

## Article

# Automatized Detection of Periodontal Bone Loss on Periapical Radiographs by Vision Transformer Networks

Helena Dujic <sup>1,\*</sup> , Ole Meyer <sup>2</sup>, Patrick Hoss <sup>1</sup>, Uta Christine Wölfle <sup>1</sup>, Annika Wülk <sup>1</sup>, Theresa Meusbürger <sup>1</sup>, Leon Meier <sup>1</sup>, Volker Gruhn <sup>2</sup>, Marc Hesenius <sup>2</sup>, Reinhard Hickel <sup>1</sup> and Jan Kühnisch <sup>1,\*</sup> 

<sup>1</sup> Department of Conservative Dentistry and Periodontology, LMU University Hospital, LMU Munich, 80336 Munich, Germany; patrick.hoss@t-online.de (P.H.); uwoelfle@dent.med.uni-muenchen.de (U.C.W.); annika.wuelk@gmx.de (A.W.); theresa.meusbuerger@hotmail.com (T.M.); meier.leon@gmail.com (L.M.); hickel@dent.med.uni-muenchen.de (R.H.)

<sup>2</sup> Institute for Software Engineering, University of Duisburg-Essen, 45127 Essen, Germany; ole.meyer@uni-due.de (O.M.); volker.gruhn@paluno.uni-due.de (V.G.); marc.hesenius@uni-due.de (M.H.)

\* Correspondence: h.dujic@med.uni-muenchen.de (H.D.); jkuehn@dent.med.uni-muenchen.de (J.K.); Tel.: +49-89-4400-59395 (H.D.)

**Abstract:** Several artificial intelligence-based models have been presented for the detection of periodontal bone loss (PBL), mostly using convolutional neural networks, which are the state of the art in deep learning. Given the emerging breakthrough of transformer networks in computer vision, we aimed to evaluate various models for automatized PBL detection. An image data set of 21,819 anonymized periapical radiographs from the upper/lower and anterior/posterior regions was assessed by calibrated dentists according to PBL. Five vision transformer networks (ViT-base/ViT-large from Google, BEiT-base/BEiT-large from Microsoft, DeiT-base from Facebook/Meta) were utilized and evaluated. Accuracy (ACC), sensitivity (SE), specificity (SP), positive/negative predictive value (PPV/NPV) and area under the ROC curve (AUC) were statistically determined. The overall diagnostic ACC and AUC values ranged from 83.4 to 85.2% and 0.899 to 0.918 for all evaluated transformer networks, respectively. Differences in diagnostic performance were evident for lower (ACC 94.1–96.7%; AUC 0.944–0.970) and upper anterior (86.7–90.2%; 0.948–0.958) and lower (85.6–87.2%; 0.913–0.937) and upper posterior teeth (78.1–81.0%; 0.851–0.875). In this study, only minor differences among the tested networks were detected for PBL detection. To increase the diagnostic performance and to support the clinical use of such networks, further optimisations with larger and manually annotated image data sets are needed.

**Keywords:** artificial intelligence; deep learning; machine learning; transformer; periapical radiographs; periodontitis; periodontal bone loss; diagnostics



**Citation:** Dujic, H.; Meyer, O.; Hoss, P.; Wölfle, U.C.; Wülk, A.; Meusbürger, T.; Meier, L.; Gruhn, V.; Hesenius, M.; Hickel, R.; et al. Automatized Detection of Periodontal Bone Loss on Periapical Radiographs by Vision Transformer Networks. *Diagnostics* **2023**, *13*, 3562. <https://doi.org/10.3390/diagnostics13233562>

Academic Editor: Francesco Inchingolo

Received: 26 October 2023

Revised: 18 November 2023

Accepted: 27 November 2023

Published: 29 November 2023



**Copyright:** © 2023 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) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Periodontitis is a chronic inflammatory disease of the supporting dental tissues and affects a relevant proportion of the world's population [1–4]. Furthermore, periodontitis can also be associated with various risk factors such as smoking and stress, as well as systemic diseases such as diabetes mellitus or pulmonary diseases. Clinically, periodontitis is associated with periodontal bone loss (PBL), tooth loosening and tooth loss. All of these factors can further impair functionality, aesthetics and quality of life [5,6]. Considering the recommendations of the latest workshop on the classification of periodontal diseases [7,8], the initial diagnosis is primarily based on clinical assessment, bleeding on probing, repeated measurements of clinical attachment loss and probing pocket depth. The early manifestations of periodontitis are only clinically recognisable. Furthermore, staging based on the radiographic assessment of PBL is considered possible only with the progression of the disease. As a result, the importance of radiographs increases as the disease progresses,

since the extent of alveolar bone changes can be visualized more accurately [9,10]. However, a reliable assessment of PBL remains susceptible to diagnostic subjectivity among dentists [11,12]. Therefore, the use of image analysis tools based on artificial intelligence (AI) methods could possibly enable the automated assessment of PBL on radiographs and potentially improve diagnostic accuracy. Interestingly, several research groups have developed AI-based algorithms and published promising results on panoramic [11,13–21] and periapical radiographs [12,22–30]. Looking at the methodology of the studies published so far, almost all research groups have used an image set of a limited size to train different types of convolutional neural networks (CNNs). This has led to heterogeneous but promising results [31,32]. In particular, more than half of the studies published to date have reported a data set of less than 1000 X-ray images [12,14–19,21,26,29,30]. In addition, some studies used different exclusion criteria for their data set, meaning that radiographs with a specific tooth group or radiographs with caries or root canal treatment were excluded (e.g., [23,28]). In addition, variability in the architecture of the CNNs used can be observed, e.g., ResNet, U-Net and faster R-CNNs were trained for PBL detection [12,13,15,17–19,25]. Accurate manual annotation also contributed significantly to the reported results, as studies reporting the annotation of radiologic features of PBL described a better diagnostic performance, e.g., [13,25]. Moreover, none of the previously mentioned studies used recently introduced transformer networks for computer vision tasks, which are the most recent available technology for automatized image analysis and may possibly outperform current CNNs in the future [33]. On the one hand, CNNs have proven their value in tasks such as image classification and segmentation by efficiently processing large data sets. Among the most significant advantages is the ability of CNNs to recognize local patterns, such as edges or shapes. This proved to be particularly helpful for recognizing features in dental X-rays, such as tooth decay, different tooth shapes, etc. On the other hand, the vision transformer's attention mechanism allows the model to learn the correlation of parts of the image that may not be in direct proximity. In the case of PBL detection, these are primarily the cemento-enamel junction, alveolar bone and apex, as well as other anatomical structures relevant for the evaluation. Notably, transformer networks usually require a larger amount of training data compared to CNNs. Following this, we aimed to compare the diagnostic performance of five different transformer networks for automatized PBL detection on periapical radiographs. Specifically, it was hypothesized that the diagnostic performance of the included transformer networks would be similar and that an overall diagnostic accuracy of 90% would be achievable.

## 2. Materials and Methods

### 2.1. Study Design

The Ethics Committee of the Medical Faculty of Ludwig Maximilian University (LMU) of Munich approved this study protocol (project number 020-798). The periapical radiographs used in this study were anonymized and obtained as part of previous clinical examinations. Consequently, we could not identify any of the patients and were therefore unable to obtain written informed consent. The reporting of this research followed the Standard for Reporting of Diagnostic Accuracy Studies (STARD) Steering Committee recommendations [34] as well as the recommendations for reporting AI studies in dentistry [35].

### 2.2. Periapical Radiographs

This study used anonymized periapical radiographs (Figure 1). All X-rays were taken at the Department of Conservative Dentistry and Periodontology (LMU University Hospital) and different dental practices. To ensure a high-quality image sample, exclusion criteria were previously defined. This involved excluding distorted radiographs, radiographs with overlapping teeth, radiographs with artifacts, and radiographs with incompletely imaged teeth for which an assessment of the periodontium was not possible. Furthermore, radiographs with implants, with endodontic treatments or photographed radiographs, were

also excluded. Further exclusion criteria were not defined. All periapical radiographs were stored in .jpg format and processed without downsizing the original resolution. Altogether, 21,819 periapical radiographs, divided into upper/lower anterior and posterior teeth, were selected for this study (Table 1). The majority of the radiographs were upper ( $N = 9461$ ) and lower posterior teeth ( $N = 8425$ ), outnumbering upper ( $N = 1944$ ) and lower anterior teeth ( $N = 1989$ ). Additionally, the radiographs were categorized according to PBL.



**Figure 1.** Examples of periapical radiographs for all categories: healthy periodontium (Score 0), mild radiographic periodontal bone loss (PBL) up to 15% of the root length (Score 1), moderate radiographic PBL between 15% and 33% of the root length (Score 2), and severe radiographic PBL extending to the mid-third of the root and beyond (Score 3).

**Table 1.** Overview of the included periapical radiographs ( $N = 21,819$ ) in relation to the corresponding regions and categories of periodontal bone loss.

Region of Periapical Radiograph	Healthy Periodontium (Score 0)	Mild PBL (Score 1)	Moderate PBL (Score 2)	Severe PBL (Score 3)	Total (N)
1st Quadrant	1701 (35.8%)	1826 (38.5%)	851 (18.0%)	367 (7.7%)	4745
2nd Quadrant	1231 (26.1%)	2080 (44.1%)	1093 (23.2%)	312 (6.6%)	4716
3rd Quadrant	1477 (34.7%)	2033 (47.7%)	593 (13.9%)	157 (3.7%)	4260
4th Quadrant	1282 (30.8%)	2027 (48.7%)	713 (17.1%)	143 (3.4%)	4165
Maxillary anteriors	653 (33.6%)	661 (34.0%)	433 (22.3%)	197 (10.1%)	1944
Mandibular anteriors	202 (10.2%)	676 (34.0%)	786 (39.5%)	325 (16.3%)	1989

### 2.3. Categorisation of Periodontal Bone Loss (Reference Standard)

All radiographs were precategorized by a group of graduate dentists (P.H., T.M., A.W. and L.M.) and later independently counterchecked by experienced examiners (H.D., U.W. and J.K.). For each of the periapical radiographs, a diagnosis was made by differentiating between healthy teeth and teeth affected by mild, moderate or severe PBL [7,8]. Clinical data were not available prior to decision making. In detail, the following diagnostic criteria were applied: 0—radiographic PBL not detectable; 1—mild radiographic PBL up to 15% of the root length; 2—moderate radiographic PBL between 15% and 33% of the root length; and 3—severe radiographic PBL extending to the mid-third of the root and beyond (Figure 1). In the case of divergent opinions, each radiograph was discussed until consensus was reached. Each dichotomized diagnostic decision (0 versus 1 to 3)—one per image—served as a reference standard for the cyclic training and repeated evaluation of the deep learning-based transformer network.

Before conducting this study, all participating dentists were trained during a 2-day workshop by the principal investigator (J.K.). Following this workshop, the effectiveness of training was determined during a calibration course. The inter- and intra-examiner reproducibility for PBL were assessed on 150 periapical radiographs. The corresponding Kappa values showed substantial reliability, ranging from 0.454 to 0.482 (inter-examiner). The intra-examiner reliability in terms of Cohen's Kappa amounted to 0.739 [36].

### 2.4. Training of the Deep Learning-Based Transformer Networks (Test Method)

A pipeline of well-established methods was used to train the transformer networks. In principle, the entire image set of 21,819 periapical radiographs was divided into a training set ( $N = 18,819$ ) and a test set. The latter included 3000 randomly selected X-rays from the overall image set and served as an independent test set that was not included in the model training. Given the high number of periapical radiographs in our data set, image augmentation and preprocessing were not necessary. Furthermore, all X-rays had a standardized size.

The previously mentioned data set was used to train five different pre-trained transformer networks (Table 2) [33,37,38]. The learning performance was evaluated with the independent test set. The used transformer networks were trained by using backpropagation to determine the gradient for learning. Furthermore, the model training was accelerated by the use of Floating Point 16 and a university-based computer (i9 10850K  $10 \times 3.60$  GHz, Intel Corp., Santa Clara, CA, USA) equipped with 64 GB RAM and a professional graphic card (RTX A6000 48 GB (Nvidia, Santa Clara, CA, USA)). The batch size amounted to 16 randomly selected images. Each transformer was trained over 5 epochs with cross entropy loss as an error function and an application of the Adam optimizer (Betas 0.9 and 0.999, Epsilon  $\times 10^{-8}$ ).

**Table 2.** Model characteristics of the used transformer networks.

	ViT-Base (Google)	ViT-Large (Google)	BEiT-Base (Microsoft)	BEiT-Large (Microsoft)	DeiT-Base (Facebook/Meta)
Neural network	Vision transformer		Bidirectional encoder representation from image transformers		Data-efficient image transformer
Epochs	5	5	5	5	5
Learning rate	0.00005	0.00005	0.00005	0.00005	0.00005
FLOS	$7.280 \times 10^{15}$	$25.735 \times 10^{15}$	$7.277 \times 10^{15}$	$25.744 \times 10^{15}$	$7.280 \times 10^{15}$
Samples per second	298.6	111.7	274.4	102.9	298.5
Parameter count	$85.8 \times 10^6$	$303.3 \times 10^6$	$85.7 \times 10^6$	$303.4 \times 10^6$	$85.8 \times 10^6$

### 2.5. Statistical Analysis

The data were analysed using Python (version 3.8.5, <http://www.python.org> accessed on 28 November 2023). The diagnostic ACC was determined by calculating the number of

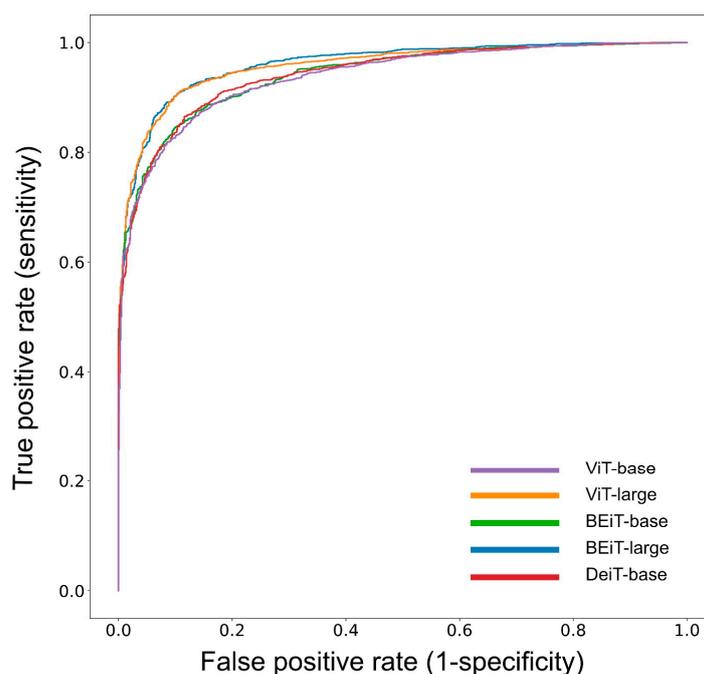
true negatives (TN), true positives (TP), false positives (FP) and false negatives (FN). In addition, the sensitivity (SE), specificity (SP), positive/negative predictive values (PPV/NPV) and area under the receiver operating characteristic (ROC) curve were calculated [39].

### 3. Results

In the present study, we calculated the diagnostic performance for automatized PBL detection on periapical radiographs for lower/upper and anterior/posterior teeth altogether (Table 3) and separately (Table 4) by using five different transformer networks. In general, when analysing the whole data set of periapical radiographs, the ACC ranged from 83.4% to 85.2%; the corresponding AUC values ranged from 0.899 to 0.918 (Figure 2). The detailed data analysis revealed generally better performance data for mandibular teeth than for maxillary teeth (Table 4). Here, the ACC ranged from 94.1% to 96.7% for mandibular anteriors and from 85.6% to 87.2% for mandibular posteriors. The corresponding data for maxillary anterior and posterior teeth varied between 86.7% and 90.2% as well as between 78.1% and 81.0%, respectively. Additionally, the AUC values tended to be similar or better for mandibular teeth (Table 4). Furthermore, the SE values were consistently higher than the SP values.

**Table 3.** Overview of the overall diagnostic performance of the five transformer neuronal networks where the independent test set ( $N = 3000$  radiographs) was evaluated by the AI-based algorithm for the assessment of periodontal bone loss. Diagnostic accuracy (ACC), sensitivity (SE), specificity (SP), negative predictive value (NPV), positive predictive value (PPV) and area under the receiver operating characteristic curve (AUC) were calculated for all types of teeth.

All Apical Radiographs	True Positive (TP)		True Negative (TN)		False Positive (FP)		False Negative (FN)		Diagnostic Performance					
	N	%	N	%	N	%	N	%	ACC	SE	SP	NPV	PPV	AUC
ViT-base	1884	62.8	673	22.4	230	7.7	213	7.1	85.2	89.8	74.5	76.0	89.1	0.918
ViT-large	1831	61.0	671	22.4	232	7.7	266	8.9	83.4	87.3	74.3	71.6	88.8	0.899
BEiT-base	1885	62.8	649	21.6	254	8.5	212	7.1	84.5	89.9	71.9	75.4	88.1	0.914
BEiT-large	1914	63.8	631	21.0	272	9.1	183	6.1	84.8	91.3	69.9	77.5	87.6	0.907
DeiT-base	1879	62.6	646	21.5	257	8.6	218	7.3	84.2	89.6	71.5	74.8	88.0	0.908



**Figure 2.** The receiver operating characteristic (ROC) curves illustrate the diagnostic performance of five different transformer networks for PBL detection.

**Table 4.** Overview of the diagnostic performance of the five transformer neuronal networks for mandibular and maxillary anterior and posterior teeth. Accuracy (ACC), sensitivity (SE), specificity (SP), negative predictive value (NPV), positive predictive value (PPV) and area under the receiver operating characteristic curve (AUC) were calculated.

		True Positive (TP)		True Negative (TN)		False Positive (FP)		False Negative (FN)		Diagnostic Performance					
		N	%	N	%	N	%	N	%	ACC	SE	SP	NPV	PPV	AUC
<b>Mandibular anterior teeth</b>	ViT-base	240	88.2	16	5.9	9	3.3	7	2.6	94.1	97.2	64.0	69.6	96.4	0.944
	ViT-large	241	88.6	18	6.6	7	2.6	6	2.2	95.2	97.6	72.0	75.0	97.2	0.960
	BEiT-base	242	89.0	21	7.7	4	1.5	5	1.8	96.7	98.0	84.0	80.8	98.4	0.963
	BEiT-large	245	90.1	18	6.6	7	2.6	2	0.7	96.7	99.2	72.0	90.0	97.2	0.952
	DeiT-base	242	89.0	15	5.5	10	3.7	5	1.8	94.5	98.0	60.0	75.0	96.0	0.970
<b>Mandibular posterior teeth</b>	ViT-base	700	61.6	287	25.3	78	6.9	70	6.2	87.0	90.9	78.6	80.4	90.0	0.937
	ViT-large	687	60.5	285	25.1	80	7.1	83	7.3	85.6	89.2	78.1	77.4	89.6	0.913
	BEiT-base	704	62.0	277	24.4	88	7.8	66	5.8	86.4	91.4	75.9	80.8	88.9	0.933
	BEiT-large	711	62.6	279	24.6	86	7.6	59	5.2	87.2	92.3	76.4	82.5	89.2	0.923
	DeiT-base	694	61.1	281	24.8	84	7.4	76	6.7	85.9	90.1	77.0	78.7	89.2	0.927
<b>Maxillary anterior teeth</b>	ViT-base	157	59.5	81	30.7	18	6.8	8	3.0	90.2	95.2	81.8	91.0	89.7	0.958
	ViT-large	156	59.1	77	29.2	22	8.3	9	3.4	88.3	94.5	77.8	89.5	87.6	0.948
	BEiT-base	158	59.8	73	27.7	26	9.8	7	2.7	87.5	95.8	73.7	91.3	85.9	0.954
	BEiT-large	157	59.5	73	27.7	26	9.8	8	3.0	87.1	95.2	73.7	90.1	85.8	0.954
	DeiT-base	154	58.3	75	28.4	24	9.1	11	4.2	86.7	93.3	75.8	87.2	86.5	0.954
<b>Maxillary posterior teeth</b>	ViT-base	787	59.2	289	21.8	125	9.4	128	9.6	81.0	86.0	69.8	69.3	86.3	0.875
	ViT-large	747	56.2	291	21.9	123	9.3	168	12.6	78.1	81.6	70.3	63.4	85.9	0.851
	BEiT-base	781	58.8	278	20.9	136	10.2	134	10.1	79.7	85.4	67.1	67.5	85.2	0.865
	BEiT-large	801	60.3	261	19.6	153	11.5	114	8.6	79.9	87.5	63.0	69.6	84.0	0.861
	DeiT-base	789	59.4	275	20.7	139	10.4	126	9.5	80.1	86.2	66.4	68.6	85.0	0.860

When comparing the metrics of the included transformer networks, only minor differences appeared in the results (Tables 3 and 4). However, the ACC and AUC values were found to be high in all scenarios, and SE was higher than SP.

#### 4. Discussion

The present study aimed to compare the diagnostic performance of five different transformer networks for automatized PBL detection on periapical radiographs. Depending on the applied network, the overall diagnostic ACC and AUC values ranged from 83.4% to 85.2% and 0.899 to 0.918, respectively (Table 3, Figure 2). On the one hand, the ACC values must be evaluated as high; on the other hand, the hypothesized overall diagnostic ACC of 90% was not achieved. Therefore, the initially formulated hypothesis must be rejected.

When comparing the documented diagnostic performance data (Tables 3 and 4) with data from the literature, the following conclusion can be drawn. In general, the majority of comparable studies presented model performances in the same or lower order of magnitude [11–15,17,20,21,23,26,28,40], whereas only a few studies registered above-average values [25,41]. In detail, Lee et al. [25] reported an ACC for staging that ranged from 88% to 99%. They further stated that the ACC for periodontitis case classification was 85%. Specifically, 693 periapical radiographs were independently annotated by examiners prior to training the model, indicating regions of interest such as the alveolar bone, presence of teeth, cemento-enamel junctions and presence of restorations. In addition, a further 644 periapical radiographs were used to assess the ACC of the model. In another study on staging, Widyaningrum et al. [41] stated that the detection rate was 95%, with the best performance shown for stage 4 periodontitis. Although the data set consisted of only 100 panoramic radiographs, two investigators annotated the previously mentioned radiographs before training the CNN. Accurate annotations were made by marking the alveolar ridge and the alveolar bone surrounding the teeth. In addition, the examiners added a number indicating the stage of periodontitis. Therefore, the few studies with better diagnostic performance seem remarkable compared to other studies with results of a lower magnitude. Here, other dental detection tasks should also be mentioned in comparison, where a higher ACC—typically approximately 90%—was usually registered with a similar methodology, e.g., in the detection of caries or periapical lesions on radiographs (e.g., [42–44]) and the detection of clinical pathologies or restorations on intraoral photographs (e.g., [45–49]). This may indicate that automatized PBL detection is more difficult to accomplish, which is supported by the fact that PBL characteristics are usually spread over the whole radiographic image and can have varying extents.

Our study revealed differences in the performance of the model in relation to the analysed group of teeth. In principle, automatized PBL detection performed better for mandibular teeth than for maxillary teeth, and better for anterior teeth compared to posterior teeth (Table 4). Only a few studies have considered this aspect thus far, e.g., by the exclusion of periapical radiographs with upper anterior and posterior teeth or by the inclusion of anterior teeth only [23,26]. To avoid the influence of data inconsistencies on the results of the trained CNN, Tsoromokos et al. [26] only considered periapical radiographs of the mandible and reported a data set with 446 radiographs. In addition, Alotaibi et al. [23] considered 1724 periapical radiographs of maxillary and mandibular anterior teeth only and excluded radiographs of teeth that had been restored with full crowns or root canal treatments, as well as radiographs of teeth that had undergone apical surgery with root resection. In this context, the study by Lee et al. [28] should also be mentioned, which included periapical radiographs of posterior teeth to identify periodontally compromised premolars and molars. Further exclusion criteria were root canal treatment and teeth with fully restorative crowns as well as moderate to severe caries and teeth with a shape deviating from the usual anatomical structure. When considering the data shown in Table 4, it must be concluded that the partial exclusion of periapical radiographs may bias the model's performance and limit the generalisability of the data shown. As is reasonable for this finding, the anatomical structures in the upper jaw in relation to the intraoral projec-

tion technique must be considered. Interestingly, this issue can be obviously downsized when using panoramic X-rays [15]. Nevertheless, a well-balanced inclusion of periapical radiographs from different groups of teeth may be relevant and should be implemented in future studies.

In this study, five well-established open-source transformer networks were trained: ViT-base and ViT-large from Google, BEiT-base and BEiT-large from Microsoft, and DeiT-base from Facebook/Meta [33,37,38]. The main differences between the transformer networks are in their size, training strategy and fine-tuning approach. “Base” and “large” models differ in size and computational complexity, whereby “large” models have more parameters. During training, ViTs process images as a sequence of patches and use an attention mechanism to learn the overall correlations within images. DeiT can achieve a high performance even with limited training data. Here, a smaller model learns to imitate a larger, already pre-trained model and benefits from a large data set without directly using it. In contrast, BEiT is trained in a two-stage process: pre-training on a large data set to capture general visual features, followed by fine-tuning for specific tasks. Transformer networks have rarely been applied for computer vision tasks in dentistry and not specifically for the detection of PBL. So far, only three studies using transformer networks were published; however, none of them focused on PBL assessment in periapical radiographs [50–52]. Nevertheless, there have been studies in which CNNs were used for PBL detection on periapical and panoramic radiographs (e.g., [11,14,15,17,21–26,40]). Here, the majority of investigations used only a low to moderate number of radiographs for model development, and most studies on periapical radiographs included a maximum of a few thousand images [13,22,23,25,27,28]. In contrast, Kim et al. [20] annotated the PBL in an extensive set of 12,179 panoramic radiographs, which may have potentially enhanced the internal study strength. The reported model-dependent AUC values ranged from 0.92 to 0.95 [20], which were slightly higher than the results from our study setup (Table 2). Therefore, it can be argued that the chosen study setup produced comparable data in the moment, which in part might be attributed to the use of transformer networks. Interestingly, we observed similar performance data with each of the included transformer networks. There was a minor tendency for less-complex transformer networks, e.g., Google’s vision transformer/base, to perform better than their more complex counterparts (Tables 2 and 3). However, further improvements might be possible, especially by employing exact annotations in a large image set. Such features could enable precise object segmentation [20].

This study has several strengths and limitations. From a methodological point of view, this study used a large and well-balanced set of periapical radiographs ( $N = 21,819$ ) in which all X-rays were diagnosed by dental professionals following the latest recommendations for PBL assessment [7,8]. Another unique feature seems to be the comparison of five transformer networks for the detection of PBL on periapical radiographs, as no other studies with the same methodology could be identified. In addition, the following limitations must be taken into account. In this study, we used categorical diagnostic scoring per image only. In detail, this means that the exact areas of PBL on periapical radiographs remained unmarked, which can be interpreted as a limitation. The exact annotation must be understood as a crucial feature to localize PBL precisely on X-rays. The exact annotation of the pathological structures would require the detection, classification and segmentation of PBL on each radiographic image. In particular, the marking of pathological segments must be understood as a time-consuming procedure that needs to be addressed in future projects. Another limitation is that only periapical radiographs were examined in this study and that panoramic radiographs have not been considered so far. However, in view of the fact that both types of radiographs are commonly used to assess PBL, but the format, size and radiographic anatomy differ, a separate analysis was justified. In addition, no clinical information was available for the anonymized radiographs in this study. Another limitation might be that we did not include any other transformer networks or CNNs in this study.

## 5. Conclusions

From the results of this study, it can be concluded that it was possible to achieve good diagnostic performance for automatized PBL detection when using a large set of periapical radiographs and several transformer networks. However, it can be hypothesized that the model performance can be improved by using exact annotations.

**Author Contributions:** Conceptualisation, project administration and supervision J.K.; study design, H.D., J.K., O.M., M.H., V.G. and R.H.; investigation, H.D., P.H., L.M., T.M., A.W., U.C.W. and J.K.; transformer network training and statistical analysis, O.M. and M.H.; writing—original draft preparation, H.D., J.K. and P.H. All authors contributed equally to the interpretation of data and reviewed, edited and approved the final manuscript version. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** This study was approved by the Ethics Committee of the Medical Faculty of the LMU Munich (project number 020-798, approved on 8 October 2020).

**Informed Consent Statement:** Procedures used in studies with human participants were all in accordance with the ethical standards of the institutional and/or national research committee and the 1964 Helsinki Declaration and its subsequent amendments or comparable ethical standards.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no potential conflicts of interest with respect to the authorship and publication of this article.

## References

1. Nazir, M.; Al-Ansari, A.; Al-Khalifa, K.; Alhareky, M.; Gaffar, B.; Almas, K. Global Prevalence of Periodontal Disease and Lack of Its Surveillance. *Sci. World J.* **2020**, *2020*, 2146160. [[CrossRef](#)] [[PubMed](#)]
2. Frencken, J.E.; Sharma, P.; Stenhouse, L.; Green, D.; Laverty, D.; Dietrich, T. Global epidemiology of dental caries and severe periodontitis—A comprehensive review. *J. Clin. Periodontol.* **2017**, *44* (Suppl. 18), S94–S105. [[CrossRef](#)]
3. Tonetti, M.S.; Jepsen, S.; Jin, L.; Otomo-Corgel, J. Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: A call for global action. *J. Clin. Periodontol.* **2017**, *44*, 456–462. [[CrossRef](#)]
4. Kassebaum, N.J.; Bernabe, E.; Dahiya, M.; Bhandari, B.; Murray, C.J.; Marcenes, W. Global burden of severe periodontitis in 1990–2010: A systematic review and meta-regression. *J. Dent. Res.* **2014**, *93*, 1045–1053. [[CrossRef](#)] [[PubMed](#)]
5. Papapanou, P.N.; Susin, C. Periodontitis epidemiology: Is periodontitis under-recognized, over-diagnosed, or both? *Periodontol. 2000* **2017**, *75*, 45–51. [[CrossRef](#)] [[PubMed](#)]
6. Petersen, P.E.; Ogawa, H. The global burden of periodontal disease: Towards integration with chronic disease prevention and control. *Periodontol. 2000* **2012**, *60*, 15–39. [[CrossRef](#)]
7. Papapanou, P.N.; Sanz, M.; Buduneli, N.; Dietrich, T.; Feres, M.; Fine, D.H.; Flemmig, T.F.; Garcia, R.; Giannobile, W.V.; Graziani, F.; et al. Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J. Periodontol.* **2018**, *89* (Suppl. 1), S173–S182. [[CrossRef](#)]
8. Tonetti, M.S.; Greenwell, H.; Kornman, K.S. Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. *J. Periodontol.* **2018**, *89* (Suppl. 1), S159–S172. [[CrossRef](#)]
9. Tonetti, M.S.; Sanz, M. Implementation of the new classification of periodontal diseases: Decision-making algorithms for clinical practice and education. *J. Clin. Periodontol.* **2019**, *46*, 398–405. [[CrossRef](#)] [[PubMed](#)]
10. Fiorellini, J.P.; Sourvanos, D.; Sarimento, H.; Karimbux, N.; Luan, K.W. Periodontal and Implant Radiology. *Dent. Clin. N. Am.* **2021**, *65*, 447–473. [[CrossRef](#)]
11. Kong, Z.; Ouyang, H.; Cao, Y.; Huang, T.; Ahn, E.; Zhang, M.; Liu, H. Automated periodontitis bone loss diagnosis in panoramic radiographs using a bespoke two-stage detector. *Comput. Biol. Med.* **2023**, *152*, 106374. [[CrossRef](#)] [[PubMed](#)]
12. 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](#)] [[PubMed](#)]
13. Kabir, T.; Lee, C.T.; Chen, L.; Jiang, X.; Shams, S. A comprehensive artificial intelligence framework for dental diagnosis and charting. *BMC Oral Health* **2022**, *22*, 480. [[CrossRef](#)] [[PubMed](#)]
14. Ertas, K.; Pence, I.; Cesmeli, M.S.; Ay, Z.Y. Determination of the stage and grade of periodontitis according to the current classification of periodontal and peri-implant diseases and conditions (2018) using machine learning algorithms. *J. Periodontol. Implant Sci.* **2022**, *53*, 38. [[CrossRef](#)]

15. Jiang, L.; Chen, D.; Cao, Z.; Wu, F.; Zhu, H.; Zhu, F. A two-stage deep learning architecture for radiographic staging of periodontal bone loss. *BMC Oral Health* **2022**, *22*, 106. [[CrossRef](#)] [[PubMed](#)]
16. Zadrozny, L.; Regulski, P.; Brus-Sawczuk, K.; Czajkowska, M.; Parkanyi, L.; Ganz, S.; Mijiritsky, E. Artificial Intelligence Application in Assessment of Panoramic Radiographs. *Diagnostics* **2022**, *12*, 224. [[CrossRef](#)] [[PubMed](#)]
17. Li, H.; Zhou, J.; Zhou, Y.; Chen, Q.; She, Y.; Gao, F.; Xu, Y.; Chen, J.; Gao, X. An Interpretable Computer-Aided Diagnosis Method for Periodontitis From Panoramic Radiographs. *Front. Physiol.* **2021**, *12*, 655556. [[CrossRef](#)]
18. Thanathornwong, B.; Suebnukarn, S. Automatic detection of periodontal compromised teeth in digital panoramic radiographs using faster regional convolutional neural networks. *Imaging Sci. Dent.* **2020**, *50*, 169–174. [[CrossRef](#)]
19. 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](#)]
20. Kim, J.; Lee, H.S.; Song, I.S.; Jung, K.H. DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs. *Sci. Rep.* **2019**, *9*, 17615. [[CrossRef](#)]
21. Krois, J.; Ekert, T.; Meinhold, L.; Golla, T.; Kharbot, B.; Wittemeier, A.; Dorfer, C.; Schwendicke, F. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. *Sci. Rep.* **2019**, *9*, 8495. [[CrossRef](#)] [[PubMed](#)]
22. Chen, C.C.; Wu, Y.F.; Aung, L.M.; Lin, J.C.; Ngo, S.T.; Su, J.N.; Lin, Y.M.; Chang, W.J. Automatic recognition of teeth and periodontal bone loss measurement in digital radiographs using deep-learning artificial intelligence. *J. Dent. Sci.* **2023**, *18*, 1301–1309. [[CrossRef](#)]
23. 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](#)]
24. Chang, J.; Chang, M.F.; Angelov, N.; Hsu, C.Y.; Meng, H.W.; Sheng, S.; Glick, A.; Chang, K.; He, Y.R.; Lin, Y.B.; et al. Application of deep machine learning for the radiographic diagnosis of periodontitis. *Clin. Oral Investig.* **2022**, *26*, 6629–6637. [[CrossRef](#)] [[PubMed](#)]
25. Lee, C.T.; Kabir, T.; Nelson, J.; Sheng, S.; Meng, H.W.; Van Dyke, T.E.; Walji, M.F.; Jiang, X.; Shams, S. Use of the deep learning approach to measure alveolar bone level. *J. Clin. Periodontol.* **2022**, *49*, 260–269. [[CrossRef](#)]
26. Tsromokos, 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](#)] [[PubMed](#)]
27. Chen, H.; Li, H.; Zhao, Y.; Zhao, J.; Wang, Y. Dental disease detection on periapical radiographs based on deep convolutional neural networks. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 649–661. [[CrossRef](#)]
28. Lee, J.-H.; Kim, D.-h.; Jeong, S.-N.; Choi, S.-H. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J. Periodontal Implant. Sci.* **2018**, *48*, 114–123. [[CrossRef](#)]
29. Lin, P.L.; Huang, P.Y.; Huang, P.W. Automatic methods for alveolar bone loss degree measurement in periodontitis periapical radiographs. *Comput. Methods Programs Biomed.* **2017**, *148*, 1–11. [[CrossRef](#)]
30. Lin, P.L.; Huang, P.W.; Huang, P.Y.; Hsu, H.C. Alveolar bone-loss area localization in periodontitis radiographs based on threshold segmentation with a hybrid feature fused of intensity and the H-value of fractional Brownian motion model. *Comput. Methods Programs Biomed.* **2015**, *121*, 117–126. [[CrossRef](#)]
31. Patil, S.; Joda, T.; Soffe, B.; Awan, K.H.; Fageeh, H.N.; Tovani-Palone, M.R.; Licari, F.W. Efficacy of artificial intelligence in the detection of periodontal bone loss and classification of periodontal diseases: A systematic review. *J. Am. Dent. Assoc.* **2023**, *154*, 795–804.e791. [[CrossRef](#)] [[PubMed](#)]
32. Scott, J.; Biancardi, A.M.; Jones, O.; Andrew, D. Artificial Intelligence in Periodontology: A Scoping Review. *Dent. J.* **2023**, *11*, 43. [[CrossRef](#)] [[PubMed](#)]
33. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image is Worth 16 × 16 Words: Transformers for Image Recognition at Scale. Available online: <https://arxiv.org/abs/2010.11929> (accessed on 23 October 2023).
34. Bossuyt, P.M.; Reitsma, J.B.; Bruns, D.E.; Gatsonis, C.A.; Glasziou, P.P.; Irwig, L.; Lijmer, J.G.; Moher, D.; Rennie, D.; de Vet, H.C.; et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ* **2015**, *351*, h5527. [[CrossRef](#)]
35. Schwendicke, F.; Singh, T.; Lee, J.H.; Gaudin, R.; Chaurasia, A.; Wiegand, T.; Uribe, S.; Krois, J. Artificial intelligence in dental research: Checklist for authors, reviewers, readers. *J. Dent.* **2021**, *107*, 103610. [[CrossRef](#)]
36. Meusburger, T.; Wulk, A.; Kessler, A.; Heck, K.; Hickel, R.; Dujic, H.; Kühnisch, J. The Detection of Dental Pathologies on Periapical Radiographs—Results from a Reliability Study. *J. Clin. Med.* **2023**, *12*, 2224. [[CrossRef](#)]
37. Bao, H.; Dong, L.; Piao, S.; Wei, F. BEiT: BERT Pre-Training of Image Transformers. *arXiv* **2022**, arXiv:2106.08254v2. Available online: <https://arxiv.org/abs/2106.08254> (accessed on 23 October 2023).
38. Touvron, H.; Cord, M.; Douze, M.; Massa, F.; Sablayrolles, A.; Jégou, H. Training Data-Efficient Image Transformers & Distillation through Attention. *arXiv* **2021**, arXiv:2012.12877v2. Available online: <https://arxiv.org/abs/2012.12877> (accessed on 23 October 2023).
39. Matthews, D.E.; Farewell, V.T. *Using and Understanding Medical Statistics*; S.Karger AG: Basel, Switzerland, 2015.
40. Kurt, S.; Celik, O.; Bayrakdar, I.S.; Orhan, K.; Bilgir, E.; Odabas, A.; Aslan, A.F. Success of artificial intelligence system in determining alveolar bone loss from dental panoramic radiography images. *Cumhuriyet. Dent. J.* **2020**, *23*, 318–324. [[CrossRef](#)]

41. Widyaningrum, R.; Candradewi, I.; Aji, N.; Aulianisa, R. Comparison of Multi-Label U-Net and Mask R-CNN for panoramic radiograph segmentation to detect periodontitis. *Imaging Sci. Dent.* **2022**, *52*, 383–391. [[CrossRef](#)]
42. Lian, L.; Zhu, T.; Zhu, F.; Zhu, H. Deep Learning for Caries Detection and Classification. *Diagnostics* **2021**, *11*, 1672. [[CrossRef](#)]
43. Moidu, N.P.; Sharma, S.; Chawla, A.; Kumar, V.; Logani, A. Deep learning for categorization of endodontic lesion based on radiographic periapical index scoring system. *Clin. Oral Investig.* **2022**, *26*, 651–658. [[CrossRef](#)]
44. Lee, J.H.; Kim, D.H.; Jeong, S.N.; Choi, S.H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.* **2018**, *77*, 106–111. [[CrossRef](#)] [[PubMed](#)]
45. Schonewolf, J.; Meyer, O.; Engels, P.; Schlickerrieder, A.; Hickel, R.; Gruhn, V.; Hesenius, M.; Kühnisch, J. Artificial intelligence-based diagnostics of molar-incisor-hypomineralization (MIH) on intraoral photographs. *Clin. Oral Investig.* **2022**, *26*, 5923–5930. [[CrossRef](#)] [[PubMed](#)]
46. Engels, P.; Meyer, O.; Schonewolf, J.; Schlickerrieder, A.; Hickel, R.; Hesenius, M.; Gruhn, V.; Kühnisch, J. Automated detection of posterior restorations in permanent teeth using artificial intelligence on intraoral photographs. *J. Dent.* **2022**, *121*, 104124. [[CrossRef](#)] [[PubMed](#)]
47. Kühnisch, J.; Meyer, O.; Hesenius, M.; Hickel, R.; Gruhn, V. Caries Detection on Intraoral Images Using Artificial Intelligence. *J. Dent. Res.* **2022**, *101*, 158–165. [[CrossRef](#)] [[PubMed](#)]
48. Zhang, X.; Liang, Y.; Li, W.; Liu, C.; Gu, D.; Sun, W.; Miao, L. Development and evaluation of deep learning for screening dental caries from oral photographs. *Oral Dis.* **2022**, *28*, 173–181. [[CrossRef](#)]
49. Schlickerrieder, A.; Meyer, O.; Schoenewolf, J.; Engels, P.; Hickel, R.; Gruhn, V.; Hesenius, M.; Kühnisch, J. Automated Detection and Categorization of Fissure Sealants from Intraoral Digital Photographs Using Artificial Intelligence. *Diagnostics* **2021**, *11*, 1608. [[CrossRef](#)]
50. Zhou, X.; Yu, G.; Yin, Q.; Yang, J.; Sun, J.; Lv, S.; Shi, Q. Tooth Type Enhanced Transformer for Children Caries Diagnosis on Dental Panoramic Radiographs. *Diagnostics* **2023**, *13*, 689. [[CrossRef](#)]
51. Gao, S.; Li, X.; Li, X.; Li, Z.; Deng, Y. Transformer based tooth classification from cone-beam computed tomography for dental charting. *Comput. Biol. Med.* **2022**, *148*, 105880. [[CrossRef](#)]
52. Ying, S.; Wang, B.; Zhu, H.; Liu, W.; Huang, F. Caries segmentation on tooth X-ray images with a deep network. *J. Dent.* **2022**, *119*, 104076. [[CrossRef](#)]

**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.