

Article

Assessing How CBCT Image Quality Influences Diagnostic Evaluability of Periodontal Bone: Establishing Human Baselines for AI Training (In Vitro Study)

Michael Moncher ^{1,*}, Vera Zimprich ¹, Jonathan von See ¹, Jörg Philipp Tchorz ², Theodor von See ¹
and Constantin von See ¹

¹ Research Center for Digital Technologies in Dentistry and Computer Aided Design/Computer-Aided Manufacturing, Department of Dentistry, Faculty of Medicine and Dentistry, Danube Private University, Steiner Landstraße 124, 3500 Krems an der Donau, Austria; vera.zimprich@dp-uni.ac.at (V.Z.); constantin.see@dp-uni.ac.at (C.v.S.)

² Division for Endodontics, Center for Operative Dentistry and Periodontology, Department of Dentistry, Faculty of Medicine and Dentistry, Danube Private University, 3500 Krems an der Donau, Austria; joerg.tchorz@dp-uni.ac.at

* Correspondence: michael.moncher@dp-uni.ac.at; Tel.: +43-6801338866

Highlights

What are the main findings?

- Systematic degradation of CBCT image quality (voxel enlargement and simulated blur) is associated with statistically significant and effect-size-quantified shifts in examiner-derived periodontal bone-level measurements and inter-examiner reliability.
- Strong reductions in spatial resolution particularly affect measurement consistency in periodontally compromised teeth, while moderate degradation shows more variable effects.

What are the implications of the main findings?

- CBCT image quality represents a critical determinant of measurement reliability in periodontal assessment and should be carefully considered when selecting acquisition protocols.
- Reliable human baseline measurements are essential for training robust and generalizable AI systems, highlighting the need for standardized image quality in CBCT-based datasets.

Abstract

Background: Cone-beam computed tomography (CBCT) is increasingly applied for the assessment of periodontal bone levels. However, its measurement reliability and consistency depend strongly on image quality parameters such as voxel size, noise, and reconstruction sharpness. With the growing use of CBCT datasets for artificial intelligence (AI)-based diagnostics, it is essential to understand how image degradation conditions affect examiner-derived measurement outcomes and the reliability of reference data used for AI training. **Methods:** An anonymized CBCT dataset containing one periodontally healthy tooth (31) and one tooth with pronounced periodontal bone loss (41) was analyzed. The original DICOM data were systematically degraded using controlled voxel enlargement (double and triple voxel size) and simulated image blur (Gaussian and median filtering). Six dentists ($n = 6$) independently performed standardized linear bone-level measurements, with three repeated measurements per tooth and image condition. Data were analyzed using the Shapiro–Wilk test for normality assessment, the Kruskal–Wallis H test for group comparisons, Bonferroni-adjusted Mann–Whitney U tests for post hoc pairwise comparisons, and intraclass correlation coefficients (ICC (2,1)) for inter-examiner reliability assessment.



Received: 16 January 2026

Revised: 4 March 2026

Accepted: 10 March 2026

Published: 16 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Results: A total of 180 measurements were evaluated. Image degradation conditions were associated with statistically significant differences in bone-level measurements for both teeth (tooth 31: $p = 0.017$; tooth 41: $p = 0.0049$). Significant pairwise differences were primarily observed between the original dataset and specific degraded conditions involving blur and reduced spatial resolution, while several comparisons remained non-significant. Inter-examiner reliability varied across image groups and decreased notably with pronounced voxel enlargement, particularly in the periodontally compromised tooth. **Conclusions:** Controlled image degradation conditions of CBCT image quality significantly affect measurement outcomes and inter-examiner reproducibility of periodontal bone measurements. These findings demonstrate that image quality is a critical determinant of measurement reliability and examiner-dependent interpretation. From both a clinical and AI-development perspective, maintaining adequate CBCT resolution may contribute to more consistent measurement behavior and more reliable training datasets.

Keywords: cone-beam computed tomography; periodontal bone loss; voxel size; image degradation conditions; diagnostic accuracy; inter-examiner reliability; artificial intelligence; dental imaging

1. Introduction

Periodontal bone loss is one of the most clinically relevant indicators for diagnosing, staging, and monitoring periodontitis, and radiographic assessment remains an indispensable component of periodontal evaluation. In recent years, cone-beam computed tomography (CBCT) has gained increasing importance due to its ability to provide three-dimensional visualization of periodontal structures, offering advantages over conventional two-dimensional imaging, such as improved detection of intrabony defects, dehiscences, and complex bone morphologies [1,2]. However, the diagnostic value of CBCT is closely dependent on image quality, including voxel size, noise levels, and reconstruction algorithms, which directly influence the visibility and measurability of fine periodontal structures [3–5].

Voxel size has been widely discussed as a key determinant of measurement precision in CBCT-based periodontal assessments. However, its diagnostic impact remains incompletely resolved. While smaller voxel sizes generally improve spatial resolution, they are associated with increased radiation exposure, whereas larger voxel sizes may reduce anatomical detail and increase the risk of measurement inaccuracy or diagnostic misclassification [6,7]. Several studies have reported that voxel enlargement compromises the detection of subtle cortical defects, external resorptions, and early-stage bone loss [8–10]. However, other investigations suggest that clinically acceptable measurement variability may still be achievable under certain low-resolution or low-dose conditions, depending on defect morphology and examiner experience [11]. Moreover, CBCT studies differ substantially in acquisition protocols, voxel ranges, reconstruction algorithms, and evaluation methods, limiting direct comparability across findings. Variations in exposure parameters, segmentation accuracy, and post-processing techniques further contribute to inconsistent conclusions regarding diagnostic fidelity [12–14]. Consequently, despite extensive research, the extent to which voxel size and image degradation systematically influence periodontal bone measurements, particularly in the context of standardized measurement workflows and emerging AI applications, remains insufficiently characterized [6].

Importantly, most previous investigations have evaluated voxel size or exposure parameters under device-specific acquisition settings rather than through systematic, experimentally controlled degradation of identical datasets. As a result, it remains difficult

to isolate the independent contribution of spatial resolution reduction or image blurring to examiner-derived measurement variability. Controlled degradation modeling offers a methodological advantage by allowing the manipulation of defined image parameters while preserving all other anatomical and acquisition-related variables. Such an approach enables the assessment of relative measurement shifts under standardized conditions and provides a reproducible experimental framework for analyzing image-quality-dependent variability.

Beyond clinical diagnostics, interest in leveraging CBCT datasets for AI-driven periodontal assessment has surged in recent years. Initial studies have indeed reported promising results—for instance, deep learning models can automatically detect, classify, and quantify periodontal bone loss from both 2D radiographs and 3D CBCT scans. Some AI systems even approach expert-level performance in identifying bony defects and measuring bone levels under ideal conditions [9,15–19]. However, these advances come with notable limitations and open questions. Critically, most algorithms have been developed and tested on relatively high-quality, controlled imaging datasets—sometimes even excluding lower-quality scans, which inherently limits their generalizability [20]. In practice, CBCT image quality varies widely due to differences in voxel size, noise levels, reconstruction algorithms and artifact burdens. While artificial intelligence (AI) models have achieved promising performance on standardized datasets, their robustness under real-world variability remains unclear. Recent research demonstrates that AI-driven noise reduction and artifact suppression can significantly enhance subjective and objective image quality parameters, such as contrast-to-noise ratio and artifact index, suggesting the potential utility of AI in dental CBCT [21]. However, these improvements are often observed in controlled settings and do not necessarily translate across heterogeneous data sources. Systematic analyses of deep learning-based image enhancement methods indicate that, although low-dose CBCT images can be processed to improve diagnostic visibility, the intrinsic variability in CBCT image quality—including noise and artifact patterns—poses ongoing challenges that have not yet been fully resolved [22]. Moreover, deep learning techniques for artifact reduction show promise but are currently limited by the need for consistent, high-quality training data and may not generalize well to all clinical CBCT acquisitions [22]. Finally, broader reviews of AI applications in dental imaging highlight that models performing well on one dataset often lack external validation and may not maintain accuracy across scanners and imaging protocols, raising concerns for generalizability in real-world practice [23]. This inconsistency underscores the need to better understand how CBCT image degradation systematically affects both human and algorithmic interpretation, and thus motivates the current study, which investigates the impact of voxel enlargement and simulated blur on periodontal bone evaluation.

From a methodological perspective, variability in image quality may introduce structured label noise into AI training datasets. If examiner-derived measurements systematically shift under specific degradation conditions, such variability may propagate into ground-truth annotations, thereby affecting model calibration and generalizability. Consequently, understanding how controlled image degradation conditions influence human measurement behavior represents a critical prerequisite for establishing reliable reference baselines in AI-based periodontal diagnostics.

Image noise and blur further complicate periodontal assessment in CBCT and reflect inherent limitations of the modality. Unlike conventional medical CT, CBCT reconstruction relies on cone-beam geometry and filtered back-projection algorithms that are particularly susceptible to noise amplification and incomplete scatter correction. Consequently, reconstruction and smoothing algorithms often fail to preserve fine anatomical detail and instead introduce image blurring—an established limitation of CBCT. Although post-processing

and noise-reduction techniques may improve subjective readability, several studies have shown that such approaches can distort edge definition and bias quantitative measurements, especially when assessing thin cortical bone or subtle periodontal defects [24–26]. Image noise is further exacerbated by artifacts caused by high-density materials. Metallic restorations, implants, and prosthetic components generate beam hardening, streak artifacts, and photon starvation, leading to localized contrast loss and artificial blurring. These effects are particularly pronounced in periodontal regions adjacent to metallic structures or complex anatomical boundaries, where partial volume effects further impair defect visualization [27]. While metal-artifact-reduction and advanced reconstruction techniques have been proposed, their effectiveness remains limited and highly dependent on scanner design and acquisition parameters. When such degradations occur, examiner-dependent variability increases substantially, as observers rely more on subjective interpretation rather than clearly defined anatomical landmarks. This has been consistently reported in studies evaluating CBCT reproducibility under varying acquisition and reconstruction conditions [28–30].

Despite the growing body of literature addressing voxel size, artifact reduction, and AI-based periodontal diagnostics, a structured experimental framework systematically modeling defined image degradation conditions within a single standardized dataset remains largely absent. Specifically, there is a lack of proof-of-concept investigations quantifying how controlled voxel enlargement and simulated blurring influence examiner-derived periodontal bone measurements under identical anatomical conditions. Addressing this gap is essential to disentangle measurement variability caused by image degradation from variability arising from anatomical heterogeneity or acquisition differences [27,31,32].

Therefore, the aim of the present study was to assess how systematic degradation of CBCT datasets via voxel enlargement and simulated blur affects periodontal bone measurements performed by clinicians. By establishing a controlled in vitro reference standard and quantifying examiner variability across multiple image degradation conditions, this study seeks to provide a reproducible methodological framework for analyzing image-quality-dependent measurement shifts. Rather than evaluating absolute diagnostic validity, the present investigation focuses on relative measurement behavior and inter-examiner reproducibility under standardized degradation scenarios.

Moreover, given the rapid proliferation of AI-based periodontal imaging tools, understanding how image quality affects human-derived measurements baselines is a critical step toward designing robust AI training pipelines capable of accommodating real-world variability. The findings are intended to serve as hypothesis-generating evidence and as a methodological foundation for future multicenter validation studies integrating human and algorithmic performance assessments under controlled image degradation conditions.

In the present study, the term “image degradation conditions” refers to systematically modified CBCT datasets generated through voxel enlargement and/or simulated blurring.

2. Materials and Methods

2.1. Study Design and Conceptual Framework

This study was designed as an exploratory in vitro proof-of-concept investigation to evaluate the influence of controlled image degradation conditions on periodontal bone-level measurements. A single CBCT dataset was systematically modified to generate predefined image degradation conditions. The primary outcome was examiner-derived linear bone-level measurement (mm). The study aimed to quantify relative measurement shifts and inter-examiner reproducibility across degradation conditions rather than to assess diagnostic accuracy against a gold standard. Given the exploratory proof-of-concept design and the use of a single CBCT dataset under controlled experimental manipulation,

the study was not intended to establish absolute diagnostic thresholds but to isolate the relative influence of defined image-quality parameters on examiner-derived measurements under standardized conditions.

2.2. CBCT Dataset and Acquisition Parameters

For systematic investigations, a CBCT dataset was acquired from an anonymized participant who exhibited one periodontally healthy tooth (31) and one tooth with clearly detectable periodontal bone loss (41). Imaging was performed using a Planmeca ProMax 3D Mid unit (Planmeca, Helsinki, Finland). The original isotropic voxel size of the acquired dataset was $0.2\text{ mm} \times 0.2\text{ mm} \times 0.2\text{ mm}$, which served as the baseline spatial resolution for all subsequent degradation procedures. The complete dataset was anonymized in Romexis 7 (Planmeca), and the patient orientation was randomized prior to export. Randomization of orientation was performed to minimize potential observer bias related to anatomical alignment. The dataset was then saved as a single-shot DICOM series, which served as the reference for all subsequent processing steps and for generating the artificially degraded comparison datasets. The CBCT acquisition was performed with a cylindrical field of view of 200 mm in diameter and 100 mm in height, using a tube voltage of 90 kV, a tube current of 8 mA, and an exposure time of 13.5 s. The recorded dose–area product (DAP) was $1745\text{ mGy}\cdot\text{cm}^2$, and the computed tomography dose index (CTDI) was 8.1 mGy. No additional reconstruction filters beyond the manufacturer’s standard reconstruction protocol were applied during initial image acquisition. All DICOM data were exported without compression, and grayscale intensity values were preserved in their native format without rescaling or histogram normalization prior to degradation processing.

2.3. Image Degradation Pipeline

The analysis was based on the exported clinical DICOM data, from which a single axial series was selected to function as the reference volume. The selected axial series contained the complete mandibular anterior region and was verified to include consistent slice spacing and intact DICOM metadata prior to further processing. Artificial image degradation was applied to this dataset using two methodological approaches: a controlled reduction in spatial resolution through voxel enlargement and a simulation of image blur using filter-based noise modification. All degradation procedures were performed on the identical reference dataset in order to isolate the independent effect of each image degradation condition while preserving anatomical constancy. No intensity normalization, contrast enhancement, histogram equalization, or additional preprocessing steps were applied before or after degradation procedures. The degradation workflow was implemented using Python-based processing (SimpleITK v2.3.1, NumPy v1.26.4, SciPy v1.11.4), and all transformation parameters were applied uniformly across datasets to ensure methodological consistency and reproducibility.

2.3.1. Voxel Enlargement

To simulate reduced spatial resolution, the original DICOM reference volume was re-sampled using a custom Python script built on the SimpleITK library. The original isotropic voxel size of 0.2 mm was increased to 0.4 mm (Group 2) and 0.6 mm (Group 3), respectively, in all three spatial axes (x, y, z), resulting in isotropic resampling. All valid slices of the input dataset were merged into a three-dimensional volume, and the image spacing was increased by a factor of either two or three in all spatial axes. Resampling was performed using trilinear interpolation (SimpleITK (v2.3.1; Insight Software Consortium, open-source software, USA) default linear interpolator) to approximate clinically realistic reconstruction behavior while avoiding edge exaggeration associated with higher-order spline interpolation. The resampled volume was reconstructed using linear interpolation and then re-sliced

into individual layers. No nearest-neighbor interpolation was used in order to prevent voxel block artifacts. The resulting DICOM series contained regenerated UID structures to ensure DICOM conformity, while essential relative metadata such as instance number and slice location were preserved. Grayscale intensity values were preserved during resampling without histogram normalization, contrast adjustment, or intensity rescaling.

2.3.2. Simulated Blurring

The second form of image degradation involved a simulated blurring process designed to reduce image sharpness and mimic noise-related quality loss. This was carried out using Jupyter Notebook (v7.0) with pydicom, numpy, and scipy.ndimage libraries. Each slice of the dataset was modified using either a Gaussian Blur filter with a value of 4.5 or a median filter with a value of 2.0. For Gaussian filtering, a standard deviation parameter of $\sigma = 4.5$ was applied isotropically across the pixel matrix. This parameter was selected based on pilot simulations to induce measurable attenuation of high-frequency edge information while preserving overall anatomical morphology. The median filter (kernel size = 2.0) was chosen to simulate clinically realistic noise-reduction post-processing while avoiding excessive morphological distortion. Pixel matrices were modified directly, and all resulting series were saved with updated DICOM metadata under standard DICOM encoding. No additional preprocessing, contrast enhancement, thresholding, intensity normalization, or grayscale compression was applied before or after blurring procedures.

2.4. Experimental Group

The degradation workflow was applied independently to the identical reference dataset to ensure anatomical constancy across all image degradation conditions.

The experimental conditions were defined as follows:

- Group 1 (Original DICOM, Figure 1): original voxel size (0.2 mm) without modification;
- Group 2 (Double Voxel, Figure 2): isotropic resampling to 0.4 mm;
- Group 3 (Triple Voxel, Figure 3): isotropic resampling to 0.6 mm;
- Group 4 (Gaussian Blur, Figure 4): $\sigma = 4.5$ applied to original voxel size;
- Group 5 (Gaussian Blur + Triple Voxel, Figure 5): $\sigma = 4.5$ combined with 0.6 mm isotropic voxel resampling.

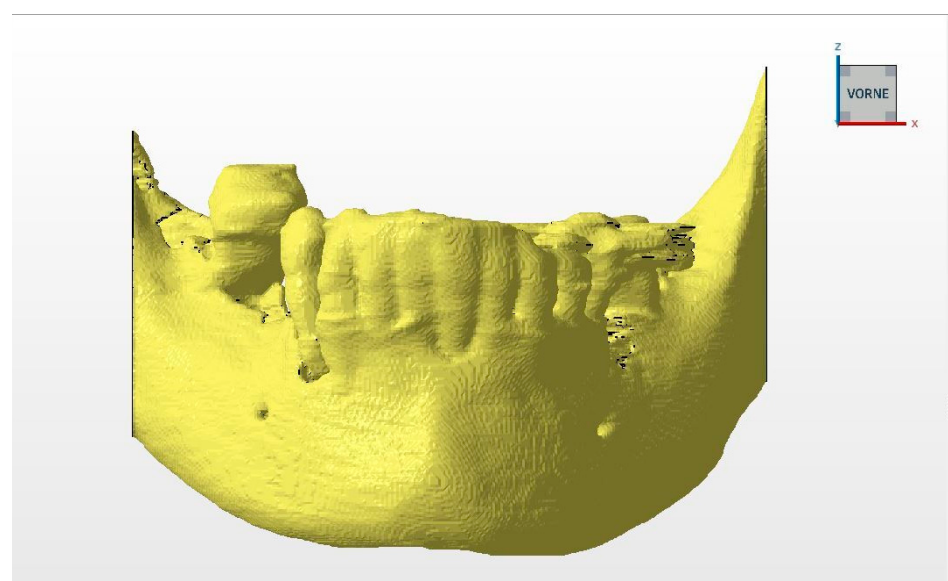


Figure 1. Group 1 (original DICOM): STL model of the mandibular anterior region after application of a median filter (2.0), showing reduced surface noise and smoothing of bone and tooth structures with partial loss of fine anatomical detail.

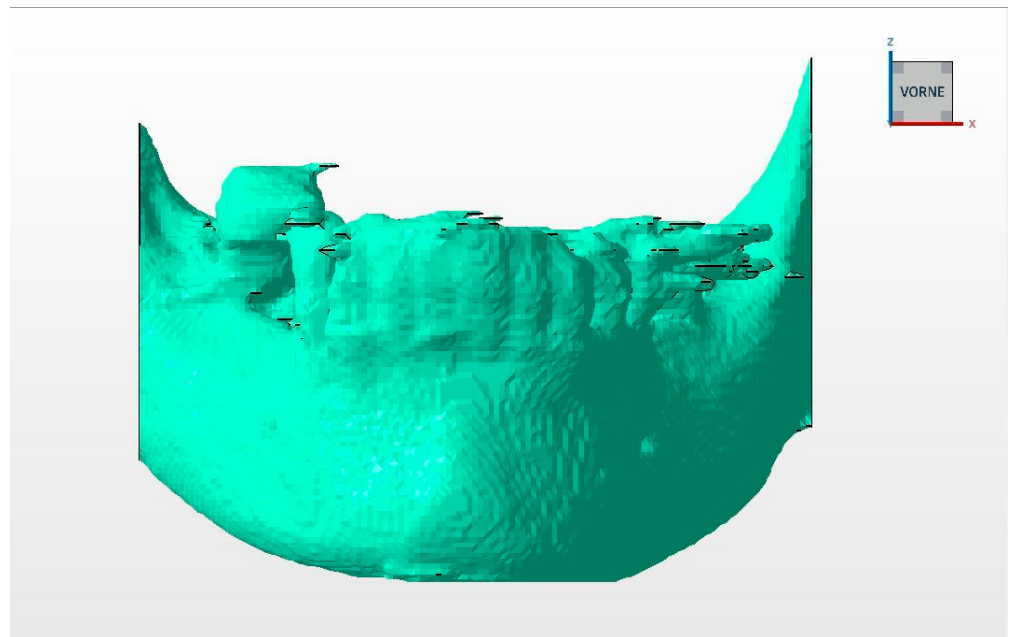


Figure 2. Group 2 (Double Voxel Size): STL model of the mandibular anterior region after voxel size doubling, showing a coarser surface structure with reduced spatial detail compared with the reference dataset. The overall anatomical morphology is preserved, while fine structural features—particularly along the alveolar bone crest and interdental regions—appear less sharply defined.

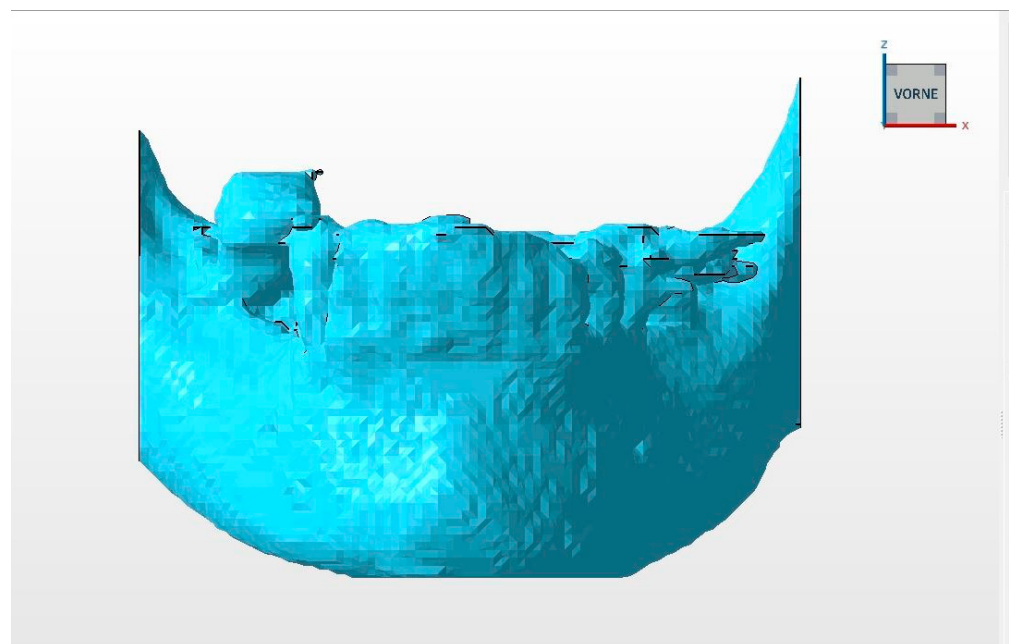


Figure 3. Group 3 (Triple Voxel Size): STL model of the mandibular anterior region after tripling of voxel size, showing a substantially simplified surface geometry with clearly reduced anatomical detail. While the overall mandibular morphology remains visible, fine anatomical features and sharp contours—particularly at the alveolar bone crest—are only limitedly represented.

These group labels were used consistently throughout the manuscript. All resulting datasets were converted into STL files and exported for subsequent geometric analysis.

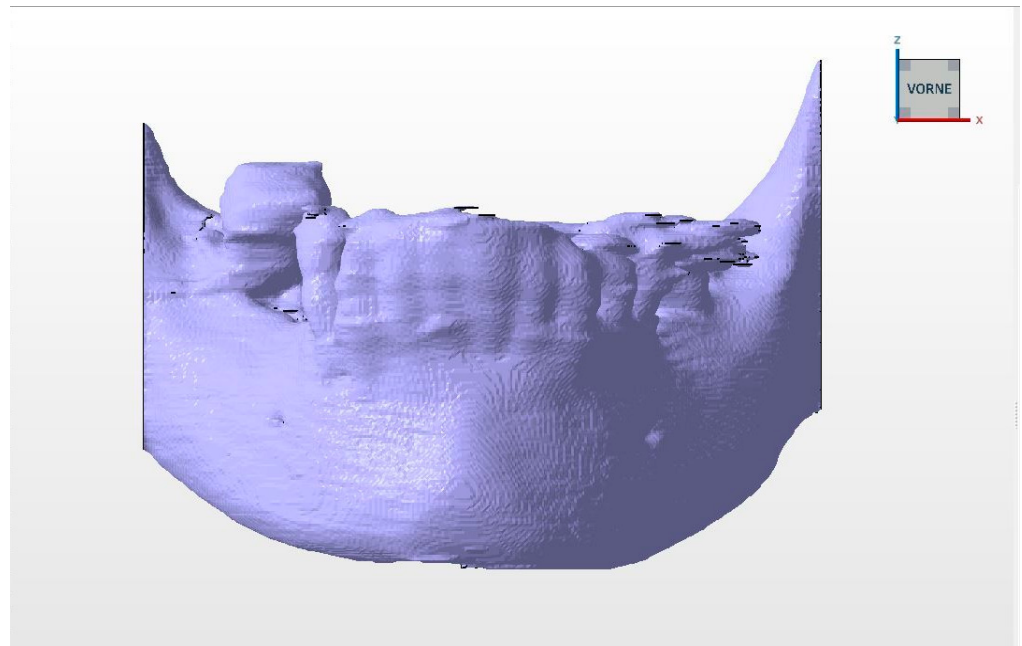


Figure 4. Group 4 (Gaussian Blur): STL model of the mandibular anterior region after application of a Gaussian blur filter ($\sigma = 4.5$), showing uniform surface smoothing and reduced high-frequency surface roughness. The macroscopic anatomy of bone and teeth remains recognizable, while fine edges and detailed structural features appear attenuated.

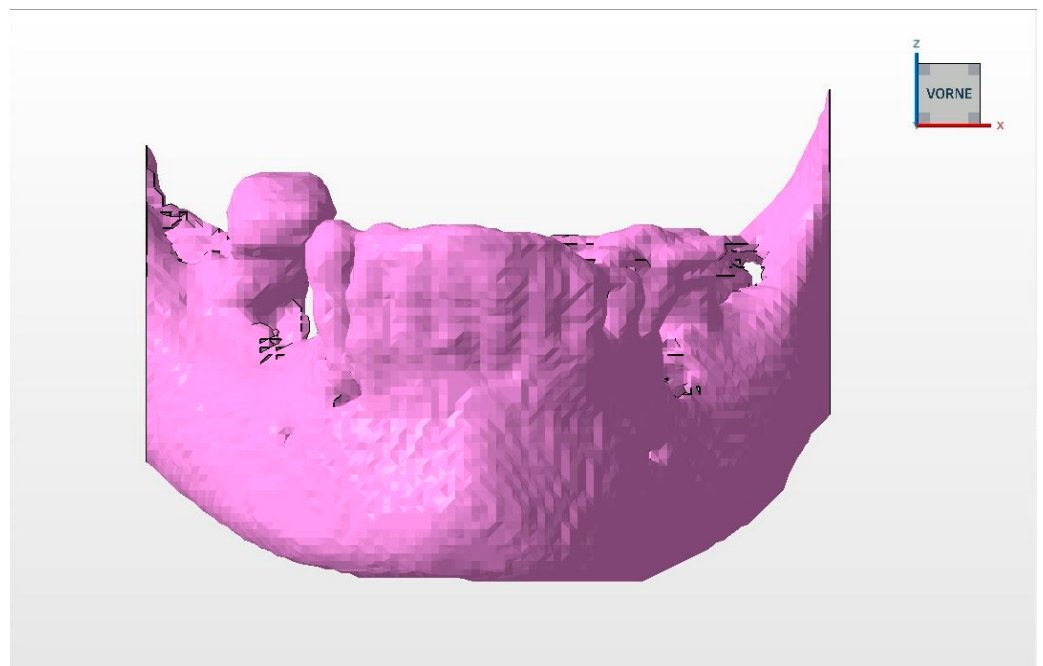


Figure 5. Group 5 (Gaussian Blur + Triple Voxel Size): STL model of the mandibular anterior region after combined Gaussian blurring ($\sigma = 4.5$) and tripled voxel size, showing pronounced surface smoothing together with reduced spatial resolution. The surface appears homogenized, while fine anatomical details and sharp transitions of bone and tooth structures are markedly diminished.

2.5. STL Conversion and Geometric Processing

The STL models were imported into Autodesk Netfabb (Version 2021; Autodesk Inc., San Rafael, CA, USA). STL generation inherently involves segmentation thresholds, triangulation, and mesh reconstruction. Segmentation was performed using a fixed global threshold value applied uniformly across all groups to avoid condition-specific segmenta-

tion bias. Mesh triangulation density was kept constant across all groups, and no mesh decimation or polygon reduction was performed following STL export. Automatic mesh-smoothing functions were deactivated to prevent additional geometric alteration beyond the predefined degradation procedures. These processing steps may introduce geometric approximation effects that are independent of the original DICOM voxel resolution. While identical export parameters were applied across all image degradation conditions to ensure methodological consistency, potential interactions between voxel enlargement, smoothing procedures, and mesh geometry cannot be fully excluded and should be considered when interpreting measurement variability. To ensure standardized comparability across all groups, each dataset was aligned in a unified vestibular orientation focused on the mandibular anterior region. Alignment was performed manually using anatomical landmarks to ensure consistent frontal orientation across datasets.

2.6. Measurement Protocol

Measurements were performed using the integrated digital measurement tool within the software. First, in the original reference dataset (Group 1), three consecutive measurements were taken on tooth 31, which exhibited no periodontal bone loss, followed by three measurements on tooth 41, which showed marked bone resorption. This procedure was repeated identically for all experimental groups.

Dentists ($n = 6$), each with a minimum of two years of clinical experience, served as examiners. Each examiner conducted all measurements independently following the same standardized protocol to ensure repeatability and inter-examiner comparability. The cemento-enamel junction (CEJ) and the most coronal alveolar bone crest were manually identified by each examiner. No automated landmark detection was used. No permanent landmark markers were stored between measurement sessions to minimize recall bias.

The primary outcome parameter was the linear distance between the cemento-enamel junction and the most coronal level of the alveolar bone, measured in millimeters (mm). All measurements were performed using the digital measurement tool (Netfabb Version 2021; Autodesk Inc., San Rafael, CA, USA), which provides sub-millimeter resolution. Prior to measurement, each STL model was spatially aligned and leveled to ensure a consistent vestibular viewing orientation across all datasets. Measurements were conducted in a standardized frontal view of the mandibular anterior region, and no alternative viewing angles were permitted during data acquisition to minimize variability related to observer-dependent repositioning.

All measurements were performed at a fixed magnification level, which could be adjusted for visualization purposes but remained constant throughout each measurement session. No additional image enhancement, smoothing, or contrast manipulation was allowed during the measurement process. Each examiner performed three consecutive measurements per tooth and image-quality condition, with a minimum temporal separation of several minutes between repeated measurements to reduce recall bias. The arithmetic mean of the three repeated measurements was used for subsequent statistical analysis. Intra-examiner reliability was not calculated separately, as repeated measurements were averaged to reduce random intra-observer variability.

2.7. Statistical Analysis

Descriptive statistics were calculated for all measurement values. Normality of the data distribution was assessed using the Shapiro–Wilk test. As the data showed significant deviations from normal distribution, non-parametric statistical methods were applied. Differences between image-quality groups were evaluated using the Kruskal–Wallis H test. In cases of significant overall group effects, pairwise comparisons were performed

using the Mann–Whitney U test with Bonferroni correction to adjust for multiple testing. Inter-examiner reliability was assessed using the intraclass correlation coefficient (ICC), calculated as a two-way random-effects model with absolute agreement (ICC (2,1)). ICC values were interpreted according to established guidelines. All statistical tests were conducted at a significance level of $\alpha = 0.05$. For significant pairwise comparisons, effect sizes were calculated using $r = Z/\sqrt{N}$ to quantify the magnitude of measurement differences. Effect sizes were interpreted according to conventional thresholds proposed for non-parametric effect size r (small ≈ 0.1 , moderate ≈ 0.3 , large ≥ 0.5), analogous to Cohen's criteria for correlation coefficients [33]. For ICC estimates, 95% confidence intervals were calculated to provide precision estimates of inter-examiner reliability. Given the exploratory proof-of-concept design and the use of a single CBCT dataset, no a priori power calculation was performed. The analysis is therefore intended to identify relative measurement trends and reproducibility patterns rather than to establish definitive clinical thresholds.

The limited sample structure (six examiners and two teeth derived from a single dataset) restricts statistical power; therefore, findings should be interpreted as indicative of relative measurement trends rather than definitive clinical effect magnitudes.

3. Results

It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

A total of 180 individual measurements were analyzed (6 examiners $\times$ 5 image-quality groups $\times$ 2 teeth $\times$ 3 repetitions). All measurement data showed significant deviations from normality in the Shapiro–Wilk test ($p < 0.05$). Therefore, non-parametric statistical methods were applied throughout the analysis. The influence of the five image-quality groups on the measurement values was first assessed using the Kruskal–Wallis H test. In case of significant group effects, pairwise post hoc comparisons were conducted using the Mann–Whitney U test (Wilcoxon rank-sum equivalent) with Bonferroni correction to adjust for multiple testing. Inter-examiner agreement among the six dentists was evaluated using the intraclass correlation coefficient (ICC), calculated as a two-way random-effects model with absolute agreement (ICC (2,1)). For significant pairwise comparisons, effect sizes were calculated using $r = Z/\sqrt{N}$ to quantify the magnitude of measurement differences. Effect sizes were interpreted descriptively ($r \approx 0.1$ small, $r \approx 0.3$ moderate, $r \geq 0.5$ large) [33]. For ICC estimates, 95% confidence intervals were calculated to provide precision estimates of inter-examiner reliability.

Tooth 31—Influence of Image Degradation on Measurement Outcomes

For the periodontally healthy tooth 31, the Kruskal–Wallis test revealed a statistically significant overall group effect ($H = 11.99$, $p = 0.017$), indicating that the different types of artificial image degradation produced measurable changes in diagnostic outcomes. Post hoc analysis revealed significant pairwise differences between group 1 and group 4 ($p_{\text{adj}} = 0.0099$, $r = 0.211$) as well as between group 2 and group 4 (Table 1) ($p_{\text{adj}} = 0.0074$, $r = 0.236$). The corresponding effect sizes indicate small-to-moderate magnitude differences. All remaining pairwise comparisons were non-significant after Bonferroni correction ($p_{\text{adj}} > 0.05$).

Inter-examiner agreement for tooth 31 was generally high across most image groups. ICC values ranged from 0.882 (group 2) to 0.770 (group 4). A marked reduction in reliability was observed in group 3 (TripleVoxel), which demonstrated substantially decreased examiner agreement (ICC = 0.407).

The 95% confidence intervals for ICC values (Table 2) demonstrate varying degrees of precision across image-quality groups. Notably, the TripleVoxel condition (group 3)

showed a wide confidence interval (-0.029 to 0.503), reflecting reduced stability and higher uncertainty in examiner agreement under pronounced spatial resolution reduction.

Table 1. Post hoc Pairwise Comparisons—Tooth 31.

Comparison 1	Comparison 2	U-Statistic	<i>p</i> (adj.)	<i>r</i> (Effect Size)	Significance
Group 1	Group 2	245.5	0.0846	0.439	not significant
Group 1	Group 5	184.0	1.6879	0.227	not significant
Group 1	Group 4	282.5	0.0099	0.211	significant
Group 1	Group 3	209.0	0.8735	0.563	not significant
Group 2	Group 5	154.0	0.3397	0.196	not significant
Group 2	Group 4	286.0	0.0074	0.236	significant
Group 2	Group 3	210.0	0.8204	0.011	not significant
Group 5	Group 4	220.0	0.3792	0.016	not significant
Group 5	Group 3	174.5	1.3087	0.151	not significant
Group 4	Group 3	222.0	0.3383	0.322	not significant

Table 2. ICC-Table—Tooth 31.

Groups	ICC (2.1)	95% CI Lower	95% CI Upper
Original DICOM (group 1)	0.855	0.019	0.959
DoubleVoxel (group 2)	0.882	0.406	0.981
TripleVoxel (group 3)	0.407	-0.029	0.503
GaussianBlur (group 4)	0.770	0.045	0.858
Gaussian Blur + TripleVoxel (group 5)	0.895	0.478	0.959

Tooth 41—Influence of Image Degradation on Measurement Outcomes

For the periodontally compromised tooth 41, a significant overall effect of image-quality modification was also detected ($H = 14.90$, $p = 0.0049$). The magnitude of measurement differences was generally greater than in tooth 31, reflecting greater variability in measurement distributions compared with the healthy reference tooth. Post hoc comparisons demonstrated significant differences between group 1 and group 4 (Table 3) ($p_{\text{adj}} = 0.0086$, $r = 0.061$), between group 2 and group 4 ($p_{\text{adj}} = 0.0190$, $r = 0.260$), and between group 4 and group 3 ($p_{\text{adj}} = 0.0332$, $r = 0.489$). The effect sizes ranged from small ($r \approx 0.06$) to moderate-to-large ($r \approx 0.49$), indicating that certain degradation conditions were associated with substantial shifts in measurement distributions.

Table 3. Post hoc Pairwise Comparisons—Tooth 41.

Comparison 1	Comparison 2	U-Statistic	<i>p</i> (adj.)	<i>r</i> (Effect Size)	Significance
Group 1	Group 2	251.0	0.0590	0.143	not significant
Group 1	Group 5	188.0	1.5125	0.388	not significant
Group 1	Group 4	284.0	0.0086	0.061	significant
Group 1	Group 3	204.0	1.0529	0.375	not significant
Group 2	Group 5	162.0	0.1390	0.304	not significant
Group 2	Group 4	274.0	0.0190	0.260	significant
Group 2	Group 3	161.0	0.1530	0.302	not significant
Group 5	Group 4	248.5	0.0622	0.456	not significant
Group 5	Group 3	166.5	8.9904	0.021	not significant
Group 4	Group 3	69.0	0.0332	0.489	significant

Inter-examiner reliability for tooth 41 also varied across groups. ICC values ranged from 0.916 (group 5) to 0.497 (group 4). Notably, the combined degradation (group 5)

yielded the highest agreement, whereas group 4 alone demonstrated the lowest inter-examiner consistency. These findings indicate that different degradation conditions influence measurement variability and examiner agreement to varying degrees.

The 95% confidence intervals for ICC values (Table 4) further illustrate variability in precision. For example, the GaussianBlur condition (group 4) demonstrated a comparatively wide interval (0.054 to 0.836), indicating increased uncertainty in reproducibility estimates under isolated blurring conditions.

Table 4. ICC-Table—Tooth 41.

Groups	ICC (2.1)	95% CI Lower	95% CI Upper
Original DICOM (group 1)	0.883	0.476	0.923
DoubleVoxel (group 2)	0.754	−0.212	0.856
TripleVoxel (group 3)	0.838	−0.255	0.961
GaussianBlur (group 4)	0.497	0.054	0.836
Gaussian Blur + TripleVoxel (group 5)	0.916	0.391	0.954

Summary of Statistical Findings

The statistical analyses demonstrate that image degradation conditions are associated with measurable differences in periodontal bone-level measurements and variability across groups. Significant group effects in the Kruskal–Wallis tests confirm that certain degradation conditions influence measurement distributions. Effect size calculations provide additional context regarding the magnitude of these differences, ranging from small to moderate-to-large effects depending on the comparison.

ICC results indicate that image quality conditions also affect measurement reproducibility. While original and moderately degraded datasets maintained relatively high inter-examiner agreement, pronounced voxel enlargement was associated with reduced reliability and wider confidence intervals.

Boxplot Description: Tooth 31

For tooth 31 (Figure 6), the boxplot shows differences in the distribution of measured distances across the five image-quality conditions. The original dataset (group 1) displays the highest median values (approximately 8.4–8.6 mm) and a comparatively narrow interquartile range (IQR $\approx$ 7.0–9.0 mm). Group 2 condition shows a lower median (approximately 7.3–7.5 mm) and a wider IQR ($\approx$ 5.8–8.1 mm). The group 5 condition presents a median of approximately 7.6–7.8 mm with a broad IQR ($\approx$ 6.3–8.7 mm). Group 4 condition shows a median similar to the group 5 condition ($\approx$ 7.6–7.8 mm) with a moderately sized IQR ($\approx$ 7.3–8.1 mm) and one low-value outlier. The group 3 condition exhibits the lowest median values ($\approx$ 6.5–6.7 mm) and a comparatively narrow IQR ($\approx$ 6.3–7.5 mm). Across all conditions, differences are observed in median values, spread, and the presence of isolated outliers.

Boxplot Description: Tooth 41

For tooth 41 (Figure 7), the boxplot demonstrates greater variability across image-quality conditions. The group 5 condition shows the highest median values (approximately 5.8–6.0 mm) and the widest interquartile range (IQR $\approx$ 4.2–6.4 mm). The group 4 condition presents the lowest median values ($\approx$ 4.0–4.2 mm) and the narrowest IQR ($\approx$ 3.8–4.3 mm). The group 3 condition shows median values of approximately 4.9–5.0 mm with a broader IQR ($\approx$ 4.2–5.3 mm) and several high-value outliers exceeding approximately 7 mm. The group 2 condition exhibits a median of approximately 4.6–4.8 mm with a moderate IQR ($\approx$ 3.7–4.8 mm). Compared with tooth 31, the distributions for tooth 41 show wider interquartile ranges and a higher frequency of outliers across multiple conditions.

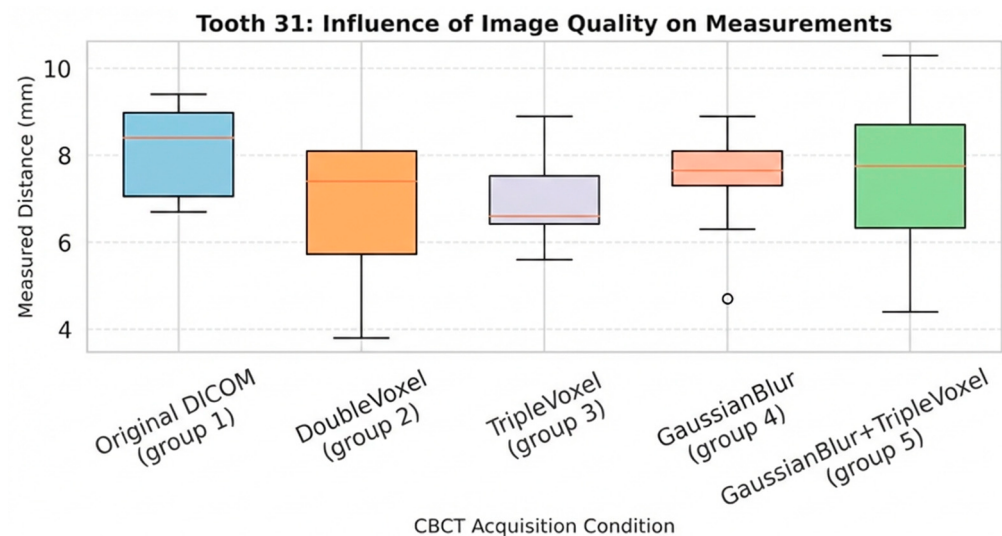


Figure 6. Boxplot representation of periodontal bone-level measurements (mm) for Tooth 31 across five image conditions (group 1–5). Boxes represent interquartile range (IQR), horizontal lines indicate medians, whiskers represent range excluding outliers. Statistical analysis was performed using the Kruskal–Wallis H test with Bonferroni-adjusted Mann–Whitney U post hoc comparisons. The significance level was set at $\alpha = 0.05$.

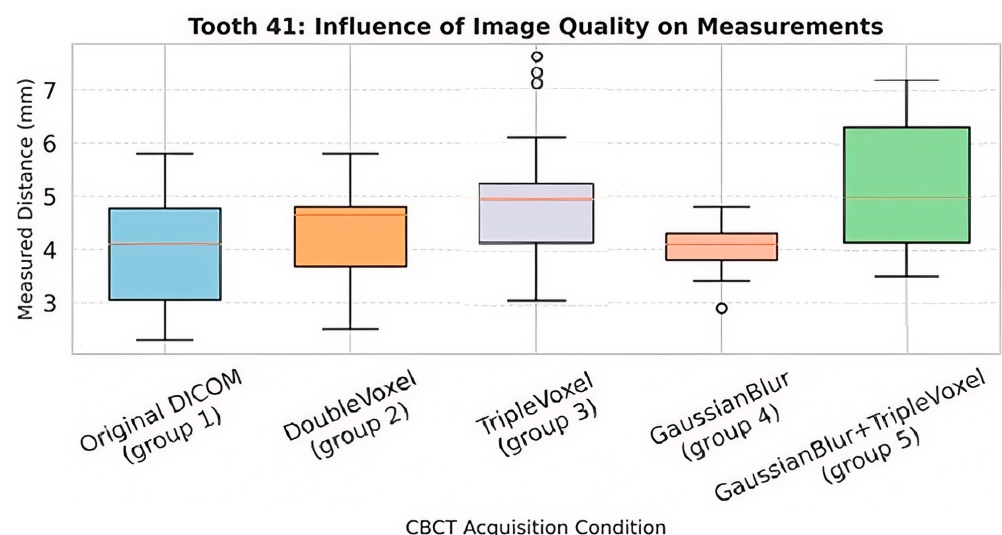


Figure 7. Boxplot representation of periodontal bone-level measurements (mm) for Tooth 41 across five image conditions (group 1–5). Boxes represent interquartile range (IQR), horizontal lines indicate medians, whiskers represent range excluding outliers. Statistical analysis was performed using the Kruskal–Wallis H test with Bonferroni-adjusted Mann–Whitney U post hoc comparisons. The significance level was set at $\alpha = 0.05$.

4. Discussion

The present study examined how systematically degraded CBCT datasets influence examiner-derived measurement outcomes of periodontal bone levels rather than diagnostic accuracy, with a specific emphasis on the implications for both clinical decision-making and the development of AI-based diagnostic systems. The results demonstrate that reductions in spatial resolution and the introduction of artificial blur substantially affect measurement outcomes, reproducibility, and examiner agreement. These effects are consistent with a growing body of literature showing that CBCT image quality is a major determinant of measurement reliability for periodontal assessments [4,5,8]. Importantly, the present analysis extends prior findings by quantifying not only statistical significance but also the

magnitude of measurement shifts using effect size calculations (r), thereby providing a more nuanced interpretation of how strongly specific degradation conditions influence examiner-derived outcomes. In particular, this study highlights how voxel enlargement and blurring filters impair measurement precision, thereby underscoring the need to understand CBCT acquisition parameters not only in clinical contexts but also when generating training data for artificial intelligence algorithms, including repetitive measurements and inter-observer.

A central observation of this study is the significant influence of voxel size on measurement outcomes, which is consistent with previous research reporting that increasing voxel dimensions reduces spatial resolution and may impair the detectability of fine periodontal structures [2,3,6]. In the present analysis, datasets with doubled or tripled voxel size showed measurable differences in bone-level assessments compared with the reference condition. While statistically significant differences were observed for specific image-quality comparisons, other voxel-related changes were characterized by consistent directional shifts in measured values without reaching statistical significance, indicating systematic tendencies rather than uniform effects. The incorporation of effect size estimates (r) revealed that certain significant comparisons were associated with small-to-moderate effects (e.g., $r \approx 0.21$ – 0.24 for Tooth 31), whereas others demonstrated moderate-to-large magnitude shifts (e.g., $r \approx 0.49$ for the comparison between Group 4 and Group 3 in Tooth 41). This differentiation is critical, as statistical significance alone does not necessarily reflect clinical or methodological relevance. These findings are in line with *ex vivo* and clinical studies demonstrating that voxel size represents an important determinant of CBCT measurement accuracy, although in the present study, no reference standard was available to assess absolute measurement correctness [28,30,34]. Notably, the magnitude and consistency of voxel-related effects were not uniform across all conditions, with more pronounced differences observed for the periodontally compromised tooth (41). This observation corresponds with prior evidence suggesting that sites with irregular defect morphology exhibit increased sensitivity to image-quality variations in CBCT-based periodontal assessment [1,35].

The influence of blur—whether Gaussian or median filtering—further underscores the sensitivity of CBCT interpretation to noise and artifact simulation. Gaussian blurring reduced sharpness in a manner that sometimes led to underestimation of bone levels, particularly in tooth 41, while the median filter produced systematic deviations in tooth 31. Effect size analysis demonstrated that while some statistically significant blur-related comparisons exhibited only small magnitude effects (e.g., $r \approx 0.06$ in Tooth 41), others reached moderate-to-large ranges ($r \approx 0.49$), indicating that the impact of blurring is not uniform but condition-dependent. These observations should be interpreted as relative measurement shifts between image degradation conditions rather than definitive evidence of reduced diagnostic accuracy. These findings align with studies suggesting that post-processing or reconstruction filters can introduce clinically relevant distortions that alter linear measurements [12,29]. Furthermore, prior investigations into metal artifact reduction and noise suppression algorithms have similarly demonstrated that image manipulation may paradoxically introduce measurement bias [24,25].

The inter-examiner agreement (ICC values) observed in this study provides insight into how image degradation conditions influence measurement reproducibility. While the reference and moderately degraded datasets under image degradation conditions maintained high reliability consistent with published data [3,28], severely degraded image groups—particularly those with tripled voxel sizes—yielded substantially lower ICC values. The addition of 95% confidence intervals for ICC estimates allows a more precise evaluation of reliability stability. Wide confidence intervals observed in certain groups (e.g., TripleVoxel in Tooth 31: -0.029 to 0.503) indicate substantial uncertainty in reproducibility estimates under pronounced spatial resolution reduction. Similarly, the GaussianBlur

condition in Tooth 41 demonstrated a broad interval (0.054 to 0.836), reflecting variability in examiner agreement under isolated smoothing conditions. These findings are consistent with previous reports indicating that voxel enlargement may have a stronger impact on reproducibility than other degradation conditions [7,36]. It is important to note that ICC values reflect inter-examiner agreement and do not provide direct information about measurement correctness or diagnostic validity. High agreement may therefore indicate consistency among examiners without necessarily reflecting the structural fidelity of the underlying image data. Interestingly, the combined degradation (Group 5) maintained high ICC values for Tooth 41. This suggests that certain degradation patterns may lead to more homogeneous examiner responses even when central measurement tendencies differ between groups. Such findings align with prior observations that subjective interpretability does not always correlate directly with quantitative measurement behavior [37].

Given the exploratory proof-of-concept design of this investigation, the clinical implications of these findings should be interpreted cautiously. Periodontal diagnosis often relies on subtle radiographic indicators, and even small deviations in bone measurements can shift clinical staging or alter treatment planning [37]. Although several comparisons yielded statistically significant differences, the observed effect sizes ranged from small to moderate in most conditions, suggesting that not all statistically detectable shifts necessarily translate into clinically meaningful diagnostic deviations. The present data suggest that lower-resolution CBCT scans may influence measurement consistency, particularly in anatomically compromised sites; however, no direct assessment of clinical misclassification was performed. Considering the continued expansion of CBCT usage in periodontics, clinicians should be cautious when interpreting scans with low voxel resolution or signs of excessive noise reduction. Prior best-evidence guidelines have similarly warned that CBCT should only be used when intraoral radiographs are insufficient, particularly due to concerns over resolution adequacy [34,37].

Beyond clinical diagnostics, this study has significant implications for artificial intelligence. AI-based systems for periodontal bone detection, staging, and volumetric quantification, such as convolutional neural networks (CNNs) and landmark-based models, depend heavily on the quality of their training data [9,16,20]. The present findings demonstrate that controlled image degradation conditions can systematically shift examiner-derived reference measurements, thereby introducing structured variability into potential ground-truth annotations. Effect size quantification further indicates that some degradation conditions may produce only minimal measurement displacement, whereas others generate moderate-to-large distributional shifts that could materially influence AI model calibration. Low-quality or inconsistent CBCT datasets may produce algorithms with poor generalizability, increased false-positive or false-negative rates, or an inability to detect nuanced bone changes. Studies have shown that AI performance is directly correlated with image clarity, voxel consistency, and noise characteristics [17,38]. Furthermore, research into enhanced-resolution AI systems confirms that deep learning models can exploit subtle texture cues that become inaccessible in datasets under image degradation conditions [18]. The present findings, therefore, provide a methodological foundation for future studies investigating how human measurement variability under controlled image degradation conditions may translate into differences in AI training robustness.

Artificial degradation studies such as the present one serve an important role in building robust AI pipelines. By characterizing precisely how degradation affects diagnostic outputs, researchers can simulate realistic low-quality conditions for data augmentation, an established technique for improving model resilience [39,40]. However, excessive degradation may distort anatomical morphology beyond clinically realistic scenarios and should therefore be applied cautiously in AI dataset generation. Several previous studies empha-

size the necessity of standardizing voxel sizes and acquisition protocols in AI research, to avoid confounding effects that might inflate or suppress algorithmic performance [41,42].

This study also aligns with concerns raised in the literature about the consequences of STL conversion and mesh resolution reduction during CBCT post-processing. Prior work has demonstrated that triangulation steps can introduce geometric inaccuracies, especially when voxel resolution is low [43,44]. In the present study, identical STL export parameters were applied across all image degradation conditions; nevertheless, interactions between voxel enlargement, smoothing procedures, and mesh reconstruction may represent an additional source of geometric approximation that should be considered when interpreting measurement variability.

Despite its contributions, this study has several limitations. First, it uses a single CBCT dataset, which restricts generalizability; however, this allowed for systematic manipulation under tightly controlled conditions. Similar methodological approaches have been used in foundational CBCT accuracy studies [31]. Second, although examiner variation was analyzed, automated segmentation tools were not directly evaluated. Future research should compare human and AI performance under identical degradation conditions to identify whether AI systems are more or less sensitive to resolution loss than clinicians. Third, this study examined only voxel enlargement and blur, whereas real-world CBCT scans may suffer from beam hardening, motion artifacts, or metal artifacts, all of which may interact with voxel parameters in unpredictable ways [25,27]. Fourth, although STL-based linear measurements were used for consistency, future studies should analyze raw volumetric CBCT data directly to minimize transformation-related errors. No a priori power calculation was performed due to the exploratory design; therefore, the findings should be interpreted as indicative of relative trends rather than definitive clinical thresholds. Furthermore, the limited sample structure (six examiners and two teeth derived from a single CBCT dataset) restricts the statistical power of the analysis. Although repeated measurements increase internal consistency, the findings should be interpreted as indicative of relative trends rather than definitive effect magnitudes. The limited sample size may reduce sensitivity to detect smaller between-group differences, and therefore, non-significant findings should not be interpreted as evidence of equivalence.

Clinically, the findings reinforce the necessity of selecting CBCT acquisition protocols that balance radiation exposure with diagnostic needs. While low-dose or large-voxel protocols can reduce patient exposure, they also risk compromising diagnostic evaluability—especially in early or mild periodontal lesions. This echoes prior recommendations that clinicians must tailor CBCT parameters to the anatomical target and specific diagnostic question [13,14,32].

From an AI perspective, this study underscores the importance of curating high-quality training datasets and documenting acquisition parameters thoroughly. AI research in dentistry increasingly depends on multicenter data collection, making cross-platform voxel standardization critical [19,40]. The present results do not directly evaluate AI performance but provide an experimental framework for future investigations linking human measurement variability and algorithmic robustness under standardized image degradation conditions.

In conclusion, the present study demonstrates that image degradation conditions in CBCT are associated with measurable effects on periodontal bone measurement outcomes and examiner reproducibility. The integration of effect size analysis and ICC confidence intervals provides a more comprehensive understanding of both the magnitude and precision of these measurement shifts. These findings should be interpreted within the exploratory scope of the study and do not establish absolute diagnostic validity. Future investigations should expand this work by including diverse patient datasets, incorporating real-world

artifact patterns, and directly comparing human and AI performance under controlled image degradation conditions.

5. Conclusions

This study demonstrates that controlled image degradation conditions of CBCT image quality through voxel enlargement and simulated blur are associated with measurable differences in examiner-derived periodontal bone-level measurements under the conditions of this exploratory investigation. Both spatial resolution and noise-related sharpness influence not only the absolute measurement values but also the reproducibility of examiner-based evaluations. The significant group effects and variability patterns observed in the Kruskal–Wallis, post hoc, and ICC analyses indicate relative measurement shifts and changes in inter-examiner agreement, particularly in periodontally compromised sites, without establishing absolute diagnostic accuracy or validity. Effect size analysis further demonstrated that the magnitude of these measurement shifts ranged from small to moderate-to-large, depending on the degradation condition and tooth morphology, indicating that image quality alterations may exert quantitatively variable impacts on examiner-derived outcomes.

Importantly, the findings highlight that image quality is not solely a technical acquisition parameter but represents a determinant of measurement reliability within the experimental framework applied here. Strong image degradation was associated with increased variability and directional measurement shifts, which may influence measurement consistency, especially when assessing subtle or early bone changes. However, the present study did not directly evaluate clinical misclassification or diagnostic correctness, and therefore, clinical implications should be interpreted as hypothesis-generating rather than definitive. Given the proof-of-concept design based on a single standardized CBCT dataset, the findings should be interpreted as indicative of relative measurement behavior under controlled degradation conditions rather than as definitive evidence of clinical performance thresholds.

From an AI perspective, the study provides methodological insights into the role of imaging fidelity in the development and training of diagnostic algorithms. Artificial intelligence systems depend on accurate, high-quality ground-truth datasets, and image degradation, whether inherent or artificially introduced, may influence the consistency of human-derived reference labels used for training purposes. Systematic degradation modeling, as applied in the present investigation, may therefore serve as a structured framework for assessing how imaging variability propagates into annotation variability, which represents a critical factor in AI model calibration and generalizability. The present study does not directly evaluate AI performance but establishes an experimental framework for subsequent validation studies.

The results emphasize that maintaining appropriate CBCT image quality may contribute to more consistent periodontal measurement behavior by human examiners. Future studies should expand on these findings using larger, more diverse datasets, incorporate real-world artifact patterns, and directly compare human and algorithmic performance under identical degradation conditions. Advancing this line of research may support the development of more standardized acquisition protocols and more methodologically robust CBCT-based assessment strategies in both clinical diagnostics and AI-assisted periodontal analysis.

Author Contributions: All authors have made substantial contributions to the present work, have approved the submitted version, and agree to be personally accountable for their own contributions. Conceptualization, M.M. and C.v.S.; Methodology, M.M., J.v.S. and T.v.S.; Validation, J.P.T., J.v.S. and T.v.S.; Formal Analysis, M.M., V.Z. and J.P.T.; Investigation, M.M. and V.Z.; Resources, C.v.S. and M.M.;

Data Curation, M.M.; Writing—Original Draft Preparation, M.M.; Writing—Review and Editing, V.Z., J.P.T. and C.v.S.; Visualization, M.M., J.v.S. and T.v.S.; Supervision, C.v.S.; Project Administration, C.v.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was reviewed and approved by the official Ethics Committee for Scientific Integrity and Ethics of Danube Private University (DPU), Krems-Stein, Austria (Protocol No. DPU-EK/084; approval date: 25 July 2024). This committee is the institutional ethics board responsible for overseeing research conducted at Danube Private University. The study did not involve any animal subjects.

Informed Consent Statement: According to the formal assessment of the Ethics Committee for Scientific Integrity and Ethics of Danube Private University (DPU), the use of fully anonymized retrospective CBCT imaging data did not require informed consent from the patient.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CBCT	Cone-Beam Computed Tomography
CEJ	Cemento-Enamel Junction
CI	Confidence Interval
CNN	Convolutional Neural Network
CTDI	Computed Tomography Dose Index
DAP	Dose–Area Product
DICOM	Digital Imaging and Communications in Medicine
FOV	Field of View
ICC	Intraclass Correlation Coefficient
IQR	Interquartile Range
MAR	Metal Artifact Reduction
STL	Standard Tessellation Language
σ	Sigma (standard deviation parameter of Gaussian filter)
2D	Two-Dimensional
3D	Three-Dimensional

References

1. Braun, X.; Ritter, L.; Jervøe-Storm, P.-M.; Frentzen, M. Diagnostic Accuracy of CBCT for Periodontal Lesions. *Clin. Oral Investig.* **2014**, *18*, 1229–1236. [[CrossRef](#)]
2. Vandenberghe, B.; Jacobs, R.; Yang, J. Detection of Periodontal Bone Loss Using Digital Intraoral and Cone Beam Computed Tomography Images: An in Vitro Assessment of Bony and/or Infrabony Defects. *Dentomaxillofac. Radiol.* **2008**, *37*, 252–260. [[CrossRef](#)]
3. Maret, D.; Telmon, N.; Peters, O.A.; Lepage, B.; Treil, J.; Inglessè, J.M.; Peyre, A.; Kahn, J.L.; Sixou, M. Effect of Voxel Size on the Accuracy of 3D Reconstructions with Cone Beam CT. *Dentomaxillofac. Radiol.* **2012**, *41*, 649–655. [[CrossRef](#)] [[PubMed](#)]
4. Menezes, C.C.; Janson, G.; da Silveira Massaro, C.; Cambiaghi, L.; Garib, D.G. Precision, Reproducibility, and Accuracy of Bone Crest Level Measurements of CBCT Cross Sections Using Different Resolutions. *Angle Orthod.* **2016**, *86*, 535–542. [[CrossRef](#)]
5. Spin-Neto, R.; Gotfredsen, E.; Wenzel, A. Impact of Voxel Size Variation on CBCT-Based Diagnostic Outcome in Dentistry: A Systematic Review. *J. Digit. Imaging* **2013**, *26*, 813–820. [[CrossRef](#)]
6. Kolsuz, M.E.; Bagis, N.; Orhan, K.; Avsever, H.; Demiralp, K.Ö. Comparison of the Influence of FOV Sizes and Different Voxel Resolutions for the Assessment of Periodontal Defects. *Dentomaxillofac. Radiol.* **2015**, *44*, 20150070. [[CrossRef](#)]

7. Mukhia, N.; Birur, N.P.; Shubhasini, A.R.; Shubha, G.; Keerthi, G. Dimensional Measurement Accuracy of 3-Dimensional Models from Cone Beam Computed Tomography Using Different Voxel Sizes. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2021**, *132*, 361–369. [\[CrossRef\]](#)
8. Liedke, G.S.; da Silveira, H.E.D.; da Silveira, H.L.D.; Dutra, V.; de Figueiredo, J.A.P. Influence of Voxel Size in the Diagnostic Ability of Cone Beam Tomography to Evaluate Simulated External Root Resorption. *J. Endod.* **2009**, *35*, 233–235. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Chang, H.-J.; Lee, S.-J.; Yong, T.-H.; Shin, N.-Y.; Jang, B.-G.; Kim, J.-E.; Huh, K.-H.; Lee, S.-S.; Heo, M.-S.; Choi, S.-C.; et al. Deep Learning Hybrid Method to Automatically Diagnose Periodontal Bone Loss and Stage Periodontitis. *Sci. Rep.* **2020**, *10*, 7531. [\[CrossRef\]](#)
10. Tayman, M.A.; Kamburoğlu, K.; Küçük, Ö.; Ateş, F.S.Ö.; Günhan, M. Comparison of Linear and Volumetric Measurements Obtained from Periodontal Defects by Using Cone Beam-CT and Micro-CT: An in Vitro Study. *Clin. Oral Investig.* **2019**, *23*, 2235–2244. [\[CrossRef\]](#)
11. Dikici, A.S.; Mihmanli, I.; Kilic, F.; Ozkok, A.; Kuyumcu, G.; Sultan, P.; Samanci, C.; Halit Yilmaz, M.; Rafiee, B.; Tamcelik, N.; et al. In Vivo Evaluation of the Biomechanical Properties of Optic Nerve and Peripapillary Structures by Ultrasonic Shear Wave Elastography in Glaucoma. *Iran. J. Radiol.* **2016**, *13*, e36849. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Pauwels, R.; Araki, K.; Siewerdsen, J.H.; Thongvigitmanee, S.S. Technical Aspects of Dental CBCT: State of the Art. *Dentomaxillofac. Radiol.* **2015**, *44*, 20140224. [\[CrossRef\]](#)
13. Mai, H.-N.; Lee, D.-H. Effects of Exposure Parameters and Voxel Size for Cone-Beam Computed Tomography on the Image Matching Accuracy with an Optical Dental Scan Image: An In Vitro Study. *BioMed Res. Int.* **2021**, *2021*, 6971828. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Ihlis, R.L.; Kadesjö, N.; Tsilingaridis, G.; Benchimol, D.; Shi, X.Q. Image Quality Assessment of Low-Dose Protocols in Cone Beam Computed Tomography of the Anterior Maxilla. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2022**, *133*, 483–491. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Krois, J.; Ekert, T.; Meinhold, L.; Golla, T.; Kharbot, B.; Wittemeier, A.; Dörfer, C.; Schwendicke, F. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. *Sci. Rep.* **2019**, *9*, 8495. [\[CrossRef\]](#)
16. Danks, R.P.; Bano, S.; Orishko, A.; Tan, H.J.; Moreno Sancho, F.; D'Aiuto, F.; Stoyanov, D. Automating Periodontal Bone Loss Measurement via Dental Landmark Localisation. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 1189–1199. [\[CrossRef\]](#)
17. Alotaibi, G.; Awawdeh, M.; Farook, F.F.; Aljohani, M.; Aldhafiri, R.M.; Aldhoayan, M. Artificial Intelligence (AI) Diagnostic Tools: Utilizing a Convolutional Neural Network (CNN) to Assess Periodontal Bone Level Radiographically—A Retrospective Study. *BMC Oral Health* **2022**, *22*, 399. [\[CrossRef\]](#)
18. Moran, M.; Faria, M.; Giraldo, G.; Bastos, L.; Conci, A. Do Radiographic Assessments of Periodontal Bone Loss Improve with Deep Learning Methods for Enhanced Image Resolution? *Sensors* **2021**, *21*, 1013. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Revilla-León, M.; Gómez-Polo, M.; Barmak, A.B.; Inam, W.; Kan, J.Y.K.; Kois, J.C.; Akal, O. Artificial Intelligence Models for Diagnosing Gingivitis and Periodontal Disease: A Systematic Review. *J. Prosthet. Dent.* **2023**, *130*, 816–824. [\[CrossRef\]](#)
20. Wajer, R.; Wajer, A.; Kazimierczak, N.; Wilamowska, J.; Serafin, Z. The Impact of AI on Metal Artifacts in CBCT Oral Cavity Imaging. *Diagnostics* **2024**, *14*, 1280. [\[CrossRef\]](#)
21. Park, H.S.; Jeon, K.; Seo, J.K. Deep Learning-Based Artefact Reduction in Low-Dose Dental Cone Beam Computed Tomography with High-Attenuation Materials. *Philos. Trans. A Math. Phys. Eng. Sci.* **2025**, *383*, 20240045. [\[CrossRef\]](#)
22. Meto, A.; Halilaj, G. The Integration of Cone Beam Computed Tomography, Artificial Intelligence, Augmented Reality, and Virtual Reality in Dental Diagnostics, Surgical Planning, and Education: A Narrative Review. *Appl. Sci.* **2025**, *15*, 6308. [\[CrossRef\]](#)
23. Bechara, B.; McMahan, C.; Geha, H.; Noujeim, M. Evaluation of a Cone Beam CT Artefact Reduction Algorithm. *Dentomaxillofac. Radiol.* **2012**, *41*, 422–428. [\[CrossRef\]](#)
24. Saati, S.; Eskandarloo, A.; Falahi, A.; Tapak, L.; Hekmat, B. Evaluation of the Efficacy of the Metal Artifact Reduction Algorithm in the Detection of a Vertical Root Fracture in Endodontically Treated Teeth in Cone-Beam Computed Tomography Images: An in Vitro Study. *Dent. Med. Probl.* **2019**, *56*, 357–363. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Bezerra, I.S.Q.; Neves, F.S.; Vasconcelos, T.V.; Ambrosano, G.M.B.; Freitas, D.Q. Influence of the Artefact Reduction Algorithm of Picasso Trio CBCT System on the Diagnosis of Vertical Root Fractures in Teeth with Metal Posts. *Dentomaxillofac. Radiol.* **2015**, *44*, 20140428. [\[CrossRef\]](#)
26. Pauwels, R.; Stamatakis, H.; Bosmans, H.; Bogaerts, R.; Jacobs, R.; Horner, K.; Tsiklakis, K.; SEDENTEXCT Project Consortium. Quantification of Metal Artifacts on Cone Beam Computed Tomography Images. *Clin. Oral Implant. Res.* **2013**, *24*, 94–99. [\[CrossRef\]](#)
27. Damstra, J.; Fourie, Z.; Huddleston Slater, J.J.R.; Ren, Y. Accuracy of Linear Measurements from Cone-Beam Computed Tomography-Derived Surface Models of Different Voxel Sizes. *Am. J. Orthod. Dentofac. Orthop.* **2010**, *137*, 16.e1–16.e6. [\[CrossRef\]](#)
28. Maroua, A.L.; Ajaj, M.; Hajeer, M.Y. The Accuracy and Reproducibility of Linear Measurements Made on CBCT-Derived Digital Models. *J. Contemp. Dent. Pract.* **2016**, *17*, 294–299. [\[CrossRef\]](#)

29. de Oliveira, V.G.B.; Queiroz, P.M.; Simões, A.R.; Alves, M.G.O.; Jardini, M.A.N.; Costa, A.L.F.; Lopes, S.L.P.d.C. Voxel Size and Field of View Influence on Periodontal Bone Assessment Using Four CBCT Systems: An Experimental Ex Vivo Analysis. *Tomography* **2025**, *11*, 74. [[CrossRef](#)] [[PubMed](#)]
30. Brüllmann, D.; Schulze, R.K.W. Spatial Resolution in CBCT Machines for Dental/Maxillofacial Applications—What Do We Know Today? *Dentomaxillofac. Radiol.* **2015**, *44*, 20140204. [[CrossRef](#)]
31. Liljeholm, R.; Kadesjö, N.; Benchimol, D.; Hellén-Halme, K.; Shi, X.-Q. Cone-Beam Computed Tomography with Ultra-Low Dose Protocols for Pre-Implant Radiographic Assessment: An in Vitro Study. *Eur. J. Oral Implantol.* **2017**, *10*, 351–359. [[PubMed](#)]
32. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Reprint; Psychology Press: New York, NY, USA, 2009.
33. Icen, M.; Orhan, K.; Şeker, Ç.; Geduk, G.; Cakmak Özlü, F.; Cengiz, M.İ. Comparison of CBCT with Different Voxel Sizes and Intraoral Scanner for Detection of Periodontal Defects: An in Vitro Study. *Dentomaxillofac. Radiol.* **2020**, *49*, 20190197. [[CrossRef](#)]
34. Koç, C.; Sönmez, G.; Yılmaz, F.; Karahan, S.; Kamburoğlu, K. Comparison of the Accuracy of Periapical Radiography with CBCT Taken at 3 Different Voxel Sizes in Detecting Simulated Endodontic Complications: An Ex Vivo Study. *Dentomaxillofac. Radiol.* **2018**, *47*, 20170399. [[CrossRef](#)]
35. Bagis, N.; Eren, H.; Kolsuz, M.E.; Kurt, M.H.; Avsever, H.; Orhan, K. Comparison of the Burr and Chemically Induced Periodontal Defects Using Different Field-of-View Sizes and Voxel Resolutions. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2018**, *125*, 260–267. [[CrossRef](#)]
36. Kim, D.M.; Bassir, S.H. When Is Cone-Beam Computed Tomography Imaging Appropriate for Diagnostic Inquiry in the Management of Inflammatory Periodontitis? An American Academy of Periodontology Best Evidence Review. *J. Periodontol.* **2017**, *88*, 978–998. [[CrossRef](#)]
37. Miller, A.; Huang, C.; Brody, E.R.; Siqueira, R.C.E.; Credit. Artificial Intelligence Applications for the Radiographic Detection of Periodontal Disease: A Scoping Review. *J. Calif. Dent. Assoc.* **2023**, *51*, 2206301. [[CrossRef](#)]
38. Kurt Bayrakdar, S.; Orhan, K.; Bayrakdar, I.S.; Bilgir, E.; Ezhov, M.; Gusarev, M.; Shumilov, E. A Deep Learning Approach for Dental Implant Planning in Cone-Beam Computed Tomography Images. *BMC Med. Imaging* **2021**, *21*, 86. [[CrossRef](#)]
39. Ezhov, M.; Gusarev, M.; Golitsyna, M.; Yates, J.M.; Kushnerev, E.; Tamimi, D.; Aksoy, S.; Shumilov, E.; Sanders, A.; Orhan, K. Clinically Applicable Artificial Intelligence System for Dental Diagnosis with CBCT. *Sci. Rep.* **2021**, *11*, 15006. [[CrossRef](#)] [[PubMed](#)]
40. Alrashed, S.; Dutra, V.; Chu, T.-M.G.; Yang, C.-C.; Lin, W.-S. Influence of Exposure Protocol, Voxel Size, and Artifact Removal Algorithm on the Trueness of Segmentation Utilizing an Artificial-Intelligence-Based System. *J. Prosthodont.* **2024**, *33*, 574–583. [[CrossRef](#)] [[PubMed](#)]
41. Tsoromokos, N.; Parinussa, S.; Claessen, F.; Moin, D.A.; Loos, B.G. Estimation of Alveolar Bone Loss in Periodontitis Using Machine Learning. *Int. Dent. J.* **2022**, *72*, 621–627. [[CrossRef](#)]
42. Elbashti, M.; Molinero-Mourelle, P.; Aswehlee, A.; Bornstein, M.M.; Abou-Ayash, S.; Schimmel, M.; Ella, B.; Naveau, A. Effect of Triangular Mesh Resolution on the Geometrical Trueness of Segmented CBCT Maxillofacial Data into STL Format. *J. Dent.* **2023**, *138*, 104722. [[CrossRef](#)] [[PubMed](#)]
43. Zhan, L.-P.; Gao, S.-Y.; Su, S.; Jia, X.-T.; He, C.; Zhang, Q.; Huang, X.-F. Influence of Cone-Beam Computed Tomography Voxel Size on the Accuracy of Periodontal Ligament Surface Area Measurements. *J. Craniofac. Surg.* **2025**, *36*, 1175–1179. [[CrossRef](#)] [[PubMed](#)]
44. Noujeim, M.; Prihoda, T.; Langlais, R.; Nummikoski, P. Evaluation of High-Resolution Cone Beam Computed Tomography in the Detection of Simulated Interradicular Bone Lesions. *Dentomaxillofac. Radiol.* **2009**, *38*, 156–162. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.