AI-Powered Multi-Label Dental Disease Detection from Panoramic Radiographs A YOLO-Based Review and Future Directions

Aradhana S. Thorat[1] and Dr. Archana Chaudhari[2]

[1] Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Pune, India
*archana.chaudhari@sitpune.edu.in

Abstract. The panoramic radiographs offer a synoptic view of the dentoalveolar anatomy but is cumbersome to analyze reliably in the presence of concurrent caries, loss of periodontal bone and impactions. By utilizing modern deeply learned object detectors (YOLO-family) that have become commoditized, the review challenges this technology to identify multiple diseases at once, give attributions of individual teeth, and combine automated object detection with patient covariates like age, oral hygiene, and medical history. YOLO models have been found to perform high-fidelity localization of teeth and lesions empirically and with high speed, compared to segmentation paradigms such as U-Net and Mask R-CNN that are highly accurate but have a very high computational latency, which limits their applications to real-time deployment. The initial multimodal frameworks that combine imaging with patient metadata are characterized by significant increases in diagnostic accuracy as well as predictive power. Nevertheless, significant challenges remain to be overcome most notably due to the lack of large well-marked datasets, the lack of automated FDI-based tooth indexing schemes, and insufficient architectural studies in the field of multi-pathology detection in a single panorama. Overall, the synthesis highlights the fact that YOLO-based pipelines are best poised to enable the use of dental radiographic interpretation, provided that standardized datasets and fine multi-label frameworks are created, and the smooth integration of clinical metadata into individual diagnostic intelligence is facilitated. The review distinctly outlines the new paradigm of YOLO in panoramic radiography in the integrated multi-label dental disease analytics.

Keywords: Panoramic radiography, Dental AI, YOLO, multi-label disease detection, tooth-level localization, FDI tooth numbering, deep learning, clinical data integration, dental imaging analytics, diagnostic automation

1 Introduction

The toothwise decision-making process is based on panoramic radiographs (pantomograms), making it difficult to assess poly-pathology as they have superposition, acquisition anisotropy, and operator-dependent interpretation. The resulting potential impact of automated odonto-analytics with explicit FDI indexing is therefore profound for clinical triage, forensics attribution and population studies, if models can link lesion localization to calibrated, scanners and population/age/gender independent deployment-grade efficiency and transportability.

Profound innovations cut across complementary paradigms. Large-scale single-tooth pipelines to estimate ages, sex, and tooth-type Milošević et al. [1] to age pipelines, sex pipelines and tooth-type pipelines with high accuracy rates, and with sensitivity to change and adult ambiguity expose sensitivity to change and adult ambiguity. Applied to entire-arch OPGs, Aljabar et al. [2] showed that a hybrid regime of transfer learning consisting of deep feature extractors and classical classifiers can achieve competitive accuracy (e.g.

A. Vaidya et al. (eds.), *Proceedings of the Conference on Bioengineering for Global Health (SYMRESEARCH 2.0 2025)*, Advances in Biological Sciences Research 55,
https://doi.org/10.2991/978-94-6239-756-9_13

ViT-level F-scores) but at significantly lower computational cost that could be deployed on a chairside device. Through the same direction of research, Bussaban et al. [3] tested age-group stratification with increasing class granularity and found that VGG19 is more powerful in coarse stratifications and EfficientNet in fine-grained partitions, thus explaining the complex-performance trade space. Jiang et al. [4] tackled orthogonally the CBCT tooth segmentation through another level-set scheme with a controlled switch achieving a Dice score of approximately 94% and strong ability to handle blurred or missing boundaries - which is essential in downstream odonto-cartography.

Nonetheless, persistent impediments include data scarcity and class imbalance, institution-specific bias, low-resolution cues confounding cavities versus restorations, overfitting to rare phenotypes, and residual manual steps (e.g., voxel seeding) that hinder full automation. Unresolved gaps encompass standardized, publicly accessible pantomograms corpora with harmonized nosology; FDI-aware, multi-label detection benchmarks with per-tooth calibration and efficiency metrics; principled clinico-radiographic fusion; and rigorous cross-vendor, age-aware validation. Building on the foregoing, the literature synthesis now progresses from single-tooth forensics to panoramic hybridization and volumetric segmentation, tracing how methodological choices, dataset ecology, and evaluation protocols co-determine generalizable, toothwise diagnostics across real-world odontometric variability.

2 Literature Survey

2.1 Computational Architectures for Dental Image Intelligence

In recent literature, the field of dental image intelligence has converged on uncertainty-sensitive representation learning, contrast-learned 2D pipelines, and geometry-sensitive 3D instance reasoning. Zhang et al. [5] reformulates biological age estimation on orthopantomograms as distribution learning, proposes UASP-BAE with Gaussian soft-label uncertainty mapper, adaptive channel-attention stem, and cross self-attention fused with CNN locality and transformer-scale global context, and a special age-composition loss and sex-prior heads at population scale. Inspired by the effect of the clinical material on the grounding algorithm, Wuersching et al. [6] show how ion releasing bioactive restoratives and adhesive protocols affect marginal integrity and bacteria leakage and local pH, motivating stable detectors and segmenters to edge leakage, texture drift and chemistry-induced contrast deformations. The translation of these limits into orthodontic planning, Cuzin et al. [7] take advantage of the customized lingual digital configurations and utilize them to measure interproximal enamel reduction in tooth-wise priors (size ratios, torque, overjet/overbite) to produce structured, anatomy-sensitive targets capable of regularizing the learning and cropping of occlusion.

Expanding on it, Chen et al. [8] operationalize a panoramic intelligence stack that initially splits jaw topology in Riemann-sum slit capture and overlapping tooth template followed by stabilization of intensity statistics with gray-level transforms, Gaussian/high-pass filtering, histogram equalization, and morphological open-close before assembling modern CNNs under fuzzy decision fusion, and attains accuracy of nearly ceiling on retained roots, implants, and endodontically treated teeth. In further development of the above, Nistelrooij et al. [9] propose ToothInstanceNet on intraoral-scanner point clouds: a two-stage pipeline using low-resolution seed map, offsets, bandwidths and labeling with quadrant-wise FDI, followed by high-resolution crops/graph-cut refinement: again, oversampling of mixed-dentition instances encourages near-perfect detection/segmentation with less than two-second throughput.

Table 1. Distinct methodological and dataset/protocol characteristics highlighting implications for computational dental image intelligence.

Ref No.	Methodological character	Distinctive ops/dataset note	Headline outcome
Zhang et al. [5]	Outputs CI-style age distributions via Gaussian soft-labels, enabling uncertainty-aware clinical reporting.	Day-level labels let the loss function exploit fine temporal granularity rather than coarse bins.	Best MAE ≈ 0.80 y on 10,703 OPGs.
Wuersching et al. [6]	Finds no direct link between intrinsic antibacterial activity and marginal microleakage.	Sustained alkaline pH (~9.6–9.8) for alkasite (with refreshed water) over months noted.	Seal quality favored adhesive-pretreated groups despite mixed antimicrobial

			trends.
Cuzin et al. [7]	Reports systematic mesial > distal planned reduction on several teeth—an asymmetry needing explanation.	Global planning dataset (49 countries) used purely as virtual prescriptions, not execution logs.	Lower canines emerge as the most frequently planned IER targets.
Chen et al. [8]	Uses data-adaptive binarization (threshold from mean & SD) to surface implants/endodontic features pre-CNN.	Prioritizes retained-root routing before multi-class CNN ensemble.	End-to-end pipeline reaches ≈98.75% overall accuracy.

Table 1 generalizes orthogonal methodological axes whose five studies prove data provenance, granularity of supervision and the informative biases of their computational pipelines. It compares the heterogeneous acquisition geometries of 2D radiography, ex vivo specimen tests, virtual orthodontic planning results and 3D intraoral point clouds to explain the determination of preprocessing by these regimes, the topology of features and the capacity to represent in these regimes. The synopsis identifies unique supervisory signals and goals (uncertainty-conscious soft labels, biofilm/pH surrogates, digitally-optimal IER targets, fuzzy ensemble adjudication and two-stage instance-label coupling) which shape optimization space, calibration requirements, and failure patterns.

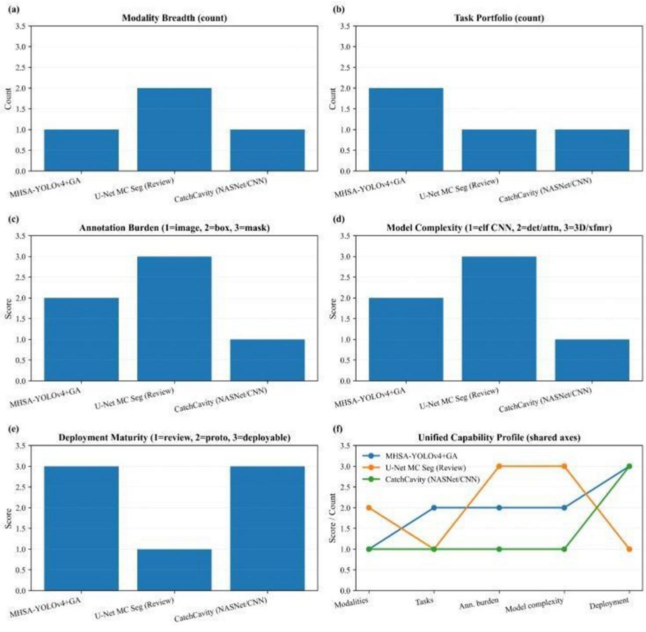

Fig 1 (a–f): Cross-study Metrics for Dental Image Intelligence

Figure 1 (a–f) is a condensed version of cross-study comparability in five orthogonal planes to reveal the operational envelope of each system. The U-Net centric MC segmentation review has the largest breadth of modalities (2D OPG + 3D CBCT), and the most annotation granularity and model complexity (mask-level supervision with 3D/transformer-heavy pipelines), but the loss of deployment maturity, a characteristic of a research-survey phenotype. MHSA-YOLOv4+GA focuses on a single modality, however, has a more general task portfolio (detection + classification), middle annotation load (bounding boxes), complexity of detectors/attention which indicates high deployment availability, and is near real-time operable. CatchCavity (NASNet/CNN) is an image-level (minimal labeling) consumer-grade color classification pipeline, based on a lightweight production pipeline, and has production-grade web-scale deployment, sacrificing breadth and architectural complexity to accessibility. These trade-offs are captured in the unified capability profile (f), which presents a Pareto-like comparison segmentation-oriented depth versus detector versatility versus pragmatic deployability, therefore, allowing principled selection when a dataset, a compute and a clinical workflow is constrained [10–12].

2.2 Dataset Ecology, Annotation Governance, and Multimodal Syntheses

To facilitate strict cross-paper synthesis for each paper along three dimensions, i.e. dataset ecology (modality, scale, heterogeneity), annotation governance (label granularity, curator strategy, QA/consensus), and multimodal synthesis (fusion topology, cross domain generation, validation). The table will be a small decision support tool that can easily map different types of experimental design into a common framework that can then be used to compare like with like without over-indexing on idiosyncratic metrics. It brings external validation practices and financial single point of failure risks to the forefront in order to elucidate translational maturity and reproducibility. It allows you to quickly isolate powerful label aggregation regimes, data-completeness approaches (e.g. synthetic modality recovery), and fusion architectures relevant to your pipeline.

Table 2. Comparative Landscape Of Dataset Ecology, Annotation Governance, And Multimodal Synthesis For Dental AI Pipelines

Ref No.	Modality & scale	Label granularity	Annotation governance	Primary limitation
Jones et al. [13]	IOS (RGB + fluor + geom); MIS BAIR 332 teeth; VASCFIND 119	Pixel masks: initial/moderate/extensive	Dual-expert masks; focal loss; practitioner agreement benchmark	Early/moderate lesions small; projection loses cues
Neeraja et al. [14]	OPG; n=153 (single site)	Pixel masks (caries ROIs)	2 senior dentists; boundary-focused loss	Small, single-center cohort
Li et al. [15]	Bitewing X-rays; class-balanced + augmentation	Boxes (teeth) + per-tooth class	Curated pipeline; CLAHE/geom transforms	Image-quality sensitivity; sparse governance detail
Wang et al. [16]	PANO; 996 cases / 29,112 instances	Instance masks (enamel/dentin/pulp) + anomaly tags	Adversarial DSIS; pseudo-labels; DPR refinement	Slow, stochastic DPR; over/under-refinement risk
Jia et al. [17]	PANO; TDD 1000 + private 200 (TDD-P)	Semantic tooth masks	2 experts; frequency-guided attention; robust preprocessing	Compute cost; artifact sensitivity
Peskersoy et al. [18]	Clinical + saliva ELISA; 160 COVID / 160 HI	Clinical indices + cytokine/Ig titers	Standard clinical + lab protocols	Not imaging; acute-phase focus
Klein et al. [19]	Radiographs + NILT; 1,007 surfaces; histology gold	Surface classes (sound/enamel/dentin)	MV/WMV/DS/MACE compared; probabilistic > MV	Few annotators; class imbalance
Wankhade et al. [20]	Raman: 1D spectra + 2D spectrogram; n=950	Case-level (benign/malignant)	Standard splits; QC'd spectra	Noise sensitivity; moderate cohort size
Pavarut et al. [21]	CT (plain/early/delayed) + MRI (ADC,T2-w); TMDU	Tumor masks; AML vs CCRCC	VGG-16 fine-tune; feature-level fusion	Modality pairing scarce; PSNR/SSIM limits for QA

2.3 Deployment Constraints: Robustness, Artifact Suppression, Self-Supervision, and Translational Interfaces

The YOLO-based pipelines are still evolving into deployable systems of dental radiography based on CAD, but the practicality depends on the lack of data, noise of labels, and imaging artefacts. Using small architectures, Razaghi et al. [22] trained a set of YOLOv8 variants with well-curated bitewing/OPG corpora with heavy augmentations to achieve a best mAP of 71.6 as a good-precision/good-recall system, while revealing a phenomenon of localized brittle behavior to classes (e.g., crowns vs. amalgams) that makes spacecrafts distribution sensitive and necessitates the relationship between two independent classes. Instead of enlarging classifier capacity because of hardware-induced corruption impairment, Pal et al. [23]

improved the sinogram inpainting by multi-band prior syntheses of high-frequency detail using the raw CT with low-frequency metal-proximal structure to facilitate downstream replacement (LB-MAR/NMAR) and the disappearance of beam-hardening streaks—an artefact-suppression condition to inpainting stability. Along the segmentation axis, Ghafoor et al. [24] combined a M-Net scaffold, Swin Transformers, and a boundary-aware Teeth Attention Block to refine the tooth edges in cluttered images and found them to be more than one-quarter the multi-scale loss, with squared-Dice loss again balancing classes to mitigate edge case situations, and Ozcelik et al. [25] implemented tooth-level identity by integrating FPN-based multi-scale features, flexible anchor, and RoI Align in Dent

Based on these advances in discriminating, it is through measurement interfaces and supervision regimes that translational reliability is limited. Esposito et al. [28] showed only mild-fair inter-observer correlation between text articulating paper and intraoral-scanner occlusal maps as requiring objective contact surrogates and emphasizing restriction of annotation governance prior to clinical roll-out. To mitigate the risk of labeled scarcity and domain shift, Almalki et al. [29] learned Masked Deep Embeddings of Patches- self-supervised ViT pretraining, which re-creates latent patch embeddings instead of pixels, to achieve better average precision on periapical and panoramic datasets, and allows fine-tuning with Mask R- CNN heads. In addition to imaging, deployment will need to be robust in its presence as a population structure and regulatory ecosystem: Nagar et al. [27] measured prevalence of anterior nonmetric traits and dimorphism, which, as well as the phenotypic heterogeneity they ingate, has a potential capacity to disrupt model priors; Sarkar et al. [26] cataloged the patent and regulatory environment of biodegradable implants, bringing to light how jurisdiction constraints moderate the translation timelines and data access. Lastly, transferring this paradigm to low-resource environments, Zaheer et al. [30] introduced a school-based toothbrushing/handwashing/SDF initiative in refugee camps and found that it offered significant caries arrest and a viable interface in which AI decision-support has the potential to jointly implement preventive workflows with exceptionally constrained operational conditions.

2.4 YOLO-Based Paradigms for Panoramic Dental Disease Detection

Recent work consolidates YOLO as the de-facto real-time detector for dental radiography, expanding from tooth localization to multi-label pathology pipelines. Complementing this, Abed et al. [31] used an enhanced YOLOv8 (PAF + C2f + SPP, tuned layer sizes) for top-view intraoral tooth detection/segmentation, reporting mAP50 $\approx$ 99.56% and ~1.66 s per image on 526 photos (augmented to 3,156), while noting limited generalization beyond clear, color, top-view images. Continuing this line, Zirek et al. [33] improved YOLOv8 to clinic-level accuracy/latency impacted-tooth localization and Winter-class angulation forecasting, and released the model as a GUI, with algorithmic response and practitioner interaction to each other-critically-important translational inflection points. Still on the paradigm, Xue et al. [34] placed YOLOv8 as the orchestration front end in a compound pipeline (Mask R- CNN to tooth outlines, TransUNet to tissue masks and key points) to automate the quantification of alveolar bone-loss and staging periodontitis with a high concordance to expert findings.

Simultaneously, population screening and cloud-based workflow deployment-oriented tests of Putra et al. [36] preferred the YOLOv8-s variant due to an ideal mAP-latency tradeoff and high sensitivity with low confidence thresholds, hence showing feasible usefulness in population screening and workflow deployment. The ability of a clinician-inspired multi-channel quadrant classifier to outperform YOLOv8 in recall with label scarcity inspired counterpoint evidence by Fang et al. [32], which encourages hybrid routers that send cases to lightweight classifiers and YOLO heads. In the context of these findings, Ragab et al. [35] generalized the architectural development of YOLO the architectural backbone/neck/head stratification, the PAN/FPN/bi-FPN interconnections, and the current loss of IoU explain why a one-stage detector is the preferred substrate in dental AI where multi-object detection, real-time constraints, and relatively small compute budgets are in the ascendancy.

2.5 Research Gaps

The main gaps in the field of dental AI research are heterogeneous datasets, weak control of annotation, low multimodal fusion, low ability to detect small lesions, panoramic artifacts, and lack of linking of FDI-compliant teeth, which prevents generalization, calibration, and clinical implementation across institutions.

Table 3. Identified Research Gaps in Feature Engineering for Botnet Detection

Category of research gap	Description	Recent advancement	Actionable steps
Multi-label, per-tooth detection	One image → many diseases per tooth not handled jointly	Single-task models; crop-then-classify pipelines	Single-stage boxes with sigmoid multi-labels per tooth in one pass
Automatic FDI numbering	Tooth IDs often manual/absent on panoramics	Post-hoc/manual mapping, task-specific labels	Add quadrant/position head + arch-curve mapping for direct FDI output
Speed for chairside use	High accuracy methods too slow	Heavy segmentation/refinement stacks	Real-time one-pass detector with PAN/FPN; tiny on-crop refiner only when needed
Data scale & annotation load	Few large, balanced multi-label OPG sets	Small/private cohorts; imbalance	Active learning on low-confidence regions; mosaic/copy-paste for rare lesions
Generalization across sites/scanners	Performance drops on external data	Single-center training; limited external tests	Strong photometric/geometry augments, domain normalization, held-out site eval, TTA
Sensitivity to small/early lesions	Early/proximal defects often missed	Focus on obvious/pathognomonic signs	High-res tiling + multi-scale heads; focal/varifocal loss; micro-lesion oversampling

3 Result & Discussions

This section covers the findings and discussion of a comparative analysis of object identification algorithms, with a particular emphasis on YOLO-based systems. The figures summarize reported accuracy, inference speed, model size, and architectural complexity from several research to demonstrate the key performance and efficiency trade-offs. We also investigate dataset size, annotation complexity, and data- source dispersion to better understand the experimental settings under which these algorithms are evaluated.

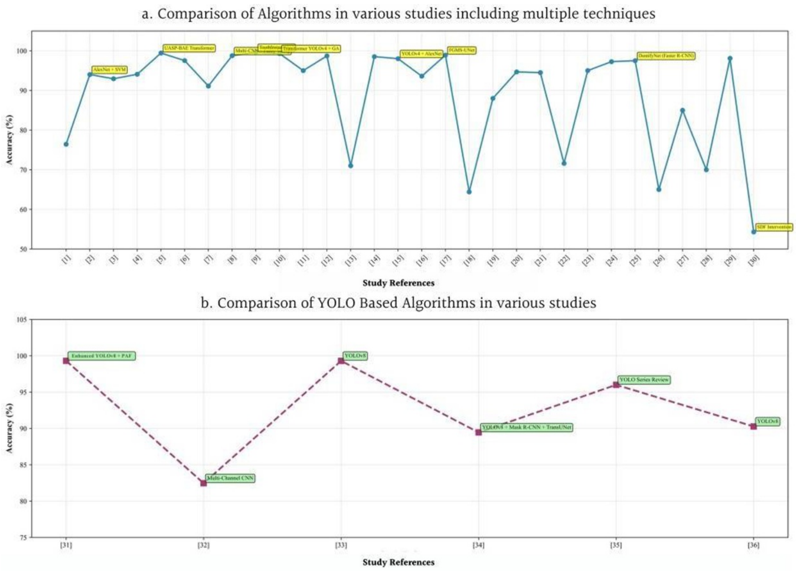

Fig 2. Panoramic radiograph accuracy: (a) diverse methods vs. (b) YOLO variants—compact comparison.

Figure 2(a) indicates that segmentation- and transformer-centric pipelines have sporadic highs with large variance and sensitivity to panoramic X-ray peculiarities (low SNR, metal artifact, overlapping teeth). It is in contrast evident in Figure 2(b) that detector-first YOLO variants maintain ≥ 90% accuracy with a narrower dispersion, which is indicative of better bias-variance trade-offs in domain shift. YOLO has one- stage architecture, which is technically equivalent to multi-scale FPN/PAN aggregation, decoupled cls/reg heads, and the IoU-aware loss (CIoU/GIoU), allowing localization of small and low-contrast lesions and dense packs of teeth with minimal latency compared to pixel-wise decoders. The further optimization of multi-label predictions per image is provided in modern label assignment (e.g., SimOTA) and optimized NMS, whereas streamlined compute enables real-time inference and small memory footprints, which are clinical throughput friendly.

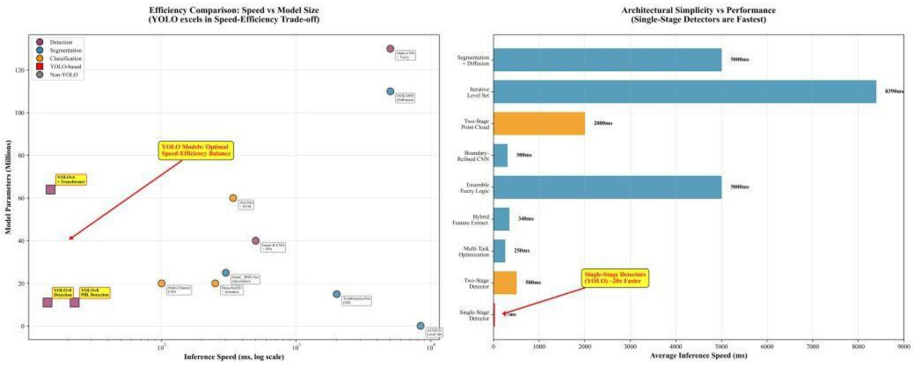

Fig 3. Efficiency–complexity trade-offs in dental imaging models: (a) speed–size scatter places single-stage detectors on the Pareto frontier; (b) architectural runtimes show single-stage detection ≈20× faster than two-stage/segmentation pipelines—enabling real-time panoramic analysis.

Figure 3 (a) contrasts inference speed (log-scale) against model size. One-stage detectors cluster on the Pareto frontier—sub-50 ms latency with ≤ 20 M parameters—while heavier pipelines (e.g., instance segmentation, diffusion refinement, point-cloud two-stage models) incur 10–100× slower runtimes and far larger footprints (60–125 M+ params). This indicates superior compute efficiency and more favorable bias–variance characteristics for clinic-grade deployment on commodity GPUs/CPUs. Figure 3 (b) benchmarks architectural latency. Single-stage detectors (~7 ms) are ~20× faster than two-stage detectors (~500 ms) and orders faster than iterative/ensemble or diffusion-augmented segmenters (2–8.4 s). Simpler detection-first designs therefore maximize throughput for panoramic multi-label triage while avoiding the memory and scheduling overheads of pixel-wise decoders and iterative refinement.

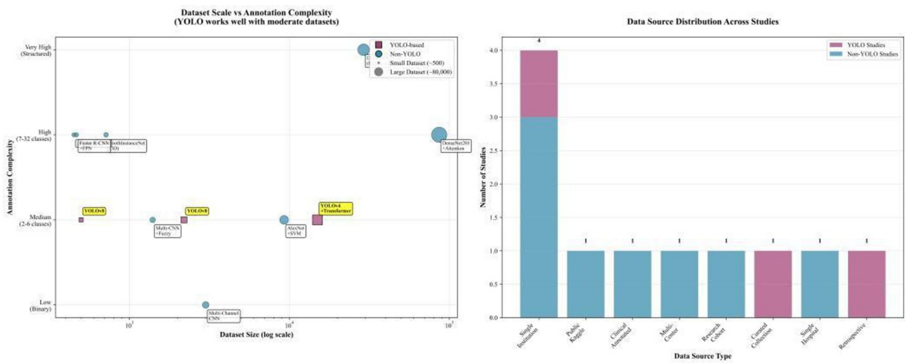

Fig 4. Data constraints in dental AI: (a) size–complexity landscape shows detector-first models remain effective with moderate datasets and class counts; (b) study sources are chiefly single-institution, underscoring the need for multi-center, public benchmarks.

Figure 4 (a) plots the scale of dataset (log) versus complexity of annotation. The single-stage detector tagged points are concentrated in the "moderately data / 2–6 classes regime", and still achieve a useful level of operational efficiency, but the heavier two-stage/segmentation pipelines are concentrated in either very large, highly structured, or small datasets with fragile high-class taxonomies, both a sign of a lack of data efficiency. This indicates that detector-first workflows can support a small amount of curation and heterogeneous labels and be scaled to larger cohorts. In Figure 4 (b) one can see data-source provenance: single-institution cohorts are the most common, and there are relatively few publics, curated, or multi- center sets of data-sets, indicating that there are a generalization bottleneck and an annotation burden mismatch that is well addressed by architectures with high inductive bias.

4 Future Directions

Through the generations, the lineage of YOLO has become increasingly fine-tuned at the triad of detection bias, variance and latency that are most relevant towards panorama radiography. YOLOv3 empowered with (please) dense scene multi-scale heads, anchor-based, developed CSPDarknet + SPP + PANet, new modern augmentations (Mosaic/ CutMix), and IoU-conscience regression, enhanced small-object recall in metal-induced stationary non-photometric. YOLOv5 trained (has been industrialized, works with sparse side-network hyper-parameters, auto-anchor, hyper-parameter evolution) and strongly throughput on trading use to have large-scale commodity GPUs. YOLOv7 (E-ELAN, re-param convs) reused sharpy without growing latency. Most importantly to our task, YOLOv8 uses an anchor-free, decoupled head that has C2f blocks, dynamic label assignment and distribution-focal regression, which provides a thinner localization on crowded teeth and can use multi-hot (sigmoid) labels per-box, which are more natural to per-tooth, multi-pathology tagging. Best empirically, transformer-augmented YOLOv4 with GD tuning achieved almost-ceiling AP in panoramic anomaly detection, and several reports demonstrated that YOLOv8 would make an efficient front-end frontier in staging/measurement downstream and a more prodigal variant, the v8-s, has the best mAP–latency trade-off in the chairside setting.

Using these trajectories and our inference guided by cross-studies (Fig. 2–4), the practical choice of managers/first timers should be YOLOv8 (s/m), the production base, which is anchor-free entirety, decoupled and PAN/FPN- aggregated and trained with heavy scale-conscious augmentation and the size-sensitive loss of IoU, with latent markups optional for outlier auditing or IoU-immune sensitivity. This layout provides (i) real time throughput with relatively small hardware, (ii) better recall of small, low contrast lesions with stride-8 pyramids and (iii) serene implementation of per-tooth ID heads and multi- label logits to generate FDI-consistent data. By assuming the detector with light attention at P4/P5, high-res tiling the proximal lesions with tiling, and mixed-precision/INT8 quantization quantities, it is possible to maintain latency, whilst adversely impacting the robustness of generalization of heteroscedastic noise. Concisely, of the YOLO variants, but most capable of occupying the Pareto frontier of panorama, multi- pathology dental analytics in terms of accuracy, speed, and deployability is YOLOv8.

5 Conclusion

Besides concern accuracy in the headlines, the conclusive standing of YOLOv8 to panoramic odonto-analytics is that it can be suited to trustworthy, operable AI: it can be calibrated to make decisions, and produced reports can be audited, reproducible by all standards. In practice, its logits can be decoupled to support post-hoc conformal risk control through texture–less deep neural networks and can be trained with under a worry about the number of bits per bit to control its overhead through low ECE calibration and its outputs can be serialized in a SNODENT/FDI-friendly JSON schema into whose slots an EHR can drive fluidly (DICOM-SR/FHIR ImagingStudy). The YOLOv8 side is then pragmatically integrated with (i) shift detector (embedding-based density measures) to quarantine out of distribution scans, (ii) test-time augment ensembling to restrict variance on potentially ambiguous proximals, and (iii) active-learning loops to active-label images and separate by labeling expert which gradually resolve annotation debt contributing to steadily better recall rates of down-sample rare lesions. The future evaluation would be conducted by coupling tooth level measurements with decision-cost curve, audits of energy to a basic analysis of clinical and environmental efficiencies, but with the governance mechanisms which include audit trails, bias stratification (age/device/site) and fallbacks to manual review in case any uncertainty thresholds were met. To the point, the choice YOLOv8 is not only an algorithmic decision, but an engineering position: it allows performing the calibrated, monitored and cross-institutional multi-label detection, which can be tested in practical experiments, scaled up, and which can be maintained with legal regulation, which are qualities that cannot only be found in a co-requesting segmentation-based and two-step pipeline.

Conflict of Interest: The authors declare that they have no conflict of interest.

Funding Declaration: This research did not receive any specific grant from funding agencies in the public, commercial or not-for profit sectors.

Ethics Declaration: Not applicable.

Clinical Trials: Not applicable.

Data Availability Statement: Data will be available upon reasonable request.

References

1. Milošević, Denis, Marin Vodanović, Ivan Galić, and Marko Subašić. "A comprehensive exploration of neural networks for forensic analysis of adult single tooth x-ray images." *IEEE access* 10 (2022): 70980-71002.
2. Aljabar, Ahmed Alabd, Zain Raisan, Mohammed Adnan, and Salam Dhou. "A Hybrid Transfer Learning Approach to Teeth Diagnosis Using Orthopantomogram Radiographs." *IEEE Access* (2024).
3. Bussaban, Limpapat, Ekkarat Boonchieng, Wannakamon Panyarak, Arnon Charuakkra, and Pornpattra Chulamanee. "Age Group Classification from Dental Panoramic Radiographs Using Deep Learning Techniques." *IEEE Access* (2024).
4. Jiang, Benxiang, Songze Zhang, Mengyao Shi, Hai-Lin Liu, and Hongjian Shi. "Alternate level set evolutions with controlled switch for tooth segmentation." *IEEE access* 10 (2022): 76563-76572.
5. Zhang, Dong, Jing Yang, Shaoyi Du, Wenqing Bu, and Yu-cheng Guo. "An uncertainty-aware and sex-prior guided biological age estimation from orthopantomogram images." *IEEE Journal of Biomedical and Health Informatics* 27, no. 10 (2023): 4926-4937.
6. Wuersching, Sabina Noreen, Lisa Kohl, Reinhard Hickel, Falk Schwendicke, and Maximilian Kollmuss. "Assessing the marginal seal of bioactive restorative materials in class II cavities with a bacterial penetration model." *Dental Materials* 41, no. 5 (2025): 553-560.
7. Cuzin, Jean-François, Dominique Gaget, Petra Maes, Peter Bottenberg, Bart Vande Vannet, and Karlien Asscherickx. "Assessment of interproximal enamel reduction planned by the digital set-up of a customized lingual orthodontic appliance: A comparison cohort study." *Heliyon* 10, no. 3 (2024).
8. Chen, Shih-Lun, Tsung-Yi Chen, Yi-Cheng Mao, Szu-Yin Lin, Ya-Yun Huang, Chiung-An Chen, Yuan-Jin Lin et al. "Automated detection system based on convolution neural networks for retained root, endodontic treated teeth, and implant recognition on dental panoramic images." *IEEE Sensors Journal* 22, no. 23 (2022): 23293-23306.
9. Van Nistelrooij, Niels, Haline Cunha de Medeiros Maia, Lingyun Cao, Shankeeth Vinayahalingam, Bas Loomans, Maximiliano Sergio Cenci, and Fausto Medeiros Mendes. "Automated detection and numbering of primary and permanent teeth in digital impressions of children using artificial intelligence." *Journal of Dentistry* (2025): 105976.

10. Singh, Sanabam Bineshwor, Anuradha Laishram, Khelchandra Thongam, and Khumanthem Manglem Singh. "Automatic Detection and Classification of Dental Anomalies and Tooth types using Transformer-based Yolo with GA Optimization." IEEE Access (2025).
11. Naufal, Mohammad Farid, Chastine Fatichah, Eha Renwi Astuti, and Ramadhan Hardani Putra. "Deep learning for mandibular Canal segmentation in digital dental radiographs: A systematic literature review." *IEEE Access* 12 (2024): 76794-76815.
12. Garine, Srija, Papolu Hema Janardhana, Lakkireddy Lakshmi Prabhanjan Reddy, Penubothu Jagadeesh Venkata Kumar, and Chapa Gagan Dwaz. "Deep learning-based cavity detection in diverse intraoral images: a web-based tool for accessible dental care." *Procedia Computer Science* 233 (2024): 882-891.
13. Jones, Bree, Mathias Lambach, Tong Chen, Stavroula Michou, Nicky Kilpatrick, Nigel Curtis, David P. Burgner, Christoph Vannahme, and Mihiri Silva. "Dental caries detection in children using intraoral scans and deep learning." *Journal of Dentistry* (2025): 105906.
14. Neeraja, R., and L. Jani Anbarasi. "DentC_BSFL Net: Boundary Refined and Spatial Adaptive Feature Learning Network for Enhanced Dental Caries Segmentation." *IEEE Access* (2025).
15. Li, Kuo-Chen, Yi-Cheng Mao, Mu-Feng Lin, Yi-Qian Li, Chiung-An Chen, Tsung-Yi Chen, and Patricia Angela R. Abu. "Detection of tooth position by YOLOv4 and various dental problems based on CNN with bitewing radiograph." *IEEE Access* 12 (2024): 11822-11835.
16. Wang, Xianyun, Linhong Wang, Zhenchen Yang, Jiacong Zhou, Yuchen Zheng, Feng Chen, Richang Hong, Jun Yu, and Fan Yang. "DSIS-DPR: Structured Instance Segmentation and Diffusion Prior Refinement for Dental Anatomy Learning." *IEEE Transactions on Multimedia* 26 (2024): 9464 9476.
17. Jia, Qihang, Xiangfu Ding, Na Tian, Youjia Shao, Xinbo Yu, and Wencang Zhao. "Efficient panoramic dental X-ray image segmentation model based on dual-domain transformation processing." *Biomedical Signal Processing and Control* 111 (2026): 108288.
18. Peskersoy, Cem, Aybeniz Oguzhan, Cagri Akcay, Beyza A. Dincturk, Hulya SE Can, Erdinc K. Kamer, and Mehmet Haciyanli. "Evaluation of oral health status and immunological parameters of hospitalized COVID-19 patients during acute and recovery phases: A randomized clinical trial." *Journal of Dental Sciences* 19, no. 3 (2024): 1515-1524.
19. Klein, Vanessa, Martha Büttner, Gerd Göstemeyer, Sarina Rolle, Antonin Tichy, Falk Schwendicke, and Noah F. Nordblom. "From inconsistent annotations to ground truth: Aggregation strategies for annotations of proximal carious lesions in dental imagery." Journal of Dentistry 157 (2025): 105728.
20. Wankhade, Dipali, Shailesh Gahane, and Mrunal Meshram. "Hybrid Spectral-Attentive Network for Enhanced Raman-based Oral Cancer Detection." In *2025 8th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 803-809. IEEE, 2025.
21. Pavarut, Srisopitsawat, Wongsakorn Preedanan, Itsuo Kumazawa, Kenji Suzuki, Masaki Kobayashi, Hajime Tanaka, Junichiro Ishioka, Yoh Matsuoka, and Yasuhisa Fuji. "Improving kidney tumor classification with multi-modal medical images recovered partially by conditional CycleGAN." *IEEE Access* 11 (2023): 146250-146261.
22. Razaghi, Marzieh, Hossein Ebrahimpour Komleh, Fereshteh Dehghani, and Zahra Shahidi. "Innovative diagnosis of dental diseases using YOLO V8 deep learning model." In *2024 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP)*, pp. 1-5. IEEE, 2024.
23. Pal, Debashish, Kriti Sen Sharma, and Jiang Hsieh. "Metal artifact correction algorithm for CT." In *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*, pp. 1-4. IEEE, 2013.
24. Ghafoor, Afnan, Seong-Yong Moon, and Bumshik Lee. "Multiclass Segmentation Using Teeth Attention Modules for Dental X-Ray Images." *IEEE Access* 11 (2023): 123891-123903.
25. Özçelik, Salih Taha Alperen, Hüseyin Üzen, Abdulkadir Sengur, Muammer Türkoğlu, Adalet Çelebi, and Nebras M. Sobahi. "Optimized Multi-Scale Detection and Numbering of Teeth in Panoramic Radiographs Using DentifyNet." *IEEE Access* (2025).
26. Sarkar, Moumita, and Tapas Kumar Bandyopadhyay. "Patenting trends in biodegradable implant: An analysis." *World Patent Information* 82 (2025): 102378.

27. Nagar, Parul, Jefferson Prince, N. Srikant, B. V. Murlimanju, Ritu Gupta, Shravan Shetty, and Dharnappa Poojary. "Prevalence of Anterior tooth traits among a sample of the coastal population of South India and comparative review of prevalence among the Eurasian population." *Translational Research in Anatomy* (2025): 100397.
28. Esposito, Rosa, Francesco Masedu, Marco Cicciù, Michele Tepedino, Martina Denaro, and Domenico Ciavarella. "Reliabilty of recording occlusal contacts by using intraoral scanner and articulating paper-A prospective study." *Journal of Dentistry* 142 (2024): 104872.
29. Almalki, Amani, Abdulrahman Almalki, and Longin Jan Latecki. "Self-Supervised Masked Deep Embeddings for Dental Caries Detection." *IEEE Access* (2025).
30. Zaheer, Khaleda, Tahmina Zerin, James Coughlan, Shuvashis Saha, Elsa Delgado-Angulo, and Elham Kateeb. "Supervised Toothbrushing and Silver Diamine Fluoride Application of Rohingya Children in a Refugee Camp in Bangladesh." *International dental journal* 75, no. 5 (2025): 100929.
31. Abed, Dhiaa Mohammed, Shuzlina Abdul-Rahman, and Sofianita Mutalib. "Dental segmentation via enhanced YOLOv8 and image processing techniques." *Mesopotamian Journal of CyberSecurity* 4, no. 3 (2024): 189-202.
32. Fang, Emma Peng, Di-Jie Liew, Yung-Chun Chang, and Chih-Yuan Fang. "Enhancing wisdom teeth detection in panoramic radiographs using multi-channel convolutional neural network with clinical knowledge." Computers in Biology and Medicine 192 (2025): 110368.
33. Zirek, Taha, Muhammet Üsame Öziç, and Melek Tassoker. "AI-Driven localization of all impacted teeth and prediction of winter angulation for third molars on panoramic radiographs: Clinical user interface design." *Computers in Biology and Medicine* 178 (2024): 108755.
34. Xue, Ting, Lei Chen, and Qinfeng Sun. "Deep learning method to automatically diagnose periodontal bone loss and periodontitis stage in dental panoramic radiograph." *Journal of Dentistry* 150 (2024): 105373.
35. Ragab, Mohammed Gamal, Said Jadid Abdulkadir, Amgad Muneer, Alawi Alqushaibi, Ebrahim Hamid Sumiea, Rizwan Qureshi, Safwan Mahmood Al-Selwi, and Hitham Alhussian. "A comprehensive systematic review of YOLO for medical object detection (2018 to 2023)." *IEEE Access* 12 (2024): 57815-57836.
36. Putra, Ramadhan Hardani, Eha Renwi Astuti, I. Komang Evan Wijaksana, Arna Fariza, Ratri Maya Sitalaksmi, and Nobuhiro Yoda. "Automated Periodontal Bone Loss Detection on Panoramic Radiographs Using You Only Look Once v8 (YOLOv8): A Retrospective AI Approach." *Journal of International Oral Health* 17, no. 3 (2025): 203-211.