Aus der

Poliklinik für Zahnärztliche Prothetik

Klinikum der Ludwig-Maximilians-Universität München

Eine externe Validierungsstudie zur diagnostischen Genauigkeit eines auf künstlicher Intelligenz basierenden Modells zur Erkennung von Karies auf Fotografien

Dissertation

zum Erwerb des Doktorgrades der Zahnmedizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Elisabeth Frenkel

aus

Moskau

Jahr

2025

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten:	Prof. Dr. Jan Kühnisch
Zweites Gutachten:	Prof. Dr. Roswitha Heinrich-Weltzien
Drittes Gutachten:	Priv. Doz. Dr. Oliver Schubert
Dekan:	Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 30.06.2025

Affidavit

Eidesstattliche Versicherung

Frenkel, Elisabeth

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

„Eine externe Validierungsstudie zur diagnostischen Genauigkeit eines auf künstlicher Intelligenz basierenden Modells zur Erkennung von Karies auf Fotografien“

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 30.06.2025

Ort, Datum

Elisabeth Frenkel

Unterschrift Doktorandin bzw. Doktorand

Inhaltsverzeichnis

Abkürzungsverzeichnis

ACC	=	Akkuratheit
AUC	=	Fläche unter der Receiver Operating Characteristic Curve
DL	=	Maschinelles Lernen
KI	=	Künstliche Intelligenz
NN	=	Neuronale Netzwerke
SE	=	Sensitivität
SP	=	Spezifität
VI	=	Visuelle Inspektion

Publikationsliste

Frenkel E, Neumayr J, Schwarzmaier J, Kessler A, Ammar N, Schwendicke F, et al. Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model-An External Validation Study. Diagnostics (Basel). 2024 Oct 14;14(20).

Schwarzmaier J, Frenkel E, Neumayr J, Ammar N, Kessler A, Schwendicke F, et al. Validation of an Artificial Intelligence-Based Model for Early Childhood Caries Detection in Dental Photographs. J Clin Med. 2024 Sep 3;13(17).

1. Ihr Beitrag zu den Veröffentlichungen

Im Rahmen des Dissertationsprojektes waren die nachfolgenden Arbeitsschritte notwendig, die wie folgt bearbeitet wurden. Der Eigenanteil ist den nachstehenden Tabellen zu entnehmen.

1.1 Beitrag zu Veröffentlichung I

Tabelle 1: Beitrag zu den Veröffentlichungen

Publikation „Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model - An External Validation Study"

	Elisabeth Frenkel	Prof. Dr. Jan Kühnisch	Co-Autoren
Ethikantrag	-	100 %	-
Projektidee/ Studiendesign	40 %	50 %	10 %
Literaturrecherche	100 %	-	-
Durchführung aller Untersuchungen	70 %	10 %	20 %
Statistische Auswertungen	80 %	10 %	10 %
Dateninterpretation und Auswahl veröffentlichungswürdiger Daten	80 %	10 %	10 %
Manuskript Management bis zur Veröffentlichung	90 %	10 %	-
Verfassen und Einreichen der Dissertation	100 %		

1.2 Beitrag zu Veröffentlichung II

Tabelle 2: Beitrag zu den Veröffentlichungen

Publikation „Validation of an Artificial Intelligence-Based Model for Early Childhood Caries Detection in Dental Photographs."

	Julia Schwarzmaier	**Prof. Dr. Jan Kühnisch**	**Co-Autoren (E. Frenkel)**
Ethikantrag	-	100 %	-
Projektidee/ Studiendesign	40 %	50 %	10 %
Literaturrecherche	100 %	-	-
Durchführung aller Untersuchungen	70 %	10 %	20 %
Statistische Auswertungen	80 %	10 %	10 %
Dateninterpretation und Auswahl veröffentlichungs-würdiger Daten	80 %	10 %	10 %
Manuskript Management bis zur Veröffentlichung	90 %	10 %	-
Verfassen und Einreichen der Dissertation	100 %		

2. Einleitung

Karies ist eine häufig auftretende Erkrankung der Zähne und betrifft weltweit rund zwei Milliarden Menschen (1). In Europa lag die Prävalenz bei ca. 33,6 % in der erwachsenen Bevölkerung. Unbehandelte Zahnkaries kann die Lebensqualität in verschiedenen Lebensphasen stark beeinträchtigen. Wiederkehrende Schmerzen, Einschränkungen beim Kauen und Schlafstörungen sind häufige Folgen, die nicht nur das Wohlbefinden, sondern auch die Produktivität erheblich mindern (2). Umso wichtiger sind eine gründliche Diagnostik, Prävention und Behandlung dieser nicht übertragbaren Erkrankung (3). Die rechtzeitige Detektion und Klassifikation erscheint daher entscheidend für die Förderung und Aufrechterhaltung einer guten Mundgesundheit und Prävention von Folgeerkrankungen (4). In der Literatur werden zunehmend Methoden zur Früherkennung von Karies vorgeschlagen, welche die herkömmlichen Diagnoseverfahren ergänzen könnten. Eine frühere Erkennung der Erkrankung könnte den Patienten die Möglichkeit einer weniger invasiven Behandlung mit geringerer Zerstörung der Zahnsubstanz ermöglichen und möglicherweise zu einer Senkung der Kosten für Patienten und Gesundheitsdienste führen (4). Die visuelle Inspektion (VI) durch den Zahnarzt stellt nach wie vor das Mittel der Wahl bei der Kariesdetektion dar (5–7). Indikationsgerecht werden bildgebende, zumeist röntgenologische Diagnostikverfahren eingesetzt und ermöglichen dem Behandler eine unterstützende Hilfestellung bei der Karieserkennung (7–9). Eine zuverlässige Kariesdiagnose führt zu einer validen Therapieentscheidung und soll die Progression der jeweiligen kariösen Läsion verhindern (10). In den letzten Jahren hat die Anwendung von künstlicher Intelligenz (KI) in der Medizin und Zahnmedizin immer mehr an Bedeutung gewonnen (8,10). Der Einsatz von KI, und maschinellem Lernen kann die zahnärztliche Erkennung und Klassifikation von Kariesläsionen unterstützen (9). Dabei verfügt die KI über die Fähigkeit, sich automatisch bestimmte Muster aus Daten anzueignen und durch wiederholte Versuche aus Erfahrungen zu lernen (11). Besonders hervorzuheben ist bei der Bildanalyse durch KI vor allem ein Teilgebiet des maschinellen Lernens, das „deep learning/maschinelles Lernen“ (DL). Es basiert auf der Grundlage künstlicher neuronaler Netze (NN) und ist inspiriert vom menschlichen Gehirn. Es weist eine Netzwerkarchitektur mit mehreren Schichten auf (10). Besonders bei komplexeren Aufgaben wie der Bildanalyse und -verarbeitung übertrifft DL mit dem Einsatz von NN das herkömmliche maschinelle Lernen (10,12). Merkmale wie Kanten, Ecken und Formen können durch gezieltes Training der KI erkannt werden (13). Es konnte gezeigt werden, dass der Einsatz von maschinellem Lernen bei der Karieserkennung auf intraoralen Fotos eine ähnliche diagnostische Leistung wie die VI erzielt (14).

Eine Arbeitsgruppe (12) konnte im Rahmen einer Metaanalyse belegen, dass die Herausforderungen bei der Karieserkennung durch Zuhilfenahme der künstlichen Intelligenz reduziert werden können. Die dokumentierten Studien zur automatisierten Bildauswertung mithilfe von KI beinhalten überwiegend röntgenologische Bildaufnahmen oder Zahnfotos (8,12). Bereits publizierte Primärstudien beschäftigten sich bislang überwiegend mit der internen Validierung eines KI-basierten Models zur Karieserkennung sowie -Klassifizierung (14–16) und verzeichneten teilweise eine hohe (81,0%-84,0%) bis sehr hohe diagnostische Güte (über 90%) (14,16). Eine Besonderheit stellt die Arbeit von Felsch und Kollegen dar (16), dessen Arbeitsgruppe sich erstmals mit der internen Validierung eines KI-basierten Models zur Kombination aus Karies-Erkennung, Klassifikation und Segmentierung beschäftigte. Der von der Arbeitsgruppe erstellte KI-Algorithmus wurde als Webapplikation zur Verfügung gestellt, so dass Interessierte klinische Zahnfotos automatisiert von einem KI-basierten Model online auswerten lassen können. Um die diagnostische Aussagekraft unter Einbeziehung der üblichen Güteparameter, wie z.B. Spezifität (SP) und Sensitivität (SE), beurteilen zu können, sind externe Validierungsstudien erforderlich. Merkmal diesen Studientyps ist die Nutzung unabhängiger Bilddaten, welche nicht Bestandteil bei der Erstellung bzw. Entwicklung des KI-Algorithmus waren.

3. Zielstellung

Das Ziel dieser Studie war die externe Validierung eines KI-gestützten Kariesererkennungsmodells (http://demo.dental-ai.de) mit Hilfe eines unabhängigen Bilddatensatzes, welcher aus frei zugänglichen Bildern aus dem Internet erstellt wurde, um die Detektion, Klassifikation, Lokalisation und Segmentierung von Karies zu determinieren. Dabei wurde die folgende Nullhypothese aufgestellt: Die externe Validierung erzielt die gleiche diagnostische Güte wie die interne Validierung (18).

4. Material und Methodik

Die vorliegende Arbeit war in das Projekt „Karieserkennung mit künstlicher Intelligenz" an der Klinik für Zahnerhaltungskunde und Parodontologie eingebettet und wurde von der Ethikkommission der Medizinischen Fakultät der Ludwig-Maximilians-Universität München genehmigt (Projektnummer 020-798).

Zur externen Validierung der Studie von Felsch et al. (16) wurden frei im Internet verfügbare Fotos von Zähnen mit oder ohne kariöse Läsionen recherchiert und verwendet. Hierfür wurde die Suchmaschine www.google.de genutzt und folgende Schlagwörter angewandt: *„Karies, Milchgebiss, bleibendes Gebiss, frühkindliche Karies, beginnende Karies, nicht-kavitierte Karies oder Kavitation*“. Alle recherchierten Fotos fanden bei der internen Validierungsstudie keine Anwendung. Alle inkludierten Fotos wurden anschließend automatisiert mit der KI-basierten Karieserkennung und in der Arbeitsgruppe ausgewertet. Letzteres Vorgehen diente als Referenzstandard. Die beiden Evaluierungen erfolgten jeweils unabhängig voneinander. Insgesamt wurden in diese Auswertung 718 Fotos in das KI-basierte Modell (http://demo.dental-ai.de) mit einbezogen, wovon nach gemeinsamen Referenzstandard bei 535 Fotos visuell eine kariöse Veränderung unterschiedlichsten Schweregrads detektiert wurde. Ausgeschlossen wurden Bilder mit Zahnrestaurationen, kieferorthopädischen Apparaturen oder seltenen Zahnerkrankungen sowie qualitativ nicht zu auszuwertende Bilder.

Das KI-basierte Modell ist in der Lage Karies zu erkennen, zu klassifizieren, zu lokalisieren und zu segmentieren. Aufgrund der Möglichkeit einer pixelgenauen Klassifizierung ist das KI-basierte Modell in der Lage mehrere Kariesinformationen pro Foto zu detektieren. Nach erfolgter KI-Analyse wurden die pixelgenau markierten Läsionen der KI der Arbeitsgruppe zur erneuten Analyse und gemeinsamen Konsensfindung vorlegt. Die nachfolgende Überprüfung der KI-gestützten Auswertung durch die Arbeitsgruppe wurde als vergleichbarer Referenzstandard festgelegt: alle inkludierten Bilder wurden auf die Korrektheit hinsichtlich der Kariesdetektion, -klassifikation, -lokalisation und -segmentierung überprüft. Die Karieserkennung erfolgte auf Grundlage etablierter Kriterien (17–21). Dabei wurde zwischen den folgenden Schweregraden unterschieden: 0 - keine Karies, 1 - nicht kariöse Läsion, 2 - gräuliche Transluzenz/ Mikrokavität, 3 - Kavitation und 4 -

zerstörter Zahn mit fast vollständigem Verlust der Zahnkrone. Bezüglich der Lokalisation musste mindestens ein Pixel des markierten Segments innerhalb der entsprechenden kariösen Läsion liegen. Befand sich kein Pixel innerhalb der eigentlichen Kariesklasse, wurde eine inkorrekte Lokalisation der Karies registriert und dokumentiert. Dabei wurde zwischen einer vollständigen (ca. >90% Übereinstimmung), einer teilweisen (ca. <90% Übereinstimmung) oder keiner Übereinstimmung unterschieden. Dabei handelte es sich um einen Schätzwert (19–23).

5. Ergebnisse

In der vorliegenden Studie von Frenkel und Kollegen (22) wurde die diagnostische Genauigkeit des KI-basierten Modells zur Karies-Erkennung, -Klassifizierung und -Segmentierung mit der des zahnärztlichen Arbeitsgruppenkonsenses überprüft. Für die Analyse wurden alle Kariesbefunde (N=991) auf den vorhandenen Zahnfotos (N=718) berücksichtigt, wobei kavitierte Zähne (N=326), nicht-kavitierte Zähne (N=300) und kariesfreie Zähne (N=192) die häufigsten Zahnbefunde ausmachten. Von allen diagnostischen Entscheidungen (N=991) konnte in 762 Fällen eine Übereinstimmung zwischen dem KI-basierten Modell und der zahnärztlichen Arbeitsgruppe erzielt werden. Die bildbezogene Karieserkennung erzielte eine Akkuratheit (ACC) von 92,0%. Die SE betrug 92,0% und die SP 91,8%. Bei der Berücksichtigung der einzelnen Kariesklassen wurde eine ACC von 85,5% (nicht-kavitierte Karies) bis 95,6% (zerstörter Zahn) erzielt. Die SE erstreckte sich von 42,9 % (gräuliche Transluzenz/Mikrokavität) bis 93,3 % (nicht kavitierte Karies), während die SP von 82,1 % (nicht kavitierte Karies) bis 99,4 % (zerstörter Zahn) reichte. Die Fläche unter der Receiver Operating Characteristic Curve (AUC) variierte zwischen 0,702 (gräuliche Transluzenz/Mikrokavität) bis 0,909 (keine Karies). Insgesamt wurde die Karieslokalisation in 755 Fällen (97,0 %) korrekt von der KI-basierten Methode vorhergesagt, wobei 52,9 % eine vollständige und 44,0 % eine teilweise korrekte Segmentierung der vorhandenen Kariesläsionen erreichten.

6. Diskussion

Bei dieser Studie handelt es sich um eine externe Validierung eines KI-gestützten Karieserkennungsmodells (http://demo.dental-ai.de) mittels eines unabhängigen Bilddatensatzes aus frei verfügbaren Bildern aus dem Internet zur Detektion, Klassifizierung, Lokalisation und Segmentierung. Die Akkuratheit bei der Kariesdetektion war mit 92,0% hoch. Im Vergleich dazu wurde im Rahmen der internen Validierung (16) ein etwas höherer Wert (97,8%) dokumentiert. Somit muss die eingangs aufgestellte Nullhypothese, dass die interne Validierung die gleiche diagnostische Güte wie die externe Validierung erzielen kann, verworfen werden. Diese Diskrepanz zwischen interner und externer Validierung wurde auch in anderen medizinischen (23,24) und zahnmedizinischen (25) Studien gezeigt. Unsere Studie untersuchte klinische Fotografien von kariösen Läsionen, die unterschiedliche Qualitäten aufwiesen und aus frei verfügbaren Internetquellen stammten. Im Vergleich dazu bestand der Bilddatensatz der internen Validierung (16) aus qualitativ hochwertigen Bildern, was den Unterschied in der Akkuratheit erklären kann. Die dennoch hohe Akkuratheit von 92% in der externen Validierung unterstreicht wiederrum die diagnostische Qualität des KI-basierten Modells für die Karieserkennung auf den unbekannten Fotografien. Bisher haben sich nur wenige Forschungsgruppen mit der Kariesdiagnostik anhand von klinischen Fotos befasst und es existieren ausschließlich Daten zur internen Validierung KI-basierter Modelle, die eine niedrigere oder vergleichbar hohe diagnostische Güte aufwiesen.

Die hier vorliegende Studie befasste sich auch mit der Korrektheit der Lokalisierung und Segmentierung kariöser Läsionen. In 97,0% der Fälle ist die Lokalisation korrekt erfolgt. Dahingegen umfasste die Beurteilung einer korrekten Segmentierung 52,9% der Fälle. Es ist allerdings kritisch anzumerken, dass es sich hierbei um subjektive Einschätzungen der Zahnärzte innerhalb der Arbeitsgruppe handelte und es derzeit keine Empfehlungen hinsichtlich der qualitativen und quantitativen Beurteilung segmentierter Flächen gibt.

Werden die KI-basierten Studiendaten mit In-vitro- und In-vivo-Studien zur alleinigen visuellen Kariesdiagnostik verglichen, so zeigte sich, dass die KI-basierte Karieserkennung und -klassifizierung auf Zahnfotos mindestens identische, wenn nicht sogar bessere Ergebnisse zeigen konnten. Die Meta-Analyse von Macey et al. zeigte geschätzte summarische SE- und SP-Werte 0,86 (95% CI 0,80-0,90) bzw. 0,77 (95% CI 0,72-0,82) (4). Eine weitere Meta-Analyse zeigte etwas niedrigere Werte für die Kariesdetektion mit SE-Werten von 0,70 (95% CI 0,59-0,80 und SP-Werten von 0,47 (95% CI 0,26-0,70) und AUC-Werte von 0,70 (5).

Bei dieser Studie handelt es sich um die vermutlich erste externe Validierung eines KI-basiertes Modells für die automatisierte Auswertung von klinischen Fotografien mit Zähnen. Als Limitationen für die Aussagekraft dieser Studie sind folgende Aspekte zu nennen: die diagnostische Leistung des KI-Modells wurde möglicherweise durch die heterogene Qualität der eingeschlossenen Bilddateien nachteilig beeinflusst, da etliche der eingeschlossenen Bilder nur eine geringe Auflösung aufwiesen. Dennoch war es überraschend, dass auch an Bildern mit einer niedrigen Qualität gute Kariesvorhersagen dokumentiert wurden. Im Umkehrschluss spricht dies für die Qualität des KI-basierten Modells. Im Hinblick auf die Verwendung von Zahnfotos aus externen Datenquellen war es außerdem unmöglich, Informationen über die individuelle Patientensituation zu erhalten. Diese fehlenden Informationen können möglicherweise dazu geführt haben, dass der Referenzstandard nicht korrekt determiniert wurde (28). Als Herausforderung wurde weiterhin wahrgenommen, dass die pixelgenaue Auswertung multiple Diagnosen pro Bild stellen. Während in konventionellen Diagnostikstudien entweder eine zahn- oder eine zahnflächen-spezifische Diagnose gestellt wird, werden mit der KI-basierten Bildauswertung deutlich mehr Informationen zur Verfügung gestellt. Dies führte zu der Notwendigkeit ein Auswertkonzept zu erarbeiten, welches mehrere Kariesklassen, Lokalisationen und/oder Segmente gleichzeitig auf einem Bild berücksichtigen kann.

7. Zusammenfassung und Ausblick

In dieser Arbeit konnte eine hohe diagnostische Gesamtgenauigkeit (92,0%) der KI-gestützten Karieserkennung heterogener Zahnfotos aus dem Internet im Rahmen einer externen Validierung nachgewiesen werden. Auch wenn die interne Validierung des KI-gestützten Modells etwas höhere diagnostische Güteparameter zeigte, kann trotzdem geschlussfolgert werden, dass die externe Validierung des KI-basierten Modells zur Karies-Erkennung, -Klassifikation, -Lokalisation und -Segmentierung vielversprechende Ergebnisse lieferte. Als positiv ist zu berichteten, dass nur ein geringes Gefälle zwischen den internen und externen Validierungsdaten zu verzeichnen war. Aus methodischer Sicht ist zu ergänzen, dass für die zuverlässige Bewertung der korrekten Detektion, Klassifikation, Lokalisation und Segmentierung Empfehlungen erarbeitet werden sollten, um die Vergleichbarkeit von unterschiedlichen Studien sicherzustellen. Weitere Studien erscheinen notwendig, um die KI-basierte Diagnostik stetig zu verbessern.

8. Abstract (English)

In this study, a high overall diagnostic accuracy (92,0%) of the AI-supported caries detection of heterogeneous dental photos from the Internet was demonstrated as part of an external validation. Even though the internal validation of the AI-based model showed slightly higher diagnostic quality parameters, it can still be concluded that the external validation of the AI-based model for caries detection, classification, localization and segmentation delivered promising results. On a positive note, there was only a slight discrepancy between the internal and external validation data. From a methodological point of view, it should be added that recommendations should be developed for the reliable evaluation of correct detection, classification, localization and segmentation in order to ensure the comparability of different studies. Further studies appear necessary in order to continuously improve AI-based diagnostics.

9. Veröffentlichung I

Journal: Diagnostics

Volume 14, 2281

Oktober 2024

Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model-An External Validation Study

Elisabeth Frenkel, Julia Neumayr, Julia Schwarzmaier, Andreas Kessler, Nour Ammar, Falk Schwendicke, Jan Kühnisch, Helena Dujic

diagnostics

IMPACT FACTOR 3.0
Indexed in: PubMed
CITESCORE 4.7

Article

Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model—An External Validation Study

Elisabeth Frenkel, Julia Neumayr, Julia Schwarzmaier, Andreas Kessler, Nour Ammar, Falk Schwendicke, Jan Kühnisch and Helena Dujic

Special Issue
Deep Learning in Medical and Biomedical Image Processing

Edited by
Dr. Shuyue Guan

MDPI

https://doi.org/10.3390/diagnostics14202281 diagnostics

MDPI

Article

Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model—An External Validation Study

Elisabeth Frenkel [1], Julia Neumayr [1], Julia Schwarzmaier [1], Andreas Kessler [1,2], Nour Ammar [1], Falk Schwendicke [1], Jan Kühnisch [1,*] and Helena Dujic [1]

[1] Department of Conservative Dentistry and Periodontology, University Hospital, Ludwig-Maximilians University of Munich, 80336 Munich, Germany; lisafrenkel04@gmail.com (E.F.); julia.neumayr97@web.de (J.N.); julia.schwarzmaier@web.de (J.S.); andreas.kessler@med.uni-muenchen.de (A.K.); nour.ammar@outlook.de (N.A.); falk.schwendicke@med.uni-muenchen.de (F.S.); h.dujic@med.uni-muenchen.de (H.D.)

[2] Department of Prosthetic Dentistry, Faculty of Medicine, Center for Dental Medicine, Medical Center-University of Freiburg, University of Freiburg, 79106 Freiburg, Germany

* Correspondence: jan.kuehnisch@med.uni-muenchen.de

Abstract: Objective: This ex vivo diagnostic study aimed to externally validate a freely accessible AI-based model for caries detection, classification, localisation and segmentation using an independent image dataset. It was hypothesised that there would be no difference in diagnostic performance compared to previously published internal validation data. Methods: For the independent dataset, 718 dental images representing different stages of carious (n = 535) and noncarious teeth (n = 183) were retrieved from the internet. All photographs were evaluated by the dental team (reference standard) and the AI-based model (test method). Diagnostic performance was statistically determined using cross-tabulations to calculate accuracy (ACC), sensitivity (SE), specificity (SP) and area under the curve (AUC). Results: An overall ACC of 92.0% was achieved for caries detection, with an ACC of 85.5–95.6%, SE of 42.9–93.3%, SP of 82.1–99.4% and AUC of 0.702–0.909 for the classification of caries. Furthermore, 97.0% of the cases were accurately localised. Fully and partially correct segmentation was achieved in 52.9% and 44.1% of the cases, respectively. Conclusions: The validated AI-based model showed promising diagnostic performance in detecting and classifying caries using an independent image dataset. Future studies are needed to investigate the validity, reliability and practicability of AI-based models using dental photographs from different image sources and/or patient groups.

Keywords: dental caries; diagnosis; validation study; artificial intelligence; deep learning

Citation: Frenkel, E.; Neumayr, J.; Schwarzmaier, J.; Kessler, A.; Ammar, N.; Schwendicke, F.; Kühnisch, J.; Dujic, H. Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model—An External Validation Study. *Diagnostics* **2024**, *14*, 2281. https://doi.org/10.3390/diagnostics14202281

Academic Editor: Shuyue Guan

Received: 24 September 2024
Accepted: 12 October 2024
Published: 14 October 2024

1. Introduction

Dental caries is a prevalent, noncommunicable disease that requires appropriate diagnostics, oral health promotion, prevention and treatment measures [1,2]. By emphasizing non-invasive or minimally invasive caries management strategies, the importance of regular diagnostic evaluation is undebatable. Here, visual examination (VE) is the method of first choice [3–7], supported by supplementary X-ray-free or radiological diagnostic procedures when indicated [8,9]. In this context, all diagnostic procedures must have good validity, reliability and practicability. Notably, with respect to validity, false positive decisions should be avoided, as they can be potentially linked to overtreatment [10].

The importance of automated image analysis in medicine and dentistry has increased in recent years [11–14]. Automation can be achieved through the application of artificial intelligence (AI) methods. One data source for automated caries detection is digital photographs of teeth, which can be understood as the machine-readable equivalent of conventional visual examination. The few available data demonstrate promising accuracy of caries detection and classification in photographs using AI [13]. Notably, any such accuracy measures were generated using internal data, i.e., testing data stemmed from the same source as data used for training the AI-based algorithm. The associated risk of bias when only testing on such internal data has been demonstrated in several recent research projects [15–20].

While most studies on AI-based caries detection do not provide access to the developed algorithms, a recent model for caries detection, classification, localisation and segmentation developed by Felsch et al. [16] is freely accessible online as a web application. Unlike AI-based models for caries detection in X-rays, most of which are only available to dentists as paid software, this web application is the first to enable professionals to have clinical photographs of teeth automatically analysed by an AI-based model. Notably, such testing by others may be conducted independently from the workgroup publishing the data, i.e., using external test data. At present, the lack of external validations for AI-based models represents a significant knowledge gap in this field, highlighting the need for independent evaluations to ensure reliability and broader applicability. The aim of this study was to determine the external accuracy of this developed AI model for caries detection using an independent sample of images from the internet. It was hypothesised that the external validity would be identical to the internal validity published by Felsch et al. [16].

2. Materials and Methods

The present ex vivo diagnostic study is part of the "Caries detection with artificial intelligence" project at the Department of Conservative Dentistry and Periodontology, which received the approval of the Ethics Committee of the Medical Faculty of the Ludwig-Maximilians-Universität München (project number 020-798). The study was implemented in accordance with the guidelines of the Standards for Reporting of Diagnostic Accuracy Studies (STARD) Steering Committee [21]. In addition, the latest recommendations for standardizing the planning, implementation and publication processes of dental studies using AI methods were considered [22].

2.1. Collection of Dental Photographs from the Web

To externally validate the recently published AI-based model (http://demo.dental-ai.de; accessed 23 September 2024) [16], independent image data not involved in the development of the model were necessary. To fulfil this requirement, photographs of teeth with or without carious lesions were obtained from freely accessible internet sources. The following keywords were entered in a web browser (www.google.de) to search for photographs: "caries", "primary dentition", "permanent dentition", "early childhood caries", "beginning caries", "noncavitated caries" and "cavity". The inclusion criteria required photographs that were clear or exhibited minimal blur or distortion, had sufficient lighting, were without artefacts and had a minimum resolution of 72 pixels per inch. Exclusion criteria were images of teeth with direct/indirect dental restorations and tooth structure disorders, such as molar incisor hypomineralisation or fluorosis, as well as images showing rare dental diseases. The search and selection process took place over a period of two weeks and was carried out by the participating dentists, whereby uncertainties regarding the eligibility of inclusion and exclusion criteria were discussed in the working group. The identified photographs featured individual anterior and posterior teeth, with posterior teeth mainly captured from the occlusal view and anterior teeth from the vestibular view. Each photograph focused on single teeth rather than multiple teeth or a single arch. Regarding the previously defined inclusion criteria, a total of 287 and 431 photographs of deciduous and permanent teeth, respectively, were identified. The image set (n = 718) included 183 photographs of healthy teeth and 535 photographs of carious teeth; the latter group exhibited caries to varying degrees of severity. The photographs were saved in the given resolution and respective format to make them available for later image analysis.

2.2. Caries Detection and Classification by the Dental Workgroup (Reference Standard)

All included photographs (n = 718) were analysed, classified and agreed upon by the workgroup (EF, JN, JS, HD and JK). In the event of differing opinions, the image in question was discussed until a consensus was reached. The visual detection and classification of carious lesions were based on established criteria for caries detection [23–27]. In the first step, it was determined whether caries could be identified in the photograph or not (score 0—no caries). In the second step, differentiation was made regarding the degree of severity: 1—noncavitated caries (presence of opacities or discolouration indicating an established noncavitated lesion on the enamel surface), 2—greyish translucency/microcavity (greyish translucency of the enamel as a sign of undermining dentin involvement/caries-induced breakdown on the natural enamel surface), 3—cavitation (visible dentin involvement) and 4—destructed tooth (extensive cavity with near-total loss of the tooth crown). The categories were based on the established diagnostic classes from the available visual diagnostic systems [23–27] and the classes included in the AI-based model. In addition to the assessment of existing carious lesions by dentists, the photograph quality was evaluated and classified as either "acceptable" or "good". This assessment included sharpness, resolution and exposure quality. When opinions in the workgroup diverged, the arguments for each classification were exchanged and discussed until a consensus was reached. The results of the dental evaluation provided a reference standard for the subsequent statistical exploration of the data. The reference standard was established before any AI-based image evaluation and can, therefore, be regarded as independent (Figure 1).

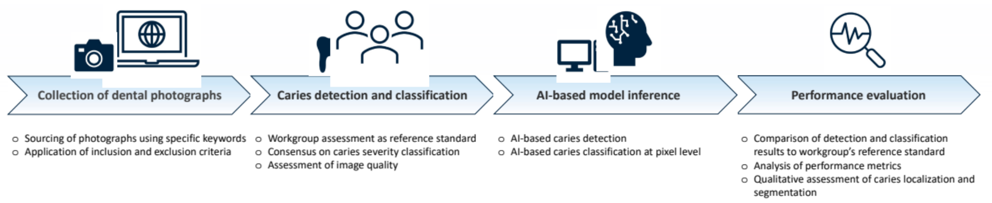

Figure 1. Workflow diagram illustrating the methodological steps.

2.3. Caries Detection and Classification by the AI-Based Model (Test Method)

The above mentioned AI-based model be applied to detect caries and classify the clinical appearance as follows: 1—noncavitated caries, 2—greyish translucency/microcavity, 3—cavitation and 4—destructed tooth with nearly complete loss of the tooth crown [16]. For images showing teeth without caries, no pixel-level class was assigned by the AI-based model. Consequently, the class 0—no caries was documented for such images. The existing caries features were classified at a pixel level, allowing multiple caries classes to be identified simultaneously for each photograph. After accessing the abovementioned website, each photograph (n = 718) was uploaded separately, and the area of interest was narrowed, as required, using the crop tool, followed by AI-based image evaluation. The analysis was performed several weeks after the initial consensus on the reference standard. Each AI-based detection and classification result was verified within the workgroup (EF, JN, JS, HD and JK) to reach a group consensus.

2.4. Caries Localisation and Segmentation by the AI-Based Model

As part of the evaluation, the segments generated using the AI-based model were evaluated for correctness regarding localisation and segmentation (Figure 2). First, a comparison was made as to whether the localisation was accurately identified. For this purpose, at least one pixel from the marked segment had to lie within the corresponding carious lesion. If there was no pixel within the actual carious lesion, it was determined that the caries was incorrectly localised. In the second step, the quality of the segmentation was evaluated, and the predicted segment was compared with the present carious lesion. A distinction was made between a complete (approximately >90% match), partial (approximately <90% match) or absent match. The match could only be qualitatively estimated based on the photograph, as no quantitative image data were provided.

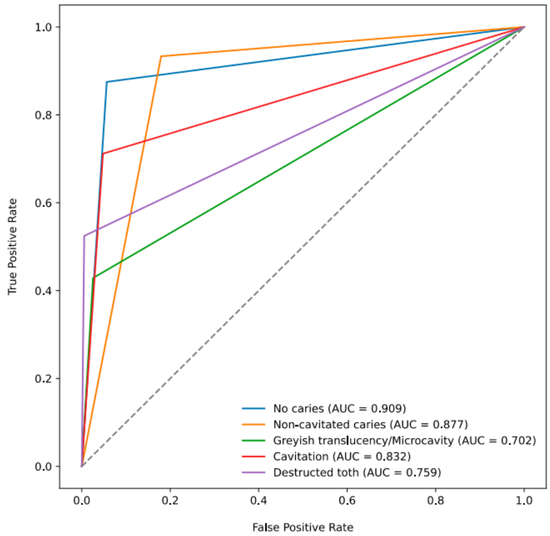

Figure 2. ROC curves and the corresponding AUCs for all caries classes.

2.5. Data Management and Statistical Analysis

All photographic information and diagnostic decisions for the test method and reference standard were recorded in a digital datasheet created for this study. For this purpose, open-access data processing software was used (EpiData Manager and EpiData Entry Client, version v4.6.0.6, EpiData Association, Odense, Denmark, http://www.epidata.dk, accessed 23 September 2024). Before descriptive and explorative data analyses, data were exported and visualised using an Excel spreadsheet (Excel 2019, Microsoft, Redmond, WA, USA) and checked for validity. Using Python (version 3.8.5, http://www.python.org, accessed 23 September 2024), and the diagnostic performance of the test method relative to the reference standard was calculated. Specifically, the numbers of true positives (TPs), false positives (FPs), true negatives (TNs) and false negatives (FNs), the sensitivity (SE), the specificity (SP), the positive and negative predictive values (PPVs and NPVs, respectively), the diagnostic accuracy (ACC = (TN + TP)/(TN + TP + FN + FP)) and the area under the receiver operating characteristic (ROC) curve (AUC) were determined [28].

3. Results

The diagnostic performance of the AI-based image analysis (test method) was determined relative to the dental workgroup consensus (reference standard). In the first step, image-based caries detection was considered, resulting in a diagnostic accuracy of 92.0%. SE and SP were 92.0% and 91.8%, respectively (Table 1).

In the next step, all recorded carious lesions (n = 991) in the available photographs (n = 718) were considered. Cavitated (n = 326) and noncavitated lesions (n = 300) and caries-free teeth (n = 192) were the most frequent dental findings (Table 2). The overall agreement was 76.9%; there were 44 false positives (4.4%) and 185 false negatives (18.7%) (Table 2). The diagnostic performance of the test method across different lesion classes is shown in Table 3. Here, the overall diagnostic accuracy ranged from 85.5% (noncavitated caries) to

95.6% (destructed tooth). SE ranged between 42.9% (greyish translucency/microcavity) and 93.3% (noncavitated caries), and SP ranged between 82.1% (noncavitated caries) and 99.4% (destructed tooth). The AUC was between 0.702 (greyish translucency/microcavity) and 0.909 (no caries). The corresponding ROC curves are shown in Figure 2.

Table 1. Cross-tabulation including diagnostic performance parameters for the AI-based image evaluation in relation to the visual consensus diagnosis by the dental work group (reference standard). In this analysis, only the ability for caries detection per image (*n* = 718) was considered.

Image-Related Caries Detection		Visual Evaluation (Reference Standard)			
		No Caries	Caries	∑	
AI-based evaluation (Test method)	No caries	168	43	211	NPV = 79.6%
	Caries	15	492	507	PPV = 97.0%
	∑	183	535	718	
		SP = 91.8%	SE = 92.0%		ACC = 92.0%

Table 2. Cross-tabulation of the AI-based image evaluation in relation to the visual consensus diagnosis by the dental work group (reference standard) for all caries diagnoses. This tabulation includes all diagnoses (*n* = 991) from all images (*n* = 718); multiple findings per image were possible.

		Visual Evaluation (Reference Standard)					
		No Caries	Noncavitated Caries	Greyish Translucency/ Microcavity	Cavitation	Destructed Tooth	∑
AI-based evaluation (Test method)	No caries	168	6	6	29	4	213
	Noncavitated caries	21	280	42	50	11	404
	Greyish translucency/Microcavity	1	9	39	13	0	62
	Cavitation	1	3	4	232	24	264
	Destructed tooth	1	2	0	2	43	48
	∑	192	300	91	326	82	991

Table 3. Summary of the diagnostic performance for all caries classes. The tabulation included all diagnoses (*n* = 991) from all images (*n* = 718).

	True Positives		True Negatives		False Positives		False Negatives		Diagnostic Performance (%)					
	n	%	*n*	%	*n*	%	*n*	%	ACC	SE	SP	PPV	NPV	AUC
No caries	168	17.0	754	76.1	45	4.5	24	2.4	93.0	87.5	94.4	78.9	96.9	0.909
Noncavitated caries	280	28.3	567	57.2	124	12.5	20	2.0	85.5	93.3	82.1	69.3	96.6	0.877
Greyish translucency/ Microcavity	39	3.9	877	88.5	23	2.3	52	5.2	92.4	42.9	97.4	62.9	94.4	0.702
Cavitation	232	23.4	633	63.9	32	3.2	94	9.5	87.3	71.2	95.2	87.9	87.1	0.832
Destructed tooth	43	4.3	904	91.2	5	0.5	39	3.9	95.6	52.4	99.4	89.6	95.9	0.759

In addition, the diagnostic performance was determined considering the image quality of each digital photograph, which was classified as either "acceptable" or "good". The calculated values can be taken from Table 4. We also evaluated the correctness of caries localisation and segmentation (Table 5). In principle, the AI-based method correctly predicted caries localisation in 755 cases (97.0%) but incorrectly in 23 cases (3.0%). Fully and partially correct segmentation was achieved for 52.9% and 44.0% of carious lesions, respectively, with 3.0% of the lesions being incorrectly segmented (Table 5).

Table 4. Summary of the diagnostic performance for all caries classes in relation to image quality. The tabulation includes all diagnoses (n = 991) from all images (n = 718).

	Image Quality	True Positives (n)	True Negatives (n)	False Positives (n)	False Negatives (n)	Diagnostic Performance (%)					
						ACC	SE	SP	PPV	NPV	AUC
No caries	Acceptable	95	277	19	17	91.2	84.8	93.6	83.3	94.2	0.892
	Good	73	477	26	7	94.3	91.3	94.8	73.7	98.6	0.930
Noncavitated caries	Acceptable	90	257	56	5	85.0	94.7	82.1	61.6	98.1	0.884
	Good	190	310	68	15	85.8	92.7	82.0	73.6	95.4	0.873
Greyish translucency/Microcavity	Acceptable	13	374	7	14	94.9	48.1	98.2	65.0	96.4	0.732
	Good	26	503	16	38	90.7	40.6	96.9	61.9	93.0	0.688
Cavitation	Acceptable	85	265	16	42	85.8	66.9	94.3	84.2	86.3	0.806
	Good	147	368	16	52	88.3	73.9	95.8	90.2	87.6	0.849
Destructed tooth	Acceptable	23	357	4	24	93.1	48.9	98.9	85.2	93.7	0.739
	Good	20	547	1	15	97.3	57.1	99.8	95.2	97.3	0.785

Table 5. Overview of the correctness of the AI-based caries localisation and segmentation. The analysis includes all caries diagnoses (n = 778).

	Localisation				Segmentation					
	Correct		Incorrect		Fully Correct		Partially Correct		Incorrect	
	N	%	N	%	N	%	N	%	N	%
Noncavitated caries	387	49.7	17	2.2	224	28.8	163	21.0	17	2.2
Greyish translucency/Microcavity	59	7.6	3	0.4	22	2.8	37	4.8	3	0.4
Cavitation	262	33.7	2	0.3	145	18.6	117	15.0	2	0.3
Destructed tooth	47	6.0	1	0.1	21	2.7	26	3.3	1	0.1
Σ	755	97.0	23	3.0	412	52.9	343	44.1	23	3.0
	778 *				778 *					

* In total, 213 images were categorised as "no caries" by the AI-based model, for which no caries localisation and segmentation could be presented.

4. Discussion

In the present ex vivo diagnostic study, a recently introduced AI-based model [16] was validated, and the results regarding its ability to automatically detect, classify, localise and segment carious lesions in digital photographs of teeth were determined. In summary, the AI-based model achieved an overall diagnostic accuracy of 92.0%, and the fraction of false-positive findings, which could be associated with overtreatment, was low (4.4%). In addition, most carious lesions were correctly localised by the test method (97.0%). The automatically drawn pixel segments were fully and partially correct in 52.9% and 44.1% of all cases, respectively.

The discussion of the documented external validity results for this new AI-based method is limited because 1) no other comparable AI-based diagnostic method is available thus far, and 2) most available diagnostic studies in the field of AI-dentistry on dental photographs involve model development and internal validation only [17–20,29,30]. Therefore, the aim of this study was to close this knowledge gap. A comparison between the internal [16] and external validity of the underlying AI-based model (Tables 1–4, Figure 2) reveals higher accuracy in internal data (97.8%) than in external data (92.0%). Internal data refer to the AI-based model's diagnostic performance on the dataset used during the development [16], which demonstrates the capability of the model to recognise patterns it has been trained on. In contrast, the performance of the AI-based model on external datasets reflects its ability to generalise to new and unseen cases. In view of the results obtained, the hypothesis that internal and external validity are identical had to be rejected. This decrease between internal and external validity has been reported in other dental [31] and medical study projects [32,33]. In a study by Fu et al. [31], the AUC for the detection of periapical lesions in 3D radiographs using an AI-based model decreased from 0.97 (internal validation) to 0.93–0.96 (external validation). The AUC values from the internal and external validations decreased from 0.949 to 0.843 in the AI-based detection of early childhood visual impairment [33]. Bora et al. [32] investigated the diagnostic performance of an AI-based model to detect diabetic retinopathy in adults; in their study, the AUC decreased from 0.81 (internal validation) to 0.71 (external validation) [32]. The abovementioned study data document the reduction in diagnostic performance from internal to external validation and thus emphasise the importance of external validation.

For the present study, clinical photographs of carious teeth of heterogenous quality were obtained from a variety of freely accessible internet sources; in contrast, model training was performed using well-standardised, high-quality, professionally captured photographs [16]. Therefore, the documented diagnostic performance (Tables 1–4, Figure 2) supported the quality of AI-based caries detection, as all the photographs supplied to the model were unknown and nevertheless assessed with a high accuracy of 92.0%. Setting this order of magnitude from external validation in relation to the published data from other internal validation studies on caries detection generally showed relatively low or comparable values for diagnostic performance [18–20,29,30]. To date, only a few workgroups have focused on caries diagnostics using clinical photographs. Bottenberg et al. [34] detected caries in dental photographs and subsequently used histological examination of the extracted teeth as a reference standard. The AUC of 0.84 determined for dental diagnostics on photographs was comparable to our results, emphasising the general possibility of using photographs of teeth for diagnostic purposes [34].

A comparative analysis of the results for caries classification from the external validation (Tables 2 and 3) and the internal validation data published by Felsch et al. [16] showed a slightly reduced diagnostic performance on our external validation data. This reduction in performance was evident in several caries groups (Table 2). For instance, in the noncavitated caries group, the ACC was 85.5%, the SE was 93.3%, and the SP was 82.1% (Table 3). Conversely, Felsch et al. [16] reported an ACC of 90.1%, SE of 88.4% and SP of 91.4%. For clinically more relevant cavitated lesions, the ACC was 87.3%, the SE was 71.2%, and the SP was 95.2, while Felsch et al. [16] reported an ACC of 95.9%, SE of 87.6% and SP of 97.8%. Similarly, Wang et al. [35] documented an ACC of 95.3% for the detection of healthy tooth surfaces versus tooth surfaces with white spot lesions. Notably, the authors integrated fluorescence data into the model, which could have increased the model performance [35]. Conversely, Thanh et al. [29] reported ACC values between 81.0% and 87.4% for the detection of healthy tooth surfaces and tooth surfaces with noncavitated and cavitated carious lesions. The internal model performance was lower than that of Felsch et al. [16] and Wang et al. [35].

As part of the external validation, the correctness of the localisation and segmentation of the AI-based model was evaluated. The recorded localisation was classified as correct in 97.0% of the cases. In contrast, segmentation was only rated as fully correct in just over half of the cases (52.9%), while in 44.1% of cases, the AI-based model achieved only partially correct segmentation of the existing carious lesions. However, these data require critical appraisal, as (1) they were subjective estimates by the participating dentists, (2) there are currently no recommendations for the qualitative and quantitative evaluation of segmented areas, (3) typical parameters, such as average precision or intersection over union, could not be determined, and (4) the selected cut-off of ~90% of the lesion area was quite strict. Therefore, these values should not be overinterpreted at this stage.

If the documented internal and external validation data are compared with the data from in vitro and/or in vivo studies on visual caries diagnostics, the data published to date for AI-based caries detection and classification on dental photographs indicate at least identical, if not improved, results. The meta-analysis by Macey et al. [7], in which the estimated summary SE and SP values were 0.86 (95% CI 0.80–0.90) and 0.77 (95% CI 0.72–0.82), respectively, should be mentioned here as a representative example. Another meta-analysis [6] confirmed this finding for caries detection on occlusal surfaces and revealed SE values of 0.70 (95% CI 0.59–0.80), SP values of 0.47 (95% CI 0.26–0.70) and AUC values of 0.70 based on the included clinical studies. On the one hand, these data show that AI methods can obviously generate accuracy similar to conventional VE methods. On the other hand, comparative clinical diagnostic studies, including meticulous VE and an independent, AI-based evaluation of the photographic tooth status, are lacking. This signals a knowledge gap that should be closed in future studies.

The present diagnostic study has strengths and limitations. Its novelty lies in being the first external validation study of the mentioned AI-based model for caries detection and classification using an independent sample of dental photographs with different caries classes as well as a range of image qualities. This aspect has to be understood as a unique feature of this study in comparison to previously published studies, which reported exclusively internal validation data [17–20,29,30]. External validation is a crucial evaluation step before recommending AI-based models for broader use in clinical practice. Therefore, this study addresses that knowledge gap and demonstrates the model's ability to evaluate images from independent sources with promising diagnostic performance. Furthermore, this diagnostic study also provides information about caries localisation and segmentation (Table 5), which represent a new aspect of caries diagnostics. An additional novel aspect is that diagnostic accuracy was correlated with meaningful image quality (Table 4).

However, the following aspects should be discussed as limitations. To create an independent image database, photographs of teeth were deliberately taken from the internet. Although these photographs were of varying image quality and included a wide range of carious lesions and tooth surfaces, image size had only a limited influence on diagnostic quality (Table 4). This probably underlines the quality of the AI-based model, as it was also able to analyse photographs of only acceptable quality. Nonetheless, the possible bias in the selection of web images and the definition and application of inclusion and exclusion criteria, as well as the decision on the reference standard—which was always based on group consensus within the workgroup—should be noted. With respect to the use of dental photographs from external data sources, it was impossible for the workgroup to differentiate between real and AI-generated photographs. Furthermore, ground truth verification, which is primarily conducted by histological methods, was not available. This lack of information should be mentioned as another limitation. However, Bottenberg et al. [34] visually evaluated dental photographs for caries detection with acceptably high diagnostic accuracy relative to that of histological examination, thereby supporting the use of VE in cases where no other methods are applicable. Further, the ratio of available photographs with carious lesions (n = 535) to photographs showing healthy teeth (n = 183) could be a limitation. However, the workgroup made an effort to include all available photographs in the dataset, taking into account the inclusion and exclusion criteria. As this study mainly used images retrieved from the internet, the selected dataset showed qualitative loss compared to the dataset used to train and internally test the AI-based model [16]. To this extent, it can be discussed as a limitation, given that the described study design does not mimic the conditions of internal validation in its entirety. This aspect should be taken into account in future studies. Furthermore, the possibility of simultaneously conducting and documenting VE of the teeth under clinical conditions should also be considered, as this would allow further conclusions to be drawn regarding validity, reliability and generalisability. One challenge that emerged during the project was that the AI-based model provided multiple (pixel-wise) diagnoses per image. During most diagnostic studies, only one diagnostic decision is made for each tooth surface and method; thus, it is necessary to derive one diagnostic decision per image (Table 1) and consider the predicted segments regarding classification (Tables 2–4), localisation and segmentation (Table 5). In addition, the decision-making process was designed as a group consensus, as a completely independent evaluation of the segments against the dental reference standard was not possible. This is because the photographs in question had mostly multiple caries findings, which could not be blinded, potentially introducing verification bias, as noted here.

5. Conclusions

In this ex vivo diagnostic study, a recently introduced AI-based model was validated, showing promising diagnostic performance for caries detection, classification, localisation and segmentation using an external dataset of heterogeneous digital photographs of teeth. Diagnostic performance was slightly lower than in internal validation, reflecting challenges from varying image quality and lack of standardisation. With respect to these findings, future studies should investigate the validity, reliability and practicability of AI-based image analysis across different image sources and/or patient groups. In addition, standardised external validation protocols and the implementation of clinical assessments could improve the applicability of these models in different clinical settings.

Author Contributions: Conceptualisation, J.K. and H.D.; investigation and validation: E.F., J.N., J.S., J.K. and H.D.; formal analysis: E.F., J.N., J.S., A.K., N.A., J.K. and H.D.; writing—original draft: E.F., J.N., J.S., J.K. and H.D.; writing—review and editing: E.F., J.N., J.S., A.K., N.A., F.S., J.K. and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: N.A. received a scholarship from the Forschungsgemeinschaft Dental e.V. (FGD) with grant number 01/2023.

Institutional Review Board Statement: This ex vivo study is part of the project "Caries detection with artificial intelligence" at the Department of Conservative Dentistry and Periodontology at the LMU University Hospital, which received the approval of the Ethics Committee of the Medical Faculty of the Ludwig-Maximilians-Universität München (project number 020-798, approved on 8 October 2020).

Informed Consent Statement: Not applicable.

Data Availability Statement: The validated AI-based model is available as a web application and can be accessed at https://dental-ai.de (accessed on 23 September 2024). The data can be made available upon reasonable request.

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

References

1. Kassebaum, N.J.; Bernabe, E.; Dahiya, M.; Bhandari, B.; Murray, C.J.; Marcenes, W. Global burden of untreated caries: A systematic review and metaregression. *J. Dent. Res.* **2015**, *94*, 650–658. [CrossRef] [PubMed]
2. Pitts, N.B.; Twetman, S.; Fisher, J.; Marsh, P.D. Understanding dental caries as a non-communicable disease. *Br. Dent. J.* **2021**, *231*, 749–753. [CrossRef] [PubMed]
3. Gimenez, T.; Piovesan, C.; Braga, M.M.; Raggio, D.P.; Deery, C.; Ricketts, D.N.; Ekstrand, K.R.; Mendes, F.M. Visual Inspection for Caries Detection: A Systematic Review and Meta-analysis. *J. Dent. Res.* **2015**, *94*, 895–904. [CrossRef]
4. Gimenez, T.; Tedesco, T.K.; Janoian, F.; Braga, M.M.; Raggio, D.P.; Deery, C.; Ricketts, D.N.J.; Ekstrand, K.R.; Mendes, F.M. What is the most accurate method for detecting caries lesions? A systematic review. *Community Dent. Oral. Epidemiol.* **2021**, *49*, 216–224. [CrossRef] [PubMed]
5. Janjic Rankovic, M.; Kapor, S.; Khazaei, Y.; Crispin, A.; Schuler, I.; Krause, F.; Ekstrand, K.; Michou, S.; Eggmann, F.; Lussi, A.; et al. Systematic review and meta-analysis of diagnostic studies of proximal surface caries. *Clin. Oral. Investig.* **2021**, *25*, 6069–6079. [CrossRef] [PubMed]
6. Kapor, S.; Rankovic, M.J.; Khazaei, Y.; Crispin, A.; Schuler, I.; Krause, F.; Lussi, A.; Neuhaus, K.; Eggmann, F.; Michou, S.; et al. Systematic review and meta-analysis of diagnostic methods for occlusal surface caries. *Clin. Oral. Investig.* **2021**, *25*, 4801–4815. [CrossRef] [PubMed]
7. Macey, R.; Walsh, T.; Riley, P.; Glenny, A.M.; Worthington, H.V.; O'Malley, L.; Clarkson, J.E.; Ricketts, D. Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries. *Cochrane Database Syst. Rev.* **2021**, *6*, CD014546. [CrossRef] [PubMed]
8. Kühnisch, J.; Anttonen, V.; Duggal, M.S.; Spyridonos, M.L.; Rajasekharan, S.; Sobczak, M.; Stratigaki, E.; Van Acker, J.W.G.; Aps, J.K.M.; Horner, K.; et al. Best clinical practice guidance for prescribing dental radiographs in children and adolescents: An EAPD policy document. *Eur. Arch. Paediatr. Dent.* **2020**, *21*, 375–386. [CrossRef]

9. Kühnisch, J.; Aps, J.K.M.; Splieth, C.; Lussi, A.; Jablonski-Momeni, A.; Mendes, F.M.; Schmalz, G.; Fontana, M.; Banerjee, A.; Ricketts, D.; et al. ORCA-EFCD consensus report on clinical recommendation for caries diagnosis. Paper I: Caries lesion detection and depth assessment. *Clin. Oral. Investig.* **2024**, *28*, 227. [CrossRef]
10. Schwendicke, F.; Stolpe, M.; Meyer-Lueckel, H.; Paris, S. Detecting and Treating Occlusal Caries Lesions. *J. Dent. Res.* **2014**, *94*, 272–280. [CrossRef]
11. Cantu, A.G.; Gehrung, S.; Krois, J.; Chaurasia, A.; Rossi, J.G.; Gaudin, R.; Elhennawy, K.; Schwendicke, F. Detecting caries lesions of different radiographic extension on bitewings using deep learning. *J. Dent.* **2020**, *100*, 103425. [CrossRef] [PubMed]
12. Mertens, S.; Krois, J.; Cantu, A.G.; Arsiwala, L.T.; Schwendicke, F. Artificial intelligence for caries detection: Randomized trial. *J. Dent.* **2021**, *115*, 103849. [CrossRef] [PubMed]
13. Moharrami, M.; Farmer, J.; Singhal, S.; Watson, E.; Glogauer, M.; Johnson, A.E.W.; Schwendicke, F.; Quinonez, C. Detecting dental caries on oral photographs using artificial intelligence: A systematic review. *Oral. Dis.* **2023**, *30*, 1765–1783. [CrossRef] [PubMed]
14. Schwendicke, F.; Samek, W.; Krois, J. Artificial Intelligence in Dentistry: Chances and Challenges. *J. Dent. Res.* **2020**, *99*, 769–774. [CrossRef] [PubMed]
15. Askar, H.; Krois, J.; Rohrer, C.; Mertens, S.; Elhennawy, K.; Ottolenghi, L.; Mazur, M.; Paris, S.; Schwendicke, F. Detecting white spot lesions on dental photography using deep learning: A pilot study. *J. Dent.* **2021**, *107*, 103615. [CrossRef] [PubMed]
16. Felsch, M.; Meyer, O.; Schlickenrieder, A.; Engels, P.; Schonewolf, J.; Zollner, F.; Heinrich-Weltzien, R.; Hesenius, M.; Hickel, R.; Gruhn, V.; et al. Detection and localization of caries and hypomineralization on dental photographs with a vision transformer model. *NPJ Digit. Med.* **2023**, *6*, 198. [CrossRef]
17. Kuhnisch, 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]
18. Moutselos, K.; Berdouses, E.; Oulis, C.; Maglogiannis, I. Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 1617–1620.
19. Park, E.Y.; Cho, H.; Kang, S.; Jeong, S.; Kim, E.K. Caries detection with tooth surface segmentation on intraoral photographic images using deep learning. *BMC Oral. Health* **2022**, *22*, 573. [CrossRef]
20. Zhang, Y.; Liao, H.; Xiao, J.; Jallad, N.A.; Ly-Mapes, O.; Luo, J. A Smartphone-Based System for Real-Time Early Childhood Caries Diagnosis. In *Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2020; pp. 233–242. [CrossRef]
21. 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]
22. Schwendicke, F.; Singh, T.; Lee, J.H.; Gaudin, R.; Chaurasia, A.; Wiegand, T.; Uribe, S.; Krois, J.; IADR e-oral health network and the ITU WHO focus group AI for Health. Artificial intelligence in dental research: Checklist for authors, reviewers, readers. *J. Dent.* **2021**, *107*, 103610. [CrossRef]
23. Kühnisch, J.; Bücher, K.; Henschel, V.; Albrecht, A.; Garcia-Godoy, F.; Mansmann, U.; Hickel, R.; Heinrich-Weltzien, R. Diagnostic performance of the universal visual scoring system (UniViSS) on occlusal surfaces. *Clin. Oral. Investig.* **2011**, *15*, 215–223. [CrossRef] [PubMed]
24. Kühnisch, J.; Goddon, I.; Berger, S.; Senkel, H.; Bücher, K.; Oehme, T.; Hickel, R.; Heinrich-Weltzien, R. Development, methodology and potential of the new Universal Visual Scoring System (UniViSS) for caries detection and diagnosis. *Int. J. Environ. Res. Public Health* **2009**, *6*, 2500–2509. [CrossRef] [PubMed]
25. Nyvad, B.; Machiulskiene, V.; Baelum, V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. *Caries Res.* **1999**, *33*, 252–260. [CrossRef]
26. Pitts, N.B. How the detection, assessment, diagnosis and monitoring of caries integrate with personalized caries management. *Monogr. Oral. Sci.* **2009**, *21*, 1–14. [CrossRef] [PubMed]
27. Pitts, N.B.; Ekstrand, K.R. International Caries Detection and Assessment System (ICDAS) and its International Caries Classification and Management System (ICCMS)—Methods for staging of the caries process and enabling dentists to manage caries. *Community Dent. Oral. Epidemiol.* **2013**, *41*, e41–e52. [CrossRef] [PubMed]
28. Matthews, D.E.; Farewell, V.T. *Using and Understanding Medical Statistics*; Karger Medical and Scientific Publishers: Basel, Switzerland, 2015. [CrossRef]
29. Thanh, M.T.G.; Van Toan, N.; Ngoc, V.T.N.; Tra, N.T.; Giap, C.N.; Nguyen, D.M. Deep Learning Application in Dental Caries Detection Using Intraoral Photos Taken by Smartphones. *Appl. Sci.* **2022**, *12*, 5504. [CrossRef]
30. 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] [PubMed]
31. Fu, W.T.; Zhu, Q.K.; Li, N.; Wang, Y.Q.; Deng, S.L.; Chen, H.P.; Shen, J.; Meng, L.Y.; Bian, Z. Clinically Oriented CBCT Periapical Lesion Evaluation via 3D CNN Algorithm. *J. Dent. Res.* **2024**, *103*, 5–12. [CrossRef] [PubMed]
32. Bora, A.; Balasubramanian, S.; Babenko, B.; Virmani, S.; Venugopalan, S.; Mitani, A.; de Oliveira Marinho, G.; Cuadros, J.; Ruamviboonsuk, P.; Corrado, G.S.; et al. Predicting the risk of developing diabetic retinopathy using deep learning. *Lancet Digit. Health* **2021**, *3*, e10–e19. [CrossRef]
33. Chen, W.; Li, R.; Yu, Q.; Xu, A.; Feng, Y.; Wang, R.; Zhao, L.; Lin, Z.; Yang, Y.; Lin, D.; et al. Early detection of visual impairment in young children using a smartphone-based deep learning system. *Nat. Med.* **2023**, *29*, 493–503. [CrossRef]

34. Bottenberg, P.; Jacquet, W.; Behrens, C.; Stachniss, V.; Jablonski-Momeni, A. Comparison of occlusal caries detection using the ICDAS criteria on extracted teeth or their photographs. *BMC Oral. Health* **2016**, *16*, 93. [CrossRef] [PubMed]
35. Wang, C.; Qin, H.; Lai, G.; Zheng, G.; Xiang, H.; Wang, J.; Zhang, D. Automated classification of dual channel dental imaging of auto-fluorescence and white light by convolutional neural networks. *J. Innov. Opt. Health Sci.* **2020**, *13*, 2050014. [CrossRef]

10. Veröffentlichung II

Journal: Journal of clinical medicine

Volume 13, 5215

September 2024

Validation of an artificial intelligence-based model for early childhood caries detection in dental photographs.

Julia Schwarzmaier , Elisabeth Frenkel , Julia Neumayr , Nour Ammar , Andreas Kessler, Falk Schwendicke , Jan Kühnisch , Helena Dujic

Journal of
Clinical Medicine

MDPI

Article

Validation of an Artificial Intelligence-Based Model for Early Childhood Caries Detection in Dental Photographs

Julia Schwarzmaier [1], Elisabeth Frenkel [1], Julia Neumayr [1], Nour Ammar [1], Andreas Kessler [1,2], Falk Schwendicke [1], Jan Kühnisch [1,*] and Helena Dujic [1]

[1] Department of Conservative Dentistry and Periodontology, University Hospital, Ludwig-Maximilians University of Munich, 80336 Munich, Germany; julia.schwarzmaier@web.de (J.S.); lisafrenkel04@gmail.com (E.F.); julia.neumayr97@web.de (J.N.); nour.ammar@outlook.de (N.A.); andreas.kessler@uniklinik-freiburg.de (A.K.); falk.schwendicke@med.uni-muenchen.de (F.S.); h.dujic@med.uni-muenchen.de (H.D.)
[2] Department of Prosthetic Dentistry, Center for Dental Medicine, University Hospital Freiburg, 79106 Freiburg, Germany
* Correspondence: jan.kuehnisch@med.uni-muenchen.de; Tel.: +49-89-4400-59343 (ext. 59301)

Abstract: Background/Objectives: Early childhood caries (ECC) is a widespread and severe oral health problem that potentially affects the general health of children. Visual–tactile examination remains the diagnostic method of choice to diagnose ECC, although visual examination could be automated by artificial intelligence (AI) tools in the future. The aim of this study was the external validation of a recently published and freely accessible AI-based model for detecting ECC and classifying carious lesions in dental photographs. **Methods**: A total of 143 anonymised photographs of anterior deciduous teeth (ECC = 107, controls = 36) were visually evaluated by the dental study group (reference test) and analysed using the AI-based model (test method). Diagnostic performance was determined statistically. **Results**: ECC detection accuracy was 97.2%. Diagnostic performance varied between carious lesion classes (noncavitated lesions, greyish translucency/microcavity, cavitation, destructed tooth), with accuracies ranging from 88.9% to 98.1%, sensitivities ranging from 68.8% to 98.5% and specificities ranging from 86.1% to 99.4%. The area under the curve ranged from 0.834 to 0.964. **Conclusions**: The performance of the AI-based model is similar to that reported for the internal dataset used by developers. Further studies with independent image samples are required to comprehensively gauge the performance of the model.

Keywords: dental caries; early childhood caries; artificial intelligence; diagnosis

Citation: Schwarzmaier, J.; Frenkel, E.; Neumayr, J.; Ammar, N.; Kessler, A.; Schwendicke, F.; Kühnisch, J.; Dujic, H. Validation of an Artificial Intelligence-Based Model for Early Childhood Caries Detection in Dental Photographs. *J. Clin. Med.* **2024**, *13*, 5215. https://doi.org/10.3390/jcm13175215

Academic Editor: Juan Martin Palomo

Received: 11 July 2024
Revised: 22 August 2024
Accepted: 30 August 2024
Published: 3 September 2024

1. Introduction

Early childhood caries (ECC) is among the most prevalent diseases worldwide [1–4]. Defined as the presence of one or more noncavitated or cavitated carious lesions, and missing or filled tooth surfaces due to caries in any deciduous tooth in a child under the age of six, ECC is often associated with a loss of dental functionality and aesthetics, pain due to abscesses, underweight, or reduced quality of life if left untreated [5]. Furthermore, the risk of caries development in the permanent dentition is increased [6]. For experienced professionals, the diagnosis of ECC is usually an immediate visual diagnosis. The diagnostic assessment is made using various classification systems, which consider either the extent of affected deciduous teeth in relation to age [4,7] or the distribution pattern of carious lesions in the dentition [8]. The established visual criteria facilitate precise diagnostic assessment of all teeth or tooth surfaces [9–13].

As part of ongoing efforts to digitise and automate caries diagnostics, several studies have reported on the use of deep learning models for caries detection on intraoral photographs [14–18] and dental radiographs [19–21]. Although these studies differ in methodology, model type, and dataset size, the reported results show promising diagnostic performance [14,15,18,19,21]. Specifically in paediatric dentistry, AI-based models have been explored not only for caries detection on intraoral photographs [17] but also for the identification and numbering of deciduous and permanent teeth [22,23] as well as the identification of mesiodens on radiographs [24–26].

Despite these advances, most studies report only internal validation data—results obtained during the initial development and validation using the test dataset—and the models developed are often not freely accessible [15–18]. Furthermore, the lack of external validation limits their applicability to different clinical settings. In contrast to this, Felsch et al. [14] published a model that uses artificial intelligence (AI) to automatically detect, classify, localise and segment caries on photographs of teeth. This AI-based model is freely accessible and can currently be used for image analysis without restriction. This accessibility is crucial as it allows for external validation using independent imaging data that were not part of the model's training. Thus, the model's diagnostic performance and, by extension, its generalisability can be tested on new, unseen data, enabling a comparison with the internal validation results. The aim of the present study was to validate the AI-based model [14] using independent image data of anterior deciduous teeth. It was hypothesised that there would be no significant differences between the internal and external validation accuracies.

2. Materials and Methods

This investigation was conducted following the recommendations of the STARD Steering Committee (Standards for Reporting of Diagnostic Accuracy Studies) and the recently published recommendations for the design and conduct of studies using AI methods in dental research [27,28].

2.1. Dental Images

A total of 143 anonymised dental photographs of anterior deciduous teeth were selected from an established external image database created for clinical documentation and training purposes, which includes a wide range of clinical presentations of ECC. The images were captured by an experienced dentist (JK) using standard procedures. Specifically, each intraoral image was taken with a professional single-reflex lens camera (D300, D7100, or D7200 with a Nikon Micro 105 mm lens, Nikon, Minato (Tokio), Japan) equipped with a macro lens and a macro flash (EM-DG 140, Sigma, Kawasaki (Kanagawa), Japan). Before taking the photographs, the teeth were cleaned and dried. All photographs had the following parameters: aspect ratio of 3:2, minimum resolution of 2784 $\times$ 1856 pixels without compression, jpeg format and RGB colour space. None of the photographs were generated using AI methods, edited or manipulated. The selection was conducted independently of all other analyses by an examiner with over 20 years of clinical experience, including extensive practice in paediatric dentistry. The process followed established clinical criteria for detecting and diagnosing ECC, utilizing classification systems that assess the distribution pattern of lesions in the dentition [8]. These criteria ensured that the selected images accurately and comprehensively reflected the relevant clinical manifestations.

2.2. Dental Image Evaluation (Reference Test)

The selected photographs (n = 143) were evaluated in detail by the dental study group of five dentists (JS, EF, JN, HD, JK). The dataset included images from different stages of ECC (n = 107) and caries-free dentitions (n = 36). In addition to the basic decision as to whether ECC was present, all existing caries findings were carefully evaluated. In detail, each carious lesion was classified according to the ICDAS/UniViSS criteria: caries-related noncavitated opacities in the enamel, caries-related breakdown of the chalky or brown-stained enamel surface in the form of a microcavity, greyish translucency indicating underlying dentin involvement, cavitation with visible dentin involvement, and extensive cavities up to the complete destruction of the tooth crown [12,29,30]. In case of differing opinions, the finding was discussed within the dental study group until a consensus decision was reached among the five dentists. The documented findings served as the reference test.

Prior to this study, the investigators had completed a one-day theoretical training session on the assessment of ECC and noncavitated and cavitated carious lesions under the supervision of an experienced principal investigator. The training provided information on the study design, indices and diagnostic principles.

2.3. AI-Based Image Evaluation (Test Method)

For automated image evaluation, the freely accessible web tool (https://demo.dental-ai.de/, accessed on 21 August 2024), which can detect, classify, localise and segment carious lesions on dental photographs, was used for AI-based analysis [14]. The carious lesion classification was based on the ICDAS and UniViSS criteria [12,29,30]. Specifically, the web tool distinguishes between the following caries scores: 1—noncavitated carious lesion, 2—greyish translucency/microcavity, 3—caries-related cavitation and 4—destructed tooth (Figure 1). All selected anonymised images (n = 143) were uploaded individually to the aforementioned website. Once uploaded, the images were cropped to optimally display the four maxillary incisors (area of interest, Figure 1). Automated image analysis was then performed, and all potential classes of findings were marked as coloured pixels. Each colour corresponds to a specific type of finding identified by the AI-based model. Multiple adjacent pixels of the same colour formed a pixel cloud or segment representing a localised area of interest. For example, green pixels indicate cavitated lesions, while blue pixels represent greyish translucencies (Figure 1). For dental images showing caries-free dentition, no segment was highlighted by the AI-based model (Figure 1). Each AI-based image analysis result was captured and saved as a screenshot for later independent dental evaluation (Figure 1). After two weeks, the findings and segments marked by the AI-based model were assessed separately by the dental study group. A total of 261 segments were identified on all included dental photographs as AI outputs. In addition, localisation and segmentation of all caries segments from the AI-based image analysis were assessed for their correctness. Caries segment localisation could be either correct or incorrect. Here, at least one pixel had to be located in the carious lesion. When assessing the marked segments (AI outputs), a distinction was made between incorrect, partially correct and fully correct segments. If >90% of the actual caries extent was recognised by the AI, this was classified as fully correct. If the AI recognised most of the caries extent (<90%), this corresponded to the partially correct segmentation and for caries segments outside the actual lesion, the segmentation was classified as incorrect. The segments were assessed as estimates, as exact values could not be determined.

2.4. Data Management and Statistics

For this project, an entry form was used to document all diagnostic findings directly (EpiData Manager and EpiData Entry Client, V4.6.0.6, EpiData Association, Odense, Denmark, http://www.epidata.dk, accessed on 21 August 2024). The dataset was exported to an Excel spreadsheet (Excel 2019, Microsoft Corporation, Redmond, WA, USA) following data collection and prepared for statistical exploration. The diagnostic performance of the test method in comparison to the reference test was calculated using Python V3.8.5 (http://www.python.org, accessed on 21 August 2024). The true-positive (*TP*), true-negative (*TN*), false-positive (*FP*) and false-negative (*FN*) rates were calculated as key figures from contingency tables [31]. Based on this, accuracy (*ACC*), sensitivity (*SE*), specificity (*SP*), and positive and negative predictive values (*PPV*, *NPV*) were determined:

$$ACC = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

$$SE = TP/(TP + FN) \tag{2}$$

$$SP = TN/(TN + FP) \tag{3}$$

$$PPV = TP/(TP + FP) \quad (4)$$

$$NPV = TN/(TN + FN) \quad (5)$$

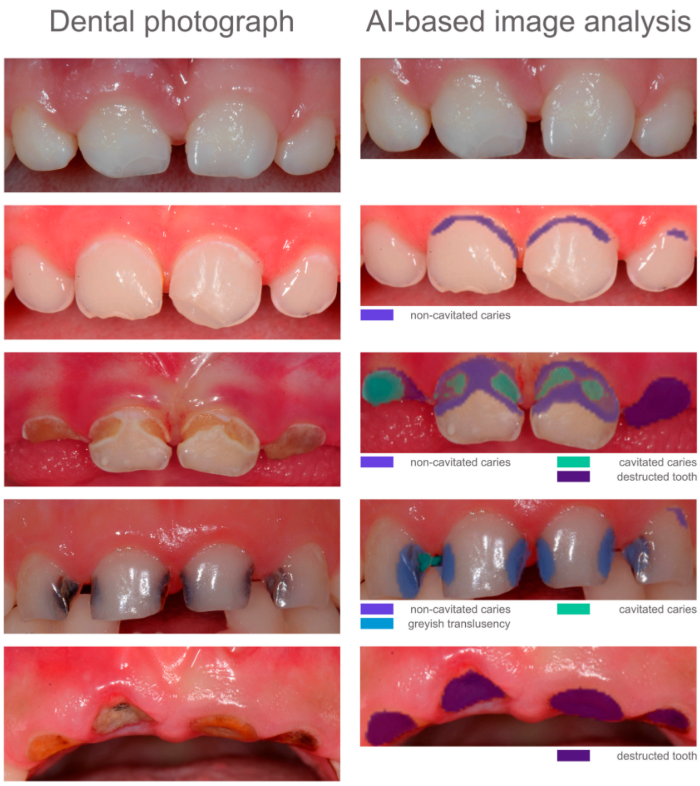

Figure 1. Exemplary photographs (**left**) and corresponding AI-based image analysis (**right**).

In addition, the area under the receiver operating characteristic (ROC) curve (*AUC*) was determined.

3. Results

The AI-based diagnostic model correctly diagnosed caries in terms of ECC on the photographs in a total of 104 images and thus achieved a *SE* of 97.2% (Table 1). Out of 36 photographs with caries-free deciduous teeth, 35 were correctly recognised, yielding a *SP* of 97.2%. This resulted in an overall diagnostic accuracy of 97.2% (Table 1).

The chosen dental photographs included multiple carious lesions which, therefore, required further analyses. The cross-tabulation of all caries diagnostic decisions from the AI-based evaluation (test method) and the visual examination (reference standard) can be taken from Table 2. The *TP*, *TN*, *FP* and *FN* rates are summarized for each caries class in Table 3. The diagnostic performance parameters of the external validation of the AI-based model are listed in Table 4. In detail, the *ACC* for the test method ranged from 88.9% (cavitation) to 98.1% (destructed tooth). The *SE* values ranged from 68.8% (greyish translucency/microcavity) to 98.5% (noncavitated carious lesion). The *SP* values ranged from 86.1% (noncavitated carious lesion) to 99.4% (cavitation).

Table 1. Cross-tabulation including diagnostic performance parameters for the image-related AI-based evaluation (test method) presented in rows in relation to the visual consensus diagnosis by the dental workgroup (reference test) presented in columns. In this analysis, the diagnostic performance for caries detection per image (n = 143) was considered.

Caries Detection		**Visual Evaluation (Reference Test)**		
		Healthy	**Caries**	
AI-based evaluation (Test method)	Healthy	35	3	*NPV* = 92.1%
	Caries	1	104	*PPV* = 99.0%
		SP = 97.2%	*SE* = 97.2%	*ACC* = 97.2%

Table 2. Contingency table of the AI-based image evaluation data (test method) presented in rows and visual consensus diagnosis data (reference test) presented in columns for all caries segments. This tabulation includes all diagnoses (n = 261) from all images (n = 143), with multiple findings per image possible.

Caries Classification		**Visual Evaluation (Reference Test)**					
		Healthy *	**Noncavitated Caries Lesion**	**Greyish Translucency/ Microcavity**	**Cavitation**	**Destructed Tooth**	**Σ**
AI-based evaluation (Test method)	Healthy *	35	1	0	2	1	39
	Noncavitated caries lesion	2	66	5	19	1	93
	Greyish translucency/Microcavity	0	0	11	5	0	16
	Cavitation	0	0	0	71	1	72
	Destructed tooth	0	0	0	2	39	41
	Σ	37	67	16	99	42	261

* Caries-free surfaces are indicated by the absence of segment markings.

Table 3. The table illustrates the true-positive (*TP*), true-negative (*TN*), false-positive (*FP*) and false-negative (*FN*) rates in relation to the used caries classes. The tabulation summarizes all diagnoses (n = 261) from all photographs (n = 143) as shown in the contingency table.

	Healthy	**Noncavitated Caries Lesion**	**Greyish Translucency/ Microcavity**	**Cavitation**	**Destructed Tooth**
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**
True positives	35 (13.4)	66 (25.3)	11 (4.2)	71 (27.2)	39 (14.9)
True negatives	220 (84.3)	167 (64.0)	240 (92.0)	161 (61.7)	217 (83.1)
False positives	4 (1.5)	27 (10.3)	5 (1.9)	1 (0.4)	2 (0.8)
False negatives	2 (0.8)	1 (0.4)	5 (1.9)	28 (10.7)	3 (1.2)
Σ	261 (100.0)	261 (100.0)	261 (100.0)	261 (100.0)	261 (100.0)

Figure 2 shows the ROC curves for the chosen caries classes and images with a caries-free dentition. The corresponding *AUC* values, which ranged from 0.834 for greyish translucency/microcavity to 0.964 for healthy dentition (Table 4), provide a summary measure of the model's overall diagnostic performance for each caries class.

Table 4. Diagnostic performance of the AI-based evaluation for each of the chosen caries classes. The calculations are based on all *TP*, *TN*, *FP* and *FN* rates (n = 261 in each caries class) from all photographs (n = 143) as shown in Table 3.

	Healthy	Noncavitated Caries Lesion	Greyish Translucency/ Microcavity	Cavitation	Destructed Tooth
ACC (in %)	97.7	89.3	96.2	88.9	98.1
SE (in %)	94.6	98.5	68.8	71.7	92.9
SP (in %)	98.2	86.1	98.0	99.4	99.1
PPV (in %)	89.7	71.0	68.8	98.6	95.1
NPV (in %)	99.1	99.4	98.0	85.2	98.6
AUC	0.964	0.923	0.834	0.855	0.960

Abbreviations: *ACC*, accuracy; *SE*, sensitivity; *SP*, specificity; *PPV*, positive predictive value; *NPV*, negative predictive value; *AUC*, area under the receiver operating characteristic (ROC) curve.

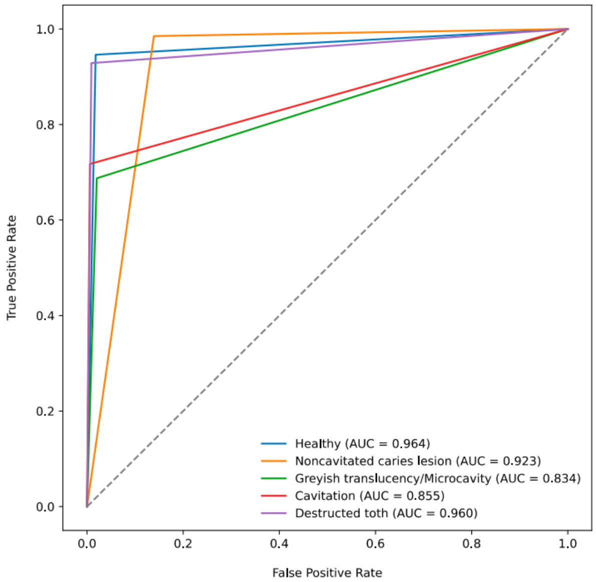

Figure 2. Receiver operating characteristic (ROC) curves and corresponding area under ROC curve (*AUC*) values for the test method and the chosen caries classes. Curves above the dashed line indicate that the proportion of correctly classified images is greater that the proportion of incorrectly classified ones.

The results for caries lesion localisation and segmentation are shown in Tables 5 and 6. A total of 143 images were analysed with the AI-based model, whereby this sample showed 226 carious lesions (86.6%) and 35 caries-free images (13.4%). In the case of existing caries, 220 segments (84.3%) were correctly localised; only 6 segments (2.3%) were incorrectly localised (Table 5). The AI-based model correctly predicted segmentation in 114 cases (43.6%). The segmentation of 101 cases (38.7%) was partially correct and in 11 cases (4.3%) the prediction was outside the existing carious lesion (Table 6).

Table 5. Results of the evaluation of AI-based caries localisation. The analysis included all diagnoses (n = 261) from all images (n = 143). * No caries localisation was possible.

	Healthy *	Noncavitated Caries Lesion	Greyish Translucency /Microcavity	Cavitation	Destructed Tooth	Σ
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Incorrect	4 (1.5)	2 (0.8)	-	-	-	6 (2.3)
Correct	-	91 (34.9)	16 (6.1)	72 (27.6)	41 (15.7)	220 (84.3)
Healthy *	35 (13.4)	-	-	-	-	35 (13.4)
Σ	39 (14.9)	93 (35.7)	16 (6.1)	72 (27.6)	41 (15.7)	261 (100.0)

Table 6. Summary of the evaluation of AI-based caries segmentation. The analysis included all diagnoses (n = 261) from all images (n = 143). * No caries segmentation was possible.

	Healthy *	Noncavitated Caries Lesion	Greyish Translucency /Microcavity	Cavitation	Destructed Tooth	Σ
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Incorrect	4 (1.5)	2 (0.8)	1 (0.4)	3 (1.2)	1 (0.4)	11 (4.3)
Partially correct	-	31 (11.9)	9 (3.4)	36 (13.8)	25 (9.6)	101 (38.7)
Fully correct	-	60 (23.0)	6 (2.3)	33 (12.6)	15 (5.7)	114 (43.6)
Healthy *	35 (13.4)	-	-	-	-	35 (13.4)
Σ	39 (14.9)	93 (35.7)	16 (6.1)	72 (27.6)	41 (15.7)	261 (100.0)

4. Discussion

The study showed that the freely accessible AI-based model can detect carious lesions on dental photographs from photographs with ECC with a high diagnostic accuracy of 97.2% (Table 1). The documented *AUC* values from the ROC curves for the classification of the included caries categories ranged from 0.834 to 0.964 (Figure 2), indicating an encouraging result for automated caries detection on photographs of anterior deciduous teeth. Furthermore, the AI-based model demonstrated high sensitivity and specificity for detecting noncavitated lesions (*SE* 98.5%, *SP* 86.1%) and destructed teeth (*SE* 92.9%, *SP* 99.1%), showing a promising performance in identifying both early-stage and advanced carious lesions. However, the model's lower sensitivity for cavitated lesions (*SE* 71.7%) and greyish translucency/microcavity (*SE* 68.8%) indicates that despite high specificity, a number of cases in these categories might be overlooked. A comparison of the documented data on external validity (Table 4) with the published results on internal validity [14] showed very good agreement based on the *ACC* values. The *ACC* values from the internal [14] and external validation (Table 4, Figure 2) datasets were 90.1% and 89.3% (noncavitated carious lesion), 99.0% and 96.2% (greyish translucency/microcavity), 95.9% and 88.9% (cavitation), and 99.0% and 98.1% (destructed tooth), respectively. Based on these findings, the initial hypothesis that there would be no difference between the external and internal accuracies can be confirmed.

Moreover, when considering other diagnostic studies on automated caries detection and classification in dental photographs, a more inconsistent picture emerges. In a study project that included smartphone images of ECC patients, the model performance was assessed but not supported by detailed validation data on caries detection and classification [32]. In another paper, Zhang et al. [33] published data on the internal validity of their approach to AI-based caries detection on dental photographs. However, the internal validation data indicated that the model performance appears to be substantially lower than that of the model presented by Felsch et al. [14]. Specifically, the model by Zhang et al. [33] achieved an imagewise *SE* of 81.9% and a boxwise *SE* of 64.6% [33]. In another project [16], four different AI-based models for caries detection were developed, and their internal validity was assessed. The *ACC* values varied depending on the deep learning model, ranging from 60.7% to 68.8% for noncavitated lesions and from 81.0% to 87.4% for cavitated lesions. Similarly, Park et al. [34] reported internal validity data of a similar magnitude for caries detection, with *ACC* values of 75.8% without object extraction and 81.3% with object extraction. In addition, further projects on caries detection and classification on dental photographs were published, which had fewer methodological similarities with the aforementioned studies. Moharrami et al. [35] published an overview of the currently available projects and noted that automatic dental caries detection using AI may provide objective verification of clinicians' diagnoses. However, future studies should use more robust designs, standardized metrics, and focus on caries detection and classification metrics.

A special feature of the freely accessible AI-based model [14] is that, in addition to caries detection and classification, the evaluated images include information on lesion localisation and segmentation. This is accompanied by the simultaneous visualisation of different caries classes in one image as part of a carious lesion (Figure 1). Compared to conventional clinical examination, this is an interesting feature of AI-based image evaluation that has not yet been considered. Evaluating the information on lesion localisation (Table 5), 220 out of a total of 226 cases were correct and thus in the range of high diagnostic performance. Only a few cases ($n = 6$) showed incorrect localisation. This situation was somewhat less favourable for the segmentation performance of the AI-based model, as only 114 of 226 lesion segments were fully correct (Table 6). The model performed best in the segmentation of noncavitated carious lesions. Fully destructed teeth and dentin cavitation were proportionally less likely to be segmented fully correctly; i.e., the displayed areas deviated from the actual extent of caries in the dentist's assessment. However, it should be noted that such evaluations are based on analysing each individual image at the pixel level, offering the dentist an unprecedented level of precision. Further improvements, which could be achieved by continuously fine-tuning the AI-based model, are desirable at this point.

Finally, the strengths and limitations of this study project should be discussed. This study represents an external validation of a recently published and freely available AI-based method for automated image analysis. For this purpose, this study utilised image data that had not been used in the development of the model. This allowed for an independent verification of the AI-based model. The images used in this study were professionally captured and of comparable quality to the training dataset of the AI-based model [14]. On the one hand, it can be assumed that high quality has a positive effect on AI-based image evaluation; on the other hand, it can be argued that high-quality images may not be available in all clinical situations. This ultimately raises the question to what extent a correct diagnosis can be made on low-quality images. However, it should also be noted that the object of interest needs to be correctly depicted [36], as otherwise, a diagnostic evaluation is potentially impossible. Provided the images are of good quality, a valid diagnostic assessment is ultimately possible [36,37]. Furthermore, the photographs only showed anterior deciduous teeth in this study. Therefore, no conclusions can be drawn about the diagnostic performance in the posterior region, for which the AI-based model was also developed [14]. Moreover, the fact that the included images were anonymized and lacked clinical data can be taken into account as a limitation. This is relevant with regard to the potential impact of remineralisation on the visual characteristics of the lesions, as remineralised areas may appear less prominent or have a different structure compared to active lesions. It remains unclear to what extent this process could affect the diagnostic performance of the AI-based model. Given the number of included and available photographs both with and without ECC for this study, it should be emphasized that although a larger number of images would be desirable, obtaining such a dataset with high-quality images seems rather challenging. Nevertheless, these points should be considered in future studies to verify the reliability and generalizability of the tested AI-based model. Another desirable feature that is not yet possible with the validated AI-based model is the analysis of images with regard to a more precise ECC type [4,7,8]. The currently available AI-based model [14] enables only the detection, classification, localisation and segmentation of caries and enamel hypomineralisations. Furthermore, it is currently unable to analyse direct or indirect restorations, dental trauma and other findings, such as plaque and discolouration. Additionally, the evaluation of dental photographs with multiple teeth is not yet fully functional, requiring the area of interest to be centred using the crop tool for valid image analysis. Ultimately, this approach was also part of the methodology of this diagnostic study, which relied on visual consensus diagnoses as the reference test, without the inclusion of any histological investigation techniques to provide details about caries characteristics. Lastly, while detailed evaluations of carious lesions cannot be performed independently of the AI-based diagnostic output, the study group aimed to minimise potential evaluation bias by making consensus decisions on all diagnoses.

5. Conclusions

In comparison to the initially published internal validation data [14], we found similar results in our external validation of dental photographs of cases with and without ECC. This underlines the diagnostic quality of the AI-based model presented by Felsch et al. [14] for caries detection, classification, localisation and segmentation. However, additional studies with independent image samples are needed to comprehensively describe the performance of the model.

Author Contributions: Supervision: J.K.; conceptualisation: J.K. and H.D.; investigation and validation: J.S., J.N., E.F., J.K. and H.D.; formal analysis: J.S., J.N., E.F., A.K., N.A., J.K. and H.D.; writing—original draft: J.S., J.N., E.F., J.K. and H.D.; writing—review and editing: E.F., J.S., J.N., A.K., N.A., F.S., J.K. and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. N.A. received a scholarship from the Forschungsgemeinschaft Dental e.V. (FGD) with grant number 01/2023.

Institutional Review Board Statement: This study is part of the project “Caries detection with artificial intelligence” at the Department of Conservative Dentistry and Periodontology at the LMU University Hospital, which received the approval of the Ethics Committee of the Medical Faculty of the Ludwig-Maximilians-Universität München (project number 020-798, approved on 8 October 2020).

Informed Consent Statement: Not applicable.

Data Availability Statement: The AI-based model is available as a web application and can be accessed at https://dental-ai.de (accessed on 21 August 2024). 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 conflicts of interest.

References

1. GBD 2017 Oral Disorders Collaborators; Bernabe, E.; Marcenes, W.; Hernandez, C.R.; Bailey, J.; Abreu, L.G.; Alipour, V.; Amini, S.; Arabloo, J.; Arefi, Z.; et al. Global, Regional, and National Levels and Trends in Burden of Oral Conditions from 1990 to 2017: A Systematic Analysis for the Global Burden of Disease 2017 Study. *J. Dent. Res.* **2020**, *99*, 362–373. [CrossRef] [PubMed]
2. Kassebaum, N.J.; Bernabe, E.; Dahiya, M.; Bhandari, B.; Murray, C.J.; Marcenes, W. Global burden of untreated caries: A systematic review and metaregression. *J. Dent. Res.* **2015**, *94*, 650–658. [CrossRef] [PubMed]
3. Uribe, S.E.; Innes, N.; Maldupa, I. The global prevalence of early childhood caries: A systematic review with meta-analysis using the WHO diagnostic criteria. *Int. J. Paediatr. Dent.* **2021**, *31*, 817–830. [CrossRef] [PubMed]
4. Dye, B.A.; Hsu KL, C.; Afful, J. Prevalence and Measurement of Dental Caries in Young Children. *Pediatr. Dent.* **2015**, *37*, 200–216.
5. Tinanoff, N.; Baez, R.J.; Diaz Guillory, C.; Donly, K.J.; Feldens, C.A.; McGrath, C.; Phantumvanit, P.; Pitts, N.B.; Seow, W.K.; Sharkov, N.; et al. Early childhood caries epidemiology, aetiology, risk assessment, societal burden, management, education, and policy: Global perspective. *Int. J. Paediatr. Dent.* **2019**, *29*, 238–248. [CrossRef]
6. Seow, W.K. Early Childhood Caries. *Pediatr. Clin. N. Am.* **2018**, *65*, 941–954. [CrossRef]
7. American Academy of Pediatric Dentistry. Policy on early childhood caries (ECC): Classifications, consequences, and preventive strategies. In *The Reference Manual of Pediatric Dentistry*; American Academy of Pediatric Dentistry: Chicago, IL, USA, 2020; pp. 79–81.

8. Wyne, A.H. Early childhood caries: Nomenclature and case definition. *Community Dent. Oral Epidemiol.* **1999**, *27*, 313–315. [CrossRef]
9. Ekstrand, K.R. Improving clinical visual detection--potential for caries clinical trials. *J. Dent. Res.* **2004**, *83*, C67–C71. [CrossRef]
10. Ekstrand, K.R.; Gimenez, T.; Ferreira, F.R.; Mendes, F.M.; Braga, M.M. The International Caries Detection and Assessment System—ICDAS: A Systematic Review. *Caries Res.* **2018**, *52*, 406–419. [CrossRef]
11. Ekstrand, K.R.; Ricketts, D.N.; Kidd, E.A. Reproducibility and accuracy of three methods for assessment of demineralization depth of the occlusal surface: An in vitro examination. *Caries Res.* **1997**, *31*, 224–231. [CrossRef]
12. Kühnisch, J.; Goddon, I.; Berger, S.; Senkel, H.; Bücher, K.; Oehme, T.; Hickel, R.; Heinrich-Weltzien, R. Development, methodology and potential of the new Universal Visual Scoring System (UniViSS) for caries detection and diagnosis. *Int. J. Environ. Res. Public Health* **2009**, *6*, 2500–2509. [CrossRef] [PubMed]
13. Nyvad, B.; Machiulskiene, V.; Baelum, V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. *Caries Res.* **1999**, *33*, 252–260. [CrossRef]
14. Felsch, M.; Meyer, O.; Schlickenrieder, A.; Engels, P.; Schonewolf, J.; Zollner, F.; Heinrich-Weltzien, R.; Hesenius, M.; Hickel, R.; Gruhn, V.; et al. Detection and localization of caries and hypomineralization on dental photographs with a vision transformer model. *NPJ Digit. Med.* **2023**, *6*, 198. [CrossRef]
15. Yoon, K.; Jeong, H.M.; Kim, J.W.; Park, J.H.; Choi, J. AI-based dental caries and tooth number detection in intraoral photos: Model development and performance evaluation. *J. Dent.* **2024**, *141*, 104821. [CrossRef]
16. Thanh, M.T.G.; Van Toan, N.; Ngoc, V.T.N.; Tra, N.T.; Giap, C.N.; Nguyen, D.M. Deep Learning Application in Dental Caries Detection Using Intraoral Photos Taken by Smartphones. *Appl. Sci.* **2022**, *12*, 5504. [CrossRef]
17. Al-Jallad, N.; Ly-Mapes, O.; Hao, P.; Ruan, J.; Ramesh, A.; Luo, J.; Wu, T.T.; Dye, T.; Rashwan, N.; Ren, J.; et al. Artificial intelligence-powered smartphone application, AICaries, improves at-home dental caries screening in children: Moderated and unmoderated usability test. *PLoS Digit. Health* **2022**, *1*, e0000046. [CrossRef] [PubMed]
18. Xiong, Y.; Zhang, H.; Zhou, S.; Lu, M.; Huang, J.; Huang, Q.; Huang, B.; Ding, J. Simultaneous detection of dental caries and fissure sealant in intraoral photos by deep learning: A pilot study. *BMC Oral Health* **2024**, *24*, 553. [CrossRef]
19. Chaves, E.T.; Vinayahalingam, S.; van Nistelrooij, N.; Xi, T.; Romero, V.H.D.; Flugge, T.; Saker, H.; Kim, A.; Lima, G.D.S.; Loomans, B.; et al. Detection of caries around restorations on bitewings using deep learning. *J. Dent.* **2024**, *143*, 104886. [CrossRef] [PubMed]
20. Azhari, A.A.; Helal, N.; Sabri, L.M.; Abduljawad, A. Artificial intelligence (AI) in restorative dentistry: Performance of AI models designed for detection of interproximal carious lesions on primary and permanent dentition. *Digit. Health* **2023**, *9*, 20552076231216681. [CrossRef]
21. Ma, T.; Zhu, J.; Wang, D.; Xu, Z.; Bai, H.; Ding, P.; Chen, X.; Xia, B. Deep learning-based detection of irreversible pulpitis in primary molars. *Int. J. Paediatr. Dent.* **2024**. *online ahead of print*. [CrossRef]
22. Caliskan, S.; Tuloglu, N.; Celik, O.; Ozdemir, C.; Kizilaslan, S.; Bayrak, S. A pilot study of a deep learning approach to submerged primary tooth classification and detection. *Int. J. Comput. Dent.* **2021**, *24*, 1–9. [CrossRef]
23. Kaya, E.; Gunec, H.G.; Gokyay, S.S.; Kutal, S.; Gulum, S.; Ates, H.F. Proposing a CNN Method for Primary and Permanent Tooth Detection and Enumeration on Pediatric Dental Radiographs. *J. Clin. Pediatr. Dent.* **2022**, *46*, 293–298. [CrossRef] [PubMed]
24. Ahn, Y.; Hwang, J.J.; Jung, Y.H.; Jeong, T.; Shin, J. Automated Mesiodens Classification System Using Deep Learning on Panoramic Radiographs of Children. *Diagnostics* **2021**, *11*, 1477. [CrossRef] [PubMed]
25. Jeon, K.J.; Ha, E.G.; Choi, H.; Lee, C.; Han, S.S. Performance comparison of three deep learning models for impacted mesiodens detection on periapical radiographs. *Sci. Rep.* **2022**, *12*, 15402. [CrossRef]
26. Kim, J.; Hwang, J.J.; Jeong, T.; Cho, B.H.; Shin, J. Deep learning-based identification of mesiodens using automatic maxillary anterior region estimation in panoramic radiography of children. *Dentomaxillofac. Radiol.* **2022**, *51*, 20210528. [CrossRef]
27. 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]
28. Schwendicke, F.; Singh, T.; Lee, J.H.; Gaudin, R.; Chaurasia, A.; Wiegand, T.; Uribe, S.; Krois, J.; on behalf of the IADR e-oral health network and the ITU WHO focus group AI for Health. Artificial intelligence in dental research: Checklist for authors, reviewers, readers. *J. Dent.* **2021**, *107*, 103610. [CrossRef]
29. Kühnisch, J.; Bücher, K.; Henschel, V.; Albrecht, A.; Garcia-Godoy, F.; Mansmann, U.; Hickel, R.; Heinrich-Weltzien, R. Diagnostic performance of the universal visual scoring system (UniViSS) on occlusal surfaces. *Clin. Oral Investig.* **2011**, *15*, 215–223. [CrossRef]
30. Pitts, N.B. How the detection, assessment, diagnosis and monitoring of caries integrate with personalized caries management. *Monogr. Oral Sci.* **2009**, *21*, 1–14. [CrossRef]
31. Kühnisch, J.; Janjic Rankovic, M.; Kapor, S.; Schüler, I.; Krause, F.; Michou, S.; Ekstrand, K.; Eggmann, F.; Lussi, A.; Neuhaus, K.; et al. Identifying and avoiding risk of bias in caries diagnostic studies. *J. Clin. Med.* **2021**, *15*, 3223. [CrossRef]
32. Zhang, Y.; Liao, H.; Xiao, J.; Jallad, N.A.; Ly-Mapes, O.; Luo, J. A Smartphone-Based System for Real-Time Early Childhood Caries Diagnosis. In *Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2020; pp. 233–242.
33. 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] [PubMed]

34. Park, E.Y.; Cho, H.; Kang, S.; Jeong, S.; Kim, E.K. Caries detection with tooth surface segmentation on intraoral photographic images using deep learning. *BMC Oral Health* **2022**, *22*, 573. [CrossRef] [PubMed]
35. Moharrami, M.; Farmer, J.; Singhal, S.; Watson, E.; Glogauer, M.; Johnson, A.E.W.; Schwendicke, F.; Quinonez, C. Detecting dental caries on oral photographs using artificial intelligence: A systematic review. *Oral Dis.* **2023**, *30*, 1765–1783. [CrossRef]
36. Ines Meurer, M.; Caffery, L.J.; Bradford, N.K.; Smith, A.C. Accuracy of dental images for the diagnosis of dental caries and enamel defects in children and adolescents: A systematic review. *J. Telemed. Telecare* **2015**, *21*, 449–458. [CrossRef]
37. Bottenberg, P.; Jacquet, W.; Behrens, C.; Stachniss, V.; Jablonski-Momeni, A. Comparison of occlusal caries detection using the ICDAS criteria on extracted teeth or their photographs. *BMC Oral Health* **2016**, *16*, 93. [CrossRef] [PubMed]

11. Literaturverzeichnis

1. World Health Organization. Global oral health status report: towards universal health coverage for oral health by 2030. 2022 Nov 18 [cited 2024 Oct 3];30–30. Available from: https://www.who.int/publications/i/item/9789240061484

2. World Health Organization. WHO/Europe calls for urgent action on oral disease as highest rates globally are recorded in European Region [Internet]. 2023 [cited 2024 Oct 3]. Available from: https://www.who.int/europe/news/item/20-04-2023-who-europe-calls-for-urgent-action-on-oral-disease-as-highest-rates-globally-are-recorded-in-european-region

3. Pitts NB, Twetman S, Fisher J, Marsh PD. Understanding dental caries as a non-communicable disease. Br Dent J. 2021 Dec;231(12):749–53.

4. Macey R, Walsh T, Riley P, Glenny AM, Worthington H V, O'Malley L, et al. Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries. Cochrane Database Syst Rev. 2021 Jun 14;6(6):CD014546.

5. Kapor S, Rankovic MJ, Khazaei Y, Crispin A, Schüler I, Krause F, et al. Systematic review and meta-analysis of diagnostic methods for occlusal surface caries. Clin Oral Investig. 2021 Aug;25(8):4801–15.

6. Gimenez T, Tedesco TK, Janoian F, Braga MM, Raggio DP, Deery C, et al. What is the most accurate method for detecting caries lesions? A systematic review. Community Dent Oral Epidemiol. 2021 Jun;49(3):216–24.

7. Gimenez T, Piovesan C, Braga MM, Raggio DP, Deery C, Ricketts DN, et al. Visual Inspection for Caries Detection. J Dent Res. 2015 Jul 20;94(7):895–904.

8. Moharrami M, Farmer J, Singhal S, Watson E, Glogauer M, Johnson AEW, et al. Detecting dental caries on oral photographs using artificial intelligence: A systematic review. Oral Dis. 2024 May;30(4):1765–83.

9. Prados-Privado M, García Villalón J, Martínez-Martínez CH, Ivorra C, Prados-Frutos JC. Dental Caries Diagnosis and Detection Using Neural Networks: A Systematic Review. J Clin Med. 2020 Nov 6;9(11).

10. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. J Dent Res. 2020 Jul;99(7):769–74.

11. Katsumata A. Deep learning and artificial intelligence in dental diagnostic imaging. Jpn Dent Sci Rev. 2023 Dec;59:329–33.

12. Mohammad-Rahimi H, Motamedian SR, Rohban MH, Krois J, Uribe SE, Mahmoudinia E, et al. Deep learning for caries detection: A systematic review. J Dent. 2022 Jul;122:104115.

13. Askar H, Krois J, Rohrer C, Mertens S, Elhennawy K, Ottolenghi L, et al. Detecting white spot lesions on dental photography using deep learning: A pilot study. J Dent. 2021 Apr;107:103615.

14. Kühnisch J, Meyer O, Hesenius M, Hickel R, Gruhn V. Caries Detection on Intraoral Images Using Artificial Intelligence. J Dent Res. 2022 Feb 20;101(2):158–65.

15. Park EY, Cho H, Kang S, Jeong S, Kim EK. Caries detection with tooth surface segmentation on intraoral photographic images using deep learning. BMC Oral Health. 2022 Dec 7;22(1):573.

16. Felsch M, Meyer O, Schlickenrieder A, Engels P, Schönewolf J, Zöllner F, et al. Detection and localization of caries and hypomineralization on dental photographs with a vision transformer model. NPJ Digit Med. 2023 Oct 25;6(1):198.

17. Kühnisch J, Bücher K, Henschel V, Albrecht A, Garcia-Godoy F, Mansmann U, et al. Diagnostic performance of the universal visual scoring system (UniViSS) on occlusal surfaces. Clin Oral Investig. 2011 Apr;15(2):215–23.

18. Kühnisch J, Goddon I, Berger S, Senkel H, Bücher K, Oehme T, et al. Development, methodology and potential of the new Universal Visual Scoring System (UniViSS) for caries detection and diagnosis. Int J Environ Res Public Health. 2009 Sep;6(9):2500–9.

19. Nyvad B, Machiulskiene V, Baelum V. Reliability of a New Caries Diagnostic System Differentiating between Active and Inactive Caries Lesions. Caries Res. 1999;33(4):252–60.

20. Pitts NB. How the detection, assessment, diagnosis and monitoring of caries integrate with personalized caries management. Monogr Oral Sci. 2009;21:1–14.

21. Pitts N, Ekstrand K. International Caries Detection and Assessment System and its International Caries Classification and Management System – methods for staging of the caries process and enabling dentists to manage caries. Community Dent Oral Epidemiol. 2013 Feb 25;41(1).

22. Frenkel E, Neumayr J, Schwarzmaier J, Kessler A, Ammar N, Schwendicke F, et al. Caries Detection and Classification in Photographs Using an Artificial Intelligence-Based Model-An External Validation Study. Diagnostics (Basel). 2024 Oct 14;14(20).

23. Bora A, Balasubramanian S, Babenko B, Virmani S, Venugopalan S, Mitani A, et al. Predicting the risk of developing diabetic retinopathy using deep learning. Lancet Digit Health. 2021 Jan;3(1):e10–9.

24. Chen W, Li R, Yu Q, Xu A, Feng Y, Wang R, et al. Early detection of visual impairment in young children using a smartphone-based deep learning system. Nat Med. 2023 Feb 26;29(2):493–503.

25. Fu WT, Zhu QK, Li N, Wang YQ, Deng SL, Chen HP, et al. Clinically Oriented CBCT Periapical Lesion Evaluation via 3D CNN Algorithm. J Dent Res. 2024 Jan 15;103(1):5–12.

Danksagung

An erster Stelle möchte ich meinem Doktorvater, Herrn Prof. Dr. Kühnisch meinen tief empfundenen Dank aussprechen. Seine unermüdliche Unterstützung, sein wissenschaftlicher Rat und sein Vertrauen in meine Arbeit haben diese Dissertation möglich gemacht. Durch seine geduldige Begleitung und seinen Enthusiasmus für Forschung hat er mir nicht nur fachliche Expertise vermittelt, sondern mich auch persönlich inspiriert.

Ein besonderer Dank gilt auch Dr. Helena Dujic, deren wertvolle Ratschläge und Unterstützung in entscheidenden Momenten einen großen Beitrag zu dieser Arbeit geleistet haben. Ihre Expertise und ihr Engagement haben mich stets motiviert und in meiner Forschung bestärkt.

Ich bin zutiefst dankbar, diese Arbeit unter der Anleitung von solch engagierten und herausragenden Wissenschaftler durchführen zu dürfen.

Mein besonderer Dank gilt meinen Eltern und meiner Oma, die mir das Studium ermöglicht haben und immer an meine Fähigkeiten glauben.

Ferner danke ich meinem Ehemann, meinen Kindern sowie meinen Schwiegereltern, die mich auf dem Weg begleitet und immer unterstützt haben.