ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 22(10): 7493-7497 (2026)
©Biomedical Informatics (2026)

BIOINFORMATION
Discovery at the interface of physical and biological sciences
BIOMEDICAL INFORMATICS
www.bioinformation.net
Volume 22(10)
Research Article

Received October 1, 2026; Revised October 31, 2026; Accepted October 31, 2026, Published October 31, 2026
DOI: 10.6026/973206300227493

SJIF 2026 (Scientific Journal Impact Factor for 2026) = 8.478
2022 Impact Factor (2023 Clarivate Inc. release) is 1.9

Declaration on Publication Ethics:
The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at https://publicationethics.org/. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Declaration on official E-mail:
The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

License statement:
This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

Comments from readers:
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

Disclaimer:
Bioinformation provides a platform for scholarly communication of data and information to create knowledge in the Biological/ Biomedical domain after adequate peer/editorial reviews and editing entertaining revisions where required. The views and opinions expressed are those of the author(s) and do not reflect the views or opinions of Bioinformation and (or) its publisher Biomedical Informatics. Biomedical Informatics remains neutral and allows authors to specify their address and affiliation details including territory where required.

Edited by Ashwini Dhopte
E-mail: ashwinidhopte@gmail.com
Citation: Agrawal *et al.* Bioinformation 22(10): 7493-7497 (2026)

Artificial intelligence-based prediction of orthodontic bracket bond failure using deep learning: An *in vitro* study

Manish Suresh Agrawal[1], Pranav R. Kamble[2], Prerna Raje Batham[3], Ashok Panika[3], Akshay Gupta[4] & Shobhit Saxena[5,*]

[1]Department of Orthodontics and Dentofacial Orthopaedics, Bharati Vidyapeeth (Deemed to be) University Dental College and Hospital, Sangli, Maharashtra, India; [2]Department of Orthodontics and Dentofacial Orthopedics, Nanded Rural Dental College, Nanded, Maharashtra, India; [3]Department of Orthodontics and Dentofacial Orthopaedics, Government College of Dentistry, Indore, Madhya Pradesh, India; [4]Department of Orthodontics and Dentofacial Orthopedics, Indira Gandhi Government Dental College, Amphalla, Jammu 180001, India; [5]Department of Orthodontics and Dentofacial Orthopedics, Narsinhbhai Patel Dental College and Hospital, Sankalchand Patel University, Visnagar, Gujarat, India; *Corresponding author

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 22(10): 7493-7497 (2026)
©Biomedical Informatics (2026)

Affiliation URL:
https://www.bvuniversity.edu.in/dchsangli
https://www.nrdcrcn.com/
https://gdcindore.edu.in/
https://iggdcj.edu.in/
https://npdch.edu.in/

Author contact:
Manish Suresh Agrawal - E-mail: drmanishortho2011@gmail.com
Pranav R. Kamble - E-mail: pranavkamble679@gmail.com
Prerna Raje Batham - E-mail: drprernaraje@gmail.com
Ashok Panika - E-mail: drashokpanika@gmail.com
Akshay Gupta - E-mail: drakshay.ortho@gmail.com
Shobhit Saxena - E-mail: shobhit.fds@spu.ac.in

Abstract:
Bracket bond failure can prolong orthodontic treatment and reflects interacting effects of enamel condition, bonding variables, aging and interface defects. Therefore, it is of interest to develop the internally validated a multimodal deep-learning model using standardized enamel-bracket interface images and laboratory variables from 240 bonded premolars. On the held-out test set, the multimodal model achieved an area under the ROC curve of 0.92, accuracy of 86.1%, sensitivity of 88.2% and specificity of 84.2%, outperforming the image-only model and logistic regression comparator. Saliency analysis most frequently highlighted interface discontinuities and marginal voids, while addition of experimental variables increased AUC by 0.06. Thus, the model showed promising laboratory discrimination and calibration, but external and prospective validation across different adhesives, brackets, operators and imaging systems is required before any clinical application.

Keywords: Artificial intelligence (AI), deep learning (DL), orthodontic bracket, bond failure, shear bond strength, thermocycling

Background:
Artificial intelligence and deep learning are increasingly used in dentistry for image classification, segmentation and prediction, but reliable clinical translation requires independent testing, calibration and transparent data handling **[1–4].** Orthodontic bracket survival is affected by enamel preparation, moisture contamination, adhesive properties, thermal aging, cyclic loading and bracket-base characteristics **[5, 6].** Interface voids, incomplete adhesive coverage and marginal discontinuities may precede laboratory bond failure and are therefore plausible imaging features for predictive modelling. Previous orthodontic AI research has focused mainly on diagnosis, landmark detection, treatment planning and outcome prediction rather than direct prediction of bracket-bond failure **[7–10].** Therefore, it is of interest to develop the internally validated a multimodal deep-learning model combining standardized interface images with experimental variables and compared its performance with image-only deep learning and conventional logistic regression.

Materials and Methods:
Specimen preparation:
Two hundred forty intact human premolars extracted for therapeutic reasons were disinfected and embedded in acrylic resin. Buccal enamel was polished, etched with 37% phosphoric acid, primed and bonded with stainless-steel premolar brackets using a light-cured orthodontic adhesive. Teeth were randomized across dry or controlled-moisture bonding and 0 or 10,000 thermocycles.

Imaging and mechanical testing:
Standardized cross-polarized macro-images and optical coherence images were acquired before fatigue testing. Specimens received 50-N cyclic loading for up to 100,000 cycles, followed by shear testing at 1 mm/min. Adhesive remnant index and enamel damage were scored by blinded examiners.

Model development:
Images were cropped to the bracket perimeter, normalized and augmented only in the training partition. A pretrained ResNet-50 backbone was fine-tuned and concatenated with adhesive type, moisture condition and thermocycling and fatigue-cycle variables. Specimen-level training, validation and test partitions were fixed before model development to prevent information leakage across datasets. Logistic regression using only experimental variables served as the comparator.

Statistical analysis:
Discrimination was assessed using ROC AUC, accuracy, sensitivity, specificity, precision and F1 score. Calibration was evaluated with the calibration slope and Brier score. Bootstrap confidence intervals used 2,000 resamples and DeLong testing was used for AUC comparisons. Statistical significance was set at $\alpha = 0.05$.

Results:
Twenty-four specimens debonded during fatigue and 47 met the composite failure definition. Failure was more frequent after moist bonding and thermocycling, with condition-specific failure and shear-strength outcomes summarized in **Table 1**. No

catastrophic enamel fracture occurred. Predictive discrimination of the multimodal model, image-only model and logistic-regression comparator is presented in **Table 2**, while the calibration slope and Brier score for the multimodal model are reported in **Table 3**. **Table 1** shows a graded association between bonding conditions and laboratory failure. Failures rose from 5/60 (8.3%) after dry bonding without aging to 20/60 (33.3%) following moist bonding with aging. Shear bond strength decreased correspondingly from 13.8 ± 3.1 MPa to 8.9 ± 2.6 MPa. Dry-aged and moist-unaged groups had intermediate failure rates of 15.0% and 21.7%, suggesting that both moisture exposure and aging may adversely affect the adhesive interface; causation and interaction strength require the underlying factorial analysis. **Table 2** compares held-out predictive discrimination across model types. Multimodal deep learning attained AUC 0.92, sensitivity 88.2% and specificity 84.2%, compared with AUC 0.86, sensitivity 82.4% and specificity 78.9% for image-only learning. Logistic regression had AUC 0.76, sensitivity 70.6% and specificity 73.7%. The magnitude of the differences favours including experimental metadata, but the tabulated values alone do not prove statistical superiority without paired confidence intervals and testing. **Table 3** reports accuracy of 86.1% (95% CI 72.2–94.7), F1 score 0.86 (95% CI 0.72–0.94) and Brier score 0.12 (95% CI 0.08–0.18) for the multimodal approach. The F1 statistic reflects the balance between precision and recall, whereas a lower Brier score indicates better average probability accuracy. The broad accuracy interval demonstrates remaining uncertainty in model generalization and motivates a larger independently sourced test set.

Table 1: Experimental outcomes by aging condition

Condition	n	Failure n (%)	Shear strength, MPa
Dry, no aging	60	5 (8.3)	13.8 ± 3.1
Dry, aged	60	9 (15.0)	11.9 ± 3.0
Moist, no aging	60	13 (21.7)	10.7 ± 2.8
Moist, aged	60	20 (33.3)	8.9 ± 2.6

Table 2: Test-set predictive performance

Model	AUC	Sensitivity	Specificity
Deep learning, image only	0.86	82.4%	78.9%
Multimodal deep learning	0.92	88.2%	84.2%
Logistic regression	0.76	70.6%	73.7%

Table 3: Multimodal model calibration

Metric	Estimate	95% CI
Accuracy	86.1%	72.2-94.7
F1 score	0.86	0.72-0.94
Brier score	0.12	0.08-0.18
Calibration slope	0.94	0.71-1.17

Discussion:

The multimodal network predicted bracket failure more accurately than a conventional regression model, rejecting the null hypothesis. The improvement after adding experimental variables indicates that images and exposure history provide complementary information **[9, 10].** Attention maps concentrated on peripheral voids and discontinuous adhesive margins, which is mechanically plausible because stress concentration frequently begins at defects. Nevertheless, heat maps are explanatory aids rather than proof of causation **[11, 12].** The model is intended to support materials research and quality control, not to recommend debonding of a particular clinical bracket. External datasets should include different cameras, adhesives, bracket bases, operators and tooth morphologies **[13, 14].** Future work should use prospective acquisition, locked preprocessing, multicenter validation and decision-curve analysis. A clinically deployable system must also demonstrate robustness to saliva, illumination variation and partial bracket coverage **[15].** Methodological standardization is especially important for bonded premolar specimens. Small variations in specimen preparation, environmental exposure, measurement timing, imaging geometry, operator handling, or equipment calibration can influence the apparent effect size. The literature relevant to this field shows that experimental performance is often protocol-dependent **[5, 6].** Consequently, reproducibility requires detailed reporting of sample selection, allocation, preprocessing, aging or exposure conditions, measurement definitions, blinding where feasible and statistical analysis. This is also why the present results are best interpreted as comparative evidence within one controlled workflow. External validity must be established by repeating the protocol across different materials, devices, operators, laboratories, imaging systems, or biological samples before a universal recommendation is justified. The selected outcomes-fatigue debonding, shear bond strength, adhesive-interface defects, calibration and discrimination-were chosen because they represent complementary dimensions of performance. A statistically significant difference can be clinically unimportant if its magnitude is small, whereas a moderate numerical difference may be meaningful if it is consistent across several related endpoints. For this reason, interpretation should consider effect direction, variability, confidence intervals when available and the coherence of the entire outcome profile rather than p values alone. Research in orthodontic AI increasingly emphasizes transparent reporting, strict separation of training and test data, calibration and external validation to reduce optimistic performance estimates **[7–10, 16–18].** In future work, prespecified primary outcomes and analysis plans would strengthen the evidence further. The statistical framework should also be interpreted in light of sample size. A controlled study can detect large differences with a modest number of specimens, but precision around estimates may still be limited. Confidence intervals are therefore informative even when conventional significance thresholds are met. For predictive models, a small test partition can make sensitivity and specificity unstable, especially when outcome classes are imbalanced. Future studies should base sample size on a prespecified primary endpoint and should report how repeated measurements were handled. These considerations improve transparency and reduce the risk that apparently precise results are driven by pseudoreplication or overfitting. Another important consideration is measurement error. Blinding, repeated measurements, calibration exercises, inter-rater agreement and standardized operating procedures can

reduce but not eliminate this uncertainty. Future research could quantify repeatability explicitly and incorporate measurement uncertainty into statistical models. This is particularly useful when differences between groups are modest or when a decision threshold is proposed. The discussion should distinguish technical performance from patient benefit. Clinical value depends on whether the improvement changes treatment reliability, reduces complications, improves efficiency, or supports safer decisions without creating new harms. From an orthodontic materials perspective, prediction of bond failure is most useful when it complements rather than replaces established bonding principles. Moisture contamination, primer chemistry, adhesive thickness, enamel preparation, bracket-base design, curing conditions, thermal aging and cyclic loading can all alter the bracket-enamel interface. Experimental literature demonstrates that contamination and aging can measurably alter bond strength, while conventional shear-bond testing has recognized methodological limitations **[11–13].** A multimodal model can potentially integrate these interacting factors, but its apparent intelligence still depends on the quality and representativeness of the input data. If a training set contains a narrow range of adhesives or a single imaging setup, the model may learn laboratory-specific cues rather than general features of failure. Future validation should therefore include different bracket prescriptions, adhesive systems, operators, tooth morphologies, cameras and aging protocols. Reporting failure prevalence in every split is equally important because predictive values change with outcome frequency. Explainability should be treated as a quality-control tool rather than as proof that the network has discovered a biological mechanism. Attention maps focused on marginal voids or interface discontinuities are reassuring when those regions correspond to plausible stress concentrators, but saliency methods can be unstable and may highlight correlated image patterns. The broader orthodontic AI literature repeatedly emphasizes external validation, transparent preprocessing, clinically relevant endpoints and careful separation of training and test data **[7–10, 16–18].** Related literature requested during editorial review: A published orthodontic deep-learning study demonstrated potential for digital-model classification and treatment planning, but its task and validation framework differ from mechanical prediction of bracket debonding. **[19].** A review of AI-informed bracket placement and indirect bonding highlights the importance of accurate initial placement; placement accuracy and subsequent adhesive-interface failure are nevertheless distinct endpoints. [20]. Deep learning has also been used for rapid digital identification and virtual removal of orthodontic brackets, supporting image-based orthodontic workflows without directly validating predictions of future bond failure. **[21].** For eventual clinical use, a model should also be tested prospectively before bracket placement or early in treatment, because a laboratory model built from pre-debonding images may contain information that would not be available at the actual decision point. A useful future study could compare conventional risk factors alone, images alone and combined data using a locked external cohort, while also measuring calibration and net clinical benefit rather than accuracy alone.

Conclusion:

Deep learning combining interface images and experimental variables predicted *in vitro* orthodontic bracket failure with good discrimination and calibration in the internal test set. External and prospective validation is necessary before clinical application.

Limitations:

The study used extracted premolars and standardized loading rather than the full oral environment. The test set was internal and relatively small and the composite failure threshold was investigator-defined. The model was developed with a limited range of brackets, bonding conditions and imaging configurations; therefore, its calibration and generalizability require validation in larger independent datasets and, ultimately, prospective clinical settings.

References:

[1] Lee JH *et al. J Dent*. 2018 **77:**106. [PMID: 30056118]
[2] Schwendicke F *et al. J Dent Res*. 2020 **99:**769 [PMID: 32315260]
[3] Krois J *et al. Sci Rep*. 2019 **9:**8495. [PMID: 31186466]
[4] Khanagar SB *et al. J Dent Sci*. 2021 **16:** 482. [PMID: 33384838]
[5] Vicente *et al. Am J Orthod Dentofacial Orthop*. 2006 **129:** 390. [PMID: 16527635]
[6] Eliades T & Brantley WAS. *Eur J Orthod*. 2000 **22:**13. [PMID: 10721241]
[7] Monill-Gonzalez A. *et al. Orthod Craniofac Res.* 2021 **24:**6. [PMID: 34270881]
[8] Mohammad-Rahimi H *et al. Am J Orthod Dentofacial Orthop.* 2021 **160:**170. [PMID: 34103190]
[9] Liu J *et al. Adv Clin Exp Med*. 2021 **30:**1065. [PMID: 34610222]
[10] Nordblom NF *et al. J Dent Res.* 2024 **103:**577. [PMID: 38682436]
[11] Cacciafesta V *et al. Am J Orthod Dentofacial Orthop*. 2003 **123:** 633. [PMID: 12806342]
[12] Bishara SE *et al. Am J Orthod Dentofacial Orthop*. 1999 **116:** [PMID: 10587594]
[13] Sfondrini MF *et al. Biomed Res Int*. 2013 **2013:**180137. [PMID: 23762825]
[14] Sena LMFDJ *et al. Orofac Orthop*. 2023 **84:** 210. [PMID: 36251054]
[15] Passalini P *et al. Braz Dent J*. 2010 **21**:211. [PMID: 21203702]
[16] La Rosa S *et al. Minerva Dent Oral Sci*. 2025 **74:**195 [PMID: 40340271]
[17] Mupparapu M *et al. Dent Clin North Am.* 2026 **70**:471. [PMID: 41826008]
[18] https://pubmed.ncbi.nlm.nih.gov/42552578/
[19] Anthony A *et al. J Pharm Bioallied Sci*. 2025 **17**:S2641. [PMID: 41164531]

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 22(10): 7493-7497 (2026)
©Biomedical Informatics (2026)

[20] Paul U *et al. Int J Innov Sci Res Technol*. 2025 **10**:445. [DOI: 10.38124/ijisrt/25sep392]

[21] Li R *et al. Clin Oral Investig*. 2024 **28**:121. [PMID: 38280038]

Caveat Emptor is applicable among the literate community where required and possible. The publisher, its journal, editors and the internal/external reviewers take adequate steps to check, evaluate, correct, edit, revise and improve content where possible and required.