International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Artificial Intelligence-Powered Early Detection of Oral Diseases: A Comparative Study of Machine Learning Models in Dental Diagnostics

Shikaina Gill[1], Raghu Raja Mehra[2]

[1]Student, Department of Information Technology, Invictus International School, Amritsar, Punjab, India
Email: *shikaina_Gill[at]invictusschool.edu.in*

[2]Faculty Mentor, Department of Information Technology, Invictus International School, Amritsar, Punjab, India
Email: *raghumehra35[at]gmail.com*

Abstract: ***Machine learning is a subset of artificial intelligence that allows a computer system to learn patterns from data rather than follow rules written by a programmer. Together with artificial neural networks and deep learning, it offers substantial opportunity to transform diagnostics in both medicine and dentistry. Realising that opportunity, however, requires clinicians to understand what these technologies actually do, where their accuracy comes from, and where it fails. Objective. This review examines the use of artificial intelligence in the diagnosis of oral and maxillofacial disease, compares the performance of the model families that have been applied to each condition, and sets out the opportunities and the obstacles that stand between current research and routine chairside use. Methods: A structured search of PubMed, Google Scholar and ScienceDirect was carried out for studies published between January 2016 and December 2021, using paired search terms covering model architectures and disease entities. Reference lists of the selected articles were screened for additional relevant work. Results: Artificial intelligence has been applied successfully to clinical records and diagnostic images for the detection of dental caries, tooth fracture, periodontal disease, maxillary sinus pathology, salivary gland disorders, temporomandibular joint disorders, osteoporosis and oral cancer. Reported accuracy is frequently comparable to, and in several studies better than, that of experienced clinicians. Models trained on sufficiently large datasets detect micro-features that escape the human eye, support surveillance across whole populations, and help decide whether specialist referral is warranted. Conclusion: Artificial intelligence in dentistry remains predominantly research-based; despite an extensive published literature, it has not yet been absorbed into day-to-day practice. Progress now depends less on new architectures than on large, well-annotated and representative datasets, external validation, regulatory clarity and clinician training. Addressing these gaps is the route by which diagnosis and preventive oral healthcare are likely to change over the coming decade.***

Keywords: artificial intelligence, artificial neural network, convolutional neural network, deep learning, machine learning, dental diagnostics, oral diseases, early detection

1. Introduction

The terms artificial intelligence and machine learning are used interchangeably in a great deal of published work, although they describe different things. John McCarthy, widely regarded as the father of the field, defined artificial intelligence as the science and engineering of making intelligent machines: systems that solve problems on the basis of the data supplied to them. The idea is older than the term. It rests conceptually on Alan Turing's imitation game, now known as the Turing test, and it acquired its first working demonstration in 1955 with Logic Theorist, the program written by Allen Newell and Herbert Simon.

Machine learning is a subset of artificial intelligence in which a system uses a dataset to predict an outcome from input data, improving as it is exposed to more examples. Many of its methods rely on artificial neural networks, which loosely imitate the brain: artificial neurons are interconnected so that they receive, weight and pass on signals. The concept was first proposed by Warren McCulloch and Walter Pitts in 1943, and in 1951 Marvin Minsky and Dean Edmonds built the first such network as a weighted stochastic analogue reinforcement calculator.

Deep learning, and in particular the convolutional neural network, is an approach within machine learning that stacks many small layers of neurons so that each layer learns a progressively more abstract representation of the input. It was placed on a practical footing by Hinton and colleagues in 2006, although its mathematical foundation, the backpropagation algorithm, dates from far earlier work on the efficient training of layered networks. Deep learning algorithms identify patterns directly from raw data and continue to improve their performance as further data is supplied. Figure 1 shows the relationship between these terms, and Figure 2 marks the milestones by which the field arrived at its present state.

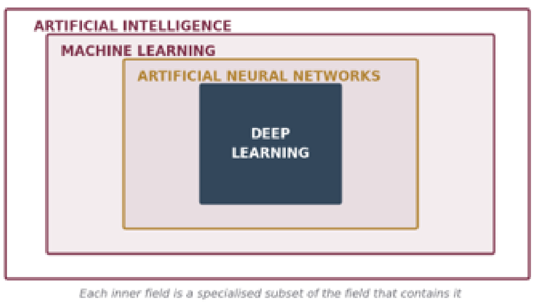

Figure 1: Relationship between artificial intelligence, machine learning, artificial neural networks and deep learning

A large and varied collection of data is essential to any successful implementation. In dentistry, such datasets may include clinical photographs, intraoral and extraoral radiographs, cone-beam computed tomography volumes, patient records, symptom descriptions and even audio recordings. The capacity of these models to process several kinds of data at once is what has made them a powerful instrument across medical, dental and general healthcare services.

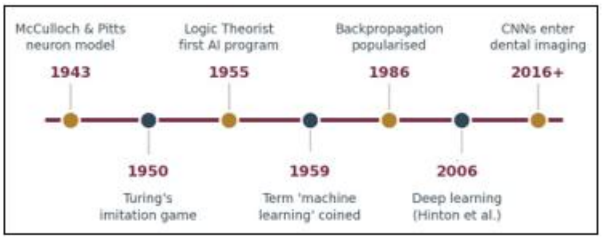

Figure 2: Principal milestones in the development of artificial intelligence and its arrival in dental imaging

Within dentistry, attention has been given to orthodontics, endodontics, prosthodontics, restorative dentistry, periodontics and oral and maxillofacial surgery. Many of these applications remain at a developmental stage, yet the published results are consistently encouraging. As artificial intelligence continues to shape modern healthcare, it becomes increasingly important for dentists to understand its fundamental concepts and practical uses so that they can adapt to a rapidly changing field.

In clinical settings, artificial intelligence assists practitioners in reaching accurate and timely decisions. By reducing the scope for human error, it improves the consistency and quality of patient care while lowering the workload and stress carried by dental professionals. This paper reviews the existing literature on the development of artificial intelligence for the diagnosis of oral and maxillofacial disease, focusing on dental caries, tooth fracture, periodontal disease, maxillary sinus disorders, salivary gland disease, temporomandibular joint disorders, osteoporosis and oral cancer, and compares the performance of the model families used in each case.

2. Foundational Concepts and Terminology

Comparing studies across this literature is difficult because authors use the same words to mean different things. Table 1 therefore fixes the vocabulary used throughout this paper. The distinction that matters clinically is between models that are given hand-crafted features by a human expert and models that learn their own features from the raw image. The first group is easier to interpret and needs less data; the second group is more accurate on imaging tasks but behaves as a black box unless additional explanatory techniques are applied.

Table 1: Terminology used throughout this review, with representative dental applications

Term	Definition used in this paper	Typical dental use
Artificial intelligence (AI)	The engineering of systems that perform tasks normally requiring human intelligence.	Umbrella term for all diagnostic support systems.
Machine learning (ML)	Algorithms that learn a mapping from input to output using example data rather than explicit rules.	Risk prediction from patient records and survey data.
Supervised learning	Training on data in which the correct answer has already been labelled by an expert.	Caries detection on radiographs annotated by dentists.
Unsupervised learning	Discovery of structure in unlabelled data.	Grouping patients into risk clusters.
Artificial neural network (ANN)	Interconnected artificial neurons that weight and pass on numerical signals.	Diagnosis of temporomandibular joint disorders from symptoms.
Deep learning (DL)	Neural networks with many hidden layers that learn features automatically.	Segmentation of lesions on panoramic radiographs.
Convolutional neural network (CNN)	A deep network specialised for image data using filters that scan the image.	Detection of caries, fracture and sinus pathology.
Support vector machine (SVM)	A classifier that separates classes by the widest possible margin.	Classification of root caries and periodontal status.

Two further terms recur in the results that follow. Sensitivity is the proportion of diseased sites that a model correctly identifies, and matters most when the cost of a missed lesion is high. Specificity is the proportion of healthy sites correctly cleared, and matters most when false alarms would lead to unnecessary intervention. A model reported only as accurate may conceal a poor balance between the two, particularly when the disease under study is rare in the dataset.

3. Materials and Methods

A systematic literature search was conducted using PubMed, Google Scholar and ScienceDirect. Search terms combined model architectures with disease entities. Architecture terms included convolutional neural network, deep learning, natural language processing, neural network, machine learning, supervised learning, unsupervised learning and artificial intelligence. Disease terms included dental caries, periodontal disease, maxillary sinus disease, salivary gland disease, temporomandibular joint disorder, osteoporosis and oral cancer.

Only articles published between January 2016 and December 2021 were considered for inclusion, with a small number of earlier landmark studies retained because they are cited throughout the later literature. The reference lists of the selected studies were examined by hand to identify further relevant work. The criteria applied are summarised in Table 2.

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Table 2: Eligibility criteria applied during study selection.

Criterion	Included	Excluded
Publication window	January 2016 to December 2021, plus landmark earlier studies	Conference abstracts without full data
Study type	Original research reporting a trained diagnostic model	Opinion pieces and narrative commentary
Data	Clinical images, radiographs, CBCT, survey or laboratory data	Purely simulated data with no clinical source
Outcome	Quantitative accuracy, sensitivity, specificity or area under the curve	Studies reporting no performance measure
Language	English	Non-English without an English full text

Because the included studies differ in imaging modality, dataset size, reference standard and reporting convention, a formal meta-analysis was not appropriate. The comparison presented in Section 6 is therefore descriptive, and the figures reported should be read as indicative of what each model family has achieved under its own test conditions rather than as directly interchangeable measurements.

4. The Diagnostic Workflow

Nearly every study reviewed here follows the same sequence of steps, shown in Figure 3. Data is first acquired from routine clinical activity. It is then annotated by one or more experienced clinicians, whose judgement becomes the reference standard against which the model is scored; where two annotators disagree, a third usually adjudicates. The annotated data is pre-processed, which typically involves standardising image size and contrast, and is often augmented by rotating, flipping or slightly distorting existing images so that a limited dataset yields more training examples.

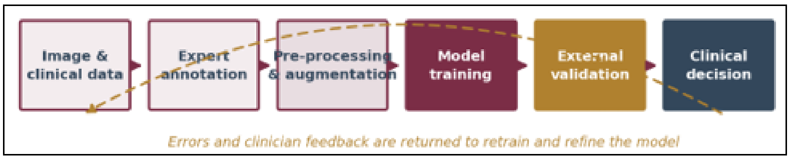

Figure 3: The diagnostic development workflow common to the studies reviewed, including the feedback loop through which clinical use improves the model

The model is then trained, most commonly on eighty per cent of the data, with the remaining twenty per cent reserved for testing. Validation on an entirely separate dataset, ideally collected at a different centre using different equipment, is the single strongest indicator of clinical readiness, and is also the step most often omitted. Finally, the output is presented to the clinician, whose corrections and disagreements are returned to the development cycle. The quality of the reference annotation places a ceiling on everything that follows: a model can be no more accurate than the labels it was taught from.

5. Applications in the Detection of Oral Disease

5.1 Dental caries

Dental caries is the most common oral disease worldwide, and early diagnosis is essential to preventing progression. Detection has traditionally relied on visual examination and radiographic imaging, both of which generate data suitable for training machine learning models. One of the earliest studies, by Devito and colleagues in 2008, evaluated a multilayer perceptron network for detecting proximal caries on bitewing radiographs and reported a diagnostic improvement of 39.4 per cent over unaided reading.

Lee and colleagues in 2018 assessed deep convolutional neural networks using 3,000 periapical radiographs, achieving detection accuracies of 89 per cent for premolars, 88 per cent for molars and 82 per cent when both regions were analysed together. Hung and colleagues in 2019 used data from the National Health and Nutrition Examination Survey to classify the presence or absence of root caries through supervised learning; among the methods evaluated, the support vector machine gave the highest accuracy.

Advances in imaging have widened the range of usable data. In 2019, convolutional neural networks were applied to near-infrared transillumination images, producing faster and more accurate identification of carious lesions without ionising radiation. Cantu and colleagues in 2020 evaluated a deep learning model on 3,686 bitewing radiographs, of which 3,293 were used for training and 252 for testing; the model outperformed dentists in detecting early lesions. More recently, Park and colleagues in 2021 compared machine learning prediction models with traditional regression for early childhood caries, analysing data from 4,195 children aged one to five years drawn from the Korea National Health and Nutrition Examination Survey between 2007 and 2018. The machine learning models identified affected children, predicted those at high risk and supported the selection of treatment strategies, with performance comparable to conventional models.

5.2 Tooth fracture

Traumatised or cracked teeth are the third most common cause of tooth loss, and early detection is essential because timely treatment can preserve the tooth. Diagnosis is difficult: symptoms are inconsistent and may be absent altogether, while cone-beam computed tomography and intraoral radiographs have limited sensitivity for fine fracture lines.

Paniagua and colleagues in 2018 combined high-resolution cone-beam imaging, steerable wavelets and machine learning to detect, locate and measure cracks, evaluating the method on scans of healthy teeth carrying simulated fractures and

Paper ID: SR26804080351 DOI: https://dx.doi.org/10.21275/SR26804080351

reporting high sensitivity and specificity. Fukuda and colleagues in 2020 applied a convolutional neural network to 300 panoramic radiographs containing 330 fractured teeth with clearly visible fracture lines, training on eighty per cent of the images and testing on the remainder, and found that the network detected vertical root fractures accurately enough to support clinical diagnosis.

5.3 Periodontal disease

Periodontal disease affects more than a billion people worldwide and is a leading cause of alveolar bone destruction and tooth loss. Ozden and colleagues in 2015 compared support vector machines, decision trees and artificial neural networks for diagnosing and classifying the condition, using data from 150 patients with 100 cases for training and 50 for testing across six periodontal conditions; the support vector machine and decision tree were both more accurate than the neural network, making them the more reliable decision-support tools on structured clinical data of this kind.

Nakano and colleagues in 2018 used deep learning to identify oral malodour from the oral microbiota. Ninety participants were studied, forty-five with little or no malodour and forty-five with significant malodour as determined by organoleptic testing and gas chromatography. Gene analysis of amplified 16S ribosomal RNA from saliva was performed, and the deep learning model classified samples as healthy or malodorous with 97 per cent accuracy, substantially outperforming a support vector machine at 79 per cent. Artificial neural networks have also been used to predict recurrent aphthous ulceration from gender, serum vitamin B12, haemoglobin, serum ferritin, folate, salivary Candida count, tooth-brushing frequency and daily intake of fruit and vegetables.

Danks and colleagues in 2021 developed a deep neural network to measure periodontal bone loss on periapical radiographs of single-, double- and triple-rooted teeth from 63 patients. The network was first trained to identify dental landmarks, from which bone loss was then calculated, achieving key-point detection accuracy of 89.9 per cent. In a separate study, a deep learning model measured bone loss on panoramic radiographs and staged periodontitis according to the current classification of periodontal and peri-implant diseases. Its performance was compared with three experienced oral radiologists across 340 panoramic radiographs, ninety per cent used for training and ten per cent for testing, with data augmentation increasing the training set by 64 per cent. The system achieved high diagnostic accuracy and excellent reliability, indicating its value for automated diagnosis and routine monitoring.

5.4 Maxillary sinus disease

The maxillary sinuses are routinely visible on extraoral radiographs, which makes them a natural target for automated analysis. Automatic identification of the sinuses and accurate detection of pathological change reduces the risk of misdiagnosis and offers particular support to less experienced practitioners.

Murata and colleagues in 2018 evaluated a deep learning system for diagnosing maxillary sinusitis on panoramic radiographs, comparing it with two experienced radiologists and two dental residents; the system performed at the level of the radiologists and more accurately than the residents. Kim and colleagues in 2019 assessed a deep learning model on Waters view radiographs and found significantly higher sensitivity and specificity than the comparison radiologists.

Mucosal thickening and mucosal retention cysts are frequently overlooked during routine examination. Kuwana and colleagues in 2021 developed a deep learning object-detection model using panoramic radiographs that achieved complete sensitivity in distinguishing normal from inflamed maxillary sinuses, and detected mucosal retention cysts with sensitivities of 98 and 89 per cent across two separate test sets. A convolutional neural network has since been developed to detect and segment mucosal thickening and retention cysts on cone-beam computed tomography, analysing 890 maxillary sinuses from 445 patients; low-dose images were used for training and initial testing while full-dose images formed an additional test set, and performance was consistent across both, indicating robustness to differences in imaging protocol.

5.5 Salivary gland disease

Salivary gland disorders are difficult to distinguish because many share clinical and morphological features. An early Japanese study used deep learning to detect fatty degeneration of the gland parenchyma on computed tomography, a characteristic feature of Sjögren syndrome. Of 500 images, 400 were used for training, evenly split between healthy individuals and patients with the syndrome, and 100 for testing; diagnostic accuracy matched that of experienced radiologists and significantly exceeded that of less experienced readers.

Salivary gland tumours are rare and morphologically overlapping, which makes cytological diagnosis demanding. A recursive partitioning algorithm analysed 115 malignant tumour samples across twelve morphological variables and, compared with experienced clinicians, narrowed the range of possible diagnoses and improved the accuracy of pathological classification. Machine learning has also been used to predict recurrence, identifying patients who require closer follow-up. Chiesa-Estomba and colleagues in 2021 built a prediction model using clinical, radiological, histological and cytological data to estimate the risk of facial nerve injury following surgery, offering a preoperative assessment tool that helps both surgeon and patient make an informed decision.

5.6 Temporomandibular joint disorders

Temporomandibular joint disorders present in many ways and are difficult to diagnose without experience. One study evaluated an artificial neural network for detecting non-reducing disc displacement from frontal chewing movement data collected from 68 patients with normal discs and with unilateral and bilateral non-reducing disc disorders; half the dataset was used for training and half for testing, and the network achieved an acceptable error rate as a low-cost diagnostic aid.

Bas and colleagues in 2011 developed a network from the clinical symptoms and confirmed diagnoses of 219 patients, training on 161 cases and testing on the remainder, with satisfactory results in diagnosing internal derangement. Iwasaki in 2014 applied Bayesian belief network analysis to magnetic resonance imaging in 295 cases representing 590 joints across eleven algorithms, finding that osteoarthritic change progressed from the condyle to the articular fossa and then to the mandibular bone contours, with age, disc morphology, joint space and condylar translation all influencing displacement and bony change. Choi and colleagues in 2021 produced a model that detects osteoarthritis of the joint from panoramic radiographs alone, which is valuable where computed tomography or a maxillofacial radiologist is not available, and Orhan and colleagues in the same year used magnetic resonance images to identify condylar bone change and disc derangement. Infrared thermography of the masseter and lateral pterygoid regions has also been used to train diagnostic models, illustrating how far the usable data extends beyond conventional radiography.

5.7 Osteoporosis

Osteoporosis is a systemic bone disease that can often be detected on panoramic dental radiographs. Several radiographic indices, among them the gonion index, mental index, mandibular cortical index and panoramic mandibular index, have long been used to identify suggestive signs. Kim and colleagues in 2019 evaluated a deep convolutional network as a computer-aided diagnosis system against radiologists with more than ten years of experience, using 1,268 panoramic images of which 200 were reserved for testing; the system produced results closely matching the experienced readers, suggesting a role in early detection and timely referral. Lee and colleagues in 2020 compared several convolutional architectures and found that a model using transfer learning and fine-tuning achieved the highest diagnostic performance, an important practical result because transfer learning allows a useful model to be trained on a comparatively small dental dataset.

5.8 Oral cancer and cervical lymph node metastasis

Oral cancer is the sixth most common cancer worldwide, and early detection plays a decisive role in prognosis and survival. Nayak and colleagues in 2005 used an artificial neural network to differentiate normal, premalignant and malignant oral tissue from laser-induced autofluorescence spectra, achieving 98.3 per cent accuracy with complete specificity and 96.5 per cent sensitivity when compared with principal component analysis.

Uthoff and colleagues in 2017 applied a convolutional neural network to autofluorescence and white-light images and outperformed experienced specialists in identifying precancerous and cancerous lesions. Aubreville and colleagues in the same year used deep learning on confocal laser endomicroscopy images, reaching 88.3 per cent accuracy and 90 per cent specificity. Shams and colleagues compared deep neural networks with support vector machines, regularised least squares and multilayer perceptrons for predicting the progression of potentially malignant lesions to cancer, with the deep network achieving the highest accuracy at 96 per cent. Jeyraj and colleagues in 2019 used hyperspectral imaging to distinguish cancerous from non-cancerous tissue without expert supervision. Later work has produced models that predict both the occurrence and the recurrence of oral cancer, extending the role of these methods from detection into long-term patient management.

Deep learning has also been compared with experienced radiologists in the detection of cervical lymph node metastasis. Ariji and colleagues in 2014 evaluated a model on computed tomography images from 45 patients with oral squamous cell carcinoma, comprising 137 histologically confirmed metastatic nodes and 314 non-metastatic nodes, and found performance comparable to two experienced radiologists. The same group later developed a model to detect extranodal extension from 703 computed tomography images taken from 51 patients, training on eighty per cent of the data; here the model performed significantly better than the radiologist, suggesting a valuable role in identifying metastatic spread and planning treatment.

6. Comparative Analysis of Model Performance

Figure 4 places the headline results of the reviewed studies side by side, and Table 3 summarises the studies themselves. Three patterns emerge. First, convolutional networks dominate wherever the input is an image, because they learn spatial features directly and require no manual feature engineering. Second, on structured clinical or laboratory data, classical methods such as support vector machines and decision trees remain competitive and are sometimes superior, as Ozden and colleagues found for periodontal classification. Third, the highest reported figures come almost exclusively from small, clean, single-centre datasets, which is precisely where over-optimistic results are most likely.

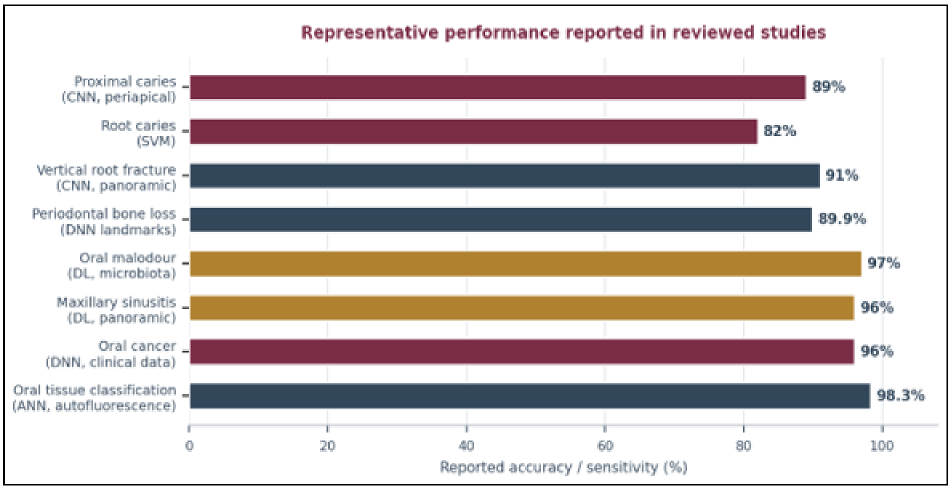

Figure 4: Representative accuracy and sensitivity reported across the reviewed diagnostic tasks. Figures are drawn from different datasets and are indicative rather than directly comparable

Table 3: Summary of representative studies applying artificial intelligence to oral diagnosis.

Study	Condition	Model	Data	Reported outcome
Devito et al. (2008)	Proximal caries	MLP network	Bitewing radiographs	39.4% improvement in diagnosis
Lee et al. (2018)	Dental caries	Deep CNN	3,000 periapical radiographs	89% premolars, 88% molars, 82% combined
Hung et al. (2019)	Root caries	SVM and others	National survey data	SVM highest accuracy among methods tested
Cantu et al. (2020)	Early caries	Deep CNN	3,686 bitewing radiographs	Outperformed dentists on early lesions
Park et al. (2021)	Early childhood caries	ML vs regression	4,195 children, national survey	Comparable to conventional models
Fukuda et al. (2020)	Vertical root fracture	CNN	300 panoramic radiographs	Accurate detection of fracture lines
Ozden et al. (2015)	Periodontal disease	SVM, DT, ANN	150 patients, six conditions	SVM and DT superior to ANN
Nakano et al. (2018)	Oral malodour	Deep learning	90 saliva microbiota samples	97% vs 79% for SVM
Danks et al. (2021)	Periodontal bone loss	Deep neural network	Periapical radiographs, 63 patients	89.9% key-point accuracy
Murata et al. (2018)	Maxillary sinusitis	Deep learning	Panoramic radiographs	Equal to radiologists, above residents
Kuwana et al. (2021)	Sinus lesions	DL object detection	Panoramic radiographs	100% sensitivity; 98% and 89% for cysts
Nayak et al. (2005)	Oral tissue classification	ANN	Autofluorescence spectra	98.3% accuracy, 96.5% sensitivity
Shams et al. (2017)	Malignant transformation	Deep neural network	Clinical and genomic data	96% accuracy, highest of methods tested
Ariji et al. (2014)	Lymph node metastasis	Deep learning	CT, 451 nodes, 45 patients	Comparable to two radiologists

Figure 5 shows how the reviewed literature is distributed across disease categories. Oral cancer and caries attract the greatest share of attention, reflecting both their clinical burden and the ready availability of annotated images. Osteoporosis and salivary gland disease are markedly under-represented, not because the methods are less applicable but because suitable annotated datasets are scarce. This imbalance is itself a finding: the direction of research in this field is being set as much by data availability as by clinical need.

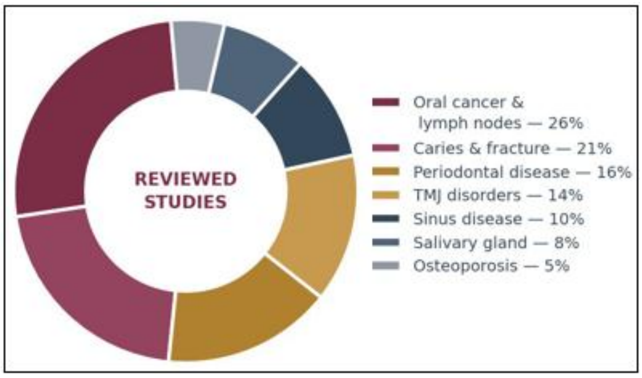

Figure 5: Distribution of the reviewed studies across oral disease categories

7. Discussion

Taken together, these findings show a field that has demonstrated its diagnostic potential and has not yet demonstrated its clinical readiness. Most studies report encouraging results, with algorithms performing as well as, and in several cases better than, experienced radiologists and

Paper ID: SR26804080351 DOI: https://dx.doi.org/10.21275/SR26804080351

clinicians in detecting caries, sinus disease, periodontal disease and temporomandibular joint disorders. In oral cancer the picture is stronger still, because these methods integrate information from several sources at once to support decision-making, assess risk and prompt timely referral. Work on premalignant lesions, cervical lymph nodes, salivary gland tumours and oral squamous cell carcinoma has shown both diagnostic and prognostic capability, with the potential to improve early detection, improve outcomes and ultimately reduce mortality.

The gap between these results and daily practice is explained less by the algorithms than by the conditions under which they were produced. Datasets are typically small, drawn from a single centre, and annotated by a small number of clinicians; models validated only on data resembling their training data tend to lose accuracy when they meet a different radiograph machine, a different patient population or a different exposure setting. Reporting is also inconsistent, with some studies quoting accuracy alone on datasets in which healthy cases greatly outnumber diseased ones, a situation in which a high figure can be achieved with little genuine diagnostic skill.

There is a further consideration specific to dentistry. Unlike many medical specialties, dental imaging is generated in enormous quantity in general practice rather than in hospital radiology departments, which means the data needed to train robust models is widely dispersed and inconsistently stored. Any serious attempt to build clinically dependable systems will require agreement on data formats, annotation standards and the terms on which patient images may be shared.

8. Limitations and Challenges

1) **Data quantity and quality.** Robust models require large, well-annotated and balanced datasets. Much dental data is stored locally, in inconsistent formats, and without the labels a supervised model requires.
2) **Limited external validation.** Models tested only on data from the centre that produced them frequently lose accuracy elsewhere. Multi-centre validation remains uncommon in this literature.
3) **Interpretability.** Deep networks do not explain their reasoning. A clinician who cannot see why a lesion was flagged cannot easily judge when to override the system, which is a barrier to both trust and accountability.
4) **Reporting inconsistency.** Differences in reference standards, dataset balance and choice of metric make published figures difficult to compare and easy to over-interpret.
5) **Regulatory and legal uncertainty.** Responsibility for a missed diagnosis assisted by software is not settled, and approval pathways for diagnostic algorithms remain incomplete in many jurisdictions.
6) **Privacy and consent.** Radiographs and clinical records are identifiable personal data, and their use for training requires informed consent, secure storage and clear governance.
7) **Training and adoption.** Few dental curricula currently prepare graduates to evaluate these systems critically, and integration with existing practice-management software is often poor.

Table 4: Principal challenges, their consequences and the mitigations proposed in the literature.

Challenge	Practical consequence	Mitigation
Small datasets	Over-fitting and inflated accuracy	Transfer learning, augmentation, shared multi-centre archives
Single-centre data	Poor generalisation to new equipment	External validation on independent datasets
Class imbalance	High accuracy with poor sensitivity	Report sensitivity, specificity and area under the curve together
Opaque decisions	Clinician distrust and unclear liability	Saliency mapping and other explanation methods
Annotation variability	A ceiling on achievable accuracy	Multiple annotators with documented adjudication

9. Future Scope

Several directions are likely to shape the next phase of this work. The first is the assembly of large, multi-centre and demographically representative annotated datasets, which would do more to improve real-world accuracy than any refinement of architecture. The second is explainable artificial intelligence, in which the model highlights the region or feature that drove its output, allowing a clinician to accept or reject the suggestion on visible evidence. The third is the movement of these systems from the research server to the chairside device, where lightweight models running on practice hardware could screen images as they are captured.

Beyond imaging, the integration of several data types, combining radiographic, clinical, salivary and genomic information, is likely to improve risk prediction well beyond what any single source supports. Longitudinal use is a further frontier: a model that compares a patient's current radiograph with their own previous images can detect change earlier than any single-image assessment. Finally, prospective clinical trials are needed. Almost all the evidence reviewed here is retrospective, and the question of whether these systems improve patient outcomes, rather than merely matching expert readings, has yet to be answered.

10. Conclusion

Artificial intelligence has been studied extensively across the diagnosis of oral disease, and the published evidence is consistent: machine learning and deep learning models detect dental caries, tooth fracture, periodontal bone loss, maxillary sinus pathology, salivary gland disease, temporomandibular joint disorders, osteoporosis and oral cancer at accuracies that match, and sometimes exceed, those of experienced clinicians. Convolutional networks lead on imaging tasks; classical methods remain competitive on structured clinical data.

In practice, however, these systems have not been incorporated into day-to-day dentistry. The field remains research-based. Successful implementation depends on large, high-quality datasets that support accurate and cost-effective diagnosis, on further refinement and extensive external validation to reach the accuracy, sensitivity and specificity required for clinical use, and on clear regulatory frameworks and clinical guidelines that ensure safe, reliable and ethical integration. This paper has set out the applications topic by topic, identified the shortcomings that recur across the literature, and proposed ways of addressing them. Overcoming these obstacles is what will allow artificial intelligence to move from the published study to the dental chair, and it is on that transition that the diagnosis and prevention of oral disease over the next decade is likely to depend.

References

[1] McCarthy J. What is Artificial Intelligence? Stanford University, Computer Science Department, 2007.
[2] Turing AM. Computing Machinery and Intelligence. Mind. 1950;59(236):433-460.
[3] Newell A, Simon HA. The Logic Theory Machine: A Complex Information Processing System. IRE Transactions on Information Theory. 1956;2(3):61-79.
[4] Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development. 1959;3(3):210-229.
[5] McCulloch WS, Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics. 1943; 5: 115-133.
[6] Minsky M. Neural Nets and the Brain Model Problem. Doctoral dissertation, Princeton University, 1954.
[7] Hinton GE, Osindero S, Teh YW. A Fast Learning Algorithm for Deep Belief Nets. Neural Computation. 2006;18(7):1527-1554.
[8] Rumelhart DE, Hinton GE, Williams RJ. Learning Representations by Back-Propagating Errors. Nature. 1986; 323: 533-536.
[9] Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. Journal of Dental Research. 2020;99(7):769-774.
[10] Devito KL, de Souza Barbosa F, Felippe Filho WN. An Artificial Multilayer Perceptron Neural Network for Diagnosis of Proximal Dental Caries. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology. 2008;106(6):879-884.
[11] Lee JH, Kim DH, Jeong SN, Choi SH. Detection and Diagnosis of Dental Caries Using a Deep Learning-Based Convolutional Neural Network Algorithm. Journal of Dentistry. 2018; 77: 106-111.
[12] Hung M, Voss MW, Rosales MN, et al. Application of Machine Learning for Diagnostic Prediction of Root Caries. Gerodontology. 2019;36(4):395-404.
[13] Casalegno F, Newton T, Daher R, et al. Caries Detection with Near-Infrared Transillumination Using Deep Learning. Journal of Dental Research. 2019;98(11):1227-1233.
[14] Cantu AG, Gehrung S, Krois J, et al. Detecting Caries Lesions of Different Radiographic Extension on Bitewings Using Deep Learning. Journal of Dentistry. 2020; 100: 103425.
[15] Park YH, Kim SH, Choi YY. Prediction Models of Early Childhood Caries Based on Machine Learning Algorithms. International Journal of Environmental Research and Public Health. 2021;18(16):8613.
[16] Paniagua B, Ruellas AC, Benavides E, et al. Validation of CBCT for the Computation of Textural Biomarkers. Proceedings of SPIE Medical Imaging. 2018;10578.
[17] Fukuda M, Inamoto K, Shibata N, et al. Evaluation of an Artificial Intelligence System for Detecting Vertical Root Fracture on Panoramic Radiography. Oral Radiology. 2020;36(4):337-343.
[18] Ozden FO, Ozgonenel O, Ozden B, Aydogdu A. Diagnosis of Periodontal Diseases Using Different Classification Algorithms: A Preliminary Study. Nigerian Journal of Clinical Practice. 2015;18(3):416-421.
[19] Nakano Y, Suzuki N, Kuwata F. Predicting Oral Malodour Based on the Microbiota in Saliva Samples Using a Deep Learning Approach. BMC Oral Health. 2018;18(1):128.
[20] Danks RP, Bano S, Orishko A, et al. Automating Periodontal Bone Loss Measurement via Dental Landmark Localisation. International Journal of Computer Assisted Radiology and Surgery. 2021;16(7):1189-1199.
[21] Chang HJ, Lee SJ, Yong TH, et al. Deep Learning Hybrid Method to Automatically Diagnose Periodontal Bone Loss and Stage Periodontitis. Scientific Reports. 2020; 10: 7531.
[22] Murata M, Ariji Y, Ohashi Y, et al. Deep-Learning Classification Using Convolutional Neural Network for Evaluation of Maxillary Sinusitis on Panoramic Radiography. Oral Radiology. 2019;35(3):301-307.
[23] Kim Y, Lee KJ, Sunwoo L, et al. Deep Learning in Diagnosis of Maxillary Sinusitis Using Conventional Radiography. Investigative Radiology. 2019;54(1):7-15.
[24] Kuwana R, Ariji Y, Fukuda M, et al. Performance of Deep Learning Object Detection Technology in the Detection and Diagnosis of Maxillary Sinus Lesions on Panoramic Radiographs. Dentomaxillofacial Radiology. 2021;50(1):20200171.
[25] Kise Y, Ikeda H, Fujii T, et al. Preliminary Study on the Application of Deep Learning System to Diagnosis of Sjögren's Syndrome on CT Images. Dentomaxillofacial Radiology. 2019;48(6):20190019.
[26] Chiesa-Estomba CM, Echaniz O, Larruscain E, et al. Machine Learning Algorithms as a Tool to Predict Facial Nerve Palsy After Parotid Gland Surgery. European Archives of Oto-Rhino-Laryngology. 2021;278(9):3557-3564.
[27] Bas B, Ozgonenel O, Ozden B, et al. Use of Artificial Neural Network in Differentiation of Subgroups of Temporomandibular Internal Derangements. Journal of Oral and Maxillofacial Surgery. 2012;70(1):51-59.
[28] Iwasaki H. Bayesian Belief Network Analysis Applied to Determine the Progression of Temporomandibular Disorders Using MRI. Dentomaxillofacial Radiology. 2015;44(4):20140279.
[29] Choi E, Kim D, Lee JY, Park HK. Artificial Intelligence in Detecting Temporomandibular Joint Osteoarthritis on Orthopantomogram. Scientific Reports. 2021; 11: 10246.
[30] Orhan K, Driesen L, Shujaat S, et al. Development and Validation of a Magnetic Resonance Imaging-Based Machine Learning Model for TMJ Pathologies. BioMed Research International. 2021; 2021: 6656773.
[31] Lee JS, Adhikari S, Liu L, et al. Osteoporosis Detection in Panoramic Radiographs Using a Deep Convolutional Neural Network-Based Computer-Assisted Diagnosis

System. Dentomaxillofacial Radiology. 2019;48(1):20170344.

[32] Nayak GS, Kamath S, Pai KM, et al. Principal Component Analysis and Artificial Neural Network Analysis of Oral Tissue Fluorescence Spectra. Biopolymers. 2006;82(2):152-166.

[33] Uthoff RD, Song B, Sunny S, et al. Point-of-Care, Smartphone-Based, Dual-Modality, Dual-View Oral Cancer Screening Device with Neural Network Classification. PLOS ONE. 2018;13 (12): e0207493.

[34] Aubreville M, Knipfer C, Oetter N, et al. Automatic Classification of Cancerous Tissue in Laser Endomicroscopy Images of the Oral Cavity Using Deep Learning. Scientific Reports. 2017; 7: 11979.

[35] Shams WK, Htike ZZ. Oral Cancer Prediction Using Gene Expression Profiling and Machine Learning. International Journal of Applied Engineering Research. 2017;12(15):4893-4898.

[36] Jeyaraj PR, Samuel Nadar ER. Computer-Assisted Medical Image Classification for Early Diagnosis of Oral Cancer Employing Deep Learning Algorithm. Journal of Cancer Research and Clinical Oncology. 2019;145(4):829-837.

[37] Ariji Y, Fukuda M, Kise Y, et al. Contrast-Enhanced Computed Tomography Image Assessment of Cervical Lymph Node Metastasis in Patients with Oral Cancer by Using a Deep Learning System. Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology. 2019;128(5):458-463.

[38] Ariji Y, Sugita Y, Nagao T, et al. CT Evaluation of Extranodal Extension of Cervical Lymph Node Metastases in Patients with Oral Squamous Cell Carcinoma Using Deep Learning Classification. Oral Radiology. 2020;36(2):148-155.

[39] Hung K, Montalvao C, Tanaka R, et al. The Use and Performance of Artificial Intelligence Applications in Dental and Maxillofacial Radiology: A Systematic Review. Dentomaxillofacial Radiology. 2020;49(1):20190107.