

Article

AI-Powered Prediction of Dental Space Maintainer Needs Using X-Ray Imaging: A CNN-Based Approach for Pediatric Dentistry

Aslıhan Yelkenci ¹ , Günseli Güven Polat ¹, Emir Oncu ²  and Fatih Ciftci ^{2,3,*} 

¹ Faculty of Dentistry, Department of Pediatric Dentistry, University of Health Sciences, Istanbul 34668, Turkey; aslihanzihni@gmail.com (A.Y.); gunseli.guvenpolat@sbu.edu.tr (G.G.P.)

² Faculty of Engineering, Department of Biomedical Engineering, Fatih Sultan Mehmet Vakıf University, Istanbul 34015, Turkey; biomedical.emr@gmail.com

³ Department of Technology Transfer Office, Fatih Sultan Mehmet Vakıf University, Istanbul 34015, Turkey

* Correspondence: fciftci@gmail.com or fciftci@fsm.edu.tr

Abstract: Space maintainers (SMs) are essential for preserving dental arch integrity after premature tooth loss. This study aimed to develop a deep learning model to predict the necessity of SMs and identify specific teeth requiring intervention. A dataset of 400 dental X-rays was preprocessed to standardize image dimensions and convert them into numerical representations for machine learning. The dataset was divided into training (80%) and testing (20%) subsets. A Convolutional Neural Network (CNN) was designed with multiple convolutional and pooling layers, followed by fully connected layers for binary classification. The model was trained using 30 epochs and evaluated with accuracy, precision, recall, F1-score, ROC AUC, and MCC. The CNN achieved 94% accuracy, with a precision of 0.93 for Class 0 (no SM needed) and 0.95 for Class 1 (SM needed). The ROC AUC was 0.94, and the MCC was 0.875, indicating strong reliability. When tested on 86 X-ray images, the model successfully identified specific teeth (showing teeth number) requiring SMs, with minimal errors. These results suggest that the proposed AI model provides high-performance predictions for SM necessity, offering a valuable decision-support tool for pediatric dentistry.

Keywords: space maintainer; convolutional neural network; X-ray imaging; machine learning; prediction



Academic Editor: Pedro Couto

Received: 24 February 2025

Revised: 19 March 2025

Accepted: 27 March 2025

Published: 3 April 2025

Citation: Yelkenci, A.; Güven Polat, G.; Oncu, E.; Ciftci, F. AI-Powered Prediction of Dental Space Maintainer Needs Using X-Ray Imaging: A CNN-Based Approach for Pediatric Dentistry. *Appl. Sci.* **2025**, *15*, 3920. <https://doi.org/10.3390/app15073920>

Copyright: © 2025 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/>).

Highlights

- Developed a CNN model for predicting the need for dental space maintainers using X-ray images.
- Achieved a high overall accuracy of 94%, with strong performance metrics including precision, recall, F1-score, ROC AUC (0.94), and MCC (0.875).
- Preprocessed a dataset of 400 dental X-rays to create standardized input data, ensuring reproducibility and efficiency in the machine learning pipeline.
- Successfully classified whether a space maintainer is needed and identified the specific tooth requiring intervention, showcasing the model's potential for personalized pediatric dental care.

1. Introduction

Primary teeth are crucial for children's overall development, providing support for speech, mastication, and aesthetics, while also maintaining space for the proper eruption of permanent teeth [1,2]. The natural physiological process involves the exfoliation of primary

teeth followed by the eruption of permanent teeth [3]. However, the premature loss of primary teeth can disrupt this process, potentially resulting in malocclusion, ectopic eruption, and midline deviation [4]. The premature loss of primary teeth is most attributed to dental caries; however, other significant contributing factors include congenital anomalies, ectopic eruption of permanent teeth, and dental trauma [5,6].

The etiology of malocclusion encompasses a complex process driven by the combined effects of genetic and environmental factors. Space loss resulting from the premature loss of primary teeth constitutes an environmental factor that may contribute to the development or exacerbation of malocclusion, ultimately increasing the need for orthodontic treatment [7]. The most effective way to prevent such issues is to ensure that primary teeth remain in the oral cavity until their natural exfoliation [8]. However, if this is not feasible, the use of SMs becomes necessary [9]. Safeguarding space following the premature loss of primary teeth has the potential to minimize or entirely negate the requirement for orthodontic intervention [10].

The selection of an appropriate SM is influenced by several critical factors, including the stage of dental development, the size of tooth loss, the number of primary teeth affected, and the degree of patient cooperation [9,11,12]. The use of fixed SMs is a well-established approach to preserving space in the dental arch following the premature loss of primary teeth [13]. Among the different types of fixed SMs, band and loop appliances are among the most utilized for maintaining arch space in cases of single-tooth loss [14].

In addition to clinical considerations, the implementation of SMs in dental practice also varies depending on dentists' clinical education, professional experience, patient acceptance of treatment, and financial considerations. Determining the necessity of SMs requires a thorough evaluation of multiple factors, with radiographs and space analysis serving as valuable tools in this process [5]. However, assessing the need for SMs using dental radiographs can be time-consuming and prone to human error.

Disease diagnosis requires clinicians to assess symptoms, interpret diagnostic test results, and consider other relevant factors [15]. However, this process can be affected by cognitive biases and reliance on memory, potentially influencing clinical judgment. Artificial intelligence (AI), trained on vast datasets, has demonstrated the ability to outperform even highly experienced specialists in certain clinical tasks [16–18]. Consequently, AI is progressively becoming an important component of contemporary healthcare, as it is in the field of medicine, with applications extending to pediatric dentistry [16,19]. In healthcare, AI is categorized into two main domains: virtual and physical. The virtual domain encompasses ML and DL [20]. Machine learning refers to a system's ability to autonomously learn from data without explicit programming [21]. It includes four primary methodologies: supervised learning, unsupervised learning, reinforcement learning, and active learning [22]. Supervised learning involves analyzing labeled input data to uncover patterns, utilizing models such as Bayesian inference, decision trees, linear discriminants, support vector machines, logistic regression, and artificial neural networks [23]. Deep learning, a more advanced subset of ML, employs multiple interconnected layers to extract features and optimize model performance [24].

AI technologies aim to develop systems and robots capable of performing tasks like pattern recognition, decision making, and adaptive problem solving—capabilities traditionally associated with human intelligence [25]. Advances in computational power, combined with innovations in machine learning techniques and neural networks, have accelerated progress in AI [26]. As a subset of AI, ML focuses on training computers to analyze large datasets, identify trends, and apply these insights for predictions or decisions [25]. AI has demonstrated transformative potential across fields such as natural language processing, autonomous vehicles, healthcare, and image recognition. In AD research, it excels at rapidly

analyzing complex datasets, identifying patterns imperceptible to humans, and providing highly accurate predictions, thereby advancing the understanding and management of the disease [27,28]. DL is centered around advanced neural network architectures, including Convolutional Neural Networks (CNNs) [29] and artificial neural networks (ANNs) [30]. CNNs are a specialized type of ANN designed to process and analyze visual data, such as images. Unlike ANNs, CNNs leverage convolutional layers that apply filters (kernels) to extract spatial and hierarchical features like edges, textures, and shapes [31].

AI is utilized in various areas of pediatric dentistry, including dental plaque detection, oral health assessment, supernumerary tooth identification, detection of early childhood caries, fissure sealant categorization, chronological age assessment, deciduous and young permanent tooth detection, ectopic eruption detection, and behavior management. Previous studies have compared the clinical success of traditional and 3D-printed space maintainers [32]. One study has evaluated the effectiveness of AI-based chatbots in providing reliable and high-quality information on space maintainers for pediatric patients and their parents [33]. A recent study developed a CNN model based on YOLOv3 to detect mesiodens in panoramic radiographs and identify specific tooth numbers, achieving high accuracy [34]. Another study has shown the effectiveness of CNNs in dental image analysis, particularly in detecting third molar angles using object detection models [35].

Unlike traditional deep learning models that focus solely on tooth type classification, such as the ZNet model proposed in [36], our CNN model not only predicts the necessity of a space maintainer (SM) but also identifies the specific tooth requiring intervention, providing a more clinically actionable solution for pediatric dentistry. This dual-function capability uniquely enhances clinical decision making by providing precise localization, which is crucial for pediatric dentists. By integrating both classification and localization, our model streamlines the treatment planning process, reducing reliance on time-consuming manual assessments. The CNN model has the potential to significantly reduce errors attributable to human oversight and may serve as a valuable resource for guiding future dental professionals [19]. This innovation bridges the gap between AI-driven dental diagnostics and real-world clinical applications, making it a valuable tool for digital dentistry.

This study aims to leverage the capabilities of AI to facilitate the automated assessment of the need for an SM through the detailed analysis of dental radiographic imagery. It is hypothesized that the AI model will accurately predict the need for a space maintainer and identify specific teeth requiring intervention, thus providing a reliable and efficient tool for pediatric dentists in clinical practice.

2. Methods

In this study, we developed an AI model to predict the need for dental SMs using X-ray images. The evaluation was conducted specifically for single-tooth loss, focusing only on a specific tooth in the radiographs, and the type of SM to be used was not assessed. The evaluation of space maintainer necessity was conducted using a dual-expert evaluation approach, considering factors such as the time elapsed since the extraction, the available space, the amount of bone covering the permanent tooth germ, the patient's dental age, and the sequential eruption pattern of the teeth. The process involved multiple stages, beginning with the preprocessing of dental X-rays into numerical data using OpenCV for compatibility with machine learning models. A CNN was then designed, trained, and tested using a labeled dataset to classify whether an SM is required and, if so, to identify the specific tooth. The methodology is detailed in the following subsections, which cover data preparation, image processing, model architecture, training and testing, and evaluation metrics. While CNNs are widely used in medical imaging, our study is innovative in applying deep learning specifically to SM prediction—an underexplored area in pediatric

dentistry. Unlike existing models that focus on diagnosing dental conditions, our approach not only determines the need for an SM but also identifies the specific tooth requiring intervention. Additionally, we implemented standardized image preprocessing and a rigorous evaluation framework. Our work contributes to preventive dentistry, offering an AI-driven tool to assist clinicians in objective decision making.

2.1. Data Preparation and Image Processing

The dataset used in this study consisted of 400 dental X-ray images, of which 195 were labeled as requiring an SM and 205 were labeled as not requiring one. The decision to use 400 images was influenced by several factors, including the need to have a balanced representation of both classes (SM needed and SM not needed) and to allow for robust model training and validation. These images were processed to standardize their format and prepare them for training and testing the artificial intelligence model. To ensure a reliable evaluation of the model's performance, the dataset was divided into training and testing subsets, with 80% of the data used for training and 20% reserved for testing. The split was performed in a stratified manner to maintain an equal representation of both labels in each subset.

The original images in the dataset are generally 2000×1000 pixels. For CNN image processing, we resized them to 100×100 pixels with three RGB color channels using Python 12.0 code. The resizing was performed using interpolation methods to maintain aspect ratio and minimize distortion. After that, each image was resized to a resolution of 100×100 pixels with three color channels (RGB format), ensuring uniformity in dimensions across the entire dataset. This resizing step was critical for the consistency required by the machine learning algorithms, as it ensured that all input data shared the same shape and scale. Following resizing, the images were flattened into one-dimensional numerical arrays. This transformation preserved the pixel intensity values while reducing the spatial complexity of the images, enabling efficient storage and processing. The flattened arrays served as a numerical representation of the images, capturing all the relevant features necessary for the learning process. The labels for the dataset were encoded numerically, with "0" representing cases where no SM was needed and "1" representing cases where an SM was required. The processed data were divided into training and testing files.

The training set, represented by "X_train", consisted of 320 flattened images (80% of the dataset), while the corresponding labels, "Y_train", comprised 165 zeros and 155 ones. Similarly, the testing set, represented by "X_test", included 80 flattened images (20% of the dataset), with "Y_test" containing 40 zeros and 40 ones. These files were stored in CSV format to ensure compatibility with the machine learning pipeline. By employing this systematic preprocessing pipeline, the study ensured that the dataset was both standardized and appropriately structured for machine learning applications. This approach not only enhanced the model's training efficiency but also improved the reproducibility of the methodology. Instead of relying solely on a separate test set, the model utilized all 400 images to predict the tooth number based on its previous classification outputs. This approach allowed us to assess the model's consistency and accuracy in identifying the correct tooth. Among these, 86 images specifically indicated the need for a space maintainer (predicting teeth number), serving as a focused subset for evaluating prediction accuracy.

2.2. AI Architecture and Performance Metrics

The predictive system for determining the need for space maintainers was built using a CNN, designed to process and analyze dental X-ray images. The architecture consisted of multiple convolutional and pooling layers, followed by fully connected layers to perform binary classification. The specific layers and their configurations were as follows:

1. Convolutional Layers:
 - The first convolutional layer consisted of 256 filters with a kernel size of 3×3 and ReLU activation. This layer extracted low-level spatial features from the input images.
 - Subsequent layers included 256 filters, also with a kernel size of 3×3 and ReLU activation. These layers progressively extracted higher-level features essential for classification.
2. Pooling Layers:
 - Max-pooling layers with a pool size of 2×2 were interspersed between convolutional layers to reduce the spatial dimensions of the feature maps, thereby controlling overfitting and improving computational efficiency.
3. Fully Connected Layers:
 - A dense layer with 256 neurons and ReLU activation was added to integrate the extracted features.
 - The final dense layer consisted of a single neuron with sigmoid activation, outputting a probability score for binary classification (0: no space maintainer needed, 1: space maintainer needed).

Figure 1 demonstrates the graphical representation of the CNN architecture, illustrating the layers and connections that enable the model to perform image classification. The model was trained using 30 epochs and a size batch of 64 to balance computational efficiency and convergence. A binary cross-entropy loss function was employed, which is well suited for binary classification tasks, and an optimizer was utilized to minimize the loss, ensuring efficient convergence during training. The model's parameters were fine-tuned to optimize its predictive capabilities and to generalize well to unseen data. The performance of the CNN model was evaluated using a comprehensive set of metrics to capture its predictive capabilities and reliability.

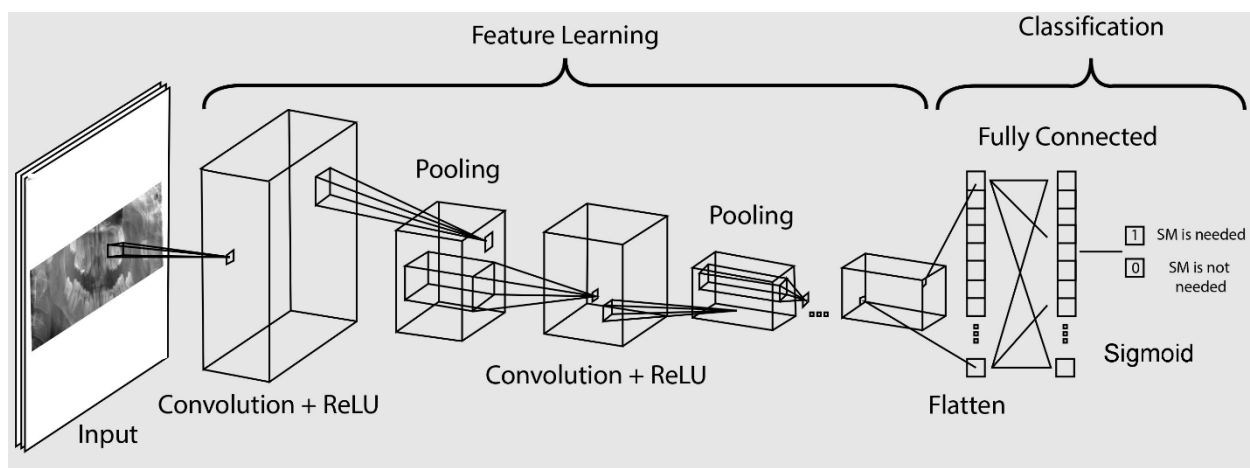


Figure 1. Graphical representation of the CNN architecture.

Table 1 shows performance metrics such as accuracy. Accuracy was used to measure the overall correctness of the model's predictions, while precision evaluated the proportion of true positive predictions among all positive predictions, indicating the model's ability to avoid false positives. Recall, or sensitivity, assessed the proportion of true positive cases detected, highlighting the model's ability to identify cases requiring space maintainers. The F1-score, a harmonic means of precision and recall, provided a balanced assessment of the model's performance. Additionally, the Receiver Operating Characteristic Area

Under Curve (ROC AUC) quantified the model's ability to distinguish between classes, with higher values indicating better discriminatory power. It is computed as the area under the ROC curve, which plots the true positive rate (TPR) against the false positive rate (FPR). The Matthews Correlation Coefficient (MCC) was also calculated to evaluate the correlation between predicted and true labels, accounting for both true and false predictions, which made it particularly useful for imbalanced datasets.

Table 1. Overview of the performance metrics used in the study and their calculation methods; TP (true positives): correctly predicted positive cases, TN (true negatives): correctly predicted negative cases, FP (false positives): incorrectly predicted positive cases, and FN (false negatives): incorrectly predicted negative cases.

Metric	Calculation
Accuracy [37]	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision [38]	$\frac{TP}{TP + FP}$
Recall [39]	$\frac{TP}{TP + FN}$
F1-Score [39]	$\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
ROC AUC [40]	$\text{TPR} = \frac{TP}{TP + FN}, \text{FPR} = \frac{FP}{FP + FN}$
Matthews Correlation Coefficient [40]	$\frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP \times FN)(TP \times FP)(TN \times FN)}}$

The comprehensive evaluation of these metrics ensured a robust assessment of the model's performance, emphasizing its clinical relevance. High precision and recall scores were critical in this application to minimize false positives, which could lead to unnecessary interventions, and false negatives, which might result in missed cases requiring intervention. The ROC AUC and MCC further validated the model's robustness, demonstrating its reliability in accurately predicting the need for space maintainers across diverse cases. By achieving high performance across these metrics, the system showed potential to assist dental professionals in making faster and more accurate clinical decisions.

3. Results

The results obtained from the experiments conducted in this study provide a comprehensive evaluation of the model's ability to predict the need for space maintainers using dental X-ray images. These results are presented to illustrate the effectiveness and reliability of the CNN in addressing the classification task. Performance metrics such as accuracy, precision, recall, F1-score, ROC AUC, and MCC were employed to assess the model's predictive capabilities. The findings highlight the model's capacity to make accurate predictions and demonstrate its potential as a valuable tool in clinical decision making for pediatric dentistry. The following sections provide a detailed analysis of these results, supported by quantitative metrics and visualizations. Figure 2 illustrates the training performance of the CNN used for predicting the need for space maintainers. The left panel shows the training accuracy over 30 epochs, with the training and validation accuracy steadily increasing as the model learns the patterns within the data. By the end of the training process, the accuracy approaches 1.0, indicating that the model has effectively captured the features necessary for accurate predictions. The right panel depicts the corresponding training loss, which consistently decreases over the epochs, reflecting the model's ability to minimize errors between predicted and actual labels. Together, these trends demonstrate

the model's capacity to learn effectively from the training data, achieving high accuracy and low loss.

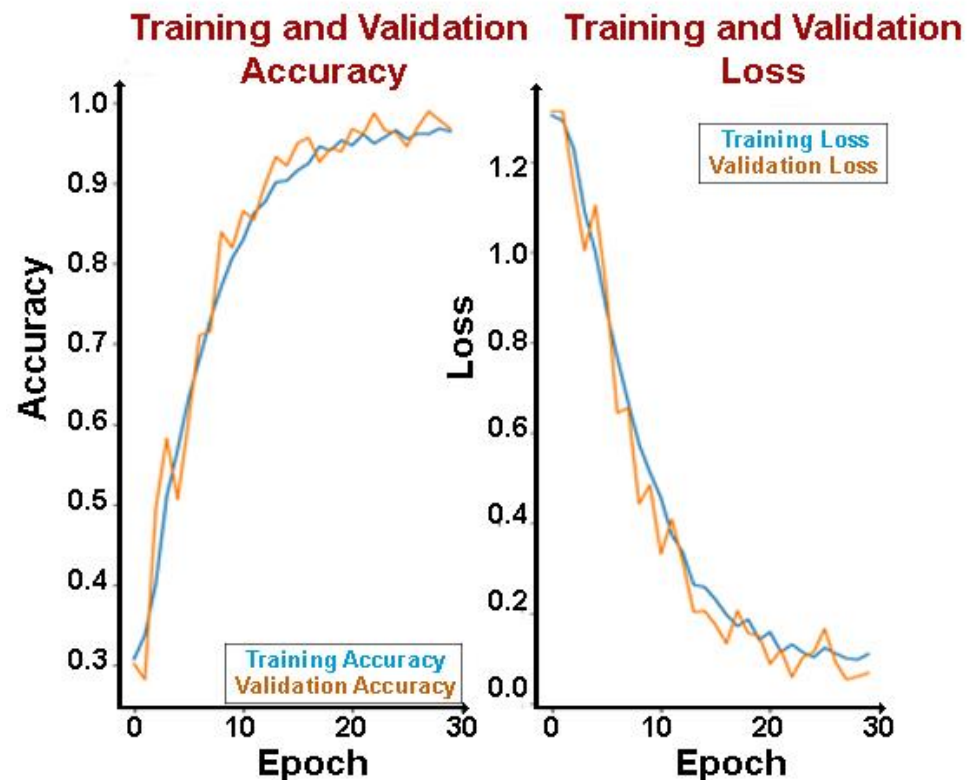


Figure 2. Training accuracy and loss curves for the space maintainer prediction model.

The ROC curve presented in Figure 3 illustrates the model's performance in distinguishing between two classes: cases where a space maintainer is required and cases where it is not. The x-axis represents the false positive rate (1-specificity), while the y-axis represents the true positive rate (sensitivity), depicting the balance between correctly identifying positive cases and minimizing false alarms. As the decision threshold changes, the curve demonstrates how the model's sensitivity and specificity vary. The ROC curve closely approaches the top-left corner, which indicates a high discriminative ability of the model. (AUC is calculated as 0.94, signifying near-perfect classification performance. AUC values range from 0 to 1, where 0.5 suggests no discriminatory power (equivalent to random guessing), and 1.0 represents a perfect classifier. An AUC of 0.94 confirms the model's strong ability to correctly identify cases requiring a space maintainer while minimizing false positives. This high AUC value highlights the robustness and reliability of the model in making accurate predictions. Given its exceptional performance, the model has significant potential as a decision-support tool in pediatric dentistry, assisting clinicians in making informed and efficient assessments regarding the necessity of space maintainers.

In Figure 4, the confusion matrix visualizes the performance of the model in predicting the need for a space maintainer using 80 flattened images. It provides a detailed breakdown of the true positives, true negatives, false positives, and false negatives.

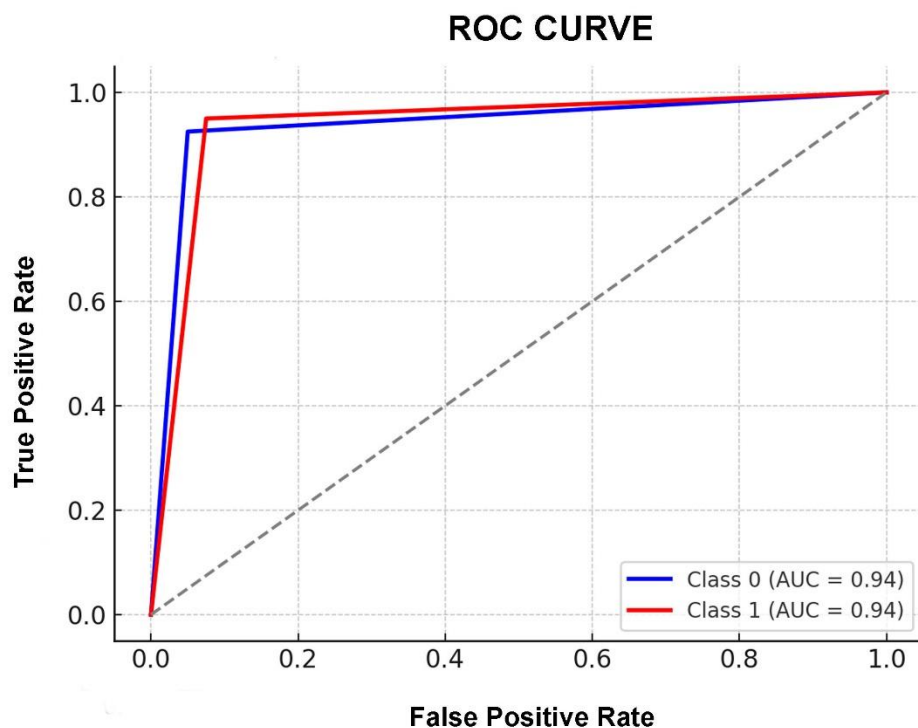


Figure 3. ROC curve illustrating the performance of each CNN model across different classes.

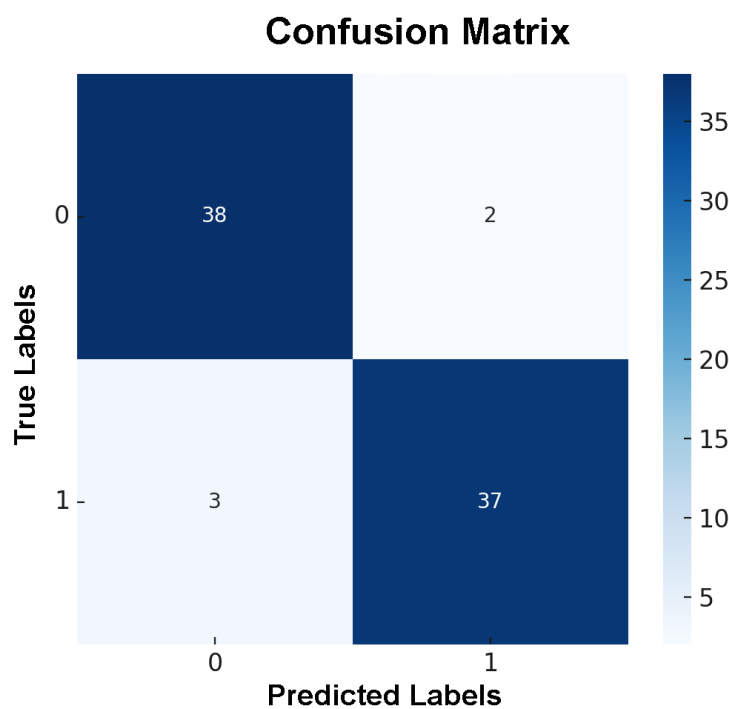


Figure 4. Confusion matrix for space maintainer prediction model.

The matrix contains four quadrants:

- The top-left quadrant shows 38 true negatives, where the model correctly predicted that no space maintainer was needed.
- The bottom-right quadrant indicates 37 true positives, representing cases where the model accurately identified the need for a space maintainer.
- The top-right quadrant reflects two false positives, where the model incorrectly predicted the need for a space maintainer when it was not required.

- The bottom-left quadrant has three false negatives, indicating that the model missed three cases that required a space maintainer.

This performance indicates that the model exhibits a high degree of accuracy and sensitivity, as it successfully identified all true cases requiring intervention (sensitivity) and minimized the occurrence of false positives.

The metrics and classification report provides a comprehensive evaluation of the model's performance in predicting the need for a space maintainer. The overall accuracy of the model is 94%, which demonstrates its strong ability to correctly classify both classes (need and no need for a space maintainer). This high accuracy reflects the model's robustness and reliability when applied to dental X-ray data.

Table 2 shows the classification report that provides key performance metrics for the model's ability to distinguish between the two classes: cases where a space maintainer is needed (Class 1) and cases where it is not needed (Class 0). For Class 0 (Not Needed), the precision is 0.93, indicating that 93% of the instances predicted as Class 0 were correct. The recall for this class is 0.95, meaning the model successfully identified 95% of actual Class 0 cases. The F1-score, which balances precision and recall, is 0.94, confirming strong performance in detecting cases where a space maintainer is unnecessary. For Class 1 (Needed), the precision is 0.95, meaning that 95% of the instances predicted as Class 1 were correctly classified. The recall is 0.93, showing that the model correctly identified 93% of actual Class 1 cases. The F1-score for this class is also 0.94, indicating an effective balance between precision and recall. The macro average precision, recall, and F1-score are all 0.94, representing the arithmetic mean of these metrics across both classes. Similarly, the weighted average precision, recall, and F1-score are also 0.94, reflecting the overall performance while accounting for class support (number of instances per class). These high values indicate that the model performs consistently well across both classes.

Table 2. Classification report summarizing the model's performance in the prediction task.

	Precision	Recall	F1-Score	Support
Class 0 (Not Needed)	0.93	0.95	0.94	40
Class 1 (Needed)	0.95	0.93	0.94	40
Macro Avg. Accuracy	0.94	0.94	0.94	80
Weighted Avg. Accuracy	0.94	0.94	0.94	80

The additional metrics, including the ROC AUC score of 0.94 and the MCC of 0.875, underscore the model's high discriminative power and reliability. The ROC AUC score reflects the model's capability to distinguish between the two classes, while the MCC provides a balanced measure of performance, particularly useful for datasets with slight class imbalances.

Figure 5 demonstrates the model's predictions for 86 X-ray images, where each image was previously annotated by a dentist to identify the teeth requiring space maintainers. The tooth number on the *x*-axis corresponds to the Universal Numbering System, which is commonly used in dentistry to label teeth in a standardized manner. Each point represents a prediction made by the CNN-based AI model, correlating the tooth number with the model's probability of requiring a space maintainer.

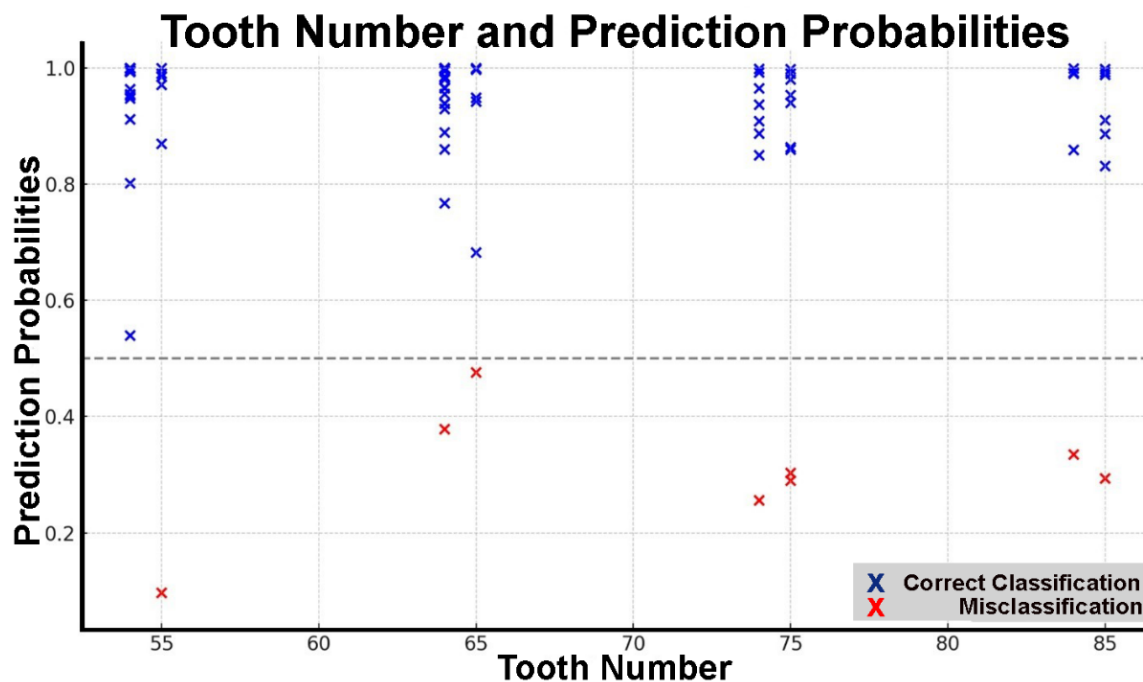


Figure 5. Prediction probabilities by tooth number for space maintainer detection. The dotted line represents the 0.5 prediction probability threshold.

The y -axis displays the prediction probabilities generated by the model, ranging from 0 to 1, where higher values indicate a stronger prediction that a space maintainer is necessary. A threshold of 0.5 was applied as follows:

- Predictions above 0.5 (marked in blue) indicate that the model predicts the need for a space maintainer.
- Predictions below 0.5 (marked in red) indicate that the model does not predict the need for a space maintainer.

The results in this figure allow for an assessment of how well the model aligns with expert annotations. The clustering of blue points near 1.0 and red points near 0.0 suggests that the model makes confident classifications in most cases. However, a few predictions close to the 0.5 threshold may indicate uncertain cases, which could require further validation. Out of the 86 images tested, the model produced eight errors.

Figure 6 illustrates a dental X-ray where the CNN model predicted that there is no need for an SM following the extraction of teeth 75 and 85. This prediction is generated by the trained CNN, which analyzes the features of the X-ray image, such as tooth alignment, spacing, and other structural patterns. For each case, the model evaluates these features and assigns a probability, determining whether an SM is necessary. In this instance, the prediction confidently indicates that the natural spacing and alignment are adequate, and no intervention is required. This demonstrates the model's ability to automate and streamline the decision-making process in pediatric dentistry.

Figure 7 illustrates the prediction result generated by the proposed CNN model, highlighting the necessity of a space maintainer following the extraction of tooth 75. The panoramic radiograph displays a clear bounding box around the affected region, emphasizing the location of the missing or soon-to-be-extracted tooth. The model successfully identifies the specific tooth requiring intervention and labels it accordingly, demonstrating its ability to detect and classify cases where space maintainers are essential. This result confirms the model's efficacy in assisting clinicians by providing automated and

accurate assessments, ultimately aiding in timely and effective treatment planning for pediatric patients.

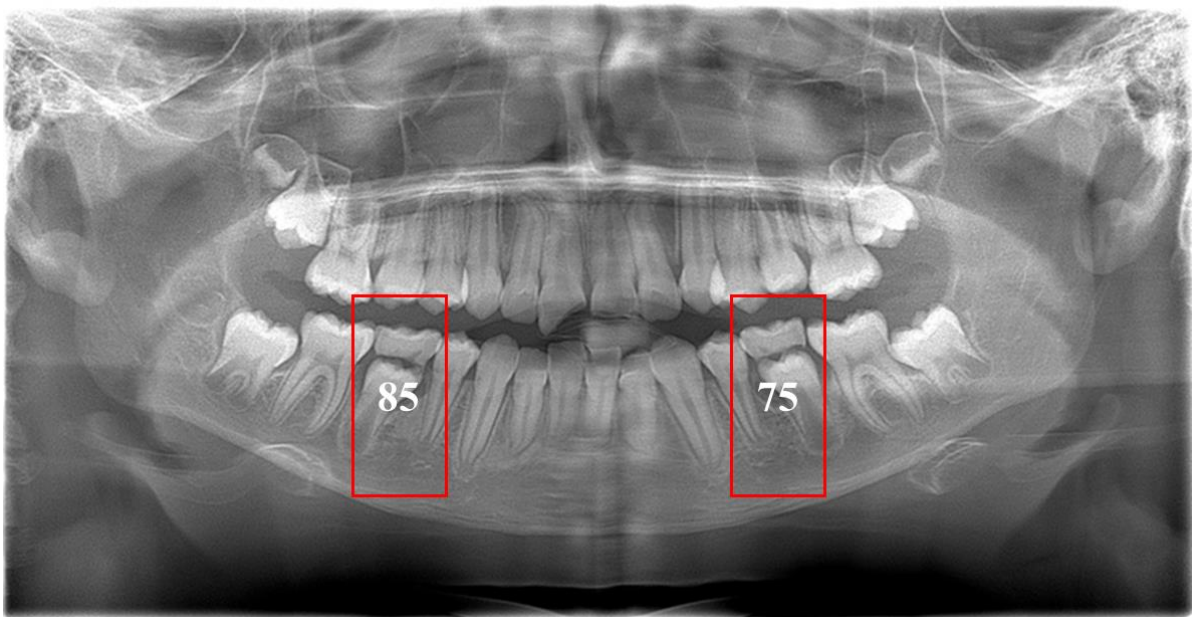


Figure 6. Prediction result indicating no need for a space maintainer after the extraction of teeth 75 and 85.

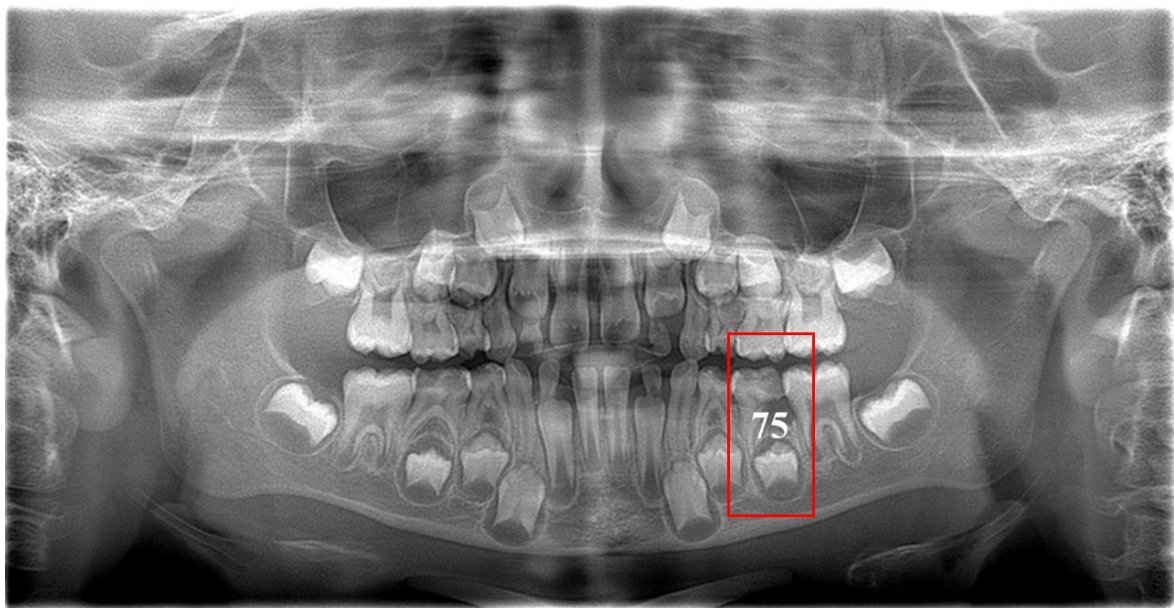


Figure 7. Prediction result indicating the necessity of a space maintainer after the extraction of tooth 75.

4. Discussion

The findings from this study emphasize the potential of the CNN model as a robust tool for predicting the need for dental SMs based on X-ray imagery. The model demonstrated a high overall accuracy of 94%, supported by other performance metrics such as precision, recall, F1-score, and ROC AUC. These results indicate that the model can effectively support pediatric dentists by automating the assessment of dental X-rays, minimizing diagnostic delays, and enhancing treatment planning. Below, we discuss the insights derived from these results, compare them with findings in similar studies, and identify limitations and opportunities for future research.

The results show the CNN model's ability to generalize effectively, as evidenced by its low false positive and false negative rates. High precision for Class 0 (0.93) indicates the model's success in avoiding unnecessary interventions, which is critical in pediatric dentistry. Similarly, the recall for Class 1 (0.93) underscores the model's ability to identify all cases requiring an SM, ensuring no critical cases are overlooked. Visual examples, such as Figures 6 and 7, highlight the practical application of the model. The accurate identification of structural dental anomalies (e.g., spacing issues) demonstrates the model's capability to process and analyze complex patterns in X-ray images. Furthermore, the model's ability to pinpoint specific tooth numbers adds a layer of precision that could streamline clinical workflows and decision-making processes.

The ROC AUC score of 0.94 obtained in this study is consistent with the results reported in other dental AI applications, such as [34], where CNN models were applied to classify orthodontic issues or identify dental caries. The confusion matrix and classification report highlight the model's ability to maintain a balance between precision and recall, like the results observed in studies like [41], where CNNs were used to predict the need for orthodontic retainers. These comparisons affirm that the model's performance is not only competitive but also consistent with the capabilities of state-of-the-art AI systems in dental diagnostics.

The focus on tooth-level predictions aligns with emerging trends in personalized dental care, as seen in [42], where AI models are tailored for patient-specific interventions. Previous studies have implemented models such as U-Net [43], SegNet [44], BiseNet [45], and Dense-ASPP [46] for dental image segmentation, each of which involves many trainable parameters. The U-Net-based approach focuses on a hybrid loss function weighted on tooth edges rather than architectural modifications. However, their method relied on hyperparameters that may not be optimal and were validated on a limited number of edge-optimized images. Compared to our approach, our model employs optimized hyperparameters, or a more efficient way to enhance prediction accuracy while maintaining computational efficiency. Recent research has shown that BERT-based classification of pediatric dental diseases achieved an accuracy of 77%, while a 1D-CNN reached 84%, outperforming other pretrained CNNs [47]. In contrast, our CNN model leverages direct radiographic image features rather than text-based transformations, enabling more precise tooth localization and treatment prediction. Recent advancements in AI-driven pediatric dental analysis have focused on automating primary teeth segmentation from CBCT scans, achieving expert-level accuracy (98%) and significantly reducing the processing time compared to manual methods [48]. Recent studies have demonstrated the effectiveness of deep learning models, such as YOLOv8, in improving diagnostic accuracy for interproximal caries detection, achieving a high precision of 96.03% for enamel caries and reducing false negatives [49]. Recent research has demonstrated the effectiveness of CNN models in diagnosing dental diseases by classifying radiographic images into different categories, such as fillings, cavities, and implants [50]. While such segmentation models enhance treatment planning efficiency, they do not predict specific clinical interventions, such as the necessity of a space maintainer. Our CNN model goes beyond segmentation by not only identifying teeth but also determining the teeth number, offering a more comprehensive AI-assisted decision-making tool for pediatric dentistry. The development of a no-code AI model for detecting primary proximal surfaces from bitewing radiographs demonstrates the increasing potential of AI in pediatric dental diagnostics, achieving high accuracy and precision with a limited dataset [51]. While this model focuses on caries detection, our CNN model targets a different aspect of pediatric dentistry by predicting the need for space maintainers and localizing specific teeth for intervention. Both models highlight the value

of AI in enhancing diagnostic and treatment planning, but our approach integrates both classification and localization for more comprehensive clinical decision support.

Despite the promising performance of the model, several limitations must be considered. Firstly, the dataset size was relatively small, with only 400 images for binary classification and 86 for tooth-level prediction. While the model performed well within this context, a larger, more diverse dataset would be crucial to improve its generalizability and robustness, ensuring that it can perform effectively across a broader range of cases. Additionally, the model was trained and tested on a single dataset, which may limit its applicability to other populations or imaging modalities. Variations in dental X-ray characteristics—such as differences in equipment, settings, and patient demographics—could influence the model's performance when applied to different clinical environments. Moreover, while the dataset was balanced for the binary classification task, there were slight imbalances in the tooth-specific predictions, which could impact the model's accuracy for underrepresented categories. Lastly, error analysis revealed higher prediction errors for tooth numbers 75 and 84, suggesting the need for further investigation to determine whether these errors are due to model limitations, dataset bias, or the inherent challenges of interpreting certain dental structures. To better understand the misclassifications, we analyzed the prediction probabilities across different tooth numbers (Figure 5). The results indicate that errors are primarily concentrated around tooth numbers 65 and 75, suggesting potential feature overlap between these teeth. Additionally, our confusion matrix (Figure 4) reveals that false positives ($n = 2$) and false negatives ($n = 3$) occur in cases with reduced contrast or anatomical similarity to adjacent teeth. Furthermore, inconsistencies in the radiographic quality, such as variations in brightness and contrast, may have contributed to the observed errors. To mitigate these issues, future improvements will focus on data augmentation strategies to enhance feature diversity, the optimization of classification thresholds, and potential architectural modifications to strengthen feature extraction in ambiguous cases.

Collaborations with multiple institutions or incorporating public datasets could improve the model's generalizability. Extending the model to handle multi-class predictions would broaden its clinical applicability by enabling it to classify additional dental conditions beyond the need for SMs. Leveraging transfer learning techniques with pretrained models on larger datasets, such as ImageNet, could further enhance performance, particularly when dealing with limited datasets. Incorporating explainability methods like Grad-CAM could provide valuable insights into the model's decision-making process, increasing trust and usability among clinicians.

5. Conclusions

This study highlights the potential of CNNs in predicting dental SM requirements from X-ray images, achieving 94% accuracy with strong performance metrics. The model effectively classifies cases needing intervention, offering a novel AI-driven approach to pediatric dental care. By automating X-ray assessments, it enhances diagnostic efficiency and supports clinical decision making, minimizing both overtreatment and missed diagnoses.

Future improvements include expanding the dataset, addressing imaging variability, and reducing errors for specific tooth numbers. Leveraging transfer learning and explainability methods like Grad-CAM could further enhance the model's robustness and clinical acceptance.

Author Contributions: Conceptualization, A.Y., G.G.P., E.O. and F.C.; methodology, A.Y., G.G.P., E.O. and F.C.; software, A.Y., G.G.P., E.O. and F.C.; validation, A.Y., G.G.P., E.O. and F.C.; formal analysis, G.G.P., E.O. and F.C.; investigation, A.Y., G.G.P., E.O. and F.C.; resources, A.Y., G.G.P., E.O. and F.C.; data curation, A.Y., G.G.P., E.O. and F.C.; writing—original draft, A.Y., G.G.P., E.O. and F.C.; writing—review & editing, A.Y., G.G.P., E.O. and F.C.; visualization, A.Y., G.G.P., E.O. and F.C.;

supervision, F.C; project administration, A.Y. 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 ethical statement of the article titled “AI-Powered Prediction of Dental Space Maintainer Needs Using X-Ray Imaging: A CNN-Based Approach for Pediatric Dentistry” has been approved by the “ethics committee” of Fatih Sultan Mehmet Vakif University. Ethics committee approval date/number: 06.02.2025/44/28.

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The writer utilized Grammarly, Quillbot, and ChatGPT to improve readability and check grammar while preparing this work. After utilizing this tool/service, the writers assumed complete accountability for the publication’s content, scrutinizing and revising it as needed.

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

References

1. Christensen, J.R.; Fields, H.W., Jr. Chapter 25: Space Maintenance in the Primary Dentition. In *Pediatric Dentistry-Pageburst E-Book on VitalSource, Infancy Through Adolescence*; 5: Pediatric Dentistry-Pageburst E-Book on VitalSource; Elsevier: Amsterdam, The Netherlands, 2012; p. 379.
2. Brothwell, D.J. Guidelines on the use of space maintainers following premature loss of primary teeth. *J. Can. Dent. Assoc.* **1997**, *63*, 753–757. [PubMed]
3. Milano, M. Conditions associated with premature exfoliation of primary teeth or delayed eruption of permanent teeth. In *Craniofacial and Dental Developmental Defects: Diagnosis and Management*; Springer: Cham, Switzerland, 2015; pp. 27–47.
4. Nadelman, P.; Magno, M.B.; Pithon, M.M.; de CASTRO, A.C.R.; Maia, L.C. Does the premature loss of primary anterior teeth cause morphological, functional and psychosocial consequences? *Braz. Oral Res.* **2021**, *35*, e092. [PubMed]
5. Durward, C.S. Space maintenance in the primary and mixed dentition. *Ann. R. Australas. Coll. Dent. Surg.* **2000**, *15*, 203–205. [PubMed]
6. Rock, W.P. UK National Clinical Guidelines in Paediatric Dentistry*. *Int. J. Paediatr. Dent.* **2002**, *1*, 151–153.
7. Mitchell, L.; Littlewood, S.J.; Nelson-Moon, Z.L.; Dyer, F. The aetiology and classification of malocclusion. *Introd. Orthod.* **2007**, *3*, 9.
8. Hu, X.; Chen, X.; Fan, M.; Mulder, J.; Frencken, J.E. What happens to cavitated primary teeth over time? A 3.5-year prospective cohort study in China. *Int. Dent. J.* **2013**, *63*, 183–188.
9. Ghafari, J. Early treatment of dental arch problems. I. Space maintenance, space gaining. *Quintessence Int.* **1986**, *17*, 423.
10. Richardson, M.E. The relationship between the relative amount of space present in the deciduous dental arch and the rate and degree of space closure subsequent to the extraction of a deciduous molar. *Dent. Pract. Dent. Rec.* **1965**, *16*, 111–118.
11. Mosharrfian, S.; Baghalian, A.; Hamrah, M.H.; Kargar, M. Clinical evaluation for space maintainer after unilateral loss of primary first molar in the early mixed dentition stage. *Int. J. Dent.* **2021**, *2021*, 3967164.
12. Kamki, H.; Kalaskar, R.; Balasubramanian, S.; Badhe, H.; Kalaskar, A. Clinical effectiveness of fiber-reinforced composite space maintainer and band and loop space maintainer in a pediatric patient: A systematic review and meta-analysis. *Int. J. Clin. Pediatr. Dent.* **2021**, *14* (Suppl. S1), S82.
13. Thakur, B.; Bhardwaj, A.; Luke, A.M.; Wahjungrum, D.A. Effectiveness of traditional band and loop space maintainer vs 3D-printed space maintainer following the loss of primary teeth: A randomized clinical trial. *Sci. Rep.* **2024**, *14*, 14081.
14. Wright, G.Z.; Kennedy, D.B. Space Control in The Primary and Mixed Dentitions. *Dent. Clin. N. Am.* **1978**, *22*, 579–601. [PubMed]
15. Croft, P.; Altman, D.G.; Deeks, J.J.; Dunn, K.M.; Hay, A.D.; Hemingway, H.; LeResche, L.; Peat, G.; Perel, P.; E Petersen, S.; et al. The science of clinical practice: Disease diagnosis or patient prognosis? Evidence about “what is likely to happen” should shape clinical practice. *BMC Med.* **2015**, *13*, 20.
16. Bouletreau, P.; Makaremi, M.; Ibrahim, B.; Louvrier, A.; Sigaux, N. Artificial intelligence: Applications in orthognathic surgery. *J. Stomatol. Oral Maxillofac. Surg.* **2019**, *120*, 347–354.
17. Giczy, A.V.; Pairolo, N.A.; Toole, A.A. Identifying artificial intelligence (AI) invention: A novel AI patent dataset. *J. Technol. Transf.* **2022**, *47*, 476–505.
18. Topol, E.J. High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* **2019**, *25*, 44–56.

19. Alharbi, N.; Alharbi, A.S. AI-driven innovations in pediatric dentistry: Enhancing care and improving outcome. *Cureus* **2024**, *16*, e69250.
20. Hamet, P.; Tremblay, J. Artificial intelligence in medicine. *Metabolism* **2017**, *69*, S36–S40.
21. Malik, P.; Pathania, M.; Rathaur, V.K. Overview of artificial intelligence in medicine. *J. Fam. Med. Prim. Care* **2019**, *8*, 2328–2331.
22. Kamel, I. Artificial intelligence in medicine. *J. Med. Artif. Intell.* **2024**, *7*, 4.
23. Goyal, H.; Mann, R.; Gandhi, Z.; Perisetti, A.; Zhang, Z.; Sharma, N.; Saligram, S.; Inamdar, S.; Tharian, B. Application of artificial intelligence in pancreaticobiliary diseases. *Ther. Adv. Gastrointest. Endosc.* **2021**, *14*, 2631774521993059. [[CrossRef](#)] [[PubMed](#)]
24. Goyal, H.; Mann, R.; Gandhi, Z.; Perisetti, A.; Ali, A.; Aman Ali, K.; Sharma, N.; Saligram, S.; Tharian, B.; Inamdar, S. Scope of artificial intelligence in screening and diagnosis of colorectal cancer. *J. Clin. Med.* **2020**, *9*, 3313. [[CrossRef](#)] [[PubMed](#)]
25. Yang, Y.J.; Bang, C.S. Application of artificial intelligence in gastroenterology. *World J. Gastroenterol.* **2019**, *25*, 1666. [[CrossRef](#)] [[PubMed](#)]
26. Hong, Y.; Hou, B.; Jiang, H.; Zhang, J. Machine learning and artificial neural network accelerated computational discoveries in materials science. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2020**, *10*, e1450. [[CrossRef](#)]
27. Burt, J.R.; Torosdagli, N.; Khosravan, N.; RaviPrakash, H.; Mortazi, A.; Tissavirasingham, F.; Hussein, S.; Bagci, U. Deep learning beyond cats and dogs: Recent advances in diagnosing breast cancer with deep neural networks. *Br. J. Radiol.* **2018**, *91*, 20170545. [[CrossRef](#)]
28. Lawson, C.E.; Martí, J.M.; Radivojevic, T.; Jonnalagadda, S.V.R.; Gentz, R.; Hillson, N.J.; Peisert, S.; Kim, J.; Simmons, B.A.; Petzold, C.J.; et al. Machine learning for metabolic engineering: A review. *Metab. Eng.* **2021**, *63*, 34–60. [[CrossRef](#)]
29. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999–7019. [[CrossRef](#)]
30. Ghali, U.M.; Usman, A.G.; Chellube, Z.M.; Degm, M.A.A.; Hoti, K.; Umar, H.; Abba, S.I. Advanced chromatographic technique for performance simulation of anti-Alzheimer agent: An ensemble machine learning approach. *SN Appl. Sci.* **2020**, *2*, 1871. [[CrossRef](#)]
31. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Umar, A.M.; Linus, O.U.; Arshad, H.; Kazaure, A.A.; Gana, U.; Kiru, M.U. Comprehensive review of artificial neural network applications to pattern recognition. *IEEE Access* **2019**, *7*, 158820–158846. [[CrossRef](#)]
32. Cengiz, A.; Karayilmaz, H. Comparative evaluation of the clinical success of 3D-printed space maintainers and band-loop space maintainers. *Int. J. Paediatr. Dent.* **2024**, *34*, 584–592. [[CrossRef](#)]
33. Cen Khan, B.A.L.; Aksoy, M.; Topsakal, K.G.; Görgülü, S. Artificial Intelligence-Based Chatbots in Providing Space Maintainer Related Information for Pediatric Patients and Parents: A comparative Study. 2024. Available online: <https://www.researchsquare.com/article/rs-4917284/v1> (accessed on 10 February 2025).
34. Ha, E.G.; Jeon, K.J.; Kim, Y.H.; Kim, J.Y.; Han, S.S. Automatic detection of mesiodens on panoramic radiographs using artificial intelligence. *Sci. Rep.* **2021**, *11*, 23061. [[CrossRef](#)]
35. Vilcapoma, P.; Parra Meléndez, D.; Fernández, A.; Váscónez, I.N.; Hillmann, N.C.; Gatica, G.; Váscónez, J.P. Comparison of Faster R-CNN, YOLO, and SSD for Third Molar Angle Detection in Dental Panoramic X-rays. *Sensors* **2024**, *24*, 6053. [[CrossRef](#)] [[PubMed](#)]
36. Çelik, B.; Genç, M.Z.; Çelik, M.E. Panoramik Radyografalarda Diş Tiplerinin Sınıflandırılması için Derin Öğrenme Yöntemlerinin Karşılaştırılması. *EMO Bilimsel Dergi* **2024**, *14*, 87–95.
37. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2016**, arXiv:1609.04747.
38. Dalianis, H.; Dalianis, H. Evaluation metrics and evaluation. In *Clinical Text Mining: Secondary Use of Electronic Patient Records*; Springer: Cham, Switzerland, 2018; pp. 45–53.
39. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *European Conference on Information Retrieval*; Springer: Cham, Switzerland, 2005; pp. 345–359.
40. Chicco, D.; Jurman, G. The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Min.* **2023**, *16*, 4.
41. Ryu, J.; Lee, Y.S.; Mo, S.P.; Lim, K.; Jung, S.K.; Kim, T.W. Application of deep learning artificial intelligence technique to the classification of clinical orthodontic photos. *BMC Oral Health* **2022**, *22*, 454. [[CrossRef](#)]
42. Khalaf, K.; Mustafa, A.; Wazzan, M.; Omar, M.; Estaitia, M.; El-Kishawi, M. Clinical effectiveness of space maintainers and space regainers in the mixed dentition: A systematic review. *Saudi Dent. J.* **2022**, *34*, 75–86. [[CrossRef](#)]
43. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015*; Proceedings, Part III 18; Springer: Cham, Switzerland, 2015; pp. 234–241.
44. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495.

45. Yu, C.; Wang, J.; Peng, C.; Gao, C.; Yu, G.; Sang, N. Bisenet: Bilateral segmentation network for real-time semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Tel Aviv, Israel, 23–27 October 2018; pp. 325–341.
46. Yang, M.; Yu, K.; Zhang, C.; Li, Z.; Yang, K. Denseaspp for semantic segmentation in street scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3684–3692.
47. Pham, T. Classification of Pediatric Dental Diseases from Panoramic Radiographs using Natural Language Transformer and Deep Learning Models. *medRxiv* **2025**, 2021–2025. [[CrossRef](#)]
48. Elsonbaty, S.; Elgarba, B.M.; Fontenele, R.C.; Swaity, A.; Jacobs, R. Novel AI-based tool for primary tooth segmentation on CBCT using convolutional neural networks: A validation study. *Int. J. Paediatr. Dent.* **2025**, *35*, 97–107.
49. Bayati, M.; Savareh, B.A.; Ahmadinejad, H.; Mosavat, F. Advanced AI-driven detection of interproximal caries in bitewing radiographs using YOLOv8. *Sci. Rep.* **2025**, *15*, 4641. [[CrossRef](#)]
50. Hasnain, M.A.; Malik, H.; Asad, M.M.; Sherwani, F. Deep learning architectures in dental diagnostics: A systematic comparison of techniques for accurate prediction of dental disease through x-ray imaging. *Int. J. Intell. Comput. Cybern.* **2024**, *17*, 161–180. [[CrossRef](#)]
51. Gonzalez, C.; Badr, Z.; Güngör, H.C.; Han, S.; Hamdan, M.H. Identifying primary proximal caries lesions in pediatric patients from bitewing radiographs using artificial intelligence. *Pediatr. Dent.* **2024**, *46*, 332–336. [[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.