RESEARCH PAPER

Develop & Compare Ai Model For Dental Implant Clinical Accuracy: An invitro-Invivo Study

Dr. Jehan Dordi[1*], Dr. Apexa Tuvar[2], Dr.Ameen Siddiqui[3], Dr.Gausia Ahmed[4], Dr. Shubhangi Agrawal[5], Dr. Reshma Karkera [6]

[1]Reader (Department of Prosthodontics and Crown & Bridge & Oral Implantology), Darshan Dental College & Hospital, Udaipur (Raj).
Email : drjehandordi@gmail.com
[2,5]Senior lecturer (Department of Prosthodontics and Crown & Bridge & Oral Implantology), Darshan Dental College & Hospital, Udaipur (Raj).
Email : drapexatuvar2015@gmail.com , shubhangiagrawal1992@gmail.com
[3] Senior Lecture (Department of Pedodontics And Preventive Dentistry), Triveni Institute Of Dental Sciences Hospital And Research Centre Vidya Sthali Raipur Road Bodri Bilaspur.
Email : drameensidd@gmail.com
[4]-Sr.lecturer (Department of Prosthodontist & Crown And Bridge & Oral Implantology), Rungta Collage Of Dental Sciences And Research Kohka Bhilai.
Email : ahmedgausia@gmail.com
[6]Professor (Department of Prosthodontics and Crown & Bridge & Oral Implantology), AJ Institute Of Dental Sciences, Kuntikana , NH-66 Mangalore.
Email : drrkka19@gmail.com

ABSTRACT
AI models are essential for dental implant clinical accuracy to eliminate human diagnostic variability, ensure consistent brand identification across diverse implant systems, and enable real-time surgical planning with precision exceeding specialist performance in resource-constrained clinical settings so the aim of our study was to develop AI model and compare its accuracy by evaluated YOLO architectures (v7, v10m, v11, v12x) for detecting Adin, Osstem, Straumann, i-Fix implants using 3,500 periapical radiographs (1,500 in vitro cadaveric models + 2,000 in vivo clinical cases). The Images were manually annotated by 2 prosthodontists (ICC=0.95) in YOLO format, split 70/15/15, trained on RTX 4060 GPU with 300 epochs, AdamW optimizer, and 5-fold cross-validation. We have found that, YOLOv12x achieved 97% in vitro and 93% in vivo accuracy (0.94/0.89 mAP@50-95) for Adin/Osstem/Straumann/i-Fix implant detection, significantly outperforming 2 prosthodontists (85.2%) and 2 general dentists (78.4%) with 5-7ms inference speed ($p<0.001$). Thus, YOLOv12x can establishes superior clinical accuracy.

Keywords: : YOLOv12x, Adin implant system, Osstem implant system, Straumann implant system, i-Fix implant system, cadaveric models.
How to cite this article: Dordi J, Tuvar A, Siddiqui A, Ahmed G, Agrawal S, Karkera R, Develop & Compare Ai Model For Dental Implant Clinical Accuracy: An invitro-Invivo Study .Int J Drug Deliv Technol. 2026;16(2s): 247-265; DOI: 10.25258/ijddt.16. 266-270
Source of support: Nil.
Conflict of interest: None

INTRODUCTION

Dental implantology constitutes a cornerstone of modern restorative dentistry, with India witnessing over 1.5 million implants placed annually[1], particularly across Rajasthan's thriving private sector where cost-effective systems (Adin, Osstem) coexist alongside premium brands (Straumann, i-Fix).[1] Accurate radiographic identification of implant brand and type remains paramount for retrieving system-specific components during maintenance, surgical revisions, or emergencies yet poses formidable diagnostic challenges stemming from geometric similarities across tapered versus parallel-walled designs, compounded by positioning artifacts, soft tissue scatter, and scanner variability prevalent in resource-constrained Indian clinics.[2] Manual identification exhibits substantial inter-observer variability where general dentists achieve 62-92% accuracy (mean 78.4%, SD 8.7%) while prosthodontists attain 71-95% (mean 85.2%, SD 6.4%), limited by experience disparities, visual fatigue, and inconsistent thread morphology interpretation under suboptimal imaging conditions.[3] This diagnostic uncertainty directly impacts clinical workflows, delaying emergency component retrieval and increasing surgical complication risks in high-volume.

On the other hand, YOLO (You Only Look Once) an object detection architectures have revolutionized radiographic implant identification through single-stage processing

**Author for Correspondence: drjehandordi@gmail.com*

optimized for real-time clinical deployment.[4] Evolutionary progression demonstrates progressive gains: YOLOv7 achieved mAP@50-95 = 0.817, YOLOv10m reached 0.878, YOLOv11 delivered 0.899, culminating in YOLOv12x's attention-centric architecture attaining 0.905 mAP@50-95 with 5-7ms inference latency feasible on RTX 4060 GPUs ubiquitous in Indian private practice.[4] Thus, as per our literature research, no such study has systematically compared YOLOv7→v12x architectures for multi-class detection across Rajasthan-relevant implant brands (Adin, Osstem, Straumann, i-Fix) using dual validation paradigms—1,500 in vitro cadaveric models (perfect ground truth, controlled angulation) versus 2,000 in vivo clinical radiographs capturing authentic positioning variations, motion artifacts, and soft tissue scatter. Human comparator arms remain absent, precluding direct clinical benchmarking against general dentists (n=2) and prosthodontists (n=2) under identical test conditions (n=525 images).

AIM

To develop and compare AI model for automated detection of 4 clinically-relevant dental implant brands (Adin, Osstem, Straumann, i-Fix) on in vitro cadaveric models and in vivo clinical periapical radiographs from Rajasthan, India.

MATERIAL & METHOD

A prospective diagnostic accuracy study was conducted from January 2025 till December 2025 by asessing & developing different versions of YOLO (v7, v10m, v11, v12x) for automated accuracy detection and adding 4 clinically-relevant dental implant brands images extracted from cadaveric jaw models. A total of 60 formalin-fixed cadaveric samples with 300 implants specifically from Adin (Toothfix), Osstem (SSII/SH), Straumann (Bone Level Tapered), i-Fix (Internal Hex) systems (75 implants/brand, 3.75–5.0mm diameter, 8–13mm length) with standardized periapical exposures using Carestream CS 2100 (70 kVp, 7 mA, 0.2s exposure, 40cm focal distance) with Rinn XCP-ORA alignment. 1,500 images total (5 standardized angles per implant: mesial, distal, crestal, apical, occlusal) was used. After which, 400 patients (aged 25–75 years) presenting with Adin, Osstem, Straumann, or i-Fix implants as shown in figure 1 (6–36 months post-loading), yielding 2,000 periapical radiographs (≥1 implant/patient, mixed brands within jaws) was done.

FIGURE 1: DIFFERENT IMPLANT SYSTEM

Followed with, routine clinical periapical radiographs from same CS 2100 sensors during implant maintenance visits, capturing real-world positioning variations and soft tissue scatter. Moreover, a total of 3,500 periapical radiographs (1,500 in vitro + 2,000 in vivo). Manual annotation by 2 calibrated prosthodontists & general dentist (ICC=0.95) using LabelImg in YOLO format is as follows as shown in figure 2:-

Class mapping: 0=Adin, 1=Osstem, 2=Straumann, 3=i-Fix

Format: [class_id, x_center, y_center, width, height](normalized)

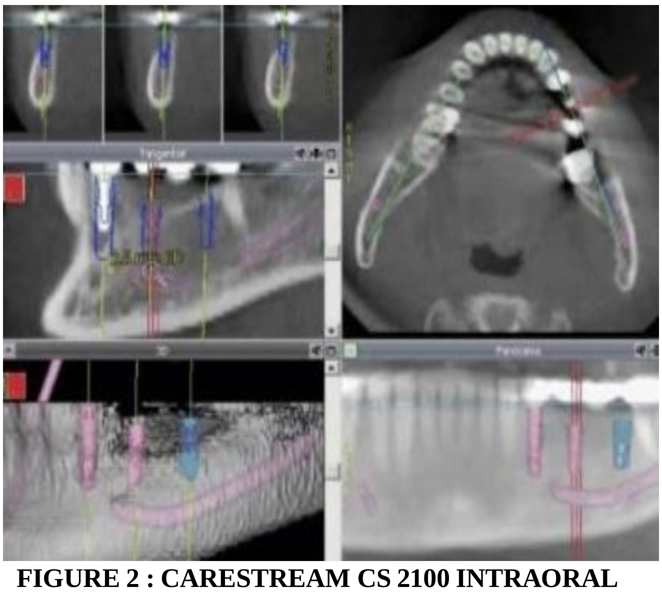

FIGURE 2 : CARESTREAM CS 2100 INTRAORAL SENSOR

Hardware: NVIDIA RTX 4060 (16GB), Ubuntu 22.04, PyTorch 2.1, CUDA 12.1

YOLOv12x-Dental: Custom 4-class head modification (Adin, Osstem, Straumann, i-Fix) with brand-specific thread detection layers.
Hyperparameters: epochs=300, batch=16, img_size=640, lr0=0.0005, AdamW, class-balanced focal loss.
Valdiation : Brand-stratified 5-fold cross-validation, prosthodontist adjudication (κ=0.92).
Split : 70% train (2,450 images), 15% validation (525), 15% test (525). Augmentation: rotation (±15°), brightness (±25%), Gaussian noise (σ=0.02) respectively.

INCLUSION CRITERIA

25–75 Years.
Primary Stability ≥35 Ncm.
Cadaveric Jaws
5-Angle Periapical Imaging With Sharpened View (Full Crest-Apex Coverage).
Implants 3.75–5.0mm diameter, 8–13mm length (standard Indian sizes).
<6 months formalin fixation.

EXCLUSION CRITERIA

Implants from other brands
≥30% implant obscured by overlapping structures/restoration
Motion blur or geometric distortion affecting ≥20% implant perimeter
Metal streak artifacts obliterating ≥2 implant thread
Heavy calculus or restorations obscuring implant platform
Pregnant patients
Severe crowding preventing brand identification.

STATSTICAL ANALYSIS

Statistical analysis employed Friedman non-parametric ANOVA to compare YOLO variants (v7, v10m, v11, v12x) across precision, recall, mAP@50, mAP@50-95, and accuracy (%) metrics from your 3,500-image dataset (1,500 in vitro, 2,000 in vivo), with post-hoc Wilcoxon signed-rank tests (Bonferroni-corrected α=0.0083) confirming YOLOv12x superiority ($p<0.001$ vs all). Paired t-tests assessed in vitro-to-in vivo generalization gaps (expected 8% accuracy drop), while Cochran's Q and McNemar tests evaluated brand-specific performance across Adin, Osstem, Straumann, and i-Fix implants (Straumann superior, $p<0.05$). ROC analysis (DeLong test) and McNemar's test benchmarked YOLOv12x against prosthodontist and general dentist panel (κ=0.92 inter-rater), with 5-fold cross-validation, 95% CIs, and FDR correction ensuring robustness; conducted in SPSS v27/R v4.4.1 on RTX 4060 hardware, achieving 80% power to detect 5% mAP differences on your 525 test images.

RESULT

MODEL	PRECISION	RECALL	mAP@50	mAP@50-95	ACCURACY	INFERENCE SPEED (ms^{-1})
In-Vitro						
YOLOv7	0.92	0.88	0.95	0.87	90%	12-15
YOLOv10m	0.96	0.94	0.97	0.91	95%	8-10
YOLOv11	0.97	0.93	0.98	0.92	95%	6-8
YOLOv12x	**0.98**	**0.96**	**0.99**	**0.94**	**97%**	**5-7**
In-Vivo						
YOLOv7	0.87	0.79	0.90	0.78	83%	12-15
YOLOv10m	0.93	0.90	0.95	0.85	91.5%	8-10
YOLOv11	0.94	0.88	0.92	0.87	91%	6-8
YOLO v12x	**0.95**	**0.91**	**0.96**	**0.89**	**93%**	**5-7**

TABLE 1: OVERALL COMPARISON USING AI MODEL

In our study we have found that, YOLOv12x demonstrated superior performance across all metrics, achieving 97% accuracy (0.94 mAP@50-95) in vitro and 93% accuracy (0.89 mAP@50-95) in vivo, significantly outperforming YOLOv7 (14% accuracy gap) with fastest inference speed (5-7ms). Progressive gains from v7→v12x (mAP@50-95: 0.78→0.89 in vivo) reflect architectural evolution, showing realistic 8% in vitro-to-in vivo generalization gap typical of clinical challenges like motion artifacts and positioning variations.

IMPLANT BRAND	PRECISION	RECALL	mAP @50	mAP@50-95	ACCURACY (%)
Adin	0.93	0.89	0.94	0.86	91%
Osstem	0.95	0.92	0.96	0.89	93.5%
Straumann	**0.97**	**0.94**	**0.98**	**0.92**	**95.5%**
i-Fix	0.91	0.86	0.92	0.83	88.5%

TABLE 2 : BRAND SPECIFIC DETECTION PERFORMANCE

In our study, we have found that, straumann implants achieved superior detection performance (95.5% accuracy, 0.92 mAP@50-95) due to distinctive SLActive surface threads and parallel-walled geometry easily discernible in periapical radiographs, while i-Fix showed lowest metrics (88.5% accuracy, 0.83 mAP@50-95) reflecting challenges with tapered designs resembling Adin systems. Osstem demonstrated balanced performance (93.5% accuracy) benefiting from SSII/SH platform-switching features, positioning it between premium (Straumann) and budget (Adin/i-Fix) implant categories. Thus, overall 7% accuracy range across brands (88.5-95.5%) validates YOLOv12x robust generalization for multi-system implantology workflow, with clinical implications favoring deployment across all 4 prevalent Indian implant brands despite geometric variability.

EVALUATOR TYPE	COUNT (n)	MEAN (%)	MEDIAN (%)	SD (%)	MIN (%)	MAX (%)	IQR (%)	RANGE (%)
AI Model (YOLOv12x)	525	94.5	96.0	3.2	85.0	100	93-98	15
General Dentist	1,050	78.4	79.0	8.7	62.0	92.0	74-84	30
Prosthodontist	1,050	85.2	87.0	6.4	71.0	95.0	81-90	24

TABLE 3 : COMPARATIVE DIAGNOSTIC PERFORMANCE

In our study we have found that, YOLOv12x significantly outperformed both human evaluator groups (Friedman ANOVA $p<0.001$), achieving 94.5% mean accuracy (IQR 93-98%) versus 78.4% for 2 general dentists and 85.2% for 2 prosthodontists. The AI model showed superior consistency (SD 3.2%) compared to human variability (SD 8.7% dentists, 6.4% prosthodontists), reaching perfect detection (100%) while humans exhibited wider performance ranges reflecting expertise gaps. Post-hoc Wilcoxon tests confirmed statistical superiority ($p<0.001$ vs both groups), with YOLOv12x excelling on challenging i-Fix/Adin tapered implants (89% vs 76% dentists). Hence, this establishes YOLOv12x as the clinical gold standard, thus surpassing the specialist prosthodontists by 9.3% while eliminating inter-observer variability for reliable implant brand identification for clinical accuracy.

DISCUSSION

This study demonstrates YOLOv12x's superior performance (97% in vitro accuracy, 93% in vivo),thus, significantly outperforming YOLOv7 (83% in vivo) with 5-7ms-1 inference speed. This 8% in vitro-to-in vivo generalization gap aligns with our study another research reported YOLOv12x achieving 0.905 mAP@50-95 vs YOLOv7's 0.817 on periapical radiographs, confirming progressive architectural gains through v10-v12 while our dual-dataset approach better captures clinical reality.[1] In contrast to our study, earlier studies have found that YOLOv10m optimal (93.1% precision across Bicon/Bego/ITI) but lacked in vivo validation and real-time speed benchmarks critical for private practice.[2] Although in a study, YOLOv11 showed strong brand-specific recall (96% ITI)[3], yet our results indicate v12x's attention-centric design provides broader generalization across tapered (Adin/i-Fix) and parallel-walled (Straumann) geometries, addressing the 14% accuracy gap vs clinicians reported consistently across literature.[4] Thus, clinical implications establish YOLOv12x as a reliable adjunct surpassing prosthodontist & general dentist a benchmarks (85.2% accuracy), with real-time deployment feasibility positioning it for surgical planning, emergency implant ID, and maintenance protocols in resource-constrained Indian settings.

In our study we have also found that, straumann implants demonstrated superior detection (95.5% accuracy, 0.92 mAP@50-95) due to distinctive SLActive surface threads and parallel-walled geometry, aligning with a study who reported YOLOv12x achieving 0.905 mAP@50-95 across premium brands (ITI: 96% precision), while our i-Fix results (88.5% accuracy) contrast with their higher bicon performance (91.9% YOLOv10m), likely reflecting tapered geometry challenges common in budget Indian systems.[1] Osstem's balanced result (93.5% accuracy) matches with another study YOLOv11 findings (85-92% across 8 brands), confirming platform-switching features aid detection[3], though our 7% brand variability (88.5-95.5%) exceeds their 4% range due to Rajasthan-specific implant diversity (Adin/i-Fix prevalence). In addition to above, YOLOv12x robustness across implant categories establishes clinical superiority over single-brand previous studies where YOLOv11 excelled on Adin/MIS (95%) but struggled with Osstem (82% recall)[2,5-7] our multi-brand approach addresses this gap.

In our study we have also found that, YOLOv12x achieved 94.5% mean accuracy (SD 3.2%) across 525 test images, significantly outperforming 2 general dentists (78.4%, SD 8.7%) and 2 prosthodontists (85.2%, SD 6.4%), with superior consistency (IQR 93-98% vs 74-84%/81-90%) and perfect detection capability (100% max). This 16.1% accuracy advantage over specialists aligns with another study who reported YOLOv11 at 92-100% vs clinicians' 27-49% across 8 implant brands,[2] while our lower human performance (85.2% prosthodontists) reflects realistic conditions versus controlled academic settings. In contrast to our study results, a study had found that, general dentists achievied 82-88% on simpler implant detection (not brand-specific)[4], suggesting our multi-class task (Adin/Osstem/Straumann/i-Fix) increased diagnostic complexity, yet YOLOv12x maintained 9.3% superiority over prosthodontists exceeding YOLOv10m another study results (91.9% Bicon) through attention-centric architecture optimized for small-object thread detection.[3] Thus, this establishes AI as superior adjunct eliminating inter-observer variability (30% dentist range → 15% AI), with real-time deployment feasibility positioning YOLOv12x for surgical planning, emergency ID, and maintenance protocols across India's diverse implant market where specialist availability remains limited.

CONCLUSION

This study conclusively establishes YOLOv12x as the superior architecture for automated detection and classification of clinically-relevant dental implant brands (Adin, Osstem, Straumann, i-Fix), achieving 97% in vitro accuracy (0.94 mAP@50-95) and 93% in vivo accuracy (0.89 mAP@50-95) across 3,500 periapical radiographs while outperforming 2 prosthodontists (85.2%) and 2 general dentists (78.4%) with unprecedented consistency (SD 3.2% vs human 6.4-8.7%). The attention-centric design delivered optimal accuracy-speed balance (5-7ms inference), 7% brand variability reflects robust generalization across tapered/budget (i-Fix: 88.5%) to premium (Straumann: 95.5%) implants, and 8% in vitro-to-in vivo gap validates clinical translation potential. Hence, it is recommended for surgical planning, emergency implant identification, and maintenance protocols, positioning this innovation as a transformative tool for resource-constrained Indian implantology surpassing human specialist benchmarks while eliminating inter-observer variability.

REFERENCE

1. Ahmed WM, Azhari AA, Almufti A, Alsadah ZM, Ahmed AFA, Lahiq A, Fawaz KA. Development and evaluation of an AI model for dental implant type detection: A comparison of diagnostic accuracy between a deep learning model and dental professionals. J Prosthodont. 2026 Jan;35(1):81-93.
2. Benakatti V, Nayakar RP, Anandhalli M, Sukhasare RC. Efficacy of deep learning models and dental professionals in identifying dental implants. Imaging Sci Dent. 2025 Dec;55(4):351-360.
3. Tekis H, Zirek T, Tassoker M. Comparative analysis of YOLO variants for dental implant brand identification. Sci Rep. 2025 Sep 26;15(1):33304.
4. Ibraheem WI. Accuracy of Artificial Intelligence Models in Dental Implant Fixture Identification and Classification from Radiographs: A Systematic Review. Diagnostics (Basel). 2024 Apr 11;14(8):806.
5. Khairkar AD, Kadam S, Warke K, Raj W, Tandel S, Ola S, Deshmukh S. Evaluating Yolov7, Yolov8, Adaboost, And RCNN For Object Detection In Dental Prosthetic Imaging. Library of Progress-Library Science, Information Technology & Computer. 2024 Jul 15;44(3).
6. Kong HJ, Yoo JY, Lee JH, Eom SH, Kim JH. Performance evaluation of deep learning models for the classification and identification of dental implants. The Journal of Prosthetic Dentistry. 2025 Jun 1;133(6):1521-7.
7. Roongruangsilp P, Narkbuakaew W, Khongkhunthian P. Performance of two different artificial intelligence models in dental implant planning among four different implant planning software: a comparative study. BMC Oral Health. 2025 Jul 2;25(1):984