Review Article

The Algorithmic Smile: Multimodal AI and the Transformation of Orthodontic Science

Mayank Shrivastava, DDS[1], Liang Ye, DDS PhD[2]*

[1]Adams School of Dentistry, University of North Carolina at Chapel Hill, NC, USA, 27599
[2]Miller School of Medicine, University of Miami, FL, USA, 33136

*Corresponding Author
Liang Ye
Miller School of Medicine, University of Miami, FL, USA, 33136

Abstract: Orthodontics is undergoing a fundamental transition from a reactive, intuition-based specialty to a proactive, data-driven discipline. The integration of deep learning (DL), computer vision, large language models (LLMs), and physics-informed biomechanical simulations has catalyzed a new era of precision medicine. Between 2024 and 2026, the field has progressed from isolated diagnostic tools to multimodal, agentic AI ecosystems capable of orchestrating clinical workflows. Breakthroughs include automated three-dimensional (3D) cephalometric landmarking with sub–1.2 mm mean radial error, transformer-based multimodal extraction prediction models achieving AUCs above 0.94, physics-informed neural networks (PINNs) for aligner staging, and generative diffusion systems that dynamically simulate soft-tissue response. This review synthesizes recent advances in orthodontic AI, emphasizing the rise of multimodal foundation models integrating skeletal, dental, soft-tissue, and emerging genomic datasets. We further examine agentic AI for remote monitoring, federated learning frameworks for data sovereignty, explainable AI (XAI) to mitigate the "black box" dilemma, and the evolving regulatory landscape. Rather than replacing clinicians, AI augments decision-making capacity, positioning orthodontists as "pilot-in-command" of increasingly autonomous digital systems.

Keywords: Deep Learning (DL), Computer Vision, Large Language Models (LLMs).

1. INTRODUCTION: FROM THE DIGITAL REVOLUTION TO THE INTELLIGENCE REVOLUTION

For over a century, orthodontic diagnosis relied on manual tracing of two-dimensional (2D) radiographs and heuristic-based interpretation. Landmark identification on lateral cephalograms—subject to inter- and intra-operator variability—formed the foundation of skeletal classification and treatment planning. The digital revolution of the early 2000s transitioned film to pixels, but decision-making remained largely human-centered.

By 2026, orthodontics has entered what may be termed the "Intelligence Revolution." Artificial Intelligence (AI), particularly deep learning, no longer serves merely as a computational assistant. Instead, AI systems increasingly operate as agentic collaborators, autonomously performing image segmentation, risk stratification, progress tracking, and even patient communication. Reviews such as Kazimierczak *et al.*, (2024) in the Journal of Clinical Medicine and subsequent analyses have chronicled this transformation.

The convergence of high-resolution intraoral scanning (IOS), cone beam computed tomography

(CBCT), facial photogrammetry, and longitudinal treatment data has created the "big data substrate" required for convolutional neural networks (CNNs), Vision Transformers (ViTs), and multimodal large models to reach clinical parity with expert orthodontists. This evolution parallels developments described in IEEE Transactions on Medical Imaging for automated segmentation and in Nature Biomedical Engineering for generative craniofacial modeling.

The orthodontic clinic of 2026 is thus not defined by digital tools alone, but by intelligent systems that perceive, predict, and adapt.

2. Automated Diagnostics: Beyond Human Perception
2.1 High-Fidelity Cephalometric Tracing
Automated landmark detection remains the cornerstone of orthodontic AI. Traditional CNN architectures (e.g., U-Net variants) achieved promising but inconsistent accuracy. The emergence of Vision Transformers—leveraging self-attention mechanisms to model global spatial dependencies—has dramatically improved performance.

Shodiyeva (2026) reported transformer-based systems identifying over 150 craniofacial landmarks on lateral cephalograms with a mean radial error (MRE) of 1.18 ± 0.12 mm, approaching expert-level reliability. Comparative analyses in Diagnostics (Koz & Uslu-Akcam, 2025) confirm reduced variability compared with manual tracings.

The ViT advantage lies in its capacity to encode long-range relationships—such as Sella–Menton orientation—without the locality constraints inherent to convolutional filters. This holistic contextualization enables robust detection even in cases with atypical morphology or orthognathic surgical alterations.

2.2 Three-Dimensional Segmentation
Orthodontics has increasingly transitioned from 2D cephalometry to 3D volumetric analysis. Automated tooth and skeletal segmentation from CBCT, once a labor-intensive task requiring hours of manual "masking," can now be completed in under 30 seconds using 3D U-Net derivatives and transformer hybrids, as reported in IEEE Transactions on Medical Imaging (2024).

Applications Include:
- Precise root morphology assessment
- Cortical bone boundary mapping
- Detection of root resorption (Dentomaxillofac. Radiol., 2025)
- Customized bracket and appliance fabrication (J. Orofac. Orthop., 2025)

These tools reduce clinician workload while enhancing anatomical precision.

2.3 Skeletal Maturation and Growth Forecasting
AI-driven skeletal age prediction from hand–wrist radiographs demonstrates high concordance with expert evaluation (Eur. J. Orthod., 2025). More significantly, generative models described in Nature Biomedical Engineering (2025) predict craniofacial growth trajectories using longitudinal datasets, marking a shift from static diagnosis to dynamic forecasting.

Orthodontics is thus moving toward anticipatory intervention rather than reactive correction.

3. Multimodal Fusion in Clinical Decision-Making
3.1 Extraction Decisions
Premolar extraction remains one of the most controversial orthodontic decisions. Historically dependent on cephalometric norms and clinician intuition, it is now informed by multimodal AI models integrating:
- Intraoral scans (crowding, arch form)
- Lateral cephalograms (incisor inclination, skeletal base)
- 3D facial photographs (soft-tissue profile)

A 2026 study (PMC) reported AUC values of 0.94 for multimodal transformer-based systems, outperforming single-modality networks. By fusing heterogeneous data streams, AI captures interactions invisible to isolated imaging modalities.

3.2 Orthognathic and Complex Cases
Narrative reviews in Journal of Oral and Maxillofacial Surgery (Taylor & Francis, 2023) emphasize AI's growing role in orthognathic planning, including surgical simulation and postoperative stability prediction. In cleft lip and palate management, AI-guided segmentation and growth prediction models (Orthod. Craniofac. Res., 2024) facilitate interdisciplinary planning.

4. Predictive Biomechanics and Clear Aligner Therapy
Clear Aligner Therapy (CAT) represents the most fertile testing ground for AI innovation.

4.1 Physics-Informed Neural Networks (PINNs)
Traditional aligner staging relies on geometric interpolation between initial and target tooth positions. However, tooth movement is governed by complex biomechanical interactions involving periodontal ligament (PDL) stress distribution and bone remodeling.

Yang & Cheng (2025) introduced force-driven PINNs incorporating physical constraints into neural network training. These systems predict lag—the discrepancy between aligner geometry and biological tooth response—and adjust future staging accordingly.

This physics–data hybrid approach enhances biological plausibility and reduces iatrogenic risk.

4.2 Root-Aware and Bone-Constrained Planning

By integrating CBCT-derived root morphology, AI systems ensure movements remain within cortical boundaries, minimizing risks of dehiscence or fenestration. Such biologically informed algorithms exemplify the shift from shape manipulation to tissue-aware treatment.

4.3 Remote Monitoring and Agentic Systems

Smartphone-based platforms such as DentalMonitoring generate weekly image streams analyzed by AI agents. These systems detect:
- Attachment loss
- Inadequate aligner seating
- Oral hygiene deterioration

Frontiers (2025) describes this transition toward dynamic orthodontic management. Agentic AI now initiates alerts, proposes staging modifications, and triages patient concerns—reducing in-office visits while preserving oversight.

5. Generative AI and the Patient Experience

Generative adversarial networks (GANs) and diffusion models have transformed orthodontic visualization.

5.1 Dynamic Soft-Tissue Simulation

Earlier smile simulations appeared artificial and static. By 2026, diffusion-based systems generate photorealistic, temporally dynamic facial predictions, incorporating lip support, muscular adaptation, and aging parameters. Reviews in Multimodal Technologies and Interaction (2026) describe "face-driven orthodontics" integrating hard- and soft-tissue prediction.

Such simulations enhance informed consent and align patient expectations with biological realities.

5.2 LLM-Driven Communication

Large language models fine-tuned on orthodontic corpora now handle routine patient inquiries—soreness management, appliance breakage, scheduling—accounting for up to 80% of digital interactions. As discussed in The Lancet Digital Health (2025), governance and safety layers remain essential to prevent misinformation.

Administrative burnout is reduced, and clinician time is redirected toward complex decision-making.

6. Toward Multimodal Foundation Models

The next frontier lies in foundation models trained on heterogeneous datasets: CBCT volumes, IOS meshes, genomic polymorphisms, skeletal age metrics, and longitudinal treatment outcomes.

Federated learning frameworks (Comput. Biol. Med., 2025) enable multi-institutional model training without transferring raw patient data—preserving privacy while expanding dataset diversity.

Future systems may predict:
- Individualized tooth movement velocity
- Risk of root resorption
- Long-term post-retention stability (Angle Orthod., 2024)

Such models represent a shift from episodic decision support to continuous, learning healthcare ecosystems.

7. Ethical Guardrails and Regulatory Landscape
7.1 Algorithmic Bias

Facial recognition systems historically underperform in underrepresented populations. Similar disparities have been documented in orthodontic AI for non-Caucasian craniofacial morphologies. Equity audits and mandatory dataset diversification are emerging best practices.

7.2 Explainability

The "black box" problem undermines clinician trust. Saliency mapping and attention heatmaps provide partial transparency, illustrating which anatomical features influenced predictions. Regulatory guidance summarized in Journal of Dental Research (2025) underscores the need for explainable frameworks in FDA-approved systems.

7.3 Data Sovereignty and Privacy

Three-dimensional facial scans carry high re-identification risk. Federated learning and differential privacy mechanisms mitigate exposure. Ethical analyses in The Lancet Digital Health advocate global governance standards balancing innovation and patient autonomy.

7.4 Autonomous Robotics

Emerging systems described in Nature Communications (2026) demonstrate robotic wire bending guided by AI planning algorithms. While promising, such autonomy necessitates stringent oversight to prevent procedural errors.

8. Limitations and Persistent Bottlenecks
Despite rapid progress, limitations remain:
- Limited longitudinal datasets for growth prediction
- Regulatory heterogeneity across jurisdictions
- High computational costs for multimodal foundation models
- Dependence on proprietary datasets restricting transparency

Furthermore, AI performance often degrades when applied to new clinical environments lacking harmonized imaging protocols.

9. The Rise of the "Super-Clinician"
Orthodontics in 2026 is not witnessing physician replacement, but augmentation. The clinician functions as "pilot-in-command," supervising autonomous subsystems handling segmentation, staging, monitoring, and communication.

Human expertise remains irreplaceable in:
- Interdisciplinary surgical coordination
- Psychosocial counseling
- Management of atypical biological responses
- Ethical judgment in ambiguous scenarios

The algorithmic smile thus symbolizes synergy rather than substitution.

10. CONCLUSION
Between 2024 and 2026, orthodontics has transitioned from digitization to intelligence integration. Automated landmarking, multimodal extraction prediction, PINN-driven biomechanics, generative facial simulation, and federated foundation models collectively redefine practice standards.

The specialty now stands at a pivotal juncture: embracing AI not as a replacement for clinical reasoning, but as an amplifier of precision, safety, and personalization. The orthodontist of the future commands a network of intelligent agents—transforming raw data into insight, and insight into biologically harmonious smiles.

The algorithm does not smile. The patient does.

REFERENCES
- Am. J. Orthod. (2026). AI in Clinical Decisions.
- Angle Orthod. (2024). Stability Prediction.
- Comput. Biol. Med. (2025). Federated Learning.
- Dentomaxillofac. Radiol. (2025). Root Resorption Detection.
- Eur. J. Orthod. (2025). Skeletal Age Prediction.
- Frontiers (2024). Mandibular Advancement Biomechanics.
- Frontiers (2025). Personalized Orthodontic Monitoring.
- IEEE Trans. Med. Imaging (2024). 3D U-Net Tooth Segmentation.
- Int. J. Comput. Dent. (2025). VR vs AI Simulations.
- J. Dent. Res. (2025). FDA-Approved AI Solutions.
- J. Orofac. Orthop. (2025). AI Workflow for Custom Brackets.
- Kazimierczak, N. et al. (2024). AI in Orthodontics. J. Clin. Med.
- Koz, S., Uslu-Akcam, O. (2025). AI Cephalometric Analysis. Diagnostics.
- Lancet Digital Health (2025). Ethical Dental AI.
- MDPI (2026a). Face-Driven Orthodontics. Multimodal Technol. Interact.
- Nanomed. Res. J. (2025). AI-Enhanced CBCT.
- Nature Biomedical Engineering (2025). Generative Craniofacial Models.
- Nature Communications (2026). Autonomous Robotic Orthodontics.
- Orthod. Craniofac. Res. (2024). AI in Cleft Care.
- PDXScholar (2025). AI in Malocclusion Review.
- PMC (2026a). Extraction Prediction Models.
- Shodiyeva, E. (2026). Multimodal Transformers for Landmark Detection.
- Taylor & Francis (2023). AI in Orthognathic Surgery. JOMS.
- World J. Orthod. (2026). AI-Monitored Aligner Outcomes.
- Yang, S., Cheng, Y. (2025). Force-Driven PINNs. Bioengineering.