

Aus der Poliklinik für Zahnerhaltung und Parodontologie  
der Ludwig-Maximilians-Universität München



***Automatisierte Erkennung von Fissurenversiegelungen mithilfe  
von künstlicher Intelligenz***

Dissertation  
zum Erwerb des Doktorgrades der Zahnmedizin  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität München

vorgelegt von  
Anne Katharina Schlickerrieder

aus  
Rosenheim

Jahr  
2022

---

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

Erster Gutachter: *Prof. Dr. med. dent. Jan Kühnisch*

Zweiter Gutachter: *Prof. Dr. Roswitha Heinrich-Weltzien\_*

Dritter Gutachter: *Prof. Dr. Nikolaos Koutsouleris*

ggf. weitere Gutachter: \_\_\_\_\_




\_\_\_\_\_

Mitbetreuung durch den  
promovierten Mitarbeiter: \_\_\_\_\_

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 25.10.2022

## Affidavit

|   |   |   |   |   |
|---|---|---|---|---|
|  | LUDWIG-<br>MAXIMILIANS-<br>UNIVERSITÄT<br>MÜNCHEN | Promotionsbüro<br>Medizinische Fakultät |  |  |
| <b>Eidesstattliche Versicherung</b>   |   |   |   |   |

Schlickenrieder, Anne

Name, Vorname

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

**„Automatisierte Erkennung von Fissurenversiegelungen mithilfe von künstlicher Intelligenz“**

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.

Köln, 18.11.2022

Ort, Datum

Anne Schlickenrieder

Unterschrift Doktorandin bzw. Doktorand

# Inhaltsverzeichnis

|   |           |
|---|-----------|
| <b>Affidavit</b> .....                                    | <b>3</b>  |
| <b>Inhaltsverzeichnis</b> .....                           | <b>4</b>  |
| <b>Abkürzungsverzeichnis</b> .....                        | <b>5</b>  |
| <b>Publikationsliste</b> .....                            | <b>6</b>  |
| <b>1. Beitrag zur Veröffentlichung</b> .....              | <b>7</b>  |
| <b>2. Einleitung</b> .....                                | <b>8</b>  |
| 2.1 Ziel des Projektes .....                              | 8         |
| 2.2 Begriffserklärungen zur künstlichen Intelligenz ..... | 8         |
| 2.3 Projektrealisierung .....                             | 9         |
| 2.4 Daten vergleichbarer Studienprojekte .....            | 11        |
| 2.5 Weiteres Vorgehen .....                               | 13        |
| <b>3. Zusammenfassung</b> .....                           | <b>14</b> |
| <b>4. Abstract (English)</b> .....                        | <b>15</b> |
| <b>5. Paper I</b> .....                                   | <b>16</b> |
| <b>6. Literaturverzeichnis</b> .....                      | <b>26</b> |
| <b>Danksagung</b> .....                                   | <b>28</b> |
| <b>Lebenslauf</b> .....                                   | <b>29</b> |

## Abkürzungsverzeichnis

|     |                                   |
|-----|-----------------------------------|
| ACC | Accuracy                          |
| AUC | Area under curve                  |
| CNN | Convolutional Neural Network      |
| FV  | Fissurenversiegelung              |
| KI  | Künstliche Intelligenz            |
| NN  | Neuronal Network                  |
| ROC | Receiver operating characteristic |

# Einleitung

## 1.1 Ziel des Projektes

Die künstliche Intelligenz (KI) gewinnt in der Medizin und Zahnmedizin immer mehr an Bedeutung, um mithilfe von mehrschichtigen, neuronalen Netzwerken (engl. «convolutional neural network», CNN) u.a. pathologische Befunde automatisiert erkennen und klassifizieren zu können (Schwendicke et al., 2020). In der Zahnmedizin wurden bislang größtenteils Röntgenbilder verwendet, um an diesen Algorithmen der KI zu entwickeln. So wurde beispielsweise die Karieserkennung mithilfe von CNNs bereits an Bissflügelaufnahmen (Bejnordi et al., 2018; Cantu et al., 2020; Geetha et al., 2020; Khan et al., 2020; Lee et al., 2018a; Park und Park, 2018), apikalen Röntgenbildern (Lee et al., 2018b) oder Orthopantomogramm-Aufnahmen (Vinayahalingam et al., 2021) durchgeführt. Dagegen wurden klinische Fotos bisher nur in einer geringen Anzahl von Studien verwendet, um diese mithilfe einer KI-Technologie auszuwerten (Askar et al., 2021; Kühnisch et al., 2021; Moutselos et al., 2019; You et al., 2020; Zhang et al., 2020). Ziel des aktuellen Dissertationsprojektes war es, eine KI zu entwickeln, die Fissurenversiegelungen (FV) erkennen und kategorisieren kann. FV dienen dem Zweck, die Kariesentstehung an anfälligen Okklusalfächen zu unterbinden bzw. vorhandene Kariesvorstufen zu arretieren (Welbury et al., 2004). Nach Kenntnis der Arbeitsgruppe wurden bislang keine diesbezüglichen Studien veröffentlicht, welche sich der automatisierten Erkennung und Kategorisierung von FV mit den Möglichkeiten der KI widmete. Hierfür wurde die KI mit klinischen Bildern und den dazugehörigen Befunden gespeist, die sie entsprechend einer vorgegebenen Klassifizierung einteilen musste.

## 1.2 Begriffserklärungen zur künstlichen Intelligenz

Mithilfe von KI sollen Computer bzw. Maschinen befähigt werden, Aufgaben zu übernehmen, die normalerweise vom Menschen ausgeführt werden. Bei dem sogenannten «machine learning», das einen Teilbereich der KI darstellt, werden Algorithmen angewendet, um die statistischen Muster und Strukturen von Daten

zu erlernen (Schwendicke et al., 2020). Dadurch können Vorhersagen von unbekannten Daten vorgenommen werden (Schwendicke et al., 2020). Das sogenannte «deep learning» ist eine Methode des maschinellen Lernens. Es kann abstrakte Datenrepräsentationen erlernen, die beispielsweise zur Datenklassifikation verwendet werden können. Bei dem «deep learning» muss hierfür kein Merkmals-Extrahierer entwickelt werden, da die wichtigsten Merkmale der gegebenen Daten automatisch durch allgemeine Lernprozeduren erst ermittelt und dann erlernt werden. Dies geschieht mit Rückwärtspropagierungs-Algorithmen, die beurteilen, wie die inneren Parameter des mehrschichtigen Modells umgeändert werden müssen. Dadurch können die Repräsentationen in einer Schicht des Modells aus der vorangehenden Schicht berechnet werden. Künstliche neuronale Netze («neural network», NN) dienen oftmals als ein solches mehrschichtiges Rechenmodell. Sogenannte CNN, also zu deutsch «faltende Netze», sind eine Weiterentwicklung von NN. CNN sind besonders zur Verarbeitung von mehrdimensionalen Daten-Arrays geeignet. (LeCun et al., 2019).

### **1.3 Projektrealisierung**

Die Umsetzung des vorliegenden Publikationsprojektes folgte den üblichen Vorgehensweisen. Um die diagnostische Genauigkeit von Beginn an sicherzustellen, wurde zunächst ein Kalibrierungstraining durchlaufen. Hierbei handelte es sich um einen zweitägigen Workshop, bei dem unter anderem die korrekte Erkennung und Kategorisierung von FV thematisiert wurde.

Mit dem Ziel der Entwicklung eines CNN für die Erkennung von FV, konnten in der Arbeitsgruppe insgesamt 2352 intraorale Fotos von Seitenzähnen identifiziert werden. Es wurden lediglich Bilder von gesunden Zähnen oder versiegelten Zahnoberflächen verwendet, die keine Pathologien aufwiesen, um ein sicheres Erlernen der KI zu ermöglichen. Gleichfalls wurden mangelhafte Aufnahmen oder Duplikate von identischen Bildern oder Zahnoberflächen im Voraus ausgeschlossen. Als Kategorisierungsschema wurden international anerkannte Kriterien gewählt (Heinrich-Weltzien und Kühnisch, 1999; Simonsen, 1991), in dem die Qualität der Fissurenversiegelungen mithilfe von vier Unterkategorien beurteilt werden. Mit «0» werden Bilder beurteilt, die eine Okklusalfäche ohne Versiegelung

abbilden. «1» zeigt Bilder von Okklusalflächen mit einer klinisch intakten FV. Bilder, die mit «2» beurteilt werden, zeigen Okklusalflächen mit einer suffizienten FV, das heißt, dass eine Retention des Fissurenversiegelungsmaterials in der Hauptfissur oder ein Materialverlust von mehr als einem Drittel des Fissurenreliefs vorliegt. Für Bilder, die mit «3» beurteilt werden, wird die Versiegelung als insuffizient wahrgenommen. Der Materialverlust ist nahezu vollständig und die Hauptfissur ist erneut freigelegt und frei von jeglichem Versiegelungsmaterial (Simonsen, 1991). Diese Unterteilung wurde vorgenommen, um der praktischen Evaluierung in einer zahnärztlichen Praxis möglichst nahe zu kommen.

Im Anschluss an die individuelle Befundung wurden alle Entscheidungen in der Gruppe gemeinsam überprüft, um mögliche Fehler zu eliminieren. Pro Bild wurde eine finale Diagnoseentscheidung festgelegt, die anschließend als Referenzstandard für das zyklische Training und die wiederholte Auswertung des «deep-learning» basierten CNN diente.

Die Programmierung des CNN erfolgte durch den Projektpartner an der Universität Duisburg–Essen. Für das Training des CNN wurde das ResNeXt-101-32x (Xie et al., 2017) gewählt. Zudem wurde der gesamte Bilddatensatz, also 2352 Bilder, von denen 461 unversiegelte Zahnflächen und 1891 versiegelte Zahnflächen zeigen, in einen Trainingsatz und einen Testatz unterteilt. Der Testatz wurde der CNN nie zur Verfügung gestellt und gilt somit als unabhängig. Im Anschluss an die Programmierung erfolgte die deskriptive und explorative Datenauswertung.

Im Ergebnis der Arbeiten konnte gezeigt werden, dass die entwickelten Algorithmen der KI in 98,7% der Fälle eine korrekte Entscheidung treffen konnten. Schwierigkeiten bestanden in der Differenzierung von einzelnen Unterkategorien. Hier konnten insgesamt nur etwa 90% der Fälle richtig zugeordnet werden. Die erreichten diagnostischen Gütedaten reihen sich in die aktuell verfügbaren Daten aus vergleichbaren Studienprojekten ein.



## 1.4 Daten vergleichbarer Studienprojekte

Im Folgenden sollen die Daten der Projekte von Askar et al. (2021), Kühnisch et al. (2021), Moutselos et al. (2019), You et al. (2020) und Zhang et al. (2020) in Bezug auf Unterschiede und Gemeinsamkeiten zur vorliegenden Arbeit (Schlickerieder et al., 2021) dargestellt werden. Im Folgenden soll die Herangehensweise und die Ergebnisse der unterschiedlichen Studien beleuchtet werden.

Moutselos et al. (2019) nutzten die sogenannte Superpixel Segmentierung. Hier lagen drei Indizes für die Rückführung auf Superpixel vor: „most common“ (häufigste Pixelklasse eines Superpixels), „centroid pixel class“ (Superpixel nimmt die Klasse des geometrischen Zentrumpixels an) und „worst class“ (Superpixel wird durch höchste in seinem Bereich gefundene Klasse charakterisiert). Bei der Superpixel Segmentierung von Moutselos et al. (2019) wird je Index eine Treffgenauigkeit („accuracy“, ACC) angegeben. Diese belief sich bei dem „most common“ Index auf 0,6, bei dem „centroid pixel class“ Index auf 0,6 und bei dem „worst class“ Index auf 0,68 (Moutselos et al., 2019).

Askar et al. (2021) untersuchten sogenannte „white spot“ Läsionen und nahmen eine Einteilung von drei Unterkategorien vor. Zum Anlernen der KI wurden Markierungen der Fotos vorgenommen, die den Zahnumriss und den white spot umrandeten. Insgesamt wurde eine ACC von 0,81 – 0,84 erzielt, wobei keine großen Unterschiede innerhalb der Einteilung der Unterkategorien festgestellt wurden (Askar et al., 2021).

Kühnisch et al. (2021) gaben der KI drei unterschiedliche Karieskategorien vor, in die die jeweiligen Läsionen auf den Bildern eingeteilt werden sollten. Es wurden keine Markierungen vorgenommen. Die entwickelte KI von Kühnisch et al. (2021) erzielte eine ACC von 92,5% für die Klassifizierung von Karies. Innerhalb der Unterkategorien kam es zu einer größeren Variabilität, die von 79,5% bis 90,6% reichte (Kühnisch et al., 2021).

In der vorliegenden Arbeit (Schlickerieder et al., 2021) wurden der KI vier unterschiedliche Kategorien vorgegeben, mithilfe derer die Qualität der Fissurenversiegelungen möglichst praxisnah beurteilt werden sollten. Es wurden keine Markierungen vorgenommen. Die entwickelte KI lag mit einer ACC von 89,6% hoch.

Die Genauigkeit bei der Einteilung der Unterkategorien unterschied sich von 83,2% ACC bis zu 92,4% (Schlickerieder et al., 2021).

You et al. (2020) untersuchten die Erkennung von Zahnplaque. Hier wurden Markierungen der Zahnflächen vorgenommen, die durch einen Plaquemarker angefärbt wurden, um so die KI zu trainieren. You et al. (2020) verglichen die Leistungsfähigkeit eines Zahnarztes und der KI mithilfe der intersection-over-union-metric. Diese berechnet vereinfacht gesagt das Verhältnis zwischen der „intersection“, zu Deutsch Schnittmenge, also den richtig positiven Ergebnissen und der „union“, zu Deutsch Vereinigung, also der Summe von richtig positiven, falschen negativen und falsch positiven Ergebnissen. In der vorliegenden Studie wurde folglich das Verhältnis zwischen der Schnittmenge und der „union“ von zwei Mengen berechnet. Die zwei Mengen waren in diesem Fall die Stellen, die tatsächlich dentale Plaque zeigen, der zuvor vom Plaquemarker markiert wurde und die Stellen, die von der KI oder dem Zahnarzt\*in als Plaque identifiziert wurden. Die KI erzielte eine höhere Metric mit  $0,74 \pm 0,17$  im Vergleich zum Zahnarzt\*in ( $0,65 \pm 0,2$ ) (You et al., 2020). Dieser Wert kann mit keinem Wert der vorliegenden Studie verglichen werden, da die Herangehensweise von You et al. (2020) sich grundlegend unterscheidet.

Zhang et al. (2020) überprüften ebenfalls klinische Fotos nach Karies mithilfe von KI. Es wurden keine Unterkategorien zur genaueren Unterteilung vorgegeben. Zum Anlernen der KI wurden die Kariesläsionen mit einer Box markiert. Die Aufgabe der KI bestand darin, anzugeben, ob Kariesläsionen in dem Bild zu erkennen sind, und, falls ja, wo sich die Läsion befindet. Zhang et al. (2020) erzielten nach einem zweimaligem sogenannten „Hard negative mining“ eine hohe Genauigkeit bei der Klassifizierung und Lokalisierung der Karies, auch im Vergleich zu den vorliegenden Ergebnissen ohne oder mit einmaligem „hard negative mining“. Insgesamt lag der sogenannte AUC („area under curve“) Wert bei 85,65% (Zhang et al., 2020). Dieser ist jedoch nicht gleichzusetzen mit der oben häufig genannten „accuracy“. Die AUC Kurve ist eine einzelne Zahl, die die Fläche unter der sogenannten ROC („receiver operator characteristic“) Kurve darstellt. Die ROC Kurve wiederum stellt die Leistung des Algorithmus dar und liefert Informationen über die Leistung bei einer Reihe von Schwellenwerten (Muschelli, 2020). Um

einen Vergleich ziehen zu können, kann der AUC Wert der vorliegenden Studie genutzt werden, der bei 0,996 lag (Schlickerieder et al. 2021).

Die Trainingszyklen oder auch Epochen reichten von 10, wie in dem vorliegenden Paper (Schlickerieder et al., 2021) und bei Askar et al. (2021) über 50 (Kühnisch et al., 2021), bis hin zu insgesamt 100 Zyklen (Moutselos et al., 2019). Eine höhere Zahl an Epochen sorgt für eine Verbesserung des Erstergebnisses der KI und so kann diese potenziell bessere Ergebnisse erzielen. Ebenso unterschied sich die Bilderanzahl. Moutselos et al. (2019) untersuchten insgesamt 88 Bilder, Askar et al. (2021) erhielten aus einem Bildsatz von 434 Bildern insgesamt 2781 einzelne Zahnbilder, You et al. (2020) untersuchten 886 Bilder, in der vorliegenden Arbeit konnten 2352 Bilder genutzt werden (Schlickerieder et al., 2021), Kühnisch et al. (2021) standen 2417 Bilder zur Verfügung, und in der Arbeit von Zhang et al. (2020) liegt der größte Bilderdatensatz mit 3932 Bildern vor. Je höher der Bilddatensatz ist, desto mehr Lernmaterial steht der KI zur Verfügung, und entsprechend genauer dürfte sie werden.

## **1.5 Weiteres Vorgehen**

Mit dem Ziel, die Genauigkeit von KI-basierten Algorithmen weiter zu verbessern, besteht die Notwendigkeit pathologische Befunde oder typische Bildmerkmale in einem Trainingsdatensatz exakt zu markieren.

So hat die KI extrem präzise Vorgaben und kann die Bilder künftig noch genauer zuordnen und auswerten. Diese pixelgenaue Markierung wird unabhängig von der vorliegenden Dissertation fortgeführt. Ebenso lassen sich bessere Ergebnisse durch einen größeren Bildersatz und eine gesteigerte Anzahl an Trainingszyklen erzielen.

Die entwickelte KI könnte beispielsweise Teil eines Computerprogramms für Zahnarztpraxen werden, das als Hilfe für Diagnosestellungen verwendet wird. Durch die Entwicklung weiterer KIs, die vor allem durch die vorherige pixelgenaue Markierung noch präziser arbeiten, könnten verschiedenste Pathologien und Diagnosestellungen berücksichtigt und in das Programm eingebaut werden und so ein umfassendes Behandlungsspektrum abdecken.

## Zusammenfassung

Das Ziel der vorliegenden Arbeit lag zuerst darin, eine KI zu entwickeln, deren Aufgabe darin besteht, Fissurenversiegelungen auf klinischen Fotos automatisiert zu erkennen und im zweiten Schritt eine Qualitätsbeurteilung vorzunehmen. Hierfür wurden 2352 Bilder verwendet, die zunächst sortiert und zugeschnitten wurden. Die anschließende Kategorisierung der Bilder wurde von drei trainierten und kalibrierten Untersucherinnen vorgenommen und anschließend von einem Experten (>20 Jahre Berufserfahrung) kontrolliert. Im Fall von Unstimmigkeiten wurde ein Konsens erarbeitet. Diese konsensbasierte Entscheidung diente als Referenzstandard für das spätere Training der KI und deren Validierung. Folgend wurden die Bilder in einen Trainings- und einen Testsatz unterteilt, und am Lehrstuhl für Software-Engineering der Universität Duisburg-Essen wurde damit das künstliche neuronale Netz trainiert. Hierfür wurde das CNN ResNeXt-101-32x8d als vortrainiertes Netzwerk genutzt, welches insgesamt 10 Epochen durchlief. Diese eher niedrige Learning-Rate könnte in Zukunft durch eine höhere Anzahl an Epochen gesteigert werden. Als statistische Parameter wurden die üblichen statistischen Güteparameter sowie die Fläche unter den ROC-Kurven (AUC) berechnet. Im Ergebnis der Arbeit wurde gezeigt, dass die diagnostische Genauigkeit zur Erkennung von Fissurenversiegelungen bei 98,7% lag, der AUC-Wert betrug 0,996. Die Akkuratheit für die Zuordnung der Unterkategorien befand sich im Bereich von etwa 90%. Auf Grundlage der gezeigten Ergebnisse konnte geschlossen werden, dass das trainierte künstliche neuronale Netzwerk eine sehr gute Präzision lieferte, um Fissurenversiegelungen automatisiert erkennen zu können. Die kategoriale Beurteilung zeigte noch Schwächen, welche mit dem gewählten Trainingsansatz allerdings erwartbar waren. Daher sind weitere Anstrengungen notwendig, um die Genauigkeit zu verbessern. Als Ansatz kann dazu die präzise Markierung von Fissurenversiegelungen auf vorhandenen Bildern gewählt werden, um der KI möglichst exakte Informationen bereit zu stellen.

## **Abstract (English)**

The aim of the present work was first to develop an AI whose task is to automatically recognise fissure sealings on clinical photos and, in a second step, to carry out a quality assessment. For this purpose, 2,352 images were used, which were first sorted and cropped. The subsequent categorisation of the images was carried out by three skilled and calibrated examiners and then controlled by an expert (>20 years of professional experience). In case of disagreement, a consensus was worked out. This consensus-based decision served as a reference standard for the later training of the AI and its validation. Subsequently, the images were divided into a training set and a test set and the artificial neural network was trained with them at the Chair of Software Engineering at the University of Duisburg-Essen. For this purpose, the CNN ResNeXt-101-32x8d was used as a pre-trained network, which ran through a total of 10 epochs. This rather low learning rate could be increased in the future by a higher number of epochs. The usual statistical quality parameters and the area under the ROC curves (AUC) were calculated as statistical parameters. As a result of the work, it was shown that the diagnostic accuracy for the detection of fissure sealants was 98.7%, the AUC value was 0.996. The accuracy for the assignment of the subcategories was in the range of about 90%. Based on the results shown, it can be concluded that the trained artificial neural network provided very good precision to automatically detect fissure sealants. The categorical assessment still showed weaknesses, which were, however, to be expected with the selected training approach. Therefore, further efforts are necessary to improve the accuracy. As an approach, the precise marking of fissure sealants on existing images can be chosen in order to provide the AI with the most accurate information possible.

## Publikationsliste

Schlickenrieder A, Meyer O, Schönewolf J, Engels P, Hickel R, Gruhn V, Hesenius M, Kühnisch J (2021). Automatized detection and categorization of fissure sealants from intraoral digital photographs using artificial intelligence. *Diagnostics* **2021**, *11*, 1608. <https://www.mdpi.com/2075-4418/11/9/1608>

## Beitrag zur Veröffentlichung

Im Rahmen des Dissertation- und Publikationsprojekts waren die nachfolgenden Arbeitsschritte notwendig. Ebenso können die Anteile, die mir und den verschiedenen Projektpartnern zuteil waren, abgelesen werden.

|  | Anne<br>Schlickenrieder | Prof. Dr.<br>Jan Kühnisch | Ko-Autoren                                       |
|--|-------------------------|---------------------------|--|
| <b>Ethikantrag</b>   | -                       | 100 %                     | -  |
| <b>Literaturrecherche</b>  | 100 %                   | -                         | -  |
| <b>Publikation "Automatized detection and categorization of fissure sealants from intraoral digital photographs using artificial intelligence"</b> |                         |                           |  |
| Projektidee  | 50 %                    | 50 %                      | -  |
| <i>Schritt 1: Training und Kalibrierung</i> der Arbeitsgruppe  | 50 %                    | 50 %                      |  |
| <i>Schritt 2: Bildselektion</i>  | 100 %                   | -                         | -  |
| <i>Schritt 3: Bildauswertung und -bearbeitung</i>  | 100 %                   | -                         | -  |
| <i>Schritt 4: Einteilung anhand des Klassifizierungssystems</i>  | 80 %                    | -                         | 10 % P. Engels;<br>10 % J. Schönewolf            |
| <i>Schritt 5: Programmierung der KI und Datenanalyse</i>   | -                       | -                         | Projektpartner an der Universität Duisburg-Essen |
| <i>Schritt 6: Datenselektion und grafische Aufarbeitung</i>  | 90 %                    | 10 %                      | -  |
| <i>Schritt 7: Erstellen und Einreichen des Manuskripts</i>   | 80 %                    | 10 %                      | 10 % Ko-Autoren                                  |
| <i>Schritt 8: Verfassen der kumulativen Dissertation</i>   | 100 %                   | -                         | -  |


## Paper I

Anne Schlickenrieder, Ole Meyer, Jule Schönewolf, Paula Engels, Reinhard Hickel, Volker Gruhn, Marc Hesenius, Jan Kühnisch. Automated detection and categorization of fissure sealants from intraoral digital photographs using artificial intelligence. *Diagnostics*, 2021;11(9): 1608



Communication

# Automatized Detection and Categorization of Fissure Sealants from Intraoral Digital Photographs Using Artificial Intelligence

Anne Schlickenrieder <sup>1</sup>, Ole Meyer <sup>2</sup>, Jule Schönewolf <sup>1</sup>, Paula Engels <sup>1</sup>, Reinhard Hickel <sup>1</sup>, Volker Gruhn <sup>2</sup>, Marc Hesenius <sup>2</sup> and Jan Kühnisch <sup>1,\*</sup> 

- <sup>1</sup> Department of Conservative Dentistry and Periodontology, University Hospital, Ludwig-Maximilians University Munich, 80336 Munich, Germany; anne.schlickenrieder@t-online.de (A.S.); juleschoenewolf@web.de (J.S.); paula.engels@icloud.com (P.E.); hickel@dent.med.uni-muenchen.de (R.H.); juleschoenewolf@web.de (J.S.); paula.engels@icloud.com (P.E.); hickel@dent.med.uni-muenchen.de (R.H.)
- <sup>2</sup> Institute for Software Engineering, University of Duisburg-Essen, 45147 Essen, Germany; ole.meyer@uni-due.de (O.M.); gruhn@adesso.de (V.G.); marc.hesenius@uni-due.de (M.H.)
- \* Correspondence: jkuehn@dent.med.uni-muenchen.de; Tel.: +49-89-4400-59301

**Abstract:** The aim of the present study was to investigate the diagnostic performance of a trained convolutional neural network (CNN) for detecting and categorizing fissure sealants from intraoral photographs using the expert standard as reference. An image set consisting of 2352 digital photographs from permanent posterior teeth (461 unsealed tooth surfaces/1891 sealed surfaces) was divided into a training set ( $n = 1881/364/1517$ ) and a test set ( $n = 471/97/374$ ). All the images were scored according to the following categories: unsealed molar, intact, sufficient and insufficient sealant. Expert diagnoses served as the reference standard for cyclic training and repeated evaluation of the CNN (ResNeXt-101-32x8d), which was trained by using image augmentation and transfer learning. A statistical analysis was performed, including the calculation of contingency tables and areas under the receiver operating characteristic curve (AUC). The results showed that the CNN accurately detected sealants in 98.7% of all the test images, corresponding to an AUC of 0.996. The diagnostic accuracy and AUC were 89.6% and 0.951, respectively, for intact sealant; 83.2% and 0.888, respectively, for sufficient sealant; 92.4 and 0.942, respectively, for insufficient sealant. On the basis of the documented results, it was concluded that good agreement with the reference standard could be achieved for automatized sealant detection by using artificial intelligence methods. Nevertheless, further research is necessary to improve the model performance.

**Keywords:** pit and fissure sealants; caries assessment; visual examination; clinical evaluation; artificial intelligence; convolutional neural networks; deep learning; transfer learning



**Citation:** Schlickenrieder, A.; Meyer, O.; Schönewolf, J.; Engels, P.; Hickel, R.; Gruhn, V.; Hesenius, M.; Kühnisch, J. Automatized Detection and Categorization of Fissure Sealants from Intraoral Digital Photographs Using Artificial Intelligence. *Diagnostics* **2021**, *11*, 1608. <https://doi.org/10.3390/diagnostics11091608>

Academic Editor: Jae-Hong Lee

Received: 2 August 2021

Accepted: 1 September 2021

Published: 3 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The availability of artificial intelligence (AI) methods has aroused increasing interest in developing convolutional neural networks (CNNs) for automatized detection and categorization of diagnostic images in medicine and dentistry to objectify the classification of pathological findings [1]. In dentistry, radiographs are mostly used as image sources for CNNs to identify pathologies. Specifically, caries detection has been trained on bitewings [2–7], apical radiographs [8] or panoramic X-rays [9]. By contrast, there have been few attempts to apply AI technology to assess clinical images, which can be interpreted as a machine-readable equivalent for visual inspection. This study is the first report of automatic detection and categorization of dental caries [10–13] or dental plaque [14] from clinical photographs. When considering the broad spectrum of pathological findings on dental hard tissue, e.g., caries, erosion or developmental disorders, as well as dental interventions, e.g., sealants, dental restorations or prosthodontic measures, it is evident that CNNs need to be trained separately for each of the aforementioned categories. The aim of this pioneering project on the automatized detection of dental materials was to identify and categorize opaque sealants, which is primarily justified by the frequent use of these sealants

in dental health services of industrialized nations [15]. Second, sealant materials constitute a uniform group of materials that are typically white and easily visually detectable on posterior teeth compared to other dental restorations. Consequently, it can be hypothesized that the learning of a CNN for detecting sealants from dental photographs represents a first step before considering other types of dental restorations. Therefore, in this diagnostic study, the ability of a CNN to detect and categorize fissure sealants was investigated (as a test method) using digital photographs of posterior teeth, and the diagnostic outcome was compared with expert evaluation (the reference standard).

## 2. Materials and Methods

### 2.1. Study Design

The reporting of this study followed the recommendations of the Standard for Reporting of Diagnostic Accuracy Studies (STARD) steering committee [16] and topic-related recommendations [17].

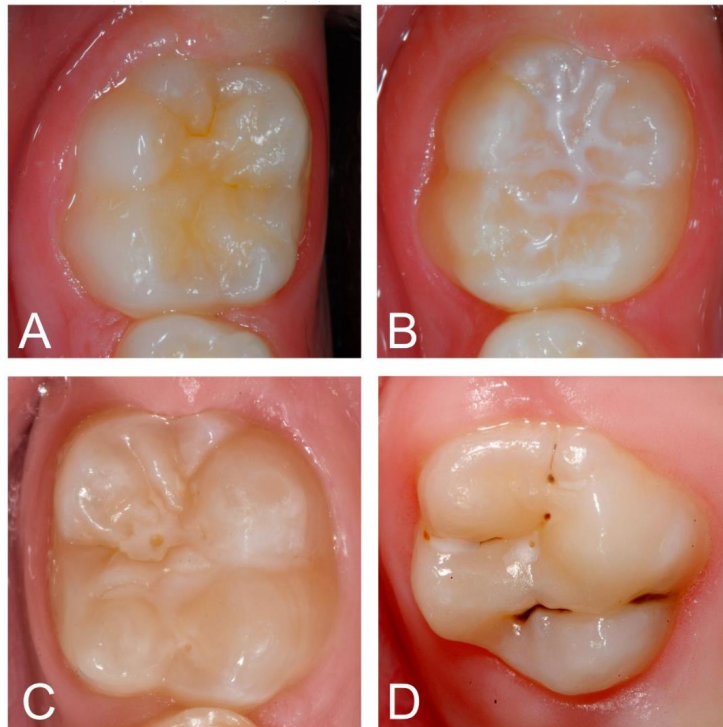
### 2.2. Photographic Images

All the images were taken for use in previous studies, as well as for clinical or teaching purposes, by an experienced dentist (J.K.). All the images were photographed using a professional single reflex lens camera (Nikon D300, D7100 or D7200 with a Nikon Micro 105-mm lens; Nikon, Tokyo, Japan) and Macro Flash EM-140 DG (Sigma, Rödermark, Germany) after tooth cleaning and drying. Molar teeth were photographed indirectly using intraoral mirrors (Reflect-Rhod, Hager and Werken, Duisburg, Germany) that were heated before being positioned in the oral cavity to prevent condensation on the mirror surface.

To ensure the best possible image quality, deficient photographs, e.g., out-of-focus images or images with saliva contamination, were excluded. Furthermore, duplicate photos from identical teeth or surfaces were removed from the dataset. This selection step ensured there were no repetitions in the included clinical photographs. All jpeg images (RGB format, resolution 1200 × 1200 pixel, no compression) were cropped to an aspect ratio of 1:1 and/or rotated in a standard manner using professional image editing software (Affinity Photo, Serif, Nottingham, UK) until, finally, the tooth surface filled most of the frame. Considering the study aim, images from healthy teeth or sealed surfaces were also included. Photographs with (additional) cavitated caries lesions or other hard tissue defects, e.g., enamel hypomineralization, hypoplasia, extensive tooth wear, and direct and indirect restorations, were excluded. Finally, 2352 anonymized, high-quality clinical photographs from permanent posterior teeth and the corresponding occlusal surfaces were included.

### 2.3. Categorization of Sealants (Reference Standard)

Each image was examined on a computer to detect and categorize fissure sealants using well-accepted international classification systems [18,19]. The following categories were used: 0—occlusal surfaces with no sealant; 1—occlusal surfaces with a clinically intact fissure sealant (up to one third loss of material in the periphery of the fissure pattern); 2—occlusal surface with a sufficient fissure sealant (retention of the material in the main fissure or loss of material exceeding one third of the fissure pattern); 3—insufficient (nearly complete loss of material and re-exposure of the main fissures) (Figure 1). Each of the given diagnostic categories is typically linked with different treatment modalities in daily dental practice and, in consequence, the quality staging appears of clinical relevance and justifies its scientific consideration in the present study. All the images were pre-labeled by a group of three graduated dentists and subsequently independently counterchecked by an experienced examiner (J.K., >20 years of clinical practice and scientific experience). In the case of divergent opinions, each image was discussed until a consensus was reached. Each diagnostic decision—one per image—served as a reference standard for cyclic training and repeated evaluation of the deep-learning-based CNN.



**Figure 1.** Example clinical images for each category: unsealed molar (A) and intact (B), sufficient (C) and insufficient fissure sealant (D).

All the annotators (A.S., J.S., P.E.) were trained during a 2-day workshop by the principal investigator (J.K.) and calibrated before beginning the study. The intra- and inter-examiner reproducibility was determined using 60 photographs, and the corresponding Kappa values showed at least a substantial capability for detecting and categorizing fissure sealants. The intra-/inter-examiner reproducibilities were 0.784/0.753 (A.S.), 0.779/0.752 (J.S.) and 0.779/0.752 (P.E.).

#### *2.4. Programming and Configuration of the Deep-Learning-Based CNN for Sealant Detection and Categorization (Test Method)*

The CNN was trained stepwise using a pipeline of established procedures, mainly image augmentation and transfer learning. Before training, the entire image set (2352 images/461 unsealed tooth surfaces/1891 sealed surfaces) was divided into a training set ( $n = 1881/364/1517$ ) and a test set ( $n = 471/97/374$ ). The latter was never made available to the CNN as training material and served as an independent test set.

Image augmentation was used to provide a large number of variable images to the CNN on a recurring basis. For this purpose, the randomly selected images (batch size = 16) were multiplied by a factor of  $\sim 5$ , altered by image augmentation (random center and margin cropping by up to 30% each, random deletion up to 30%, random affine transformation up to 180 degrees, random perspective transformation up to a distortion of 0.5, and random changes in brightness, contrast and saturation up to 10%) and resized (to  $300 \times 300$  pixels) by using torchvision (version 0.9.1, <https://pytorch.org>) in conjunction with the PyTorch

library (version 1.8.1, <https://pytorch.org>). All the images were normalized to compensate for under- and overexposure.

ResNeXt-101-32x8d [20] was used as the basis for the continuous adaptation of CNN for sealant detection and categorization. The CNN was trained using backpropagation to determine the gradient for learning. Backpropagation was repeated iteratively for images and labels using the abovementioned batch size and parameters. Overfitting was prevented by first selecting a low learning rate (0.00005) and then performing dropout (at a rate of 0.5) on the final linear layers as a regularization technique [21]. To train the CNN, this step was repeated for 10 epochs. The cross entropy loss as an error function and the Adam optimizer (Betas 0.9 and 0.999, Epsilon  $10^{-8}$ ) were applied.

To accelerate the training process of the CNN, an open-source neural network with pretrained weights was employed (ResNeXt-101-32x8d pretrained on ImageNet, Stanford Vision and Learning Lab, Stanford University, Palo Alto, CA, USA). This step enabled the transfer of existing learning results to increase the efficiency of recognition of basic structures in the existing image set. The training of the CNN was executed on a university-based server with the following specifications: RTX A6000 48 GB (Nvidia, Santa Clara, CA, USA), i9 10850K 10x3.60 GHz (Intel Corp., Santa Clara, CA, USA) and 64 GB RAM.

### 2.5. Statistical Analysis

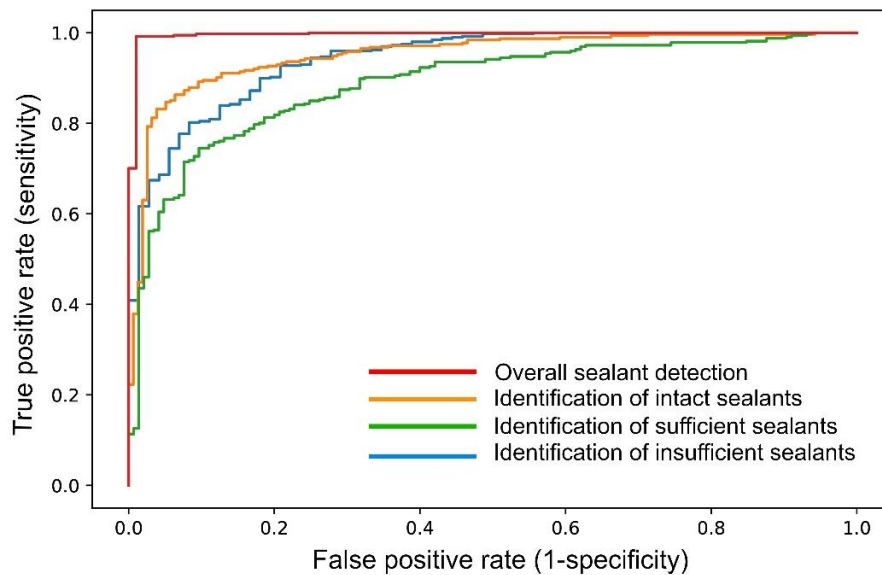
The data were analyzed using Python (<http://www.python.org>, version 3.8). The overall diagnostic accuracy ( $ACC = (TN + TP) / (TN + TP + FN + FP)$ ) was determined by calculating the number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) after using 25%, 50%, 75% and 100% of the images of the training data set. The sensitivity (SE), specificity (SP), positive and negative predictive values (PPV and NPV, respectively), and the area under the receiver operating characteristic (ROC) curve (AUC) were computed for the selected types of teeth and surfaces [22]. Saliency maps were plotted to identify image areas that are important for the CNN to make individual decisions. We calculated the saliency maps [23] by backpropagating the CNN prediction and visualized the gradient of the input of the resized images ( $300 \times 300$  pixels).

### 3. Results

The trained deep-learning-based CNN detected sealants correctly in 98.7% of all the test cases, corresponding to an AUC of 0.996 (Table 1, Figure 2). Additionally, the SE (96.9), SP (99.2), PPV (96.9) and NPV (99.2) were documented to be close to perfect (Table 1). By comparison, the model diagnostic performance was lower for the sealant subcategories (Table 1, Figure 2). Here, the AUC values were highest for the identification of intact sealants (0.951), followed by insufficient sealants (0.942) and sufficient sealants (0.888). These numbers, as well as the other performance data (Table 1), indicate that the automated identification of the subcategories in the present stage was less accurate than the simple detection of opaque sealant material from clinical photographs. The detailed case distribution was obtained from the confusion matrix (Figure 3). Here, the majority of incorrect decisions by the CNN occurred for categories other than the true classification, which indicates there were no major misclassifications. Most incorrect decisions were made for sufficient sealants. This observation is in line with the diagnostic parameters shown in Table 1. In addition to the descriptive and explorative data presentation, saliency maps (Figure 4) were plotted to illustrate the parts of each image that were used by the CNN for decision making.

**Table 1.** Overview of the diagnostic performance of the developed convolutional neural network (CNN), where the independent test set ( $n = 471$ ) was compared against independent expert evaluation of the caries detection level. The calculations were performed for different types of teeth, surfaces and training steps. In this context, the overall diagnostic accuracy (ACC), sensitivity (SE), specificity (SP), negative predictive value (NPV), positive predictive value (PPV) and area under the receiver operating characteristic curve (AUC).

| Diagnostic Categories                   | True Positives (TP) |      | True Negatives (TN) |      | False Positives (FP) |     | False Negatives (FN) |     | Diagnostic Performance |      |      |      |      |       |
|---|---------------------|------|---------------------|------|----------------------|-----|----------------------|-----|------------------------|------|------|------|------|-------|
|   | <i>n</i>            | %    | <i>n</i>            | %    | <i>n</i>             | %   | <i>n</i>             | %   | ACC                    | SE   | SP   | NPV  | PPV  | AUC   |
| Overall sealant detection               | 94                  | 20.0 | 371                 | 78.8 | 3                    | 0.6 | 3                    | 0.6 | 98.7                   | 96.9 | 99.2 | 99.2 | 96.9 | 0.996 |
| Identification of intact sealants       | 141                 | 29.9 | 281                 | 59.7 | 33                   | 7.0 | 16                   | 3.4 | 89.6                   | 89.8 | 89.5 | 94.6 | 81.0 | 0.951 |
| Identification of sufficient sealants   | 99                  | 21.0 | 293                 | 62.2 | 33                   | 7.0 | 46                   | 9.8 | 83.2                   | 68.3 | 89.9 | 86.4 | 75.0 | 0.888 |
| Identification of insufficient sealants | 52                  | 11.0 | 383                 | 81.3 | 16                   | 3.4 | 20                   | 4.3 | 92.4                   | 72.2 | 96.0 | 95.0 | 76.5 | 0.942 |



**Figure 2.** The ROC curves illustrate the model performance of the developed CNN for overall and categorical sealant detection.

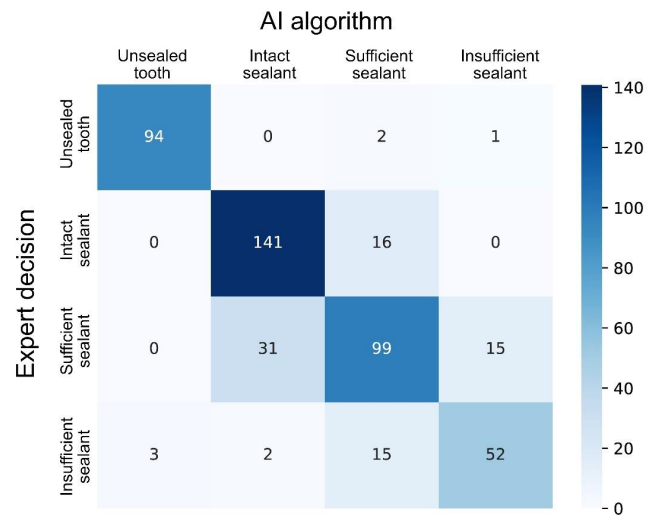


Figure 3. Confusion matrix showing the CNN classification performance for the test sample.

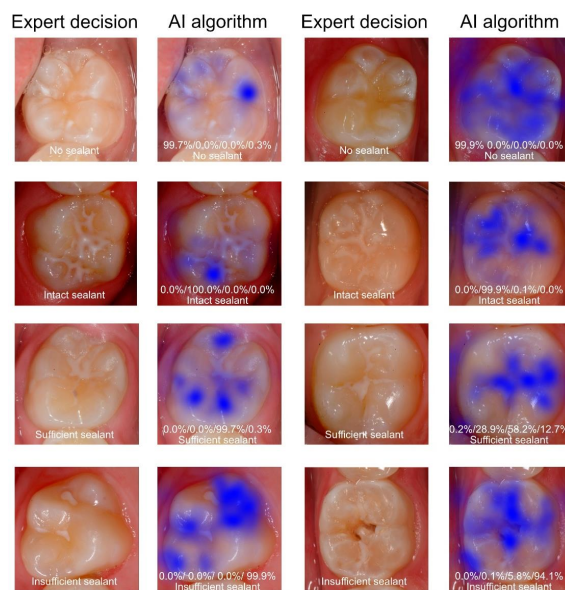


Figure 4. Examples of clinical images showing the reference decision and paired images with saliency maps visualizing the image areas (in blue) used in decision making by the AI method. The corresponding test results by the AI method are given for each example (unsealed tooth/intact fissure sealant/sufficient sealant/insufficient sealant).

#### 4. Discussion

The results of the present diagnostic study demonstrated that AI algorithms can detect and categorize sealants from machine-readable intraoral photographs. A high diagnostic accuracy of 98.7% and AUC of 0.996 were found (Table 1). Unlike this promising result, the CNN classified subcategories less accurately. Here, a diagnostic performance of approximately 90% accuracy was achieved (Table 1). In particular, sufficiently sealed occlusal surfaces were identified less reliably than the two other categories, which illustrates that further improvement is needed.

In addition, it can be concluded that the developed CNN can be used in future software applications and can identify sealants accurately with a high probability from intraoral photographs. To our knowledge, no comparable studies have been carried out thus far on the evaluation of fissure sealants using artificial intelligence, which should be recognized as a unique feature of our study. The current diagnostic performance data fit into the overall context of existing dental studies. For example, studies with a similar methodology have documented an accuracy of up to 90% for the detection of caries lesions from clinical images [10,11] or radiographs [2–9]. Considering earlier published data from methodologically similar projects, it can be concluded that our most recent results (Table 1, Figures 2 and 3) are in line with an expected outcome. Our data need to be critically assessed from different methodological perspectives. First, it should be highlighted that the pipeline used for image augmentation, transfer learning and the chosen CNN architecture (ResNeXt-101-32x8d) represents an up-to-date approach that may have enhanced the documented results. Second, as our study was performed on good quality professional clinical photographs, the results may have been positively influenced by this factor. None of the images used were overexposed or underexposed, and the teeth investigated were mostly free of plaque, calculus and saliva. All the images were normalized, cropped and standardized before processing. Third, only unsealed posterior teeth and sealed teeth of varying quality were included in the study materials. Cases with caries lesions, developmental defects, and direct or indirect dental restorations were excluded from the project to enable unbiased learning of the CNN. Another methodological advantage in this context appears to be the use of single tooth images, because interfering information from adjacent teeth or margins was mostly excluded. Consequently, it can be expected that the use of other image formats, e.g., clinical images with multiple teeth or the whole jaw, will result in a lower model performance. The number of available clinical photographs is a limitation that must be critically examined. Here, several thousand images at best should be includable, as the number of images is a crucial consideration for this type of study. In the present analysis, we were able to include 2352 clinical images, which should be interpreted as the minimum number. This fact should not be underestimated, because increasing the number of images will extend the training of the CNN and could improve the CNN precision. Further improvements in the model performance can be expected by extending the number of image samples and using the image segmentation technique. The latter approach results in precise image labeling and could be considered as the method of choice to reach the long-term goal of almost perfect detection and assessment of fissure sealants from clinical photographs by AI methods.

#### 5. Conclusions

The clinical application of AI methods in software applications may be feasible but fundamental dental research needs to be performed first. The results of the present study show that a trained CNN detected sealant intraoral photographs with an agreement of 98.7% with reference decisions. The categorical classification into intact, sufficient and insufficient sealants was performed with a diagnostic accuracy of approximately 90%. Considering the complexity of intraoral findings, it can be concluded that further training of AI-based detection, as well as categorization of prevalent and less-prevalent dental diseases and all types of restorations, is required before clinical use can be recommended.



**Author Contributions:** J.K., R.H., O.M., V.G. and M.H. designed the study protocol. J.K. provided all the intraoral photographs. A.S., J.S., P.E. and J.K. performed all the examinations. O.M. and M.H. programmed the CNN and conducted the statistical analyses. All the authors drafted the manuscript, contributed equally to the interpretation of data and approved the final manuscript version. All authors have read and agreed to the published version of the manuscript.

**Funding:** No funding was received for this study.

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

**Informed Consent Statement:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Data Availability Statement:** The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

**Conflicts of Interest:** The authors declare no potential conflict of interest with respect to the authorship and publication of this article, including financial interests and the provision of study materials by the manufacturer for free or at a discount.

## References

1. Schwendicke, F.; Samek, W.; Krois, J. Artificial Intelligence in Dentistry: Chances and Challenges. *J. Dent. Res.* **2020**, *99*, 769–774. [[CrossRef](#)] [[PubMed](#)]
2. Bejnordi, B.E.; Litjens, G.; van der Laak, J.A. Machine learning compared with pathologist assessment-reply. *JAMA* **2018**, *319*, 1726. [[CrossRef](#)] [[PubMed](#)]
3. 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, online ahead of print. [[CrossRef](#)] [[PubMed](#)]
4. Geetha, V.; Aprameya, K.S.; Hinduja, D.M. Dental caries diagnosis in digital radiographs using back-propagation neural network. *Health Inf. Sci. Syst.* **2020**, *8*, 8, (eCollection Dec). [[CrossRef](#)] [[PubMed](#)]
5. Khan, H.A.; Haider, M.A.; Ansari, H.A.; Ishaq, H.; Kiyani, A.; Sohail, K.; Muhammad, M.; Khurram, S.A. Automated feature detection in dental periapical radiographs by using deep learning. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2020**, *131*, 711–720. [[CrossRef](#)] [[PubMed](#)]
6. Lee, J.H.; Kim, D.H.; Jeong, S.N.; Choi, S.H. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J. Periodontal Implant Sci.* **2018**, *48*, 114–123. [[CrossRef](#)] [[PubMed](#)]
7. Park, W.J.; Park, J.B. History and application of artificial neural networks in dentistry. *Eur. J. Dent.* **2018**, *12*, 594–601. [[CrossRef](#)]
8. Lee, J.H.; Kim, D.H.; Jeong, S.N.; Choi, S.H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.* **2018**, *77*, 106–111. [[CrossRef](#)]
9. Vinayahalingam, S.; Kempers, S.; Limon, L.; Deibel, D.; Maal, T.; Hanisch, M.; Bergé, S.; Xi, T. Classification of caries in third molars on panoramic radiographs using deep learning. *Sci. Rep.* **2021**, *11*, 12609. [[CrossRef](#)] [[PubMed](#)]
10. 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*, online ahead of print. [[CrossRef](#)]
11. Kühnisch, J.; Meyer, O.; Hesenius, M.; Hickel, R.; Gruhn, V. Caries detection on intraoral images using artificial intelligence. *J. Dent. Res.* **2021**, online ahead of print. [[CrossRef](#)] [[PubMed](#)]
12. Moutselos, K.; Berdouses, E.; Oulis, C.; Maglogiannis, I. Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2019**, *2019*, 1617–1620. [[CrossRef](#)] [[PubMed](#)]
13. 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.* **2020**, *26*, online ahead of print. [[CrossRef](#)]
14. You, W.; Hao, A.; Li, S.; Wang, Y.; Xia, B. Deep learning-based dental plaque detection on primary teeth: A comparison with clinical assessments. *BMC Oral Health* **2020**, *20*, 141. [[CrossRef](#)] [[PubMed](#)]
15. Jordan, R.A.; Micheelis, W. *Fünfte Deutsche Mundgesundheitsstudie (DMS V)*; Institut der deutschen Zahnärzte (IDZ): Köln, Germany, 2016; ISBN 978-3-7691-0020-4.
16. 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 Group. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ* **2015**, *351*, h5527. [[CrossRef](#)] [[PubMed](#)]
17. 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*. [[CrossRef](#)] [[PubMed](#)]



18. Heinrich-Weltzien, R.; Kühnisch, J. Häufigkeit und Qualität der Fissurenversiegelung bei 8- und 14-jährigen. *Prophyl. Impuls* **1999**, *3*, 6–14.
19. Simonsen, R.J. Retention and effectiveness of dental sealants after 15 years. *JADA* **1991**, *122*, 34–42. [[CrossRef](#)] [[PubMed](#)]
20. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. *IEEE Conf. Comput. Vis. Pattern Recognit.* **2017**, 5987–5995. [[CrossRef](#)]
21. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *JMLR* **2014**, *15*, 1929–1958.
22. Matthews, D.E.; Farewell, V.T. *Using and Understanding Medical Statistics*, 5th ed.; Karger: Basel, Switzerland, 2015; ISBN 3318054585.
23. Simonyan, K.; Vedaldi, A.; Zisserman, A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. Available online: <https://arxiv.org/pdf/1312.6034.pdf>. (accessed on 26 July 2021).

## Literaturverzeichnis

Askar, H.; Krois, J.; Rohrer, C.; Mertens, S.; Elhennawy, K.; Ottolenghi, L.; Mazur, M.; Paris, S.; Schwendicke, F. (2021). Detecting white spot lesions on dental photography using deep learning: A pilot study. *J. Dent.*, 107. online ahead of print.

Bejnordi, B.E.; Litjens, G.; van der Laak, J.A. (2018). Machine learning compared with pathologist assessment-reply. *JAMA*, 319, 1726.

Cantu, A.G.; Gehrung, S.; Krois, J.; Chaurasia, A.; Rossi, J.G.; Gaudin, R.; Elhennawy, K.; Schwendicke, F. (2020). Detecting caries lesions of different radiographic extension on bitewings using deep learning. *J. Dent.*, 100, 103425, online ahead of print.

Geetha, V.; Aprameya, K.S.; Hinduja, D.M. (2020). Dental caries diagnosis in digital radiographs using back-propagation neural network. *Health Inf. Sci. Syst.*, 8, 8, (eCollection Dec).

Heinrich-Weltzien, R.; Kühnisch, J. (1999). Häufigkeit und Qualität der Fissurenversiegelung bei 8- und 14-jährigen. *Prophyl. Impuls*, 3, 6–14.

Khan, H.A.; Haider, M.A.; Ansari, H.A.; Ishaq, H.; Kiyani, A.; Sohail, K.; Muhammad, M.; Khurram, S.A. (2020). Automated feature detection in dental periapical radiographs by using deep learning. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.*, 131, 711–720.

Kühnisch, J.; Meyer, O.; Hesenius, M.; Hickel, R.; Gruhn, V. (2021). Caries detection on intraoral images using artificial intelligence. *J. Dent. Res.* online ahead of print.

LeCun Y.; Bengio Y.; Hinton G. (2015). Deep learning. *Nature*, 521, 436-444.

Lee, J.H.; Kim, D.H.; Jeong, S.N.; Choi, S.H. (2018a). Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J. Periodontal Implant Sci.*, 48, 114–123.

Lee, J.H.; Kim, D.H.; Jeong, S.N.; Choi, S.H. (2018b). Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.*, 77, 106–111.

Matthews, D. E.; Farewell, V. T. (2015) Using and understanding medical statistics. 5<sup>th</sup> ed., Karger. Basel, Switzerland, ISBN 3318054585.

Moutselos, K.; Berdouses, E.; Oulis, C.; Maglogiannis, I. (2019). Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2019, 1617–1620.

Muschelli J. (2020). ROC and AUC with a binary predictor: a potentially misleading metric. *J. Classif.*, 37, 696-708.

Park, W.J.; Park, J.B. (2018). History and application of artificial neural networks in dentistry. *Eur. J. Dent.*, 12, 594–601.

Schwendicke, F.; Samek, W.; Krois, J. (2020). Artificial Intelligence in Dentistry: Chances and Challenges. *J. Dent. Res.*, 99, 769–774.

Simonsen, R.J. (1991). Retention and effectiveness of dental sealants after 15 years. *JADA*, 122, 34–42.

Vinayahalingam, S.; Kempers, S.; Limon, L.; Deibel, D.; Maal, T.; Hanisch, M.; Bergé, S.; Xi, T. (2021). Classification of caries in third molars on panoramic radiographs using deep learning. *Sci. Rep.*, 11, 12609.

- Welbury R.; Raadal M.; Lygidakis N.A. (2004). EAPD guidelines for the use of pit and fissure sealants. *Eur J Paediatr Dent.*, 5, 179-184.
- Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. (2017). Aggregated residual transformations for deep neural networks. *IEEE Conf. Comput. Vis. Pattern Recognit.* 5987–5995
- You, W.; Hao, A.; Li, S.; Wang, Y.; Xia, B. (2020). Deep learning-based dental plaque detection on primary teeth: A comparison with clinical assessments. *BMC Oral Health*, 20, 141.
- Zhang, X.; Liang, Y.; Li, W.; Liu, C.; Gu, D.; Sun, W.; Miao, L. (2020). Development and evaluation of deep learning for screening dental caries from oral photographs. *Oral Dis.*, 26. online ahead of print

## **Danksagung**

Besonders danken möchte ich allen, die mich während der Erarbeitung meiner Dissertation unterstützt haben. Der größte Dank gilt meinem Doktorvater Herrn Prof. Dr. Jan Kühnisch für seine tatkräftige Unterstützung, seine Motivation in den richtigen Momenten, sein stets offenes Ohr und sein enormes Engagement. Ebenso danken möchte ich den Co-Autoren Paula Engels, Jule Schönewolf, Ole Meyer und Marc Hesenius für ihren unentwegten Einsatz. Für die stete und bestimmt nicht immer einfache moralische Unterstützung und die vielen Ermutigungen danke ich meinen Freunden Natalie und Michael sowie meinen Eltern Barbara und Wolfgang.