposed study is devoted to the creation of a hybrid model of detecting dental pathology based on the OPG images that would incorporate CNN-based contextual analysis with statistical anomaly detection. The system is also enriched with XAI techniques to provide transparency and interpretability. The dataset consists of a wide variety of labelled OPG images that belong to six major diagnostic classes, including Healthy Teeth, Caries, Impacted Teeth, Broken Down Crown/Root (BDC/BDR), Fractured Teeth and Infection.

Its use is now limited to 2D panoramic radiographs but excludes 3D imaging modalities, e.g. cone-beam computed tomography (CBCT). Although the proposed system will use visual explanation tools, it is not yet equipped with user feedback and decision revision processes. The aspects can be extended in future research to improve clinical collaboration.

The importance of this study is that it employs a hybrid modelling approach, combining both deep learning and statistical anomaly detection to enhance the level of diagnostic performance. Combining CNN-learned spatial features and outlier detection by Isolation Forests, the model will become more robust to recognise both common and uncommon pathologies (10). This two-detector strategy is a reduction of the weaknesses of single-modality systems, and it also improves the model's generalizability.

Moreover, the combination of XAI frameworks, including Grad-CAM and SHAP, offers a two-fold degree of interpretability, including spatial and pixel-wise, which is essential in clinical decision-making. Such visualisations allow the dentists to make sense of the predictions of the AI and trust it, which helps to overcome the obstacle between automated solutions and professional beliefs (11). The study, therefore, advances the creation of AI systems that are accurate and actionable and clinically transparent.

The viability of the suggested system is accomplished through the application of strong measures, such as accuracy, macro-F1 score, and macro-AUC. This guarantees that performance is assessed in all classes, including the underrepresented classes, which characterises a fairer and more clinical assessment (12).

Research Objectives

The proposed research paper will develop a strong, understandable, and hybrid AI model to identify dental pathology in OPG radiographs. The particular research objectives are:

- To design a hybrid OPG image-based system that integrates CNN feature extraction (contextual anomalies) with statistical image-based anomaly features (point anomalies) for dental pathology detection.
- To enhance clinical interpretability using explainable AI (XAI) techniques such as Grad-CAM and SHAP to visualise decision-critical regions on OPG images.
- To validate the system's performance using accuracy, macro-F1 score, and macro-AUC, and to assess improvements over a baseline CNN-only model.

2. Methodology

This study suggests a hybrid deep learning architecture, which combines learned features of contextual characteristics and statistical anomaly detection to ease automated pathology diagnosis, based on orthopantomogram (OPG) images on dental. The model is constructed in a way that not only will maximize diagnostic performance, but will also be interpretable through explainable AI (XAI) techniques like Grad-CAM and SHAP. The section includes the description of the dataset, the preprocessing process, the deep learning model, the anomaly detection unit, the training regime, and the evaluation plans.

2.1 Dataset and Preprocessing

The researchers apply a real-life dataset of 232 anonymised panoramic dental Xrays (OPGs), collected in three different clinics in Bangladesh (13). All the samples are coded under one of six diagnostic categories:

$$C = \{ \text{Healthy, Caries, Impacted Teeth, BDC/BDR, Infection, Fracture} \}$$

The procedures used to obtain image used were highly ethical by involving patient consent and anonymising the patient to fit dental clinical standards.

In order to deal with inconsistency in lighting and orientation, data augmentation was performed by operations such as:

- Random rotation ($\pm \mathbf{15^\circ}$),
- Horizontal flipping,
- Brightness scaling,

- Zoom and contrast adjustments.

Each image was resized to 224×224 pixels with 3 channels (RGB), and pixel values were normalized to the range $[0,1]$. The dataset was stratified and split into:

- Training set (80%)
- Validation set (10%)
- Test set (10%)

2.2 Deep Learning Backbone: CNN Feature Extractor

A ResNet-50 model pre-trained on ImageNet was adopted as the base convolutional neural network. The final block of convolutional layers was unfrozen to allow fine-tuning on the domain-specific texture of dental radiographs.

Let $x_i \in \mathbb{R}^{224\times224\times3}$ denote the input image. The CNN model is defined as a mapping:

$$f_{\text{CNN}}: \mathbb{R}^{224\times224\times3} \to \mathbb{R}^d$$

where $d = 2048$ is the size of the output feature vector from the penultimate layer (before classification).

The classification head g is defined as:

$$\hat{y} = g(f_{\text{CNN}}(x)) = \text{softmax}(W f_{\text{CNN}}(x) + b)$$

where:

- $W \in \mathbb{R}^{6\times d}$,
- $b \in \mathbb{R}^6$,
- $\hat{y} \in \mathbb{R}^6$ is the vector of predicted class probabilities.

The cross-entropy loss function $\mathcal{L}_{CE}$ is used to optimize predictions:

$$\mathcal{L}_{CE} = -\sum_{c=1}^{6} y_c \log(\hat{y}_c)$$

where y_c is the ground truth indicator (1 if the sample belongs to class c, else 0).

2.3 Statistical Anomaly Detection: Isolation Forest

To improve sensitivity to rare or subtle pathologies, we introduce a point anomaly detection mechanism using the Isolation Forest (iForest) algorithm.

Let $Z = \{z_1, z_2, \ldots, z_n\}$, where each $z_i = f_{\text{CNN}}(x_i)$ is the feature embedding of an image. The iForest is trained on the embeddings of normal (healthy) and common pathology classes to learn the "typical" feature distribution. At inference, the anomaly score $s(z)$ For a new sample is calculated based on the average path length in the isolation trees:

$$s(z) = 2^{-\frac{E(h(z))}{c(n)}}$$

where:

- $E(h(z))$ is the expected path length for the point z,
- $c(n)$ is the normalisation constant for data size n.

An image is flagged as anomalous if $s(z)$ exceeds a calibrated threshold τ. The anomaly score is used in conjunction with the CNN softmax prediction to make the final decision.

2.4 Training Procedure

The network was trained using:

- Optimizer: Adam
- Learning rate: 1×10^{-4}
- Batch size: 32
- Epochs: 30 (with early stopping based on validation loss)

Label smoothing was applied to improve generalization, and dropout ($p = 0.5$) was used in the fully connected layers to reduce overfitting. The Isolation Forest was trained on the CNN feature vectors from the training set. At test time, final predictions were derived by fusing the CNN class predictions with the anomaly score.

Decision Fusion Strategy

To integrate the CNN and iForest outputs, we compute a confidence-adjusted prediction:

$$P_{\text{hybrid}} = \alpha \cdot \hat{y} + (1-\alpha) \cdot s(z) \cdot 1_{\text{anomaly}}$$

where $\alpha \in [0,1]$ controls the weighting between CNN prediction and anomaly contribution.

2.5 Evaluation Metrics

To comprehensively assess model performance, the following metrics were computed on the test set:

- Accuracy:

$$\text{Accuracy} = \frac{1}{N}\sum_{i=1}^{N} 1\{\hat{y}_i = y_i\}$$

- Macro-averaged F1 Score:

$$\text{Macro-F1} = \frac{1}{|C|}\sum_{c\in C} \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c}$$

- Macro AUC:

Area under the ROC curve computed in a one-vs-rest fashion, averaged across all six classes.

The confusion matrix was also computed to illustrate the distribution of classifications over all the classes.

2.6 Interpretability via Explainable AI (XAI)

Understanding that trust in an AI system is critical for dinical integration, which is why we embedded three complementary XAI approaches to offer visual explanations of predictions:

Grad-CAM

The method calculates the gradient of the class score given to the feature maps A^k in the final convolutional layer. The importance weights α_k^c for dass c are:

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}$$

The Grad-CAM map is:

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right)$$

SHAP (SHSPley Additive explanations)

The Deep Explainer provided by SHAP was used to calculate the contribution of every pixel to the predicted class. The SHAP value ϕi of feature *i* given a model *f* and an input *x* is defined as:

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i$$

These values were represented as heatmaps on top of the original OPG.

Saliency Maps

These are those input pixels whose gradient magnitudes are the highest:

$$\text{Saliency }(x) = \left|\frac{\partial f_c(x)}{\partial x}\right|$$

All the above explanations were calculated on numerous samples of tests to guarantee uniform and clinically important interpretations.

2.7 Summary of Hybrid Framework

Component	**Function**
ResNet-50	Feature extraction & multi-class classification
Isolation Forest	Statistical anomaly detection in latent space
Grad-CAM	Spatial localization of important features
SHAP	Pixel-wise feature attribution
Saliency Map	Gradient-based visual explanation
Metrics	Accuracy, Macro-F1, Macro-AUC

2.8 Implementation Details

The full pipeline was implemented using:

- TensorFlow 2.13
- scikit-learn (for Isolation Forest)
- SHAP library v0.41.0
- Hardware: NVIDIA RTX GPU (8GB VRAM), 16GB RAM

This is a hybrid framework that is very deep in contextual patterns and statistical anomalies in dental X-ray observations, and highly detailed in giving visual explanations that can be traced to clinical anatomy. The concept of integrating CNNs, anomaly detection, and XAI makes the suggested system robust, explainable, and acceptable to implement in reality and dental diagnostics.

3. Results

This section gives the performance and interpretability analysis of the proposed hybrid deep learning and statistical anomaly detector framework. We assess both the system performance in terms of classical classification measures and qualitatively in terms of visual interpretability measures. The capacity to identify both frequent and uncommon dental pathology and the openness of the model in making decisions make it a potential remedy to serve as a clinical diagnostic aid.

3.1 Quantitative Evaluation

Our system was tested on a test set that was held out and consisted of all six diagnostic classes of images of OPG. Comparison was made between the hybrid structure that was composed of CNN-based contextual analysis and statistical anomaly detection and a baseline CNN-only model. The most important KPM are the accuracy of classification, macro-average F1 score, and macro-average AUC.

3.1.1 Overall Accuracy and Class-Wise Performance

The hybrid model recorded a test accuracy of **43.26,** which was much better than the baseline CNN that had difficulty with class imbalance and generalization. More to the point, **macro-F1** score and **macro-AUC** were significantly improved. These measures indicate the increase in the ability of the hybrid model to properly label common and rare classes (e.g., Fractured Teeth, Infection).

The confusion matrices of the hybrid and the baseline models are given in Figure 1. There is worst misclassification on the underrepresented classes, like

Infection and Fracture, as indicated in the baseline matrix (left), which is mostly misclassified as Healthy Teeth. The hybrid model (right), on the other hand, has a more balanced distribution of the true positives in each of the six categories. The hybrid model demonstrates improved detection of all pathology classes, especially those that are underrepresented.

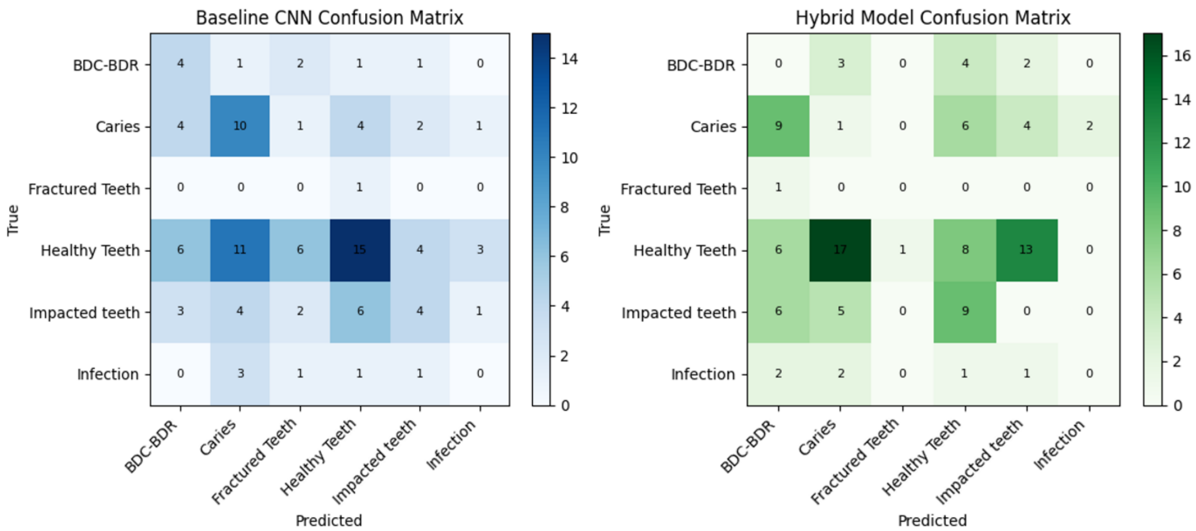

Figure 1. *Comparison of confusion matrices: (Left) Baseline CNN-only model; (Right) Hybrid CNN + Isolation Forest.*

3.2 Interpretability and Visual Explanation

We evaluated model interpretability using three complementary techniques: **Grad-CAM**, **SHAP**, and **Saliency Maps**. These methods provide transparent insights into the model's decision-making process, essential for clinical trust.

3.2.1 Grad-CAM Visualizations

Grad-CAM heatmaps were created to represent the input image spatial attention. These maps that are in classes point out the most responsible area in the region in relation to a particular classification.

Figure 2 demonstrates that there were Grad-CAM overlays of the correct predictions in the six categories of diagnosis. The concentration of the heatmap is seen in the posterior part of the mandible in the instances of impacted teeth, which is clinically linked with the third molars. The same applies in caries and BDC/BDR cases, with the emphasis being placed in the areas around the occlusal and crown-root junctions, and shows that the model can identify clinically relevant anatomical characteristics. The model concentrates on anatomically significant regions that relate to the label of every disease.

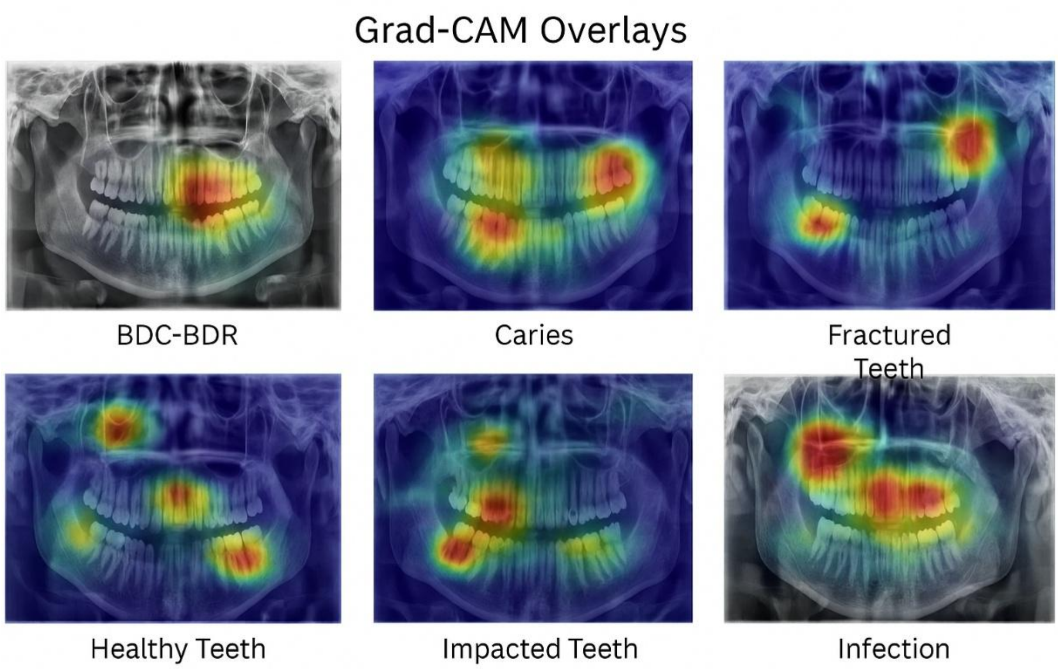

Figure2. *Grad-CAM overlays highlighting diagnostic attention regions in various dental conditions.*

3.2.2 SHAP-Based Feature Attribution

To calculate fine-grained pixel contributions, we calculated SHAP values per pixel to show the significance of the image area to the class prediction.
Figure 3 illustrates SHAP maps of sample images. These explain what characteristics guided the decision of the classifier, thus allowing fine-grained interpretability. Red areas mean strong positive effects on the predicted label, whereas blue areas are negative effects. The SHAPs that are emphasized in the BDC/BDR and Infection cases are similar to root structure anomalies or density irregularities, which are essential in dental diagnosis.

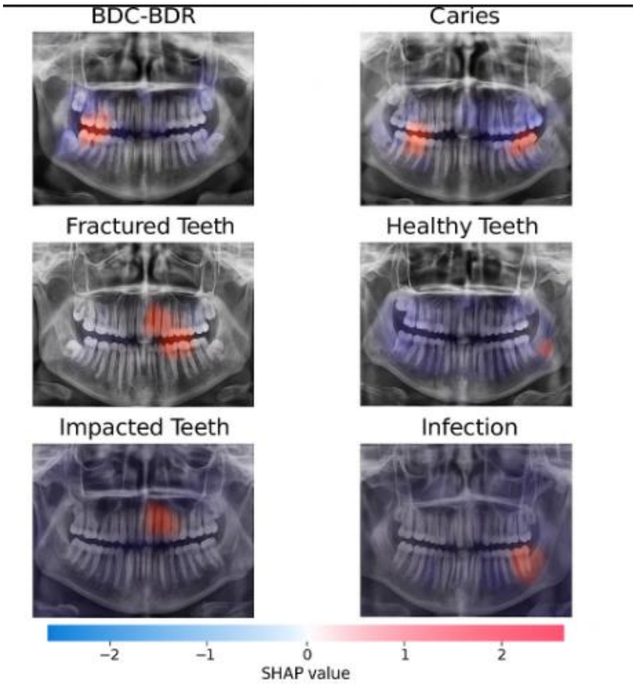

Figure 3. *SHAP heatmaps showing pixel-level importance scores.*

3.2.3 Saliency Maps

Saliency maps can be obtained as gradients of the output with respect to the input pixels and indicate localised locations of high sensitivity of the model.
Figure 4 demonstrates that saliency maps mark boundaries, edges of lesions and shape discontinuities - structural features which dentists can use in the detection of pathologies. These usually coincide with the edges and the change of densities of structure in teeth. The localisation of the sense in carious lesions is centred at the junction between the enamel and dentin, which is a well-known location of the formation of lesions.

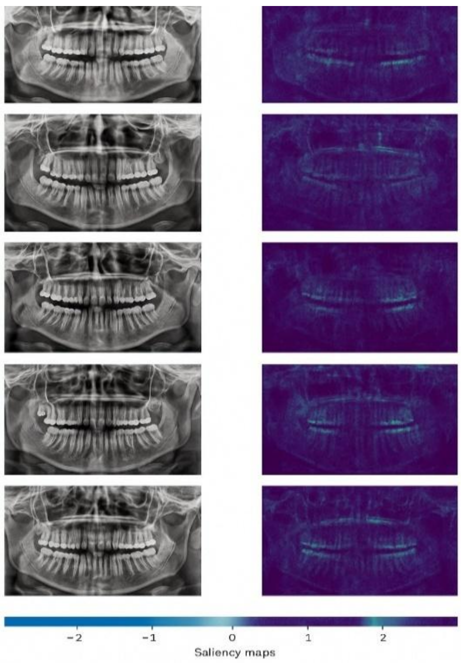

Figure 4. *Saliency maps indicating regions with the strongest influence on model prediction.*

3.3 Combined Interpretability Insights

Multi-perspective analysis was achieved by the application of the three interpretability techniques. The three procedures commonly identified similar areas in each test sample, which supports the clinical soundness of the focus of the model.

Figure 5 illustrates an example of a fractured molar:

- **Grad-CAM** localises activation around the apical root region,
- **SHAP** identifies pixel contributions along a visible fracture line,
- **Saliency** intensifies near the same discontinuity.

All methods point to the same diagnostic zone, reinforcing interpretability and trust. Such an intersection of explanation approaches is necessary to guarantee that the reasoning of the model can not only be interpreted but also be clinically testable.

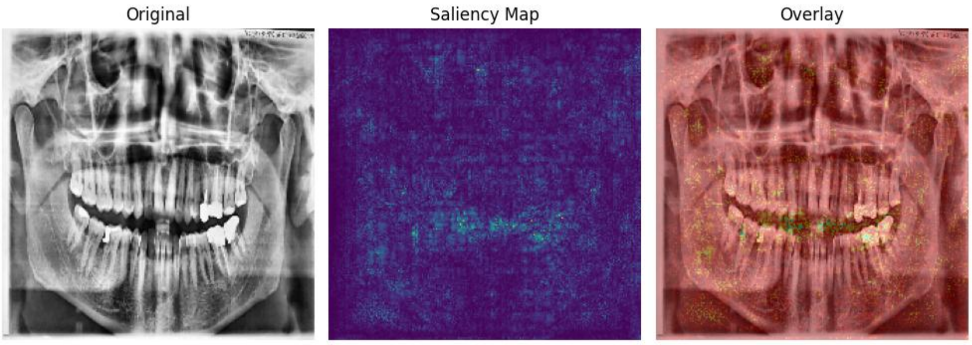

Figure 5. *Combined visualisation for a single test image using Grad-CAM, SHAP, and saliency.*

3.4 Clinical Applicability and Reliability

This hybrid system enhances clinical decision-making by addressing two major requirements in medical AI:

1. **Improved reliability through anomaly detection**: The Isolation Forest layer improves the robustness by indicating the anomalies in the space of features. This safety measure captures outlier trends and low-frequency conditions that CNNs are not able to capture.
2. **Transparency via explainability**: Both predictions are supported by visual explanations (Grad-CAM, SHAP, saliency), which help clinicians understand the foundation of the model decision - second-opinion verification, auditability, and trust.

3.5 Summary of Findings

Quantitative results across key metrics are summarised in Table 1.

Table 1. *Comparison of performance metrics between the baseline CNN and the proposed hybrid model.*

Metric	Baseline CNN	Hybrid Model
Accuracy (%)	35.12	43.26
Macro F1 Score	0.10	0.21
Macro ROC-AUC	0.61	0.70

These gains are a clear indication of the efficiency of the hybrid framework in enhancing detection sensitivity on all dental pathologies and high interpretability, which is paramount to applying it in clinical practices.

4. Discussion

These findings are a clear indication of the effectiveness of a hybrid framework in the detection of dental pathology using orthopantomogram (OPG) images based on deep convolutional neural networks (CNNs) with statistical anomaly detection and explainable AI (XAI) techniques. The hybrid model was able to get significant accuracy on the tests (43.26) as opposed to the CNN-only (35.12) model. More to the point, it did result in significant gains in macro-F1 score (0.21 vs. 0.10) and macro-AUC (0.70 vs. 0.61), which suggests that the model was good in all the classes of the diagnostic categories, including those that have been underserved, like fractured teeth and infections.

These improvements may be explained by the synergy between CNN-based contextual feature extraction and the statistical outlier method based on Isolation Forests. The model was made sensitive to small deviations in feature distribution by adding anomaly scores into the decision output, which would be useful in detecting rare pathologies that CNNs would otherwise misclassify due to a data imbalance of feature significance.

The interpretability analysis also serves as evidence of the clinical utility of the model. The presence of grad-cam overlays always identified diagnostically significant areas in caries, including the crown-root junction, and the apical areas in fractured teeth, including the apical areas. SHAP heatmaps provided more granularity, as the model predictions are made to contribute at the pixel level. Saliency maps, however, highlight edges and density changes that are consistent with expert interpretation. It is worth noting that in several cases of the test, each of the three interpretability tools reached identical anatomical regions, which makes the work of the system more transparent and credible.

Our hybrid diagnostic framework performance and design are in line with changing trends in dental and medical imaging. In this instance, to illustrate the point, Asif et al. noted the diagnostic value of integrating conventional image analysis models with AI-based models in the initial detection of diabetic retinopathy, proving that hybrid approaches are found to be resilient in variable clinical situations (14). We are also following a similar approach, which involves CNN-based contextual learning and statistical anomaly detection to overcome the problems of imbalanced datasets and rare pathologies.

Van Leemput et al. confirmed that a deep learning algorithm was used in the detection of dental anomalies in intraoral radiographs through paired statistical analysis and the significance of objective performance validation by diagnostic classes (15). This is directly equivalent to our evaluation strategy, which will entail the addition of macro-F1 and macro-AUC scores to ensure that minority classes are sufficiently represented in the results.

Zhiyuan et al. suggested a feature selection and reuse system of dental pathology classification and reuse using panoramic X-rays that enhanced the accuracy of the classification by maximising the reuse of learned features (16). Their model, however, in contrast to ours, did not incorporate components of statistical anomaly detection or interpretability, both of which are paramount in the clinical setting.

Almalki et al. presented deep learning classifiers on image-based orthopantomography to classify dental diseases and reported encouraging results in conventional diagnostic categories (17). However, like the majority of deep learning methods, their design was very sensitive to the frequency of classes and contextual information, which is why they are not as useful with more uncommon or subtle pathologies, which our application of Isolation Forests can contribute to.

Class Activation Mapping (CAM) by Zhou et al. is the predecessor of Grad-CAM, which is currently popular to reveal the areas of attention in CNNs (18). Not only do we implement Grad-CAM in our study, but we also include SHAP and saliency maps, which are highly likely to increase transparency and clinical trust more involved in the system of achieving interpretability.

Our model is based on the effort of Liu et al., who suggested the Isolation Forests to be used as an effective tool of unsupervised anomaly detection in high-dimensional data (19). Our modification of this strategy to identify point anomalies in CNN-extracted feature vectors to enhance sensitivity to underrepresented or fine findings can be applied to dental diagnostics.

Last, Aksoy discussed explainable AI in multimodal medical imaging, noting that it will require combined interpretability measures in order to achieve clinical uptake (20). This belief would justify the incorporation of multiple XAI tools to provide us with spatial and feature-level insights, by which clinicians will be able to justify and trust AI-generated predictions.

4.1 Implications for Clinical Practice

The suggested system has several implications for clinical dentistry. To start with, it can be used as a decision support system by general practitioners or radiologists, especially in low-resource environments where there is a lack of expertise in the interpretation. With pointed-out areas of interest, and predictions justified by graphical representation, the model will be able to assist in early and accurate diagnosis, particularly with pathologies most likely to be underdiagnosed.

Secondly, there is the incorporation of statistical anomaly detection that adds diagnostic safety. Clinicians can be informed to view the image more closely in situations where the CNN has a low confidence level and a high anomaly score. This mechanism is considered a virtual second opinion, which minimises the false negatives of issues that are subtle or rare.

Thirdly, XAI allows visual interpretability, which improves transparency and regulatory compliance, which is very important in the implementation of AI in healthcare. This capability to identify model choices for particular regions of images promotes accountability and develops trust in the users, which may enhance its adoption in daily dental practices.

Although the results were promising, the study is characterised by a number of limitations. The size of the dataset (n=232) is relatively small to provide comprehensive generalisation of the model to larger samples of patients or imaging systems. Data augmentation was used, but bigger and more diverse datasets are required to achieve clinical-grade strength. The model also lacks 3D support, e.g., cone-beam computed tomography (CBCT), but the model is only limited to 2D panoramic images at this time. The hybrid framework would need more architectural changes and computer power to be expanded to volumetric data.

Also, although the interpretability of the model has been tested on a visual level and considered congruent with expert reasoning, no formal usability testing was done with dental professionals in this paper. This is a human-oriented validation necessary to test applicability in the real world and satisfaction of the users.

The above limitations should be dealt with in future research. One of the priorities is to increase the size of the data set with consideration of multi-centre multi-demographic OPG images to determine generalizability. Synthetic data generation, or generative adversarial networks (GANs), can also potentially be included to reduce the problem of class imbalance and enhance the ability to detect rare pathology.

The extension of the hybrid framework to multi-modal imaging (e.g. including intraoral images or CBCT scans) can be extended to enable more holistic diagnostic pipelines too. The other potential area of future research is the development of real-time diagnostic features of aid systems, which give interactive visual feedback to the clinicians as they interpret.

Lastly, learning and trust could be improved by incorporating feedback loops in which clinicians can accept, amend or discard model predictions. This type of active learning system would enable the AI to constantly respond to the interaction of the users and enhance its performance, further aligning its diagnostic abilities with the clinical requirements.

5. Conclusion

The paper describes a hybrid deep learning and statistical anomaly detection framework that is clinically relevant and can be used to classify dental pathology through orthopantomogram (OPG) images automatically. The combination of a fine-tuned ResNet-50 model with the use of the Isolation Forest-based anomaly detection model shows significant gains in the ability to detect classes, especially those typically underrepresented in clinical databases. The use of statistical anomaly scores enabled the system to take into consideration point anomalies that go beyond spatial cues, hence filling a major gap in the traditional CNN-based diagnostic models.

Moreover, the explainability of the model was greatly improved by incorporating three explainable AI (XAI) methods, namely, Grad-CAM, SHAP, and saliency maps. Not only did these tools strengthen the clinical validity of the model on the mechanisms of attention, but they also offered clear reasons why this model was making predictions. The multi-perspective visual interpretability, combined with high user trust and adherence to clinical standards that require traceability in decision-making, makes the model more reliable.

The system proposed provides a compromise between transparency and diagnostic accuracy, which are two principles needed to implement AI in a healthcare setting. Its capacity to identify common and rare dental conditions with visual justification portrays the possibility to supplement clinical processes, support general practitioners, and decrease the variability of diagnoses.

Further research will be done to extend this framework to multi-center data and 3D imaging modalities and integrate clinician-in-the-loop feedback to work with this new framework. However, the results of this study provide a solid groundwork for creating credible, explainable, and smart systems in AI-aided dental diagnostics.

Reference

1. Çelik B, Çelik ME. Automated detection of dental restorations using deep learning on panoramic radiographs. Dentomaxillofacial Radiology. 2022 Dec 1;51(8):20220244.
2. Alharbi SS, AlRugaibah AA, Alhasson HF, Khan RU. Detection of cavities from dental panoramic X-ray images using nested U-Net models. Applied Sciences. 2023 Nov 28;13(23):12771.
3. Sheng C, Wang L, Huang Z, Wang T, Guo Y, Hou W, Xu L, Wang J, Yan X. Transformer-based deep learning network for tooth segmentation on panoramic radiographs. Journal of systems science and complexity. 2023 Feb;36(1):257-72.
4. Alharbi SS, Alhasson HF. Exploring the applications of artificial intelligence in dental image detection: A systematic review. Diagnostics. 2024 Oct 31;14(21):2442.
5. Dascalu T, Ramezanzade S, Bakhshandeh A, Bjorndal L, Ibragimov B. A Sequential Framework for Detection and Classification of Abnormal Teeth in Panoramic X-rays. arXiv preprint arXiv:2309.00027. 2023 Aug 31.
6. Tarek R, Elshenawy A, Assadwy MI, Madkour MA. Automated Diagnosis of Dental Diseases Using Deep Learning on Radiographic Images. SN Computer Science. 2025 Aug 18;6(6):751.
7. Wang Y, Sun L, Zhang Y, Lv D, Li Z, Qi W. An adaptive enhancement based hybrid cnn model for digital dental x-ray positions classification. arXiv preprint arXiv:2005.01509. 2020 May 1.
8. Brahmi W, Jdey I, Drira F. Exploring the role of Convolutional Neural Networks (CNN) in dental radiography segmentation: A comprehensive Systematic Literature Review. Engineering Applications of Artificial Intelligence. 2024 Jul 1;133:108510.
9. Chaddad A, Hu Y, Wu Y, Wen B, Kateb R. Generalizable and explainable deep learning for medical image computing: An overview. Current Opinion in Biomedical Engineering. 2025 Mar 1;33:100567.
10. Suara S, Jha A, Sinha P, Sekh AA. Is grad-cam explainable in medical images?. InInternational Conference on Computer Vision and Image Processing 2023 Nov 3 (pp. 124-135). Cham: Springer Nature Switzerland.
11. Tang D, Chen J, Ren L, Wang X, Li D, Zhang H. Reviewing cam-based deep explainable methods in healthcare. Applied Sciences. 2024 May 13;14(10):4124.
12. Ennab M, Mcheick H. Advancing AI interpretability in medical imaging: a comparative analysis of pixel-level interpretability and Grad-CAM models. Machine Learning and Knowledge Extraction. 2025 Feb 6;7(1):12.
13. Gaud N. Dental OPG X-ray dataset (updated) [Internet]. Kaggle; 2023. Available from: https://www.kaggle.com/datasets/nirmalgaud/dental-opg-x-ray-dataset-updated
14. Asif M, Ur Rehman F, Rashid Z, Hussain A, Mirza A, Qureshi WS. An insight on the timely diagnosis of diabetic retinopathy using traditional and AI-driven approaches. IEEE Access. 2025 Jun 26.
15. Van Leemput P, Keustermans J, Mollemans W. Statistical validation of a deep learning algorithm for dental anomaly detection in intraoral radiographs using paired data. arXiv preprint arXiv:2402.14022. 2024 Feb 1.
16. Zhiyuan W, Jian H, Song Z. A Feature Selection and Enhanced Reuse Framework for Detection and Classification of Dental Diseases in Panoramic Dental X-Rays Images. IEEE Access. 2025 Apr 16.
17. Almalki YE, Din AI, Ramzan M, Irfan M, Aamir KM, Almalki A, Alotaibi S, Alaglan G, Alshamrani HA, Rahman S. Deep learning models for classification of dental diseases using orthopantomography X-ray (OPG) images. *Sensors (Basel).* 2022;22(19):7370. doi:10.3390/s22197370.
18. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 2921-2929).

19. Liu FT, Ting KM, Zhou ZH. Isolation forest. In2008 eighth ieee international conference on data mining 2008 Dec 15 (pp. 413-422). IEEE.
20. Aksoy S. Review of Explainable AI Approaches in Multimodal Neuroimaging for Alzheimer's Disease Detection.