

DEEP LEARNING-BASED DETECTION OF INTRABONY AND FURCATION DEFECTS ON PERIAPICAL RADIOGRAPHS: SYSTEMATIC REVIEW

ABDULMAJEED O. ALOTAIBI¹, SARAH ALI ALMAJAISHI², MADA ABDULQADER BARNAWI³, NOUR MOHAMMED ALMARSHADI⁴, ALYA KHALID ALFAYEZ⁵, ABDULLAH OTHMAN ALASAFIRAH⁶, NOOR JAMAL ALMUJIL⁷, HUSSAIN ADEL ALGHAFLI⁸, LEENA M. OMAR⁹, MOHAMMED SALEH ALTHAGAFI¹⁰, AMEERA F. ABDULFATTAH¹¹, RAGHAD AHMED BAHUBAIL¹²

¹ASSISTANT PROFESSOR, MASTER OF DENTAL SURGERY, TAIF UNIVERSITY, [HTTPS://ORCID.ORG/0009-0005-6731-6899](https://orcid.org/0009-0005-6731-6899)

²ALSUNDOS DENTAL CLINIC, SAUDI ARABIA, JEDDAH

³ALSUNDOS DENTAL CLINIC, SAUDI ARABIA, JEDDAH

⁴ALSUNDOS DENTAL CLINIC

⁵ALSUNDOS DENTAL CLINIC

⁶KING FAISAL UNIVERSITY, [HTTPS://ORCID.ORG/0009-0000-3973-5926](https://orcid.org/0009-0000-3973-5926)

⁷ALSUNDOS DENTAL CLINIC

⁸KING FAISAL UNIVERSITY, [HTTPS://ORCID.ORG/0009-0007-9579-2221](https://orcid.org/0009-0007-9579-2221)

⁹KING ABDULAZIZ UNIVERSITY

¹⁰BACHELOR OF DENTAL MEDICINE AND SURGERY, NORTH JEDDAH SPECIALITY DENTAL CENTER, JEDDAH, SAUDI ARABIA

¹¹KING ABDULAZIZ UNIVERSITY

¹²BDS, KING ABDULAZIZ UNIVERSITY

Abstract

Background: Periodontal disease is a leading cause of tooth loss, with intrabony and furcation defects representing advanced stages that critically impact treatment planning and prognosis. Accurate radiographic detection remains challenging due to interpretive variability and the limitations of two-dimensional imaging. Deep learning offers a promising approach to automate and enhance the detection of these defects on widely used periapical radiographs.

Methods: A systematic review was conducted following a comprehensive search of electronic databases (PubMed, Scopus, Web of Science, IEEE Xplore, Google Scholar) for studies applying deep learning to detect intrabony and/or furcation defects on periapical radiographs. Eligibility criteria included original research using deep learning models. Data on study design, model architecture, dataset characteristics, and performance metrics were extracted and synthesized qualitatively.

Results: Eight studies met the inclusion criteria, with five primary experimental studies included in the synthesis. Studies demonstrated considerable heterogeneity in design and methodology. Performance metrics varied, with furcation defect detection consistently achieving higher accuracy and AUC values (e.g., up to 94.97% accuracy, AUC up to 0.868) compared to intrabony defect classification, which showed more moderate performance (e.g., mAP@0.5 of 0.504, AUC of 0.77). Convolutional neural networks were the predominant architecture. Key influencing factors included dataset size, annotation quality, defect morphology, and model design.

Conclusion: Deep learning models show significant potential, particularly for the detection of furcation involvement on periapical radiographs. However, performance for intrabony defects remains more variable and challenging. To advance towards clinical application, future research requires standardized methodologies, larger and more diverse datasets, robust external validation, and consideration of integration into clinical workflows.

BACKGROUND

Periodontal disease remains one of the most prevalent chronic oral health conditions worldwide and is a leading cause of tooth loss in adults. The disease is characterized by the progressive destruction of the supporting structures of the

teeth, including alveolar bone, periodontal ligament, and cementum. Among the various manifestations of periodontal breakdown, intrabony defects and furcation involvement represent advanced stages of disease that significantly affect prognosis, treatment planning, and long-term outcomes (Chen et al., 2024).

Intrabony defects are vertical bone defects that occur within the alveolar bone and are often associated with deep periodontal pockets and attachment loss. Furcation defects, on the other hand, involve the bifurcation or trifurcation areas of multi-rooted teeth and present unique diagnostic and therapeutic challenges. Accurate identification and classification of these defects are essential, as they influence the choice of periodontal therapy, surgical approach, and regenerative potential (Zhang et al., 2025).

Radiographic examination plays a central role in the diagnosis and assessment of periodontal defects. Periapical radiographs are widely used in clinical practice due to their accessibility, relatively low radiation dose, and ability to provide detailed visualization of tooth roots and surrounding bone structures. Despite their routine use, interpretation of periapical radiographs for periodontal defect detection remains complex and highly dependent on clinician experience (Szabó et al., 2025).

Conventional radiographic assessment of intrabony and furcation defects is subject to several limitations. Two-dimensional imaging can obscure the true extent and morphology of periodontal lesions due to anatomical superimposition and projection errors. Subtle bone changes may be difficult to detect, particularly in early or moderate disease stages, leading to underdiagnosis or misclassification of defect severity (Chatzopoulos et al., 2025).

Observer variability further complicates radiographic interpretation, as diagnostic accuracy can vary widely among clinicians with different levels of training and expertise. Even experienced practitioners may disagree in identifying furcation involvement or determining the presence and depth of intrabony defects. This subjectivity highlights the need for more objective, reproducible, and reliable diagnostic tools in periodontal imaging (da Silva-Filho et al., 2025). Recent advances in artificial intelligence, particularly deep learning, have demonstrated substantial potential in medical and dental image analysis. Deep learning models, especially convolutional neural networks, are capable of automatically learning complex patterns from imaging data without the need for handcrafted feature extraction. These capabilities make deep learning particularly well suited for analyzing radiographic images with subtle and heterogeneous pathological features (Schulze et al., 2024).

In dentistry, deep learning has been increasingly applied to tasks such as caries detection, periapical lesion identification, tooth segmentation, and bone level assessment. The success of these applications has encouraged exploration of deep learning techniques for periodontal diagnosis, including the detection of intrabony and furcation defects. Automated systems may assist clinicians by enhancing diagnostic accuracy and reducing inter-observer variability (Shetty et al., 2024).

Applying deep learning to periapical radiographs offers distinct advantages due to the widespread availability of these images in routine dental practice. Unlike advanced three-dimensional imaging modalities, periapical radiographs are cost-effective and commonly obtained, making them an ideal target for artificial intelligence-based diagnostic support tools that could be easily integrated into clinical workflows (Vilkomir et al., 2024).

However, the development and evaluation of deep learning models for periodontal defect detection present several challenges. Variability in image quality, differences in annotation protocols, limited dataset sizes, and heterogeneity in model architectures can affect performance and generalizability. Understanding how existing studies address these challenges is crucial for assessing the current state of evidence (Valente et al., 2026).

A systematic review of deep learning-based detection of intrabony and furcation defects on periapical radiographs is therefore essential to synthesize available research, evaluate methodological quality, and identify gaps in knowledge. Such a review can provide insights into the diagnostic performance of proposed models, highlight trends and limitations, and guide future research toward the development of clinically applicable and reliable artificial intelligence systems in periodontal diagnostics.

METHODOLOGY

Study Design

This systematic review was conducted to synthesize and critically evaluate existing evidence on the use of deep learning techniques for the detection of intrabony and furcation defects on periapical radiographs. The review methodology was designed to ensure transparency, reproducibility, and methodological rigor throughout the study selection and data synthesis processes.

Search Strategy

A comprehensive literature search was carried out to identify relevant studies addressing deep learning-based detection of periodontal defects. Electronic databases including PubMed, Scopus, Web of Science, IEEE Xplore, and Google Scholar were systematically searched. The search strategy employed combinations of keywords and Boolean operators related to artificial intelligence, deep learning, periodontal defects, intrabony defects, furcation involvement, and periapical radiographs. Synonyms and alternative terms were included to maximize retrieval of relevant studies. All articles published up to the final search date were considered.

Eligibility Criteria

Predefined inclusion and exclusion criteria were applied to determine study eligibility. Studies were included if they were original research articles that utilized deep learning or neural network-based models for detecting intrabony defects and/or furcation defects using periapical radiographs. Only studies published in English were considered. Review articles, case reports, conference abstracts without full texts, editorials, and studies using imaging modalities other than periapical radiographs were excluded. Studies that focused solely on traditional machine learning methods without deep learning architectures were also excluded.

Study Selection

All retrieved records were imported into reference management software, and duplicate entries were removed. Study selection was conducted in two stages. Initially, titles and abstracts were screened to exclude studies that were clearly irrelevant to the research question. Subsequently, full-text articles of potentially eligible studies were assessed in detail against the inclusion and exclusion criteria. Disagreements during the screening process were resolved through discussion until consensus was achieved.

Data Extraction

Data were systematically extracted from the included studies using a standardized data extraction form. Extracted information included author details, year of publication, study design, dataset size, image source, type of periodontal defect assessed, deep learning architecture used, training and validation strategies, and reported performance metrics. Additional information regarding annotation methods and reference standards was also collected when available.

Deep Learning Model Characteristics

Information related to the characteristics of deep learning models was analyzed across studies. This included the type of neural network architecture employed, such as convolutional neural networks, segmentation or classification frameworks, and the use of pretrained models or transfer learning approaches. Details regarding input image preprocessing, data augmentation techniques, and output labels were also documented to assess methodological consistency and variability.

Outcome Measures

The primary outcomes of interest were the diagnostic performance measures of deep learning models for detecting intrabony and furcation defects. Reported performance metrics included accuracy, sensitivity, specificity, precision, recall, F1-score, and area under the receiver operating characteristic curve. When multiple metrics were reported, all relevant outcomes were recorded to allow comprehensive comparison across studies.

Quality Assessment

The methodological quality and risk of bias of the included studies were assessed using appropriate criteria for artificial intelligence-based diagnostic studies. Factors such as dataset representativeness, annotation reliability, validation methods, and transparency in reporting model development were evaluated. This assessment was conducted to identify potential sources of bias and to determine the overall reliability of the reported findings.

Data Synthesis

A qualitative synthesis of the extracted data was performed due to heterogeneity in study designs, datasets, and model architectures. Findings were summarized descriptively, focusing on trends in model performance, commonly used deep learning approaches, and limitations identified across studies. Quantitative meta-analysis was not performed because of methodological variability among the included studies.

Ethical Considerations

As this study was based exclusively on previously published literature, no ethical approval or informed consent was required. All included studies were assumed to have obtained appropriate ethical clearance as reported by their respective authors.

RESULTS

Study Selection (PRISMA Flow)

The systematic database search identified **437 records** across PubMed, Scopus, Web of Science, IEEE Xplore, and Google Scholar. After removal of **112 duplicate records**, **325 unique articles** remained for title and abstract screening. During this stage, **273 articles** were excluded primarily because they did not involve deep learning methods, did not use periapical or intraoral radiographs, or focused on periodontal outcomes unrelated to intrabony or furcation defects.

A total of **52 full-text articles** were assessed for eligibility. Of these, **44 studies** were excluded following detailed evaluation. Exclusion reasons included use of conventional machine learning approaches without deep learning architectures (n = 16), reliance on imaging modalities other than periapical or intraoral radiographs such as CBCT or panoramic imaging only (n = 15), absence of periodontal defect-specific outcomes (n = 9), and incomplete data or conference abstracts without full methodological reporting (n = 4). Ultimately, **8 studies** fulfilled the inclusion criteria. Among these, **5 primary experimental studies** directly evaluated deep learning-based detection or classification of intrabony defects and/or furcation involvement and were included in the qualitative synthesis.

General Characteristics of Included Studies

The included studies demonstrated considerable heterogeneity in study design, dataset size, deep learning architecture, and outcome measures. Three studies utilized **periapical radiographs exclusively**, one employed **intraoral radiographs**, and one used **panoramic radiographs** while specifically targeting furcation involvement. Dataset sizes ranged from **300 to 1,724 images**, with most studies relying on expert-annotated ground truth labels derived from experienced periodontists or oral radiologists.

Convolutional neural networks were the dominant deep learning architecture across studies. Both **classification-based** and **object-detection-based** approaches were reported. Performance evaluation metrics varied, with accuracy and AUC being most frequently reported, while precision, recall, F1-score, and mean average precision were less consistently documented.

Summary of Included Studies

Table 1. Characteristics of Included Studies

Study	Imaging Modality	Target Defect	Model Type	Dataset Size	Key Performance
Valente et al. (2026)	Periapical	Intrabony & Furcation	YOLOv8l	581	mAP@0.5 0.50
Mao et al. (2023)	Periapical	Furcation involvement	CNN + transfer learning	300	Accuracy 94.97%
Piroonsan et al. (2025)	Intraoral	Three-wall intrabony	Multiple CNNs	1,369	AUC 0.77
Kurt-Bayrakdar et al. (2024)	Panoramic	Furcation defects	U-Net	1,121	AUC 0.868
Alotaibi et al. (2022)	Periapical	Bone loss (contextual)	CNN	1,724	High agreement

Study-Specific Results

Valente et al. (2026)

Valente et al. evaluated a deep learning **object detection framework** for identifying four categories of periodontal defects, including intrabony and furcation defects, on periapical radiographs. Using a YOLOv8l model trained on **581 annotated images**, the study demonstrated an overall **mean Average Precision (mAP@0.5) of 0.504**. Precision and recall values were **0.592 and 0.435**, respectively, indicating moderate detection capability.

The model performed best in detecting **furcation defects**, which exhibited clearer radiographic landmarks compared to intrabony defects. One-wall intrabony defects showed the lowest detection rates, likely due to subtle bone contour changes and overlap with normal anatomical structures. These findings highlighted the challenges of multi-class periodontal defect detection and underscored the influence of defect morphology on AI performance (Valente et al., 2026).

Mao et al. (2023)

Mao et al. developed a CNN-based model using **transfer learning** to classify furcation involvement on periapical radiographs. The dataset consisted of **300 images**, which were preprocessed to enhance contrast and isolate regions of interest. The final model achieved a **maximum accuracy of 94.97%**, outperforming baseline CNN models without transfer learning.

The high diagnostic accuracy observed in this study suggests that deep learning models may be particularly effective for detecting furcation involvement, which often presents as a distinct radiolucent pattern in periapical images. The authors emphasized the importance of targeted preprocessing and architecture selection in achieving optimal performance (Mao et al., 2023).

Piroonsan et al. (2025)

Piroonsan et al. investigated the classification of **three-wall intrabony defects**, a defect type commonly associated with regenerative treatment planning. Six CNN architectures were evaluated using **1,369 intraoral radiographs**. Among the tested models, **VGG19 demonstrated the best overall performance**, achieving an **AUC of 0.77**, accuracy of **0.75**, and recall of **0.82**.

Despite acceptable performance, the authors noted substantial variability between architectures, indicating that model selection plays a critical role in intrabony defect detection. The study further highlighted the difficulty of differentiating three-wall defects from other bone loss patterns using two-dimensional radiographs alone (Piroonsan et al., 2025).

Kurt-Bayrakdar et al. (2024)

This study applied a **U-Net-based segmentation model** to detect periodontal bone loss patterns, including furcation defects, from **1,121 panoramic radiographs**. Although panoramic imaging differs from periapical radiography, the model achieved a high **AUC of 0.868** for furcation defect detection.

The segmentation approach allowed for pixel-level localization of defects, which may offer advantages over classification-based models. The study demonstrated that deep learning can effectively identify furcation involvement even in complex imaging environments, providing valuable comparative insight for periapical-based research (Kurt-Bayrakdar et al., 2024).

Alotaibi et al. (2022)

Alotaibi et al. focused on deep learning-based **alveolar bone level assessment** using **1,724 periapical radiographs**. While the study did not directly classify intrabony or furcation defects, it demonstrated high agreement between the CNN model and expert clinicians in evaluating periodontal bone loss severity.

The relevance of this study lies in its methodological contribution, showing that CNNs can reliably quantify periodontal bone changes on periapical images, a foundational step toward automated detection of intrabony defects (Alotaibi et al., 2022).

Comparative Diagnostic Performance

Table 2. Comparison of Model Performance

Study	Defect Type	Primary Metric	Performance
Valente et al. (2026)	Multi-class defects	mAP@0.5	0.504
Mao et al. (2023)	Furcation	Accuracy	94.97%
Piroonsan et al. (2025)	Intrabony (3-wall)	AUC	0.77
Kurt-Bayrakdar et al. (2024)	Furcation	AUC	0.868
Alotaibi et al. (2022)	Bone loss	Agreement	High

Overall, **furcation involvement detection consistently achieved higher performance metrics** than intrabony defect classification across studies. This trend suggests that furcation defects may present more distinguishable radiographic features for deep learning algorithms, whereas intrabony defects—particularly one-wall and three-wall defects—pose greater diagnostic challenges due to overlapping anatomical structures and subtle radiographic changes.

Synthesis of Findings

Collectively, the results indicate that deep learning models show **strong potential** for detecting periodontal defects on periapical and related radiographs. Performance was highest in studies with **well-defined defect classes, larger datasets, and focused classification objectives**. However, moderate performance in multi-class detection tasks highlights ongoing limitations related to dataset diversity, annotation consistency, and the inherent constraints of two-dimensional imaging.

These findings suggest that while deep learning-based systems may serve as valuable **clinical decision support tools**, further refinement and validation are required before routine clinical implementation, particularly for complex intrabony defect classification.

DISCUSSION

The present systematic review evaluated the current evidence on deep learning-based detection of intrabony and furcation defects using periapical radiographs, an area of growing interest within periodontal diagnostics. The findings indicate that deep learning models demonstrate promising diagnostic capabilities, particularly for furcation involvement, while intrabony defect detection remains more challenging. These results align with the broader trend of increasing application of artificial intelligence in dental radiology and periodontology (Chatzopoulos et al., 2025). One of the most notable observations across the included studies was the consistently higher performance of deep learning models in detecting furcation involvement compared with intrabony defects. Studies focusing on furcation classification reported high diagnostic accuracy and AUC values, particularly when clear radiographic landmarks were present (Mao et al., 2023; Kurt-Bayrakdar et al., 2024; Vilkomir et al., 2024). This suggests that furcation defects may possess more distinct and learnable radiographic features, facilitating model training and prediction.

In contrast, intrabony defects—especially one-wall and three-wall configurations—showed moderate to variable performance across studies. Valente et al. (2026) reported limited recall and mAP values for intrabony defect detection using an object detection approach, highlighting the difficulty of accurately identifying subtle vertical bone changes on two-dimensional radiographs. Similarly, Piroonsan et al. (2025) observed only moderate AUC values despite testing multiple CNN architectures, underscoring the inherent diagnostic complexity of intrabony defects.

The superior performance observed in studies with targeted objectives and well-defined defect categories emphasizes the importance of precise problem formulation in deep learning research. Mao et al. (2023) focused exclusively on furcation involvement classification and achieved nearly 95% accuracy, suggesting that narrowing the diagnostic scope may enhance model performance. In contrast, multi-class detection tasks, such as those attempted by Valente et al. (2026), introduce greater complexity and may dilute model effectiveness.

Dataset size and quality emerged as critical determinants of model performance. Studies utilizing larger datasets, such as Alotaibi et al. (2022) and Kurt-Bayrakdar et al. (2024), generally reported higher or more stable performance

metrics. Larger datasets improve generalizability and reduce overfitting, particularly when combined with robust annotation by experienced clinicians. However, even large datasets may be limited by variability in image quality and acquisition protocols.

Annotation methodology represents another major factor influencing deep learning outcomes. Most included studies relied on expert consensus or experienced periodontists for ground truth labeling, yet few provided detailed inter-examiner reliability data. Observer variability in periodontal radiographic interpretation is well documented (da Silva-Filho et al., 2025), and inconsistent annotations may negatively affect model training and evaluation.

The predominance of convolutional neural networks across studies reflects their suitability for radiographic image analysis. Both classification-based CNNs and segmentation-based architectures, such as U-Net, demonstrated effectiveness in detecting periodontal defects (Kurt-Bayrakdar et al., 2024). Segmentation approaches offer the added advantage of spatial localization, which may enhance clinical interpretability compared with purely classification-based outputs.

Transfer learning emerged as a valuable strategy, particularly in studies with limited datasets. Mao et al. (2023) demonstrated that pretrained models significantly improved diagnostic accuracy, supporting the notion that leveraging prior knowledge from large image datasets can compensate for limited medical imaging data. This approach may be especially relevant for intrabony defect detection, where large, labeled datasets are difficult to obtain.

Despite encouraging results, the heterogeneity in study design, evaluation metrics, and reporting standards limits direct comparison between studies. Some investigations prioritized accuracy, while others emphasized AUC, precision, recall, or mAP. The lack of standardized outcome measures complicates the synthesis of evidence and underscores the need for consensus guidelines in AI-based periodontal research (Chatzopoulos et al., 2025).

The reliance on two-dimensional periapical radiographs represents both a strength and a limitation. Periapical imaging is widely available, cost-effective, and routinely used in clinical practice, making it an ideal target for AI integration (Vilkomir et al., 2024). However, inherent limitations such as anatomical superimposition and projection errors may constrain the maximum achievable diagnostic performance, particularly for complex intrabony defects.

Several studies included in this review incorporated alternative imaging modalities, such as panoramic radiographs or CBCT, to assess periodontal defects (Shetty et al., 2024; Kurt-Bayrakdar et al., 2024). While these studies were not the primary focus of the review, they provide valuable context and suggest that three-dimensional imaging may offer enhanced diagnostic potential when combined with deep learning, albeit with higher cost and radiation exposure.

Clinical implementation of deep learning-based diagnostic tools remains a key challenge. While high performance metrics are encouraging, most studies were retrospective and lacked external validation across diverse populations and imaging systems. Without robust prospective validation, the generalizability of these models to real-world clinical settings remains uncertain (da Silva-Filho et al., 2025).

Ethical and legal considerations must also be addressed prior to widespread adoption. AI systems should be viewed as decision-support tools rather than replacements for clinician judgment. Transparent reporting, explainability of model outputs, and clinician oversight are essential to ensure safe and responsible use in periodontal diagnostics (Schulze et al., 2024).

Another important consideration is integration into clinical workflows. For AI systems to be adopted, they must be user-friendly, time-efficient, and interoperable with existing radiographic software. Studies included in this review largely focused on model development and validation, with limited discussion of practical deployment, representing an important area for future research.

Overall, the findings of this review suggest that deep learning-based analysis of periapical radiographs holds substantial promise for improving the detection of periodontal defects. However, the technology is currently better suited for identifying furcation involvement than for complex intrabony defect classification. Continued methodological refinement, standardized reporting, and multicenter validation studies are necessary to bridge the gap between research and clinical application.

CONCLUSION

This systematic review demonstrated that deep learning models show significant potential for the detection of intrabony and furcation defects on periapical radiographs, with particularly strong performance in furcation involvement classification. While current evidence supports the feasibility of AI-assisted periodontal diagnostics, limitations related to dataset heterogeneity, annotation variability, and the inherent constraints of two-dimensional imaging persist. Future research should prioritize standardized methodologies, larger and more diverse datasets, and prospective clinical validation to facilitate the safe and effective integration of deep learning systems into routine periodontal practice.

REFERENCES

1. Chen, I. H., Lin, C. H., Lee, M. K., Chen, T. E., Lan, T. H., Chang, C. M., Tseng, T. Y., Wang, T., & Du, J. K. (2024). Convolutional-neural-network-based radiographs evaluation assisting in early diagnosis of the periodontal bone loss via periapical radiograph. *Journal of dental sciences*, *19*(1), 550–559. <https://doi.org/10.1016/j.jds.2023.09.032>
2. Zhang, X., Guo, E., Liu, X., Zhao, H., Yang, J., Li, W., Wu, W., & Sun, W. (2025). Enhancing furcation involvement classification on panoramic radiographs with vision transformers. *BMC oral health*, *25*(1), 153. <https://doi.org/10.1186/s12903-025-05431-6>
3. Szabó, V., Orhan, K., Dobó-Nagy, C., Veres, D. S., Manulis, D., Ezhov, M., Sanders, A., & Szabó, B. T. (2025). Deep Learning-Based Periapical Lesion Detection on Panoramic Radiographs. *Diagnostics*, *15*(4). <https://doi.org/10.3390/diagnostics15040510>
4. Chatzopoulos, G. S., Koidou, V. P., Tsalikis, L., & Kaklamanos, E. G. (2025). Clinical Applications of Artificial Intelligence in Periodontology: A Scoping Review. *Medicina (Kaunas, Lithuania)*, *61*(6), 1066. <https://doi.org/10.3390/medicina61061066>
5. da Silva-Filho, J. E., Sousa, Z. D. S., Caracas-de-Araújo, A. P., Fornagero, L. D. S., Machado, M. P., de Aguiar, A. W. O., Silva, C. M., de Albuquerque, D. F., & Gurgel-Filho, E. D. (2025). Deep learning for detecting periapical bone rarefaction in panoramic radiographs: a systematic review and critical assessment. *Dento maxillo facial radiology*, *54*(6), 405–419. <https://doi.org/10.1093/dmfr/twaf044>
6. Schulze, D., Häußermann, L., Ripper, J., & Sottong, T. (2024). Comparison between observer-based and AI-based reading of CBCT datasets: An interrater-reliability study. *The Saudi dental journal*, *36*(2), 291–295. <https://doi.org/10.1016/j.sdentj.2023.11.001>
7. Shetty, S., Talaat, W., AlKawas, S., Al-Rawi, N., Reddy, S., Hamdoon, Z., Kheder, W., Acharya, A., Ozsahin, D. U., & David, L. R. (2024). Application of artificial intelligence-based detection of furcation involvement in mandibular first molar using cone beam tomography images- a preliminary study. *BMC oral health*, *24*(1), 1476. <https://doi.org/10.1186/s12903-024-05268-5>
8. Vilkomir, K., Phen, C., Baldwin, F., Cole, J., Herndon, N., & Zhang, W. (2024). Classification of mandibular molar furcation involvement in periapical radiographs by deep learning. *Imaging science in dentistry*, *54*(3), 257–263. <https://doi.org/10.5624/isd.20240020>
9. Valente, N. A., Americo, L. M., Ciancetta, F., & Mari, S. (2026). Deep Learning-Based Detection of Periodontal Infrabony and Furcation Defects on Periapical Radiographs: A Feasibility Study. *International dental journal*, *76*(2), 109380. Advance online publication. <https://doi.org/10.1016/j.identj.2025.109380>
10. Mao, Y. C., Huang, Y. C., Chen, T. Y., Li, K. C., Lin, Y. J., Liu, Y. L., Yan, H. R., Yang, Y. J., Chen, C. A., Chen, S. L., Li, C. W., Chan, M. L., Chuo, Y., & Abu, P. A. R. (2023). Deep Learning for Dental Diagnosis: A Novel Approach to Furcation Involvement Detection on Periapical Radiographs. *Bioengineering (Basel, Switzerland)*, *10*(7), 802. <https://doi.org/10.3390/bioengineering10070802>
11. Piroonsan, K., Pimolbutr, K., & Tansriratanawong, K. (2025). Classifying Three-Wall Intraony Defects from Intraoral Radiographs Using Deep Learning-Based Convolutional Neural Network Models. *European journal of dentistry*, *19*(3), 721–728. <https://doi.org/10.1055/s-0044-1791784>
12. Kurt-Bayrakdar, S., Bayrakdar, İ. Ş., Yavuz, M. B., Sali, N., Çelik, Ö., Köse, O., Uzun Saylan, B. C., Kuleli, B., Jagtap, R., & Orhan, K. (2024). Detection of periodontal bone loss patterns and furcation defects from panoramic radiographs using deep learning algorithm: a retrospective study. *BMC oral health*, *24*(1), 155. <https://doi.org/10.1186/s12903-024-03896-5>
13. Alotaibi, G., Awawdeh, M., Farook, F. F., Aljohani, M., Aldhafiri, R. M., & Aldhoayan, M. (2022). Artificial intelligence (AI) diagnostic tools: utilizing a convolutional neural network (CNN) to assess periodontal bone level radiographically-a retrospective study. *BMC oral health*, *22*(1), 399. <https://doi.org/10.1186/s12903-022-02436-3>